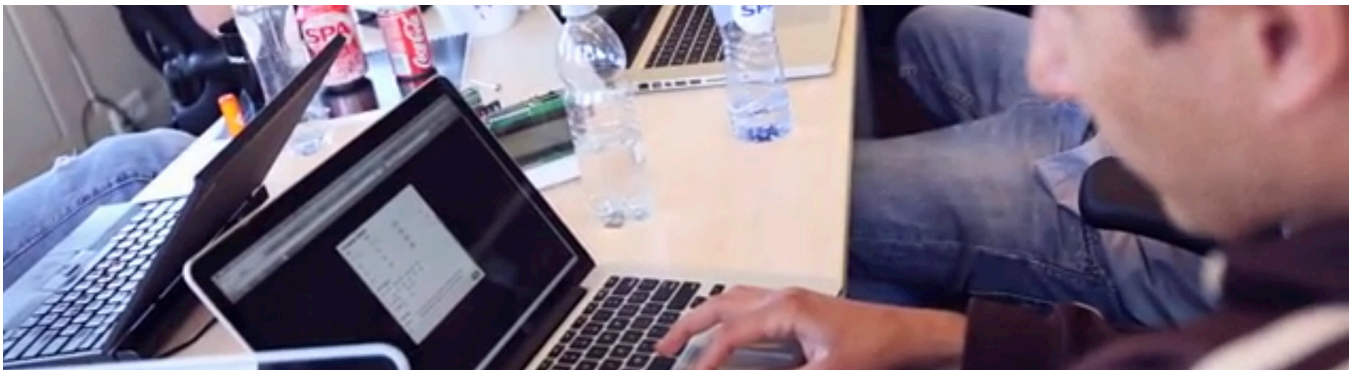
Elasticsearch on Cisco Unified Computing System:

Optimizing your UCS infrastructure for Elasticsearch's analytics software stack



HIGHLIGHTS

Real-Time Results
Elasticsearch on Cisco UCS enables a deeper and richer understanding of data converged from multiple sources within the enterprise in real time.

Integrates with Cisco UCS CPA for Big Data
Elasticsearch on Cisco UCS can coexist with other Cisco UCS CPA for Big Data solutions in the same system, sharing high-bandwidth connectivity to speed the use of analytic results.

Ease of Deployment
Cisco UCS Manager automates deployment and scaling, reducing the risk of configuration errors that can cause downtime.

Architectural Scalability
The solution is designed to grow to maximum scale without adding complex layers of switching infrastructure.

Choice of Configurations
Whether you want to optimize for high IOPs and low capacity or low outputs and high capacity, the Cisco UCS Elasticsearch bundle has you covered.

Elasticsearch on Cisco UCS® combines hardware and software to deliver real-time actionable insights from your data

One challenge that organizations face is how quickly they can glean actionable insight from their large volumes of data. Having a scalable and performant compute, storage and networking infrastructure to support Big Data projects is only part of the equation – for organizations to get value out of their data they need to be able to quickly search and analyze it. Elasticsearch bridges that gap with a real-time search and analytics engine. Deployed as part of a comprehensive data center architecture, the Elasticsearch on Cisco Unified Computer System (Cisco UCS) solution delivers real-time business insights over a powerful and flexible infrastructure.

Elasticsearch Solutions

Elasticsearch is a company dedicated to leveraging the collective power of three massively popular open source projects: Elasticsearch, Logstash, and Kibana. Each tool can be used in isolation for search, logging, and data visualization respectively. However, when all three are combined, they create an end-to-end solution known as the Elasticsearch ELK stack, a powerful analytics platform that can help you glean actionable insights in real time from almost any type of structured and unstructured data source. The ELK stack can be configured as a stand-alone application, it can serve as a platform for custom web applications, or it can be integrated with your organization's existing data applications.

Elasticsearch

Elasticsearch is a powerful open source, distributed, real-time search and analytics engine built on top of Apache Lucene. Architected from the

ground up for use in distributed environments where reliability and scalability are must-haves, Elasticsearch gives you the ability to easily move beyond simple, full-text search. It enables real-time analytics and full text search over stored data. With its robust set of APIs, query DSLs, and clients for the most popular programming languages, Elasticsearch is gaining wide adoption as an open platform for all of your organization's log data.

Logstash

Logstash aggregates log data as well as time series data, csv, and over 40 data inputs from any system into a single place for additional transformation and processing. Using Elasticsearch as its backend data store, Logstash creates a powerful pipeline for storing, querying, and analyzing log files.

Kibana

Kibana is Elasticsearch's data visualization engine, allowing users to natively interact with all of their data in Elasticsearch via pre-built and custom dashboards.

Elasticsearch for Hadoop

Elasticsearch for Apache Hadoop, called es-Hadoop, is an adapter that facilitates Hadoop jobs to interact with Elasticsearch through a bi-directional flow of data. It enables real-time search for Hadoop HDFS data. The adapter currently supports Map/Reduce, Cascading, Spark, Pig, Hive, and Storm. This might be a good solution if you want to search on top of your long-term data store.

Cisco UCS

The Cisco UCS is a next-generation data center platform that unites compute, network, and storage access into a cohesive system designed to meet a variety of scale-out application demands with transparent data and management integration capabilities.

Scalability

Cisco UCS brings simplified management, greater deployment flexibility, and easier scalability to the scaled-out nature of many of today's applications. Every aspect of server personality, configuration, and connectivity is set on demand, through the Cisco UCS Manager. Using Cisco's powerful service profiles, you can configure Elasticsearch clusters rapidly and automatically without the risk of configuration drift, which can lead to errors that cause downtime. Unified management in Cisco UCS enables greater agility and more rapid deployment.

Unified Fabric

This solution uses important Cisco UCS innovations of unified fabric, unified I/O, and unified management for all connected devices. Its low-latency, lossless 10-Gbps unified fabric is fully redundant and, through its active-active configuration, delivers higher performance compared to other vendor solutions.

Reduced TCO and Improved Staff Efficiency

This simplified, intelligent infrastructure reduces your total cost of ownership (TCO) with fewer management endpoints, switches, adapters, cables, power, and cooling components. Through the use of embedded management and automation, your staff's ability to quickly and efficiently deploy and troubleshoot Cisco UCS will lower operating expenses and allow staff to focus on strategic business initiatives rather than infrastructure maintenance.

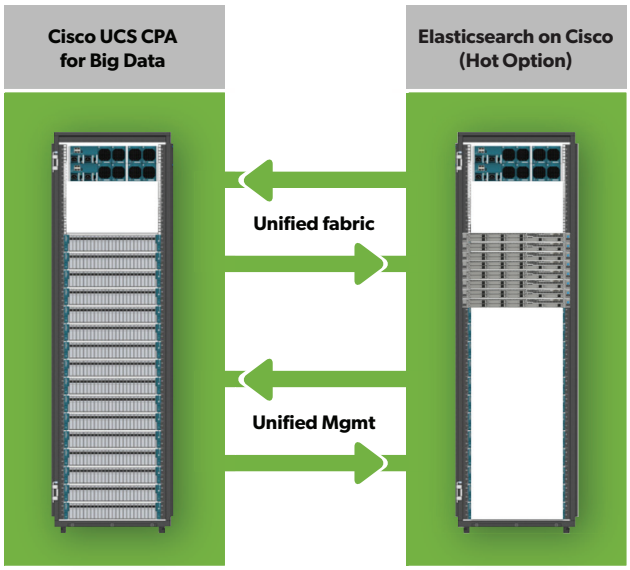


Figure 1

Elasticsearch on Cisco UCS

Elasticsearch on Cisco UCS delivers high performance with exceptional scalability. Cisco UCS unified fabric architecture provides fully redundant, highly scalable, lossless 10-Gbps unified fabric connectivity for Elasticsearch's query, index, and replication data traffic. The reference architecture can easily scale to support a large number of nodes to meet additional demands.

This solution can connect across the same management plane to other Cisco UCS deployments such as those deployed with Cisco UCS Common Platform Architecture (CPA) for Big Data, thereby radically simplifying datacenter management and connectivity.

The solution is built using the following components:

Cisco UCS 6200 Series Fabric Interconnects

Cisco's UCS 6200 Series Fabric Interconnects provide high-bandwidth, low-latency connectivity for servers, with Cisco UCS Manager providing integrated, unified management for all connected devices. Deployed in redundant pairs, Cisco fabric interconnects offer the full active-active redundancy, performance, and exceptional scalability needed to support the large number of nodes that are typical in clusters serving Big Data applications. Cisco UCS Manager enables rapid and consistent server configuration using service profiles, automating ongoing system maintenance activities such as firmware updates across the entire cluster as a single operation. Cisco UCS Manager also offers advanced monitoring with options to raise alarms and send notifications about the health of the entire cluster.

Cisco UCS C220 M4 Rack Servers

Cisco UCS C220 M4 Rack Servers are designed for performance and density over a wide range of business workloads in a 1-rack unit (1RU) form factor. Cisco UCS C220 M4 servers are powered by dual Intel Xeon E5-2600 v3 series CPUs, and they support up to 768 GB of main memory. These servers support four or eight SAS/SATA/SSD drives as well as Cisco UCS virtual interface cards (VICs) optimized for high-bandwidth and low-latency cluster connectivity, with support for up to 256 virtual devices.

Cisco UCS C240 M4 Rack Servers

Cisco UCS C240 M4 Rack Servers are enterprise-class servers that deliver an outstanding combination of performance, flexibility, and efficiency for storage. The Cisco UCS C240M4 servers are 2-socket, 2-rack-unit (2RU) servers based on Intel Xeon E5-2600 v3 series processors supporting up to 768 GB of DDR4 main memory. These servers support up to 26 SFF SAS/SATA/SSD drives or 12 LFF SAS/SATA drives plus 2 SFF SSD drives. Their expandability and exceptional performance makes them an ideal fit for big data analytics, virtualization, graphics-rich and bare-metal applications.

Configuration Options

Systems that are optimized for performing many fast queries look different than systems optimized for just fast indexing. Figure 1 below shows how your use case impacts how you might want to configure your UCS hardware. The UCS-Elasticsearch bundles (Table-1) come in two flavors - hot and warm - and can be used separately or combined together to meet your deployment needs.

The UCS-Elasticsearch bundles (Table-1) come in two flavors, each according to these common use cases:

Hot Option

Fast Indexing, Many Queries, Low Capacity Use Cases -- This option is for when you want to run fast indexing on today's data and only need to retain it in storage for a short period of time (~2 days). The system is optimized for high IOPS and low storage capacity.

Warm Option

Many Queries, Medium Capacity Use Cases -- This option is for when you need to retain data for ~3-8+ days and require medium-to-less frequent querying. This option balances I/O bandwidth with medium storage capacity.

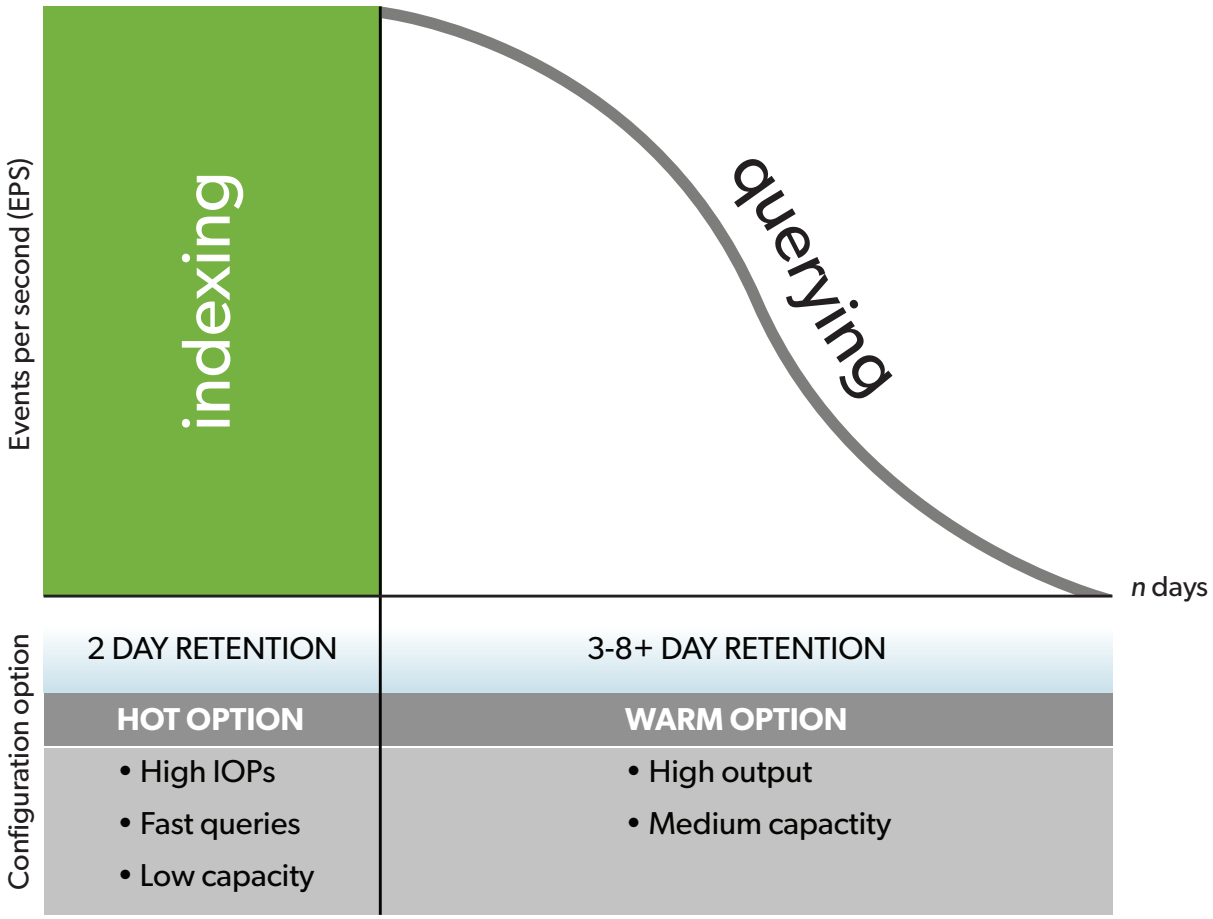


Figure 2

Rack Solution By Use Case	Hot Option	Warm Option
Network	2 Cisco UCS 6248UP 48-Port Fabric Interconnects	2 Cisco UCS 6248UP 48-Port Fabric Interconnects
Management	Cisco UCS Manager	Cisco UCS Manager
Servers	**8 Cisco UCS C220 M4 Rack Servers, each with:** • 2 Intel Xeon processors E5-2620 v3 at 2.4 GHz • 256 GB of memory • Cisco 12G-2GB RAID Controller • 8 1.2TB 10K SFF SAS drives (76 TB total)	**8 Cisco UCS C240 M4 Rack Servers, each with:** • 2 Intel Xeon processors E5-2620 v3 at 2.4 GHz • 256 GB of memory • Cisco 12G-2GB RAID Controller • 24 1.2TB 10K SFF SAS drives (230 TB total) • 2 120GB SSD drives

Table 1: Elasticsearch on Cisco UCS Solution

Conclusion

Big Data technology has become a valuable and compelling asset for organizations of all sizes. Harnessing the value of Big Data analytics in real time from a stable, manageable, and scalable infrastructure has been a long-standing challenge. However, the Elasticsearch on Cisco UCS solution rises above this challenge by delivering a real-time data analytics platform that can be rapidly deployed, scaled on demand, and is easy to use. It also reduces the total cost of ownership by requiring fewer infrastructure components and reducing operating expenses associated with staff time. Organizations can rely on this solution to provide real-time data intelligence and make critical business decisions.

For More Information:

www.elasticsearch.com

www.cisco.com/go/ucs

For more information about the Elasticsearch support, please visit:
http://www.elasticsearch.com/support

For more information about Cisco UCS big data solutions, please visit:
http://www.cisco.com/go/bigdata