



BNL-221188-2021-JAAM

Methods and Experiences for Developing Abstractions for Data-intensive, Scientific Applications

A. Luckow, S. Jha

To be published in "2020 IEEE 34TH INTERNATIONAL PARALLEL AND DISTRIBUTED PROCESSING
SYMPOSIUM WORKSHOPS (IPDPSW 2020)"

May 2020

Computational Science Initiative
Brookhaven National Laboratory

U.S. Department of Energy
USDOE Office of Science (SC), Advanced Scientific Computing Research (SC-21)

Notice: This manuscript has been authored by employees of Brookhaven Science Associates, LLC under Contract No. DE-SC0012704 with the U.S. Department of Energy. The publisher by accepting the manuscript for publication acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party's use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Methods and Experiences for Developing Abstractions for Data-intensive, Scientific Applications

Andre Luckow^{1,2, } and Shantenu Jha^{2,3}

¹*Ludwig-Maximilian University, Munich, Germany*

²*RADICAL, ECE, Rutgers University, Piscataway, NJ 08854, USA*

³*Brookhaven National Laboratory, Upton, NY, USA*

Abstract—Scientific applications need to integrate various programming, computational and data models, over diverse, evolving and heterogeneous infrastructure at scale. Pervasive and effective abstractions are thus critical; however, the process of developing abstractions for scientific applications and infrastructures is not well understood. While formal approaches for system development are suited for well-defined, closed environments, they are less effective in designing abstractions for scientific systems and applications. The design science research (DSR) method provides the basis for designing practical systems that can handle real-world complexities. In contrast to formal approaches, DSR emphasizes both practical relevance and knowledge creation by building and rigorously evaluating all artifacts. We show how DSR provides a well-defined framework for developing abstractions and middleware systems for distributed systems. Specifically, we address the critical problem of distributed resource management on heterogeneous infrastructure over a dynamic range of scales, a challenge that currently limits many scientific applications. We use the pilot-abstraction, a widely used resource management abstraction for high-performance, high throughput, big data, and streaming applications, as a case study for evaluating the DSR activities. For this purpose, we analyze the research process and artifacts produced during the design and evaluation of the pilot-abstraction. We find DSR provides a concise framework for iteratively designing and evaluating systems. Finally, we capture our experiences and formulate different lessons learned.

I. INTRODUCTION

New scientific applications and discoveries are enabled by advanced data and compute infrastructures, algorithms, and tools. Scientific progress increasingly depends on driving forward the ability to support large-scale computational and data demands of simulations in conjunction with data processing, analytics, and machine learning [1], [2]. The complexity of developing, deploying and scaling scientific applications arises from various sources, in particular, the increasing heterogeneity that exists at all levels, from hardware, infrastructure, middleware to software [3].

Abstractions are crucial for scalable systems that hide internal complexities and expose simple interfaces [4], [5]. Designing useful abstractions is challenging: hiding complexity does not automatically lead to simple interfaces. The possible design space for abstractions is typically vast, and there is no consensus on what constitutes effective abstractions. Further, there are no accepted recipes to design and develop abstractions for large-scale scientific distributed compute and data infrastructures.

The particular challenge addressed in this paper is the design of abstractions for resource management on distributed and heterogeneous compute and data infrastructure. Currently, the scale and uptake of scientific, data-intensive applications are hindered by a reliance on proprietary application and systems-level resource management systems. These are often implemented using rigid and ad-hoc approaches [6]. A generalized abstraction that helps overcome these limitations and enable scalable applications is needed.

While formal approaches maybe suitable for closed systems, they have limitations for designing open, scientific distributed systems. Iivari emphasizes that the “*theory-with-practical-implications research strategy has seriously failed to produce results that are of real interest in practice* [7]”. DSR is an iterative approach to building, evaluating, and refining software systems. While many research approaches solely focus on theory and knowledge, DSR emphasizes practical relevance. It realizes that complex systems need to be designed and evaluated in real-world settings. By introducing a rigorous evaluation of the produced artifacts, DSR provides generalizable knowledge that informs future design iterations, but can also be transferred to other problems.

We propose the application of the design science research method (DSR) [8] to the design of an abstraction and middleware for distributed resource management. Specifically, we apply the DSR method to the process of designing the pilot-abstraction [6]. Based on in-depth studies of different applications, we define the design objective. Using rigorous build/evaluate cycles, we design, evaluate, and evolve the abstraction from a compute-centric to an integrated abstraction for managing compute and data resources and applications. In this paper, we demonstrate the suitability of DSR for creating well-defined abstractions and implement these in a real-world system.

As part of DSR, we define different evaluation methods and criteria for assessing the abstraction. For example, we investigate the usability and versatility of the abstraction in several case studies, e. g., in ensemble-based simulations, MapReduce, and stream processing applications. We use conceptual modeling to provide and validate our understanding of the pilot-abstraction and the underlying mechanisms. Further, we study different implementations of the abstraction concerning the performance and scalability using different types of ap-

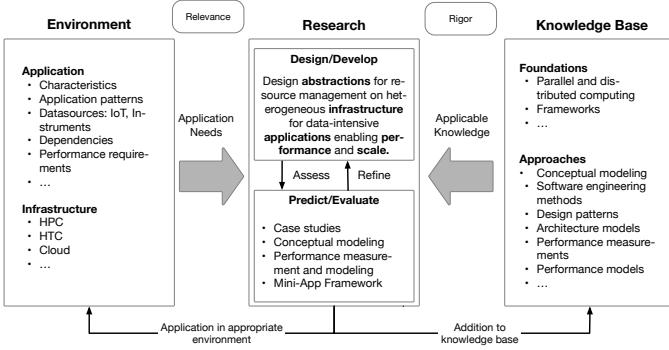


Fig. 1: Design Science Research Method (adapted from Hevner [8]): To address the complexity of the problem space, we follow an iterative research approach of continuously building and evaluating abstractions.

lications, e. g., from the domains of genome sequencing and light source sciences.

This paper makes the following contributions: (i) it uses the DSR framework to assemble a set of methods for the design and evaluation of abstractions, (ii) it demonstrates the validity of DSR for designing and evaluation abstractions, such as the pilot-abstraction, (iii) it surveys publications related to the pilot-abstraction and investigates the used methods for design and evaluation, and (iv) it synthesizes the experiences gathered during this process in a set of lessons learned. DSR was initially introduced in the domain of information system research; we believe this is its first application to scientific distributed computing.

This paper is structured as follows: We begin with an introduction of the methodology in Section II, and continue with an investigation of scientific applications and their characteristics in Section III. The result is five application scenarios that the abstraction needs to address. We present the pilot-abstraction in Section IV. In Section V, we discuss the methods used for evaluating the system. We discuss our learnings and experiences of applying DSR in Section VI.

II. METHODOLOGY

The objective of this section is to provide an introduction to *design science research (DSR)* [8]. DSR avoids the limitations of theory-based approaches, in particular, their inability to capture complex, real-world systems. It emphasizes the iterative creation, evaluation, and refinement of systems. The complexity of scientific applications and infrastructure make DSR suitable for designing abstractions that enable applications to scale across heterogeneous infrastructure. For this purpose, we customize DSR and apply it the first time to the problem of abstraction development (see Figure 1).

The build-evaluate-refine cycle has two primary inputs: The environment provides essential context for the problem, in particular, concerning application requirements, characteristics, and infrastructure. The knowledge base defines, in particular, the foundations and methodologies used the evaluation. In the following, we give an overview of the different design



Fig. 2: Application Analysis (adapted from [12]): Patterns emerge by observing application characteristics and implementations.

science research activities (adapted from Peffer [9]): (i) problem identification, (ii) definition of objectives, (iii) design and implementation, (iv) demonstration, and (v) evaluation.

A. Problem Identification and Objectives

Before starting the design process, an understanding of the problem and design objectives is essential. Common methods for this activity are literature reviews, expert interviews, focus groups, and surveys [10]. In scientific application development, requirements frequently only emerge during the creation of the systems leading [11]. Thus, iterative methods, such as DSR, are instrumental. The challenge addressed by this paper is the design and development of effective abstractions that provide the right level of detail and support a variety of application scenarios and infrastructure while retaining ease-of-use.

B. Design

1) Abstraction Design: Abstractions are a fundamental method of computer science, enabling reasoning about a problem at the right level while allowing the underlying system to implement a solution [4]. Shaw defines an abstraction as *“a simplified description or specification of a system that emphasizes some of the system’s details or properties while suppressing others [5].”*

To develop efficient abstractions, an understanding of applications and infrastructure usage modes is instrumental. An important foundation for the development of abstractions are patterns. Patterns are suitable solutions to recurring problems in a particular context that can be applied multiple times without doing it exactly the same ways [13], [12]. Patterns can be discovered by observing common problem decompositions (e. g., task and data partitioning), communication, and coordination structures in applications. Jha et al. utilize this process to study patterns and abstractions for distributed applications [12]. Mattson et al. [13] investigate patterns for parallel and distributed applications. Figure 2 illustrates the relationship between applications, patterns, abstractions, and systems. Discovered patterns serve as candidates for the development of abstractions and their implementation in a middleware system.

An abstraction represents the external interface of the system. Thus, careful design is essential. The desired properties of an abstraction are generality and simplicity [14]. Generality refers to the ability of the abstraction to be broadly used.

Simplicity is reflected in multiple properties, e. g., ease-of-use, maintainability, and extensibility [15].

The development of abstractions is a difficult task and requires the identification of essential concepts, properties, and relationships. Conceptual modeling enables abstract thinking and reasoning about a system and its abstraction. Conceptual models represent and describe systems, e. g., applications, systems, and infrastructure. They can be used to formulate concepts about the system, explaining how a system works. Mylopoulos defines a conceptual model as “*a description of an aspect of the physical and social world around us for the purposes of understanding and communication* [16].” Johnson defines a conceptual model as “a high-level description of how a system is organized and operates [17].” Conceptual models were introduced to computer science in 1984 by Brodie, Mylopoulos, and Schmidt to overcome the increasing specialization of computer science disciplines to describe high-level aspects and interactions better [18]. Conceptual models are used in different areas, e. g., for software architectures [19], and programming languages [5].

2) *System Design*: Software architecture and engineering research is the study of useful system organization [20], i. e., the design of the composition/decomposition of systems and subsystems and the communication between these. Common objectives of the system design are flexibility, maintainability, and re-usability [19]. Patterns are an important aspect of designing software systems. Initially, introduced in the domain of building architecture [21], were adopted to the domain of software architecture and engineering by Beck/Cunningham [22].

A fundamental principle of system design is modularization and decomposition. Modularization has many benefits, e. g., flexibility, comprehensibility, and maintainability [23]. Further, the development time can be reduced by the ability to distribute work across different groups. According to Parnas, the most critical criteria for organizing a system is information hiding [24], i. e., the ability to carefully control the information exposed by a component using well-defined external interfaces and hiding information that is likely to change.

A common type of modularization used by complex systems is a *layered architecture* [25]. The layered architecture model partitions the system in distinct hierarchical layers. Each layer encapsulates a defined set of functions and provides services to the layer above. This pattern is widely used in system-level software, such as databases, operating systems, middleware, and distributed system.

Similar to a layered architecture model is the hourglass model [14], which relies on a central bottleneck layer at the waist of the hourglass that connects a wide range of lower-level and higher-level services. Resource management is commonly described using hourglass models [26], [27].

C. Evaluation

A rigorous evaluation of all artifacts is a crucial part of the DSR process. Sonnenberg/Brocke [10] propose four evaluation activities. Eval 1 evaluates the problem statements, using

methods such as literature review and surveys. Eval 2 investigates the design specifications, e. g., using expert interviews, demonstrations, simulations, and benchmarking. Eval 3 is concerned with the instantiation of an artifact, e. g., a prototype, using methods, such as experimentation. Last, Eval 4 observes the artifact in the real world. We utilize different evaluation methods for different activities, in particular, case studies [28], performance characterization, and modeling [29]. In the following, we particularly focus on methods for performance and scalability evaluations, i. e., Eval 3 and 4.

1) *Performance Characterization*: Performance measurements and characterizations are common methods for describing a system in artificial and natural settings (Eval 3/4). Performance measurements can have different objectives: (i) workload and system characterization, (ii) performance improvements, and (iii) to evaluate design alternatives [30]. An essential component of a performance evaluation is the workload defined as the set of all inputs (programs and data) that a system receives from its environment [30]. A *benchmark* refers to a workload that is used to compare computer systems. A workload used in performance evaluations should be representative; in the best case, it should reflect an actual, real-world workload. *Benchmarking* refers to the process of comparing two or more computer systems [29]. A measure describes the performance of a system, e. g., the runtime or throughput of a system. More complex metrics can capture cost/price or quality/runtime trade-offs [31].

Scientific applications are complex, unique, and not well-represented by standard benchmarks. The chosen metrics often do not provide a comprehensive view of the system, and thus, are not a proxy for real-world performance [31]. Further, most benchmarks neglect application-level quality metrics and focus mostly on runtime and scaling performance. However, it is often challenging to obtain real-world performance data that provide useful insights. For example, for data-intensive applications, there are, e. g., complex infrastructure components, such as data source, broker, and processing applications, that need to be carefully controlled. To account for that, often simplified, synthetic workloads are used to study performance (e. g., the Mini-App framework [32]). Similar techniques are commonly used for generating reproducible data and compute workloads, see [33], [34], [35].

2) *Performance Models*: Performance models [36] are a way of abstracting performance-related insights into an analytical model. An analytical model is a precise formulation of a model using mathematical logic, entities, and relations to describe concepts [37]. Analytical models are white-box and can quantify the relationship between the different concepts. Statistical models, in contrast, derive insights and predictions from data [38]. The advantage of statistical models is that they do not require domain knowledge and can model highly complex domains. However, they are often black-box models, i. e., they are more difficult to interpret.

Many computer science domains use performance models, e. g., for programming languages, operating systems, database systems, system components (e. g., schedulers), as well as par-

allel and distributed systems. For example, database systems utilize cost-based optimizer to generate an optimal query execution plan [39]. A well-known performance model for distributed applications is Amdahl's Law [40].

III. APPLICATION SCENARIOS, CHARACTERISTICS AND REQUIREMENTS

Understanding the problem domain is an essential step in the design process. In this section, we discuss important application characteristics and requirements.

A. Application Characteristics

The requirements of scientific applications are growing more diverse and complex [41], [42]. An increasing number of instruments, such as light sources, telescopes, and genome sequence machines, generate vast volumes of data. Applications are becoming more sophisticated and increasingly require the combination of various processing types, e.g., simulation, analytics, and machine learning. These processing types impose different requirements on abstractions, middleware, and infrastructure.

While the integration of different processing types is challenging, it yields many benefits, e.g., it has been demonstrated that machine learning-based approximation techniques can improve simulations (e.g., by quickly identifying regions of interests). Another example is the guidance of experiments using machine learning, e.g., to find interesting events and regions, and to adapt sampling accordingly [41].

While data-intensive, scientific applications are highly diverse, they often share common computational and data characteristics. Early studies, e.g., the Berkeley Dwarfs [43], focused on the understanding of parallel algorithms based on their computation and data movement patterns. Jha et al. [12] study distributed applications. In D3 science [44], we conducted a survey consisting of 9 questions and a series of workshops to understand the distributed and dynamic data aspects of 13 scientific applications. The Big Data Ogres [45], [46] introduce a multi-dimensional framework, so-called facets, which represent key characteristics of big data applications and use them to define a set of Mini Apps based on a study of more than 50 use cases collected by NIST [47].

Based on an investigation of 50+ applications and their characteristics in [44], [45], [52], we derived five applications scenarios: task-parallel, data-parallel, dataflow, iterative, and streaming (see Table I). In the following, we discuss important characteristics and patterns found in these application scenarios.

An important characteristic is the decomposition pattern: Task-level parallelism describes the execution of diverse compute tasks on multiple compute resources. In contrast, data-parallelism creates tasks by partitioning the data. Abstractions, such as MapReduce [56], enable the data-parallel processing and aggregation of data using high-level primitives. The runtime system then handles the implementation of the parallelism, i.e., the partitioning of the data, the mapping of data

to tasks, the orchestration and synchronization of tasks and data movements.

The dataflow model further generalizes the data-parallel model by supporting applications comprising of multiple stages of processing. The abstraction is based on directed acyclic graphs, where nodes represent multiple stages of processing and the flow of data between these stages. It was invented in the 1960s at MIT [57] and later adapted to the domain of data-intensive computing (LGDF2 [58], Dryad [59]) as a way to describe data processing pipelines comprising of multiple stages, e.g., map, reduce, shuffling. A stage can also be comprised of an external application (e.g., a simulation).

Iterative computation is a scenario applicable in particular to model training in machine learning applications. An important requirement of these types of applications is the need to cache data to facilitate reading and processing data multiple times [60]. In machine learning applications, this pattern is often found as many optimization techniques require multiple passes on the data to compute and update model parameters.

The last scenario is stream processing, defined as the ability to process unbounded data feeds and provide near-realtime insights [61]. The processing patterns for streaming are similar. However, the amount of data is typically smaller, as messages are processed in small batches. The management of state between individual messages can be required. Stream processing is used to analyze data streams from scientific experiments, e.g., light source sciences [32].

B. Application Requirements

To support these scenarios, an efficient resource management abstraction and middleware that can support highly diverse task-based workloads is required. The heterogeneity of the tasks in these application scenarios is high, i.e., often tasks with diverse runtime, resource, and data requirements need to be efficiently managed. For example, complex scenarios require the management of both long-running tasks, e.g., simulation tasks, and short-running tasks, e.g., data-parallel tasks arising from analytics applications. Further, data-intensive applications can be highly unpredictable due to data-dependencies and as a result very complex task graphs. The necessity to respond to dynamic events demands support for task creation at runtime. The requirements for the abstraction and middleware can be summarized as follows:

R1 Abstractions: Provide a higher-level abstraction that hides the details of complex distributed infrastructure, but allows reasoning about trade-offs. The abstractions should be *simple* and easy-to-use, while supporting as many application scenarios as possible (*applicable*). Further, it should be *generalizable* to multiple systems and implementations.

R2 Middleware for Application-Level Resource Management: Provide the ability to manage highly diverse parallel and dependent tasks and associated data on heterogeneous infrastructure comprising of complex hardware and software stacks. The system should support the *interoperable* use of infrastructures and should be *extensible* to new frameworks and applications.

| | Task-Parallel | Data-Parallel/MapReduce | Dataflow | Iterative | Streaming |
|---------------------|--|---|---|--|--|
| Description | Focus on functional decomposition into tasks and control flow | Decomposition based on data with minimal communication between tasks | Multiple processing stages modeled with a directed acyclic graph | Multiple generations of tasks with sharing of data between the generations | Processing of unbounded data feeds in near-realtime |
| Characteristics | Decomposition of a problem into a diverse set of dependent and parallel tasks | Embarrassingly parallel, loosely-coupled with minimal communication. Details, such as communication and synchronization hidden from the application | Multiple stages, loosely-coupled parallelism, global communication for shuffle operation | Loosely coupled parallelism with global communication for updating machine learning model parameters | Data is processed in small batches often using data-parallel algorithms. For many algorithms, a global state needs to be maintained across batches of data |
| Application Example | Molecular Dynamics [48], [49], Ensemble-Kalman Filter [50], Scientific Gateways and Workflows [51] | Map-Only analytics [52], Molecular Data analysis Hausdorff Distance [53] | MapReduce for sequence alignment [54], Molecular Data analysis leaflet finder and RMSD [53] | Machine learning algorithms, K-Means [55] | Streaming for light source data [32] |

TABLE I: **Data-Intensive Application Scenarios – Characteristics and Patterns:** Data-intensive applications are more complex than compute-oriented applications and require the management of data, I/O and compute resources.

R3 **Dynamism and Adaptivity:** Ability to respond to changes in the environment at runtime. Both middleware and abstraction need to support this capability.

R4 **Performance, Scalability, and Efficiency:** The system should provide adequate performance, mainly high-throughput and low latencies, for highly diverse task-based workloads. By doing so, the system supports the strong and weak *scaling* of applications while ensuring efficient resource usage.

IV. PILOT-ABSTRACTION: AN ABSTRACTION AND MODEL FOR DISTRIBUTED RESOURCE MANAGEMENT

The need for tools and high-level abstractions to support application development and the extreme heterogeneity of infrastructures has been widely recognized [12]. Resource management is a fundamental challenge in distributed and parallel computing. The current state is characterized by highly heterogeneous and fragmented systems, rigid point solutions, and a lack of a unified model for expressing data and compute tasks. The advent of data-intensive, machine learning, and streaming applications complicated the state even further. Infrastructure is getting more complicated by introducing new storage and memory tiers as well as accelerators.

Data-intensive applications exhibit complex characteristics and demand a highly flexible abstraction for allocating resources and managing highly diverse workloads of tasks. Balancing application characteristics and infrastructure requires careful consideration of application-level and infrastructure-level concerns. High-level abstractions are critical to retain developer productivity and to scale applications. For example, the ability to manage resources efficiently, taking into account application objectives is important. By infusing application knowledge (e.g., about the data, compute and I/O characteristics) into scheduling decisions, the runtime and scalability can be significantly improved [62]. Thus, data needs to be integrated into these abstractions as first-class citizen.

A. Pilot-Abstraction and Conceptual Model

The pilot-abstraction [6] is a unified abstraction for resource management on heterogeneous infrastructure from high-performance computing, high-throughput computing, big data, and cloud for distributed applications. In the following,

we discuss the experience of developing and extending the abstraction and underlying middleware systems to support the described application scenarios. In this section, we focus on the DSR activities and artifacts related to the design and creation of the abstraction and middleware system.

We follow the iterative design approach of DSR closely aligning the abstraction design to real application needs. The first system focused on the design of an application-internal resource management framework for replica-exchange simulations [48]. The resource management capabilities were evaluated using different application scenarios as case studies and performance measurements on different HPC infrastructures focusing on the internal resource management subsystem. Based on the positive evaluation, we generalized the abstraction and created a re-usable, application-agnostic, standalone pilot-job system called BigJob [63].

The observation of similar concepts for other infrastructures and applications [64] motivated the development of the pilot-abstraction [6], a general abstraction for the re-occurring concept of utilizing a placeholder job as a container for a set of computational tasks. The abstraction comprises of two main concepts, the pilot that represents a placeholder for a specific set of resources and the compute-unit, a self-contained task. The implementation of the pilot-job system conceals details about the resource management systems of the different infrastructures (e.g., HPC, HTC, and clouds). Thus, the user can focus on the composition of tasks rather than dealing with infrastructure specific aspects.

The pilot-abstraction addresses the need to efficiently and flexibly manage resources on application-level across distributed, heterogeneous infrastructure. Pilot-jobs provide two key capabilities: (i) they support the late binding of resources and workloads, and (ii) they provide a higher-level abstraction for the specification of application workloads removing the need to manage the execution of the workload manually. At the same time, they provide critical capabilities to compose task- and data-parallel workloads while providing optimal scalability and performance by managing task granularities, data dependencies, and I/O via the abstraction. The design of the pilot-abstraction aims to offer a simple as possible and general interface to these capabilities.

Another artifact of the design process is a conceptual model

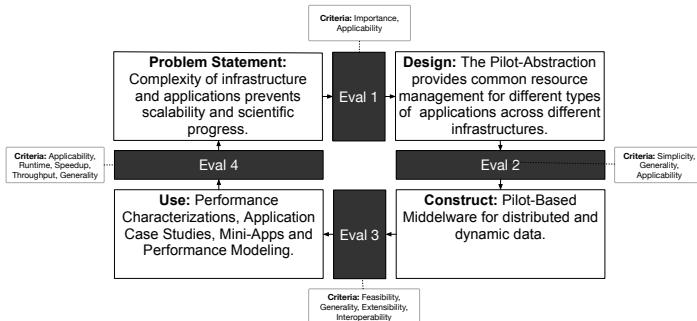


Fig. 3: DSR Evaluation Activities and Criteria (adapted from Sonnenberg [10]): The incremental evaluation provides valuable input for refinement and valuable knowledge that can be transferred to other problems.

for understanding pilot-based systems. The P* model [6] aims to provide a common framework to understand the abstraction, as well as commonly used pilot-job systems. The P* model defines the high-level concepts and mechanisms found in most pilot-systems. The model specification is done using a functional description of the components and interactions, as well as various components and interaction diagrams. The characteristics and interactions of all concepts and analysis of different pilot-systems using the model is available in [6]. Turilli et al. [65] refined the framework.

While the pilot-job concept was developed for HPC and HTC, the need to manage data in conjunction with pilots and tasks became apparent. Pilot-Data [66] extends the pilot-abstraction and provides the ability to manage storage and data, and couple these effectively with computational tasks. With the emergence of big data frameworks, such as Hadoop, Spark, and Dask, the ability to couple HPC applications to specialized data processing engines become increasingly important, which lead to the development of Pilot-Hadoop [67], [68]. Further, extensions for in-memory processing [68], and streaming [32] have been designed and implemented.

B. Middleware: System Design and Architecture

The objective of the system design phase is to create a system design and implementation that can support the desired abstractions. We applied methods and practices described earlier to achieve a flexible, maintainable, and comprehensible architecture. The system architecture is based on well-known design patterns [69], e.g., the adaptor pattern for abstracting specific resource types, i.e., HPC, cloud, and data infrastructures, such as Hadoop and Spark. For some of these infrastructures, we utilize the SAGA [70] as an access layer for local resource management systems. The design artifacts of the architecture model are created using block diagrams inspired by UML [71] to visualize system layers, composition, and interactions. Examples of architectural models artifacts can be found here: Pilot-Job [63], Pilot-Data [66], Pilot-Hadoop [67], and Pilot-Streaming [32].

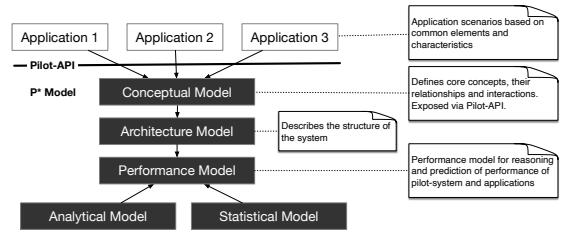


Fig. 4: Understanding the DSR Artifacts using Different Methods: Modeling techniques for characterization and evaluation of the pilot-abstraction and system.

V. EVALUATION

Evaluation is an essential part of the DSR process and ensures that the designed system achieves the desired purpose. We explain the distinct types of evaluation conducted on different artifacts produced throughout the Eval 1-4 activities proposed by Sonnenberg [10]. Figure 3 illustrates the four main activities: problem statement, design, construction, and use. We discuss in-depth the used methods and criteria used for evaluating the output of every stage.

Table II summarizes the evaluation methods used for the different DSR activities and the evolution of the pilot-abstraction. As proposed in [10], we evaluate the system interior, i.e., the architecture, as well as the exterior, i.e., the usage of the abstraction and system.

Figure 4 summarizes the modeling methods used. We use conceptual modeling to provide high-level intuition and to allow reasoning about inevitable trade-offs. Architecture models enable the evaluation of the internal structure of the systems. Performance models are used to describe the dynamic properties while using the abstraction and system. Insights from the conducted evaluations inform the abstraction design and to provide generalizable knowledge. In the following, we discuss the applied methods and criteria in detail.

A. Problem Identification and Design Evaluation (Eval 1/2)

Eval 1 activity, i.e., the justification of the problem statement and research gap, has been performed in the introduction and Section III. The results of the literature and application survey define the design objectives for the pilot-abstraction. The main criteria applied for evaluation of the problem was the importance and applicability of the design idea to a broad set of applications.

The design of the pilot-abstraction and middleware system is evaluated according to three main criteria: simplicity, generality, and applicability. An important artifact of the design phase is the P* conceptual model. The model defines the elements, characteristics, and interactions. The objective of P* is to provide a minimal but complete model that provides an intuition of the system. A metric for the simplicity of the model is the number of elements of the model, which is very low with four main concepts. The design of the pilot-abstraction reduces the amount of code necessary significantly while providing interoperability across different infrastructures. Further,

| | Pilot-Job [63], [6] | Pilot-Data [66] | Pilot-Hadoop [67] | Pilot-Memory [68] | Pilot-Streaming [32] |
|--|---|---|--------------------------------|--|---|
| Description | Management of computational tasks on heterogeneous infrastructure | Management of data and compute tasks | Management of Hadoop and Spark | Management of in-memory runtimes for iterative tasks | Streaming data sources and processing |
| Infrastructure | HPC, HTC, Cloud | HPC, Cloud, Hadoop/Yarn | HPC, Cloud, Hadoop/Yarn | HPC, Cloud, Hadoop/Yarn | HPC, Cloud, Serverless |
| System Design (Eval 2) | Conceptual model [6], architecture model [63] | Conceptual model [6], architecture model [66] | Architecture model [67] | Architecture model [68] | Architecture model [32] |
| Performance, Scalability and Efficiency (Eval 3) | Pilot overhead, application and task runtimes, strong scaling, analytical model for replica-exchange simulations [72] | Pilot overhead, application and task runtimes, strong scaling | Runtime, strong scaling | Runtime, strong scaling | Throughput, latency, scalability, statistical performance model for throughput [73] |
| Case Studies (Eval 4) | Adaptive Replica Exchange [48], [72], Ensemble Kalman Filter simulations [50], HIV binding [49], science portals [51], Pilot-MapReduce [54] | Genome Sequencing, K-Mean [66], [55] | Wordcount, K-Means | K-Means | Light source data reconstruction, K-Means |

TABLE II: **Evaluation:** Overview of Case Studies, Modeling Approaches and Performance Evaluation Methods Used.

we demonstrate the model’s generality by comparing and mapping different implementations of the pilot-abstraction [6].

B. System Implementation (Eval 3)

The Eval 3 activity evaluates the pilot-abstraction in artificial settings. The developed conceptual models provide an important basis for the construction of the system and the performance evaluation by offering essential information about the structure and expected behavior of the system.

The prototype implementation of the pilot-system is evaluated using an architecture model comprising of several component and interaction diagrams. The main criteria are feasibility, extensibility, interoperability.

The feasibility and generality of the abstraction is shown in various prototype and production implementations [63], [74]. Various extensions, e.g., for data management, in-memory processing, and in support of new infrastructures, such as cloud and serverless, demonstrate the extensibility of the system. The implementation maps the pilot-abstraction to the different infrastructures enabling interoperability. We verified the interoperability by various experiments with a broad set of distributed HPC and data-intensive applications.

C. Performance and Case Studies (Eval 4)

An important objective of the pilot-abstraction is to overcome barriers to scaling. Thus, performance and scalability are essential evaluation criteria as both are instrumental for the many scientific applications. We use three approaches: (i) performance characterization of the pilot-system and several applications, and (ii) analytical performance modeling and (iii) statistical performance modeling for selected use cases.

As benchmarks do not correctly reflect the requirements of scientific applications, we rely on custom experiments for evaluations. A challenge for performance characterizations and modeling is the experimental design and data collection. The experimental design is the process of determining the factors, factor levels, and combinations of these for an experiment to understand the effect of each factor while minimizing the number of experiments [29], [75]. A good experimental design is essential to capture essential characteristics while minimizing data collection efforts.

We propose the Mini App framework [32] to address these challenges and to automate and accelerate the build-assess-

refine cycle. The Mini App framework helps to evaluate abstractions, middleware, and infrastructure in real-world conditions. Further, the data collected can serve as a basis for statistical models and predictions. It was designed to support an excellent experimental design following best practices defined by Gray [31] and Waller [76]: (i) **Simplicity**: Easy-to-use and setup via high-level APIs and configurations. (ii) **Relevance**: It gives the developer full control of the application workload and metrics necessary for the application scenario. (iii) **Scalability**: Support for distributed resources and datasets at various scale levels and data rates. (iv) **Portability**: Infrastructure and application-agnostic by design. Different types of infrastructure supported via pilot-abstraction. (v) **Reproducibility**: It provides comprehensive automation of performance experiments ensuring repeatability and reproducibility.

Another important aspect of DSR is the ability to derive knowledge and insights. We use different modeling approaches to generalize abstractions, systems, and applications. For example, we provide analytical models for the performance of the application and pilot-systems [72], [66]. These models capture the significant components of the runtime and allow users to understand the impact of input data volume and parallelism on the runtime. Further, it enables the assessment of the system overheads and their ratio to the overall runtime of the application. Further, we use statistical modeling, e.g., for the prediction of the throughput of streaming systems for different infrastructure configurations [73].

Further, we evaluate the applicability of the abstraction in a natural setting, e.g., in various applications [50], [77], and frameworks [54]. In these investigations, we assess whether the pilot-abstraction meets the defined requirements concerning its capabilities, simplicity, and the feasibility to implement, deploy, and execute applications. In particular, we focus on the resource management requirements, such as the ability to adapt to changing resource needs, while providing adequate performance and scalability. The abstraction proved useful to capture the critical parameters necessary to express task and data decompositions and the associated performance trade-offs. In various case studies, we demonstrated that the abstraction allows a suitable control of the compute and data movements.

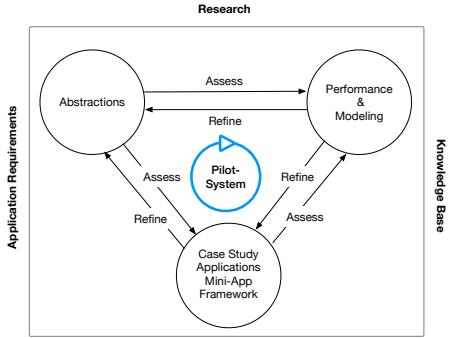


Fig. 5: **Iterative Research Approach:** Using an iterative feedback loop of abstraction design and evaluation using real-world and synthetic applications to refine design and system.

VI. DISCUSSION

Abstractions are vital for handling complexity and building systems at an unprecedented scale. We present the design science research method to design and evaluate the pilot-abstraction—an abstraction for enabling resource management across heterogeneous, distributed resources. By iteratively addressing real-world application and system challenges using the DSR method, we were able to design, develop and refine the pilot-abstraction, and evaluated it against the defined requirements:

- R1 **Abstractions:** The Pilot-Abstraction’s capabilities and simplicity have been evaluated and validated in several application scenarios, e.g., ensemble simulations, data-intensive applications, and streaming. Further, the extensive usage of the Pilot-Abstraction for higher-level building blocks, e.g., a workflow framework [78], an ensemble simulations management framework, and a MapReduce framework [54], demonstrates its viability and usefulness.
- R2 **Middleware for Application-Level Resource Management:** The pilot-system provides interoperable use of HPC, cloud, and data infrastructures. In [79], we explore the interoperable use of HPC, HTC, and clouds. In [66], we use and characterize the use of Pilot-Data on HPC and HTC resources. The system is extensible to new infrastructures, such as Hadoop [67], streaming [32], and serverless [73].
- R3 **Dynamism and Adaptivity:** An important capability of the pilot-abstraction is the ability to respond to changes in the environment at runtime. In [63], we explore the usage of additional cloud resources at runtime to meet application demands. In [73], we demonstrated a model for throughput prediction to determine the optimal set of resources for a given workload.
- R4 **Performance, Scalability, and Efficiency:** We demonstrated in various studies that the pilot-abstraction enables the creation of scalable application by given fine-grained control on data/task composition while hiding the details [63], [66], [53], [73].

The incremental evaluation of the artifacts of the DSR pro-

cess provides valuable input for future iterations and generalizable knowledge for similar problems. In the following, we describe and synthesize our experiences of the development of the pilot-abstraction in a set of lessons learned to inform the design process of future systems.

Iteration: The iterative design and evaluation process of DSR is instrumental in creating appropriate abstractions and middleware systems. Building real systems and applications is instrumental in discovering new usage modes and further requirements. Implementing smaller working systems is instrumental before scaling to more extensive resources and further applications. Specifically, we iteratively grew the pilot-job system from supporting coarse-grained ensembles of simulation tasks on single infrastructures to support for high-volume, fine-grained data-parallel tasks, and streaming.

Automation: Collecting data on the design is an instrumental part of the process. Automating experiments for performance characterizations and measurements is important to enable the exploration of larger parameter spaces and to ensure reproducibility. We developed the Mini Apps framework to formalize and automate the experiments and data collection. Figure 5 illustrates the feedback loop used for the design of pilot-abstraction and the implementation in the pilot-system. By using continuous evaluations, partially automated with the Mini App framework, valuable inputs for the abstraction and experimental design and modeling process are generated.

Abstraction Design: The design process is complex and requires the careful trade-off of capabilities, simplicity, and generality. The more application-specific knowledge can be induced via abstractions into middleware systems, the better the decision the system can make, e.g., concerning scheduling. However, the more application-specific the abstraction, the less general is its utility. Balancing simplicity, generality, and capability is challenging and requires a careful evaluation of the abstraction in different applications and settings.

Compute and Data: Managing heterogeneous compute tasks at scale is challenging by itself. The addition of data complicates the problem significantly. There is a significant amount of heterogeneity and dynamism in the way data can be stored, transferred, and used. Typically, a great extent of the data lifecycle is external to the applications. We address these challenges, particularly by focusing on defined application scenarios (see Table I) and by supporting and optimizing for important patterns, e.g., MapReduce.

Optimize Application Algorithms: A universal abstraction and system for resource management can help to scale applications by simplifying and standardizing the process of resource and task management. In many cases, an improvement of algorithms can lead to even more significant improvements compared to scaling out a non-optimal algorithm to more resources (see e.g. [53]).

Limitations of Abstractions: In many cases, systems are not limited by conceptual abstraction, but by the implementation of the system and infrastructure. Further, abstractions can exhibit undesirable behaviors. Leaky abstractions describe the phenomena that abstractions frequently fail in real-world set-

tings exposing complexities from underlying systems that it meant to abstract [80].

Re-Use and Interoperability: A well-designed abstraction is a minimal requirement for developing robust and scalable software systems. By abstracting commonalities between systems, interoperability can be achieved. However, significant investments into the stability and robustness of the system are required to support real-world applications.

REFERENCES

- [1] Anthony J. G. Hey, Stewart Tansley, and Kristin M. Tolle. *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Microsoft, 2009.
- [2] Geoffrey Fox, Judy Qiu, David Crandall, Gregor Von Laszewski, Oliver Beckstein, John Paden, Ioannis Paraskevakos, Shantenu Jha, Fusheng Wang, Madhav Marathe, Anil Vullikanti, and Thomas Cheatham. Contributions to high-performance big data computing. In L. Grandinetti, G.R. Joubert, K. Michielsen, S.L. Mirtaheri, M. Taufer, and R. Yokota, editors, *Future Trends of HPC in a Disruptive Scenario*. IOS Press Volume 34 of Advances in Parallel Computing, 2019.
- [3] Jeffrey S. Vetter, Ron Brightwell, Maya Gokhale, ..., and Jeremiah Wilke. Extreme heterogeneity 2018 - productive computational science in the era of extreme heterogeneity: Report for doe ascr workshop on extreme heterogeneity.
- [4] Alfred V. Aho and Jeffrey D. Ullman. *Foundations of Computer Science*. Computer Science Press, Inc., USA, 1992.
- [5] Mary Shaw. *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages*, chapter The Impact of Modelling and Abstraction Concerns on Modern Programming Languages. In *Topics in Information Systems* [18], 1984.
- [6] Andre Luckow, Mark Santcroos, Andre Merzky, Ole Weidner, Pradeep Mantha, and Shantenu Jha. P*: A model of pilot-abstractions. *IEEE 8th International Conference on e-Science*, pages 1–10, 2012. <http://dx.doi.org/10.1109/eScience.2012.6404423>.
- [7] Juhani Iivari. A paradigmatic analysis of information systems as a design science. *Scandinavian Journal of Information Systems*, 19:39–, 01 2007. <https://aisel.aisnet.org/sjis/vol19/iss2/51>.
- [8] Alan R. Heyner, Salvatore T. March, Jinsoo Park, and Sudha Ram. Design science in information systems research. *MIS Quarterly*, 28(1):75–105, 2004.
- [9] Ken Peffers, Tuure Tuunanen, Marcus Rothenberger, and Samir Chattee. A design science research methodology for information systems research. *J. Manage. Inf. Syst.*, 24(3):45–77, December 2007.
- [10] Christian Sonnenberg and Jan vom Brocke. Evaluations in the science of the artificial – reconsidering the build-evaluate pattern in design science research. In Ken Peffers, Marcus Rothenberger, and Bill Kuechler, editors, *Design Science Research in Information Systems. Advances in Theory and Practice*, pages 381–397, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.
- [11] Judith Segal. Models of scientific software development. In *SECSE 08, First International Workshop on Software Engineering in Computational Science and Engineering*, May 2008. Workshop co-located with ICSE 08 <http://icse08.upb.de/>.
- [12] Shantenu Jha, Murray Cole, Daniel S. Katz, Manish Parashar, Omer Rana, and Jon Weissman. Distributed computing practice for large-scale science and engineering applications. *Concurrency and Computation: Practice and Experience*, 25(11):1559–1585.
- [13] Timothy Mattson, Beverly Sanders, and Berna Massingill. *Patterns for Parallel Programming*. Addison-Wesley Professional, first edition, 2004.
- [14] Micah Beck. On the hourglass model. *Commun. ACM*, 62(7):48–57, June 2019.
- [15] Joshua Bloch. How to design a good api and why it matters. In *Companion to the 21st ACM SIGPLAN Symposium on Object-Oriented Programming Systems, Languages, and Applications, OOPSLA '06*, page 506–507, New York, NY, USA, 2006. ACM.
- [16] John Mylopoulos. Conceptual modelling and telos. <http://www.cs.toronto.edu/~jm/2507S/Readings/CM+Telos.pdf>, 1992.
- [17] Jeff Johnson and Austin Henderson. Conceptual models: Begin by designing what to design. *Interactions*, 9(1):25–32, January 2002.
- [18] Michael L. Brodie, John Mylopoulos, and Joachim W. Schmidt. *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages*. Topics in Information Systems. Springer New York, 1984.
- [19] Bernd Bruegge and Allen H. Dutoit. *Object-Oriented Software Engineering Using UML, Patterns, and Java*. Prentice Hall Press, Upper Saddle River, NJ, USA, 3rd edition, 2009.
- [20] M. Shaw. The coming-of-age of software architecture research. In *Proceedings of the 23rd International Conference on Software Engineering, ICSE 2001*, pages 657–664a, May 2001.
- [21] C. Alexander, P.D.A.C. Alexander, S. Ishikawa, M. Silverstein, M. Jacobson, Center for Environmental Structure, I. Fiksdahl-King, and A. Shlomo. *A Pattern Language: Towns, Buildings, Construction*. Center for Environmental Structure Berkeley, Calif: Center for Environmental Structure series. OUP USA, 1977.
- [22] Kent Beck and Ward Cunningham. Using pattern languages for object oriented programs. In *Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA)*, 1987.
- [23] David Lorge Parnas. On the criteria to be used in decomposing systems into modules. *Commun. ACM*, 15(12):1053–1058, December 1972.
- [24] David Lorge Parnas. Information distribution aspects of design methodology. *Methods*, 4(5):6–7, 1971.
- [25] David Garlan and Mary Shaw. An introduction to software architecture. Technical report, Pittsburgh, PA, USA, 1994.
- [26] I. Foster and C. Kesselman. *The Grid 2: Blueprint for a New Computing Infrastructure*. ISSN. Elsevier Science, 2003.
- [27] Judy Qiu, Shantenu Jha, Andre Luckow, and Geoffrey C. Fox. Towards hpc-abds: An initial high-performance big data stack. In *Proceedings of ACM Big Data Interoperability Framework Workshop*, 2015.
- [28] Kathleen M. Eisenhardt. Building theories from case study research. *The Academy of Management Review*, 14(4):532–550, 1989.
- [29] Raj Jain. *The art of computer systems performance analysis - techniques for experimental design, measurement, simulation, and modeling*. Wiley professional computing. Wiley, 1991.
- [30] D. Ferrari. *Computer Systems Performance Evaluation*. Prentice-Hall, 1978.
- [31] Jim Gray. *Benchmark Handbook: For Database and Transaction Processing Systems*. Morgan Kaufmann, San Francisco, CA, USA, 1992.
- [32] André Luckow, George Chantzialexiou, and Shantenu Jha. Pilot-streaming: A stream processing framework for high-performance computing. *IEEE eScience International Conference*, abs/1801.08648, 2018.
- [33] W. Buchholz. A synthetic job for measuring system performance. *IBM Systems Journal*, 8(4):309–318, 1969.
- [34] J. W. Anderson, K. E. Kennedy, L. B. Ngo, A. Luckow, and A. W. Apon. Synthetic data generation for the internet of things. In *2014 IEEE International Conference on Big Data (Big Data)*, 2014.
- [35] Andre Merzky, Ming Tai Ha, Matteo Turilli, and Shantenu Jha. Synapse: Synthetic application profiler and emulator. *Journal of Computational Science*, 27:329 – 344, 2018.
- [36] Adolfy Hoisie. Performance modeling overview. Talk at PAM 2018: Performance Analysis and Modeling Workshop: https://indico.bnl.gov/event/3950/contributions/12021/attachments/10817/13215/Talk_at_the_Perf_Workshop_Feb_2018.pdf, 2018.
- [37] A. Bordgida, J. Mylopoulos, and H. K. T. Wong. *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages*, chapter Generalization/Specialization as a Basis for Software Specification. In *Topics in Information Systems* [18], 1984.
- [38] Danilo Bzdok, Naomi Altman, and Martin Krzywinski. Statistics versus machine learning. *Nature Methods*, 15:233 EP –, 04 2018.
- [39] P. Griffiths Selinger, M. M. Astrahan, D. D. Chamberlin, R. A. Lorie, and T. G. Price. Access path selection in a relational database management system. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD '79*, NY, NY, USA, 1979. ACM.
- [40] Gene M. Amdahl. Validity of the single processor approach to achieving large scale computing capabilities. In *Proceedings of the April 18-20, 1967, Spring Joint Computer Conference, AFIPS '67 (Spring)*, pages 483–485, New York, NY, USA, 1967. ACM.
- [41] Geoffrey C. Fox, James A. Glazier, J. C. S. Kadupitiya, Vikram Jadhao, Minje Kim, Judy Qiu, James P. Sluka, Endre T. Somogyi, Madhav Marathe, Abhijin Adiga, Jiangzhuo Chen, Oliver Beckstein, and Shantenu Jha. Learning everywhere: Pervasive machine learning for effective high-performance computation. *CoRR*, abs/1902.10810, 2019.
- [42] Engineering National Academies of Sciences and Medicine. *Future Directions for NSF Advanced Computing Infrastructure to Support U.S.*

Science and Engineering in 2017-2020. The National Academies Press, Washington, DC, 2016.

[43] Krste Asanović, Ras Bodik, Bryan Christopher Catanzaro, Joseph James Gebis, Parry Husbands, Kurt Keutzer, David A. Patterson, William Lester Plishker, John Shalf, Samuel Webb Williams, and Katherine A. Yellick. The landscape of parallel computing research: A view from berkeley. Technical Report UCB/EECS-2006-183, EECS Department, University of California, Berkeley, Dec 2006.

[44] Shantenu Jha, Daniel S. Katz, Andre Luckow, Neil Chue Hong, Omer Rana, and Yogesh Simmhan. Introducing distributed dynamic data-intensive (d3) science: Understanding applications and infrastructure. *Concurrency and Computation: Practice and Experience*, 29(8), 2017.

[45] Geoffrey C. Fox, Shantenu Jha, Judy Qiu, and Andre Luckow. Towards an understanding of facets and exemplars of big data applications. In *Proceedings of Beowulf'14*, Annapolis, MD, USA, 2014. ACM.

[46] Geoffrey C. Fox, Shantenu Jha, Judy Qiu, and Andre Luckow. A systematic approach to big data benchmarks. In Lucio Grandinetti, Gerhard Joubert, Marcel Kunze, and Valerio Pascucci, editors, *Big Data and High Performance Computing*, volume 24, pages 47–66. IOS Press, München, 2015. <http://dx.doi.org/10.3233/978-1-61499-583-8-47>.

[47] NIST BigData Working Group. <http://bigdatawg.nist.gov/usecases.php>, 2019.

[48] Andre Luckow, Shantenu Jha, Joohyun Kim, Andre Merzky, and Bettina Schnor. Adaptive Replica-Exchange Simulations. *Royal Society Philosophical Transactions A*, pages 2595–2606, jun 2009.

[49] David W. Wright, Benjamin A. Hall, Owain A. Kenway, Shantenu Jha, and Peter V. Coveney. Computing clinically relevant binding free energies of hiv-1 protease inhibitors. *Journal of Chemical Theory and Computation*, 10(3):1228–1241, 2014. PMID: 24683369.

[50] Yaakoub El-Khamra and Shantenu Jha. Developing autonomic distributed scientific applications: A case study from history matching using ensemble kalman-filters. In *Proceedings of the 6th International Conference Industry Session on Grids Meets Autonomic Computing*, GMAC '09, pages 19–28, New York, NY, USA, 2009. ACM.

[51] Sharath Maddineni, Joohyun Kim, Yaakoub El-Khamra, and Shantenu Jha. Distributed application runtime environment (dare): A standards-based middleware framework for science-gateways. *Journal of Grid Computing*, 10(4):647–664, 2012.

[52] G. C. Fox, J. Qiu, S. Kamburugamuve, S. Jha, and A. Luckow. Hpc-abds high performance computing enhanced apache big data stack. In *2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pages 1057–1066, May 2015.

[53] Ioannis Paraskevakos, Andre Luckow, Mahzad Khoshlessan, George Chantzialexiou, Thomas E. Cheatham, Oliver Beckstein, Geoffrey C. Fox, and Shantenu Jha. Task-parallel analysis of molecular dynamics trajectories. In *Proceedings of the 47th International Conference on Parallel Processing*, ICPP 2018, New York, NY, USA, 2018. ACM.

[54] Pradeep Kumar Mantha, Andre Luckow, and Shantenu Jha. Pilot-MapReduce: An Extensible and Flexible MapReduce Implementation for Distributed Data. In *Proceedings of third international workshop on MapReduce and its Applications*, MapReduce '12, pages 17–24, New York, NY, USA, 2012. ACM.

[55] Shantenu Jha, Judy Qiu, André Luckow, Pradeep Kumar Mantha, and Geoffrey Charles Fox. A tale of two data-intensive paradigms: Applications, abstractions, and architectures. *Proceedings of 3rd IEEE International Congress of Big Data*, abs/1403.1528, 2014.

[56] Jeffrey Dean and Sanjay Ghemawat. Mapreduce: simplified data processing on large clusters. *Commun. ACM*, 51(1):107–113, January 2008.

[57] William Robert Sutherland. *The on-line graphical specification of computer procedures*. PhD thesis, MIT, 1966.

[58] D. C. DiNucci and R. G. Babb. Design and implementation of parallel programs with Igdf2. In *Digest of Papers. COMPCON Spring 89. Thirty-Fourth IEEE Computer Society International Conference: Intellectual Leverage*, pages 102–107, Feb 1989.

[59] Michael Isard, Mihai Budiu, Yuan Yu, Andrew Birrell, and Dennis Fetterly. Dryad: distributed data-parallel programs from sequential building blocks. *SIGOPS Oper. Syst. Rev.*, 41(3):59–72, 2007.

[60] Jaliya Ekanayake, Hui Li, Bingjing Zhang, Thilina Gunarathne, Seung-Hee Bae, Judy Qiu, and Geoffrey Fox. Twister: A runtime for iterative mapreduce. In *Proceedings of the 19th ACM International Symposium on High Performance Distributed Computing*, HPDC '10, pages 810–818, New York, NY, USA, 2010. ACM.

[61] Geoffrey Fox, Shantenu Jha, and Lavanya Ramakrishnan. Stream 2015 final report. <http://streamingsystems.org/finalreport.pdf>, 2015.

[62] Anand Jayarajan, Jinliang Wei, Garth Gibson, Alexandra Fedorova, and Gennady Pekhimenko. Priority-based parameter propagation for distributed dnn training. In *Proceedings of SysML*, 05 2019.

[63] Andre Luckow, Lukas Lacinski, and Shantenu Jha. SAGA BigJob: An Extensible and Interoperable Pilot-Job Abstraction for Distributed Applications and Systems. In *The 10th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pages 135–144, 2010.

[64] James Frey, Todd Tannenbaum, Miron Livny, Ian Foster, and Steven Tuecke. Condor-g: A computation management agent for multi-institutional grids. *Cluster Computing*, 5(3):237–246, July 2002.

[65] Matteo Turilli, Mark Santcroos, and Shantenu Jha. A comprehensive perspective on pilot-job systems. *ACM Comput. Surv.*, 51(2):43:1–43:32, April 2018.

[66] Andre Luckow, Mark Santcroos, Ashley Zebrowski, and Shantenu Jha. Pilot-data: An abstraction for distributed data. *Journal of Parallel and Distributed Computing*, 2014.

[67] Andre Luckow, Pradeep Kumar Mantha, and Shantenu Jha. Pilot-abstraction: A valid abstraction for data-intensive applications on hpc, hadoop and cloud infrastructures? *CoRR*, abs/1501.05041, 2015.

[68] A. Luckow, I. Paraskevakos, G. Chantzialexiou, and S. Jha. Hadoop on HPC: Integrating Hadoop and Pilot-based Dynamic Resource Management. *IEEE International Workshop on High-Performance Big Data Computing in conjunction with The 30th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2016)*, 2016.

[69] Erich Gamma, Richard Helm, Ralph Johnson, and John M. Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley Professional, 1 edition, 1994.

[70] Andre Merzky, Ole Weidner, and Shantenu Jha. SAGA: A standardized access layer to heterogeneous distributed computing infrastructure. *Software-X*, 2015. DOI: 10.1016/j.softx.2015.03.001.

[71] Object Management Group (OMG). Unified modeling language specification version 2.5.1. <https://www.omg.org/spec/UML/2.5.1/>, 2017.

[72] Abhinav Thota, Andre Luckow, and Shantenu Jha. Efficient large-scale Replica-Exchange Simulations on Production Infrastructure. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1949):3318–3335, 2011.

[73] Andre Luckow and Shantenu Jha. Performance characterization and modeling of serverless and hpc streaming applications. In *Proceedings of StreamML Workshop at IEEE International Conference on Big Data (IEEE BigData 2019)*, 2019.

[74] Andre Merzky, Matteo Turilli, Manuel Maldonado, Mark Santcroos, and Shantenu Jha. Using pilot systems to execute many task workloads on supercomputers. In *Job Scheduling Strategies for Parallel Processing*, pages 61–82, Cham, 2019. Springer International Publishing.

[75] M. Hauck. *Automated Experiments for Deriving Performance-relevant Properties of Software Execution Environments*. The Karlsruhe Series on Software Design and Quality. KIT Scientific Publishing, 2014.

[76] Jan Waller. Performance benchmarking of application monitoring frameworks. https://macau.uni-kiel.de/receive/diss_mods_00016245, 2015.

[77] Jack A. Smith, Melissa Romanus, Pradeep Kumar Mantha, Yaakoub El Khamra, Thomas C. Bishop, and Shantenu Jha. Scalable online comparative genomics of mononucleosomes: A bigjob. In *Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery*, XSEDE '13, NY, NY, USA, 2013. ACM.

[78] Matteo Turilli, Vivek Balasubramanian, Andre Merzky, Ioannis Paraskevakos, and Shantenu Jha. Middleware building blocks for workflow systems. *Computing in Science & Engineering (CiSE) special issue on Incorporating Scientific Workflows in Computing Research Processes*. 10.1109/MCSE.2019.2920048 (2019).

[79] Shantenu Jha, Daniel S. Katz, Andre Luckow, Andre Merzky, and Katerina Stamou. *Understanding Scientific Applications for Cloud Environments*, pages 345–371. John Wiley and Sons, 1 2011.

[80] Joel Spolsky. The law of leaky abstractions. *Joel on Software: And on Diverse and Occasionally Related Matters*, 01 2002.