White Paper

Cisco Tetration Analytics Platform: A Dive into Blazing Fast Deep Storage

What You Will Learn

A Cisco Tetration Analytics™ appliance bundles computing, networking, and storage resources in one easy-to-consume unit. The Cisco Tetration Analytics platform uses a carefully selected combination of spinning and solid-state disk drives to provide a powerful blend of deep storage and efficient data retrieval speed.

This document explains:

- The technical specifications of the storage devices that are in use
- How the storage in the system is allocated to different components
- How different operating environments consume this storage
- What affects the total predicted data retention of the platform
- Some sample data ingestion and retention figures

Introduction

How much data does your data center retain? To find out, you need to record the entire amount of data that you generate: an easy task with the Cisco Tetration Analytics platform, and a nearly impossible without it.

If you are an existing Cisco Tetration Analytics customer, you can use the guidance in this document to help with your ongoing operations. If you are a prospective customer, the included examples, gleaned from production environments, should give you an idea of the potential amount of data that your Cisco Tetration Analytics appliance may retain.

Note that everything discussed in this document is relevant to Cisco Tetration Analytics OS (TetOS) Release 1.102.0. However, the Cisco Tetration Analytics engineers are constantly striving to add more functions to the same hardware for you, introducing new features that will almost certainly affect the validity of this document.

Storage Devices in Use

The Cisco Tetration Analytics platform uses two disk types and three node types.

Disk Types

Cisco Tetration Analytics uses two types of drives:

- Spinning hard-disk drives (HDDs)
- Solid-state disks (SSDs)

HDDs and SSDs differ in a number of ways. Here are the high-level reasons that the Cisco Tetration Analytics platform uses both:

- Hard-disk drives
  - Used as the primary deep storage devices

© 2017 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

    - Support high capacity
    - Are cheaper per megabyte, gigabyte, and terabyte
    - Provide medium-speed data retrieval
- Sold-state disks
    - Used as the primary caching devices
    - Support medium-level capacity
    - Are more expensive per megabyte, gigabyte, or terabyte
    - Provide extremely fast data retrieval

Node Types

The Cisco Tetration Analytics platform uses three node types:

- Compute nodes
- Serving nodes
- Base nodes

The different node types are invisible to the user when the user is installing, operating, and maintaining the Cisco Tetration Analytics platform, with the orchestration managed entirely by the platform. This section discusses the differences among the nodes.

Note that all nodes have eight disks (the maximum capacity of a Cisco UCS® C-Series Server).

- Computing nodes: Computing nodes exclusively use HDDs, because the ingestion of raw telemetry data and data batching, stitching, annotating, and compressing is performed in memory, with few disk read or write operations, so optimization for space is important. Computing nodes are configured with one 1.2-TB disk, which is used as the primary boot drive. The storage partition uses seven 1.8-TB 4-KB-sector drives, for a total of 12.6 TB of raw space. With 16 computing nodes, the total capacity is 201.6 TB. This space is used as the deep storage location. The allocation of this space is discussed later in this document.
- Serving nodes: Serving nodes exclusively use SSDs. Serving nodes need access to data at extremely fast speeds to be able to retrieve millions of flows from among billions in the data pool in less than a second. The entire server is provisioned with eight 400-GB SSDs, for a total of 3.2 TB of raw space per server. With 8 serving nodes, the total capacity is 25.6 TB.
- Base nodes: Part of the extended data lake, base nodes have eight 1.2-TB disks, for a total of 9.6 TB of raw storage. With 12 base nodes, the total capacity is 115.2 TB.

The platform thus has a total of 316.8 TB of HDD space, referred to as the data lake, plus 25.6 TB of SSD space in the serving layer.

How Storage Devices Are Allocated

The allocation of storage devices takes into account the data lake and the serving layer (Figure 1).

Data Lake

Raw space is only one factor to consider when determining capacity. Any highly resilient system must include layers of redundancy on top of the physical hardware to provide system integrity. Hardware RAID controllers and software solutions are commonly used, with trade-offs between the different approaches.

© 2017 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

A newer solution, Apache Hadoop Distributed File System (HDFS), has quickly become the technology of choice for many big data platforms. Pioneered in early work by Google, with its Google File System, HDFS provides an extremely scalable, distributed, and resilient file system layer on top of raw disks, without the need for an expensive (and often error-prone) hardware RAID controller.

The Cisco Tetration Analytics platform uses HDFS for all HDDs, which are pooled in an enormous data lake that is redundantly distributed across computing nodes and base nodes.

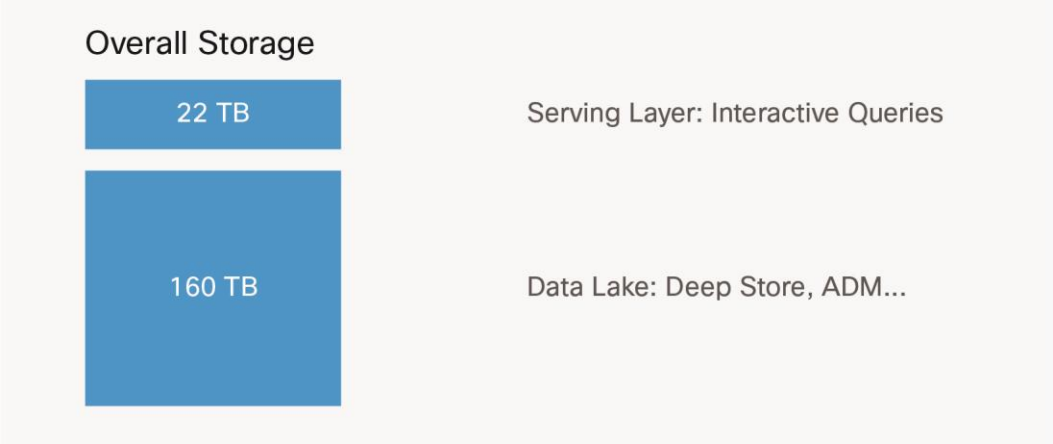
Adding any form of redundancy consumes a percentage of the raw space available. HDFS makes three replicas of the same block of data, storing each replica on a unique node. In the Cisco Tetration Analytics platform, the block size uses the HDFS default of 128 MB. This redundancy allows the loss of two computing or base nodes without any data loss.

After taking into account the replication overhead, the platform has 160 TB of usable, distributed, redundant storage in its primary data lake.

Serving Layer

The SSDs in the Cisco Tetration Analytics platform do not operate in a redundant fashion, because they are in place to provide a temporal hot cache layer to the same data that is redundantly stored in the data lake. Essentially, nearly the entire amount of raw SSD storage space is available to the serving layer. If a disk fails, it is simply replaced, and the necessary data is drawn from the data lake.

Figure 1. Overall Storage Space Presented to the System



How Flows Are Ingested

You now know the amount of storage available to the system. Now consider how that storage is used as data flows in from sensors.

© 2017 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

Flow Records

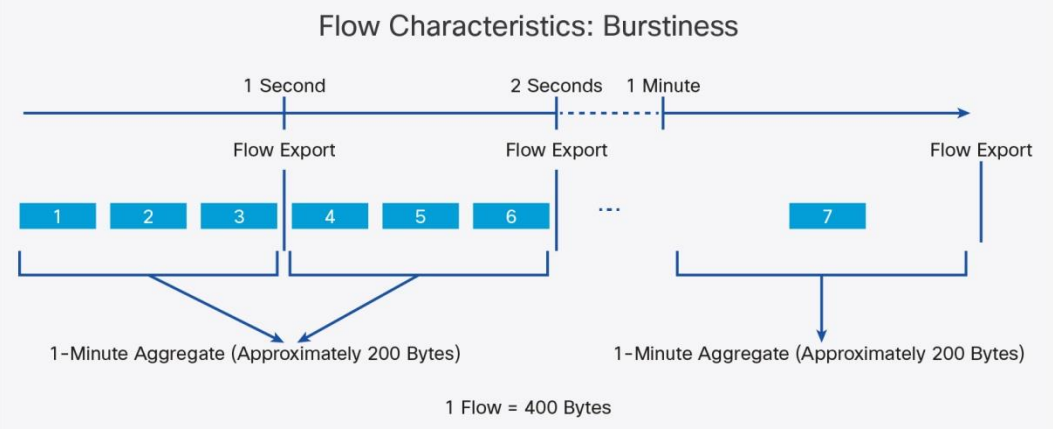
To understand how retention is calculated in the Cisco Tetration Analytics platform, you need to understand the raw unit of consumption: the flow record.

As discussed in detail in other documents, hardware and software sensors report detailed metadata about every single packet traversing the sensor's domain. For the purposes of this discussion, you need to know the number of bytes that each flow record consumes after it reaches the collector and has been stitched, annotated, compressed, and stored.

Whereas sensors stream telemetry data in one-second batches to the cluster, flows are stored in minute-long batches at the collector. Each direction of a flow is allocated 200 bytes per minute, regardless of whether the flow has been observed once or 60 times in that one-minute interval. If the flow is bidirectional (for example, with TCP), that figure is doubled, to 400 bytes.

Figure 2 shows that a flow will consume 200 bytes per minute, regardless of whether multiple packets have been observed or only one packet. This pattern favors bursty flows that have a duration of less than a minute (which are most common). Flows that have small bursts stretched out over longer than one-minute periods can have a negative effect on storage.

Figure 2. Flow Bursting Affects Storage Requirements



© 2017 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

Figure 3 shows that many short-lived flows can have an adverse effect on storage requirements. Each flow has a static overhead of 200 bytes no matter how many packets are monitored.

Figure 3. Many Short-Lived Flows Can Contribute to Increased Storage Requirements

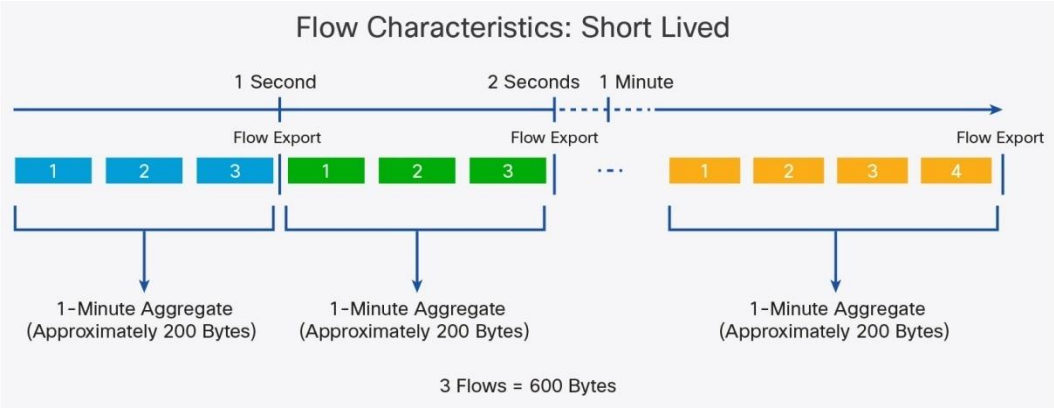
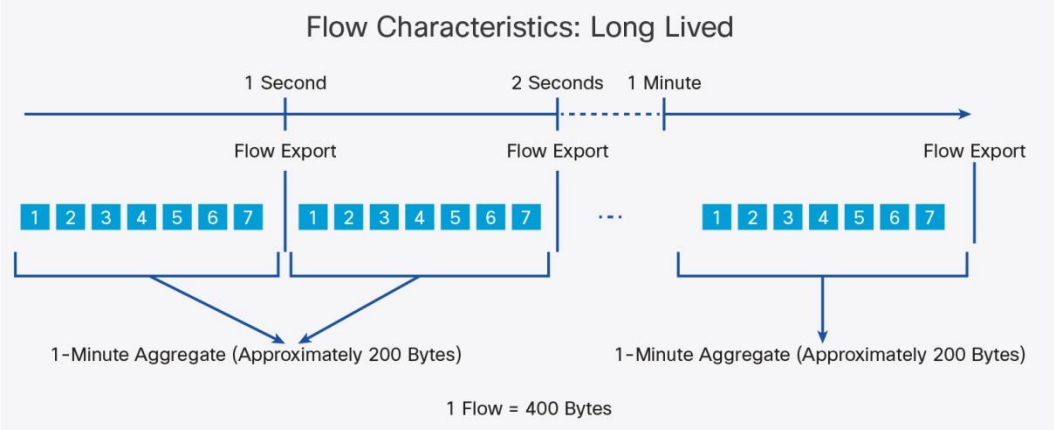


Figure 4 shows that the increasing number of packets in a one-minute interval does not increase the storage requirements. The Cisco Tetration Analytics platform thus optimizes storage for long-lived flows.

Figure 4. Long-Lived Flows Are Stored Efficiently Because They Benefit from the Fixed 200-Byte Overhead



Flow Annotation

A powerful feature of the Cisco Tetration Analytics platform is the capability to tag flows with contextual annotations. The sensors themselves can be configured to send annotation data, including process strings, user IDs, buffer details, drop reasons, and more. Depending on the annotations that you choose to enable, annotations can substantially increase the amount of storage space required.

Annotations are stored at one-minute intervals.

Process Annotation

Note that process strings can vary greatly depending on the type of application you are running, and they are difficult to compress because the entropy is high.

© 2017 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

Consider an environment hosting multiple database applications using the MySQL server. The process string might look like this:

```
/usr/sbin/mysqld --basedir=/usr --datadir=/var/lib/mysql --plugin-
dir=/usr/lib/mysql/plugin --user=mysql --wsrep-new-cluster --log-
error=/var/lib/mysql/percona-2.err --pid-file=/var/lib/mysql/percona-2.pid --
wsrep_start_position=a243ecb6-ff07-11e5-958e-5719d8e72c74:5288
```

Total = 270 bytes

Alternatively, a Java application (ZooKeeper, in this case) often has much longer process strings:

```
/usr/lib/jvm/java-7-openjdk-amd64/jre/bin/java -
Dzookeeper.log.dir=/opt/mapr/zookeeper/zookeeper-3.4.5/logs -
Dzookeeper.root.logger=INFO, ROLLINGFILE -
XX:ErrorFile=/opt/mapr/zookeeper/zookeeper-3.4.5/logs/hs_err_pid%p.log -cp
/opt/mapr/zookeeper/zookeeper-
3.4.5/bin/../build/classes:/opt/mapr/zookeeper/zookeeper-
3.4.5/bin/../build/lib/*.jar:/opt/mapr/zookeeper/zookeeper-
3.4.5/bin/../lib/slf4j-log4j12-1.6.1.jar:/opt/mapr/zookeeper/zookeeper-
3.4.5/bin/../lib/slf4j-api-1.6.1.jar:/opt/mapr/zookeeper/zookeeper-
3.4.5/bin/../lib/netty-3.2.2.Final.jar:/opt/mapr/zookeeper/zookeeper-
3.4.5/bin/../lib/log4j-1.2.15.jar:/opt/mapr/zookeeper/zookeeper-
3.4.5/bin/../lib/jline-0.9.94.jar:/opt/mapr/zookeeper/zookeeper-
3.4.5/bin/../zookeeper-3.4.5-mapr-1503.jar:/opt/mapr/zookeeper/zookeeper-
3.4.5/bin/../src/java/lib/*.jar:/opt/mapr/zookeeper/zookeeper-
3.4.5/conf::/opt/mapr/lib/maprfs-5.1.0-mapr.jar:/opt/mapr/lib/protobuf-java-
2.5.0.jar:::/opt/mapr/lib/json-20080701.jar:/opt/mapr/lib/flexjson-
2.1.jar:/opt/mapr/lib/commons-codec-1.5.jar -Djava.library.path=/opt/mapr/lib -
Dzookeeper.sasl.serverconfig=Server_simple -
Djava.security.auth.login.config=/opt/mapr/conf/mapr.login.conf -
Dzookeeper.sasl.clientconfig=Client_simple -
Dzookeeper.saslprovider=com.mapr.security.simplesasl.SimpleSaslProvider -
Dcom.sun.management.jmxremote -Dcom.sun.management.jmxremote.local.only=false
org.apache.zookeeper.server.quorum.QuorumPeerMain /opt/mapr/zookeeper/zookeeper-
3.4.5/conf/zoo.cfg
```

Total = 1466 bytes

If maximizing data retention is an important deployment requirement, you should turn on process annotation sparingly.

Factors That Contribute to Storage Retention

The Cisco Tetration Analytics platform stores three main types of data:

- Deep storage data
- Flow search data
- Computed results

The three data types have different storage requirements.

Deep Storage Data

Every flow record that has been compiled into batches of 200 bytes (plus any annotations that have been enabled) per minute is committed to deep storage. Deep storage resides in the primary HDFS data lake and allows you to confidently keep a record of your flows. Currently not directly accessible by the user, it is in place to allow expansion of software features that may benefit from access to raw flow data at slower retrieval speeds.

© 2017 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

Flow Search Data

Flow search data is effectively a mirror of the latest deep storage flows, and it has a maximum capacity bound by the total amount of serving-layer storage: 22 TB.

In the Cisco Tetration Analytics web user interface, flow search storage defines the total time range of flows that can be explored using the powerful filtering tools at speeds of less than a second. It does not represent the entire pool of flows that the platform has recorded.

Computed Results

All the flow information that the Cisco Tetration Analytics platform captures is put to work through the computing nodes to generate actionable information and visibility into the structure and operation of the data center network.

The main consumers of computed results are:

- Application insight: Machine-learning-generated maps of the applications that run on your network
- Policy compliance: Near-real-time visibility into the health of traffic that is traversing your network, with detailed application-level tracking of packet drops and potentially malicious traffic flows

Because the computed results do not contain raw flow data, the amount of time they can be retained is considerably longer because the disk space they occupy is negligible. One example of a common computed result is the output from application dependency maps (ADMs), which can be stored for years.

Sample Data Retention Prediction

Estimating flows per second in a complex environment is practically impossible without a detailed understanding of the network. Nevertheless, Table 1 provides some sample statistics from large-scale data centers.

Table 1. Examples of Data Retention Statistics: Sampled at Peak Time of Day

Sensor Count	Searchable Observation Count	Flows per Second	Flow Search	Deep Storage (Estimated)	Description
1365	37, 629, 964, 895	91,000	18.36 weeks	2.46 years	Government hosting platform (production)
1929	61, 388, 413, 808	154,000	4.7 weeks*	7.58 months	Enterprise hosting platform (staging)
3966	82, 426, 102, 512	378,000	2.71 weeks*	4.37 months	Enterprise hosting platform (production)
72	755, 523, 474	7,000	24.26 weeks*	3.26 years	Staging lab
36	24, 472, 721	5,600	32.56 weeks*	4.35 years	Development lab
976	4, 207, 237, 935	65,800	5.12 weeks*	8.25 months	Healthcare internal applications
2792	11, 1152, 358, 693	67,200	9.24 weeks*	1.23 years	Banking applications

* Process annotation is active.

Note: Computed results are so negligible in size that they are not tracked in this table.

Conclusion

The Cisco Tetration Analytics platform has a purpose-designed storage setup that provides excellent storage capacity and access speeds, paired with optimized software that efficiently caches and compresses data to achieve generous retention periods and blazing fast flow retrieval.

© 2017 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.

For More Information

For additional information, see:

- Cisco Tetration Analytics at-a-glance
- Cisco Tetration Analytics data sheet

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: www.cisco.com/go/trademarks. Third party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)

Printed in USA C11-738477-00 02/17

© 2017 Cisco and/or its affiliates. All rights reserved. This document is Cisco Public Information.