Cisco HyperFlex Systems for Splunk Enterprise

Implement a private cloud-in-a-box solution for Splunk workloads.

Highlights

Cisco HyperFlex™ systems: This flexible, agile, efficient, and scalable next-generation hyperconverged platform is powered by the Cisco Unified Computing System™ (Cisco UCS®). It provides customers with unified fabric, unified management, and advanced monitoring capabilities. It also provides consistent and rapid deployment for out-of-the-box performance using service profiles.

Cisco HyperFlex HX Data Platform: This high-performance, flash-optimized distributed file system delivers a wide range of enterprise-class data management and optimization services without compromising data management, storage efficiency, or latency.

Operational intelligence with Splunk Enterprise: Splunk software monitors and analyzes data from any source, including computing, storage, and networking activities; service health; firewall access; and customer click streams and call records. It turns machine-generated data into business insight.

Intelligent private cloud-in-a-box solution: For new applications and operational models, you need a solution that helps you scale capabilities as you need them. This solution simplifies scaling at precise levels for a scale-out design, enabling you to spend money only when you need to do so.

Powerful search, analysis, and visualization capabilities with Splunk Enterprise: Gain an easy, fast, and secure way to analyze massive streams of data generated by IT systems, security devices, and technical infrastructure.

Validated solutions: This solution is built to complement the widely deployed converged infrastructure solution Splunk on Cisco UCS Integrated Infrastructure for Big Data.

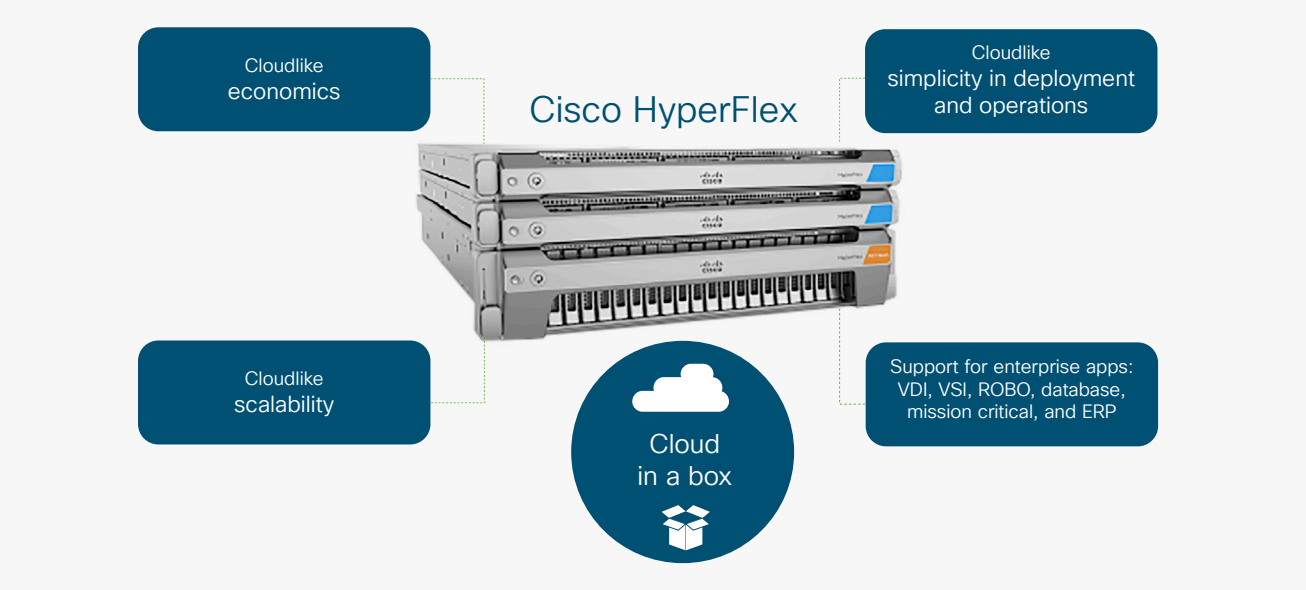
Cisco HyperFlex systems: Fast, flexible hyperconverged systems

Engineered on the Cisco Unified Computing System™ (Cisco UCS®), Cisco HyperFlex™ systems unlock the full potential of hyperconverged solutions to deliver the agility, scalability, security, and lifecycle management capabilities you need for operational simplicity. Cisco HyperFlex systems support the pay-as-you-grow economics of the cloud with the benefits of on-premises infrastructure.

Cisco HyperFlex systems combine software-defined computing in the form of Cisco UCS servers, software-defined storage with powerful Cisco HyperFlex HX Data Platform software, and Software-Defined Networking (SDN) with Cisco® unified fabric.

In Cisco HyperFlex Systems, the data platform spans three or more Cisco HyperFlex HX-Series nodes to create a highly available cluster. Each node includes an HX Data Platform controller that implements the scale out and distributed file system using internal flash-based Solid State Disks (SSDs) or a combination of flash-based SSDs and high-capacity Hard Disk Drives (HDDs) to store data. The controllers communicate with each other over 10 or 40 Gigabit Ethernet to present a single pool of storage that spans the nodes in the cluster (Figure 2). Nodes access data through a data layer using file, block, object, and API plug-ins. As nodes are added, the cluster scales linearly to deliver computing, storage capacity, and I/O performance.

Figure 1. Cisco HyperFlex systems: Cloud in a box

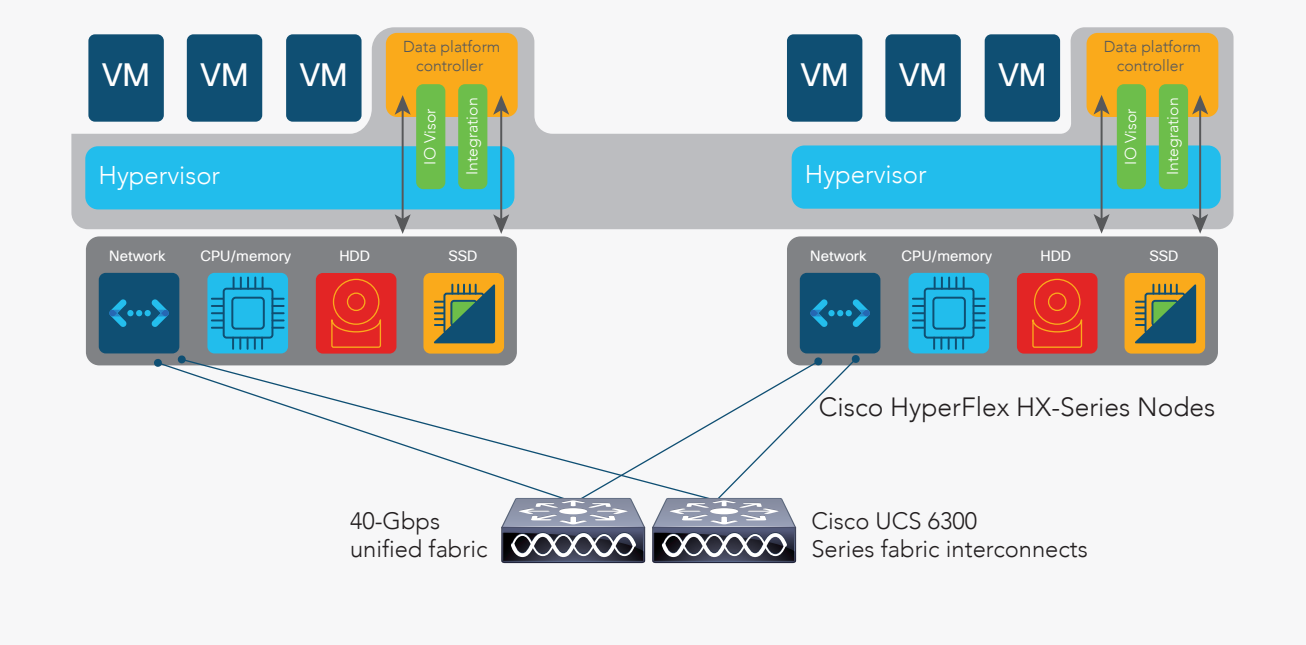


Cisco HyperFlex HX Data platform

The unique data demands imposed by applications on virtual machines have resulted in many storage silos. A foundation of Cisco HyperFlex systems, the HX Data Platform (Figure 2) is a purpose-built, high-performance, log-structured, scale-out file system that is designed for hyperconverged environments. The data platform's innovations redefine scale-out and distributed storage technology, going beyond the boundaries of first-generation hyperconverged infrastructure and offering a wide range of enterprise-class data management services.

The Cisco HyperFlex HX Data Platform has been demonstrated to be the industry-leading platform for hyperconvergence. It is a viable platform for applications such as virtual desktop infrastructure (VDI), Virtual Server Infrastructure (VSI), and databases. It can be used in production, test, and development environments in which multiple application instances can coexist and be managed from a single management pane. For more information, refer to the Cisco HyperFlex systems white paper.

Figure 2. Cisco HyperFlex HX Data platform architecture

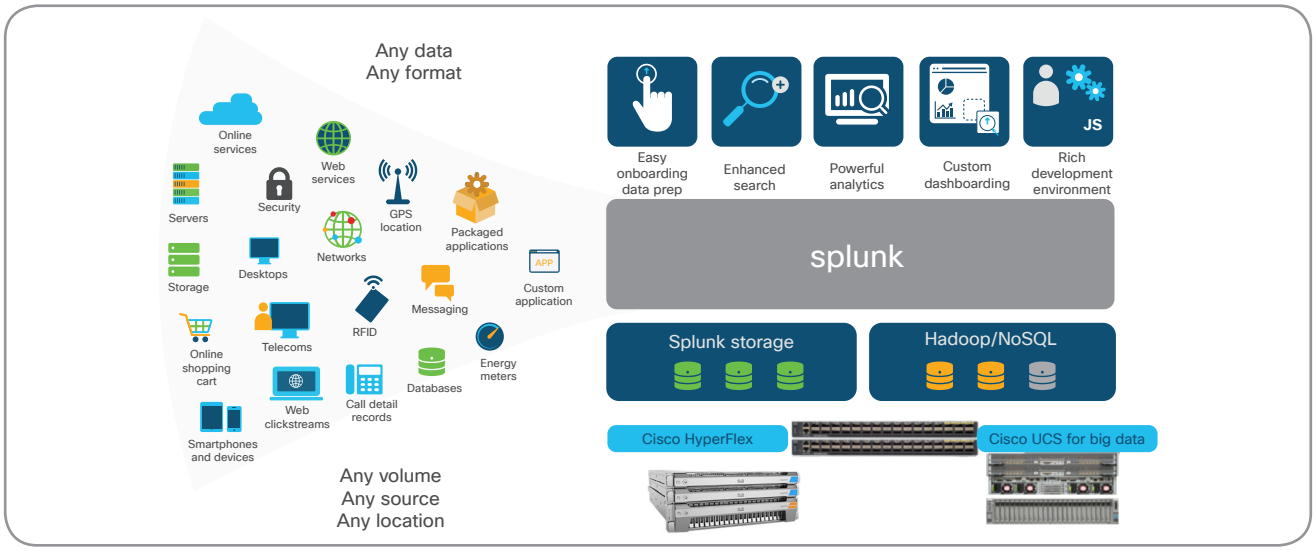


Machine data: The pulse of your digital infrastructure

The interconnected systems powering the on-premises digital data infrastructure constantly provide volumes of information about the status and details of a system's health, operations, results, and intrusions. The application of operational analytics to this machine-generated data is essential to keep the digital machines of modern enterprises operating at high efficiency. Visibility into the operation of various IT systems at multiple levels—network switches, routers, firewalls, Internet of Things (IoT) devices, virtual machines, containers, applications, and clouds—is of paramount importance for modern enterprises.

The data generated by various systems is disparate in nature, so the traditional approach was to use custom-built tools to analyze each specific system or group of systems. This approach to analysis results in data silos, which require much co-ordination and manual correlation to derive insights for decision makers. The repetitive and manual nature of this process limits agility, scalability, and focus.

Figure 3. Splunk Enterprise on Cisco Hyperflex



Splunk Enterprise for IT operational intelligence and security analytics

Splunk Enterprise is a leading platform for IT operational analytics. It can monitor a variety of machine data and help with the monitoring, analysis, and correlation of the operational data from the entire digital infrastructure. It thus can offer competitive and productive insights to the decision makers and business leaders. It offers custom insightful dashboards for key business decision makers, enabling organizations to achieve value fast. It also empowers engineers to drill down and perform additional targeted automated and impromptu searches for correlation across multiple tiers of machine data.

Reference architecture for Splunk Enterprise

The proven converged infrastructure solution Cisco UCS Integrated Infrastructure for Splunk Enterprise is recommended for large-scale, highly available distributed Splunk deployments. The hyperconverged all-flash Cisco HyperFlex configuration summarized in Table 1 has been developed as a complementary offering for those customers who want their Splunk workloads to be virtualized and who are interested in a cloud like-architecture on their premises. This solution has been tested and validated with Splunk workloads with various virtual machine specifications.

Table 1. Configuration details

Server nodes	4 x Cisco HyperFlex 240c M5 All Flash Nodes, each with: • 1 x Cisco UCS Virtual Interface Card (VIC) 1387 modular LAN on motherboard (mLOM) • 2 x Intel Xeon processor 6148 CPUs (40 physical cores) • 384 GB of DDR4 RAM • 1 x 1.6TB 2.5in U.2 NVMe High Performance High Endurance caching drive • 23 x 3.8 TB SSD Enterprise Value (SSD-EV) all-flash (capacity) SSD drives
Available storage	107 TB
Replication Factor	3 (tolerates 2 node failures)
Deduplication	Enabled
Compression	Enabled
Connectivity	2 x Cisco UCS 6332 Fabric Interconnects with 32 x 40 Gigabit Ethernet ports

Cisco HyperFlex HX Data Platform solutions are built on the Cisco UCS platform. They offer faster deployment and greater flexibility and efficiency at a competitive price, while lowering risk for the customer. This approach reduces or eliminates the need for planning and configuration decisions, while allowing the customization needed to address customer workload needs. The platform and management model adopted represents an extension of established Cisco UCS data center strategy, in which familiar components are managed in a consistent manner through a policy-based framework with Cisco UCS Manager.

Splunk Enterprise on Cisco HyperFlex all-flash systems

Splunk deployments typically start small and expand rapidly to address additional use cases. An infrastructure that can scale quickly to meet this need is thus of paramount importance. Cisco HyperFlex all-flash systems can be used to host a number of virtual machines that can support any of the following Splunk software roles:

- **Splunk indexers:** An indexer is an instance of Splunk Enterprise that parses, transforms, indexes, and stores data in a distributed manner. It searches the indexed data in response to search requests from search heads. It can also allow data input in the absence of forwarders.
- **Splunk search heads:** A search head is a Splunk instance specifically configured to perform only search operations in a distributed Splunk configuration. It sends search requests to the appropriate set of indexers and merges the results. Multiple search heads can be configured in a cluster for high availability.
- **Splunk heavy forwarders:** A heavy forwarder is an instance of Splunk Enterprise configured specifically to gather data from multiple sources and forward them to Splunk indexers.
- **Multiple Splunk standalone Splunk instances:** The standalone instances are Splunk indexers that are meant to index less than 5 GB of data per day. They perform data indexing and searching and offer complete dashboard capabilities.

The HX Data Platform allows computing and storage resources to be managed at very precise levels. Table 2 lists the recommended virtual machine configurations for general-purpose IT operational analytics, enterprise security, and IT service intelligence use cases.

Table 2. Virtual machine configurations for general-purpose use cases

Minimum Specifications of virtual machine[1] building block	Splunk Enterprise for IT Operations Analytics (ITOA)
CPU: 24 virtual CPUs (vCPUs)[2]	Index capacity per day: Up to 250 GB[3]
Memory: 64 or 96 GB	
Storage: According to the need	

Notes:

[1] The virtual machine can serve as an indexer, search head, or heavy forwarder. Please see guidelines virtualizing Splunk deployments.

[2] For premium solutions such as Splunk Enterprise Security (ES) and Splunk IT Services Intelligence (ITSI), plan to increase the number of vCPUs and memory as needed.

[3] The suggested maximum indexing capacities per indexer node are up to 250 GB per day for ITOA, and up to 100 GB per day for ITSI and ES solutions.

- Up to 2 TB daily indexing when used for ITOA use cases
- Up to 800 GB daily indexing when the system is used for ES or ITSI use cases

Figure 4. Configured solution

This solution provides the following benefits to your Splunk workloads:

- **Hyperconverged platform:** Built on Cisco UCS, the Cisco HyperFlex system combines networking, storage, and virtualization resources in a single converged platform.
  - Enterprise-class storage and maintenance features are included, such as inline deduplication, compression, snapshots, and single-button nondisruptive rolling upgrades.
  - Intelligent data distribution promotes the balanced growth of Splunk indexes across the cluster, thereby helping ensure the operation of all indexers at peak performance.
  - Native replication helps reduce the number of replications needed at the Splunk layer.
  - Hyperconverged and converged (Cisco UCS Integrated Infrastructure) Splunk Enterprise deployments can coexist in the same Cisco UCS domain, thus providing a single management plane for virtual, physical, scale-up, and scale-out needs.
  - A single administrator can manage all aspects of Cisco UCS and the Cisco HyperFlex system through Cisco UCS Manager and VMware vCenter Web Client, making tasks much easier and faster to complete.
  - Resources can be carved out logically into multiple pools, thereby enabling seamless multitenancy.
- **Rapid deployment:** The programmability and ease of use of Cisco UCS Manager allow Cisco HyperFlex systems to be deployed quickly and consistently.
  - Deployment and scaling is easy, with deployment taking an hour or less, thus accelerating time to value.
  - Adaptive infrastructure offers with pay-as-you-grow efficiency for production, development, test, and Remote-Office and Branch-Office (ROBO) deployments.
  - This feature helps operationalize Splunk deployments at a fast pace.
  - Organizations experience Splunk as a ready-to-use solution.
- **Consistent performance:** Organizations gain consistent performance and significantly better utilization of the hardware and software platform resources for Splunk workloads.

Conclusion

Machine data offers a trove of insights, leading to organizational success and efficiency—but mining that data can be complicated without the right data analytics platform. Splunk Enterprise enables customers to derive real-time insights from this data, and Cisco HyperFlex systems' agility, consistency, and resiliency addresses the complexities of hardware resource management and the need for rapid deployment and operationalization. Splunk Enterprise and Cisco HyperFlex systems help organizations more quickly build and maintain a next-generation digital data center that can provide smarter business outcomes.

For more information

For additional information, see:

- www.cisco.com/go/bigdata
- www.cisco.com/go/bigdata_design
- www.cisco.com/go/HyperFlex

C22-739511-02 07/18