Cisco HyperFlex Systems

Hyperconvergence for the core, cloud, and edge

You need infrastructure that can follow your data and increase the speed of business regardless of where it takes place: from your core data center (enterprise applications, big data, and deep learning) to private and public clouds (virtualized and containerized applications) and edge locations (remote offices, branch offices, retail and industrial sites). Cisco HyperFlex™ systems with Intel® Xeon® Scalable processors deliver hyperconvergence with power and simplicity for any application, on any cloud, anywhere. Engineered on the Cisco Unified Computing System™ (Cisco UCS®), Cisco HyperFlex systems deliver the agility, scalability, and pay-as-you-grow economics of the cloud with the benefits of multisite, distributed computing at global scale.

The solution

Our platform includes hybrid, all-flash, all-NVMe, and edge configurations, an integrated network fabric, and powerful data optimization features that bring the full potential of hyperconvergence to a wide range of workloads and use cases. These faster-to-deploy, simpler-to-manage, and easier-to-scale systems provide a unified pool of infrastructure resources to power applications as your business needs dictate.

© 2017–2019 Cisco and/or its affiliates. All rights reserved.
What's new?

Platform version 4.0 implements Cisco HyperFlex Anywhere capabilities:

- **Flexibility and scalability at the edge.** 2-, 3-, and 4-node edge configurations support a wider set of use cases. Template-based lights-out deployment, configuration, management, and monitoring speeds anywhere deployment at scale, anywhere.

- **Cisco Intersight™ invisible cloud witness.** Automatic, cloud-based witness for 2-node clusters eliminates the complexity of configuring and maintaining a witness node for each edge site.

- **Inferencing at the edge.** You can perform deep learning on GPU-only nodes in the data center and drive inferencing with up to two NVIDIA® Tesla® T4 and P6 GPUs in edge nodes and up to six NVIDIA Tesla GPUs in Cisco HyperFlex HX240c nodes.

- **All-NVMe nodes.** These deliver the highest performance for mission-critical data center workloads. We provide an architectural performance edge with NVMe drives connected directly to the CPU rather than through a latency-inducing PCIe switch. Intel Optane™ SSDs also connect to the PCIe bus to accelerate caching for even greater performance than NVMe drives alone.

Any application, anywhere

Cisco HyperFlex systems include a purpose-built, high-performance, low-latency hyperconverged platform that adapts to support any application, in any cloud, anywhere (Figure 1). The platform supports multiple hypervisors and virtualized environments (Microsoft Windows Server 2019 Hyper-V, VMware vSphere), Docker containers with Kubernetes, multicloud services, and edge deployments to efficiently and economically deploy, monitor, and manage applications.

Cisco® Validated Designs give you the benefit of pretested enterprise application deployment using the best practices developed by Cisco engineers. These guidebooks for implementation help you accelerate deployment and reduce risk for virtual desktop environments (Citrix or VMware), Oracle Database, Microsoft SQL Server, big data applications including Splunk and SAP HANA, and graphics-accelerated high-performance computing, artificial intelligence (AI), and machine learning (ML) applications.

Complete solution

Designed with an end-to-end software-defined infrastructure, the platform eliminates the compromises found in other hyperconverged products. It combines software-defined computing using Cisco UCS servers, software-defined storage using the Cisco HyperFlex HX Data Platform, and software-defined networking (SDN) using Cisco Unified Fabric, which integrates smoothly with the Cisco Application Centric Infrastructure (Cisco ACI™).

With hybrid, all-flash, and all-NVMe configurations, self-encrypting drive options, and a new acceleration engine, the systems deliver a preintegrated cluster that is up and running in an hour or less and that scales resources independently to closely match your application resource needs (Figure 2). The systems support virtualized and containerized applications, multicloud services, and edge deployments with a simple, low-cost option for remote and branch-office locations. Optional GPU acceleration speeds AI and ML software for model training, machine learning, and inference generation.

Figure 1. Cisco HyperFlex systems support data center core, multicloud, and edge use cases.
What’s new? (continued)

- **Cisco HyperFlex Acceleration Engine.** Improves performance and efficiency in the data center with faster and lower-latency data compression operations. More compression makes more efficient use of storage resources.

- **Enhanced Cisco Intersight management as a service.** End-to-end lifecycle management lets you install, configure, manage, and monitor with a worldwide reach. We integrate the entire hardware stack for lights-out zero-touch deployment; no other vendor offers this capability. Full-stack provisioning and upgrades keep firmware, hypervisor, and data platform revisions at the level you specify. Parallel, heterogeneous deployment lets you handle massive scale. So do cluster profiles that make it as easy to deploy hundreds of sites as it is to deploy a single one. Connected Cisco TAC can automatically initiate support cases based on cloud-based monitoring.

Engineered on Cisco UCS

Cisco UCS fabric interconnects provide a single point of connectivity integrating Cisco HyperFlex HX-Series all-flash, all-NVMe, or hybrid nodes and other Cisco UCS servers into a single unified cluster. You can choose the combination of CPU, flash memory, graphics acceleration, and disk storage resources you need to deliver an optimal infrastructure for your applications. Incremental scalability allows you to start small and scale up and out as your needs grow.

Powered by next-generation data technology

The Cisco HyperFlex HX Data Platform combines the cluster’s SSDs, HDDs, and NVMe drives into a single distributed, multiter, object-based data store. A self-healing architecture replicates data for high availability, remediates hardware failures, and alerts IT administrators so that problems can be resolved quickly and your business can continue to operate.

Figure 2. Cisco HyperFlex systems support virtualized and containerized applications and let you choose the exact combination of resources to power your enterprise applications
“HyperFlex’s approach ensures high performance of Microsoft SQL and Oracle databases and critical applications with faster delivery of the environment, lower costs, and more effective management.”

Edivaldo Rocha
CEO
CorpFlex
Read the story

- **In-cluster synchronous replication** stripes and replicates data across the cluster. Data remains available if one or more components fail (depending on the replication factor).

- **Compression and deduplication** are always on with dedicated resources for consistent performance. The Cisco HyperFlex Acceleration Engine can improve performance and efficiency through lower-latency data compression operations and higher compression ratios.

- **Space-efficient, pointer-based snapshots and clones** facilitate backup operations.

- **Logical availability zones** increase availability for larger clusters by automatically partitioning the physical cluster into logical zones and then intelligently placing data to increase cluster resiliency to node and component failures.

- **Stretch clusters** support deployment into two geographically split locations for active/active operations even through a data center failure.

- **Thin provisioning** allows large data volumes to be created without dedicated storage, enabling a "pay-as-you-grow" procurement model.

- **Self-encrypting drive options** securely store data at rest in coordination with enterprise key management software.

- **Native replication** transfers virtual machine data to local or remote clusters for backup or disaster-recovery purposes. You can script, test, and execute failover plans with PowerShell scripting or integrate with third-party products.

- **Data protection API** integration enables enterprise backup tools to protect your data.

**Management at scale**

Cisco Intersight™ management as a service extends computing from the core data center to the cloud and edge, at virtually any scale. A single interface lets you manage all of your clusters at once with support for installation, inventory management, data platform configuration, and health status. A recommendation engine can tell you when you vary from best configuration practices, and a connection to the Cisco Technical Assistance Center (TAC) enables it to automatically open service requests when the management platform detects a problem. Unique in the industry, the update feature lets you upgrade your node firmware, hypervisor, and the HX Data Platform software with rolling, nondisruptive updates.

If you choose to host local management tools, all of your cluster operations can be managed with locally hosted Cisco HyperFlex Connect software. Virtual-machine-level management is supported in Microsoft System Center Virtual Machine Manager (SCVMM), Microsoft Hyper-V Manager, or the VMware vSphere plug-in.

**Next steps**

To deploy any application, in any cloud, anywhere, contact your Cisco sales representative or authorized partner.

Learn how Cisco HyperFlex systems with Intel Xeon Scalable processors can enable your digital transformation at cisco.com/go/hyperflex.