SAND2018-1519R

ANALYSIS REPORT FOR EXASCALE STORAGE REQUIREMENTS FOR SCIENTIFIC DATA

Abstract

Over the next 10 years, the Department of Energy will be transitioning from Petascale to Exascale Computing resulting in data storage, networking, and infrastructure requirements to increase by three orders of magnitude. The technologies and best practices used today are the result of a relatively slow evolution of ancestral technologies developed in the 1950s and 1960s. These include magnetic tape, magnetic disk, networking, databases, file systems, and operating systems.

These technologies will continue to evolve over the next 10 to 15 years on a reasonably predictable path. Experience with the challenges involved in transitioning these fundamental technologies from Terascale to Petascale computing systems has raised questions about how these will scale another 3 or 4 orders of magnitude to meet the requirements imposed by Exascale computing systems.

This report is focused on the most concerning scaling issues with data storage systems as they relate to High Performance Computing- and presents options for a path forward. Given the ability to store exponentially increasing amounts of data, far more advanced concepts and use of metadata will be critical to managing data in Exascale computing systems.

Thomas Ruwart

t.maurice.ruwart@gmail.com

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Contents

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1 Introduction

This report examines the requirements for Exascale Data Storage of Scientific Data at the Department of Energy as High-Performance Computing moves to Exascale Computing and beyond. Much of the data storage hardware and software architectures and implementations in use today were designed 20 or more years ago when GigaFLOP computing was a $20 Million problem and 16GigaBytes of data storage was a luxury at most HPC data centers. One TeraByte of disk-based data storage was unheard of[1]. These data storage systems formed the basis of software and best practices used at HPC data centers today.

As High-Performance Computing evolved from GigaFLOPs to TeraFLOPs to PetaFLOPs, the supporting data storage hardware and software expanded to meet the requirements of these bigger, faster, better supercomputers, but they did not necessarily *evolve*. Hence, the current data storage hardware and software used to support current-day PetaFLOP supercomputers is simply larger (higher-capacity), somewhat faster, and no better versions of that which supported the GigaFLOP supercomputers of the 1990s.

The subsequent sections of this report address the requirements for Exascale Data Storage Systems outlined in the previous report (See Section 24 for the full report. Concluding this report will be recommendations for technologies to investigate that can provide one or more parts of the overall Exascale data storage system).

2 Requirements for Exascale Data Storage Systems

A set of 14 requirements set forth in the prior report are listed in Section.24 The 14 requirements were condensed to a smaller set of 7 requirements that were used as a basis for this report (refer to Section25 for details about how the 14 requirements were condensed):

1. Hierarchical Storage Management System(s)
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. Data security

The next section will provide recommendations based on the condensed requirement set for: 1, 3, 5 and 10 years out:

- 1 year – Products and services currently being deployed or in beta testing.
- 3 years – Products and services in alpha testing and not yet ready for prime time.
- 5 years – Products and services in the final stages of design and initial development.
- 10 years – Speculation on the growth and evolution of data storage technologies and best practices. A "vision" of what the future might look like based on prior experience and current trends.

[1] In 1995, a one-TeraByte disk subsystem consisted of more than 250 enterprise-class 4GB disk drives in several equipment racks.

2.1 Hierarchical Storage Management Systems

Hierarchical Storage Management (HSM) systems within the context of this document, refer to the software and hardware that manages movement and storage of data between external active storage systems and internal archival storage systems. This section will describe HSM responsibilities and the increasing importance of rich metadata for archived data.

2.1.1 The Data Storage Hierarchy

The Data Storage Hierarchy is best described by the pyramid in Figure 1. It serves as an important reference when discussing subjects such as Hierarchical Storage Management Systems (HSM). HSMs can conceivably manage data movement between any number of adjacent tiers. In practice however, HSMs typically manage data movement from Capacity HDD (Tier 2) to Near Online (aka Nearline) HDD (Tier 3) and subsequently to Offline Removable Media (Tier 4). Data movement between Offline Removable Media (Tier 4) and Warehoused Removable Media (Tier 5) involves physically moving the media to an offsite warehouse for example.

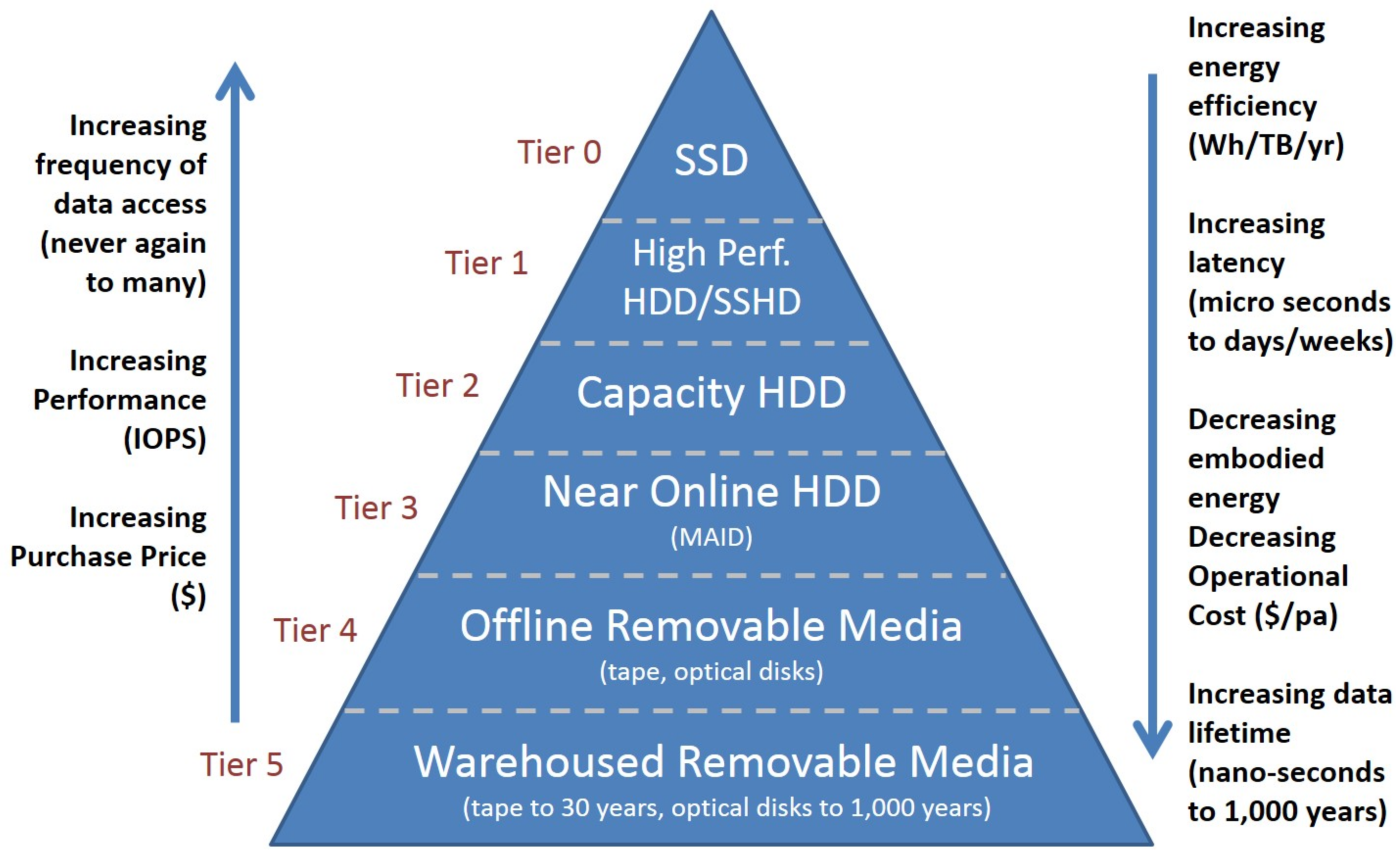


Figure 1 - The Data Storage Hierarchy. Tier 4, Offline Removable Media, is also referred to as "Cold" storage and Tier 5, Warehoused Removable Media is referred to as "Frozen" storage. HSMs generally manage data movement between Tiers 2 and 5. Source: Image Gallery of Data Storage Hierarchy, lacasamorett.com.

Many of the DoE HPC data centers use an HSM called the High-Performance Storage System (HPSS) developed in the 1990s when the idea of an *exabyte scale* system was never part of the design discussions. As a result, these aging HSMs may have unforeseen scaling limits that can manifest themselves in unpredictable ways. Therefore, it is necessary to talk with companies that found it necessary to design, build, and manage exascale data storage systems. These include companies such as Google Cloud Services, Amazon Web Services, and Microsoft Azure to name a few. These are companies that are already well into exascale data management and can serve as valuable resources to the DoE.

That said, it is not being suggested that a cloud infrastructure be employed around the DoE Exascale data storage systems. Rather, it is important to examine and adopt *cloud best practices* where appropriate to provide an abstraction of the underlying data storage systems and technologies. This abstraction is tailored specifically for web-based access where the users do not need to know or care about what or how their data is stored so long as they can access it on-demand. In other words, the user does not need to know anything about the underlying HSM being used to store their data.

There are many development efforts focused on HSMs responsible for:

- Storing large quantities of data where *large* is 10EB or more
- Accessing the associated storage devices such as disks, tapes, libraries, RAIDs, MAIDs, RAITs, …etc.
- Managing data migration through the tiers of the data storage hierarchy using either explicit or policy-based migration methods
- Migrating data to/from the lowest tiers of the storage hierarchy such as *frozen* data

Furthermore, some companies are devoted to enhancing the metadata that describes data at any tier of the data storage hierarchy, including the lowest tiers where most of the data resides. As the amount of archived data increases and ages in the archive, its metadata becomes increasingly critical in aiding the user to more efficiently use *archived* data.

It is important to note that accessing blobs[2] in frozen[3] storage is generally difficult and time consuming. Accessing data in frozen storage should only be done as required. Metadata becomes critical when browsing frozen data. This leads to rich metadata requirements to generate and tag datasets that describe things like:

- what the datasets represent
- how the datasets are structured
- the dataset provenance
- methods that can be used with the datasets

This seems to be beyond what is available with current file systems and HSMs and is key to addressing part of the metadata problem. HSM systems must innovate to provide significantly improved metadata capabilities. The integrated component solution, pairing HSM systems and meta-data systems, allows each to address related but disparate problem sets. This allows HSM systems to more specifically meet extreme scale performance and storage. It also allows the meta-data component solution to address user meta-data requirements, without direct imposition of the HSM solution. It is important to note that HSM systems should not be entirely responsible for the generation and maintenance of rich metadata. Rather, they need to be *metadata-aware* and use metadata to efficiently limit physically accessing data on tape or any other removable media.

[2] A "blob" is a Binary Large Object or a large group of bits independent of structure

[3] The term "Frozen" data refers to data on removable media that has been physically stored someplace outside the media library, such as a warehouse.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

2.2 Storage Media and Associated Hardware

Several reports and discussions with industry experts on the future of magnetic tape, magnetic disk, optical disc, and solid-state storage devices (SSD) all suggest the following:

- Magnetic tape is not dead. (See Sections 9, 10, and 11for details)
  - Usage is growing for nearline, offline data storage and is evolving to increase areal bit density and reduce the Bit Error Rate. (See Section 9.5 for details)
  - Tape areal bit density tracks closely behind that of magnetic disk by a couple generations. *Figure 2 Magnetic Tape Areal Density Trends 1991 to 2025"* represents the current predictions regarding magnetic recording density. Chart provided courtesy of the Information Storage Industry Consortium (INSIC). ©2016 Information Storage Industry Consortium - All Rights Reserved
  - Current roadmaps and recent technology demonstrations suggest at least one order of magnitude increase in areal density within the next 3-5 years. (See Section 9.2 for details)
  - Linear Tape Open (LTO) will be the sole tape standard going forward.
  - IBM, HPE, and Quantum are three primary suppliers of high-end LTO tape drives

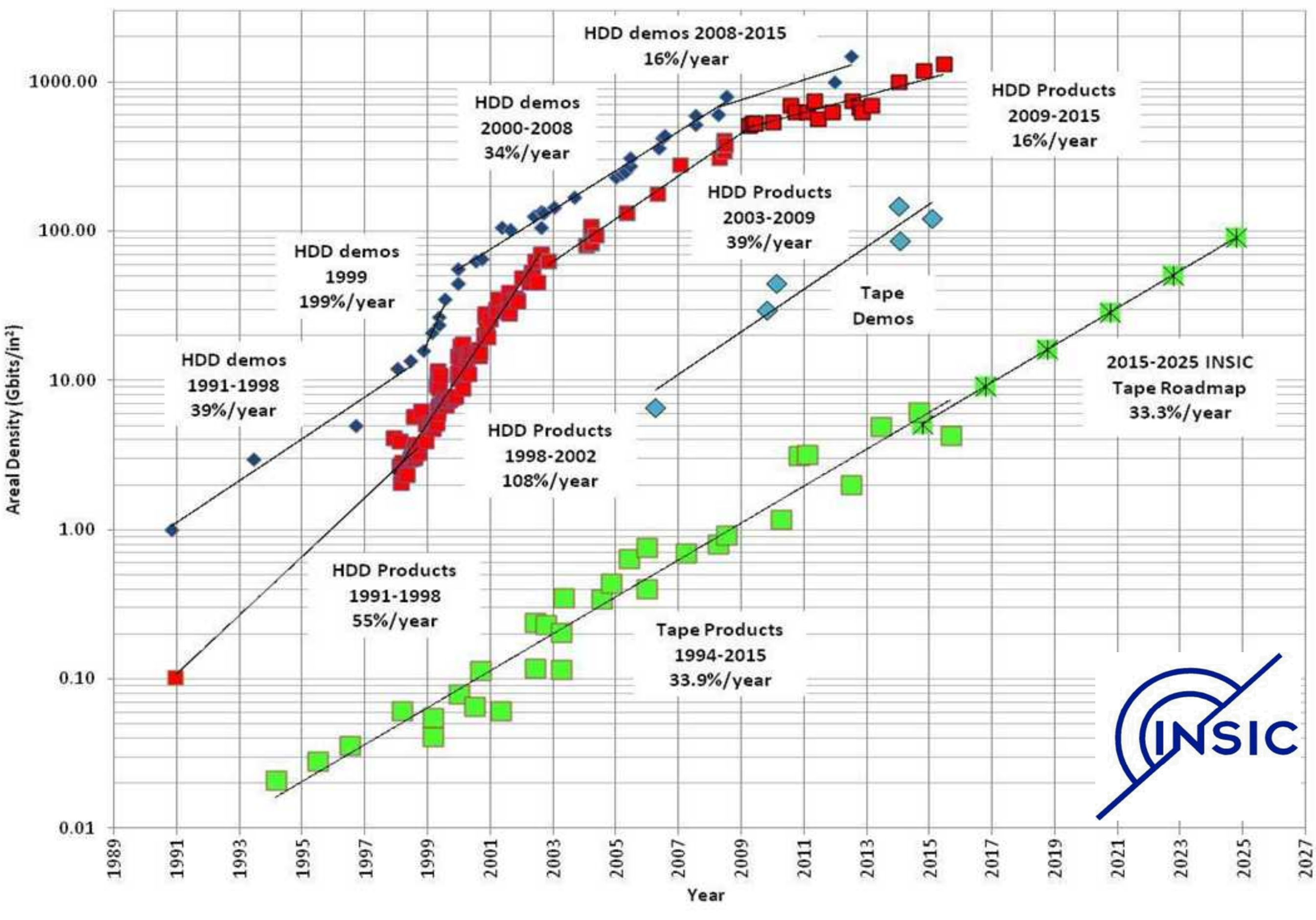


Figure 2 Magnetic Tape Areal Density Trends 1991 to 2025

- Magnetic tape media is extremely reliable. LTO-7 tape drives and media have a Bit Error Rate (BER) of 1 in 10^{19}, an improvement over the LTO-6 drive that has a BER of 1 in 10^{17}.
  - Sony and Fujifilm continue to develop high-density recording media including [See Section 8 for details].

  - According to Fujifilm, tape-based magnetic recording will be viable for at least two to three more decades.
- It is expected that traditional disk drives will continue to provide high-capacity, high bandwidth, and moderate latency data storage.
  - Magnetic disk is expected to be around for at least another 10-15 years with one to two orders of magnitude increase in areal density *(see Figure 3 Magnetic Disk Areal Density Trends 2013 to 2025)*
  - The rotational latency and seek times are unlikely to change much for magnetic disk drives
  - Data bandwidth will increase about 40% for every doubling of areal density.

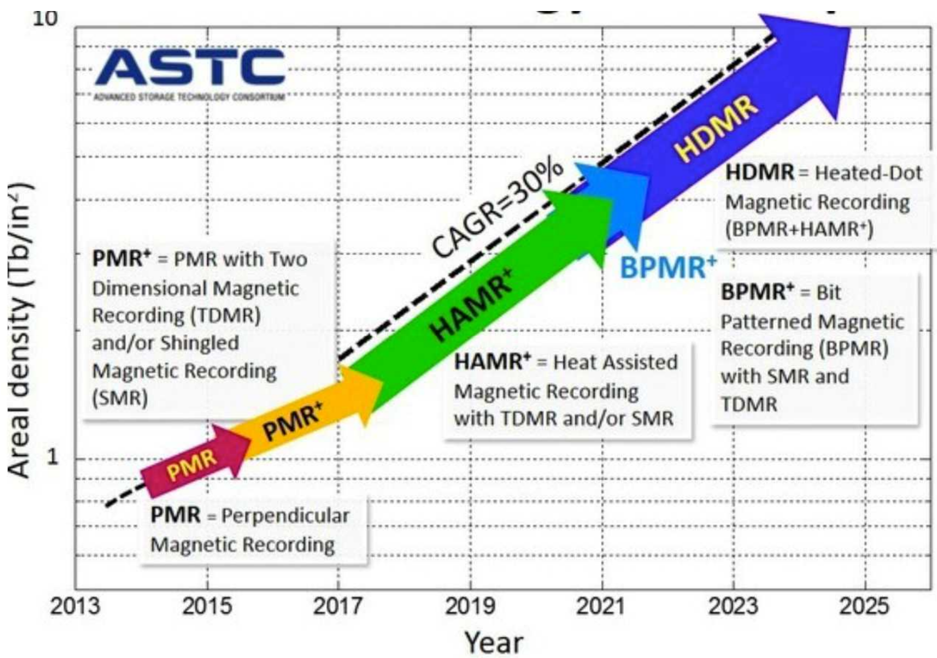


Figure 3 Magnetic Disk Areal Density Trends 2013 to 2025

- Solid State Storage Device (SSD) technology is on a steep development curve.
  - SSD chip manufacturers have developed ways to stack SSD chips within a single package overcoming density limitations. This provides two or three orders of magnitude more capacity and performance in the same footprint as a single-chip package.
  - Stacking chips improves the “unit capacity” of an SSD, already surpassing that of magnetic disk (albeit at 10 to 100 times the price).
  - FLASH technology still has issues with wear. It is also demonstrably poor at long-term offline data storage and is not a good choice for archive data.
  - As of 2017 there is not enough total capacity for chip fabrication to make enough FLASH to replace any significant amount of disk-based data storage capacity.
- Future solid-state data storage technologies may appear as viable competitors to FLASH within the next 10-15 years. These include technologies such as:
  - Memristors (HP) – See Section 13 for more details
  - Racetrack Memory (IBM) – See Section 14 for more details
  - Phase Change Memory (PCM) – See Section 15 for more details.

It is important to note that the underlying data storage technology is not as critical as the higher-level access and data management methods employed to use the storage hardware. Furthermore, SSD technologies are making their way into new system architectures (e.g. non-volatile RAM) thereby significantly improving rich metadata management. Given the low latency, non-volatile, and relatively high capacity, although expensive, they can serve as ideal physical storage for rich metadata.

Lastly, it should be mentioned that the use of any removeable media implies the use of a robotic library that contains the media and devices necessary to read/write the media. Robotic libraries are well understood and do not need to be mentioned further in this document.

2.3 Convergence of Interconnect Technologies

Interconnect technologies in the context of this report refer to various storage technologies such as Fibre Channel, Ethernet, and InfiniBand. There are two distinct aspects of interconnect technologies, physical communication technology and the communication protocols. These should be considered independently.

Data transmission speed is the primary metric for physical communication technology. Currently, 100Gbits/sec is the fastest network technology available for transmitting data between systems. It is expected that 10 TB/sec will be available in the 2023-2027 time frame. [See Section 16 for details on the Ethernet Roadmap.]

The term "convergence", in the context of communications, concerns the ability to use multiple protocols on a single communication medium. This implies that the hardware endpoints and intermediate equipment be compatible with differing protocols. For example, ISCSI consists of running the SCSI protocol (specifically, the disk drive command subset of the SCSI protocol) over ethernet using IP. This allows a user to effectively have a Storage Area Network (SAN) without the expense or complexity of a Fibre Channel SAN while sacrificing only a relatively small amount of performance. Other converged networks include Fibre Channel Over Ethernet (FCoE) that allows for the proper use of the Fibre Channel SCSI Protocol (FCP) to run over an IP connection without any Fibre Channel switches, but it does require the use of a Converged Network Adapter at each end to handle the Fibre Channel and TCP protocols simultaneously. Similarly, IP over InfiniBand (IPoIB) allows an InfiniBand fabric to look like a very fast, very low latency IP network.

A recent example of a converged network is the NVME Fabric. This allows for any computer on the NVME Fabric to read/write NVME spaces in any other computer system on the fabric. The physical interconnecting fabric can be either InfiniBand or XXGbE[4]. The significance of this, particularly with an IB fabric, is the ability for data transfers to happen with little or no CPU intervention on either end of the transfers (aka Remote Direct Memory Access or RDMA). Furthermore, by definition the NVME data storage devices are non-volatile memory. An NVME fabric with X nodes, each containing some amount of NVME, appears very similar to a NUMA (Non-Uniform Memory Access) architecture. This becomes highly useful when designing a metadata engine that needs to search a large amount of metadata (think petabytes of metadata that, in a sense, appears to be in system memory).

To summarize, it is unlikely that interconnect technologies will become an I/O bottleneck for exascale data storage systems. Furthermore, there are some new converged communication technologies that allow for more advanced system architectures in the area of metadata management that will be useful to the exascale data storage systems.

2.4 Metadata, Data, and Information Access at Exascale

The storage system needs to support considerably more detailed information about the data it is storing in order to minimize physical access to the data itself. Typically, "metadata" is much smaller than the data it represents. It should be structured to grow at a slower rate than the underlying data. For example, a single tag associated with a million file simulation would describe content without physically reading all the files to understand some aspect of the dataset. This holds true for each order of magnitude: millions, billions, trillions.

As new information and/or knowledge is gained about a frozen dataset, new metadata or changes to existing metadata may be desirable. Adding new metadata tags and/or values to existing tags is considered Growth in this case. Simply modifying values of existing tags without adding anything is considered "Evolution" of the dataset as a whole (i.e.

[4] XXGbE where XX=10 is 10GbE or XX=40 is 40GbE

dataset plus associated metadata). It is important to note that the Growth and Evolution of this metadata may require a mechanism for updating the frozen copy of the metadata for consistency. Ideally, the end result is a database of sorts that contains more highly detailed information about the [archived] data that users interested in some aspect of the [archived] data need only search this database rather than physically extract the [archived] data for the same information. This is a general description of making the metadata more usable.

There are several aspects to the metadata "Usability" that need to be addressed. This includes defining and developing the mechanism by which a user can most easily create, manage, and search metadata for their data. Overall guidelines, standards, or best practices should be employed. These have yet to be defined, but as stated in the Introduction, this is an area that the DoE can have the greatest impact.

Another development opportunity for DOE is to help provide rich meta-data, especially for frozen datasets. This may be software or user driven. It may also be performed as data migrates between storage tiers or storage devices (e.g., technology upgrades.) This must be a simple, additive process of meta-data improvement.

Finally, one aspect of metadata that involves extreme scaling is that of the *database* where the metadata resides. The term *database* within the context of this report refers to a *generic* system used to manage and access the metadata. It is assumed here that the collective *amount* of metadata for any specific installation is nearly the same regardless of the specific data management system (i.e. HSM) employed. The important point here is that these databases can get exceedingly large presenting a challenge for the database software and/or hardware to perform adequately. For that reason, it is important to investigate metadata storage, search, and access techniques employed by organizations with similarly large quantities of metadata. These include but are not limited to Google, Amazon, Microsoft, …etc.

2.5 Storage System Management

The term "Storage Management" is overloaded. Therefore, Storage System Management in the context of this report consists of the following:

- Storage Hardware management – day-to-day care and feeding of the physical storage devices, including upgrades and technology refresh
- Storage Capacity management – managing how the physical storage devices are segmented and assigned to systems
- Storage Space management – Assigning data storage capacity to systems and users and recovering unused storage capacity.
- Storage Software management – managing all the software that falls under the storage systems administrative domain, such as HSM software, operating system (Linux, Windows, …etc.) storage components, vendor-specific storage tools, …etc.
- Data management – management of directories and files on storage systems
- Information management – management of groups of related data that make up identifiable information

As these systems grow in size, they will also grow correspondingly in complexity. It will become increasingly important to abstract the underlying storage technology from the "user" such that it can be managed independently and transparently to the "user". By abstracting the data storage infrastructure from the user, the Management and Reliability, Availability, and Serviceability (RAS) characteristics of the user's data become the responsibility of the storage provider. The user makes the assumption that the data will always be there. In the case of the DoE labs, they are the storage providers as well as the consumers.

The concept of Service Level Agreements (SLA) should be developed as a way to manage user expectations with respect to I/O performance, data integrity, data life span, …etc. The reason for this is to decouple the user from the physical system details. For example, if the user knows there are 1,000 hard drives that can transfer data at 100MB/sec each, they may expect to achieve 1,000x100=100,000MB/sec or 100GB/sec. That's not how these systems work.

There are many reports that demonstrate the reliability of various classes of disk drives (enterprise vs consumer), tapes, and optical media for both short-term and long-term (archival) data storage. It comes down to:

- Where in the storage hierarchy to use magnetic disk and how much (one aspect of availability)
- Where to use tape and how much (one aspect of availability)
- How to manage the effects of hardware/software failures on availability and performance
- How to manage rolling technology refresh (one aspect of availability)
- How to avoid vendor lock-in with the storage management software and/or hardware
- How to avoid being so splintered that when something does not work, none of the vendors of the various parts will accept responsibility for fixing it. This becomes a reliability, availability, and serviceability problem.

Storage hardware management primarily involves the installation, configuration, maintenance, and removal of physical hardware. Once the initial hardware is installed, future additions, modifications, and/or removal of this hardware is referred to as "technology refresh." In general, technology refresh for any specific piece of hardware occurs on a timescale of 3-5 years. The process of decommissioning equipment that has reached the end of its useful life and the installation of the new equipment is assumed to be continuous. Old equipment is removed, and new equipment is installed on a regular basis. These activities must be transparent and non-disruptive to the overall system.

As the HPC systems evolve and transition toward exascale, **storage capacity management** will increase by three orders of magnitude (a thousand times larger.) The potential number of physical storage system components under management will also grow significantly, albeit not necessarily by the same factor - factors of tens or hundreds. It is likely that the new equipment and/or software will be managed *differently* than the older technology. Thus, there will be old and new management practices for similar equipment and software.

Significantly larger computer and applications exacerbates the need for **storage space management**. supercomputer systems generate more data from one generation to the next. How the I/O paths from the processors to external storage from one generation to the next would usually result in subsequent changes in the overall organization of the storage systems. In the late 1990s the data storage systems were *decoupled* from the processing system so that they could evolve independently. For very high-performance systems, decoupling was accomplished by employing Storage Area Networks (SANs), followed closely by shared file systems (GFS, cXFS, StoreNext, …etc.), distributed file systems (GPFS, …etc.), and ultimately by object-based distributed file systems (Lustre, PanFS, CEPH, …etc.) There are many other file systems in use and in development at any given time and the precise meaning of the term "files system" has subsequently been blurred.

It is the evolution of these file systems that is key to providing access to data that meets or exceeds the requirements in the SLA. The file system is responsible for data placement and its design should be cognizant of the underlying storage device characteristics in order to efficiently use the device. With the increase in unit-capacity of storage devices such as magnetic disk drives, it becomes increasingly difficult to use them as "dumb" block devices. It is necessary that high-capacity storage devices have some knowledge about the data being stored on them. This was the original intent of *object-based storage devices*. For example, individual disk drives would store and retrieve "objects" rather than fixed sized data blocks. Each storage device requires a rudimentary file system to manage the underlying data. For example, Solid State Devices (SSDs) and the LTO tapes have embedded FLASH memory for an LTO File System (LTFS)

Data management in the context of this report, refers to Hierarchical Storage Management (HSM) systems. As previously mentioned, HSMs attempt to automate the process of placing data on appropriate devices using policy engines. While this worked to an extent, there never seemed to be any reasonably "well defined" or "standard" set of policies that the user (i.e. creator of the data) could use to determine the long-term fate of the data as it aged. Furthermore, the HSM systems available to date are more often than not proprietary and do not necessarily work and

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

play well together. Any one vendor's HSM system usually makes it relatively easy to migrate data from another vendor's HSM system, but there may be issues mapping attributes and policies between the two HSM systems because there are no well-defined standards. To address the issues with legacy HSM systems, the current trend is toward:

- Appliances that virtualize the storage devices they manage
- Software-defined storage
- Storage as a service

Each of these use policies defined in part by the users to manage the migration of data from one tier down to the next. With respect to the specific requirement of data placement on devices appropriate to subsequent use[5], the user can specify when specific datasets can be moved from fast storage to slower storage (i.e. tape or long-term archive) or in the reverse. However, data placement on storage systems or devices that closely match the access (IOPs) or bandwidth (MB/s) requirements of the user processing that data is difficult at best. This is due, in part, to the lack of an ability to specify parametrically when the programs will need at any given time from any given set of data.

The use of SLAs can provide more deterministic access to data and can also present the user with illusion of being able to store an infinite amount of data. There are several aspects to this that can scale independently of each other. The first is simply the bits of actual user data that need be stored. This is normally the bulk of the overall data (data+metadata). The second is the metadata or the information that describes the data being stored (see section 4) and is normally much smaller than the bulk data. That said, given the ability to specify rich metadata that can conceivably be updated as needed over the lifetime of the data need be considered only by the storage provider – the user should always assume the ability to access an arbitrary amount of metadata.

2.6 Data and Information Management

Data is the tangible output of work done by a supercomputer for any given user. Users generate and subsequently "own" sets of data (aka datasets) that they will move between storage systems and/or data centers for analysis and visualization. *Information* is the output from analyzing data. Generally speaking, information is orders of magnitude smaller than the data from whence it came.

Over time the "value" of the "data" decreases. This is very subjective because each set of data (aka dataset[6]) has its own "value curve" or the rate at which the [perceived] value of the data decreases over time. Within the context of this report, the value of any dataset asymptotically approaches zero but in theory, never reaches zero. Incidentally, the area under asymptote out to infinity is referred to as "the long tail".

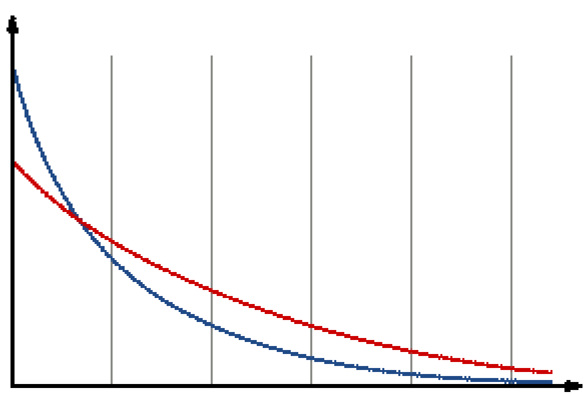


Figure 4. In this figure, the x-axis is time increasing from left to right and the y-axis is subjective value of the data. The "blue" curve starts at a higher "value" than the "red" curve, but it decreases and approaches zero value more rapidly than the "red" curve.

As a result of the long tail, retention period of the data for any given campaign, is considered to be infinite. The reason for the relatively infinite retention period is a property of how the data was generated. Observational data such as astronomical observations or

[5] Requirement #3 in Section 25

[6] The term "dataset" in this context refers to the cumulative data directly related to a specific campaign.

data from a collider, is unique and not reproducible thereby giving it some non-zero value forever. In some cases, simulation data can be regenerated caused by computing system dependencies and hardware architecture used to generate the data. Data that can be regenerated may have an "expiration date" or finite retention period after which the data can be deleted.

Information Management refers to information associated with a given dataset. Compared to data, this information is smaller. It can be represented in various forms such as: binary data, text files, charts and graphs. Data management is primarily an administrative responsibility whereas information management is entirely the user's responsibility.

2.7 Data Security

Here, Data Security is restricted to data being transmitted over external networks between secure sites (aka data in-flight). Encryption algorithms are largely serial in nature and do not parallelize to any significant degree, reducing transfer speed. Network transmission speeds outpace the ability to encrypt at those speeds, thereby limiting the realized end-to-end data transmission of any one stream of data.

To address the end-to-end performance problem, it is necessary to use alternative data transport mechanisms, such as a truck full of tapes. The efficiency of physically moving tapes improves as the amount of data required to be moved increases. From a data security perspective, the tapes can be encrypted to prevent unauthorized use. Because tapes can be written in parallel (not RAIT, just moving independent dataset to their own tape), the problem of parallel encryption is an inherent property of this method. Parallel Encryption means that each tape being written has its own encryption engine that can [presumably] run at the peak streaming speed of the tape drive. Therefore, each of the tapes in the RAIT group is encrypted without a performance penalty. This does, however, inflate the key-management associated with encryption because there needs to be one unique key for each tape in the RAIT group rather than a single key for the entire dataset.

3 A word about Clouds and Cloud-like Services

The latest buzzwords in large-scale data storage is *cloud*. It has become an overloaded description of a variety of services performed by a service provider somewhere on the Internet at large. Cloud Services are basically the evolution of what used to be services on The Grid, formerly known as things like Storage Service Providers, Application Service Providers, …etc. The overall benefit that the Cloud provides cloud users is the ubiquity of the data center, which is of little or no concern to the users themselves so long as it provides the necessary services at reasonable cost and performance. Sections 20 thru 23 contain copious information about all things Cloud.

3.1 Cloud-like Services

Within the context of this report, the use of a Cloud, either external or internal, is *not* being recommended or suggested. Rather, the use of *cloud-like services* is being recommended to provide software technologies and protocols that have been shown to scale to the extent required by the DoE. This is particularly true for data storage services.

Cloud services cover a very wide range of items as seen in Section 7 . Fortunately, the scope of Cloud Storage Services proposed here is very narrow and does not require an actual cloud. The basic Cloud-like Storage Services consist of the following:

- Requesting Storage and agreeing to specific a specific level of service (an SLA)
- Uploading data to the storage
- Downloading data from the storage – either copying the data or moving it out
- Defining and assigning metadata for objects – this includes predefined or user-defined key-value pairs
- Searching the data or metadata

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

- Removing objects from storage – permanent file deletion
- Releasing Storage – making the storage space available to allocate to another user

The idea is to make it appear to the user that they are using Cloud Storage for their data archiving needs. The user experience of requesting data storage, filling, draining it, searching it, and releasing it should look similar to performing the same tasks on Google Cloud Storage or Amazon Storage Service.

For example, Google Cloud Storage can provide a large amount of storage for a user to fill up with, say, simulation data from the Google compute service. The user specifies the level of service expected during the simulation, after the simulation, and after the completion of the project. Amazon, in turn, provides the user with a Service Level Agreement (SLA) detailing things like bandwidth, access rates, and retention period. If the retention period is essentially until the end of time, the SLA will lay out access times to recover the data from the Amazon archive. The use of an SLA, whether explicit or implicit, is one example of a "*Cloud-like service*" for data storage. Refer to the Google Cloud Storage SLA in Section 7 for an example of what the SLA actually looks like.

The motivation for adopting Cloud-like Practices is to assist in making the storage, searching, and retrieval of large data stores more like accessing data on the Internet which is what users are more accustomed to. It will also provide a layer of abstraction behind which is the more traditional HSM that can move into the future at its own pace.

4 VENDORS AND RESPECTIVE PRODUCTS

There are many companies and organisations offering products and services that deserve investigation. They can be divided into the following categories:

- Software: Anything that comes on a piece of storage media (DVD, FLASH, …etc.) or downloadable from the vendor. An example would be HSM software.
- Hardware: Strictly hardware with minimal support software. An example would be an LTO tape drive or a tape library robot excluding any tape library software.
- Integrated Solutions: Products that combine software with hardware purchased from a single vendor. An example would be an appliance designed to be inserted into the data flow of data being ingested into an HSM strictly for adding metadata to objects as they pass through. Generally speaking, an *appliance* is a physical, turn-key system that includes all the necessary hardware and software to operate independently of any external systems.
- Services: This pertains mainly to Cloud Design and Development services that are not required but may be useful at some point. They are listed here for completeness.

4.1 SOFTWARE

The principle software technologies of interest are HSM products that run on generic platforms. Other software products include advanced metadata management software and other capabilities not mentioned in this report. The criteria used to select these specific vendors is the list of 14 requirements set forth in the previous report. Each of the HSMs listed support all 14 requirements to some degree. The degree to which they support the individual requirements would be the goal of an evaluation lab at the DoE. The current vendors of HSM and other useful software products include:

- **IBM** with High Performance Storage System (HPSS) as well as General Purpose File System (GPFS)
- **Quantum** with StorNext and their Storage Manager. The extent to which Quantum supports libraries other than their own is not understood at this point.
- **Versity** is a startup by the author of the QFS/SAMFS HSM product that Sun Microsystems marketed. After Oracle purchased Sun, Versity obtained the right to use the open-source version of QFS/SAMFS as the base of a new, more capable HSM. Version 1.0 is currently available. Version 2.0 will be available Q1 of 2018 with significant enhancements.
- **Oracle** currently has an HSM marketed as OHSM (Oracle HSM). This is just a rebranding of the QFS/SAMFS HSM they inherited when they purchased Sun Microsystems. Reliable sources indicate Oracle is encouraging their OHSM users to migrate their data to [Oracle] cloud storage in preparation for dropping the OHSM product and associated support.
- **Intel** provides HSM software that will interface readily with Lustre.
- **Atavium** is a startup in the process of developing software that will sit between the user and the back-end data storage system. Given the early stages of development, it is not appropriate to state what their software does, other than to say that it is within the scope of software products that should be evaluated.
- **StrongBox Data Solutions** has a product called *StrongLink* that will allow for automatic or automated metadata generation and data tagging before the data is committed to an archive. Strong-Link is available as just software or as an integrated appliance from StrongBox Data Solutions.

4.2 HARDWARE

- **Tape Drives**: The scope of hardware described here is restricted to tape drives, tape libraries, and specialized appliances. Furthermore, for practical reasons, only LTO tape drives are considered in this report. There are currently only three LTO tape drive manufacturers: **HPE**, **IBM**, and **Quantum**. By definition, LTO devices from these three vendors are completely interoperable – i.e. a tape written on any one vendor's LTO drive can be read on any of the vendor's LTO drives.

- **Tape Libraries: Spectra Logic** is the premier supplier of tape libraries consisting of tape drives, physical tape storage and associated robotics. They also offer a *management appliance* called Black Pearl that provides additional, higher-level functionality for their tape libraries. Spectra Logic does not provide HSM software. Rather, they are more of a hardware and middleware provider. It is important to note that Spectra Logic is also the most vendor-neutral of the companies that provide tape libraries, such as Quantum and IBM.

- **Specialized Appliances:** It would be useful to evaluate preconfigured, turn-key appliances from the vendors that offer them, such as Atavium or StrongBox Data Solutions. Even though these appliances are not required by any HSM, they could provide useful functions that augment an HSM.

4.3 Services

Services are divided into two broad categories: services provided to users or services sought by designers/developers.

- **Cloud-like services provided to the users:** Providing cloud-like services is not the same as providing actual cloud services. It is important to emphasize that building an actual cloud it is not suggested or recommended. Rather, the user interface to the storage and management of the users' data should appear to a user to *look* and *behave similarly* to that provided by the Google Cloud Storage or Amazon Simple Storage Service or Microsoft Azure Storage.

- **Cloud Services sought by designers/developers:** There are only a few companies that could provide guidance on designing and implementing the restricted set of cloud services that pertain to this report. These include Google, Amazon, Microsoft, and Oracle. For example, Oracle provides a service whereby they will design and build an entire private cloud for a customer, for a price. Even though building an entire cloud is not within the scope of recommendations, they could provide guidance on the cloud-like services that are within the scope of recommendations. Similarly, Microsoft could provide software and guidance on the cloud-like services that are within the scope of recommendations. Microsoft does provide a complete set of cloud software to build out small to large private clouds. Information from Google and Amazon is still forthcoming

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

5 SUMMARY

The intent of this document is to identify the basic components of a data storage system used to store large collections of data for an indefinite period. The hardware and software for such data storage systems has historically been referred to as Hierarchical Storage Management systems. The HSMs in use today were designed and developed in the 1990s and there is concern that they may have serious issues when adapted to an exascale system. Therefore, it is necessary to identify HSM software and hardware vendors that will scale to meet the needs of exascale systems. This report identified vendors to work with to achieve the desired HSM scaling.

Additionally, this report identified vendors providing technologies of interest that deal with the metadata problem at large. Metadata is something that is currently not a significant problem but will become more problematic as the archive scales up. This is the most appropriate time to address the metadata problem, before the archive gets so large that it becomes significantly more difficult to fix.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

6 RECOMMENDATIONS

The following recommendations are meant to address designing, developing, and deploying an Exascale HSM that meets the 14 requirements set forth as the basis of this document. First and foremost are recommendations for HSM software vendors and products to evaluate/compare to the incumbent HSM. These are followed by hardware recommendations for a facility in which to perform evaluations of software and hardware technologies including the ones previously mentioned. All recommendations following the Software and Hardware recommendations are suggestions for useful research and evaluation related specifically to exascale long-term data storage systems, namely, the metadata problem at large. Finally, a few suggestions are offered to help maintain up to date focus and understanding of the above technologies and industry/market direction

It is assumed that there is a facility capable of performing the evaluations and comparisons mentioned in the recommendations.

6.1 SOFTWARE RECOMMENDATIONS

It is assumed that HPSS is the incumbent HSM to which comparisons will be made. Thus, recommending HPSS is redundant. Oracle OHSM, as previously mentioned is essentially an older version of the initial product from Versity. Furthermore, it is believed that Oracle will discontinue OHSM in the near future as they are moving away from HSM archival systems to cloud storage. Any one of these HSM products could conceivably replace HPSS if and when necessary. It is important to note that cost structure of storing data in these HSMs is not addressed in this report because that is a negotiable item.

1. The first HSM to evaluate is Versity. The justification is that the current version of the Versity HSM, Version 1.0, is very mature, widely used, and open sourced so it is not likely to go away if, for some reason, the vendor, Versity, disappears. It will also serve as a precursor to the second version of the Versity HSM, Version 2.0, that is purported to have greatly improved capabilities, particularly in scaling performance. The second version of Versity is scheduled to be available before Q2 in 2018.

2. The second HSM to evaluate is Storage Manager from Quantum. This too is very mature software, but its scaling properties have not been verified. The Storage Manager marketing information makes claims that it will scale to meet exascale requirements, but the current understanding of its architecture is that it may contain performance bottlenecks and/or single points of failures that need to be confirmed one way or another.

3. The third HSM to evaluate is the Lustre HSM, offered by Intel, that is simply named Lustre HSM. The reason this is third in the list is because of its relation to Lustre, a file system used by many but not necessarily all the DoE labs. Given the wide adoption of Lustre at DoE computing facilities, an evaluation of the Lustre HSM is warranted in the event Lustre is adopted at a site where it previously did not exist.

6.2 HARDWARE RECOMMENDATIONS

The selection of hardware vendors is limited by the nature of the technology itself because the market for tape-based technology is rather narrow but deep. There are three essential hardware components to consider:

- Tape drive: The range of tape drive technologies is currently at one: LTO. There are three LTO tape drive manufacturers: IBM, HPE, and Quantum. As previously stated, their tape drives are completely interoperable. The superiority of one brand over another is purely subjective. The recommendation here is to obtain several tape drives from each of the three vendors and evaluate them all.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

- Tape Library: Similar to the tape drives, the number of tape libraries capable of supporting an exascale system are limited to a few companies. A recommendation would be to favor getting a tape library from Spectra Logic because it is HSM-neutral and affords scaling properties applicable to an exascale environment. It is worth mentioning that the HSM vendors IBM and Quantum both offer their own tape library hardware as part of an HSM bundle.

6.3 Research and Evaluation Recommendations

1. Adjacent to any particular HSM is the metadata management software and/or appliances. The first one of these to evaluate is StrongLink from StrongBox Data Solutions because it is currently available. A close second metadata management tool is the one from Atavium, that is believed to be available in early 2018. Both are intended to provide experience with tools that will allow adding rich metadata to files as they move into an HSM. These also provide advanced search and retrieval capabilities not available in any currently available HSM.

2. With respect to clouds, it is worth understanding what *data storage cloud-like practices* would make sense and work in the DoE HPC data centers if for no other reason than to provide a layer of abstraction from the underlying HSM for future flexibility. The goal is to provide a user with an experience similar to one they would have with the Google Storage Cloud, for example. The cloud-like services are restricted to the storage cloud services that are presented to the user when they want to store, retrieve, or remove data from the back-end storage. Using the Google Cloud Storage service as an example, the user would fill out and agree to an online form that describes a Service Level Agreement between the user and the DoE computing facility storage system so to speak. This includes but is not limited to things like reliability, uptime, time to retrieval, …etc. The Storage Cloud Service experience also allows for the movement of data using drag and drop or command-line interfaces to manage anything from small to extremely large data collections. The point is that it needs to *look* like Google Cloud Storage. It does not have to *be* Google Cloud Storage, but it must behave like that. Again, it is very important to understand that providing cloud-like services does not imply building a [private] cloud at any DoE HPC data center.

3. Focus on understanding extreme scaling properties of the overall metadata problem: generating, storing, and searching rich metadata that represents several exabytes of data or several trillion files. Granted, this may not be done at scale. But this laboratory can provide researchers with the ability to test algorithms that can be theoretically scaled. The previously mentioned companies, StrongBox and Atavium, each have or will have product offerings poised to address these metadata issues. It was recommended to obtain and evaluate products from each of these companies.

4. Continually evaluate new technologies as they become available. As mentioned in the section on NVME fabrics, it is a relatively new technology with unique properties that could prove valuable to a large, metadata rich HSM. For example, one recommendation would be to obtain and evaluate an NVME fabric that would tie several servers together via XXGbE or InfiniBand. Then apply knowledge gained from this experience to design and build a tightly coupled distributed metadata search engine for an HSM.

6.4 Other Recommendations and Suggestions

The following is a list of things that would be nice to have but not critical

1. Visit organisations that have already deployed Exascale data storage systems. These include Google, Microsoft, Amazon, Oracle, …etc. It is critical to leverage their experience and expertise to successfully design and deploy such systems at the DoE. These organisations are believed to have been running Exascale systems for several years now. For example, Google (with Google Drive) and Microsoft (with One Drive) provide cloud storage for

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

any of their users that care to use it. Considering the number of users each one has, it is easy to see that the amount of data in terms of numbers of things (files, pictures, documents, videos, …etc.) is enormous. Furthermore, there does not seem to be an expiration date on this data. The intent with these visits is to understand what their cloud storage best practices are and how they would map to the HPC environment at the DoE. A future contract could consist of visiting these organizations for the sole purpose of discussing their approach to their Exascale data challenges. This can also be done with a contract for attending storage conferences and workshops.

2. Work with other Exascale research projects around the US. The Exascale Computing Project (https://exascaleproject.org/) and the MARFS project at LANL are two such projects. The reason for engaging with these groups is to avoid duplicating efforts in any specific area with respect to Exascale system development.

3. Get involved with the various data storage industry organisations and associations such as but not limited to:

    a. Storage Networking Industry Association (SNIA) The SNIA covers many areas in data storage and their events are usually free or very low cost.

    b. Information Storage Industry Consortium (INSIC) This organisation is largely responsible for all things tape. They have small periodic meetings that offer opportunities to network with people in the tape industry.

4. Identify and attend relevant conferences and workshops. This is not an exhaustive list, rather, just a few of the more prominent events that might be of interest. A more detailed and complete list could be generated if needed. These include but not limited to:

    - Supercomputing, usually held the week before Thanksgiving. This conference is in Denver, CO in 2017. It is a good venue to network and potentially meet with people from companies like Microsoft, Intel, Google, Amazon, …etc. about Exascale data storage.

    - Events sponsored by SNIA These events occur periodically throughout the year.

    - IEEE Mass Storage Systems (MSS) This is a more academically-oriented conference held every 18 months. It is also a good place to network with users from other government organisations dealing with similar data storage challenges.

The first recommendation is the most important. Granted, some of these recommendations may be obvious but need to be included to emphasize the importance of developing working relationships with commercial and government organisations with similar data storage issues as well as staying up to date with the developments in the data storage industry at large.

Articles

7 GOOGLE SERVICE LEVEL AGREEMENT

During the term of the Google Cloud Platform License Agreement, Google Cloud Storage License Agreement, or Google Cloud Platform Reseller Agreement (as applicable, the "Agreement"), the Covered Service will provide a Monthly Uptime Percentage to Customer as follows (the "Service Level Objective" or "SLO"):

Covered Service	Monthly Uptime Percentage
Multi-Regional Storage class of Google Cloud Storage	>= 99.95%
Regional Storage class of Google Cloud Storage	>= 99.9%
Nearline, Coldline and Durable Reduced Availability Storage classes of Google Cloud Storage	>= 99.0%

If Google does not meet the SLO, and if Customer meets its obligations under this SLA, Customer will be eligible to receive the Financial Credits described below. This SLA states Customer's sole and exclusive remedy for any failure by Google to meet the SLO. Capitalized terms used in this SLA, but not defined in this SLA, have the meaning set forth in the Agreement. If the Agreement is the Google Cloud Platform Reseller Agreement, then all references to “Customer” in this SLA mean “Reseller,” and any Financial Credit(s) will only apply for impacted Reseller order(s) under the Agreement.

Definitions

The following definitions apply to the SLA:

- **"Back-off Requirements"** means, when an error occurs, the Application is responsible for waiting for a period of time before issuing another request. This means that after the first error, there is a minimum back-off interval of 1 second and for each consecutive error, the back-off interval increases exponentially up to 32 seconds.
- **"Covered Service"** means Google Cloud Storage.
- **"Error Rate"** means the number of Valid Requests that result in a response with HTTP Status 500 and Code “Internal Error” divided by the total number of Valid Requests during that period. Repeated identical requests do not count towards the Error Rate unless they conform to the Back-off Requirements.
- **"Financial Credit"** means the following for the Multi-Regional Storage class of Google Cloud Storage:

Monthly Uptime Percentage	Percentage of monthly bill for the Multi-Regional Storage class of Google Cloud Storage which does not meet SLO that will be credited to future monthly bills of Customer
99.0% – < 99.95%	10%
95.0% – < 99.0%	25%
< 95.0%	50%

- **"Financial Credit"** means the following for Regional Storage class of Google Cloud Storage:

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.

Monthly Uptime Percentage	Percentage of monthly bill for the Regional Storage class of Google Cloud Storage which does not meet SLO that will be credited to future monthly bills of Customer
99.0% – < 99.9%	10%
95.0% – < 99.0%	25%
< 95.0%	50%

- **"Financial Credit"** means the following for the Nearline, Coldline and Durable Reduced Availability Storage classes of Google Cloud Storage:

Monthly Uptime Percentage	Percentage of monthly bill for the Nearline, Coldline or Durable Reduced Availability Storage classes of Google Cloud Storage which does not meet SLO that will be credited to future monthly bills of Customer
98.0% – < 99.0%	10%
95.0% – < 98.0%	25%
< 95.0%	50%

- **"Monthly Uptime Percentage"** means 100%, minus the average of Error Rates measured over each five minute period during a monthly billing cycle.
- **"Valid Requests"** are requests that conform to the Documentation, and that would normally result in a non-error response.

Customer Must Request Financial Credit

In order to receive any of the Financial Credits described above, Customer must notify Google technical support within thirty days from the time Customer becomes eligible to receive a Financial Credit. Failure to comply with this requirement will forfeit Customer's right to receive a Financial Credit.

Maximum Financial Credit

The aggregate maximum number of Financial Credits to be issued by Google to Customer in a single billing month will not exceed 50% of the amount due by Customer for the applicable Covered Service for the applicable month. Financial Credits will be made in the form of a monetary credit applied to future use of the Service and will be applied within 60 days after the Financial Credit was requested.

SLA Exclusions

The SLA does not apply to any: (a) features or Services designated Alpha or Beta (unless otherwise set forth in the associated Documentation), (b) features or Services excluded from the SLA (in the associated Documentation) or (c) errors: (i) caused by factors outside of Google's reasonable control; (ii) that resulted from Customer's software or hardware or third party software or hardware, or both; (iii) that resulted from abuses or other behaviors that violate the Agreement; or (iv) that resulted from quotas listed in the Admin Console.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

8 IBM AND SONY CRAM UP TO 330 TERABYTES INTO TINY TAPE CARTRIDGE

Sputtered magnetic layer, lubricant, and new heads enable massive 200Gb/inch density.
SEBASTIAN ANTHONY (UK) - 8/2/2017, 8:00 AM

IBM and Sony have developed a new magnetic tape system capable of storing 201 gigabits of data per square inch, for a max theoretical capacity of 330 terabytes in a single palm-sized cartridge.

For comparison, the world's largest hard drives—which are about twice the physical size of a Sony tape cartridge—are the 60TB Seagate SSD or 12TB HGST helium-filled HDD. The largest commercially available tapes only store 15TB. So, 330TB is quite a lot.

To achieve such a dramatic increase in areal density, Sony and IBM tackled different parts of the problem: Sony developed a new type of tape that has a higher density of magnetic recording sites, and IBM Research worked on new heads and signal processing tech to actually read and extract data from those nanometre-long patches of magnetism.

IBM's Tale of the Tape
More than 60 years of tape innovation

	2006	2010	2014	2015	2017
Aerial Density (bits per sq inch)	6.67 Billion	29.5 Billion	85.9 Billion	123 Billion	201 Billion
Cartridge Capacity (Terabytes)	8	35	154	220	330
# of Books Stored	8 Million	35 Million	154 Million	220 Million	330 Million
Track Width	1.5 µm	0.45 µm	0.177 µm	0.140 µm	103 nm
Linear Density (bits per inch)	400'000	518'000	600'000	680'000	818'000
Tape Material	Barium Ferrite	Barium Ferrite	Barium Ferrite	Barium Ferrite	Sputtered Media
Tape Thickness (micrometers)	6.1	5.9	4.3	4.3	4.7
Tape Length (meters)	890	917	1255	1255	1098

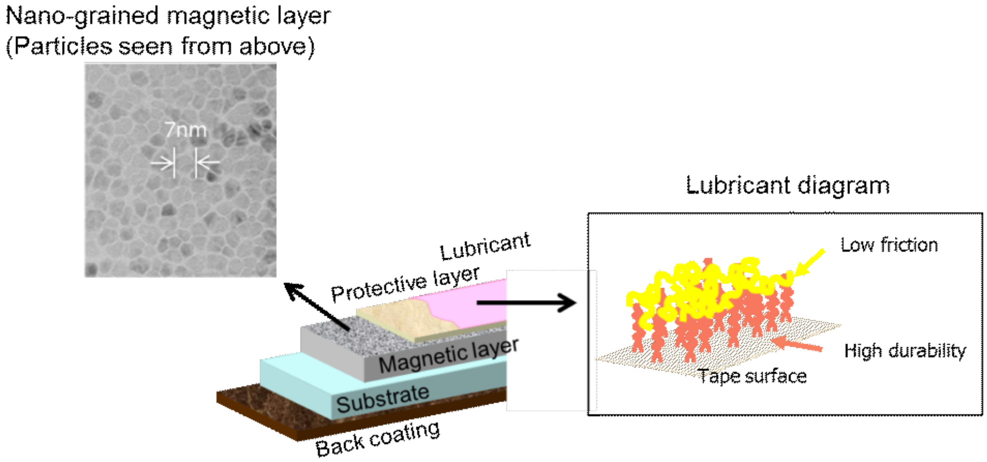
#5thtaperecord

A quick rundown of IBM's various tape storage density records. Note that commercial tape cartridges max out at 15TB—so, less than the theoretical amount enabled by the 2010 breakthrough. - Source: IBM Research

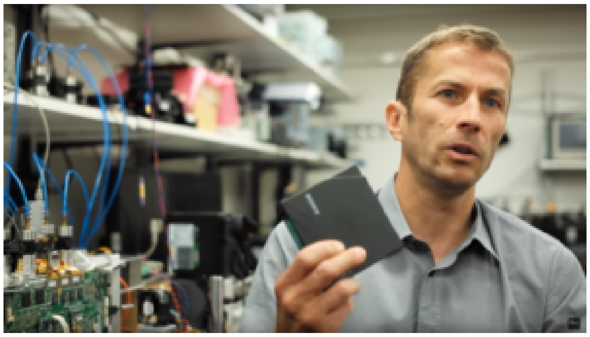
Sony's new tape is underpinned by two novel technologies: an improved built-in lubricant layer, which keeps it running smoothly through the machine, and a new type of magnetic layer. Usually, a tape's magnetic layer is applied in liquid form, kind of like paint—which is one of the reasons that magnetic tape is so cheap and easy to produce in huge quantities. In this case, Sony has instead used sputter deposition, a mature technique that has been used by the semiconductor and hard drive industries for decades to lay down thin films.

The main upshot of sputtering—a cool process that you should probably read about—is that it produces magnetic tape with magnetic grains that are just a few nanometres across, rather than tens or hundreds of nanometres in the case of commercially available tape.

The new lubrication layer, which we don't know much about, makes sure that the tape streams out of the cartridge and through the machine extremely smoothly. Some of the biggest difficulties of tape recording and playback are managing friction and air resistance, which cause wear and tear and chaotic movements. When you're trying to read a magnetic site that is just 7nm across, with the tape whizzing by at almost 10 metres per second, even the smallest of movements can be massively problematic.



A close-up look at Sony's new magnetic tape. – Source: IBM Research


IBM Research Zurich's Mark Lantz, their tape storage expert.
Modern tape cartridges are small, just four inches across. Source: IBM Research

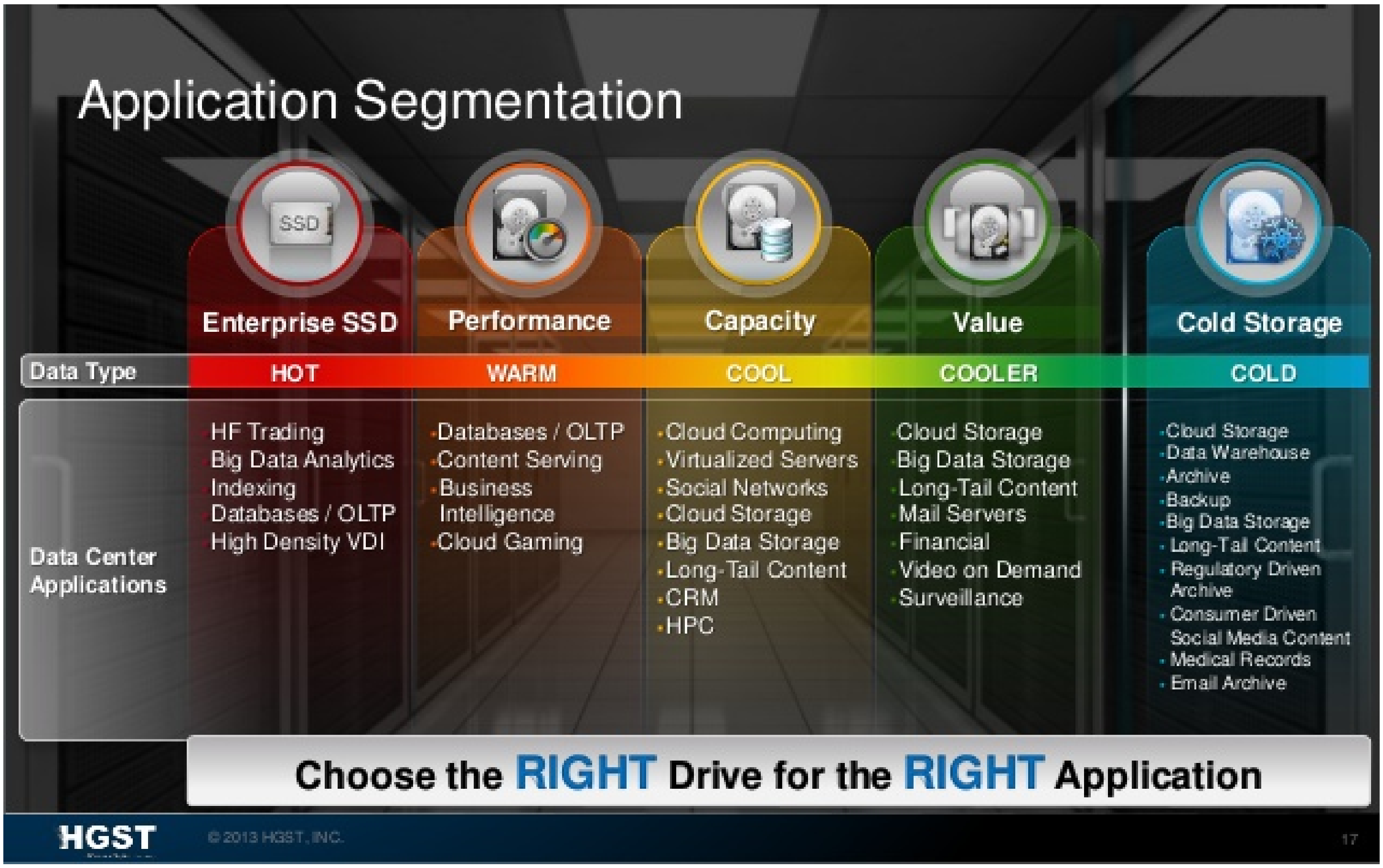
We know a little more about IBM's new read head, which appears to be a 48nm-wide tunneling magneto-resistive head that would usually be found in a hard disk drive—which makes sense, given the tape's sputtered medium is very similar to the surface of a hard drive platter. This new head, combined with new servo tech that precisely controls the flow of tape through the system, allows for a positional accuracy of under 7nm. A new signal processing algorithm helps the system make sense of the tiny magnetic fields that are being read by the head.

The new cartridges, when they're eventually commercialised, will be significantly more expensive because of the tape's complex manufacturing process. Likewise, a new tape drive (costing several thousand pounds) would be required. Still, given the massive increase in per-cartridge capacity, the companies that still use tape storage for backups and cold storage will be quite excited.

Some more details of IBM's side of the work are available in a paper published in a recent issue of *IEEE Transactions on Magnetics.* DOI: 10.1109/TMAG.2017.2727822.
This post originated on Ars Technica UK

9 WHY USE TAPE? (2017)

For the past decade, the tape industry has been re-architecting itself and the renaissance is well underway. Several new and important technologies for both LTO and enterprise tape products have yielded unprecedented cartridge capacity increases, much longer media life, improved bit error rates, and vastly superior economics compared to any previous tape or disk technology. This progress has enabled tape to effectively address many new data intensive market opportunities in addition to its traditional role as a backup device such as archive, big data, compliance, entertainment and surveillance. Clearly disk technology has been advancing, but the progress in tape has been even greater over the past 10 years. Today's modern tape technology is nothing like the tape of the past.



9.1 GROWTH IN TAPE

Demand for tape is being fueled by unrelenting data growth, significant technological advancements, tape's highly favourable economics, the growing requirements to maintain access to data 'forever' emanating from regulatory, compliance or governance requirements, and the big data demand for large amounts of data to be analysed and monetized in the future. The Digital Universe study suggests that the world's information is doubling every two years and much of this data is most cost-effectively stored on tape.

Enterprise tape has reached an unprecedented 10TB native capacity with data rates reaching 360MB/s. Enterprise tape libraries can scale beyond one exabyte. Enterprise tape manufacturers IBM and Oracle StorageTek have signaled future cartridge capacities far beyond 10TB with no limitations in sight. Open systems users can now store more than 300 Blu-ray quality movies with the LTO-6 2.5TB cartridge. In the future, an LTO-10 cartridge will hold over 14,400 Blu-ray movies. Nearly 250 million LTO tape cartridges have been shipped since the format's inception. This equals over 100,000PB of data protected and retained using LTO technology. The innovative active archive solution combining tape with low-cost NAS storage and LTFS is gaining momentum for open systems users.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

9.2 Recent Announcements and Milestones

Tape storage is addressing many new applications in today's modern data centres while offering welcome relief from constant IT budget pressures. Tape is also extending its reach to the cloud as a cost-effective deep archive service. In addition, numerous analyst studies confirm the TCO for tape is much lower than disk when it comes to backup and data archiving applications. See TCO studies section below.

- On Jan. 16, 2014 Fujifilm Recording Media USA, Inc. reported it has manufactured over 100 million LTO Ultrium data cartridges since its release of the first generation of LTO in 2000. This equates to over 53,000PB of storage and more than 41 million miles of tape, enough to wrap around the globe 1,653 times.
- April 30, 2014, Sony Corporation independently developed a soft magnetic under layer with a smooth interface using sputter deposition, created a nano-grained magnetic layer with fine magnetic particles and uniform crystalline orientation. This layer enabled Sony to successfully demonstrate the world's highest areal recording density for tape storage media of 148GB/in2. This areal density would make it possible to record more than 185TB of data per data cartridge.
- On May 19, 2014 Fujifilm in conjunction with IBM successfully demonstrated a record areal data density of 85.9 Gb/in² on linear magnetic particulate tape using Fujifilm's proprietary NANOCUBIC and Barium Ferrite (BaFe) particle technologies. This breakthrough in recording density equates to a standard LTO cartridge capable of storing up to 154TB of uncompressed data, making it 62 times greater than today's current LTO-6 cartridge capacity and projects a long and promising future for tape growth.
- On Sept. 10, 2014 the LTO Program Technology Provider Companies (TPCs) HP, IBM and Quantum, announced an extended roadmap which now includes LTO generations 9 and 10. The new generation guidelines call for compressed capacities of 62.5TB for LTO-9 and 120TB for generation LTO-10 and include compressed transfer rates of up to 1,770MB/s for LTO-9 and a 2,750MB/s for LTO-10. Each new generation will include read-and-write backwards compatibility with the prior generation as well as read compatibility with cartridges from two generations prior to protect investments and ease tape conversion and implementation.
- On Oct. 6, 2014 IBM announced the TS1150 enterprise drive. Features of the TS1150 include a native data rate of up to 360MB/s versus the 250MB/s native data rate of the predecessor TS1140 and a native cartridge capacity of 10TB compared to 4TB on the TS1140. LTFS support was included.

9.3 Technology Innovations Fuel Tape's Future

Development and manufacturing investment in tape library, drive, media and management software has effectively addressed the constant demand for improved reliability, higher capacity, power efficiency, ease of use and the lowest cost per GB of any storage solution.

9.4 Summary of tape's value proposition followed by key metrics for each:

- Tape drive reliability has surpassed disk drive reliability
- Tape cartridge capacity (native) growth is on an unprecedented trajectory
- Tape has a faster device data rate than disk
- Tape has a much longer media life than any other digital storage medium
- Tape's functionality and ease of use is now greatly enhanced with LTFS
- Tape requires less energy consumption than any other digital storage technology
- Tape storage has a much lower acquisition cost and TCO than disk

9.5 Reliability

Tape reliability levels have surpassed HDDs. Reliability levels for tape exceeds that of the most reliable disk drives by one to three orders of magnitude. The BER (Bit Error Rate – bits read per hard error) for enterprise tape is rated at 1×1019 and 1×1017 for LTO tape. This compares to 1×1016 for the most reliable enterprise FC disk drive.

9.6 Capacity and Data Rate

LTO-6 cartridges provide 2.5TB capacity and more than double the compressed capacity of the preceding LTO-5 drive with a 14% data rate performance boost to 160MB/s. Enterprise tape has reached 8.5TB native capacity and 252MB/s on the Oracle StorageTek T10000D and 10TB native capacity and 360MB/s on the IBM TS1150. Tape cartridge capacities are expected to grow at unprecedented rates for the foreseeable future.

9.7 Media Life

Manufacturers specifications indicate that enterprise and LTO tape media has a life span of 30 years or more while the average tape drive will be deployed 7 to 10 years before replacement. By comparison, the average disk drive is operational 3 to 5 years before replacement.

9.8 LTFS Changes Rules for Tape Access

Compared to previous proprietary solutions, LTFS is an open tape format that stores files in application-independent, self-describing fashion, enabling the simple interchange of post_content across multiple platforms and workflows. LTFS is also being deployed in several innovative 'Tape as NAS' active archive solutions that combine the cost benefits of tape with the ease of use and fast access times of NAS. The SNIA LTFS Technical Working Group has been formed to broaden cross-industry collaboration and continued technical development of the LTFS specification.

9.9 TCO Studies

Tape's widening cost advantage compared to other storage mediums makes it the most cost effective technology for long-term data retention. The favourable economics (TCO, low energy consumption, reduced raised floor) and massive scalability have made tape the preferred medium for managing vast volumes of data. Several tape TCO studies are publicly available and the results consistently confirm a significant TCO advantage for tape compared to disk solutions.

According to the Brad Johns Consulting Group, a TCO study for an LTFS-based 'Tape as NAS' solution totaled $1.1 million compared with $7.0 million for a disk-based unified storage solution. This equates to a savings of over $5.9 million over a 10-year period, which is more than 84% less than the equivalent amount for a storage system built on a 4TB HDD drive unified storage system. From a slightly different perspective, this is a TCO savings of over $2,900/TB of data. (Source: Johns, B. A New Approach to Lowering the Cost of Storing File Archive Information, Brad Johns Consulting Group, April 2013)

Another comprehensive TCO study by ESG (Enterprise Strategies Group) comparing an LTO-5 tape library system with a low-cost SATA disk system for backup using de-duplication (best case for disk) shows that disk deduplication has a 2-4x higher TCO than the tape system for backup over a 5-year period. The study revealed that disk has a TCO of 15x higher than tape for long-term data archiving. (Source: A Comparative TCO Study: VTLs and Physical Tape, by ESG)

9.10 Select Case Studies Highlight Tape and Active Archive Solutions

CyArk (case study) is a non-profit foundation focused on the digital preservation of cultural heritage sites including places such as Mt. Rushmore, and Pompeii. It predicted that their data archive would grow by 30% each year for the foreseeable future reaching one to two petabytes in five years. They needed a storage solution that was secure, scalable, and more cost-effective to provide the longevity required for these important historical assets. To meet this challenge CyArk implemented an active archive solution featuring LTO and LTFS technologies.

Dream Works Animation (case study) a global Computer Graphic (CG) animation studio has implemented a reliable, cost-effective and scalable active archive solution to safeguard a 2PB portfolio of finished movies and graphics, supporting a long-term asset preservation strategy. The studio's comprehensive, tiered and converged active archive architecture, which spans software, disk and tape, saves the company time, money and reduces risk.

McDonald 's primary challenge (case study) was to create a digital video workflow that streamlines the management and distribution of their global video assets for their video production and post-production environment. McDonald's implemented the Spectra Logic Corp.'s T200 tape library with LTO-6 providing 250TB of video production storage. Nightly, incremental backup jobs store their media assets into separate disk and LTO-6 storage pools for easy backup, tracking and fast retrieval. This system design allows McDonald's to effectively separate and manage their assets through the use of customized automation and data service policies.

NCSA (case study) employs an Active Archive solution providing 100% of the nearline storage for the NCSA Blue Waters supercomputers, which is one of the world's largest active file repositories stored on high capacity, highly reliable enterprise tape media. Using an active archive system along with enterprise tape and RAIT (Redundant Arrays of Inexpensive Tape) eliminates the need to duplicate tape data, which has led to cost savings.

The oil industry has used tape for seismic acquisition for decades, storing critical field tapes in long term storage. The TS1150 is the latest IBM Magstar technology that continues serving this demanding industry.

9.11 Looking Ahead and Beyond

The role tape serves in today's modern data centres is expanding as IT executives and cloud service providers address new applications for tape that leverage its significant operational and cost advantages. This recognition is driving investment in new tape technologies and innovations with extended roadmaps, and it is expanding tape's profile from its historical role in backup to one that includes long-term archiving requiring cost-effective access to enormous quantities of stored data. Given the current and future trajectory of tape technology, data intensive markets such as big data, broadcast and entertainment, archive, scientific research, oil and gas exploration, surveillance, cloud, and HPC are expected to become significant beneficiaries of tape's continued progress. Clearly the tremendous innovation, compelling value proposition and development activities demonstrate tape technology is not sitting still; expect this promising trend to continue in 2015 and beyond.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

10 KEEPING DATA FOR A LONG TIME

Tom Coughlin , *FORBES CONTRIBUTOR*
June 29, 2014
Opinions expressed by Forbes Contributors are their own.

Keeping information for a long time has always been a challenge. Thermodynamics doesn't favor information lasting a long time and so to make that happen people have to spend effort and energy. Deciding how to create a long-term archive involves choosing the right storage system with the right technology under the proper environmental conditions. This can be combined with migration and replication practices to improve the odds of keeping content useful and accessible for an extended period of time. Note that Coughlin Associates organized the 2014 Creative Storage Conference.

At the 2014 IEEE Massive Storage Systems and Technology (MSST) Conference and at the 2014 Creative Storage Conference there were a number of interesting talks about digital storage for long term archiving and preservation. The storage technologies examined included flash memory, hard disk drives, magnetic tape and optical discs. These are the most common current digital storage technologies on the market and thus are readily available for storage system architects. Lets look at these technologies and some applications.

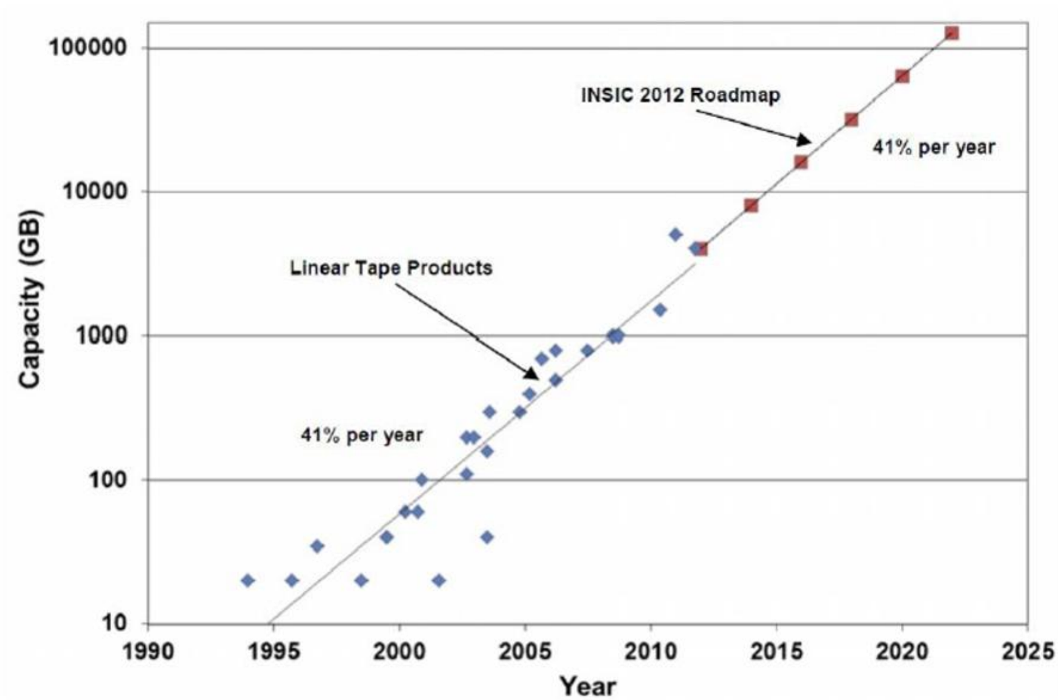
First consider flash memory in archiving. At the 2013 Flash Memory Summit Jason Taylor from Facebook, in a keynote talk, presented the idea of using really bad low endurance flash memory for a cold storage archive. According to Marty Czekalski of Seagate at the MSST conference flash writing is best done at elevated temperatures while data retention and data disturb favor storage at lower temperatures. The JEDEC JESD218A endurance specification states that if flash power off temperature is at 25 degrees C then retention is 101 weeks—that isn't quite 2 years. So it appears conventional flash memory may not have good media archive life and should only be used for storing transitory data.

Cold storage should also be inexpensive, but the demand for flash memory to fill consumer and other computer application combined with the cost of building flash memory factories will make storing long term data on flash relatively expensive. However Bruce Moxon from Samsung pointed out that flash memory may play an important role in an overall active archive where data is available for access since flash caching can help meet access requirements and flash memory can also be use for object caching of keys and metadata. Thus flash memory may not be suitable for the main storage in an archive system but for an active archive, flash memory can provide better overall system access to content kept in a lower cost media storage library.

Hard disk drives are often used in active archives because hard disk drive arrays can be continually connected to the storage network, allowing relatively rapid access to content. Hard disk drive active archives can also be combined with flash memory to provide better overall system performance. However hard disk drives do not last forever. They can wear out with continued use and even if the power is turned off the data in the hard disk drive will eventually decay due to thermal erasure (again we run into the enemy of data retention, thermodynamics). In practice hard disk drive arrays have built in redundancy and data scrubbing to help retain data for a long period. It is probably good advice to assume that HDDs in an active archive will last only 3-5 years and will need to be replaced over time. Using commodity HDDs, open source software and commodity computer components companies such as BackBlaze report that they build a 180 TB HDD system for $9,300. That comes to $51.67 per TB or about $0.05 per GB.

Less active archives where data is stored for longer periods of time will be interested in storage media that can retain the information stored on them for an extended period of time. There are two common digital storage media that are used for long term cold storage applications. They are magnetic tape and optical discs. Lets look at these two storage technologies and compare them for long term cold storage applications.

Magnetic tapes used for archiving come in half-inch tape cartridges. The popular formats used to day are the LTO format supported by the Ultrium LTO Program, the T10000 series tapes from Oracle/StorageTek and the TS series enterprise tapes from IBM. Modern magnetic tapes have a storage life under low temperature/humidity storage conditions and low usage of several decades and currently native storage capacities per cartridges as high as 8.5 TB. iNSIC projections from 2012 indicate that we could see 100 TB+ magnetic tapes sometime after 2020.

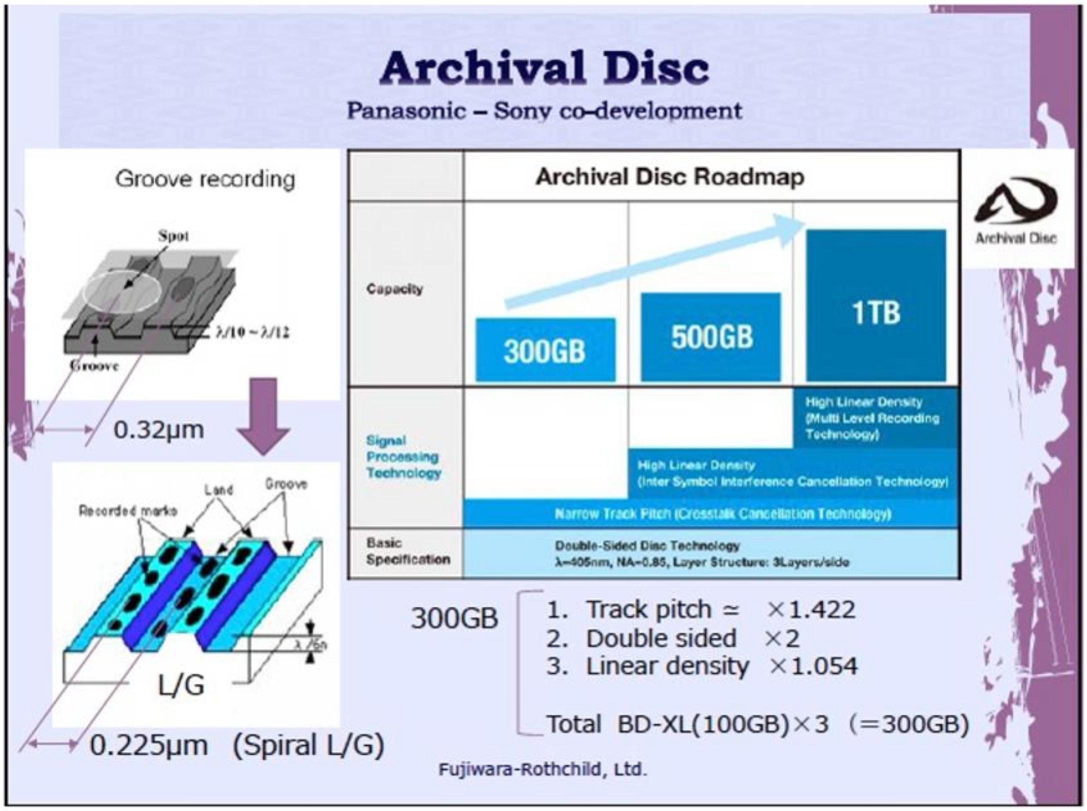


When not actively being written or read, magnetic tape cartridges can sit in a library system consuming no power. Digital magnetic tape is thus a good candidate for long-term data retention and has a long history of use in many industries for this application. According to Chris Kehoe at the Creative Storage Conference, magnetic tape archiving can cost as low as $0.003/GB/month or $0.04/GB/year.

Optical storage has also been used for long-term data retention and environmental stress tests indicate that the latest generation of Blu-ray optical media should have an expected life-time of at least several decades. At the Creative Storage Conference Bill Cubellis from Sony said that properly made archival grade optical discs should have a shelf life of 50 years. At the Open Compute Project Summit in January 2014 Facebook presented a 1 PB Blu-ray disc storage system prototype with 10,000 discs. Facebook estimated that this system would reduce the storage costs by 50% and the energy consumption by 80% compared to their current cold storage system (probably HDD based).

Ken Wood from Hitachi Data Systems at the MSST Conference presented data that the migration/remastering costs for 5 PB of content over 75 years is much less for an optical system with the media replaced every 50 years rather than more frequent tape and HDD replacement (however I wonder if a 50 year replacement rate makes sense considering the advantages in storage footprint with ever evolving more dense storage capacities).

A lot of digital data has persistent value and so long term retention of that data is very important. In an Oracle talk at MSST they estimated that storage for archiving and retention is currently a $3B market growing to over $7B by 2017. Several storage technologies can play a role in an archive system depending upon the level of activity expected in the archive. Flash memory can provide caching of frequently used or anticipated content to speed retrieval times while HDDs are often used for data that is relatively frequently accessed.



Magnetic tape and optical disks provide low cost long-term inactive storage with additional latency for data access vs. HDDs due to the time to mount the media in a drive. Thus depending upon the access requirements for an archive it may be most effective to combine two or even three technologies to get the right balance of performance and storage costs. As the total content that we keep increases these considerations will become more important drive new generations of storage technologies geared toward protecting valuable content and bringing it to the future.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.

11 Why Tape is Poised for a George Foreman-Like Comeback

Source: Wikibon Blog

Posted by David Vellante in Compliance, Data Protection, Storage, Wikibon on June 24, 2014

11.1 Tape is Dead, Not!

The combination of tape and flash will yield much better performance and substantially lower cost than spinning disk. This statement will prove true for long-term data retention use cases storing large data objects. The implications of this forecast are: 1) Tape is relevant in this age of Big Data; 2) Certain tape markets may actually show growth again; 3) Spinning disk is getting squeezed from the top by flash and from below by a disk/tape mashup we call "flape."

11.2 Spinning Disk: Slow and Getting Slower

For decades we've heard that the amount of capacity under a disk actuator will increasingly be problematic for application performance. Specifically, as disk capacities increase at the rate prescribed by Moore's Law, placing more data under a mechanical arm will serve to relentlessly decelerate performance and eventually the spinning rust business will hit a wall. Guess what? The predictions were right. We're finally at the breaking point. For a long time the storage industry somewhat masked disk performance problems by spinning the platters faster, increasing track capacities, using larger caches, writing better algorithms, creating massive backend disk farms, wide striping data and doing unnatural acts like short-stroking devices. More recently the industry began to jam flash into legacy storage arrays, which breathed new life into disk performance. For a while…But the end is near for so-called "high performance" spinning disk. Ironically, the very flash technology that has extended the runway of legacy SAN is the reason why tape, for certain applications, could grow again. There are three main reasons that explain this contrarian view:

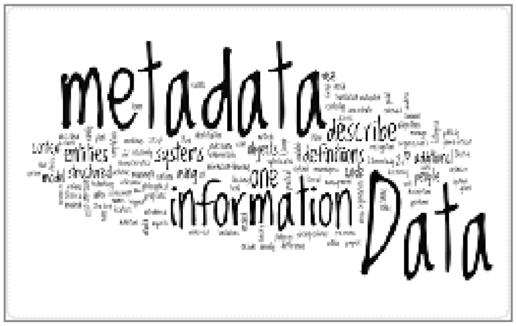
1. The number of disk vendors is down to the single digits. In the 1980's there were more than 70 disk drive players…now we have three or four. The investment going into mechanical disk drives simply isn't there any more.
2. Head technology investments, the lynchpin of disk technology for years, are in a long, slow, managed decline cycle. Head technology used to be a strategic advantage. It was a main reason Al Shugart bought Control Data's Imprimis disk division in 1989. In 2014, the ROI of innovating in disk head technology is minimal.
3. The days of the so-called high-speed spinning disk are numbered. High performance spinning disk is an oxymoron. Flash will replace high spin speed disks within a year. Flash is cheaper and now more economical than the segment formerly known as high-performance disk drives.

The lack of investments in spinning disk tech, the fact that spin speeds have hit their physical limits and the reality that track capacities aren't increasing that fast, combine to give you slow as molasses disk bandwidth. To be clear, we're talking about the speed at which data comes off the disk internally. While time-to-first-byte is faster than tape, because disk is a random-access device…once you find the data, if it's a large object like a video file or archived email blob, it takes a long time to get data off the disk. This is not the case for tape.

Tape heads aren't housed in a hermetically sealed disk unit. Their tolerances are much less stringent than disk heads. Tape heads are fixed…a stationary component that can more easily be replicated and staggered.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

11.3 METADATA IS THE KEY TO TAPE'S FUTURE



Today, metadata, the data about the data (i.e. what files live where), is locked inside the contents of an individual tape cartridge. If you take all the metadata that's locked on tapes and surface that onto a flash layer as a front-end to tape – Wikibon calls this flape – you'll get way lower costs and much better performance than with a spinning disk system…even a disk system with flash. It's happening today. Wikibon practitioners (several in media and entertainment) are writing metadata taxonomies and surfacing metadata to a high-speed layer. Many still use "fast" disk for that layer but soon they will be using flash. This infrastructure is connected to an application server, which writes both metadata and data directly to the high-speed layer. The data is then grouped into objects that are categorized based on the taxonomy and trickled asynchronously to the backend tape for cold storage. These customers are also writing algorithms that fetch data intelligently based on the data request. For example, if the fifth request for a piece of data in a chain can be serviced more quickly than the first, they'll complete the fifth request while in parallel servicing the first and then re-order the chain at the backend.

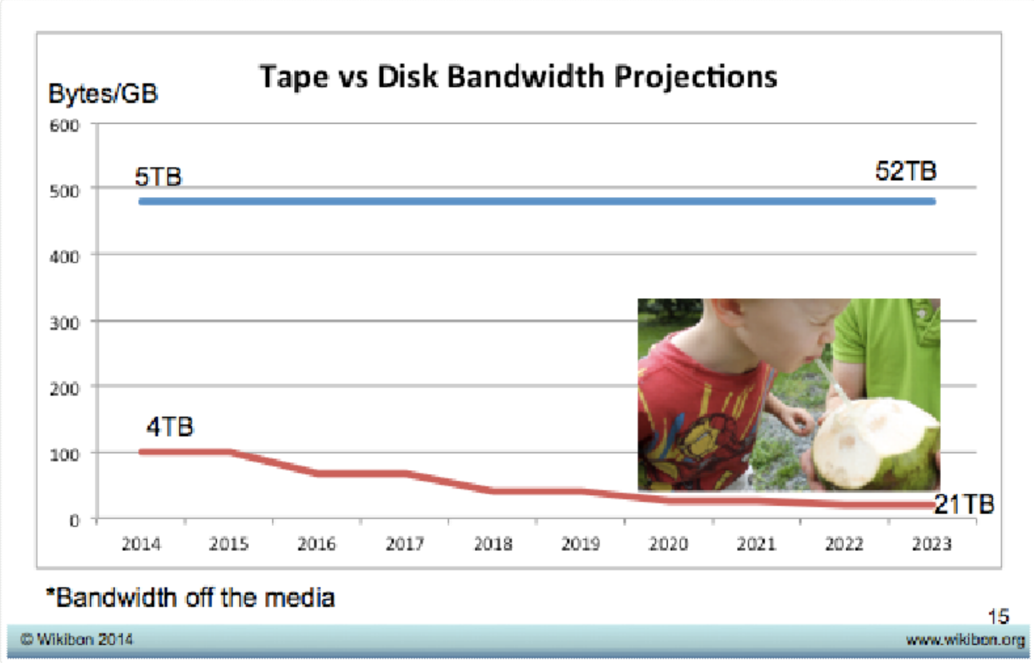
These future flape architectures are delivering new business value for customers because they're able to monetize complex and previously unattainable (in a decent amount of time) seek requests for information. An example might be give me all the Lady Gaga video clips where she performed live; with Crayon Pop opening for her…between 2011 and 2012. Think about applying this to facial recognition applications or email archive blobs or any large object data set.

The spinning disk cartel will cringe at the premise that tape is faster/cheaper/better. They'll make faces and tell you why this is nonsense. They will tell you tape is not cool and you are not cool if you use tape. Don't just trust their skepticism. Do the research yourself and come to your own conclusions. You may find that you can drive significant value for your organization.

David Floyer explains all the gory details and technology assumptions in this Research Note. Read it carefully and think about what flape could do for your business.

11.4 TAPE VERSUS DISK PERFORMANCE

Let's look at performance more closely. How can tape be faster than disk? Tape is a serial medium whereas disk is a random access device. Disk must be faster. Well this turns out not to always be true – especially where bandwidth is the primary measure of performance – as it is for large objects (e.g. audio, video, facial recognition, large email blobs, etc.). The slide below shows a comparison of disk versus tape bandwidth– specifically the internal bandwidth of the device (normalized by capacity). Key takeaways from this data include:

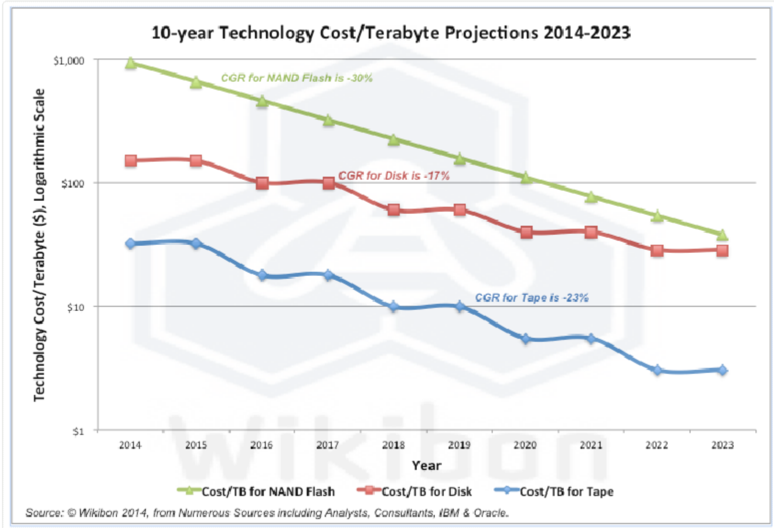


- Tape bandwidth is nearly 5X that of disk.
- The tape blue line stays flat over time, while the disk red line declines dramatically.
- The implication is that the time it takes to scan a 5TB tape cartridge today will be the same as it takes to scan a 50+TB cartridge in 2023.
- See the kid sucking coconut milk from the straw? The disk straw is tiny and really isn't getting bigger.

The fact is disk is slower than tape when large files are involved because the internal bandwidth of disk devices are limited, especially as capacity grows. As implied in the points above, the time it takes (for example) to rebuild a disk drive on a failure is escalating with every new generation of capacity increase. To rebuild a 21TB disk could take more than a month—which by the way is why no one should use RAID 5 anymore because during a rebuild you're exposed for far too long, risking a second drive failure and data loss. Another reason tape can be faster is file systems like Linear Tape File System (LTFS) and SAM-QFS. Tape file systems allow separation of metadata in a self-describing tape format providing direct access to file content and metadata, independent of any external database or storage system. This capability presents tape in a standard file system view of the data stored making accessing files stored on the tape, logically similar to accessing files stored on disk. LFTS was introduced by IBM and SAM-QFS by Sun Microsystems. Oracle is now the steward of the latter format. We believe these types of innovations, combined with flash will breathe new life into tape and solve a nagging user problem – how to contain data growth, cost effectively.

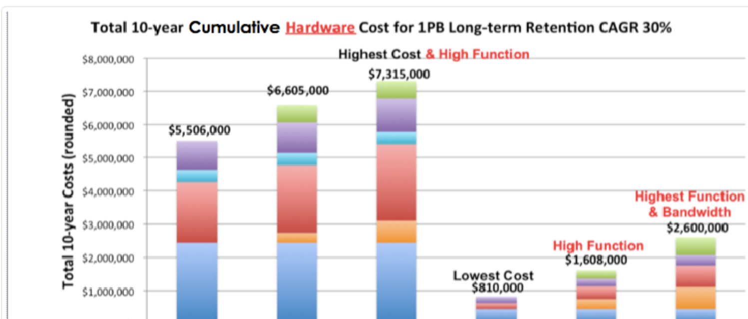
11.5 What about Costs?

Is the demise of tape greatly exaggerated? We believe yes. There are two reasons in addition to the performance scenario put forth above that tape continues to be around, including: 1) Tape is cheaper than disk for long term retention; and 2) As a last resort disaster recovery medium, tape is still the most cost effective (and fastest) way to move data from point A to point B. We call this, "CTAM," the Chevy Truck Access Method. When it comes to compliance for DR, tape still is viable. From a cost perspective, HDDs have been unable to keep pace with the areal density curve of tape. As an example, overall $/TB declines for disk are forecast at roughly 17% per annum whereas tape is tracking at a 23% decline. Tape of course is starting at a much lower cost per bit than disk and is likely to remain an order of magnitude cheaper. Flash meanwhile is on the steepest price decline curve of the three storage technologies and is expected to continue to close the gap on disk.



11.6 Quantifying the Flape Effect

We believe a new architecture called "Flape" will emerge, combining flash storage and tape. The data below show an economic model developed by Wikibon around flape as compared with alternative spinning disk architectures. Notably, the best value is flape with a large flash metadata layer. That scenario will be 3-5 times faster (for large objects) than disk-based alternatives.



	Disk	Disk + Small Flash	Disk + Large Flash	Tape	Flape (Small Flash)	Flape (Large Flash)
Metadata Management Software	$-	$532,000	$532,000	$-	$249,000	$532,000
Power/Space	$871,000	$930,000	$1,007,000	$195,000	$236,000	$312,000
Technology Migration	$368,000	$368,000	$368,000	$-	$-	$-
Hardware Maintenance	$1,830,000	$2,037,000	$2,305,000	$178,000	$385,000	$653,000
NAND Hardware	$-	$301,000	$666,000	$-	$301,000	$666,000
Magnetic Hardware	$2,437,000	$2,437,000	$2,437,000	$437,000	$437,000	$437,000
Total Hardware	$5,506,000	$6,605,000	$7,315,000	$810,000	$1,608,000	$2,600,000

Source: © Wikibon 2014, from Numerous Sources including Analysts, Consultants, IBM & Oracle.

17

© Wikibon 2014 www.wikibon.org

What about smaller files? Flape may still be the way to go in such use cases because archiving software is often able to group smaller files into larger objects. Today that is done on disk-based systems but there's no reason it couldn't be done on flash with some backend integration to tape. With flape, data and metadata would be written to the flash layer and the data moved asynchronously to tape. There are a few items to consider as one thinks about a flape architecture. First, is it realistic that every application or data set can reside on either flash/SSD or tape? Second, can the application that may need to leverage the data that resides on tape, easily access the data that lives on the tape? This is why a middleware layer is needed to be able to find the data in an "appropriate" amount of time. Flape architectures must be

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

designed to read metadata that lives on the high-speed layer. This metadata informs the system as to where the data lives and can retrieve it in the most efficient timeframe. As long as an application can work in this fashion, a flape architecture can help any organization achieve the right balance between performance and cost which in turn helps to drive business value.

11.7 Who Will Lead the Flape Revolution?

Our view is that in order to scale, flape needs leadership that can set standards, point the way for customers to develop metadata taxonomies, provide middleware technology to exploit flape and finally entice ISVs to recognize flape as a viable platform. Today, customers are leading the innovation charge. Wikibon spoke with several IT practitioners who have developed early instantiations of flape using a high spin speed disk metadata layer. As indicated, it’s only a matter of time before this becomes flash-based storage. These customers are writing metadata / data classification taxonomies and automating the placement of metadata on the high- speed layer at the point of data creation. They’ve also written request optimization algorithms as described earlier. In order for flape to be adopted widely however, the vendor community must step up to the plate.

In our view, the two companies best positioned to execute on flape are IBM and Oracle. HP could also be a player, but Oracle in particular, with its StorageTek heritage, could emerge as an early innovator. This is not to say that IBM could not participate in this trend. In fact we believe IBM will. Both IBM and Oracle sell flash storage products, IBM with the TMS acquisition and Oracle with the ZFS line of flash-first arrays. In addition, several Oracle customers we spoke with are focusing the ZFS appliance on backup applications as an alternative to NetApp targets. Meanwhile, IBM with its Tivoli line of management software has a background in this market. We believe both companies are actively working on flape-like products. Clearly the market will be much better off with two suppliers and ironically, we feel IBM and STK will further collaborate on tape standards to accelerate market adoption. It would be in their mutual interests to do so as market leader EMC is vocal in its marketing about the death of tape.

Like the famous boxer who once lost his luster, tape is re-inventing itself and can become a prominent mainstream player again. Tape is no longer a good backup medium. It has been re-positioned for long-term retention. Tape’s superior economics relative to disk and its better performance for large files (when combined with flash) make it attractive for retention and big data repository apps. Not to mention that tape lasts longer. The bit error rates for tape are two orders of magnitude better than spinning disk – meaning tape is a much more suitable platform for long-term storage. If the industry steps up and buyers keep an open mind, flape is a winner that will cut costs for the “bit bucket,” improve archiving performance for large object workloads and deliver substantial incremental business value.

12 JUST TWO LTO TAPE CARTRIDGE MANUFACTURERS REMAINING: FUJIFILM AND SONY

Maxell and TDK quitting
By Jean Jacques Maleval on 2014.05.14

Two Japanese companies decided last months to stop manufacturing tape cartridge media (especially LTO): Maxell and TDK.

Maxell no more produces magnetic tapes representing 21% of global sales in FY08 and 7% in FY13. It also stopped its optical media business.

TDK withdraws LTO tape during the first quarter of 2014. That's bad news for partner Imation, no more manufacturer of tapes and reseller of TDK media. But TDK continues to be in storage producing disk heads for HDDs.

So now just two manufacturers remain on the market, Fujifilm and Sony, OEMs of IBM and HP respectively according to US source.

In the past other computer tape media makers include 3M, Ampex/Quantegy, BASF/Emtec, Graham Magnetics, Kodak/Verbatim, Memorex, Pyral, Rhone-Poulenc Systems.

Magnetic technology business is really shrinking with now only two in tape media, three in tape drives (HP, IBM, Oracle StorageTek) and three in HDDs (Seagate, Toshiba, Western Digital).

As of 2017.09.15 Magnetic [data storage] business is growing rather than shrinking. Also, these are the three LTO drive manufacturers: IBM, HPE, and Quantum.

Read also:
TDK to Withdraw From LTO Tape Business
Media Technology Corporation to be dissolved
LTO-6 Cartridges Available From Maxell
Following Fujifilm, Sony and TDK (Imation)

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

13 MEMRISTORS

Source: Memristors.org

What is a memristor? Memristors are basically a fourth class of electrical circuit, joining the resistor, the capacitor, and the inductor, that exhibit their unique properties primarily at the nanoscale. Theoretically, Memristors, a concatenation of “memory resistors”, are a type of passive circuit elements that maintain a relationship between the time integrals of current and voltage across a two terminal element. Thus, a memristors resistance varies according to a devices memristance function, allowing, via tiny read charges, access to a “history” of applied voltage. The material implementation of memristive effects can be determined in part by the presence of hysteresis (an accelerating rate of change as an object moves from one state to another) which, like many other non-linear “anomalies” in contemporary circuit theory, turns out to be less an anomaly than a fundamental property of passive circuitry.

Until recently, when HP Labs under Stanley Williams developed the first stable prototype, memristance as a property of a known material was nearly nonexistent. The memristance effect at non-nanoscale distances is dwarfed by other electronic and field effects, until scales and materials that are nanometers in size are utilized. At the nanoscale, such properties have even been observed in action prior to the HP Lab prototypes.

But beyond the physics of electrical engineering, they are a reconceptualizing of passive electronic circuit theory first proposed in 1971 by the nonlinear circuit theorist Leon Chua. What Leon Chua, a UC Berkeley Professor, contended in his 1971 paper Transactions on Circuit Theory, is that the fundamental relationship in passive circuitry was not between voltage and charge as assumed, but between changes-in-voltage, or flux, and charge. Chua has stated: “The situation is analogous to what is called “Aristotle’s Law of Motion, which was wrong, because he said that force must be proportional to velocity. That misled people for 2000 years until Newton came along and pointed out that Aristotle was using the wrong variables. Newton said that force is proportional to acceleration–the change in velocity. This is exactly the situation with electronic circuit theory today. All electronic textbooks have been teaching using the wrong variables–voltage and charge–explaining away inaccuracies as anomalies. What they should have been teaching is the relationship between changes in voltage, or flux, and charge.”

As memristors develop, its going to come down to, in part, who can come up with the best material implementation. Currently IBM, Hewlett Packard, HRL, Samsung and many other research labs seem to be hovering around the titanium dioxide memristor, but there are quite a few other types of memristorswith vectors of inquiry.

14 RACETRACK MEMORY

Racetrack Memory - Ferroelectric Domain Wall Memory Shows Its Promise

By Dexter Johnson
Posted 26 Jun 2017 | 20:00 GMT

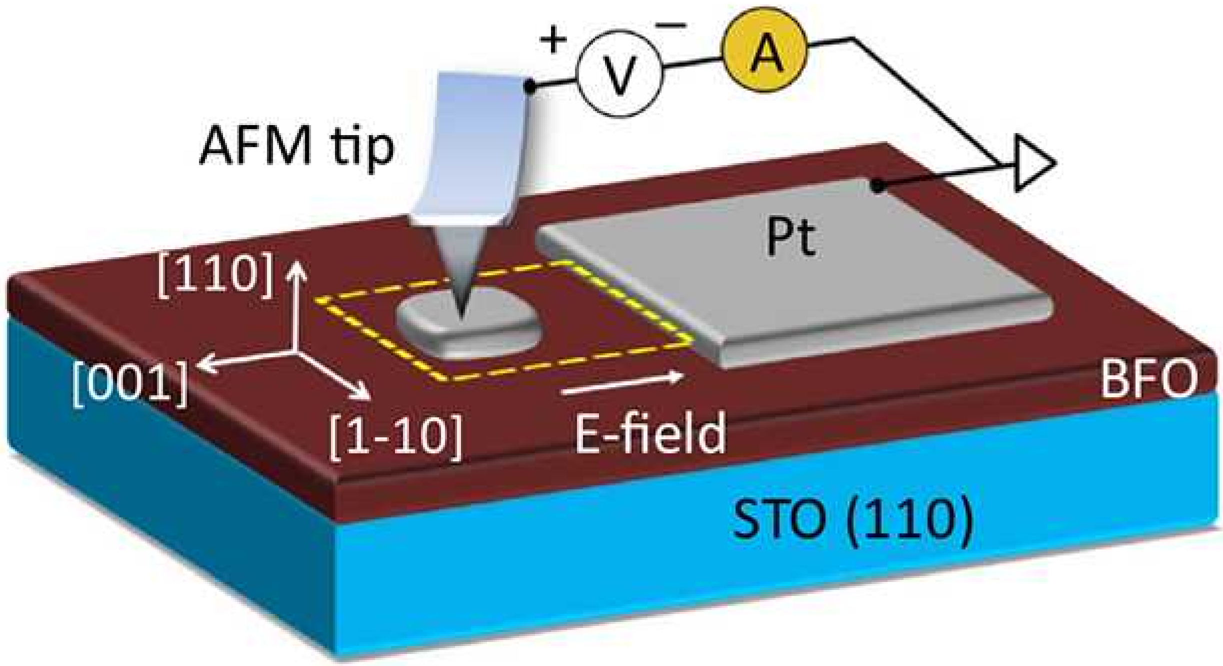


Illustration: Science Advances

Researchers at the University of New South Wales (UNSW) in Australia have taken a significant step in the development of so-called ferroelectric domain wall (FEDW) memories. These results could lead to a non-volatile memory with a higher storage density than traditional memory devices and be the realization of the unfulfilled promise presented by magnetic domain wall memory, also known as "racetrack memory".

Nearly a decade ago, Stuart Parkin and his colleagues at IBM Almaden Research demonstrated a three-bit version of so-called "racetrack memory." Racetrack memory is a solid-state non-volatile memory that promised much higher storage density than conventional solid-state memory devices.

It was expected to be a truly disruptive alternative memory technology, but it turned out to be about a thousand times slower than anticipated, putting it on par with the performance of traditional hard drives.

While racetrack memories are based on nanowires made from ferromagnetic materials that can have their magnetization switched through an applied magnetic field, FEDW memories are made from ferroelectric materials whose electric field polarization is changed via an electric field.

In research described in the journal *Science Advances*, the UNSW scientists fabricated a memory device from these ferroelectric materials and have demonstrated that it can operate at relatively low voltages (less than 3 volts). It also has a relatively high OFF-ON ratio, making it easier to assign a "0" and "1" to distinct ferroelectric states.

"The domain walls we investigate and use here are ferroelectric walls, not magnetic walls," said Jan Seidel, associate professor at UNSW and co-author of the research, in an e-mail interview with *IEEE Spectrum*. "While you have oriented spins in a magnet, you have oriented electric dipoles in a ferroelectric materials, so this is, in some sense, the electric equivalent of a ferromagnet."

One of the benefits of ferroelectric domain walls is that their width is anywhere from 10 to 100 times smaller than conventional magnetic walls. That makes them the ultimate nanoscale controllable feature in solid materials, according to Seidel. Because the walls are so thin, you can encounter structural changes and symmetry changes of the material, which can drastically alter the material's property.

"You can think of the wall potentially being a completely different material," explained Seidel. "This is the key point we use in our memory cells. The domain walls are electrically conductive, while the wall's surroundings (the bulk of the material) is insulating."

Because the electric dipoles (positive and negative charges) in ferroelectrics can be influenced and reoriented by external electric fields (applied voltage), walls can be created, erased or relocated in the material. This method of creating and erasing walls makes it possible to form or take away conductive channels in which data is stored.

Seidel and his UNSW colleagues fabricated their prototype FEDW device with bismuth ferrite (BFO) thin films that are grown on a strontium titanate (STO) substrate. The BFO thin films—at least the type produced here—possess a near mono-domain state, providing a clean slate for creating and erasing different domain walls.

To ensure that a steady current was applied to the device, the researchers placed nanofabricated platinum/titanium electrodes (a big one and a small one) on the same plane as the BFO thin film. This in-plane geometry of the electrodes makes it possible to encode and retrieve information through moderate electric fields rather than electric currents, leading to low-energy operation.

"A key point in our research is that we have found a way to stabilize the domain walls on the nanoscale by using the right electrode geometry tailored to the specific materials properties at hand," explained Seidel.

In this configuration, the large electrode is grounded, while the small electrode serves as the target for the tip of a conductive atomic force microscope. This provides the steady current that triggers a specific FEDW configuration. This design is a major departure from previous FEDW memory designs that exploited in-plane geometry.

While magnetic domain walls have been investigated heavily for memory and logic applications over the past 10 years, work on ferroelectric domain walls is more recent, but is now seeing significant attention, according to Seidel.

Seidel adds: "Much of the work at the moment is fundamental research, but proposals for commercialization have been discussed and are being actively explored. I think we are laying some of the groundwork at the moment for potential industrial uptake of our ideas and concepts."

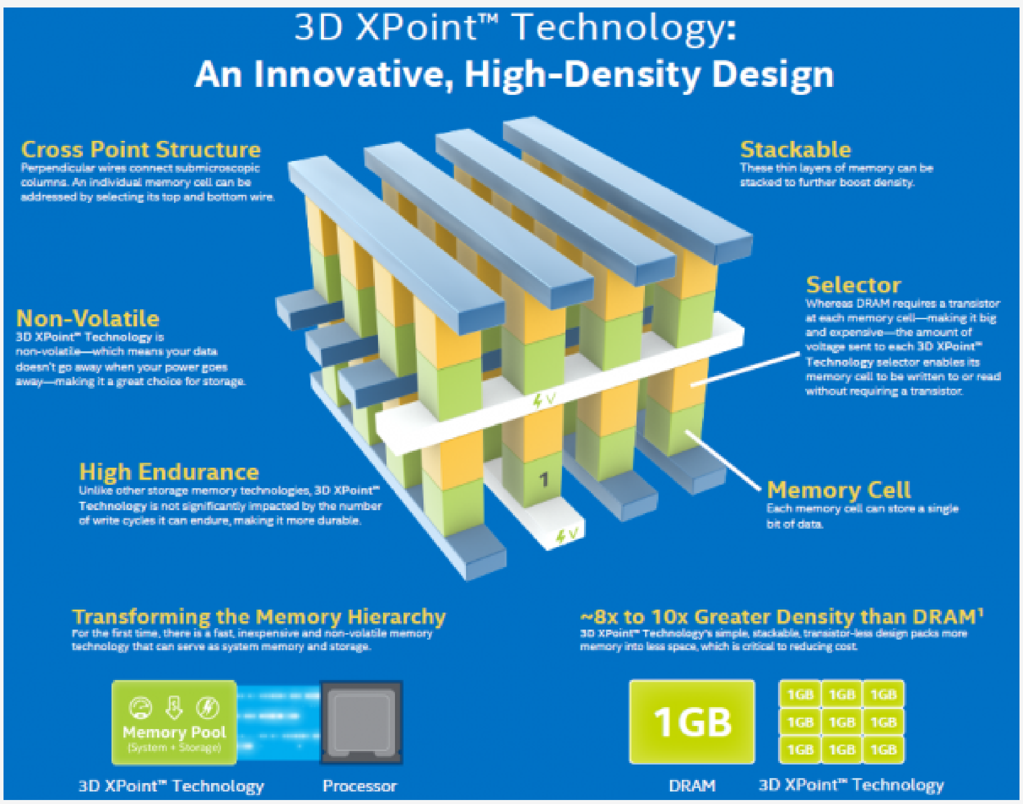
Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

15 3D XPOINT PHASE-CHANGE MEMORY

Author: **Allyn Malventano**
Date: **June 2, 2017**
Manufacturer: **PC Perspective**

15.1 INTRODUCTION

I've seen a bit of flawed logic floating around related to discussions about 3D XPoint technology. Some are directly comparing the cost per die to NAND flash (you can't - 3D XPoint likely has fewer fab steps than NAND - especially when compared with 3D NAND). Others are repeating a bunch of terminology and element names without taking the time to actually explain how it works, and far too many folks out there can't even pronounce it correctly (it's spoken 'cross-point'). My plan is to address as much of the confusion as I can with this article, and I hope you walk away understanding how XPoint and its underlying technologies (most likely) work. While we do not have absolute confirmation of the precise material compositions, there is a significant amount of evidence pointing to one particular set of technologies. With Optane Memory now out in the wild and purchasable by folks wielding electron microscopes and mass spectrometers, I have seen enough additional information come across to assume XPoint is, in fact, PCM based.



XPoint memory. Note the shape of the cell/selector structure. This will be significant later.

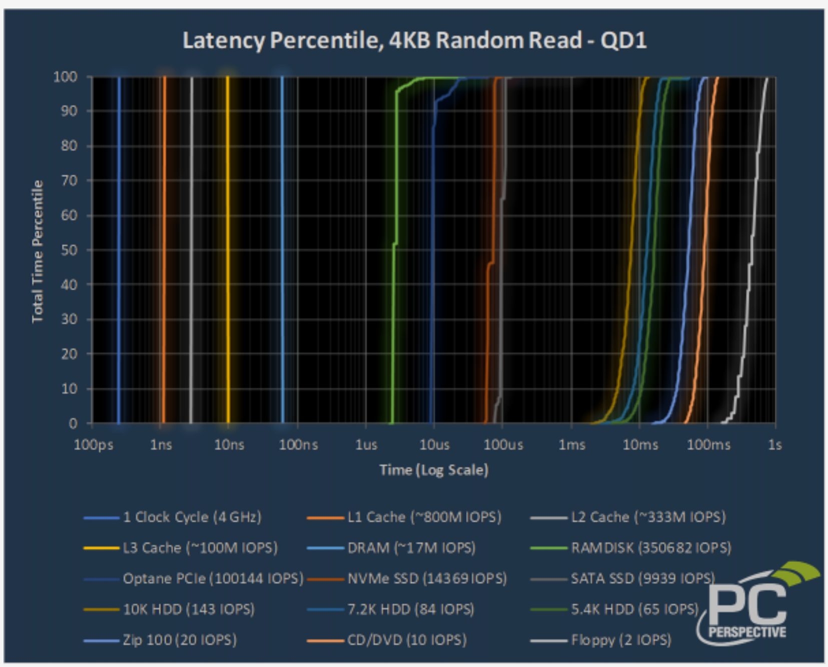
While we were initially told at the XPoint announcement event Q&A that the technology was not phase change based, there is overwhelming evidence to the contrary, and it is likely that Intel did not want to let the cat out of the bag too early. The funny thing about that is that both Intel and Micron were briefing on PCM-based memory developments five years earlier, and nearly everything about those briefings lines up perfectly with what appears to have ended up in the XPoint that we have today.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

	DRAM	PCM	NAND Flash
Page size Page read latency Page write latency Write bandwidth Erase latency	64B 20-50ns 20-50ns ~ GB/s per die N/A	64B ~ 50ns ~ 1 µs 50-100 MB/s per die N/A	4KB ~ 25 µs ~ 500 µs 5-40 MB/s per die ~ 2 ms
Endurance	∞	$10^6 - 10^8$	$10^4 - 10^5$
Read energy Write energy Idle power	0.8 J/GB 1.2 J/GB ~ 100 mW/GB	1 J/GB 6 J/GB ~ 1 mW/GB	1.5 J/GB [28] 17.5 J/GB [28] 1–10 mW/GB
Density	1×	2 – 4×	4×

Some die-level performance characteristics of various memory types. source

The above figures were sourced from a 2011 paper and may be a bit dated, but they do a good job putting some actual numbers with the die-level performance of the various solid state memory technologies. We can also see where the ~1000x speed and ~1000x endurance comparisons with XPoint to NAND Flash came from. Now, of course, those performance characteristics do not directly translate to the performance of a complete SSD package containing those dies. Controller overhead and management must take their respective cuts, as is shown with the performance of the first generation XPoint SSD we saw come out of Intel:

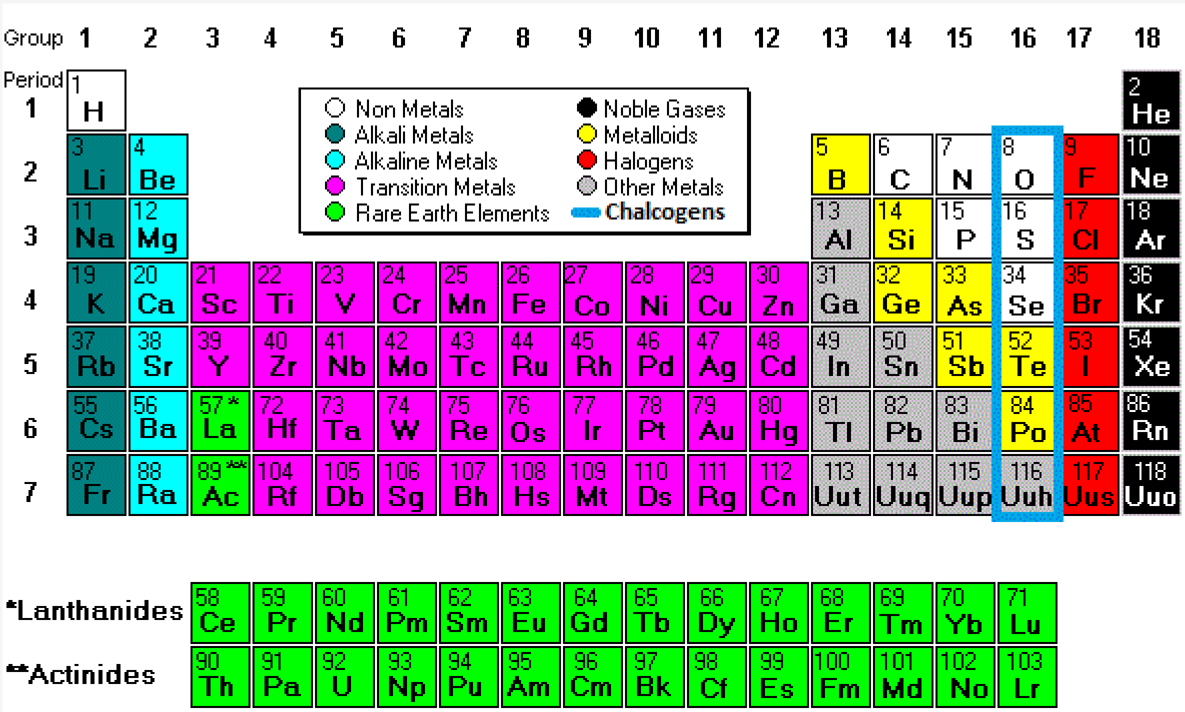


The ‘bridging the gap’ Latency Percentile graph from our Intel SSD DC P4800X review.
(The P4800X comes in at 10us above).

There have been a few very vocal folks out there chanting 'not good enough', without the basic understanding that the first publicly available iteration of a new technology never represents its ultimate performance capabilities. It took NAND flash decades to make it into usable SSDs, and another decade before climbing to the performance levels we enjoy today. Time will tell if this holds true for XPoint, but given Micron's demos and our own observed performance of Intel's P4800X and Optane Memory SSDs, I'd argue that it is most certainly off to a good start!

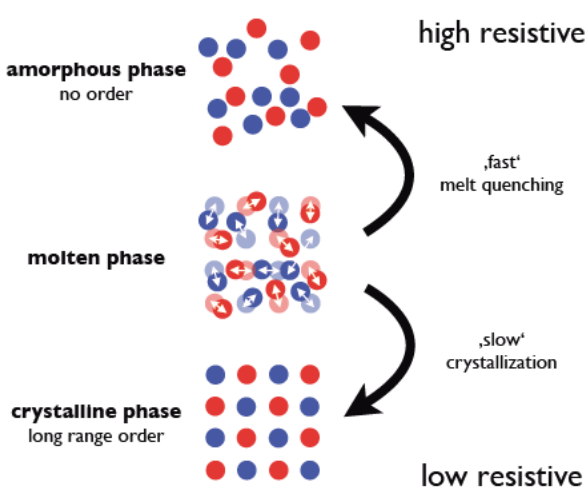
15.2 How PCM Works

To understand how XPoint reads and writes bits, let's start with how phase change materials work, and to do that we need to know what makes a material PCM capable in the first place:



Periodic Table. Metalloids in yellow. Chalcogens blue boxed. source

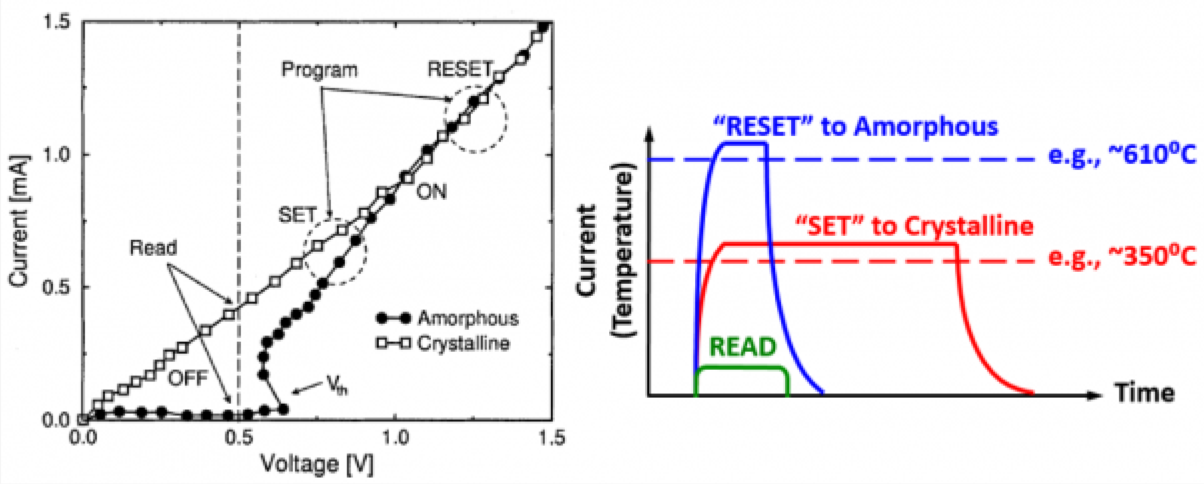
Phase change materials are generally alloys of metalloids. Metalloids are elements that share properties with metals and non-metals. They act as insulators at room temperature and as conductors when heated (or when doped). Alloys of varying mixtures of the semimetals have been experimented with for decades. Boron is mainly used for doping, and Polonium is unstable and radioactive, so we won't be seeing much of that one :). Silicon is great for standard transistors and other semiconductors, but less than optimal as a phase change material. That leaves Germanium (Ge), Arsenic (As), Antimony (Sb), and Tellurium (Te). Alloying these together results in a chalcogenide , which in the context of this article is a compound containing a Tellurium anion (Te is the only stable metalloid belonging to the chalcogen group of the periodic table). Once mixed in the proper proportions, these materials offer some rather unique properties:



Specifically, metalloid alloys have multiple stable states that each come with their own distinct resistance characteristics. These can be manipulated by heating and cooling the material in various ways. The amorphous state resembles a glass, while the crystalline state more closely resembles a metal.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

15.3 Reading and Writing Phase Change Memory



Attributes of PCM materials. Sources: 1,2

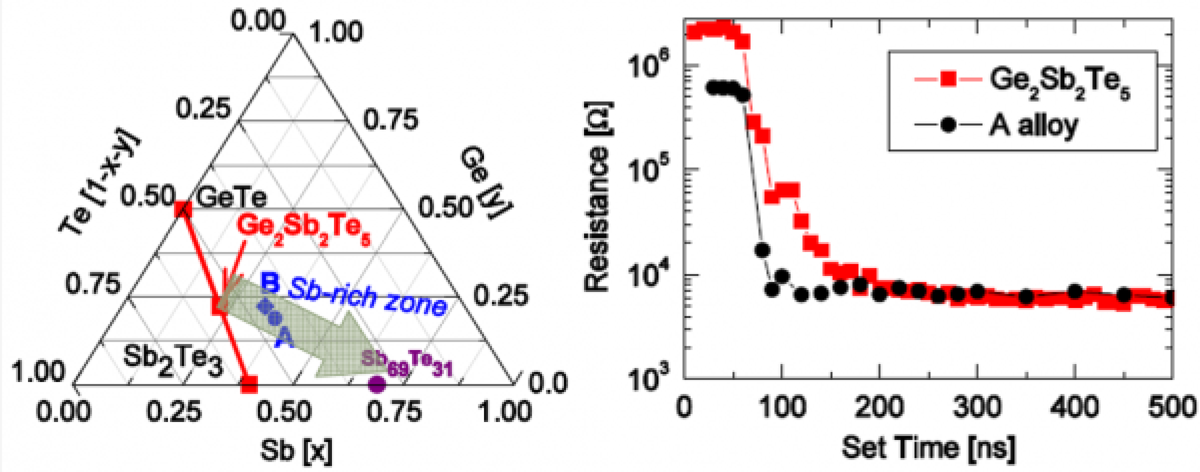
Ok, let's explain what is going on here. Voltage is applied across a section of chalcogenide material. If the material is in an amorphous (mixed) state, it does not begin to conduct until the threshold voltage (V_{th}) is exceeded. Once conducting, as voltage increases further, so does the current. Since the material is now acting resistive, it dissipates heat and therefore increases in temperature. If held at the 'set' (1) voltage, the material reaches ~350C, which is not hot enough to become molten, but *is* warm enough for its molecules to realign into a crystalline structure if the temperature is maintained for ~100 nanoseconds. Once formed, the crystalline structure behaves like a resistor and remains even after the voltage is removed and the material cools. Once in the set state, applying 0.5V would result in ~0.5mA (using the above example). The voltage no longer needs to meet a threshold in order for the material to conduct, and its response follows the plot line marked 'crystalline'.

To reset the cell, we apply a much higher voltage, pushing currents and temperatures high enough (~600C) to heat the material to a molten state. This melts down the crystalline structure. The voltage is then removed and the material rapidly cools, passing through the crystallization temperature region too quickly to form any crystal structure, 'freezing' it in the amorphous state. It is now 'reset' (0), and applying that same 0.5V will result in near zero current. I should point out that we don't need nearly that high of a voltage to perform a read, as even 0.1V would produce a readable difference in current between the two states used in our example.

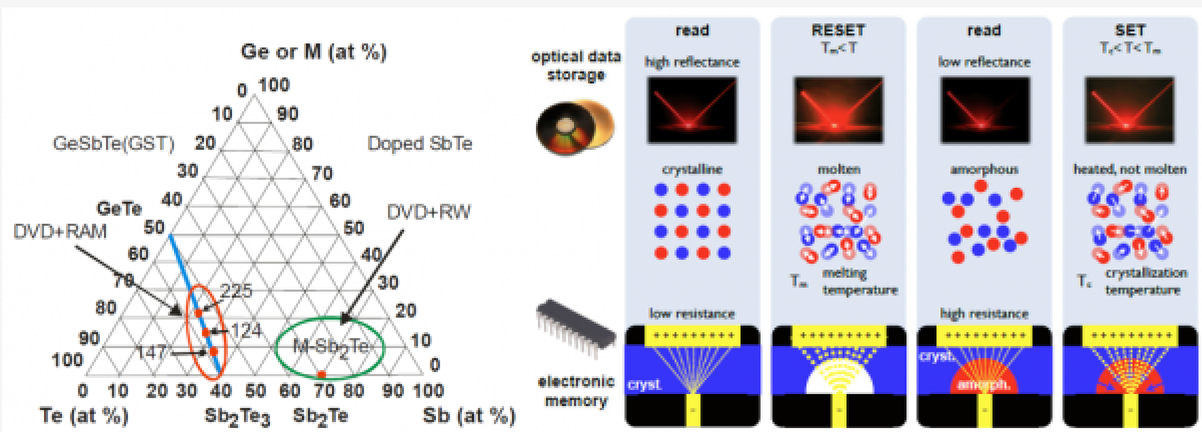
An interesting thing to note about the above is that there is no 'erase' required before programming a cell as is the case with NAND flash. With PCM cells, we can perform a set or reset operation by simply applying the associated voltage/time profile without regard for the previously set state. Unlike NAND which must be written in pages (KB) and erased in even larger blocks (MB), PCM data can be overwritten 'in-place', and single bit overwrites are possible without disturbing adjacent cells.

15.4 Tweaking the Formula

Intel and Micron would have you believe that the stuff that makes up XPoint is an ancient Chinese secret. Well, it's not. The common phase change alloy is a 2:2:5 stoichiometric ratio of Germanium, Antimony, and Tellurium. $Ge_2Sb_2Te_5$, dubbed 'GST' for short. As is with most alloys, there are many slight variations possible to the recipe, and that is where the manufacturer-specific secrets come into play. That said, we do have some clues as to what might have been tweaked from a 2010 Micron presentation:



Those developing PCM technology will naturally finely tune the mixture to try and improve performance. Above we see an excerpt from a Micron brief showing how slightly increasing the concentration of Antimony (Sb) helped reduce the reset resistance (reducing the voltage needed), as well as reducing the time needed for a set operation. There are also external factors related to cell selection that might require tweaking the ratios further, which we will touch on shortly.



source

You might think phase change alloys are so exotic that you have never seen or held them, but you are likely wrong. Rewritable optical discs (CD-RW/DVD-RW) are extremely close cousins to the materials found in XPoint. Optical discs used Silver and Indium in place of the Germanium found in GST, which naturally changed the properties of the alloy. ‘Blank’ media was crystalline, and pits were written by heating spots by pulsing the write laser. The spots then quickly cooled without a chance to recrystallize, forming darker areas that could later be read as differences in reflectivity. Discs were erased by applying a lower power laser which started the recrystallization process (these alloys could continue crystalizing after the laser passed the area). Other metalloid alloy variants were used in various optical media technologies, aiming to improve the number of erase cycles and other performance characteristics. DVD-RAM actually used GST compounds but relied on its changing optical properties as opposed to electrical conductivity.

16 SIGNPOSTS ON THE ROADMAP OUT TO 10 TB/SEC ETHERNET

September 8, 2017 Timothy Prickett Morgan

The world of Ethernet switching and routing used to be more predictable than just about any other part of the datacenter, but for the past decade the old adage – ten times the bandwidth for three times the cost – has not held. While 100 Gb/sec Ethernet was launched in 2010 and saw a fair amount of uptake amongst telecom suppliers for their backbones, the hyperscalers decided, quite correctly, that 100 Gb/sec Ethernet was too expensive and opted for 40 Gb/sec instead.

Now, we are sitting on the cusp of the real 100 Gb/sec Ethernet rollout among hyperscalers and enterprise datacenters, which John D'Ambrosia, chairman of the Ethernet Alliance trade group, says "will be the largest rollout that we have ever seen," and that is true for a bunch of reasons. For one thing, the cost of 100 Gb/sec Ethernet switches, which often include routing functions and therefore allow standardization of iron across switching and routing workloads, is coming down fast as new ASICs enter the field based on the 25G signaling standard that the hyperscalers (primarily Microsoft and Google) rammed down the IEEE's throat a few years back for the good of the entire industry. For another thing, there are machine learning and IoT workloads that are dependent on gathering up immense amounts of telemetry from every device known to man, from blenders to cars, and chewing on it back in the datacenter for insight, and that is putting bandwidth pressure on networks. And then, of course, there is the ever-embiggening media files that we use in our business and personal lives, the increasing cross connection between people, the increasing distributed nature of applications, and the increasing population of the world.

There are no surprises, then, that with 100 Gb/sec Ethernet now at an affordable price, seven years since it entered the field, it is finally ready to take off. It is beyond overdue, based on the pressure from compute and storage, which has been growing capacity faster than networking bandwidth rates in the past decade.

Gigabit Ethernet came out in 1998, running at the blazing speed of 1 Gb/sec then, when the hyperscalers were not anywhere near the technology giants they are today. The 10 Gb/sec Ethernet standard came out in 2002, and it ramped pretty quickly in the datacenter and then down into other use cases, but the jump to 100 Gb/sec did not take the typical three to four years. That is because it 25 Gb/sec lane signaling was not developed and vendors instead decided to gang up the 10 Gb/sec lanes used in the intermittent 40 Gb/sec Ethernet standard, which was really done to give hyperscalers a stopgap on the way to 100 Gb/sec. The telcos did not mind paying the price, because they can pass it on to us in our exorbitant phone and data plan bills. Those providing essentially free services at a massive scale cannot afford to do this. Hence the 25G Ethernet standard in 2014, which has led not only to 25 Gb/sec, 50 Gb/sec, and 100 Gb/sec switching and sometimes routing, but also these speeds at a port density and price point that hyperscalers demand and that enterprises, lagging a few years behind, will be able to leverage.

Anyone who thinks that enterprises will embrace older 40 Gb/sec technologies when they make the jump from 10 Gb/sec gear is silly. They will go to 25 Gb/sec on their server ports and either 50 Gb/sec or 100 Gb/sec on their switches, and if we were them, and they were smart, they would be using cable splitters to use 25 Gb/sec on the switch today and then have the option to move up to 50 Gb/sec and 100 Gb/sec in the future just by changing the splitters and doubling and quadrupling their switches.

Enough cannot be said about how the 25G standard changed things. The networking folks at Alcatel-Lucent (the former Bell Labs where so many great technologies were invented) cooked up this comparison that makes it plain:

3.2 Tb/s Switch	Servers	100 GE Uplinks	Capacity Utilization (Tb/s)	Capacity Utilization (%)	ToR Switches for100K Servers
25 GE (1 x 25 Gb/s)	96	8	3.2	100	1042
40 GE (4 x 10 Gb/s)	28	4	1.52	47.5	3572

Port Speed (Gb/s)	Lane Speed (Gb/s)	Lanes per Port	Usable Ports	Total Capacity (Tb/s)
10	10	1	128	1.28
25	25	1	128	3.2
40	10	4	32	1.28
100	25	4	32	3.2

- Connections to switch ASICs is limited by SERDES count and bandwidth
- Single higher speed 25 Gb/s lanes maximize bandwidth and switch fabric utilization vs. 4 x 10 Gb/s lanes
- A single lane per physical port maximizes the number of connected servers or uplinks per switch
- Overall higher port count, utilization and total server interconnect bandwidth vs. 40 GE

This is better at both the server level, where hyperscalers had to aggregate four 10 Gb/sec ports together or buy more expensive 40 Gb/sec adapter cards, on the servers (they did the former at first and the latter after a few years of rolling out 40 Gb/sec on the switches) and at the switch level. To be sure, 25 Gb/sec ports coming out of the server does not

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

provide as much bandwidth as the 40 Gb/sec option, but look at the difference in the switching, which is what really matters in the total cost of ownership calculation. Using 25 Gb/sec signaling, a typical ASIC with 3.2 Tb/sec of aggregate bandwidth can support 96 server ports running at 25 Gb/sec and eight 100 Gb/sec uplinks, and every bit per second of that switch is utilized; moreover, you only need 1,042 top of rack switches to cross-connect 100,000 servers, which is about the capacity of a hyperscale datacenter, give or take. (A region has multiple datacenters, of course.) This is compared to the 40 Gb/sec setup, which would require 3,572 switches for those 100,000 servers. And if you wanted to double up the port speeds to 50 Gb/sec on the servers, yielding 25 percent more bandwidth than the 40 Gb/sec ports, you would still only need 2,084 switches using 25G devices, still a lot fewer than the 40 Gb/sec switch farm.

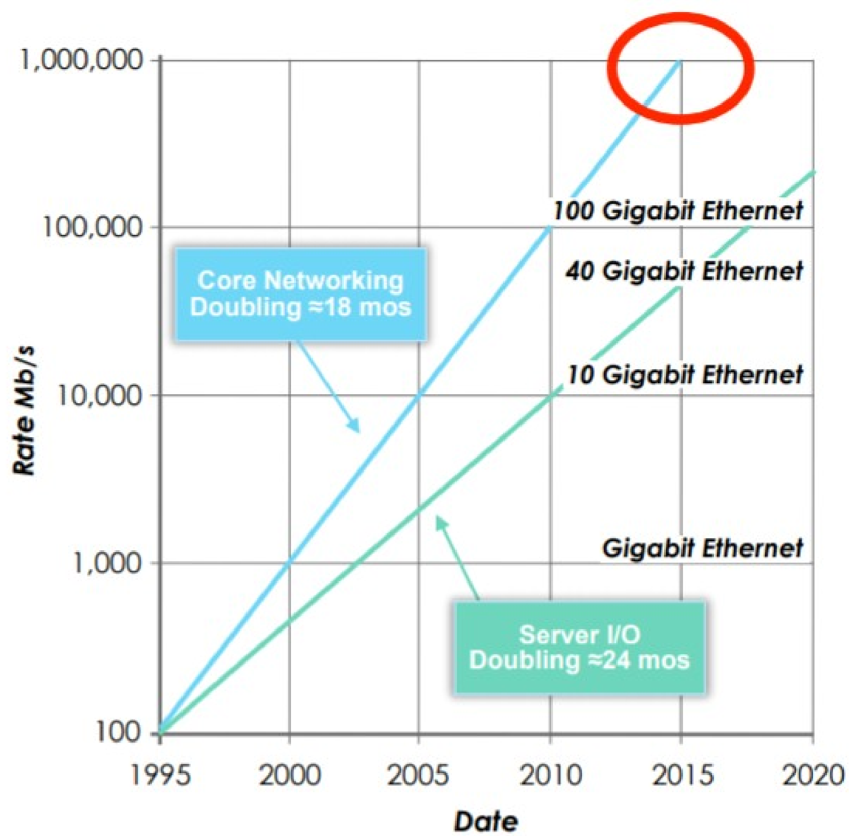
The 40 0Gb/sec compromise was just that – a compromise until 25 Gb/sec signaling was available, and the issue was that the switch ASIC makers and the IEEE were willing to wait for 25 Gb/sec signaling at the 200 Gb/sec Ethernet bump, and the hyperscalers were not.

16.1 The Unfolding Ethernet Roadmap

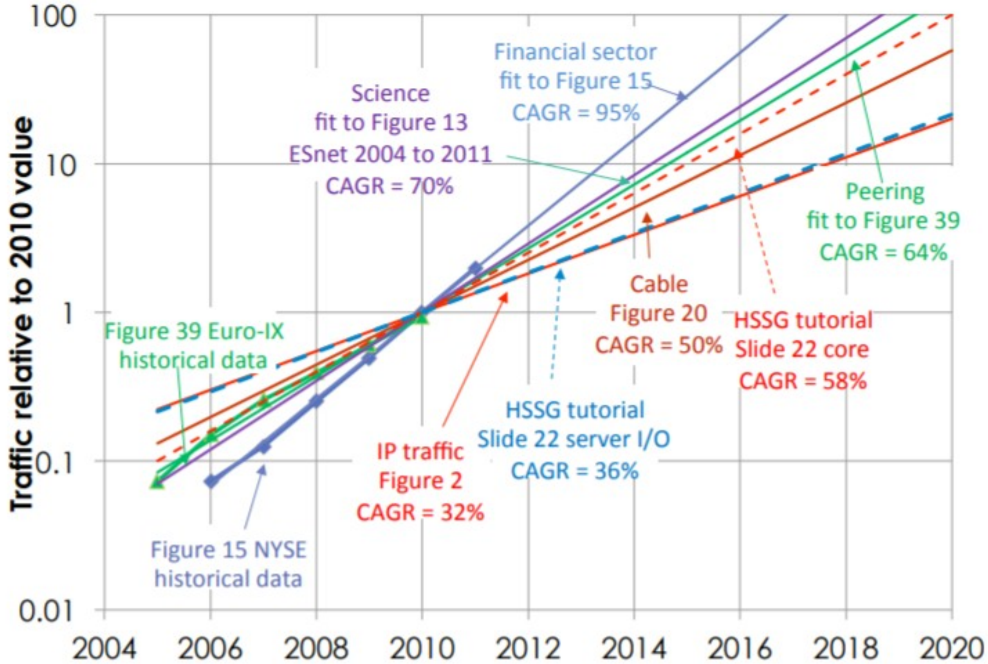
As we said above, predicting the progression of increasing Ethernet speeds used to be fairly easy, and then it got tough for a while when the base ten upgrades stopped happening in a predictable fashion. Now, rather than focusing on 10X upgrades, the industry is settling in to do 2X, 4X, and maybe 10X or 16X bandwidth improvements over the coming years.

Making such predictions is hard, though, because there is a lot of clever engineering that needs to be done in the switches, in the transceivers, and in the cables to make it all work. There are no sure bets here, because Moore's Law issues affect switch chips every bit as much as they affect compute chips.

Just for fun, here is an Ethernet roadmap from 2007, showing the steady state that was expected a decade ago:

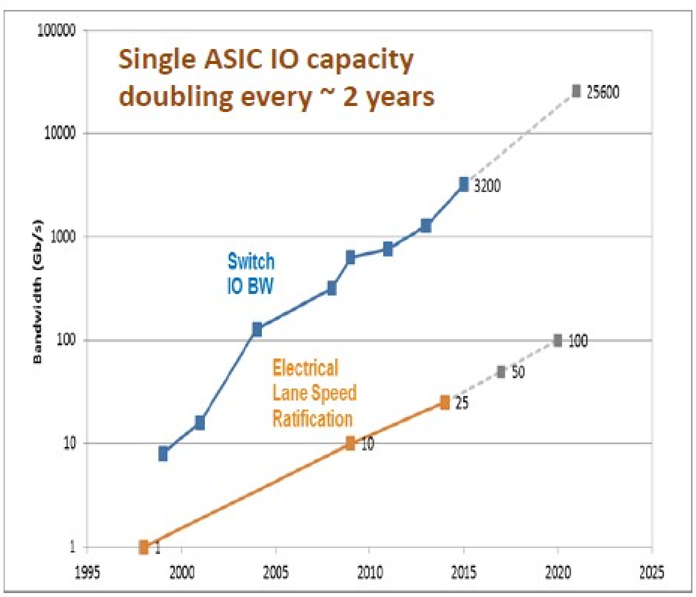


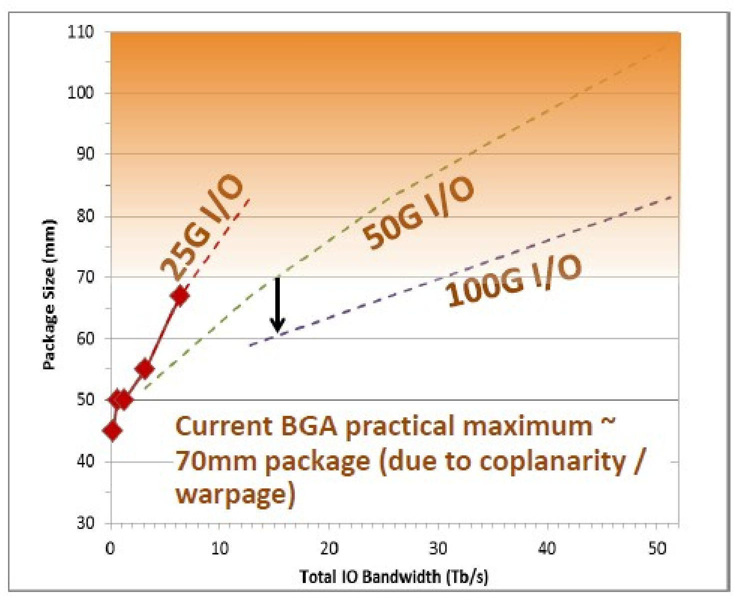
Based on this roadmap, we should have seen 400 Gb/sec networking in 2013 and 1 Tb/sec networking in 2015. The demand projections were certainly there. In 2012, the IEEE Ethernet Working Group published a bandwidth assessment report that showed key types of workloads – scientific computing, financial services, telecom are the biggies – growing such that 1 Tb/sec Ethernet would be needed by 2015 and 10 Tb/sec by 2020. The normal cadence:



We obviously do not have 400 Gb/sec, much less 1 Tb/sec, switching available today. But as we have pointed out before, networking is getting back on track and is now accelerating faster than compute again. (It is very likely that compute and networking will not grow faster than storage, but the shift to faster media is mitigating this gap while also putting pressure on the network at the same time.)

In a talk about the future Ethernet roadmap, D'Ambrosia demonstrated that the I/O capacity of a single network ASIC has been doubling every 18 months or so, as expected, but also showed how the capacity in a switch ASIC is being constrained by practical limits on the size of chips. Take a look:

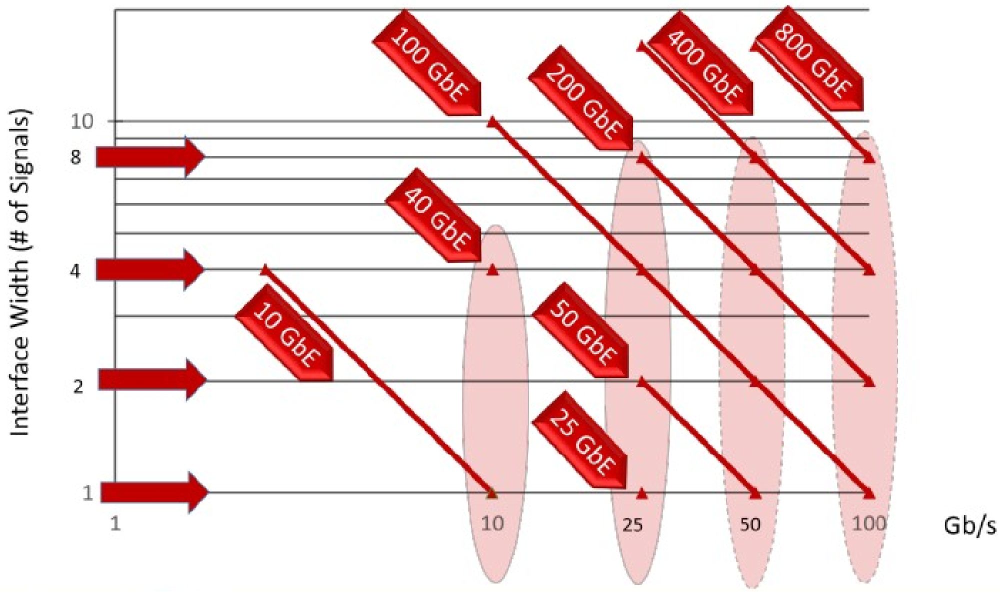




The upshot is this: To cram more bandwidth onto a switch ASIC, the only way to do it is to boost the lane signaling on the SERDES. We just jumped from 10 Gb/sec to 25 Gb/sec, and that was tough enough, but to keep on the Moore's Law pace, that means switch ASIC makers have to get to 50 Gb/sec signaling in 2018 or so (and we have talked about some vendors who are trying to do that) and to 100 Gb/sec signaling by 2020. That would mean a switch ASIC that could drive a whopping 25.6 Tb/sec of aggregate bandwidth.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

That would be pretty sweet to have about right now, and the fact that such capacity is not here is why chip makers and the hyperscalers that are egging them on are pushing the technology ahead of the standards. They can read a SERDES chart as well as the ASIC makers, and they can make their own assessments of how to best make use of signaling rates and lane aggregation to get a desired bandwidth per device and per port. Like this one:



Anywhere you want to take Ethernet in the next couple of years has to intersect a point on that chart above.

“In the past four years or so, we have come to recognize the importance of the SERDES,” explains D’Ambrosia, who is senior principal engineer at FutureWei Technologies and who was chief evangelist for Ethernet at Force10 Networks before Dell bought it several years ago and had a similar position at Dell. “So with 25 Gb/sec you get 100 Gb/sec Ethernet, with 50 Gb/sec you get 200 Gb/sec, and with 100 Gb/sec you get 400 Gb/sec Ethernet. There is a basic relationship that is going to be driving things, and we are following the SERDES. As we do this 1X and 4X approach, we are also finding that we are going to have to do 8X. We are not doing eight times 25 Gb/sec to get to 200 Gb/sec switching, but today we are doing eight by 50 Gb/sec to get to 400 Gb/sec and in the future, we will do four by 100 Gb/sec to get to 400 Gb/sec as well. That begs the question of what the next Ethernet speed is going to be after that. If we follow this logic, eight lanes is what we would jump to after four, and sixteen lanes is a little wide.”

No one is talking about ten lanes at 100 Gb/sec to hit 1 Tb/sec, which would have been the classical bandwidth bump beyond 100 Gb/sec. Everyone assumes that we will be filling in the gaps, making the best price/performance choices to drive down the cost per bit as the hyperscalers have compelled the industry to do. And that probably means that the next speed after 400 Gb/sec is probably going to be 800 Gb/sec or 1.6 Tb/sec at the switch level, very likely sometime way out beyond 2020. That hardly better than the 1 Tb/sec that was originally expected in 2015 and the 10 Tb/sec expected in 2020, mind you. Networking might be getting back to a Moore’s Law curve, as we have observed, but it did not make up for lost time.
The Ethernet Alliance posted a roadmap in 2015, and then updated it again in 2016, and D’Ambrosia says that it will take in all the new developments in ASICs and transceivers and update the roadmap again in 2018. It would be interesting to see far out into the hills of 2025 and the mountain ranges of 2030 and have the industry take a stab at what networking might look like way far out. Compute and networking could hit the Moore’s Law wall at about the same time, and that is precisely what we expect.

17 THE HARD DRIVE SIZE GROWS AS 20TB DRIVES EXPECTED IN 2020

by Cask J. Thomson
From TheWholeStory.news

Hard drive manufacturer Seagate has laid out its near-term roadmap, with plans to increase drive capacities over the next few years. Seagate's 10TB hard drives are now readily available, but the company has planned to reach production of 16TB magnetic drives in 2018 with 20TB drives to be available by 2020.

According to unit sales, SSD's are taking over, with Seagate themselves revealing a 60TB Solid State to be released this year but the focus is still on magnetic drives according to the company.

In 2015, Western Digital revealed their 10TB hard disk drive (HDD) based on its HelioSeal Platform, which hermetically seals helium gas in the drive to decrease friction. At the time of publication, Western Digital's HelioSeal 14TB is the highest capacity drive available to consumers. 4TB being the highest 2.5" drive (small external and laptop drives)

Western Digital, Seagate and Toshiba are currently the only manufacturers of drives. It is expected that we will reach 40TB by 2030 according to computer science analysts.

Innovation slowed in 2011 after the "Global Hard Disk Famine" caused by the 2011 Thailand Floods which drove up prices and dropped the availability of drives. Thailand is one of the world's largest producers of HDD's.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

18 The Next Platform

The Next Platform is published by Stackhouse Publishing Inc and formally launched February 23, 2015, in partnership with the UK's top technology publication, *The Register*. It offers in-depth coverage of high-end computing at large enterprises, supercomputing centers, hyperscale data centers, and public clouds.
If you're looking for a more detailed overview of the vision, read the Editors' Introduction or, if that's too wordy, a) you might be in the wrong place and b) reach out to the editorial staff for more specific info if you need it.

Some companies are building their own platforms for internal use, while others are building platforms for others to use to run their applications as a service.

Inspiration for platform designs is coming from the hyperscale data center operators that have pushed the boundaries of scalability for new kinds of analytics, as well as from supercomputing centres that have been scaling up simulation and modelling workloads for decades.

Regardless of where the inspiration originates, all platforms have some common characteristics. They are based on an integrated stack of hardware and software, tuned to run specific workloads, and outfitted with orchestration tools to automatically react to changes in those workloads.

The idea is to break down as many silos in the data center as possible, virtualising system components and mashing them up in interesting ways to increase the efficiency of the underlying servers, storage, and networks.

The Next Platform will step behind the headlines and provide analysis to help readers understand what technologies are used to solve particular problems, how they are integrated with other systems and applications, why organisations choose particular technologies to solve their problems, and why they do not pick something else.

The new publication will cover the key elements of the modern system, from processors, main memory, storage, and networking up through operating systems, middleware, and other key systems software such as databases and data stores, systems management tools, and cluster and cloud controllers.

It will look at the myriad clustering technologies available today to bring compute to bear, from hyperconverged systems for virtualised enterprise workloads, to shared memory NUMA machines, all the way up to massively parallel supercomputing systems.

The technologies created to solve one set of problems can often be used to solve another set, and *The Next Platform* will examine how this is actually happening in the data centers of the world. If a technology can be used to rejig a system to have more throughput, deliver lower latency, or to be easier to manage or program, then *The Next Platform* will drill down into it.

The Next Platform will be inclusive about underlying hardware and systems software, it will also be broad in its coverage of the applications that drive the design of systems from the beginning.

Transaction processing systems are still important and evolving with in-memory databases, but they are augmented by layers of analytics software that wraps around transactions before they start, as they are running, and after they are done.

Modelling and simulation are also key aspects of the manufacturing and the financial sectors and these applications are increasingly being integrated with other kinds of analytic workloads to make better products. The important thing to consider about the distinct but related markets – large enterprise, high performance computing, hyperscale data centers, and large-scale clouds – is that technologies developed in one arena are being adopted by the others.

For example, Facebook's Open Compute server and storage designs and Rackspace's OpenStack cloud controller are making headway into the HPC centers of the world and among financial services firms, just to name two early adopters. GPU, FPGA, DSP, and other kinds of accelerators are being deployed by enterprises after many years of development and deployment in academia and in segments of the HPC and financial services sectors.

Various analytics tools that got their start at hyperscale and media companies (Hadoop being the obvious one) and file systems and middleware that have an HPC heritage are being mashed up and used by enterprises, too. (Everybody is looking to sell a replacement for the Hadoop Distributed File System and other layers of the Hadoop stack because of their inherent limitations.)

Tracking this interplay and interchange of technology between these different parts of the high-end of the IT market is one of the core missions of *The Next Platform*. How risk-adverse enterprises adopt these technologies for competitive advantage, and why, is what is interesting.
"We believe it is time to create a single publication that brings several different parts of the high-end of the IT market together to reflect the increasing convergence of systems and interdependence of workloads that are being brought to bear to solve tough IT problems," explains Timothy Prickett Morgan, co-editor of *The Next Platform*. "We also want to get back to the idea that depth and insight matter. It takes more than sound bites to make sound decisions."

According to co-editor, Nicole Hemsoth, "All told, there are a few tens of millions of programmers, administrators, architects, and managers in the IT world, and probably somewhere on the order of a third of them work at hyperscale, HPC, and high-end enterprise shops as defined above. That is the broadest definition of our intended audience — and it is one that we know well."

The Next Platform will, Hemsoth notes, focus on the IT challenges of companies with more than several thousand employees and in excess of $250m and their equivalents in the public sector.

"Over the course of a year we will document these trends by tapping into user-specific stories that reflect the global economy at large, with stories from the manufacturing, distribution, retail, energy, financial services, public, media, and other sectors."

19 THE ASCENDANCY OF ETHERNET STORAGE FABRICS

September 26, 2017 Timothy Prickett Morgan

It is hard to remember that for decades, whether a system was large or small, its storage was intricately and inescapably linked to its compute.

Network attached storage, as pioneered by NetApp, helped break those links between compute and storage in the enterprise at the file level. But it was the advent of storage area networks that allowed for storage to still be reasonably tightly coupled to servers and work at the lower block level, below file systems, while at the same time allowing that storage to scale independently from the number of disks you might jam into a single box.

SANs were a kind of precursor to the disaggregated and composable system that we often talk about here at *The Next Platform*. And they are still a very big business in the enterprise, even if hyperscalers and cloud builders do not use such storage except, perhaps ironically, in the back office systems where they count their money and pay their bills. (You don't think Google and Microsoft and Facebook and Amazon write their own code for that or co-design systems to support these functions, do you? They most certainly do not.) But with the advent of clustered storage and the rise of unstructured data and the object storage that not only contains it, but can be the underpinning of block and file storage as well, as you might expect the era of the SAN has long since peaked even if it will take a long, long time before the last SAN and its Fibre Channel switching is unplugged in the glass house.

Some kinds of change come to the datacenter explosively, some at the pace of molasses in the winter. The ascent of InfiniBand and now Ethernet as a storage fabric and the decline of Fibre Channel is measured in decades, not years, but it looks like the enterprise will build storage that looks more like that seen in HPC centers and at hyperscalers and cloud builders, with their parallel and clustered file systems and high speed, lossless InfiniBand or Ethernet networks, and looks a lot less like the massive aggregations of spinning rust, now sometimes front ended by flash to speed up accesses and still using Fibre Channel switching, that have dominated the datacenter.

19.1 A WALK DOWN HISTORY LANE

At this point in the history of networking, Fibre Channel, Ethernet, and InfiniBand are on parallel tracks that seem to overlap, with InfiniBand usually out on the bow wave of bandwidth expansion and Fibre Channel riding on the tail. But it wasn't always this way, which is why Fibre Channel became the interconnect of choice for high performance SAN array attachment for servers and, because of its popularity and the vast installed base of SAN storage in the enterprise, this is why it persists today even as, in some cases, it is being supplanted by Ethernet and InfiniBand.

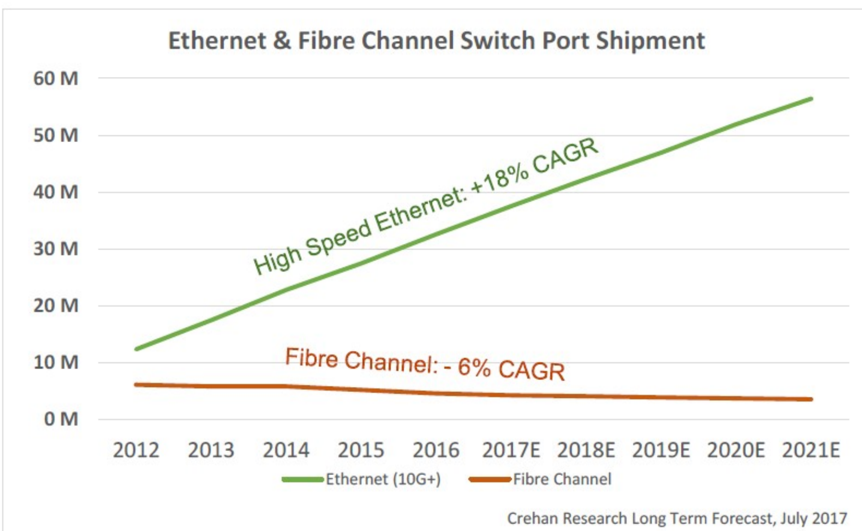
Back in the mid-1990s, when decoupling storage from servers and aggregating it in a SAN first took off, Fibre Channel had a serious bandwidth advantage with its blazing 1 Gb/sec speed, and it quickly moved up to 2 Gb/sec because it did not have to go through Ethernet standards to blaze ahead. At the time, Ethernet was humming along at 100 Mb/sec, which was fine for corporate networks and maybe NAS devices linked at the file system layer where performance was not critical. But this setup not good for linking storage arrays to servers at the block level where databases and then virtual machines – yes, proprietary and Unix machines had hypervisors a decade before the X86 platform went crazy for them – needed more oomph.

InfiniBand would not even be invented until the late 1990s, and even though it was created to be a universal fabric for linking servers to each other, to storage, and to clients, that never really happened and InfiniBand was largely relegated to being a high speed interconnect for HPC clusters until fairly recently, when it has also seen the interconnect fabric for clustered storage systems as well as niche uses in data analytics.

Perhaps more importantly for the emergence of Fibre Channel, the parallel SCSI interface for disk drives was limited to a handful of disk drives because of its limited fan out. To scale out SCSI, the interface was serialized and the SCSI protocol was put on a wire and became Fiber Channel. So not only did Fibre Channel have more bandwidth, but it allowed for many more devices to hang off a channel. In effect, it made distant disks look local and it allowed more of them to be allocated to a single server. That SAN was more an aggregation of independent, logical storage arrays than a shared storage utility, mind you, but the beautiful thing was that all of the operating system and application software could speak the same SCSI block storage and it would work over the Fibre Channel network. And unlike Ethernet-attached storage, Fibre Channel SANs were running over a lossless fabric, features that have only been added to Ethernet in recent years – Data Center Bridging is a decade old, but has only recently matured and taken off – and the Fibre Channel protocol was also absorbed into Ethernet, starting with Fibre Channel over Ethernet (FCoE) with the Unified Computing System servers from Cisco Systems back in early 2009.

The rise of Fibre Channel was a little more complicated than that, but it was a relatively easy transition that allowed volume economics to come into play. Yes, SANs were more expensive, but they were also a lot more flexible, and the net-net – as is so common in the IT industry – is that a new technology was not really cheaper, but it was better.

That said, even if SAN was better than direct or network attached storage – DAS and NAS – it was not something that could be deployed affordably at hyperscale, which is why you don't see SANs at Google, Amazon, Microsoft, Facebook, and the like. They had to invent something else that used commodity parts in a different way, and they decided to build massive Ethernet networks that spanned a datacenter of 100,000 server and storage nodes and interconnected the whole shebang, allowing anything to talk to anything else. Now, that disaggregated storage approach, which is most definitely not hyperconverged storage, no matter what people say, is trickling down from on high and it is changing the Ethernet and Fibre Channel markets with it.



The Fibre Channel switch and adapter market is flat to declining slightly, as the market data for port counts from Creehan Research above suggests, but Adarsh Viswanathan, senior manager of datacenter product management at Cisco, says that it has a very, very long tail. Viswanathan adds that among enterprise customers, 80 percent of the flash arrays sold have linked to servers over Fibre Channel switching infrastructure, and for mission critical applications, customers still prefer to deploy their flash arrays or mixed disk and flash arrays in SANs, which are familiar. Moreover, as much as Cisco loves hyperconverged storage, and recently bought Springpath to have its own offering here, Viswanathan says that hyperconverged is not going to replace SANs for Tier 0 or Tier 1 mission critical applications because companies want to rock solid accuracy and predictability of Fibre Channel. Flash and other kinds of persistent storage are driving the bandwidth between server and storage and are pushing Fibre Channel just as hard as Ethernet and InfiniBand are being pushed.

To that end, back in April, Cisco announced 32 Gb/sec Fibre Channel cards for its MDS 9700 storage switches and matching 32 Gb/sec host bus adapters for its UCS C Series rack servers, At the moment, the virtual network interface cards on the UCS B Series blade servers top out at 40 Gb/sec of total bandwidth and only 8 Gb/sec and 16 Gb/sec FCoE ports are supported out to storage across the integrated switch, but you can anticipate that the UCS iron will be updated to support 32 Gb/sec FCoE ports. Just for a sense of scale, 16 Gb/sec Fibre Channel was rolled out in late 2013 and ramped in 2014. So the speed bumps are not fast and furious here.

The 32G Fibre Channel module for the MDS 9700 has 48 ports running at 32 Gb/sec, and the chassis supports up to 768 ports running at line rate for an aggregate bandwidth of 1.5 TB/sec, with support for NVM-Express over Fibre Channel being a seamless upgrade in the future. The 32G adapters that go into the C Series machines come from Broadcom (which bought Emulex) and Cavium (which bought QLogic). Dell, IBM, Hitachi, and NetApp are big partners for the MDS 9700 line; Hewlett Packard Enterprise rebadges Brocade Fibre Channel switching gear.

While this is still a big business, the big question is: *When does Fibre Channel go away?*

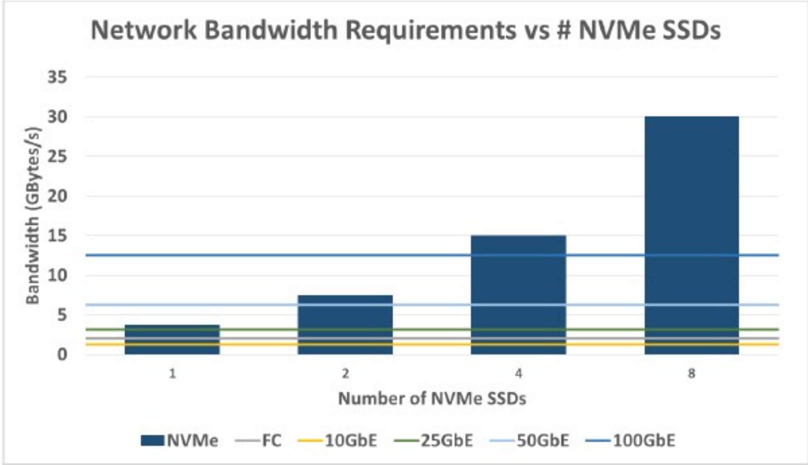
“You might say it has already gone away in that there are really no standalone Fibre Channel companies left,” says Kevin Deierling, vice president of marketing at Mellanox. “They have all gotten sucked into larger companies. The innovation is already gone, and we don’t see lots of new protocols. Most vendors are farming this as a cash cow. They will support NVM-Express over Fabrics, maybe a year or two behind, but this is no longer a driving force. That said, people are still buying IBM mainframes thirty years after the end of growth in that market.”

The real issue for large enterprises, says Deierling, is that they are hitting the same limits of scale as the hyperscalers, and it is more on their budgets than on any inherent aspects of their storage technology. If their storage is growing at 30 percent a year, but their budgets are flat, that quickly becomes a problem. Existing applications might stay on a Fibre Channel platform, but all new applications will end up on disaggregated, converged, or hyperconverged storage and all three of these drive the adoption of Ethernet fabrics.

It is not that companies made the wrong decision to either adopt Fibre Channel SANs or to continue to use them here in the 2010s. These were the best answers for large enterprises that could not, unlike the hyperscalers, engineer their own software-defined storage stacks, and for certain workloads, continuing to invest in SANs is the easiest and most cost-effective strategy. But now, as companies look at adding new platforms and applications, they can buy Ceph object/block/file storage from Red Hat or use the open source software; deploy any number of hyperconverged storage stacks like Nutanix Enterprise Computing Platform and its several competitors; or go with upstarts like DriveScale, Datrium, Excelero, and myriad other contenders who have their own twists on scale-out storage inspired by the hyperscalers.

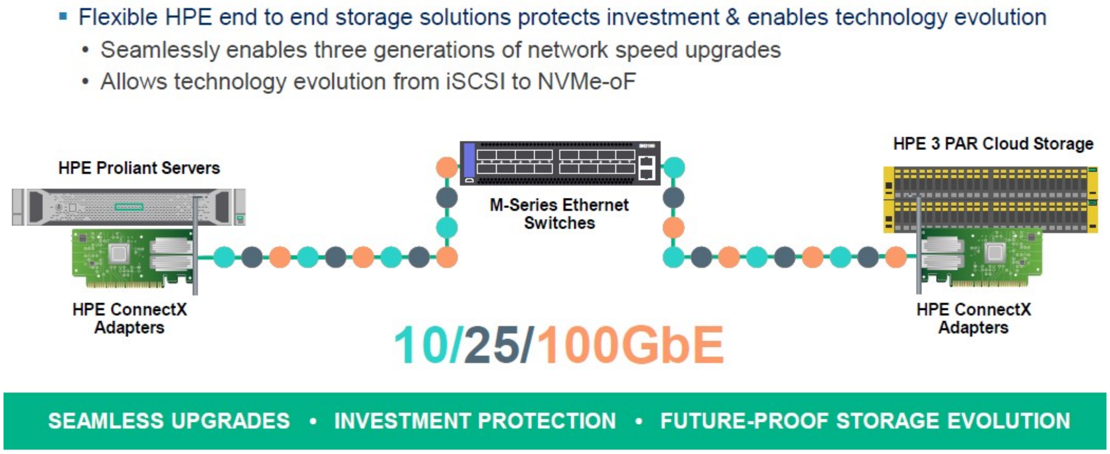
As far as Ethernet fabrics are concerned, the goal, says Deierling, is for key features of Fibre Channel –management, automation, zoning, and security – will be exposed on top of Ethernet. “The vision is that everything you can do on Fibre Channel, you will be able to do on Ethernet, only faster and cheaper,” says Deierling, and to put some numbers on that, he adds it will have three times the performance for one third the cost, which is nearly an order of magnitude better bang for the buck. You understand now why hyperscalers moved to disaggregated storage linked by the fastest and widest Ethernet they could build.

The beauty of this approach is that all of the goodness of Ethernet, such as accelerators to support VXLAN overlay networks or Open vSwitch virtual switching, just to name two features, now comes to the storage fabric, which presents some interesting possibilities for innovation.



The bandwidth and latency that supports such innovation are going to be issues. Back in the day, when Fibre Channel was new, a disk drive had an access time of maybe 10 milliseconds, and today, the fastest disk drives, two decades later, are maybe around 7 milliseconds. That is just the limit of a physical device that has mechanical rather than electrical access.

By contrast, a flash-based SSD has an access time on the order of 100 microseconds, and looking forward, there are low latency SSDs and persistent memory and NVDIMMs that offer access times below 10 microseconds and the envelope is pushing down to under 1 microsecond. That factor of 10,000X different in access time means the network linking compute to storage is key, and with Ethernet at 100 Gb/sec today and on track for 200 Gb/sec and 400 Gb/sec shortly and perhaps hitting 1.6 Tb/sec early in the next decade, the odds are that Fibre Channel is going to lag, just as it is today. Only now is 32 Gb/sec Fibre Channel coming to market, and the gap is going to widen, and today, if you use NVM-Express over fabrics on top of Ethernet switching, this only adds around 10 microseconds of additional latency over running an SSD local in the chassis. So even when NVM-Express comes to Fibre Channel, which is great for those still deploying SANs, the combination of Ethernet and software defined storage running on server nodes will have a big advantage.



20 CLOUD COMPUTING WIKIPEDIA

From Wikipedia, the free encyclopedia: https://en.wikipedia.org/wiki/Cloud_computing

20.1 INTRODUCTION TO CLOUD COMPUTING - WIKIPEDIA

Cloud computing is an information technology (IT) paradigm, a model for enabling ubiquitous access to shared pools of configurable resources (such as computer networks, servers, storage, applications and services),[1][2] which can be rapidly provisioned with minimal management effort, often over the Internet. Cloud computing allows users and enterprises with various computing capabilities to store and process data either in a privately-owned cloud, or on a third-party server located in a data center - thus making data-accessing mechanisms more efficient and reliable.[3] Cloud computing relies on sharing of resources to achieve coherence and economy of scale, similar to a utility.

Advocates note that cloud computing allows companies to avoid or minimize up-front IT infrastructure costs. As well, third-party clouds enable organizations to focus on their core businesses instead of expending resources on computer infrastructure and maintenance.[4] Proponents also claim that cloud computing allows enterprises to get their applications up and running faster, with improved manageability and less maintenance, and that it enables IT teams to more rapidly adjust resources to meet fluctuating and unpredictable business demand.[4][5][6] Cloud providers typically use a "pay-as-you-go" model. This could lead to unexpectedly high charges if administrators are not familiarized with cloud-pricing models.[7]

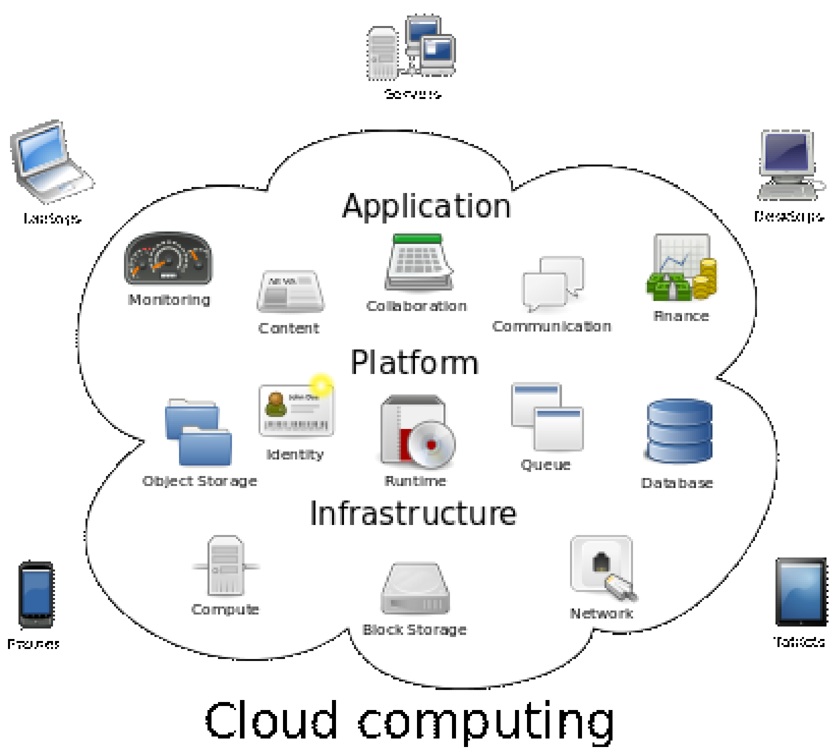


Figure 5Cloud computing metaphor: For a user, the network elements representing the provider-rendered services are invisible, as if obscured by a cloud.

In 2009 the availability of high-capacity networks, low-cost computers and storage devices as well as the widespread adoption of hardware virtualization, service-oriented architecture, and autonomic and utility computing led to a growth in cloud computing.[8][9][10] Companies can scale up as computing needs increase and then scale down again when demands decrease.[11] In 2013 it was reported that cloud computing had become a highly demanded service or utility due to the advantages of high computing power, cheap cost of services, high performance, scalability, and accessibility - as well as availability. Some cloud vendors experience growth rates of 50% per year,[12] but while cloud computing remains in a stage of infancy, it has pitfalls that need to be addressed to make cloud-computing services more reliable and user-friendly.[13][14]

20.2 HISTORY

20.2.1 Origin of the term

The origin of the term *cloud computing* is unclear. The word *cloud* is commonly used in science to describe a large agglomeration of objects that visually appear from a distance as a cloud and describes any set of things whose details are not further inspected in a given context.[15] Another explanation is that the old programs that drew network schematics surrounded the icons for servers with a circle, and a cluster of servers in a network diagram had several overlapping circles, which resembled a cloud.[16] In analogy to the above usage, the word *cloud* was used as a metaphor for the Internet and a standardized cloud-like shape was used to denote a network on telephony schematics. Later it was used to depict the Internet in computer network diagrams. With this simplification, the implication is that the specifics of how the end points of a network are connected are not relevant for the purposes of understanding the diagram. The cloud symbol was used to represent networks of computing equipment in the original ARPANET by as early as 1977,[17] and the CSNET by 1981[18]—both predecessors to the Internet itself.

The term *cloud* has been used to refer to platforms for distributed computing. In *Wired's* April 1994 feature "Bill and Andy's Excellent Adventure II" on the Apple spin-off General Magic, Andy Hertzfeld commented on General Magic's distributed programming language Telescript that:

> *"The beauty of Telescript ... is that now, instead of just having a device to program, we now have the entire Cloud out there, where a single program can go and travel to many different sources of information and create sort of a virtual service. No one had conceived that before. The example Jim White [the designer of Telescript, X.400 and ASN.1] uses now is a date-arranging service where a software agent goes to the flower store and orders flowers and then goes to the ticket shop and gets the tickets for the show, and everything is communicated to both parties[19]."*

References to "cloud computing" in its modern sense appeared as early as 1996, with the earliest known mention in a Compaq internal document.[20] The popularization of the term can be traced to 2006 when Amazon.com introduced its Elastic Compute Cloud.[21]

1970s

During the 1960s, the initial concepts of time-sharing became popularized via RJE (Remote Job Entry);[22] this terminology was mostly associated with large vendors such as IBM and DEC. Full time-sharing solutions were available by the early 1970s on such platforms as Multics (on GE hardware), Cambridge CTSS, and the earliest UNIX ports (on DEC hardware). Yet, the "data center" model where users submitted jobs to operators to run on IBM mainframes was overwhelmingly predominant.

1990s

In the 1990s, telecommunications companies, who previously offered primarily dedicated point-to-point data circuits, began offering virtual private network (VPN) services with comparable quality of service, but at a lower cost. By switching traffic as they saw fit to balance server use, they could use overall network bandwidth more effectively. They began to use the cloud symbol to denote the demarcation point between what the provider was responsible for and what users were responsible for. Cloud computing extended this boundary to cover all servers as well as the network infrastructure.[23] As computers became more diffused, scientists and technologists explored ways to make large-scale computing power available to more users through time-sharing. They experimented with algorithms to optimize the infrastructure, platform, and applications to prioritize CPUs and increase efficiency for end users.[24]

2000s

Since 2000, cloud computing has come into existence. In early 2008, NASA's OpenNebula, enhanced in the RESERVOIR European Commission-funded project, became the first open-source software for deploying private and hybrid clouds, and for the federation of clouds.[25] In the same year, efforts were focused on providing quality of service guarantees (as required by real-time interactive applications) to cloud-based infrastructures, in the framework of the IRMOS European Commission-funded project, resulting in a real-time cloud environment.[26][27] By mid-2008, Gartner saw an opportunity for cloud computing "to shape the relationship among consumers of IT services, those who use IT services and those who sell them"[28] and observed that "organizations are switching from company-owned hardware and software assets to per-use service-based models" so that the "projected shift to computing ... will result in dramatic growth in IT products in some areas and significant reductions in other areas."[29]

In August 2006 Amazon introduced its Elastic Compute Cloud.[21] Microsoft Azure was announced as "Azure" in October 2008 and was released on 1 February 2010 as Windows Azure, before being renamed to Microsoft Azure on 25 March 2014.[30] In July 2010, Rackspace Hosting and NASA jointly launched an open-source cloud- software initiative known as OpenStack. The OpenStack project intended to help organizations offering cloud- computing services running on standard hardware. The early code came from NASA's Nebula platform as well as from Rackspace's Cloud Files platform. As an open source offering and along with other open-source solutions such as CloudStack, Ganeti and OpenNebula, it has attracted attention by several key communities. Several studies aim at comparing these open sources offerings based on a set of criteria.[31] [32] [33] [34] [35] [36] [37]

On March 1, 2011, IBM announced the IBM SmartCloud framework to support Smarter Planet.[38] Among the various components of the Smarter Computing foundation, cloud computing is a critical part. On June 7, 2012, Oracle announced the Oracle Cloud.[39] While aspects of the Oracle Cloud are still in development, this cloud offering is poised to be the first to provide users with access to an integrated set of IT solutions, including the Applications (SaaS), Platform (PaaS), and Infrastructure (IaaS) layers.[40][41][42]

In April of 2008, Google released Google App Engine in beta.[43] In May of 2012, Google Compute Engine was released in preview, before being rolled out into General Availability in December of 2013.[44]

20.3 Similar Concepts

Cloud computing is the result of the evolution and adoption of existing technologies and paradigms. The goal of cloud computing is to allow users to take benefit from all of these technologies, without the need for deep knowledge about or expertise with each one of them. The cloud aims to cut costs, and helps the users focus on their core business instead of being impeded by IT obstacles.[45] The main enabling technology for cloud computing is virtualization. Virtualization software separates a physical computing device into one or more "virtual" devices, each of which can be easily used and managed to perform computing tasks. With operating system–level virtualization essentially creating a scalable system of multiple independent computing devices, idle computing resources can be allocated and used more efficiently. Virtualization provides the agility required to speed up IT operations, and reduces cost by increasing infrastructure utilization. Autonomic computing automates the process through which the user can provision resources on-demand. By minimizing user involvement, automation speeds up the process, reduces labor costs and reduces the possibility of human errors.[45] Users routinely face difficult business problems. Cloud computing adopts concepts from Service-oriented Architecture (SOA) that can help the user break these problems into services that can be integrated to provide

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

a solution. Cloud computing provides all of its resources as services, and makes use of the well-established standards and best practices gained in the domain of SOA to allow global and easy access to cloud services in a standardized way.

Cloud computing also leverages concepts from utility computing to provide metrics for the services used. Such metrics are at the core of the public cloud pay-per-use models. In addition, measured services are an essential part of the feedback loop in autonomic computing, allowing services to scale on-demand and to perform automatic failure recovery. Cloud computing is a kind of grid computing; it has evolved by addressing the QoS (quality of service) and reliability problems. Cloud computing provides the tools and technologies to build data/compute intensive parallel applications with much more affordable prices compared to traditional parallel computing techniques.[45]
Cloud computing shares characteristics with:

- Client–server model—*Client–server computing* refers broadly to any distributed application that distinguishes between service providers (servers) and service requestors (clients).[46]
- Computer bureau—A service bureau providing computer services, particularly from the 1960s to 1980s. Grid computing—"A form of distributed and parallel computing, whereby a 'super and virtual computer' is composed of a cluster of networked, loosely coupled computers acting in concert to perform very large tasks."
- Fog computing—Distributed computing paradigm that provides data, compute, storage and application services closer to client or near-user edge devices, such as network routers. Furthermore, fog computing handles data at the network level, on smart devices and on the end-user client side (e.g. mobile devices), instead of sending data to a remote location for processing.
- Dew computing—In the existing computing hierarchy, the Dew computing is positioned as the ground level for the cloud and fog computing paradigms. Compared to fog computing, which supports emerging IoT applications that demand real-time and predictable latency and the dynamic network reconfigurability, Dew computing pushes the frontiers to computing applications, data, and low level services away from centralized virtual nodes to the end users.[47]
- Mainframe computer—Powerful computers used mainly by large organizations for critical applications, typically bulk data processing such as: census; industry and consumer statistics; police and secret intelligence services; enterprise resource planning; and financial transaction processing.
- Utility computing—The "packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility, such as electricity."[48][49]
- Peer-to-peer—A distributed architecture without the need for central coordination. Participants are both suppliers and consumers of resources (in contrast to the traditional client–server model).
- Green computing
- Cloud sandbox—A live, isolated computer environment in which a program, code or file can run without affecting the application in which it runs.

20.4 CHARACTERISTICS

Cloud computing exhibits the following key characteristics:

- **Agility** for organizations may be improved, as cloud computing may increase users' flexibility with re-provisioning, adding, or expanding technological infrastructure resources.
- **Cost reductions** are claimed by cloud providers. A public-cloud delivery model converts capital expenditures (e.g., buying servers) to operational expenditure.[50] This purportedly lowers barriers to entry, as infrastructure is typically provided by a third party and need not be purchased for one-time or infrequent intensive computing tasks. Pricing on a utility computing basis is "fine-grained", with usage-based billing options. As well, less in-

house IT skills are required for implementation of projects that use cloud computing.[51] The e-FISCAL project's state-of-the-art repository[52] contains several articles looking into cost aspects in more detail, most of them concluding that costs savings depend on the type of activities supported and the type of infrastructure available in-house.

- **Device and location independence**[53] enable users to access systems using a web browser regardless of their location or what device they use (e.g., PC, mobile phone). As infrastructure is off-site (typically provided by a third-party) and accessed via the Internet, users can connect to it from anywhere.[51]
- **Maintenance** of cloud computing applications is easier, because they do not need to be installed on each user's computer and can be accessed from different places (e.g., different work locations, while travelling, etc.).
- **Multitenancy** enables sharing of resources and costs across a large pool of users thus allowing for:
    - **centralization** of infrastructure in locations with lower costs (such as real estate, electricity, etc.)
    - **peak-load capacity increases** (users need not engineer and pay for the resources and equipment to meet their highest possible load-levels)
    - **utilisation and efficiency** improvements for systems that are often only 10–20% utilised.[54][55] Performance is monitored by IT experts from the service provider, and consistent and loosely coupled architectures are constructed using web services as the system interface.[51][56][57]
- **Resource pooling** is the provider’s computing resources are commingle to serve multiple consumers using a multi-tenant model with different physical and virtual resources dynamically assigned and reassigned according to user demand. There is a sense of location independence in that the consumer generally have no control or knowledge over the exact location of the provided resource.[1] (https://www.uengr.com/cloud-computing-introduction-and-characteristics/)
- **Productivity** may be increased when multiple users can work on the same data simultaneously, rather than waiting for it to be saved and emailed. Time may be saved as information does not need to be re-entered when fields are matched, nor do users need to install application software upgrades to their computer.[58]
- **Reliability** improves with the use of multiple redundant sites, which makes well-designed cloud computing suitable for business continuity and disaster recovery.[59]
- **Scalability and elasticity** via dynamic ("on-demand") provisioning of resources on a fine-grained, self- service basis in near real-time[60][61] (Note, the VM startup time varies by VM type, location, OS and cloud providers[60]), without users having to engineer for peak loads.[62][63][64] This gives the ability to scale up when the usage need increases or down if resources are not being used.[65]
- **Security** can improve due to centralization of data, increased security-focused resources, etc., but concerns can persist about loss of control over certain sensitive data, and the lack of security for stored kernels. Security is often as good as or better than other traditional systems, in part because service providers are able to devote resources to solving security issues that many customers cannot afford to tackle or which they lack the technical skills to address.[66] However, the complexity of security is greatly increased when data is distributed over a wider area or over a greater number of devices, as well as in multi-tenant systems shared by unrelated users. In addition, user access to security audit logs may be difficult or impossible. Private cloud installations are in part motivated by users' desire to retain control over the infrastructure and avoid losing control of information security.

The National Institute of Standards and Technology's definition of cloud computing identifies "five essential characteristics":

- *On-demand self-service.* A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.

- *Broad network access.* Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).

- *Resource pooling.* The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.

- *Rapid elasticity.* Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear unlimited and can be appropriated in any quantity at any time.
- *Measured service.* Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

— National Institute of Standards and Technology[2]

20.5 Service Models

Though service-oriented architecture advocates "everything as a service" (with the acronyms **EaaS** or **XaaS**,[67] or simply **aas**),[68] cloud-computing providers offer their "services" according to different models, of which the three standard models per NIST are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).[2] These models offer increasing abstraction; they are thus often portrayed as a *layers* in a stack: infrastructure-, platform- and software-as-a-service,[69] but these need not be related. For example, one can provide SaaS implemented on physical machines (bare metal), without using underlying PaaS or IaaS layers, and conversely one can run a program on IaaS and access it directly, without wrapping it as SaaS.

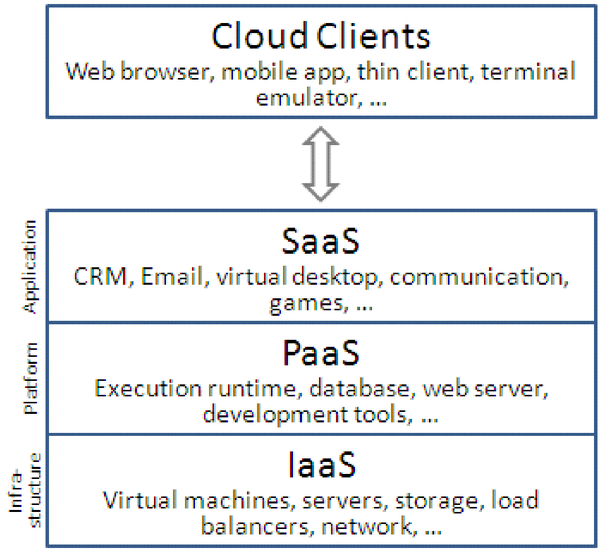
The NIST's definition of cloud computing defines the service models as follows:[2]

Software as a Service (SaaS). The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

Platform as a Service (PaaS). The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

Infrastructure as a Service (IaaS). The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls).

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.



Cloud computing service models arranged as layers in a stack

20.5.1 Infrastructure as a service (IaaS)

According to the Internet Engineering Task Force (IETF), the most basic cloud-service model is that of providers offering computing infrastructure – virtual machines and other resources – as a service to subscribers. **Infrastructure as a service** (IaaS) refers to online services that provide high-level APIs used to dereference various low-level details of underlying network infrastructure like physical computing resources, location, data partitioning, scaling, security, backup etc. A hypervisor, such as Xen, Oracle VirtualBox, Oracle VM, KVM, VMware ESX/ESXi, or Hyper-V, LXD, runs the virtual machines as guests. Pools of hypervisors within the cloud operational system can support large numbers of virtual machines and the ability to scale services up and down according to customers' varying requirements. Linux containers run in isolated partitions of a single Linux kernel running directly on the physical hardware. Linux groups and namespaces are the underlying Linux kernel technologies used to isolate, secure and manage the containers. Containerisation offers higher performance than virtualization, because there is no hypervisor overhead. Also, container capacity auto-scales dynamically with computing load, which eliminates the problem of over-provisioning and enables usage-based billing.[70] IaaS clouds often offer additional resources such as a virtual-machine disk-image library, raw block storage, file or object storage, firewalls, load balancers, IP addresses, virtual local area networks (VLANs), and software bundles.[71] IaaS-cloud providers supply these resources on-demand from their large pools of equipment installed in data centers. For wide-area connectivity, customers can use either the Internet or carrier clouds (dedicated virtual private networks). To deploy their applications, cloud users install operating-system images and their application software on the cloud infrastructure.[72] In this model, the cloud user patches and maintains the operating systems and the application software. Cloud providers typically bill IaaS services on a utility computing basis: cost reflects the amount of resources allocated and consumed.[73][74][75][76]

20.5.2 Platform as a service (PaaS)

PaaS vendors offer a development environment to application developers. The provider typically develops toolkit and standards for development and channels for distribution and payment. In the PaaS models, cloud providers deliver a computing platform, typically including operating system, programming-language execution environment, database, and web server. Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers. With some PaaS offers like Microsoft Azure and Google App Engine, the underlying computer and storage resources scale automatically to match application demand so that the cloud user does not have to allocate resources manually. The latter has also been proposed by an architecture aiming to facilitate real-time in cloud environments.[77] Even more specific application types can be provided via PaaS, such as media encoding as provided by services like bitcodin.com[78] or media.io.[79]

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Some integration and data management providers have also embraced specialized applications of PaaS as delivery models for data solutions. Examples include **iPaaS (Integration Platform as a Service)** and **dPaaS (Data Platform as a Service)**. iPaaS enables customers to develop, execute and govern integration flows.[80] Under the iPaaS integration model, customers drive the development and deployment of integrations without installing or managing any hardware or middleware.[81] dPaaS delivers integration—and data-management—products as a fully managed service.[82] Under the dPaaS model, the PaaS provider, not the customer, manages the development and execution of data solutions by building tailored data applications for the customer. dPaaS users retain transparency and control over data through data-visualization tools.[83] Platform as a Service (PaaS) consumers do not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but have control over the deployed applications and possibly configuration settings for the application-hosting environment.

A recent specialized PaaS is the **Blockchain as a Service (BaaS)**, that some vendors such as IBM Bluemix have already included in their PaaS offering.[84]

20.5.3 Software as a service (SaaS)

In the software as a service (SaaS) model, users gain access to application software and databases. Cloud providers manage the infrastructure and platforms that run the applications. SaaS is sometimes referred to as "on-demand software" and is usually priced on a pay-per-use basis or using a subscription fee.[85] In the SaaS model, cloud providers install and operate application software in the cloud and cloud users access the software from cloud clients. Cloud users do not manage the cloud infrastructure and platform where the application runs. This eliminates the need to install and run the application on the cloud user's own computers, which simplifies maintenance and support. Cloud applications differ from other applications in their scalability—which can be achieved by cloning tasks onto multiple virtual machines at run-time to meet changing work demand.[86] Load balancers distribute the work over the set of virtual machines. This process is transparent to the cloud user, who sees only a single access-point. To accommodate a large number of cloud users, cloud applications can be *multitenant*, meaning that any machine may serve more than one cloud-user organization.

The pricing model for SaaS applications is typically a monthly or yearly flat fee per user,[87] so prices become scalable and adjustable if users are added or removed at any point.[88] Proponents claim that SaaS gives a business the potential to reduce IT operational costs by outsourcing hardware and software maintenance and support to the cloud provider. This enables the business to reallocate IT operations costs away from hardware/software spending and from personnel expenses, towards meeting other goals. In addition, with applications hosted centrally, updates can be released without the need for users to install new software. One drawback of SaaS comes with storing the users' data on the cloud provider's server. As a result, there could be unauthorized access to the data. For this reason, users are increasingly adopting intelligent third-party key-management systems to help secure their data.

20.5.4 Security as a service (SECaaS)

Security as a service (SECaaS) is a business model in which a large service provider integrates their security services into a corporate infrastructure on a subscription basis more cost effectively than most individuals or corporations can provide on their own, when total cost of ownership is considered. In this scenario, security is delivered as a service from the cloud,[89] without requiring on-premises hardware avoiding substantial capital outlays. These security services often include authentication, anti-virus, anti-malware/spyware, intrusion detection, and security event management, among others.[90]

20.5.5 Mobile "backend" as a service (MBaaS)

In the mobile "backend" as a service (m) model, also known as backend as a service (BaaS), web app and mobile app developers are provided with a way to link their applications to cloud storage and cloud computing services with application programming interfaces (APIs) exposed to their applications and custom software development kits (SDKs). Services include user management, push notifications, integration with social networking services[91] and more. This is

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

a relatively recent model in cloud computing,[92] with most BaaS startups dating from 2011 or later[93][94][95] but trends indicate that these services are gaining significant mainstream traction with enterprise consumers.[96]

20.5.6 Serverless computing

Serverless computing is a cloud computing code execution model in which the cloud provider fully manages starting and stopping virtual machines as necessary to serve requests, and requests are billed by an abstract measure of the resources required to satisfy the request, rather than per virtual machine, per hour.[97] Despite the name, it does not actually involve running code without servers.[97] Serverless computing is so named because the business or person that owns the system does not have to purchase, rent or provision servers or virtual machines for the back-end code to run on.

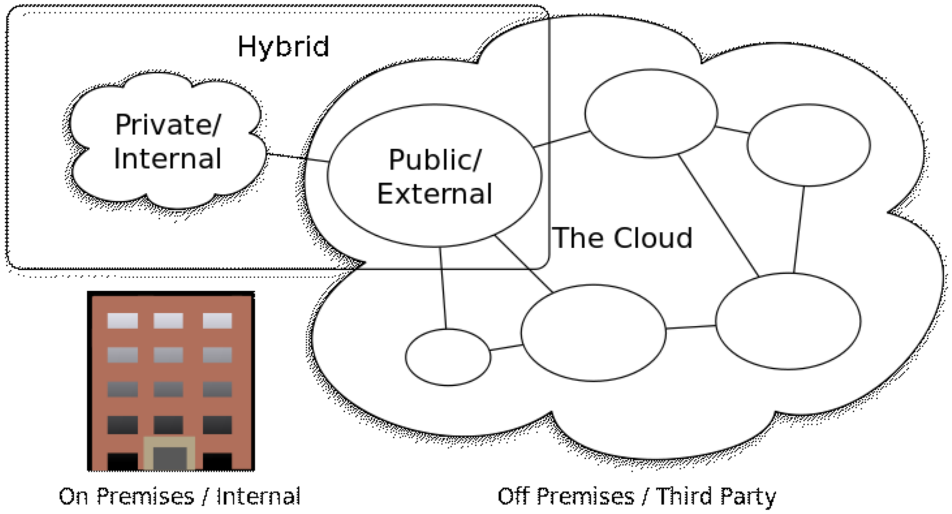
20.6 CLOUD CLIENTS

Users access cloud computing using networked client devices, such as desktop computers, laptops, tablets and smartphones and any Ethernet-enabled device such as Home Automation Gadgets. Some of these devices—cloud clients—rely on cloud computing for all or a majority of their applications so as to be essentially useless without it. Examples are thin clients and the browser-based Chromebook. Many cloud applications do not require specific software on the client and instead use a web browser to interact with the cloud application. With Ajax and HTML5 these Web user interfaces can achieve a similar, or even better, look and feel to native applications. Some cloud applications, however, support specific client software dedicated to these applications (e.g., virtual desktop clients and most email clients). Some legacy applications (line of business applications that until now have been prevalent in thin client computing) are delivered via a screen-sharing technology.

20.7 DEPLOYMENT MODELS

20.7.1 Private cloud

Private cloud is cloud infrastructure operated solely for a single organization, whether managed internally or by a third-party, and hosted either internally or externally.[2] Undertaking a private cloud project requires significant engagement to virtualize the business environment, and requires the organization to reevaluate decisions about existing resources. It can improve business, but every step in the project raises security issues that must be addressed to prevent serious vulnerabilities. Self-run data centers[98] are generally capital intensive. They have a significant physical footprint, requiring allocations of space, hardware, and environmental controls. These assets have to be refreshed periodically, resulting in additional capital expenditures. They have attracted criticism because users "still have to buy, build, and manage them" and thus do not benefit from less hands-on management,[99] essentially " [lacking] the economic model that makes cloud computing such an intriguing concept".[100][101]



20.7.2 Public cloud

A cloud is called a "public cloud" when the services are rendered over a network that is open for public use. Public cloud services may be free.[102] Technically there may be little or no difference between public and private cloud architecture, however, security consideration may be substantially different for services (applications, storage, and other resources) that are made available by a service provider for a public audience and when communication is effected over a non-trusted network. Generally, public cloud service providers like Amazon Web Services (AWS), Microsoft and Google own and operate the infrastructure at their data center and access is generally via the Internet. AWS and Microsoft also offer direct connect services called "AWS Direct Connect" and "Azure ExpressRoute" respectively, such connections require customers to purchase or lease a private connection to a peering point offered by the cloud provider.[51]

20.7.3 Hybrid cloud

Hybrid cloud is a composition of two or more clouds (private, community or public) that remain distinct entities but are bound together, offering the benefits of multiple deployment models. Hybrid cloud can also mean the ability to connect collocation, managed and/or dedicated services with cloud resources.[2] Gartner, Inc. defines a hybrid cloud service as a cloud computing service that is composed of some combination of private, public and community cloud services, from different service providers.[103] A hybrid cloud service crosses isolation and provider boundaries so that it can't be simply put in one category of private, public, or community cloud service. It allows one to extend either the capacity or the capability of a cloud service, by aggregation, integration or customization with another cloud service.

Varied use cases for hybrid cloud composition exist. For example, an organization may store sensitive client data in house on a private cloud application, but interconnect that application to a business intelligence application provided on a public cloud as a software service.[104] This example of hybrid cloud extends the capabilities of the enterprise to deliver a specific business service through the addition of externally available public cloud services. Hybrid cloud adoption depends on a number of factors such as data security and compliance requirements, level of control needed over data, and the applications an organization uses.[105]

Another example of hybrid cloud is one where IT organizations use public cloud computing resources to meet temporary capacity needs that can not be met by the private cloud.[106] This capability enables hybrid clouds to employ cloud bursting for scaling across clouds.[2] Cloud bursting is an application deployment model in which an application runs in a private cloud or data center and "bursts" to a public cloud when the demand for computing capacity increases. A primary advantage of cloud bursting and a hybrid cloud model is that an organization pays for extra compute resources only when they are needed.[107] Cloud bursting enables data centers to create an in- house IT infrastructure that supports average workloads, and use cloud resources from public or private clouds, during spikes in processing demands.[108] The specialized model of hybrid cloud, which is built atop heterogeneous hardware, is called "Cross-platform Hybrid Cloud". A cross-platform hybrid cloud is usually powered by different CPU architectures, for example, x86-64 and ARM, underneath. Users can transparently deploy and scale applications without knowledge of the cloud's hardware diversity.[109] This kind of cloud emerges from the raise of ARM-based system-on-chip for server-class computing.

20.7.4 Others

Community cloud

Community cloud shares infrastructure between several organizations from a specific community with common concerns (security, compliance, jurisdiction, etc.), whether managed internally or by a third-party, and either hosted internally or externally. The costs are spread over fewer users than a public cloud (but more than a private cloud), so only some of the cost savings potential of cloud computing are realized.[2]

Distributed cloud

A cloud computing platform can be assembled from a distributed set of machines in different locations, connected to a single network or hub service. It is possible to distinguish between two types of distributed clouds: public- resource computing and volunteer cloud.

- **Public-resource computing**—This type of distributed cloud results from an expansive definition of cloud computing, because they are more akin to distributed computing than cloud computing. Nonetheless, it is considered a sub-class of cloud computing, and some examples include distributed computing platforms such as BOINC and Folding@Home.

- **Volunteer cloud**—Volunteer cloud computing is characterized as the intersection of public-resource computing and cloud computing, where a cloud computing infrastructure is built using volunteered resources. Many challenges arise from this type of infrastructure, because of the volatility of the resources used to built it and the dynamic environment it operates in. It can also be called peer-to-peer clouds, or ad- hoc clouds. An interesting effort in such direction is Cloud@Home, it aims to implement a cloud computing infrastructure using volunteered resources providing a business-model to incentivize contributions through financial restitution.[110]

Intercloud

The Intercloud[111] is an interconnected global "cloud of clouds"[112][113] and an extension of the Internet "network of networks" on which it is based. The focus is on direct interoperability between public cloud service providers, more so than between providers and consumers (as is the case for hybrid- and multi-cloud).[114][115][116]

Multicloud

Multicloud is the use of multiple cloud computing services in a single heterogeneous architecture to reduce reliance on single vendors, increase flexibility through choice, mitigate against disasters, etc. It differs from hybrid cloud in that it refers to multiple cloud services, rather than multiple deployment modes (public, private, legacy).[117][118][119]

20.8 Architecture

Cloud architecture,[120] the systems architecture of the software systems involved in the delivery of cloud computing, typically involves multiple *cloud components* communicating with each other over a loose coupling mechanism such as a messaging queue. Elastic provision implies intelligence in the use of tight or loose coupling as applied to mechanisms such as these and others.

20.8.1 Cloud engineering

Cloud engineering is the application of engineering disciplines to cloud computing. It brings a systematic approach to the high-level concerns of commercialization, standardization, and governance in conceiving, developing, operating and maintaining cloud computing systems. It is a multidisciplinary method encompassing contributions from diverse areas such as systems, software, web, performance, information, security, platform, risk, and quality engineering.

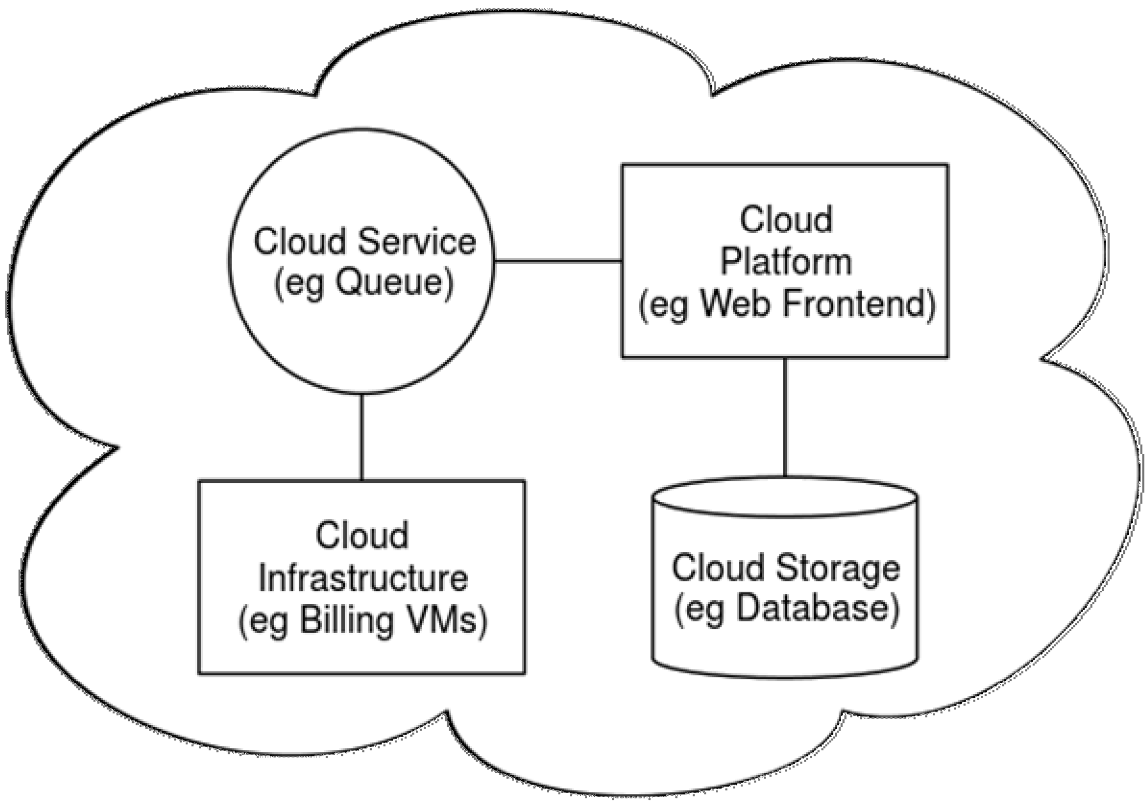


Figure 6. Cloud computing sample architecture

20.9 SECURITY AND PRIVACY

Cloud computing poses privacy concerns because the service provider can access the data that is in the cloud at any time. It could accidentally or deliberately alter or even delete information.[121] Many cloud providers can share information with third parties if necessary for purposes of law and order even without a warrant. That is permitted in their privacy policies, which users must agree to before they start using cloud services. Solutions to privacy include policy and legislation as well as end users' choices for how data is stored.[121] Users can encrypt data that is processed or stored within the cloud to prevent unauthorized access.[3][121]

According to the Cloud Security Alliance, the top three threats in the cloud are *Insecure Interfaces and API's*, *Data Loss & Leakage*, and *Hardware Failure*—which accounted for 29%, 25% and 10% of all cloud security outages respectively. Together, these form shared technology vulnerabilities. In a cloud provider platform being shared by different users there may be a possibility that information belonging to different customers resides on same data server. Therefore, Information leakage may arise by mistake when information for one customer is given to other.[122] Additionally,

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Eugene Schultz, chief technology officer at Emagined Security, said that hackers are spending substantial time and effort looking for ways to penetrate the cloud. "There are some real Achilles' heels in the cloud infrastructure that are making big holes for the bad guys to get into". Because data from hundreds or thousands of companies can be stored on large cloud servers, hackers can theoretically gain control of huge stores of information through a single attack—a process he called "hyperjacking". Some examples of this include the Dropbox security breach, and iCloud 2014 leak.[123] Dropbox had been breached in October 2014, having over 7 million of its users passwords stolen by hackers in an effort to get monetary value from it by Bitcoins (BTC). By having these passwords, they are able to read private data as well as have this data be indexed by search engines (making the information public).[123] There is the problem of legal ownership of the data (If a user stores some data in the cloud, can the cloud provider profit from it?). Many Terms of Service agreements are silent on the question of ownership.[124] Physical control of the computer equipment (private cloud) is more secure than having the equipment off site and under someone else's control (public cloud). This delivers great incentive to public cloud computing service providers to prioritize building and maintaining strong management of secure services.[125] Some small businesses that don't have expertise in IT security could find that it's more secure for them to use a public cloud. There is the risk that end users do not understand the issues involved when signing on to a cloud service (persons sometimes don't read the many pages of the terms of service agreement, and just click "Accept" without reading). This is important now that cloud computing is becoming popular and required for some services to work, for example for an intelligent personal assistant (Apple's Siri or Google Now). Fundamentally, private cloud is seen as more secure with higher levels of control for the owner, however public cloud is seen to be more flexible and requires less time and money investment from the user.[126]

20.10 LIMITATIONS AND DISADVANTAGES

According to Bruce Schneier, "The downside is that you will have limited customization options. Cloud computing is cheaper because of economics of scale, and – like any outsourced task – you tend to get what you get. A restaurant with a limited menu is cheaper than a personal chef who can cook anything you want. Fewer options at a much cheaper price: it's a feature, not a bug." He also suggests that "the cloud provider might not meet your legal needs" and that businesses need to weigh the benefits of cloud computing against the risks.[127] In cloud computing, the control of the back end infrastructure is limited to the cloud vendor only. Cloud providers often decide on the management policies, which moderates what the cloud users are able to do with their deployment.[128] Cloud users are also limited to the control and management of their applications, data and services.[129] This includes data caps, which are placed on cloud users by the cloud vendor allocating certain amount of bandwidth for each customer and are often shared among other cloud users.[130]

Privacy and confidentiality are big concerns in some activities. For instance, sworn translators working under the stipulations of an NDA, might face problems regarding sensitive data that are not encrypted.[131]

20.11 EMERGING TRENDS

Cloud computing is still a subject of research.[132] A driving factor in the evolution of cloud computing has been chief technology officers seeking to minimize risk of internal outages and mitigate the complexity of housing network and computing hardware in-house.[133] Major cloud technology companies invest billions of dollars per year in cloud Research and Development. For example, in 2011 Microsoft committed 90 percent of its $9.6 billion R&D budget to its cloud.[134] Research by investment bank Centaur Partners in late 2015 forecasted that SaaS revenue would grow from $13.5 billion in 2011 to $32.8 billion in 2016.[135] .[136]
https://www.pcworld.idg.com.au/article/614885/how-virtual-data-room-boosting-mergers-aquisitions/

20.12 SEE ALSO

- Category: Cloud computing providers Category: Cloud platforms

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

- Cloud computing security Cloud computing comparison Cloud management
- Cloud research Cloud storage Edge computing eScience Microservices
- Mobile cloud computing
- Personal cloud Robot as a Service
- Service-Oriented Architecture Ubiquitous computing
- Web computing Cloud collaboration

20.13 REFERENCES

1. Hassan, Qusay (2011). "Demystifying Cloud Computing" (http://www.crosstalkonline.org/storage/issue-archives/2011/201101/201101-Hassan.pdf) (PDF). The Journal of Defense Software Engineering. CrossTalk. **2011** (Jan/Feb): 16–21. Retrieved 11 December 2008.
2. Peter Mell and Timothy Grance (September 2011). *The NIST Definition of Cloud Computing* (https://dx.doi.org/10.6028/NIST.SP.800-145) (Technical report). National Institute of Standards and Technology: U.S. Department of Commerce. doi:10.6028/NIST.SP.800-145 (https://doi.org/10.6028%2FNIST.SP.800-145). Special publication 800-145.
3. M. Haghighat, S. Zonouz, & M. Abdel-Mottaleb (2015). CloudID: Trustworthy Cloud-based and Cross- Enterprise Biometric Identification (https://dx.doi.org/10.1016/j.eswa.2015.06.025). Expert Systems with Applications, 42(21), 7905–7916.
4. "What is Cloud Computing?" (http://aws.amazon.com/what-is-cloud-computing/). *Amazon Web Services*. 2013-03-19. Retrieved 2013-03-20.
5. Baburajan, Rajani (2011-08-24). "The Rising Cloud Storage Market Opportunity Strengthens Vendors" (http://it.tmcnet.com/channels/cloud-storage/articles/211183-rising-cloud-storage-market-opportunity-strengthen s-vendors.htm). It.tmcnet.com. Retrieved 2011-12-02.
6. Oestreich, Ken, (2010-11-15). "Converged Infrastructure" (https://web.archive.org/web/20120113094920/http://www.thectoforum.com/content/converged-infrastructure-0). *CTO Forum*. Thectoforum.com. Archived from the original (http://www.thectoforum.com/content/converged-infrastructure-0) on 2012-01-13. Retrieved 2011-12-02.
7. "Where's The Rub: Cloud Computing's Hidden Costs" (https://www.forbes.com/sites/centurylink/2014/02/2 7/wheres-the-rub-cloud-computings-hidden-costs/). 2014-02-27. Retrieved 2014-07-14.
8. "Cloud Computing: Clash of the clouds" (http://www.economist.com/displaystory.cfm?story_id=14637206). *The Economist*. 2009-10-15. Retrieved 2009-11-03.
9. "Gartner Says Cloud Computing Will Be As Influential As E-business" (http://www.gartner.com/it/page.jsp?i d=707508). Gartner. Retrieved 2010-08-22.
10. Gruman, Galen (2008-04-07). "What cloud computing really means" (http://www.infoworld.com/d/cloud-computing/what-cloud-computing-really-means-031). *InfoWorld*. Retrieved 2009-06-02.
11. Dealey, C. "Cloud Computing Working Group" (https://www.ncoic.org/technology/technical-team/cloud-co mputing-wg), *Network Centric Operations Industry Consortium - NCOIC* (http://www.ncoic.org/), 2013
12. "The economy is flat so why are financials Cloud vendors growing at more than 90 percent per annum?" (http://www.fsn.co.uk/channel_outsourcing/the_economy_is_flat_so_why_are_financials_cloud_vendors_growi ng_at_more_than_90_percent_per_annum#.UbmtsPlJPGA/). FSN. March 5, 2013.
13. "Realization of Interoperability & Portability Among Open Clouds by Using Agent's Mobility & Intelligence - TechRepublic" (http://www.techrepublic.com/resource-library/whitepapers/realization-of-interoperability-p ortability-among-open-clouds-by-using-agent-s-mobility-intelligence/). *TechRepublic*. Retrieved 2015-10-24.
14. "Interoperability and Portability among Open Clouds Using FIPA Agent / 978-3-659-24863-4 / 9783659248634 / 3659248630" (https://www.lap-publishing.com/catalog/details//store/gb/book/978-3-659-2 4863-4/interoperability-and-portability-among-open-clouds-using-fipa-agent). *www.lap-publishing.com*. Retrieved 24 October 2015.
15. Hassan, Qusay F.; Riad, laa M.; Hassan, Ahmed E. (2012). "Software reuse in the emerging cloud computing era" (http://www.igi-global.com/chapter/software-reuse-emerging-cloud-computing/65173). In Yang, Hongji; Liu, Xiaodong. *Understanding Cloud Computing* (PDF). Hershey, PA: Information Science Reference. pp. 204–227. ISBN 978-1-4666-0897-9. doi:10.4018/978-1-4666-0897-9.ch009 (https://doi.org/1 0.4018%2F978-1-4666-0897-9.ch009). Retrieved 11 December 2014.
16. Schmidt, Eric; Rosenberg, Jonathan (2014). *How Google Works* (http://www.howgoogleworks.net/). Grand Central

Publishing. p. 11. ISBN 978-1-4555-6059-2.
17. "Internet History 1977"(http://www.computerhistory.org/internet_history/internet_history_70s.html).
18. "National Science Foundation, "Diagram of CSNET," 1981" (http://gu.friends-partners.org/Bookwriting/PAR T_I/Chapter_I/Total/Insertions/NSF/CSNET/CSNET.html).
19. Steven Levy (April 1994). "Bill and Andy's Excellent Adventure II" (http://archive.wired.com/wired/archive/ 2.04/general.magic_pr.html). *Wired*.
20. Antonio Regalado (31 October 2011)."Who Coined 'Cloud Computing'?" (http://www.technologyreview.co m/news/425970/who-coined-cloud-computing/). *Technology Review*. MIT. Retrieved 31 July 2013.
21. "Announcing Amazon Elastic Compute Cloud (Amazon EC2) - beta" (https://aws.amazon.com/about-aws/w hats-new/2006/08/24/announcing-amazon-elastic-compute-cloud-amazon-ec2---beta/). Amazon.com. 24 August 2006. Retrieved 31 May 2014.
22. White, J.E. "Network Specifications for Remote Job Entry and Remote Job Output Retrieval at UCSB" (http s://tools.ietf.org/html/rfc105). *tools.ietf.org*. Retrieved 2016-03-21.
23. "July, 1993 meeting report from the IP over ATM working group of the IETF" (https://archive.is/201207101 70149/http://mirror.switch.ch/ftp/doc/ietf/ipatm/atm-minutes-93jul.txt). CH: Switch. Archived from the original (http://mirror.switch.ch/ftp/doc/ietf/ipatm/atm-minutes-93jul.txt) on 2012-07-10. Retrieved 2010-08-22.
24. Corbató, Fernando J. "An Experimental Time-Sharing System" (http://larch-www.lcs.mit.edu:8001/~corbato/ sjcc62/). *SJCC Proceedings*. MIT. Retrieved 3 July 2012.
25. Rochwerger, B.; Breitgand, D.; Levy, E.; Galis, A.; Nagin, K.; Llorente, I. M.; Montero, R.; Wolfsthal, Y.; Elmroth, E.; Caceres, J.; Ben-Yehuda, M.; Emmerich, W.; Galan, F. "The Reservoir model and architecture for open federated cloud computing". *IBM Journal of Research and Development*. **53** (4): 4:1–4:11. doi:10.1147/JRD.2009.5429058 (https://doi.org/10.1147%2FJRD.2009.5429058).
26. Kyriazis, D; A Menychtas; G Kousiouris; K Oberle; T Voith; M Boniface; E Oliveros; T Cucinotta; S Berger (November 2010). "A Real-time Service Oriented Infrastructure". *International Conference on Real-Time and Embedded Systems (RTES 2010)*. Singapore.
27. Gogouvitis, Spyridon; Konstanteli, Kleopatra; Waldschmidt, Stefan; Kousiouris, George; Katsaros, Gregory; Menychtas, Andreas; Kyriazis, Dimosthenis; Varvarigou, Theodora (2012). "Workflow management for soft real-time interactive applications in virtualized environments". *Future Generation Computer Systems*. **28** (1): 193–209. ISSN 0167-739X (https://www.worldcat.org/issn/0167-739X). doi:10.1016/j.future.2011.05.017 (h ttps://doi.org/10.1016%2Fj.future.2011.05.017).
28. Keep an eye on cloud computing (http://www.networkworld.com/newsletters/itlead/2008/070708itlead1.htm l), Amy Schurr, Network World, 2008-07-08, citing the Gartner report, "Cloud Computing Confusion Leads to Opportunity". Retrieved 2009-09-11.
29. Gartner (2008-08-18). "Gartner Says Worldwide IT Spending on Pace to Surpass Trillion in 2008" (http://w ww.gartner.com/it/page.jsp?id=742913).
30. "Windows Azure General Availability" (http://blogs.technet.com/b/microsoft_blog/archive/2010/02/01/wind ows-azure-general-availability.aspx). The Official Microsoft Blog. Microsoft. 2010-02-01. Retrieved 2015-05-03.
31. DZone, "Apache CloudStack vs. OpenStack: Which Is the Best?", https://dzone.com/articles/apache- cloudstack-vs-openstack-which-is-the-best
32. SoftwareInsider, "OpenNebula vs OpenStack", http://cloud-computing.softwareinsider.com/compare/70-322/OpenNebula-Project-vs-OpenStack-Foundation
33. Kostantos, Konstantinos, et al. "OPEN-source IaaS fit for purpose: a comparison between OpenNebula and OpenStack." International Journal of Electronic Business Management 11.3 (2013)
34. L. Albertson, "OpenStack vs. Ganeti", LinuxFest Northwest 2017
35. Qevani, Elton, et al. "What can OpenStack adopt from a Ganeti-based open-source IaaS?." Cloud Computing (CLOUD), 2014 IEEE 7th International Conference on. IEEE, 2014
36. Von Laszewski, Gregor, et al. "Comparison of multiple cloud frameworks.", IEEE 5th International Conference on Cloud Computing (CLOUD), 2012.
37. Diaz, Javier et al. " Abstract Image Management and Universal Image Registration for Cloud and HPC Infrastructures ", IEEE 5th International Conference on Cloud Computing (CLOUD), 2012
38. "Launch of IBM Smarter Computing" (https://web.archive.org/web/20130420162543/https://www-304.ibm.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

com/connections/blogs/IBMSmarterSystems/date/201102?lang=en_us). Archived from the original (https:// www-304.ibm.com/connections/blogs/IBMSmarterSystems/date/201102?lang=en_us) on 20 April 2013. Retrieved 1 March 2011.

39. "Launch of Oracle Cloud" (https://www.theregister.co.uk/2012/06/07/oracle_cloud_rehash_platinum_service s/). Retrieved 28 February 2014.
40. "Oracle Cloud, Enterprise-Grade Cloud Solutions: SaaS, PaaS, and IaaS" (https://cloud.oracle.com/home). Retrieved 12 October 2014.
41. "Larry Ellison Doesn't Get the Cloud: The Dumbest Idea of 2013" (https://www.forbes.com/sites/oracle/201 2/10/09/larry-ellison-doesnt-get-the-cloud-the-dumbest-idea-of-2013/). Forbes.com. Retrieved 12 October 2014.
42. "Oracle Disrupts Cloud Industry with End-to-End Approach" (https://www.forbes.com/sites/oracle/2012/10/ 03/oracle-disrupts-cloud-industry-with-end-to-end-approach/). Forbes.com. Retrieved 12 October 2014.
43. "Introducing Google App Engine + our new blog" (http://googleappengine.blogspot.nl/2008/04/introducing- google-app-engine-our-new.html). Google Developer Blog. 2008-04-07. Retrieved 2017-03-07.
44. "Google Compute Engine is now Generally Available with expanded OS support, transparent maintenance, and lower prices" (http://googledevelopers.blogspot.nl/2013/12/google-compute-engine-is-now-generally.ht ml). Google Developers Blog. 2013-12-02. Retrieved 2017-03-07.
45. HAMDAQA, Mohammad (2012). *Cloud Computing Uncovered: A Research Landscape* (http://www.stargro up.uwaterloo.ca/~mhamdaqa/publications/Cloud_Computing_Uncovered.pdf) (PDF). Elsevier Press. pp. 41– 85. ISBN 0-12-396535-7.
46. "Distributed Application Architecture" (http://java.sun.com/developer/Books/jdbc/ch07.pdf) (PDF). Sun Microsystem. Retrieved 2009-06-16.
47. Skala, Karolj; Davidović, Davor; Afgan, Enis; Sović, Ivan; Šojat, Zorislav (2015-12-31). "Scalable Distributed Computing Hierarchy: Cloud, Fog and Dew Computing" (http://www.ronpub.com/publications/o jcc/OJCC_2015v2i1n03_Skala.html). *Open Journal of Cloud Computing*. RobPub. **2** (1): 16–24. ISSN 2199- 1987 (https://www.worldcat.org/issn/2199-1987).
48. "It's probable that you've misunderstood 'Cloud Computing' until now" (http://portal.acm.org/citation.cfm?id=1496091.1496100&coll=&dl=ACM&CFID=21518680&CFTOKEN=18800807). TechPluto. Retrieved 2010-09-14.
49. Danielson, Krissi (2008-03-26). "Distinguishing Cloud Computing from Utility Computing" (http://www.ebi zq.net/blogs/saasweek/2008/03/distinguishing_cloud_computing/). Ebizq.net. Retrieved 2010-08-22.
50. "Recession Is Good For Cloud Computing – Microsoft Agrees" (http://www.cloudave.com/link/recession-is- good-for-cloud-computing-microsoft-agrees). CloudAve. Retrieved 2010-08-22.
51. "Defining 'Cloud Services' and "Cloud Computing" " (http://blogs.idc.com/ie/?p=190). IDC. 2008-09-23. Retrieved 2010-08-22.
52. "e-FISCAL project state of the art repository" (http://www.efiscal.eu/state-of-the-art).
53. Farber, Dan (2008-06-25). "The new geek chic: Data centers" (http://news.cnet.com/8301-13953_3-9977049-80.html). CNET News. Retrieved 2010-08-22.
54. "Jeff Bezos' Risky Bet" (http://www.businessweek.com/magazine/content/06_46/b4009001.htm). *Business Week*.
55. He, Sijin; Guo, L.; Guo, Y.; Ghanem, M. "Improving Resource Utilisation in the Cloud Environment Using Multivariate Probabilistic Models" (http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6253553). 2012 2012 IEEE 5th International Conference on Cloud Computing (CLOUD): 574–581. ISBN 978-1-4673-2892- 0. doi:10.1109/CLOUD.2012.66 (https://doi.org/10.1109%2FCLOUD.2012.66).
56. He, Qiang, et al. "Formulating Cost-Effective Monitoring Strategies for Service-based Systems." (2013): 1-1.
57. A Self-adaptive hierarchical monitoring mechanism for Clouds Elsevier.com (https://dx.doi.org/10.1016/j.js s.2011.11.1043)
58. Heather Smith (23 May 2013). *Xero For Dummies* (https://books.google.com/books?id=drOF19aBKfgC&pg=PT37). John Wiley & Sons. pp. 37–. ISBN 978-1-118-57252-8.
59. King, Rachael (2008-08-04). "Cloud Computing: Small Companies Take Flight" (http://www.businessweek. com/technology/content/aug2008/tc2008083_619516.htm). *Bloomberg BusinessWeek*. Retrieved 2010-08-22.
60. Mao, Ming; M. Humphrey (2012). "A Performance Study on the VM Startup Time in the Cloud" (http://ieee xplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6253534&isnumber=6253471). *Proceedings of 2012 IEEE 5th International Conference on Cloud Computing (Cloud2012)*: 423. ISBN 978-1-4673-2892-0. doi:10.1109/CLOUD.2012.103 (https://doi.org/10.1109%2FCLOUD.2012.103).
61. Dario Bruneo, Salvatore Distefano, Francesco Longo, Antonio Puliafito, Marco Scarpa: Workload-Based Software

Rejuvenation in Cloud Systems (http://doi.ieeecomputersociety.org/10.1109/TC.2013.30). IEEE Trans. Computers 62(6): 1072–1085 (2013)

62. "Defining and Measuring Cloud Elasticity" (http://digbib.ubka.uni-karlsruhe.de/volltexte/1000023476). KIT Software Quality Departement. Retrieved 13 August 2011.
63. "Economies of Cloud Scale Infrastructure" (https://www.youtube.com/watch?v=nfDsY3f4nVI). Cloud Slam 2011. Retrieved 13 May 2011.
64. He, Sijin; L. Guo; Y. Guo; C. Wu; M. Ghanem; R. Han. "Elastic Application Container: A Lightweight Approach for Cloud Resource Provisioning" (http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6184989). 2012 IEEE 26th International Conference on Advanced Information Networking and Applications (AINA): 15–22. ISBN 978-1-4673-0714-7. doi:10.1109/AINA.2012.74 (https://doi.org/10.1109%2FAINA.2012.74).
65. Marston, Sean; Li, Zhi; Bandyopadhyay, Subhajyoti; Zhang, Juheng; Ghalsasi, Anand (2011-04-01). "Cloud computing – The business perspective" (http://www.sciencedirect.com/science/article/pii/S01679236100023 93). *Decision Support Systems*. **51** (1): 176–189. doi:10.1016/j.dss.2010.12.006 (https://doi.org/10.1016%2F j.dss.2010.12.006).
66. Mills, Elinor (2009-01-27). "Cloud computing security forecast: Clear skies" (http://news.cnet.com/8301-10 09_3-10150569-83.html). CNET News. Retrieved 2010-08-22.
67. Duan, Yucong; Fu, Guohua; Zhou, Nianjun; Sun, Xiaobing; Narendra, Nanjangud; Hu, Bo. "Everything as a Service (XaaS) on the Cloud: Origins, Current and Future Trends" (http://ieeexplore.ieee.org/document/7214 098/). IEEE.
68. Kurdi, Heba; Li, Maozhen; Al-Raweshidy, H. S. (2010). "Taxonomy of Grid Systems". In Antonopoulos, Nick. *Handbook of Research on P2P and Grid Systems for Service-Oriented Computing: Models, Methodologies and Applications* (https://books.google.com/books?id=WWsNz5V6aEcC). IGI Global research collection. IGI Global. p. 34. ISBN 978-1-61520-687-2. Retrieved 2015-07-29. "Nowadays Service-Oriented Architecture (SOA) has become as [sic] the main architectural model of many IT initiatives including grid, cloud and everything as a service (Essa\XaaS\aas) computing."
69. Alcaraz Calero, Jose M.; König, Benjamin; Kirschnick, Johannes (2012). "Cross-Layer Monitoring in Cloud Computing". In Rashvand, Habib F.; Kavian, Yousef S. *Using Cross-Layer Techniques for Communication Systems* (https://books.google.com/books?id=0KWeBQAAQBAJ). Premier reference source. IGI Global. p. 329. ISBN 978-1-4666-0961-7. Retrieved 2015-07-29. "*Cloud Computing* provides services on a stack composed of three service layers (Hurwitz, Bloor, Kaufman, & Halper, 2009): *Infrastructure as a Service* (*IaaS*), *Platform as a Service* (*PaaS*), and *Software as a Service* (*SaaS*)."
70. "ElasticHosts Blog" (https://www.elastichosts.com/blog/elastic-containers/#thetechnicaldetails). *Elastichosts*. 2014-04-01. Retrieved 2016-06-02.
71. Amies, Alex; Sluiman, Harm; Tong, Qiang Guo; Liu, Guo Ning (July 2012). "*Infrastructure as a Service Cloud Concepts*". *Developing and Hosting Applications on the Cloud* (http://www.ibmpressbooks.com/boo kstore/product.asp?isbn=9780133066845). IBM Press. ISBN 978-0-13-306684-5.
72. Ananich, Anthony (February 20, 2016). "What is IaaS?" (http://ananich.pro/2016/02/what-is-iaas/). *ananich.pro*. Retrieved 2016-02-20.
73. "Amazon EC2 Pricing" (http://aws.amazon.com/ec2/pricing/). *aws.amazon.com*. Retrieved 7 July 2014.
74. "Compute Engine Pricing" (https://cloud.google.com/products/compute-engine/#pricing). *cloud.google.com*. Retrieved 7 July 2014.
75. "Microsoft Azure Virtual Machines Pricing Details" (http://azure.microsoft.com/en-us/pricing/details/virtual- machines/). *azure.microsoft.com*. Retrieved 7 July 2014.
76. "cloud.ca" (https://cloud.ca/).
77. Boniface, M.; et al. (2010), *Platform-as-a-Service Architecture for Real-Time Quality of Service Management in Clouds* (https://dx.doi.org/10.1109/ICIW.2010.91), 5th International Conference on Internet and Web Applications and Services (ICIW), Barcelona, Spain: IEEE, pp. 155–160, doi:10.1109/ICIW.2010.91 (https://doi.org/10.1109%2FICIW.2010.91)
78. "bitcodin – cloud based transcoding and streaming" (http://www.bitcodin.com). Retrieved 22 April 2015.
79. "media.io" (http://www.media.io/). *Media.io*. Retrieved 21 July 2017.
80. Gartner. "Gartner IT Glossary" (http://www.gartner.com/it-glossary/information-platform-as-a-service- ipaas). Retrieved 6 July 2015.
81. Gartner; Massimo Pezzini; Paolo Malinverno; Eric Thoo. "Gartner Reference Model for Integration PaaS" (h ttp://www.gartner.com/id=1729256). Retrieved 16 January 2013.
82. Loraine Lawson. "IT Business Edge" (http://www.itbusinessedge.com/blogs/integration/liaisons-data-platfor m-as-a-service-includes-data-mapping.html). Retrieved 6 July 2015.

83. Enterprise CIO Forum; Gabriel Lowy. "The Value of Data Platform-as-a-Service (dPaaS)" (https://web.archi ve.org/web/20150419045605/http://www.enterprisecioforum.com/en/blogs/gabriellowy/value-data-platform- service-dpaas). Archived from the original (http://www.enterprisecioforum.com/en/blogs/gabriellowy/value- data-platform-service-dpaas) on 19 April 2015. Retrieved 6 July 2015.
84. "Blockchain as a Service (BaaS) | Microsoft Azure" (https://azure.microsoft.com/en-gb/solutions/blockchai n/). *azure.microsoft.com*. Retrieved 2016-08-22.
85. "Definition of: SaaS" (https://www.pcmag.com/encyclopedia/term/56112/saas). *PC Magazine Encyclopedia*. Ziff Davis. Retrieved 14 May 2014.
86. Hamdaqa, Mohammad. *A Reference Model for Developing Cloud Applications* (http://www.stargroup.uwater loo.ca/~mhamdaqa/publications/A%20REFERENCEMODELFORDEVELOPINGCLOUD%20APPLICATI ONS.pdf) (PDF).
87. Chou, Timothy. *Introduction to Cloud Computing: Business & Technology* (https://www.scribd.com/doc/646 99897/Introduction-to-Cloud-Computing-Business-and-Technology).
88. "HVD: the cloud's silver lining" (http://www.intrinsictechnology.co.uk/FileUploads/HVD_Whitepaper.pdf) (PDF). Intrinsic Technology. Retrieved 30 August 2012.
89. Furfaro, A.; Garro, A.; Tundis, A. (2014-10-01). "Towards Security as a Service (SecaaS): On the modeling of Security Services for Cloud Computing" (http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6986 995). *2014 International Carnahan Conference on Security Technology (ICCST)*: 1–6. doi:10.1109/CCST.2014.6986995 (https://doi.org/10.1109%2FCCST.2014.6986995).
90. "Definition of Security as a Service" (http://searchsecurity.techtarget.com/definition/Security-as-a-Service).
91. Carney, Michael. "AnyPresence partners with Heroku to beef up its enterprise mBaaS offering" (http://pando daily.com/2013/06/24/anypresence-partners-with-heroku-to-beef-up-its-enterprise-mbaas-offering/). PandoDaily. Retrieved 24 June 2013.
92. Alex Williams (11 October 2012). "Kii Cloud Opens Doors For Mobile Developer Platform With 25 Million End Users" (https://techcrunch.com/2012/10/11/kii-cloud-opens-doors-for-mobile-developer-platform-with-2 5-million-end-users/). *TechCrunch*. Retrieved 16 October 2012.
93. Aaron Tan (30 September 2012). "FatFractal ups the ante in backend-as-a-service market" (http://www.techg oondu.com/2012/09/30/fatfractal-ups-the-ante-in-backend-as-a-service/). *Techgoondu.com*. Retrieved 16 October 2012.
94. Dan Rowinski (9 November 2011). "Mobile Backend As A Service Parse Raises $5.5 Million in Series A Funding" (http://readwrite.com/2011/11/09/mobile-backend-as-a-service-pa). *ReadWrite*. Retrieved 23 October 2012.
95. Pankaj Mishra (7 January 2014). "MobStac Raises $2 Million in Series B To Help Brands Leverage Mobile Commerce" (https://techcrunch.com/2014/01/07/mobstac-raises-2-million-in-series-b-to-help-brands-leverag e-mobile-commerce/). *TechCrunch*. Retrieved 22 May 2014.
96. "built.io Is Building an Enterprise MBaas Platform for IoT" (http://blog.programmableweb.com/2014/03/03/ built-io-is-building-an-enterprise-mbaas-platform-for-iot/). *programmableweb*. Retrieved 3 March 2014.
97. Miller, Ron (24 Nov 2015). "AWS Lambda Makes Serverless Applications A Reality" (https://techcrunch.co m/2015/11/24/aws-lamda-makes-serverless-applications-a-reality/). *TechCrunch*. Retrieved 10 July 2016.
98. "Self-Run Private Cloud Computing Solution – GovConnection"(http://www.govconnection.com/IPA/PM/In fo/Cloud-Computing/Self-Run-Private-Cloud.htm). *govconnection.com*. 2014. Retrieved April 15, 2014.
99. Foley, John. "Private Clouds Take Shape" (http://www.informationweek.com/news/services/business/showA rticle.jhtml?articleID=209904474). *InformationWeek*. Retrieved 2010-08-22.
100. Haff, Gordon (2009-01-27). "Just don't call them private clouds" (http://news.cnet.com/8301-13556_3-10150 841-61.html). CNET News. Retrieved 2010-08-22.
101. "There's No Such Thing As A Private Cloud" (http://www.informationweek.com/cloud-computing/theres-no-such-thing-as-a-private-cloud/229207922). *InformationWeek*. 2010-06-30. Retrieved 2010-08-22.
102. Rouse, Margaret. "What is public cloud?" (http://searchcloudcomputing.techtarget.com/definition/public-clo ud). Definition from Whatis.com. Retrieved 12 October 2014.
103. "Mind the Gap: Here Comes Hybrid Cloud – Thomas Bittman" (http://blogs.gartner.com/thomas_bittman/20 12/09/24/mind-the-gap-here-comes-hybrid-cloud/). *Thomas Bittman*. Retrieved 22 April 2015.
104. "Business Intelligence Takes to Cloud for Small Businesses" (http://www.cio.com/article/753708/Business_I ntelligence_Takes_to_Cloud_for_Small_Businesses). CIO.com. 2014-06-04. Retrieved 2014-06-04.
105. Désiré Athow. "Hybrid cloud: is it right for your business?" (http://www.techradar.com/news/internet/cloud-s

ervices/hybrid-cloud-is-it-right-for-your-business--1261343). *TechRadar*. Retrieved 22 April 2015.

106. Metzler, Jim; Taylor, Steve. (2010-08-23) "Cloud computing: Reality vs. fiction" (http://www.networkworld.com/newsletters/frame/2010/082310wan1.html), Network World.

107. Rouse, Margaret. "Definition: Cloudbursting" (http://searchcloudcomputing.techtarget.com/definition/cloud- bursting), May 2011. SearchCloudComputing.com.

108. "How Cloudbursting "Rightsizes" the Data Center" (http://insights.dice.com/2012/06/21/how-cloudbursting- rightsizes-the-data-center/).

109. Kaewkasi, Chanwit (3 May 2015). "Cross-Platform Hybrid Cloud with Docker" (https://medium.com/@chan wit/cross-platform-hybrid-cloud-with-docker-ded000f792fb).

110. Vincenzo D. Cunsolo, Salvatore Distefano, Antonio Puliafito, Marco Scarpa: Volunteer Computing and Desktop Cloud: The Cloud@Home Paradigm (http://doi.ieeecomputersociety.org/10.1109/NCA.2009.41). IEEE International Symposium on Network Computing and Applications, NCA 2009, pp 134–139

111. Bernstein, David; Ludvigson, Erik; Sankar, Krishna; Diamond, Steve; Morrow, Monique (2009-05-24). "Blueprint for the Intercloud – Protocols and Formats for Cloud Computing Interoperability" (http://www2.c omputer.org/portal/web/csdl/doi/10.1109/ICIW.2009.55). IEEE Computer Society: 328–336. ISBN 978-1- 4244-3851-8. doi:10.1109/ICIW.2009.55(https://doi.org/10.1109%2FICIW.2009.55).

112. "Kevin Kelly: A Cloudbook for the Cloud" (http://www.kk.org/thetechnium/archives/2007/11/a_cloudbook_ for.php). Kk.org. Retrieved 2010-08-22.

113. "Intercloud is a global cloud of clouds" (http://samj.net/2009/06/22/the-intercloud-is-a-global-cloud-of-cloud s/). Samj.net. 2009-06-22. Retrieved 2010-08-22.

114. "Vint Cerf: Despite Its Age, The Internet is Still Filled with Problems" (http://www.readwriteweb.com/archiv es/vint_cerf_despite_its_age_the.php?mtcCampaign=2765). Readwriteweb.com. Retrieved 2010-08-22.

115. "SP360: Service Provider: From India to Intercloud" (http://blogs.cisco.com/sp/comments/from_india_to_int ercloud/). Blogs.cisco.com. Retrieved 2010-08-22.

116. Canada (2007-11-29). "Head in the clouds? Welcome to the future" (https://www.theglobeandmail.com/techn ology/head-in-the-clouds-welcome-to-the-future/article698926/). *The Globe and Mail*. Toronto. Retrieved 2010-08-22.

117. Rouse, Margaret. "What is a multi-cloud strategy" (http://searchcloudapplications.techtarget.com/definition/ multi-cloud-strategy). SearchCloudApplications. Retrieved 3 July 2014.

118. King, Rachel. "Pivotal's head of products: We're moving to a multi-cloud world" (http://www.zdnet.com/pivo tals-head-of-products-were-moving-to-a-multi-cloud-world-7000030737/). ZDnet. Retrieved 3 July 2014.

119. Multcloud manage multiple cloud accounts (http://www.groovypost.com/reviews/multcloud-manage-multipl e-cloud-accounts/). Retrieved on 06 August 2014

120. "Building GrepTheWeb in the Cloud, Part 1: Cloud Architectures" (http://developer.amazonwebservices.co m/connect/entry.jspa?externalID=1632&categoryID=100). Developer.amazonwebservices.com. Retrieved 22 August 2010.

121. "Cloud Computing Privacy Concerns on Our Doorstep" (http://cacm.acm.org/magazines/2011/1/103200-clou d-computing-privacy-concerns-on-our-doorstep/fulltext).

122. Chhibber, A (2013). "SECURITY ANALYSIS OF CLOUD COMPUTING" (http://garph.co.uk/IJAREAS/M ar2013/6.pdf) (PDF). *International Journal of Advanced Research in Engineering and Applied Sciences*. **2** (3): 2278–6252. Retrieved 27 February 2015.

123. "Google Drive, Dropbox, Box and iCloud Reach the Top 5 Cloud Storage Security Breaches List" (https://ps g.hitachi-solutions.com/credeon/blog/google-drive-dropbox-box-and-icloud-reach-the-top-5-cloud-storage-s ecurity-breaches-list). *psg.hitachi-solutions.com*. Retrieved 2015-11-22.

124. Maltais, Michelle (26 April 2012). "Who owns your stuff in the cloud?" (http://articles.latimes.com/2012/ap r/26/business/la-fi-tech-savvy-cloud-services-20120426). *Los Angeles Times*. Retrieved 2012-12-14.

125. "Security of virtualization, cloud computing divides IT and security pros" (https://www.networkworld.com/a rticle/2244954/virtualization/security-of-virtualization--cloud-computing-divides-it-and-security-pros.html). Network World. 2010-02-22. Retrieved 2010-08-22.

126. "The Bumpy Road to Private Clouds" (http://www.computerworld.com/article/2549867/data-center/the-bum py-road-to-private-clouds.html). Retrieved 8 October 2014.

127. "Should Companies Do Most of Their Computing in the Cloud? (Part 1) - Schneier on Security" (https://ww w.schneier.com/blog/archives/2015/06/should_companie.html). *www.schneier.com*. Retrieved 2016-02-28.

128. "Disadvantages of Cloud Computing (Part 1) - Limited control and flexibility" (https://cloudacademy.com/blog/disadvantages-of-cloud-computing/). *www.cloudacademy.com*. Retrieved 2016-11-03.

129. "The real limits of cloud computing" (https://www.itworld.com/article/2726566/cloud-computing/the-real-li mits-of-cloud-computing.html). *www.itworld.com*. Retrieved 2016-11-03.

130. "The real limits of cloud computing" (https://www.itworld.com/article/2726566/cloud-computing/the-real-li mits-of-cloud-computing.html). *www.itworld.com*. Retrieved 2016-11-03.

131. Karra, Maria. "Cloud solutions for translation, yes or no?" (https://www.iapti.org/articles/art34-using-cloud-s olutions-for-translation-yes-or-no.html). *IAPTI.org*. Retrieved 23 February 2017.

132. Smith, David Mitchell. "Hype Cycle for Cloud Computing, 2013" (https://www.gartner.com/doc/2573318/hy pe-cycle-cloud-computing-). Gartner. Retrieved 3 July 2014.

133. "The evolution of Cloud Computing" (http://www.hello-cirro.co.uk/evolution-of-cloud-computing/). Retrieved 22 April 2015.

134. "Microsoft Says to Spend 90% of R&D on Cloud Strategy" (http://cloudtimes.org/2011/04/12/microsoft-says-to-spend-90-of-rd-on-cloud-strategy/). Retrieved 22 April 2015.

135. "Roundup of Cloud Computing Forecasts And Market Estimates, 2014" (https://www.forbes.com/sites/louiscolumbus/2014/03/14/roundup-of-cloud-computing-forecasts-and-market-estimates-2014/). *Forbes*. Retrieved 2015-11-22.

136. "How the Virtual Data Room is boosting Mergers and Acquisitions" (https://www.pcworld.idg.com.au/article/614885/how-virtual-data-room-boosting-mergers-aquisitions//). PcWorld. Retrieved 2017-02-27.

20.14 FURTHER READING

- Millard, Christopher (2013). Cloud Computing Law (https://global.oup.com/academic/product/cloud-computing-law-9780199671687?cc=gb&lang=en&). Oxford University Press. ISBN 978-0-19-967168-7.

- Singh, Jatinder; Powles, Julia; Pasquier, Thomas; Bacon, Jean (July 2015). "Data Flow Management and Compliance in Cloud Computing". IEEE Cloud Computing. 2 (4): 24–32. doi:10.1109/MCC.2015.69 (https://doi.org/10.1109%2FMCC.2015.69).

- Armbrust, Michael; Stoica, Ion; Zaharia, Matei; Fox, Armando; Griffith, Rean; Joseph, Anthony D.; Katz, Randy; Konwinski, Andy; Lee, Gunho; Patterson, David; Rabkin, Ariel (1 April 2010). "A view of cloud computing". Communications of the ACM. 53 (4): 50. doi:10.1145/1721654.1721672 (https://doi.org/10.1145%2F1721654.1721672).

- Hu, Tung-Hui (2015). A Prehistory of the Cloud. MIT Press. ISBN 978-0-262-02951-3.

- Mell, P. (2011, September 31). The NIST Definition of Cloud Computing. Retrieved November 1, 2015, from National Institute of Standards and Technology website: http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

21 SEARCHCLOUDCOMPUTING.COM CLOUD COMPUTING

By Margaret Rouse - 21 Jul 2017

Cloud computing is a general term for the delivery of hosted services over the internet.

Cloud computing enables companies to consume a compute resource, such as a virtual machine (VM), storage or an application, as a utility -- just like electricity -- rather than having to build and maintain computing infrastructures in house.

21.1 CLOUD COMPUTING CHARACTERISTICS AND BENEFITS

Cloud computing boasts several attractive benefits for businesses and end users. Five of the main benefits of cloud computing are:

- **Self-service provisioning:** End users can spin up compute resources for almost any type of workload on demand. This eliminates the traditional need for IT administrators to provision and manage compute resources.
- **Elasticity:** Companies can scale up as computing needs increase and scale down again as demands decrease. This eliminates the need for massive investments in local infrastructure, which may or may not remain active.
- **Pay per use:** Compute resources are measured at a granular level, enabling users to pay only for the resources and workloads they use.
- **Workload resilience:** Cloud service providers often implement redundant resources to ensure resilient storage and to keep users' important workloads running -- often across multiple global regions.
- **Migration flexibility:** Organizations can move certain workloads to or from the cloud -- or to different cloud platforms -- as desired or automatically for better cost savings or to use new services as they emerge.

21.2 CLOUD COMPUTING DEPLOYMENT MODELS

Cloud computing services can be private, public or hybrid.

Private cloud services are delivered from a business's data center to internal users. This model offers the versatility and convenience of the cloud, while preserving the management, control and security common to local data centers. Internal users may or may not be billed for services through IT chargeback.

Common private cloud technologies and vendors include VMware and OpenStack.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

In the public cloud model, a third-party cloud service provider delivers the cloud service over the internet. Public cloud services are sold on demand, typically by the minute or hour, though long-term commitments are available for many services. Customers only pay for the CPU cycles, storage or bandwidth they consume.

Leading public cloud service providers include Amazon Web Services (AWS), Microsoft Azure, IBM and Google Cloud Platform.

A hybrid cloud is a combination of public cloud services and an on-premises private cloud, with orchestration and automation between the two. Companies can run mission-critical workloads or sensitive applications on the private cloud and use the public cloud to handle workload bursts or spikes in demand.

The goal of a hybrid cloud is to create a unified, automated, scalable environment that takes advantage of all that a public cloud infrastructure can provide, while still maintaining control over mission-critical data.

In addition, organizations are increasingly embracing a multicloud model, or the use of multiple infrastructure-as-a-service providers. This enables applications to migrate between different cloud providers or to even operate concurrently across two or more cloud providers.

Organizations adopt multicloud for various reasons. For example, they could do so to minimize the risk of a cloud service outage or to take advantage of more competitive pricing from a particular provider.

Multicloud implementation and application development can be a challenge because of the differences between cloud providers' services and application program interfaces (APIs). Multicloud deployments should become easier, however, as providers' services and APIs converge and become more homogeneous through industry initiatives such as the Open Cloud Computing Interface.

21.3 TYPES OF CLOUD COMPUTING SERVICES

Although cloud computing has changed over time, it has been divided into three broad service categories: infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS).

IaaS providers, such as AWS, supply a virtual server instance and storage, as well as APIs that enable users to migrate workloads to a VM. Users have an allocated storage capacity and can start, stop, access and configure the VM and storage as desired. IaaS providers offer small, medium, large, extra-large and memory- or compute-optimized instances, in addition to customized instances, for various workload needs.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.

In the PaaS model, cloud providers host development tools on their infrastructures. Users access these tools over the internet using APIs, web portals or gateway software. PaaS is used for general software development, and many PaaS providers host the software after it's developed. Common PaaS providers include Salesforce's Force.com, AWS Elastic Beanstalk and Google App Engine.

SaaS is a distribution model that delivers software applications over the internet; these applications are often called web services. Users can access SaaS applications and services from any location using a computer or mobile device that has internet access. One common example of a SaaS application is Microsoft Office 365 for productivity and email services.

21.4 EMERGING CLOUD TECHNOLOGIES AND SERVICES

Cloud providers are competitive, and they constantly expand their services to differentiate themselves. This has led public IaaS providers to offer far more than common compute and storage instances.

For example, serverless, or event-driven computing is a cloud service that executes specific functions, such as image processing and database updates. Traditional cloud deployments require users to establish a compute instance and load code into that instance. Then, the user decides how long to run -- and pay for -- that instance.

With serverless computing, developers simply create code, and the cloud provider loads and executes that code in response to real-world events, so users don't have to worry about the server or instance aspect of the cloud deployment. Users only pay for the number of transactions that the function executes. AWS Lambda, Google Cloud Functions and Azure Functions are examples of serverless computing services.

Public cloud computing also lends itself well to big data processing, which demands enormous compute resources for relatively short durations. Cloud providers have responded with big data services, including Google BigQuery for large-scale data warehousing and Microsoft Azure Data Lake Analytics for processing huge data sets.

Another crop of emerging cloud technologies and services relates to artificial intelligence (AI) and machine learning. These technologies build machine understanding, enable systems to mimic human understanding and respond to changes in data to benefit the business. Amazon Machine Learning, Amazon Lex, Amazon Polly, Google Cloud Machine Learning Engine and Google Cloud Speech API are examples of these services.

21.5 CLOUD COMPUTING SECURITY

Security remains a primary concern for businesses contemplating cloud adoption -- especially public cloud adoption. Public cloud service providers share their underlying hardware

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.

infrastructure between numerous customers, as public cloud is a multi-tenant environment. This environment demands copious isolation between logical compute resources. At the same time, access to public cloud storage and compute resources is guarded by account login credentials.

Many organizations bound by complex regulatory obligations and governance standards are still hesitant to place data or workloads in the public cloud for fear of outages, loss or theft. However, this resistance is fading, as logical isolation has proven reliable, and the addition of data encryption and various identity and access management tools has improved security within the public cloud.

21.6 A BRIEF HISTORY OF CLOUD COMPUTING

Cloud computing traces its origins back to the 1960s, when the computer industry recognized the potential benefits of delivering computing as a service or a utility. However, early computing lacked the connectivity and bandwidth needed to implement computing as a utility. It wasn't until the broad availability of internet bandwidth in the late 1990s that computing as a service became practical.

In the late 1990s, Salesforce offered one of the first commercially successful implementations of enterprise SaaS. This was followed closely by the arrival of AWS in 2002, offering a range of services, including storage and computation -- and now embracing databases, machine learning and other services. Today, Microsoft Azure, Google Cloud Platform and other providers have joined AWS in providing cloud-based services to individuals, small businesses and global enterprises.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

22 WHAT IS CLOUD COMPUTING?

The 'cloud' is a real buzzword, but what is it, how does it impact what you do, and is it anything really new?

- 
- *By Eric Griffith*
- *May 3, 2016 12:01AM EST*

What is the cloud? Where is the cloud? Are we in the cloud now? These are all questions you've probably heard or even asked yourself. The term "cloud computing" is everywhere.

In the simplest terms, cloud computing means storing and accessing data and programs over the Internet instead of your computer's hard drive. The cloud is just a metaphor for the Internet. It goes back to the days of flowcharts and presentations that would represent the gigantic server-farm infrastructure of the Internet as nothing but a puffy, white cumulus cloud, accepting connections and doling out information as it floats.

What cloud computing is not about is your hard drive. When you store data on or run programs from the hard drive, that's called local storage and computing. Everything you need is physically close to you, which means accessing your data is fast and easy, for that one computer, or others on the local network. Working off your hard drive is how the computer industry functioned for decades; some would argue it's still superior to cloud computing, for reasons I'll explain shortly.

The cloud is also not about having a dedicated network attached storage (NAS) hardware or server in residence. Storing data on a home or office network does not count as utilizing the cloud. (However, some NAS will let you remotely access things over the Internet, and there's at least one brand from Western Digital named "My Cloud," just to keep things confusing.)

For it to be considered "cloud computing," you need to access your data or your programs over the Internet, or at the very least, have that data synced with other information over the Web. In a big business, you may know all there is to know about what's on the other side of the connection; as an individual user, you may never have any idea what kind of massive data processing is happening on the other end. The end result is the same: with an online connection, cloud computing can be done anywhere, anytime.

22.1 CONSUMER VS. BUSINESS

Let's be clear here. We're talking about cloud computing as it impacts individual consumers—those of us who sit back at home or in small-to-medium offices and use the Internet on a regular basis.

There is an entirely different "cloud" when it comes to business. Some businesses choose to implement Software-as-a-Service (SaaS), where the business subscribes to an application it accesses over the Internet. (Think Salesforce.com.) There's also Platform-as-a-Service (PaaS), where a business can create its own custom applications for use by all in the company. And don't forget the mighty Infrastructure-as-a-Service (IaaS), where players like Amazon, Microsoft, Google, and Rackspace provide a backbone that can be "rented out" by other companies. (For example, Netflix provides services to you because it's a customer of the cloud services at Amazon.)

Of course, cloud computing is big business: The market generated $100 billion a year in 2012, which could be $127 billion by 2017 and $500 billion by 2020.

22.2 COMMON CLOUD EXAMPLES

The lines between local computing and cloud computing sometimes get very, very blurry. That's because the cloud is part of almost everything on our computers these days. You can easily have a local piece of software (for instance, Microsoft Office 365) that utilizes a form of cloud computing for storage (Microsoft OneDrive).

That said, Microsoft also offers a set of Web-based apps, Office Online, that are Internet-only versions of Word, Excel, PowerPoint, and OneNote accessed via your Web browser without installing anything. That makes them a version of cloud computing (Web-based=cloud).

Some other major examples of cloud computing you're probably using:

Google Drive: This is a pure cloud computing service, with all the storage found online so it can work with the cloud apps: Google Docs, Google Sheets, and Google Slides. Drive is also available on more than just desktop computers; you can use it on tablets like the iPad or on smartphones, and there are separate apps for Docs and Sheets, as well. In fact, most of Google's services could be considered cloud computing: Gmail, Google Calendar, Google Maps, and so on.

Apple iCloud: Apple's cloud service is primarily used for online storage, backup, and synchronization of your mail, contacts, calendar, and more. All the data you need is available to you on your iOS, Mac OS, or Windows device (Windows users have to install the iCloud control panel). Naturally, Apple won't be outdone by rivals: it offers cloud-based versions of its word processor (Pages), spreadsheet (Numbers), and presentations (Keynote) for use by any iCloud subscriber. iCloud is also the place iPhone users go to utilize the Find My iPhone feature that's all important when the handset goes missing.

Amazon Cloud Drive: Storage at the big retailer is mainly for music, preferably MP3s that you purchase from Amazon, and images—if you have Amazon Prime, you get unlimited image storage. Amazon Cloud Drive also holds anything you buy for the Kindle. It's essentially storage for anything digital you'd buy from Amazon, baked into all its products and services.

Hybrid services like Box, Dropbox, and SugarSync all say they work in the cloud because they store a synced version of your files online, but they also sync those files with local storage. Synchronization is a cornerstone of the cloud computing experience, even if you do access the file locally.

Likewise, it's considered cloud computing if you have a community of people with separate devices that need the same data synced, be it for work collaboration projects or just to keep the family in sync. For more, check out the The Best Cloud Storage and File-Syncing Services for 2016.

22.3 Cloud Hardware

Right now, the primary example of a device that is completely cloud-centric is the Chromebook. These are laptops that have just enough local storage and power to run the Chrome OS, which essentially turns the Google ChromeWeb browser into an operating system. With a Chromebook, most everything you do is online: apps, media, and storage are all in the cloud.

Or you can try a ChromeBit, a smaller-than-a-candy-bar drive that turns any display with an HDMI port into a usable computer running Chrome OS.

Of course, you may be wondering what happens if you're somewhere without a connection and you need to access your data. This is currently one of the biggest complaints about Chrome OS, although its offline functionality (that is, non-cloud abilities) are expanding.

The Chromebook isn't the first product to try this approach. So-called "dumb terminals" that lack local storage and connect to a local server or mainframe go back decades. The first Internet-only product attempts included the old NIC (New Internet Computer), the Netpliance iOpener, and the disastrous 3Com Ergo Audrey(pictured). You could argue they all debuted well before their time—dial-up speeds of the 1990s had training wheels compared to the accelerated broadband Internet connections of today. That's why many would argue that cloud computing works at all: the connection to the Internet is as fast as the connection to the hard drive. (At least it is for some of us.)

22.4 Arguments Against the Cloud

In a 2013 edition of his feature What if?, xkcd-cartoonist (and former NASA roboticist) Randall Monroe tried to answer the question of "When—if ever—will the bandwidth of the Internet surpass that of FedEx?" The question was posed because no matter how great your broadband connection, it's still cheaper to send a package of hundreds of gigabytes of data via Fedex's "sneakernet" of planes and trucks than it is to try and send it over the Internet. (The answer, Monroe concluded, is the year 2040.)

Cory Doctorow over at boingboing took Monroe's answer as "an implicit critique of cloud computing." To him, the speed and cost of local storage easily outstrips using a wide-area network connection controlled by a telecom company (your ISP).

That's the rub. The ISPs, telcos, and media companies control your access. Putting all your faith in the cloud means you're also putting all your faith in continued, unfettered access. You might get this level of access, but it'll cost you. And it will continue to cost more and more as companies find ways to make you pay by doing things like metering your service: the more bandwidth you use, the more it costs.

Maybe you trust those corporations. That's fine, but there are plenty of other arguments against going into the cloud whole hog. Apple co-founder Steve Wozniak decried cloud computing in 2012, saying: "I think it's going to be horrendous. I think there are going to be a lot of horrible problems in the next five years."

In part, that comes from the potential for crashes. When there are problems at a company like Amazon, which provides cloud storage services to big name companies like Netflix and Pinterest, it can take out all those services (as happened in the summer of 2012). In 2014, outages afflicted Dropbox, Gmail, Basecamp, Adobe, Evernote, iCloud, and Microsoft; in 2015 the outages hit Apple, Verizon, Microsoft, AOL, Level 3, and Google. Microsoft had another this year. The problems typically last for just hours.

Wozniak was concerned more about the intellectual property issues. Who owns the data you store online? Is it you or the company storing it? Consider how many times there's been widespread controversy over the changing terms of service for companies like Facebook and Instagram—which are definitely cloud services—regarding what they get to do with your photos. There's also a difference between data you upload, and data you create in the cloud itself—a provider could have a strong claim on the latter. Ownership is a relevant factor to be concerned about.

After all, there's no central body governing use of the cloud for storage and services. The Institute of Electrical and Electronics Engineers (IEEE) is trying. It created an IEEE Cloud Computing Initiative in 2011 to establish standards for use, especially for the business sector. The Supreme Court ruling against Aereo could have told us a lot about copyright of files in the cloud... but the court side-stepped the issue to keep cloud computing status quo.

Cloud computing—like so much about the Internet—is a little bit like the Wild West, where the rules are made up as you go, and you hope for the best.

For more, check out our roundups of the Business Choice Awards for Cloud Computing Services and the Cloud Storage area of the PCMag Business Software Index.

23 WHAT IS CLOUD COMPUTING? – ACCORDING TO MICROSOFT

23.1 A BEGINNER'S GUIDE

Simply put, cloud computing is the delivery of computing services—servers, storage, databases, networking, software, analytics, and more—over the Internet ("the cloud"). Companies offering these computing services are called cloud providers and typically charge for cloud computing services based on usage, similar to how you're billed for water or electricity at home.
Still foggy on how cloud computing works and what it's for? This beginner's guide is designed to demystify basic cloud computing jargon and concepts and quickly bring you up to speed.

23.2 USES OF CLOUD COMPUTING

You're probably using cloud computing right now, even if you don't realize it. If you use an online service to send email, edit documents, watch movies or TV, listen to music, play games, or store pictures and other files, it's likely that cloud computing is making it all possible behind the scenes. The first cloud computing services are barely a decade old, but already a variety of organizations—from tiny startups to global corporations, government agencies to non-profits—are embracing the technology for all sorts of reasons. Here are a few of the things you can do with the cloud:

- Create new apps and services
- Store, back up, and recover data
- Host websites and blogs
- Stream audio and video
- Deliver software on demand
- Analyze data for patterns and make predictions

23.3 TOP BENEFITS OF CLOUD COMPUTING

Cloud computing is a big shift from the traditional way businesses think about IT resources. What is it about cloud computing? Why is cloud computing so popular? Here are 6 common reasons organizations are turning to cloud computing services:

23.3.1 Cost

Cloud computing eliminates the capital expense of buying hardware and software and setting up and running on-site datacenters—the racks of servers, the round-the-clock electricity for power and cooling, the IT experts for managing the infrastructure. It adds up fast.

23.3.2 Speed

Most cloud computing services are provided self service and on demand, so even vast amounts of computing resources can be provisioned in minutes, typically with just a few mouse clicks, giving businesses a lot of flexibility and taking the pressure off capacity planning.

23.3.3 Global scale

The benefits of cloud computing services include the ability to scale elastically. In cloud speak, that means delivering the right amount of IT resources—for example, more or less computing power, storage, bandwidth—right when its needed, and from the right geographic location.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

23.3.4 Productivity

On-site datacenters typically require a lot of “racking and stacking”—hardware set up, software patching, and other time-consuming IT management chores. Cloud computing removes the need for many of these tasks, so IT teams can spend time on achieving more important business goals.

23.3.5 Performance

The biggest cloud computing services run on a worldwide network of secure datacenters, which are regularly upgraded to the latest generation of fast and efficient computing hardware. This offers several benefits over a single corporate datacenter, including reduced network latency for applications and greater economies of scale.

23.4.6 Reliability

Cloud computing makes data backup, disaster recovery, and business continuity easier and less expensive, because data can be mirrored at multiple redundant sites on the cloud provider’s network.

23.5 Types of cloud services: IaaS, PaaS, SaaS

Most cloud computing services fall into three broad categories: infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (Saas). These are sometimes called the cloud computing stack, because they build on top of one another. Knowing what they are and how they’re different makes it easier to accomplish your business goals.

23.5.1 Infrastructure-as-a-service (IaaS)

The most basic category of cloud computing services. With IaaS, you rent IT infrastructure—servers and virtual machines (VMs), storage, networks, operating systems—from a cloud provider on a pay-as-you-go basis.

23.5.2 Platform as a service (PaaS)

Platform-as-a-service (PaaS) refers to cloud computing services that supply an on-demand environment for developing, testing, delivering, and managing software applications. PaaS is designed to make it easier for developers to quickly create web or mobile apps, without worrying about setting up or managing the underlying infrastructure of servers, storage, network, and databases needed for development.

23.5.3 Software as a service (SaaS)

Software-as-a-service (SaaS) is a method for delivering software applications over the Internet, on demand and typically on a subscription basis. With SaaS, cloud providers host and manage the software application and underlying infrastructure, and handle any maintenance, like software upgrades and security patching. Users connect to the application over the Internet, usually with a web browser on their phone, tablet, or PC. To learn more,

23.6 Types of cloud deployments: public, private, hybrid

Not all clouds are the same. There are three different ways to deploy cloud computing resources: public cloud, private cloud, and hybrid cloud.

23.6.1 Public cloud

Public clouds are owned and operated by a third-party cloud service provider, which deliver their computing resources like servers and storage over the Internet. Microsoft Azure is an example of a public cloud. With a public cloud, all hardware, software, and other supporting infrastructure is owned and managed by the cloud provider.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.

You access these services and manage your account using a web browser.

23.6.2 Private cloud

A private cloud refers to cloud computing resources used exclusively by a single business or organization. A private cloud can be physically located on the company's on-site datacenter. Some companies also pay third-party service providers to host their private cloud. A private cloud is one in which the services and infrastructure are maintained on a private network.

23.6.3 Hybrid cloud

Hybrid clouds combine public and private clouds, bound together by technology that allows data and applications to be shared between them. By allowing data and applications to move between private and public clouds, hybrid cloud gives businesses greater flexibility and more deployment options.

23.7 HOW CLOUD COMPUTING WORKS

Cloud computing services all work a little differently, depending on the provider. But many provide a friendly, browser-based dashboard that makes it easier for IT professionals and developers to order resources and manage their accounts. Some cloud computing services are also designed to work with REST APIs and a command-line interface (CLI), giving developers multiple options.

23.8 MICROSOFT AND CLOUD COMPUTING

Microsoft is a leading global provider of cloud computing services for businesses of all sizes. To learn more about our cloud platform, Microsoft Azure, and how it compares to other cloud providers, see What is Azure? and Azure vs. AWS.

23.9 AZURE IS PRODUCTIVE FOR DEVELOPERS

Get your apps to market faster. Azure integrated tools, from mobile DevOps to serverless computing support your productivity. Build the way you want to, using the tools and open source technologies you already know. Azure supports a range of operating systems, programming languages, frameworks, databases, and devices.
Continuously innovate and deliver high-quality apps.
Provide cross-device experiences with support for all major mobile platforms.
Run any stack, Linux-based or Windows-based, and use advanced capabilities such as Kubernetes cluster in Azure Container Service.

23.10 AZURE IS THE ONLY CONSISTENT HYBRID CLOUD

Build and deploy wherever you want with Azure, the only consistent hybrid cloud on the market. Connect data and apps in the cloud and on-premises—for maximum portability and value from your existing investments. Azure offers hybrid consistency in application development, management and security, identity management, and across the data platform.

- Extend Azure on-premises and build innovative, hybrid apps with Azure Stack.
- Connect on-premises data and apps to overcome complexity and optimize your existing assets.
- Distribute and analyze data seamlessly across cloud and on-premises.
- Learn more about our consistent hybrid cloud

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

23.11 AZURE IS THE CLOUD FOR BUILDING INTELLIGENT APPS

Use Azure to create data-driven, intelligent apps. From image recognition to bot services, take advantage of Azure data services and artificial intelligence to create new experiences—that scale—and support deep learning, HPC simulations, and real-time analytics on any shape and size of data.

- Develop breakthrough apps with built-in AI.
- Build and deploy custom AI models at scale, on any data.
- Combine the best of Microsoft and open source data and AI innovations.

23.12 AZURE IS THE CLOUD YOU CAN TRUST

Ninety percent of Fortune 500 companies trust the Microsoft Cloud. Join them. Take advantage of Microsoft security, privacy, transparency, and the most compliance coverage of any cloud provider.

- Achieve global scale on a worldwide network of Microsoft-managed datacenters across 42 announced regions.
- Detect and mitigate threats with a central view of all your Azure resources through Azure Security Center.
- Rely on the cloud with the most comprehensive compliance coverage (50 compliance offerings), and recognized as the most trusted cloud for U.S. government institutions.

23.13 AZURE VS. AWS – OR MIRCOSOFT VERSUS AMAZON WEB SERVICES

Why Azure is the right choice

Organizations all over the world recognize Azure over AWS as the most trusted cloud, because it offers:

More regions than any other cloud provider

Unmatched hybrid capabilities

The strongest intelligence

23.14 TRUST THE CLOUD THAT HELPS PROTECT YOUR WORK

When you compare AWS versus Azure, you'll find that Azure has more comprehensive compliance coverage with more than 60 compliance offerings, and was the first major cloud provider to contractually commit to the requirements of the General Data Protection Regulation (GDPR). To protect your organization, Azure embeds security, privacy, and compliance into its development methodology, and has been recognized as the most trusted cloud for U.S. government institutions, earning a FedRAMP High authorization that covers 18 Azure services. In addition, Azure IP Advantage provides best-in-industry intellectual property protection, so you can focus on innovation, instead of worrying about baseless lawsuits.

> *"It felt like companies that needed to be compliant were being punished for it by having to pay more. Now, with Microsoft Cloud's scalable and pay-as-you-go model, we can be competitive while offering world-class service, flexibility, and security."*
>
> *Ed Anderson, Managing Director, Dyrand Systems*

23.15 INNOVATE WITH UNMATCHED INTELLIGENCE

Build intelligent solutions at scale using cognitive APIs, bots, machine learning, and blockchain as a service (BaaS)capabilities that you'll only find with Azure. By pairing these capabilities with powerful GPU-based compute, you'll accelerate deep learning, enable high-performance computing simulations, and conduct real-time data analytics using NVIDIA GPUs in Azure.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

"Thousands of partners sign in to our platform every hour. The response time from the Face API is incredible, enabling us to verify our drivers without slowing them down."

Dina Kovalev, Product Manager, Uber

23.16 EXPAND GLOBALLY WITH THE MOST REGIONS

Achieve global scale with 42 announced Azure regions—more than any other cloud provider. Our priority on geographic expansion means you can choose the datacenter and region that's right for you and your customers, with the performance and support you need, where you need it.

"All the traffic that we had—and it was more than we ever had before—went through the new Azure architecture and it was remarkably trouble-free."

Bob Strudwick, Chief Technology Officer, ASOS

23.17 GET IT FLEXIBILITY WITH A TRULY CONSISTENT HYBRID CLOUD

Optimize your existing assets by taking a hybrid approach to the cloud. Azure offers hybrid consistency everywhere—in application development, management and security, identity management, and across the data platform. This helps reduce the risk and cost of a hybrid cloud environment by enabling a common set of skills and offering portability of applications and workloads. Plus, save up to 40 percent when migrating Windows Server virtual machines to Azure using the Azure Hybrid Benefit.

"With the hybrid environment, we have the flexibility to place our applications and data on the appropriate cloud, based on our business, technology, and regulatory needs."

Tan Kok Meng, Chief Information Officer, Malaysia Airlines

23.18 BUILD ON THE LEADING CLOUD PLATFORM

Rely on the only cloud provider recognized in the industry as having leading solutions in infrastructure as a service (IaaS), software as a service (SaaS), and platform as a service (PaaS)—in fact, according to this Forrester Total Economic Impact study, you'll be more productive and increase your ROI with Azure PaaS services. Azure also offers the broadest support for container platforms, supporting DC/OS, Docker Swarm, and Kubernetes.

"In my opinion, Microsoft has a far stronger PaaS offering than Amazon does, and PaaS was always our ultimate destination."

Brian McFadden, Vice President and General Manager of Technology Services, MedPoint Digital

23.19 DEVELOP AND DEPLOY YOUR OPEN SOURCE TECHNOLOGY IN THE CLOUD

Develop and build the way you want in Azure, with your choice of tools, applications, and frameworks. As a leading open source contributor on GitHub, Microsoft actively supports the open source community and has a unique partnership with Red Hat to offer coordinated, multi-lingual support.

"Microsoft support for open source is very helpful. It means we can use whatever technologies make sense for us and have them fit in seamlessly with the rest of our stack."

Hanna Landrus, Data Scientist, BitTitan

23.20 DIG INTO BIG DATA WITH A COMPLETE BUSINESS ANALYTICS SOLUTION

Bring limitless elastic scale to your applications with Azure Cosmos DB. Then, turn your data into a competitive advantage using business analytics solutions, such as demand forecasting and inventory optimization. Rapidly build, customize, and deploy best practices with solution templates and Cortana Intelligence.

> *"Microsoft Cortana Intelligence capabilities are helping us filter the signal from the noise across large data sets, so we can focus on finding the real value in the data. Our vision of future digital capability will need to aggregate many sources of data and provide a platform for collaboration with customers."*
>
> Nick Farrant, Senior Vice President, Rolls-Royce

23.21 EASILY IMPLEMENT READY-TO-USE IOT

Quickly start with the most common Internet of Things (IoT) scenarios, such as remote monitoring and predictive maintenance, using preconfigured solutions in Azure IoT Suite. It's open and customizable by design, and 46 percent of Azure Certified for IoT devices run on Linux, Android, or other open source technologies.

> *We compared different cloud platforms and how they could support devices, connectivity, infrastructure services, IoT, and so on. We found that Microsoft Azure is the best cloud for this solution based on its flexibility. Azure also has a huge portfolio of very rich services that we can use, especially for IoT."*
>
> Rene Grohman, Azure Platform Architect, T-Systems

23.22 TAKE YOUR AZURE INVESTMENTS FURTHER WITH THE MICROSOFT CLOUD

Get even more value out of your investment in Azure by using one of our industry-leading SaaS services, including Office 365, Dynamics 365, and Enterprise Mobility + Security. Azure shares many foundational capabilities with these services, such as identity management through Azure Active Directory and mobile device management through Intune.

> *"Digital transformation is not an end, it's a means for continual improvement. We use Azure cloud analytics to gather operational and customer insights, and we use Office 365 to share that information to drive business value. The future for Hershey is very exciting—with Microsoft cloud solutions, we are in control of our digital journey, connecting everyone to work better together and share the goodness."*
>
> Carlos Amesquita, Chief Information Officer, The Hershey Company

24 APPENDIX A – ANALYSIS REPORT FOR EXASCALE STORAGE REQUIREMENTS FOR SCIENTIFIC DATA

Analysis Report for Exascale Storage Requirements for Scientific Data

1 INTRODUCTION

This report examines the breath of requirements for Exascale Data Storage of Scientific Data at the Department of Energy. The initial set of requirements is intended to define the overall scope of exascale data storage systems that will meet or exceed the needs of DoE researchers over the next 10 years. In an effort to contain the overall scope the requirements are referred to as either "must have" or "desirable". Of the fourteen proposed requirements, ten are "must have" and four are "desirable".

Section 2 of this report presents the fourteen requirements as initially stated. Section 3 defines four functional categories, each of which contains one or more of the fourteen requirements and expands on each of the four functional categories and its constituent requirements in order to outline the areas of interest that will be more fully researched in a subsequent report. Finally, Section 4 summarizes the four functional categories and areas of interest for the subsequent report.

This report is intended to be a foundation on which to build a more complete and comprehensive set of requirements that meet the needs of the DoE labs that need exascale data storage capabilities.

2 INITIAL SET OF REQUIREMENTS

The initial set of requirements proposed are as follows:

1. **Must** have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. **Must** be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. **Must** use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. **Must** be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. **Must** be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. **Must** integrate into network and data lifecycle management software.
7. **Must** provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. **Must** have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. **Must** include a standard interface to exchange data between existing and future storage solutions
10. **Must** include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. ***Desirable***: *While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. ***Desirable***: *POSIX or Relaxed POSIX compatibility*
13. ***Desirable***: *Space management capabilities (e.g. Quotas)*
14. ***Desirable***: *Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

To reiterate, these fourteen requirements are broadly divided into "must have" and "desirable" in order to limit the overall scope of this and the subsequent report.

3 FUNCTIONAL CATEGORIES

The functional categories (and *estimated time to research*) are defined as:

1. Scalability (*2 weeks*)
2. Reliability, Availability, and Serviceability (RAS) (*1.5 weeks*)
3. Growth and Evolution (*1 week*)
4. Usability (*1.5 weeks*)

3.1 SCALABILITY

With respect to data storage systems the term "scalability" can mean several things. In the context of this report, "scalability" will describe how a data storage system increases in each of several fundamental dimensions:

1. Capacity – Refers to increasing the total amount of data storage space managed by a data storage system or system(s). Capacity scaling is achieved through increases in the capacity of individual storage media (e.g. disk drives, tape drives, …etc.) and/or the ability to attach additional storage units, storage subsystems, storage systems, …etc. to a particular point of control (e.g. a storage controller or file system).
2. Performance – There are several aspects of data storage system performance that need to be considered:
    a. Metric – For storage systems there are basically three performance "metrics"
        i. Access operations per second – typically small operations
        ii. Bytes per second commonly referred to as bandwidth or data rate
        iii. Time to first byte – this is more commonly used in tape-based archival systems
    b. Raw performance of each "component" of a storage system (e.g. disk drives, disk array controller, interconnects, …etc.) for any particular metric
    c. Perspective
        i. Performance of the hardware subsystems (e.g. an individual disk drive, the array controller, …etc.) – the rate of data access requests coming in, rate of data access requests processed, data transfer rate, …etc.
        ii. Performance of the overall data storage system – the aggregate rate of [higher level] data access requests processed by the data storage system, aggregate data transfer rate, …etc.
        iii. Performance realized by an individual user or application – rate of data requests processed, aggregate data transfer rate, file access rate, metadata performance, …etc.
    d. Efficiency – Percentage of the aggregate performance of the individual data storage system components versus the observed or delivered performance of the overall system. For example, a data storage system with a single 20Gbits/sec InfiniBand host interface in front of an array controller managing 1,000 disk drives, each capable of a sustained data transfer rate of 100MBytes/sec, has an aggregate data rate of 1,000x100MBytes/sec = 100,000MBytes/sec or 100GBytes/sec. However, in practice, the "delivered" data rate is limited by the host interface to 20Gbits/sec or 2GBytes/sec. This results in a performance efficiency of 2%.
3. Interconnects – The communication hardware, software, and protocol(s) used to move data ***between*** storage devices and control systems. Examples include Fibre Channel, SAS, SATA, InfiniBand, Ethernet …etc. For the purposes of this report, "interconnects" are defined as communications ***within*** a storage system. For communications outside the storage system see "access".
4. Access – The communication hardware, software, and protocol(s) outside the storage system that facilitates access to data inside a storage system. Examples include Fibre Channel, Ethernet, InfiniBand, …etc. For the purposes of this report, "access" is defined as communications outside the storage system.
5. Metadata – the data that describes the data being storage and accessed. Metadata performance for one million files is currently reasonably well understood and acceptable. However, when the number of "files" increases to billions or trillions as it can in an exascale system, the current methods used to manage metadata break down and significantly limit the scaling of metadata performance. For example, let us suppose that the metadata required to adequately describe a single file is 1024 bytes to be generous. This includes things like the name of the file, access rights, data tags, …etc. Therefore one trillion (10^{12}) such files requires one petabyte (10^{15}) of storage space just for the metadata. Hence the metadata performance is limited by the access mechanisms used to process and store metadata.
6. Cost – There are several aspects to cost scaling that need to be considered:
    a. Purchase cost – includes hardware, software, licensing, …etc.
    b. Operational cost – facility space, energy usage, staff, …etc.
    c. Maintenance cost – this is normally included with the Purchase Cost but needs to be listed as a separate item because it is not always necessary to purchase maintenance with the system.
    d. Total Cost of Ownership (TCO) – essentially the sum of a, b, and c. For the purposes of this report TCO will be used as an indicator of how "cost" scales.

It is important to note that for every factor of 10 an object is scaled (up or down), things that previously never caused any issues can become issues in the scaled object in one or more of each scaling dimension. For example, scaling a 100-disk storage system up to 1,000 disks increases space utilization, produces higher heat and power loads, and presents greater interconnect complexity… etc. just to name a few. It is easy to see that any one or more of these can exceed that required for the smaller storage system.

Finally, one aspect of storage system performance critical to the overall mission of an exascale data center is that of End-to-End Performance. This usually refers to the speed at which data can be moved from one storage system to another. In general scaling End-to-End performance to involves scaling the systems managing the data movement, the network infrastructure between these systems, and storage systems at both ends. The maximum End-to-End performance realized is limited by the slowest component in the overall data path.
With respect to this report, requirements 2 and 3 fit into the Scalability category.

3.2 Reliability, Availability, and Serviceability (RAS)

in an exascale data storage system is somewhat different than RAS in more traditional (smaller) storage systems. In an ideal data storage system, it is assumed that all components are operational all the time and component failures (e.g. disk drives, interface transceivers, …etc.) are treated as anomalies. The basic rule is stated as:

A system must be designed to operate at stated performance levels regardless of what is failing at any given time.

In an exascale data storage system it should be assumed that something is always in the process of failing. Furthermore, given that components are always in the process of failing, it is also necessary to have the ability to service these components without interruption to the system. This is a very complex problem that needs to be addressed at every level in a system including but not limited to:

- Servers (compute, I/O , …etc.)
- Storage Systems
- Networks
- Software, firmware, and hardware

Catastrophic failures or failures that negatively impact the availability of a significant portion of the system cannot be tolerated.

Finally, there is the growing issue of Silent Data Corruption: *Undetected* errors in computer data that occur during writing, reading, storage, transmission, or processing, that introduce *unnoticed changes to the original data*. Results range from a minor loss of data to a catastrophic system crash. In exascale data storage systems, the probability of silently corrupting data increases significantly because the overall amount of data stored grows faster than the mechanisms used to detect data errors. Therefore additional methods are required to detect *silently corrupted data* as the capacity of exascale data storage systems increase.

3.3 Growth and Evolution

Growth is defined to be how a data storage system grows over time in any one of several dimensions. For example, given a data storage system that has the ability to scale from 100TB to 1000TB, how, exactly, is the additional capacity added to the storage system over time? This can be accomplished through increases in data storage density or simply attaching and incorporating more storage devices.

Evolution refers to how the capabilities of a storage system change (improve) over time either through incremental updates or through the explicit addition of new features and capabilities. For example, Solid State Devices (aka SSDs) started out as devices that mimicked the behavior of traditional spinning disk drives in form and function (e.g. 2.5-inch disk form factor packaging and use of the SAS/SATA protocol). SSDs later evolved to allow them to be placed directly on PCI express thereby eliminating both the traditional spinning disk packaging, protocols, and associated infrastructure. The SAS/SATA protocol was replaced by the NVMe protocol that treats an SSD more as nonvolatile memory than as a spinning disk drive. The end result was far better performance with significantly less support infrastructure.

It is critical that new capabilities can be added to exascale data storage systems without interrupting the operation of the system. Again, capabilities include but are not limited to:

- Storage architecture
- Better, faster, more appropriate data storage protocols
- Metadata processing offload engines
- Self healing

3.4 Usability

Generally speaking, data representation and access methods have not changed significantly since the 1950s. Data is still organized in hierarchical sets of "directories" and "files" with limited metadata that represents the contents of the directories and files. At scales appropriate for exascale systems it becomes abundantly clear that new, fundamentally different data representation and access methods are needed to allow users and applications to more effectively utilize data in exascale data storage systems. For example, given a data set containing a billion files, how does a user or application locate one or more files within that dataset based on the contents of each file? In other words, it is one thing to find any one file, it is quite another to find one thing that may exist in any one or more of the billion files.

Thus, usability refers to how easy or difficult it is to search and access data in an exascale data storage system. There can be several metrics for usability but in the end, it comes down to “productivity” or how much “work” a user can get done in a given amount of time. To assist in data access scaling much work needs to be done with data tagging and rich metadata capabilities of both storage systems (file systems) and applications.

One example that comes to mind about the “usability” of a storage system is that of the human brain. It is truly astonishing how quickly a random piece of information can be retrieved even if it has not been retrieved for a very long time. The example here is mentioning the phrase “A three-hour tour…” In many people raised in the US during the 1960s, this will immediately be recognized as coming from the theme song of an American sitcom circa 1964-1967. Sorry for the ear worm.

4 SUMMARY

Areas to be addressed in subsequent report are the primary functional categories (and *estimated time to research*) are:

1. Scalability (*2 weeks*)
2. Reliability, Availability, and Serviceability (RAS) (*1.5 weeks*)
3. Growth and Evolution (*1 week*)
4. Usability (*1.5 weeks*)

Overall the report will cover topics of particular interest to the DoE and will be given priority. Topics include but not limited to:

8. Hierarchical Storage Management Systems such as HPSS
9. Storage media and associated hardware such as tapes, disks, and solid state storage
10. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
11. Metadata and data access at exascale
12. Information access at exascale
13. Storage system management including capacity management, hardware management, software management, upgrades, normal maintenance …etc.
14. Data management
15. Information management
16. Data security
17. How to deal with technology obsolescence, rolling upgrades, retirement of old equipment, …etc.

These topics will be discussed as they currently exist within the context of the four Functional Areas. This will be followed by informed speculation on what could be possible in 1, 3, 5, and 10 years out.

25 ORIGINAL FOURTEEN REQUIREMENTS

The original set of 14 requirements for Exascale Data Storage Systems were condensed to 7 more general requirements. The following pages describes how each of the 14 requirements was condensed into one of the 7 general requirements.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. Hierarchical Storage Management System(s)
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. Data security

1. **Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).**
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. **Metadata, Data, and Information Access at Exascale**
5. Storage System Management
6. Data and Information Management
7. Data security

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. **Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.**
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. **Storage media and associated hardware such as tapes, disks, and solid state storage**
3. **Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet**
4. **Metadata, Data, and Information Access at Exascale**
5. Storage System Management
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. **Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).**
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. **Storage media and associated hardware such as tapes, disks, and solid state storage**
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. Data security

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. **Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.**
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. **Storage media and associated hardware such as tapes, disks, and solid state storage**
3. **Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet**
4. Metadata, Data, and Information Access at Exascale
5. **Storage System Management**
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. **Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.**
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. **Storage media and associated hardware such as tapes, disks, and solid state storage**
3. **Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet**
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. Data security

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. **Must integrate into network and data lifecycle management software.**
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. **Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet**
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. **Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.**
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. **Storage System Management**
6. **Data and Information Management**
7. **Data security**

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. **Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).**
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. **Storage media and associated hardware such as tapes, disks, and solid state storage**
3. **Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet**
4. **Metadata, Data, and Information Access at Exascale**
5. Storage System Management
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. **Must include a standard interface to exchange data between existing and future storage solutions**
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. **Storage media and associated hardware such as tapes, disks, and solid state storage**
3. **Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet**
4. **Metadata, Data, and Information Access at Exascale**
5. Storage System Management
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. **Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).**
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. **Metadata, Data, and Information Access at Exascale**
5. Storage System Management
6. Data and Information Management
7. Data security

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. ***Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.***
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. Hierarchical Storage Management System(s)
2. **Storage media and associated hardware such as tapes, disks, and solid state storage**
3. **Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet**
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. ***Desirable: POSIX or Relaxed POSIX compatibility***
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. **Data and Information Management**
7. **Data security**

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. ***Desirable: Space management capabilities (e.g. Quotas)***
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. **Hierarchical Storage Management System(s)**
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. **Storage System Management**
6. **Data and Information Management**
7. Data security

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. ***Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).***

1. **Hierarchical Storage Management System(s)**
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. **Storage System Management**
6. **Data and Information Management**
7. Data security

1. **Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).**
2. **Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.**
3. **Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).**
4. **Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.**
5. **Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.**
6. **Must integrate into network and data lifecycle management software.**
7. **Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.**
8. **Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).**
9. **Must include a standard interface to exchange data between existing and future storage solutions**
10. **Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).**
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. ***Desirable: POSIX or Relaxed POSIX compatibility***
13. ***Desirable: Space management capabilities (e.g. Quotas)***
14. ***Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).***

1. **Hierarchical Storage Management System(s)**
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. **Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.**
3. **Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).**
4. **Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.**
5. **Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.**
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. **Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).**
9. **Must include a standard interface to exchange data between existing and future storage solutions**
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. ***Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.***
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. Hierarchical Storage Management System(s)
2. **Storage media and associated hardware such as tapes, disks, and solid state storage**
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. Data security

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. **Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.**
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. **Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.**
5. **Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.**
6. **Must integrate into network and data lifecycle management software.**
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. **Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).**
9. **Must include a standard interface to exchange data between existing and future storage solutions**
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. ***Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.***
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. Hierarchical Storage Management System(s)
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. **Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet**
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. **Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).**
2. **Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.**
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.
8. **Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).**
9. **Must include a standard interface to exchange data between existing and future storage solutions**
10. **Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).**
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. ***Desirable: POSIX or Relaxed POSIX compatibility***
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. *Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).*

1. Hierarchical Storage Management System(s)
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. **Metadata, Data, and Information Access at Exascale**
5. Storage System Management
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. **Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.**
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. **Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.**
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. *Desirable: POSIX or Relaxed POSIX compatibility*
13. ***Desirable: Space management capabilities (e.g. Quotas)***
14. ***Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).***

1. Hierarchical Storage Management System(s)
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. **Storage System Management**
6. Data and Information Management
7. Data security

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. **Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.**
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. ***Desirable: POSIX or Relaxed POSIX compatibility***
13. ***Desirable: Space management capabilities (e.g. Quotas)***
14. ***Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).***

1. Hierarchical Storage Management System(s)
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. **Data and Information Management**
7. Data security

1. Must have the ability to assist the user in categorizing and locating necessary data (e.g. data tagging).
2. Must be scalable with respect to size, performance, and cost. As data set sizes increase, capacity and performance of the underlying storage system must also increase.
3. Must use devices consistent with their storage characteristics (e.g., avoid random access on serial devices).
4. Must be able to incorporate current and developing storage technologies including but not limited to disk, tape, and solid state storage. Technology refresh at DoE sites occurs approximately every five years. Every year something is being upgraded and something is becoming stable. This study should include a decade roadmap.
5. Must be able to incorporate current and developing network technologies favoring non-proprietary and more commodity based solutions.
6. Must integrate into network and data lifecycle management software.
7. **Must provide storage manageability tools to reserve storage, migrate storage within a hierarchy.**
8. Must have the ability to identify and potentially correct data corruption, including device bit error and failure rates (e.g. may include RAIT).
9. Must include a standard interface to exchange data between existing and future storage solutions
10. Must include a flexible metadata framework to describe data (e.g. find data without downloading or maintaining external indices).
11. *Desirable: While power and cooling are out of scope, we would like to know what the new proposed solution(s) would require.*
12. ***Desirable: POSIX or Relaxed POSIX compatibility***
13. *Desirable: Space management capabilities (e.g. Quotas)*
14. ***Desirable: Ability to integrate with existing data transfer tools (Command Line Interface (CLI) and/or Graphical User Interface (GUI) APIs).***

1. Hierarchical Storage Management System(s)
2. Storage media and associated hardware such as tapes, disks, and solid state storage
3. Convergence of various storage related technologies such as the convergence of Fibre Channel and Ethernet
4. Metadata, Data, and Information Access at Exascale
5. Storage System Management
6. Data and Information Management
7. **Data security**

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.