LA-UR-17-22359
Approved for public release; distribution is unlimited.

Title: Packet Capture Solutions: PcapDB Benchmark for High-Bandwidth Capture, Storage, and Searching

Author(s): Steinfadt, Shannon Irene
Ferrell, Paul Steven

Intended for: Report

Issued: 2017-03-21

Packet Capture Solutions

PcapDB Benchmark for High-Bandwidth Capture, Storage, and Searching

Dr. Shannon Steinfadt and Paul Ferrell

February 27, 2017

Los Alamos
NATIONAL LABORATORY
EST.1943

Overview of PcapDB

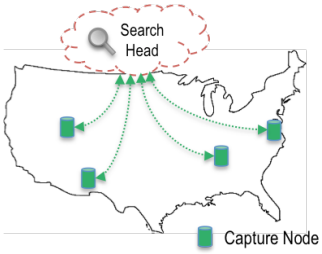
PcapDB is designed first and foremost as a full packet capture system built for distributed deployment across geographically disparate networks. While it lacks the analytic capability available in Bro and others, it directly provides high-speed search and packet indexing capabilities.

PcapDB was designed to not require load balancing based upon flows, and can utilize much simpler and less error prone schemes. While PcapDB does organize packets by flow on ingest, it

does not require that these packets be routed to the same processor or even the same host. As a result, PcapDB installations need only be concerned with the overall network speed.

1.1 PcapDB General Advantages

PcapDB provides several features outlined below:



1.1.1 Geological Distribution

PcapDB is designed to be deployed, managed, and searched across geologically disparate locations with minimal configuration requirements at each site. Distributing Bro requires that all nodes, both worker and manager, be fully accessible to each other through password-free SSH. PcapDB, in contrast, only requires that the management node (Search Head) be accessible from the worker Capture Nodes. This allows for the setup of distributed capture systems that cross both site and organizational boundaries without giving unnecessary access to a managing node.

Users in PcapDB are also managed with mult-site distribution in mind. Searching and limited admin privileges are configurable on a site-by-site basis. The search system itself can group results by site, or merge them into a single view as needed.

PcapDB sets the bar for management of multiple sets of capture boxes with a single, integrated search head. This feature is one of many that set PcapDB apart from other packet capture solutions for large, geographically disparate organizations like the DOE. Pcap is not being pushed over the wire to another location, cutting bandwidth in half. It is stored locally at the Capture Node(s) and searching occurs through the Search Head. Pcap data does not need to be moved or accessed for searching, only the stored indexes. Pcap is directly accessed only when a pull request for that specific pcap is made.

1.1.2 Built-In Flow Indexing

The ability to quickly search network flows in a distributed manner forms the basis of PcapDB's overall design. Its patent-pending flow indexing technology provides a specialized database system specifically optimized around providing fast flow searches. Once relevant flows are found, pulling the related Pcap data requires minimal disk I/O. As mentioned, these indexes are tiny relatively to the overall size of the captured packets; small enough to be affordably stored on relatively expensive SSD's.

1.1.3 Ultrafast Searching

In order to accelerate kill chain reconstruction, it is imperative to quickly locate and decode traffic and session during and after a security event. This access to the network data is a necessity. PcapDB provides ultrafast searching with it's highly optimized, highly efficient, patent pending indexing and searching algorithms.

1.1.4 PcapDB 100 Gb

A PcapDB 100 Gb solution would consist of a load-balanced set of 10 Gb capture nodes, and a single search head host. These capture nodes are relatively affordable: each capture system costing around $3k, since only commodity (COTS) hardware is required. While storage comprises the primary cost, commodity JBOD storage is one of the most affordable options. With PcapDB's built-in disk management, using JBODs with this disk management keep costs to

a bare minimum. The search head has minimal hardware requirements and can easily be virtualized or put in a cloud environment where it can search across multiple sites of varying size.

1.1.5 A Unified System

PcapDB is not a collection of scripts. It is a full system, designed and developed to be scalable, geographically distributed, and highly targeted for fast storage, patent-pending indexing approaches and very efficient searching. PcapDB's Web interface for the system provides full system management capabilities, including disk management for JBOD enclosures, user access controls, and searching. This is in addition to the RESTful API that can be used to pipeline searches and return results to and from existing tools. The JBOD storage option removes any cost and capacity barriers that some hardware vendors impose for "special" storage requirements. This can remove a significant amount of overhead spent on the hardware, again the majority of which is spent on disk capacity.

PcapDB is open source, eliminating costly licensing agreements. Hardware upgrades are made on the users' time schedule since there is no planned end-of-life for your investment. And unlike some approaches, PcapDB handles both IPv4 and the ever-growing IPv6.

1.2 PcapDB Next Steps

Continued development of PcapDB under iJC3 would focus on the following areas:

- Flow/Packet compression
    - Further minimize storage costs and maximize disk usage for a longer history of network traffic. This is essential when many threats are not detected for days, weeks, or even longer.
    - Given 30 days of capture, every 3% reduction in overall captured data size produces a sizable one day gain in capacity.
    - Organizing by flows already offers ample opportunities for applying novel compression strategies that will be explored with these next steps.
    - Additional compression strategies could be selectively applied dependent upon system load.
- Incorporation of WireCap as an alternative to PFRing to support
    - PFRing is license burdened under certain capture modes. WireCap can potentially eliminate the PFRing dependency.
    - WireCap potentially offers better performance than PFRing, which we propose to evaluate and integrate it into PcapDB.
- Further interface refinements and indexing on additional flow fields.

2 Other Packet Capture Solutions

The remainder of this document gives a brief overview of several other competing solutions, with the majority of space devoted to those that are most viable for the high-bandwidth, geographically distributed DOE domain. These comparisons are primarily based on publically available information, and some direct experience with several of them. Since we have not had the opportunity to directly evaluate the myriad of possibilities, we give sales and promotional material the benefit of the doubt in their performance and ability claims.

In these summaries, we are generally looking for the following capabilities in comparison to PcapDB:

1. **Cost**: High-speed packet capture no longer requires specialty hardware. The hardware that typically comes with a commercial solution is the same hardware we use for PcapDB, only it is rebranded with a significant markup.
2. **Maximum per box capture rates**: Many network security tools can only handle 10Gb/s in short bursts. Sustained 10Gb/s capture rates per capture host is an assumed minimum.
3. **Distributed Capture and Search**: We need to capture at rates higher than what a single 10 Gb/s box can handle, so capture must be distributed. See the load balancing section below to understand why.
4. **Web User Interface (UI)**: Most existing open source solutions are command line only. A web interface can provide flexibility and ease of use.
5. **Connectivity**: If the capture nodes need to communicate with each other *and* the capture nodes are widely distributed, this becomes significantly more difficult.
6. **Geographic Distribution**: Can multiple sets of capture boxes be managed and search from them via a search head?

3 BRO

(open-source, https://www.bro.org/index.html)

Bro is a dynamic, capable IDS system designed to selectively analyze network flows, and can take a variety of actions in response to that analysis. Since that analysis requires the full capture of the flow and those actions can include storing and recording flow records, Bro can be used as a selective packet capture system.

PcapDB was designed to not require load balancing based upon flows, and can utilize much simpler and less error prone schemes. While PcapDB does organize packets by flow on ingest, it does not require that these packets be routed to the same processor or even the same host. As a result, PcapDB installations need only be concerned with the overall network speed.

3.1 Commonalities

- Both Open Source, DOE developed technologies
- Similar hardware requirements per box (though capacity may differ).
- Both utilize the PFRing library for high-speed capture.

3.2 Load Balancing

Both PcapDB and Bro can handle loads distributed via a load balancer, but Bro's requirements for load balancing are significant.

Bro requires that all packets of the same flow be sent to the same Bro sensor, as it must have these packets collected at a single source in order to process them. Load balancing according to flow is achievable using a variety of software and hardware solutions, all of which have the same base challenges. Once a flow is committed to be balanced to a given output, it must continue to

be sent to that output regardless of the resources the flow consumes. In the case where this exceeds the capacity of either the base network layer the Bro host connected to it, packets will be lost or redistributed. Building sufficient excess capacity can alleviate this issue, but never eliminate it entirely.

This is opposed to the design of PcapDB, which does not require flow-based load balancing, thus utilizing simpler and less error-prone schemes. While PcapDB does organize packets by flow on ingest, it does not require that these packets be routed to the same processor or even the same Capture Node. As a result, PcapDB installations need only be concerned with the overall network speed.

3.3 Selective Packet Capture

Bro is designed for selective full packet capture, in that expects to save only those flows that match a certain set of criteria only. While these flows can be indexed in a separate database, the system simply is not designed to capture process searches for *every* flow it encounters and will quickly run into capacity issues when trying to do so. Bro's greatest strength: individual flow analysis also makes it difficult to group storage and index transactions to increase efficiency. PcapDB is designed to capture, index, and search *every* flow and packet on high network speeds. Its primary limitation is providing the I/O bandwidth for storing those packets, which is a concern Bro also has to contend with.

Captured packets and index data are stored in bulk as efficiently as possible in PcapDB. In addition, the indexes are compact, typically less than 0.5% the size of the captured packets, and designed specifically for the ultra fast searching of network data.

To increase efficiency in Bro, the decision to *not* capture every flow creates a different conundrum. One of the great strengths of packet capture is that it allows us find and analyze events we did not expect, such as command and control packet over ports we would normally ignore. Selectively capturing packets requires that we predict in advance the methods adversaries will use in advance. While Bro and its deep packet inspection may be enough to partially alleviate these concerns, the blanket filtering rules necessary to cover unforeseen traffic requires considerable foresight and configuration complexity.

4 Time Machine

(open-source, https://www.bro.org/community/time-machine.html)

Time machine comes out of the Bro community. It is an open-source tool. There are indexing similarities for both tools in that a 5-tuple of network information is used for indexing. Unlike PcapDB, the system is not set up to handle full capture on fully utilized Gbps links. Instead, it uses a mechanism called a "connection cutoff," to only capture the first X bytes and reduce the amount of data to process. For space savings, they state "this approach it does not impair the analysis capabilities (unless the cutoff is set to low) because most of the 'interesting' data is located in the first few packets of a connection."

When Time Machine is coupled with the Bro IDS, "the IDS can directly interact with the Time Machine and request historic traffic to represent it to a security analyst or to do

retrospective analysis." This type of interaction with PcapDB can be achieved using the RESTful interface to the system.

As with Bro, the system is not designed for connectivity and geographic distribution with different levels of access permitted.

Users have reported turning off indexing in Time Machine due to packet drops / packet loss when it is enabled. In response, there are work-arounds that have been developed that require using a virtual machine and brute force searching using GNU parallel using command line arguments.

PcapDB indexing is one of the keys to its success. While both API and web interface enable fast searching, the interactive web interface handles syntax checking of those searches. Additionally, searches in PcapDB are very fast, allowing for a responsive system that is a unified, single system that does full packet capture, indexing, and search with a smart, methodical approach. As users refine their searches, recently executed searches and the indices to those results are cached, increasing the system response speed to searching and reducing duplication. The underlying indexing structure for Time Machine is unknown to the authors at the time of writing.

5 Endace

(commercial, https://www.endace.com/endace-high-speed-network-recorders.html)

Endace is a commercial company that has an appliance and software application structure. They have multiple types of solutions and applications available. Focusing on EndaceProbe, they are described as "a family of network recorders capable of capturing, indexing and recording network traffic with 100% accuracy on even the fastest, most complex networks."

Endace is a capable system, with the richest feature set and can match or exceed PcapDB for packet capture. Endace and PcapDB mirror each other in their high-bandwidth capacity with their DAG hardware cards. The concepts reflect each other in centralized management, scalability, and API availability. The drawback with Endace is that each of capabilities is an add-on with additional hardware or software required, with a potentially high price tag. The Endace business model and cost are what make PcapDB an affordable alternative with no planned end of life, and no special hardware required.

A user reported great difficulty with consistency and support.

5.1 Centralized Management

For centralized management, EndaceCMS (Central Management Server) enables "connected fabric of hundreds of EndaceProbes (and other Endace appliances) to be centrally managed through a powerful GUI interface and a suite of command-line tools."

Central management in PcapDB is through the Search Head node. This can be a separate server, or virtualized into a cloud environment. The Search Head software is included in the open-source PcapDB software. Users in PcapDB are also managed with multi-site distribution in mind.

Searching and limited admin privileges are configurable on a site-by-site basis. The search system itself can group results by site, or merge them into a single view as needed.

5.2 100GbE Network Speeds

In order to have the high-bandwidth capability for 100GbE, EndaceAccess is required, for load balancing and splitting it across multiple 10GbE egress ports. The PcapDB solution is outlined in Section 1.1.4.

5.3 Deep Packet Inspection

Endace does have "Deep Packet Inspection" (DPI), with Layer 7 Application awareness. This is where the main trade-off occurs between Endace with DPI and PcapDB's fast, light-weight, disk-efficient capture. PcapDB, like Stenographer (see Section 11), utilize the Transport Layer. PcapDB mirrors how many responders use PcapDB: time-based searching utilizing IP addresses, port numbers, and protocol. This trade-off of DPI enables index sizes less than 0.5% of the size of captured packets, enabling a much longer history at a fraction of the disk capacity.

5.4 Cost Savings

There are two main cost savings that come from using PcapDB over a commercial solution such as Endace: hardware and licensing. Many of these considerations can be applied to other commercial solutions discussed elsewhere in this report.

5.4.1 Hardware

There is a significant expense due to the Endace requirement to use their own JBOD hardware, at a significantly higher cost than the equivalent hardware purchased directly. Additionally, their indexing is less efficient than PcapDB's indexing. In PcapDB, indexes less than 0.5% of the data size of captured packetsmeans that +99% of the no-markup JBOD hardware can be utilized for packet capture history. A much larger disk array is needed to store the same amount of packet capture, i.e. 30 days, under Endace. This could be nearly double when compared to PcapDB.

Note that the JBOD hardware pictured below is of the same manufacturer and is nearly identical to what PcapDB at LANL uses, albeit without the branding or vendor markup. Our former Solera based solution was also rebranded hardware from the same manufacturer.

5.4.2 Licensing

Endace does offer vendor support, unlike PcapDB at this time. PcapDB does not have any licensing costs that add additional financial barriers to using the system.

6 Moloch

(open-source, http://molo.ch/)

"Moloch is a large scale, open source, full packet capturing, indexing, and database system.
Moloch is not meant to replace Intrusion Detection Systems (IDS). Moloch augments your current security infrastructure by storing and indexing network traffic in standard PCAP format, while also providing fast indexed access. Moloch is built with an intuitive UI/UX which reduces the analysis time of suspected incidents."

Moloch is a packet capture system built on top of Elastic Search technology. While it provides deeper indexing of packets (thus requiring flow based load balancing) and some nice graph utilities, it has several disadvantages:

- Search is offloaded to entirely separate Elastic Search nodes. According the Moloch Architecture Guide, 10 Gb/s of capture with 30 days of history would require 300 additional Elastic Search nodes. While a single host can have multiple Elastic Search nodes, the system does not appear to be particularly scalable.
- Configuration and management of Moloch occurs manually through the Linux command line. For PcapDB, after the initial install, management is entirely through the Search Head (with the exception of system upgrades).
- No centralized host or user management, or segregation of multiple capture sites.

7 FireEye PX

(commercial, https://www.fireeye.com/products/enterprise-forensics/network-forensics-platform-datasheet.html, https://community.fireeye.com/docs/DOC-6168)

This commercial solution is multi-faceted. There is a network forensics platform, known as the FireEye Network Forensics Platform (PX series), and an Investigation Analysis (IA) System. It is the network forensics platform that is comparable to PcapDB.

- The PX Series captures packets and handles the query forensics, while the IA Series "extends that functionality with application contextualization, activity visualization, and campaign management."
- The Investigation Analysis System (IA Series) is an appliance that works with the PX Series to accelerate the investigative process.
- Real-time indexing. IA handles the Layer 7 (application layer) that does post-capture analysis to produce L7 data. This can be combined with another piece of hardware to have global search.
- "Web-based, drill-down GUI for search and inspection of packets, connections and sessions"
- "Ultrafast search and retrieval of target connections and packets using patent-pending indexing architecture"
- FireEye PX is a single host solution that can capture at up to 20 Gb/s.
- Does not include distributed searching or management capabilities.
- Appliance-based with hardware purchasing requirements

8 Solera/Bluecoat DeepSee

(commercial, https://www.bluecoat.com/products-and-solutions/security-analytics-and-incident-response)

DeepSee

- DeepSee treats disks as a big ring buffer and you can optionally partition the disk to hold a different amount of index versus capture (i.e. you can keep 365 days of index and 180 days of pcap).
- ThreatBLADES is a product to bolt analytics on top of the capture, storage, search.
- Functionality is added via their results display page (i.e. click on an extracted executable and upload it to VirusTotal).
- Distributed search can be enabled with multiple appliances combined.
- User reported a 40GB DeepSolution paired with Gigamon would scale to 100GB.
- Not required to purchase their hardware for the storage solution
- High cost, esp. as you scale the retention up. Users are charged for both the software license/maintenance and file system license/maintenance.

Similar cost considerations discussed in Section 5.4 are applicable here.

Solera was purchased by Bluecoat a few years ago, and this appears to be an evolution Solera's former product line. Negative past experience from another user with Solera appliances and indexes that consumed almost 50% of available capture disk were what led to the minimal indexing ideas used in PcapDB.

9 VAST

(open-source, http://vast.io and http://www.icir.org/robin/papers/nsdi16-vast.pdf)

Vast has not yet been released. Many of the same features that VAST is focused on are the very same ones that sparked PcapDB's creation: interactivity, responsiveness, and scalability with a distributed architecture. This is "to address a deep-running operational need of large-scale network monitoring and incident response: archiving and searching massive amounts of structured data." PcapDB also grew out of the operational needs of the Los Alamos National Laboratory's CSIRT team for better packet capture storage and searching. The criteria discussed in Section 2 were all driving factors in the development of

PcapDB, unlike VAST, does not introduce its own type-safe query language. Instead, Boolean logic across the indexed network tuple (IP addresses [v4 and v6], port numbers, and date time) are syntax checked, reducing error likelihood in searching in PcapDB's web interface.

10 OpenFPC

(open-source, http://www.openfpc.org/)

An open source, single host packet capture system.

- Their website offers no information on maximum capture rates.

- While each capture system is independent, they can be easily individually searched through the OpenFPC interface.
- The OpenFPC interface is command line only, and requires direct connections to each capture host.

11 Google Stenographer

(open-source, https://github.com/google/stenographer)

Google Stenographer was designed with an almost identical set of design requirements as PcapDB. As a result, its underlying capture engine is extremely similar:

- Produces tiny indexes that index up to the transport layer only.
- Provides fast, multi-threaded search.
- Provides a simple, logical search language.

11.1 Disadvantages

Unfortunately, Google Stenographer is not as fully-featured as PcapDB.

- Single host solution (10 Gb max)
- Command line interface only.

12 N2disk

(http://www.ntop.org/products/traffic-recording-replay/n2disk/)/ nbox-recorder http://www.ntop.org/products/traffic-recording-replay/nbox-recorder/)

N2Disk from Ntop is a packet capture solution built around the PFRing library, which is used for high speed capture by both Bro and PcapDB.

- Single, 10Gb/s node solution (no distributed capture or interface)
- Indexes packet-by-packet, resulting in comparatively large index files and slow packet retrieval.

13 Conclusion

PcapDB stands alone when looking at the overall field of competitors, from the cost-effective COTS hardware, to the efficient utilization of disk space that enables a longer packet history. A scalable, 100GbE-enabled system that indexes every packet and indexes flow data without complicated load-balancing requirements. The Transport Layer search and indexing approach led to patent-pending flow indexing technology, providing a specialized database system specifically optimized around providing fast flow searches.

While there are a plethora of options in network packet capture, there are very few that are able to effectively manage capture rates of more than 10 Gb/s, distributed capture and querying, and a responsive user interface. By far, the primary competitor in the market place is Endace and DeepSee; in addition to meeting the technical requirements we set out in this document, they provide technical support and a fully 'appliance like' system. In terms of cost,

however, our experience has been that the yearly maintenance charges alone outstrip the entire hardware cost of solutions like PcapDB.

Investment in cyber security research and development is a large part of what has enabled us to build the base of knowlegable workers needed to defend government resources in the rapidly evolving cyber security landscape. We believe projects like Bro, WireCap, and Farm do more than just fill temporary gaps in our capabilities. They give allow us to build the firm foundation needed to tackle the next generation of cyber challenges. PcapDB was built with loftier ambitions than simply solving the packet capture of a single lab site, but instead to provide a robust, scaleable packet capture solution to the DOE complex and beyond.