Cisco Data Intelligence Platform: Modernizing Data Lake with AI

Highlights

Cloud-scale architecture brings together big data, Artificial Intelligence (AI), and tiered storage.

Bringing together siloed workloads in a cloud-scale architecture allows a more robust cloud-like experience at both infrastructure and platform levels.

Paradigm shift in enabling diverse computing constructs to work on data.

The architecture of the Cisco® Data Intelligence Platform solution enables data to be operated by different computing constructs—whether CPU or Graphics Processing Unit (GPU) or Field-Programmable Gate Array (FPGA)—based on application demand.

Independently scale storage and computing resources based on demand.

This architecture enables both data-intensive and computation-intensive workloads to work together. Customers can start small and expand non-disruptively, independently scaling storage and computing resources to thousands of nodes.

Enable AI and Machine-Learning (ML) and microservices on Kubernetes and YARN.

Cisco Data Intelligence Platform enables Data Lake and Tiered storage to work closely with Industry leading containerization platform such as Kubernetes on Cisco UCS.

Get the Cisco UCS and Cisco ACI advantage

Architecture built on Cisco UCS

The Cisco Data Intelligence Platform is built on Cisco UCS, a proven platform for enterprise analytics applications. Cisco UCS offers a broad portfolio of Cisco Validated Designs with industry-leading Independent Software Vendor (ISV) partners in each of these areas: big data, AI, and object storage.

Ease of deployment

Cisco UCS Manager together with Cisco Intersight™ software simplifies infrastructure deployment with an automated, policy-based mechanism that helps reduce configuration errors and system downtime. The Cisco UCS and Cisco Intersight solution enables users to manage multiple Cisco UCS instances or domains across disparate locations and environments.

Single-pane management of hundreds of switches with Cisco Application Centric Infrastructure (Cisco ACI™)

With Cisco ACI, the Cisco Data Intelligence Platform can be scaled to thousands of nodes and handle multiple petabytes of data. Cisco ACI treats the network as a single entity instead of as a collection of switches and offers simple and easy network management through a single management pane.

Data landscape and architectural evolution

Data scientists are constantly searching for newer techniques and methodologies that can unlock the value of big data and distill this data further to identify additional insights which could transform productivity and provide business differentiation.

One such area is Artificial Intelligence/Machine Learning (AI/ML), which has seen tremendous development with bringing in new frameworks and new forms of compute (CPU, GPU and FPGA) to work on data to provide key insights. While data lake has historically been data intensive workloads, these advancement in technologies has led to a new growing demand of compute intensive workloads to operate on the same data.

While data scientists want to be able use these latest and greatest advancement in AI/ML software and hardware technologies on their datasets, the IT is also constantly looking at enabling these data scientists to be able to provide such a platform to a data lake. This has led to an architecturally siloed implementations where data which is ingested into a data lake is often worked and processed in a data lake and when needs to be operated by AI/ML frameworks often leaves the platform and has to be onboarded to a different platform to process it. This would be fine if this demand is seen only on a small percentage of workloads. However, AI/ML workloads working closely on the data in a data lake are seeing an increase in adoption. For instance, data lakes in customer environment is seeing deluge of data from new use cases such as IoT, Autonomous driving, smart cities, genomics and financials, who are all seeing more and more demand of AI/ML processing of this data.

IT is demanding newer solutions to enable the data scientists operate on both data lake and a AI/ML platform (or a compute farm) without worrying about the underlying infrastructure. IT also needs this to seamlessly grow to cloud scale while reducing the TCO of this infrastructure and without affecting utilization. Thus, driving a need to plan a data lake along with an AI/ML platform in a systemic fashion.

illuluillu
CISCO

The bridge to possible

Cisco Data Intelligence Platform

The Cisco® Data Intelligence Platform brings big data, AI compute farms, and tiered storage to work together as a single entity, but with each element still able to be scaled independently. This architecture supports:

• Extremely fast ingestion and engineering of data performed at the data lake

• An AI computing farm, allowing different types of AI frameworks and computing resources (GPU, CPU, and FPGA) to work on this data for additional analytics processing

• A storage tier, allowing the gradual retirement of data that has been worked on to a dense storage system with a lower cost per terabyte, reducing TCO

The Cisco Data Intelligence Platform supports today's evolving architecture (Figure 1). It brings together a fully scalable infrastructure with centralized management and a fully supported software stack (in partnership with industry leaders in the relevant areas) to each of these three independently scalable components of the architecture: the data lake, AI/ML technologies, and object stores.

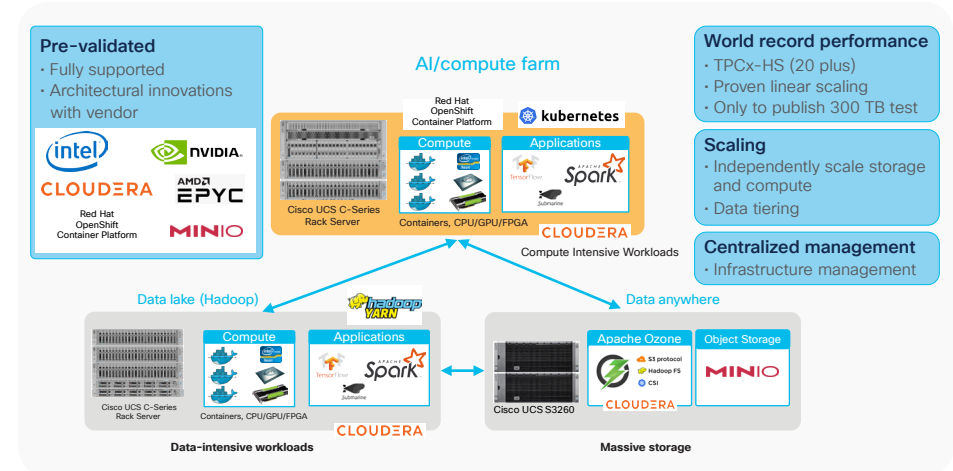Figure 1.   Cisco Data Intelligence Platform
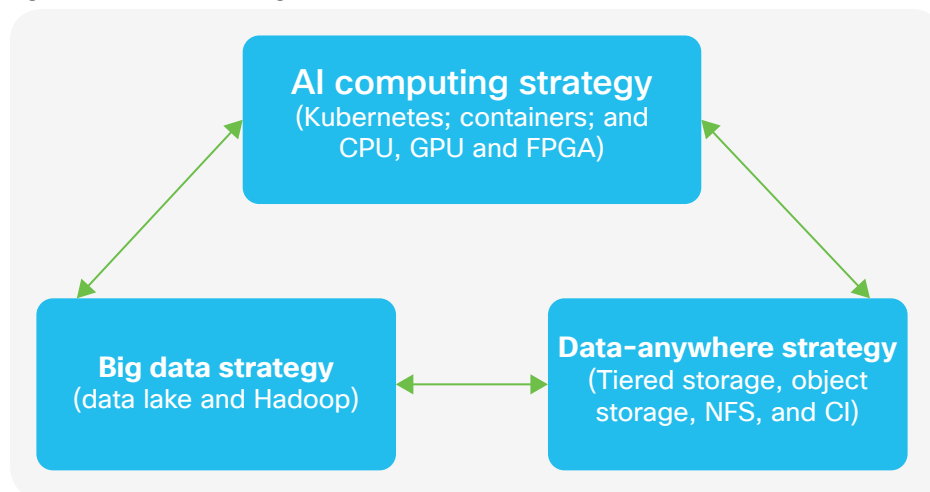
Cisco has developed numerous industry-leading Cisco Validated Designs (reference architectures) with architectures for big data (with Cloudera, Hortonworks, and MapR), computing farms with Kubernetes (with Red Hat OpenShift), and object stores (with Scality, and Cloudian)

Hortonworks merged with Cloudera now they are single company and with CDP PC release it all comes under one architecture.

Figure 2 shows the platform's software stack and ISV partners.

Figure 2.   Cisco Data Intelligence Platform with software stack and ISV partners

ıllıılı
CISCO

The bridge to possible

Cisco Data Intelligence Platform deployed on Hadoop

With Hadoop, most of these concepts are organic, as shown in Figure 3. The data lake consists of Apache Kafka (data retention) and Hadoop nodes for data-intensive workloads and YARN-only nodes for the AI computing farm and tiered storage for massive storage.

Figure 3.   Cisco Data Intelligence Platform deployed with Hadoop
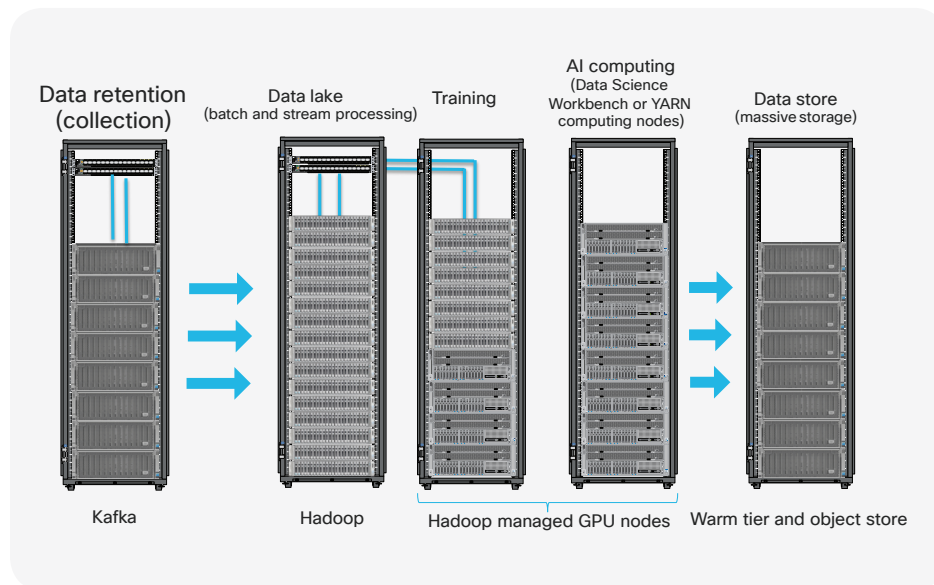
Data-intensive workloads (Hadoop)

For data-intensive workloads, the Cisco Data Intelligence Platform is built on Hadoop. Hadoop enables data engineering, providing very fast ingestion of data and Extract, Transform, and Load (ETL) processing. In a data-intensive workload, computing moves to the data to enable faster, distributed processing of the data.

As organizations scale their Hadoop deployments, they want to run more analytics applications with various data-access models, such as batch, interactive, real-time, and streaming workloads, that all need to access the data simultaneously (lambda architecture).

Enterprises and service providers collect data from various sources such as IoT, web, social media, logs, and sensors. This data is both growing at an increasingly fast rate and is stored using totally different file structures and formats. For the Hadoop Distributed File System (HDFS) to process it, this data first needs to be ingested in an efficient and reliable manner through the lambda architecture for a data lake, as shown in Figure 4.

Figure 4.   Lambda architecture for a data lake supporting high-speed, batch, and server layer processing
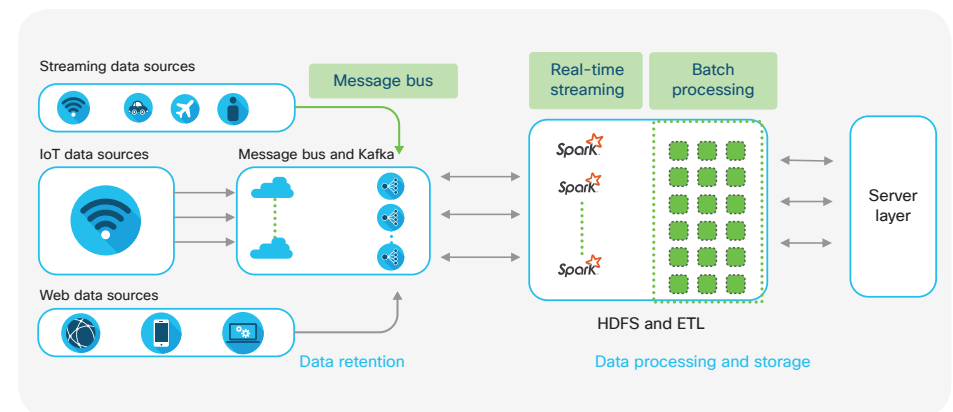
Building a data pipelines that receives data flows from different data sources at higher velocities, performs ETL on this data to land in HDFS, and makes it available for serving layer either for real-time streaming or batch processing is an extremely IO intensive operation.

atualtılı
CISCO

The bridge to possible

Big data relies heavily on I/O bandwidth and requires high network bandwidth utilization. As a result, network designs for Hadoop in the data center must involve little or no oversubscription. Modern data centers for big data require networks that scale out and support interhost communication with significant east-west traffic. To make best use of the scale-out capabilities—with designs scaling to thousands of nodes—storage and computing resources should be deployed in spine-and-leaf network fabric. Such a design reduces the number of network hops and the addition of oversubscription at different layers. In a spine-and-leaf network fabric, all nodes are equally distant, and the network can be scaled out by adding more spine switches as needed.

> In the Cisco Data Intelligence Platform, AI computing farm supports computation-intensive workloads in a multi-tenant containerized environment in a Kubernetes cluster.

AI computing farm (computation-intensive workloads)

Today, deep learning is fueling many scenarios, including autonomous driving, computer vision, health care (cancer diagnosis, drug discovery, etc.), speech and image recognition, and video analytics. But processing these kinds of analytics for a very large data set with millions of simulations and several thousand containers to achieve high precision requires a lot of computation. These new use cases and applications need a lot of computing power (often expressed in teraflops—one trillion floating-point operations per second—or TFLOPs), and GPUs along with CPUs are required to power the AI/ML algorithms.

Furthermore, with containerization, organizations can now manage computing resources elastically (on demand), deploy applications in microservices architectures, and run multiple versions of applications for AI/ML and deep-learning workloads.

The Cisco Data Intelligence Platform integrates an AI computing farm with Hadoop clusters, enabling organizations to easily run AI/ML containerized workloads in the computing farm while accessing data on HDFS. The computing farm provides a large pool of memory, CPU, and GPU resources as a whole to Hadoop cluster. The computing farm enables logical separation between data and computing, thus allowing massive linear scaling without disruption.

> The data lake in the Cisco Data Intelligence Platform is designed with servers that support high I/O and network bandwidth with little or no network oversubscription to help prevent bottlenecks even when the network is scaled out to thousands of servers.

AI computing farm deployed on Hadoop

The AI computing farm can be deployed on Hadoop either as Cloudera Data Science Workbench or as Hadoop compute-only nodes.

Cloudera Data Science Workbench

Cloudera Data Science Workbench is a web-based application that allows data scientists to use their favorite open-source libraries and languages, including R, Python, and Scala, directly in secure environments, accelerating analytics projects from research to production.

With the help of Data Science Workbench, data scientists can share, collaborate, and manage their data and libraries with Hadoop, resulting in an easier and faster path to production that is secure for the enterprise.

Data Science Workbench provides the following features:

- **CPU and GPU as a resource:** Data Science Workbench provides basic support for the use of existing general-purpose CPUs for each stage of the workflow and, optionally, accelerates the math-intensive steps with the selective application of special-purpose GPUs all through a Docker container, with Kubernetes scheduling these resources in the back end.

The bridge to possible

- **Self-service portal:** The Data Science Workbench web user interface console provides a self-service portal for data scientists to create an environment for their workloads (Figure 5). Currently, R, Python, and Scala are supported.
- **Jupyter Notebook:** Most data scientists use Jupyter Notebooks for AI/ML analysis and development. Data Science Workbench provides a Jupyter Notebook environment when data scientists create a portal, and these notebooks can be shared or worked in a collaborative manner.

Figure 5.   Example of Cloudera Data Science Workbench web user interface
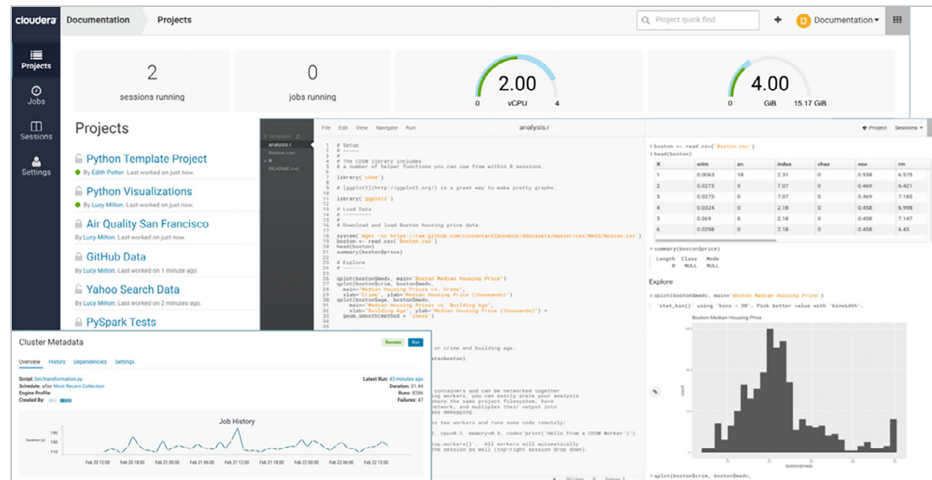
Hadoop computing-only nodes (YARN-only nodes)

In Hadoop, typically a data node is also a computing node. However, Hadoop also allows computing-only nodes, which the user can use to separate storage and computing entities. Although this model is not common for data-intensive workloads, for new evolving workloads of Docker containers on Hadoop with Hadoop 3.0, it offers the capability to separate these rich AI/ML container-based application stacks onto independent clusters with different CPU and GPU specifications than those used for data nodes (Figure 6).

Figure 6.   Hadoop dashboard showing CPU, GPU, and memory as managed resources
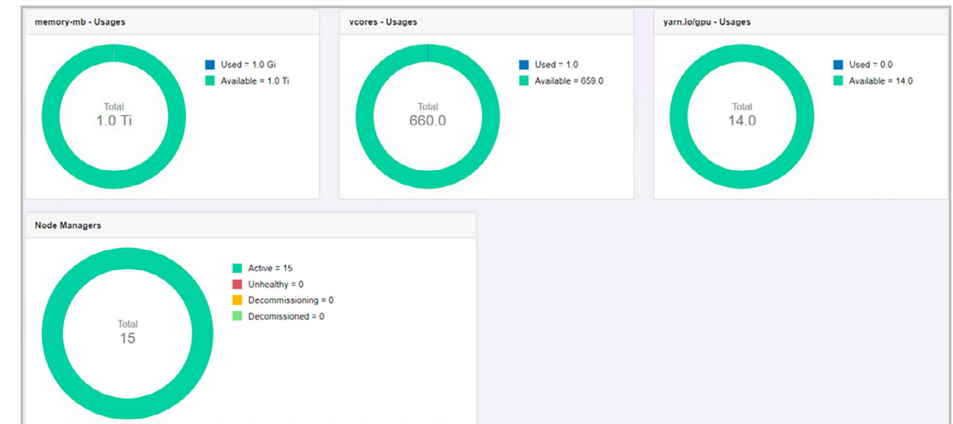
Computing farm with Kubernetes

Kubernetes is rapidly becoming, in effect, the standard platform as it gains enormous adoption for data science workloads. Data science often uses interactive coding and visualization environments such as Jupyter Notebook, Apache Zeppelin, and RStudio. These environments provide capabilities to data scientists for writing code in Python and R. Computation-intensive tasks such as deep learning require GPU resources, and for larger data sets data science tasks quickly reach the limits of system resources. An AI computing farm with a Kubernetes cluster requires abundant CPU, memory, and GPU resources for large-scale data science projects consisting of thousands of containers. Spark, which will bring in support for native Kubernetes, can be used for processing large data sets. Also, by using Kubeflow, Tensorflow training models can be implemented by provisioning Tensorflow clusters within Kubernetes.

To augment these features, external storage is needed to store end results and output files. The Kubernetes persistent volume framework allows you to provision persistent storage using networked storage available in your environment. Persistent volumes can be provisioned according to your application needs using the appropriate volume drivers of the storage back end, such as Network File System (NFS), Ceph, and Gluster, giving users a way to request those resources without having any knowledge of the underlying infrastructure.

Red Hat OpenShift Container Platform

Red Hat OpenShift Container Platform (OCP) is the leading platform for containerized applications. OpenShift uses Kubernetes to schedule and orchestrate resources within the cluster. OCP stitches Dockers and Kubernetes together and provides the foundation to quickly build, develop, and deploy applications and services.

NVIDIA GPU Cloud

The NVIDIA GPU Cloud (NGC) container registry provides GPU-optimized software tools for AI. Red Hat OpenShift Enterprise Linux (RHEL) is officially supported through the NGC-Ready validation program. Red Hat OpenShift provides a platform for orchestrating these containers at scale.

> Cisco and NVIDIA have partnered to provide fully supported NGC containers on Cisco UCS C480 ML M5 Rack Servers, providing customers with the assurance of a validated solution when they perform AI/ML processing on Cisco UCS.

Data anywhere (or massive storage)

New data is more frequently accessed than older data. Therefore, over time the frequency of read operations on a given data set naturally decreases, with data becoming less frequently accessed as it ages. New data is deemed "hot," and old data (data that has already been analyzed and may not be part of 95 percent of the queries or workloads) is deemed "cold" or archival. An in-between type of data is "warm" data. As enterprises collect increasing volumes of data of all three types, they have a growing need to retire data to a more cost-effective storage with better $/TB. Below are examples of Hot, Warm, and Cold data tiers:

- The **hot data tier** delivers a storage tier that consists of Cisco UCS C240 M5 Rack Servers to store data sets that require high-speed storage access.
- The **warm data tier** uses Cisco UCS S3260 M5 Storage Servers, providing high-capacity storage within HDFS clusters by defining storage types and policies in HDFS. Each Cisco UCS S3260 has two nodes, and each node supports up to 100 TB of HDFS data.
- The **cold data tier** is provided through the object store, which provides only one copy of data, with processing on this data performed outside the storage unit.
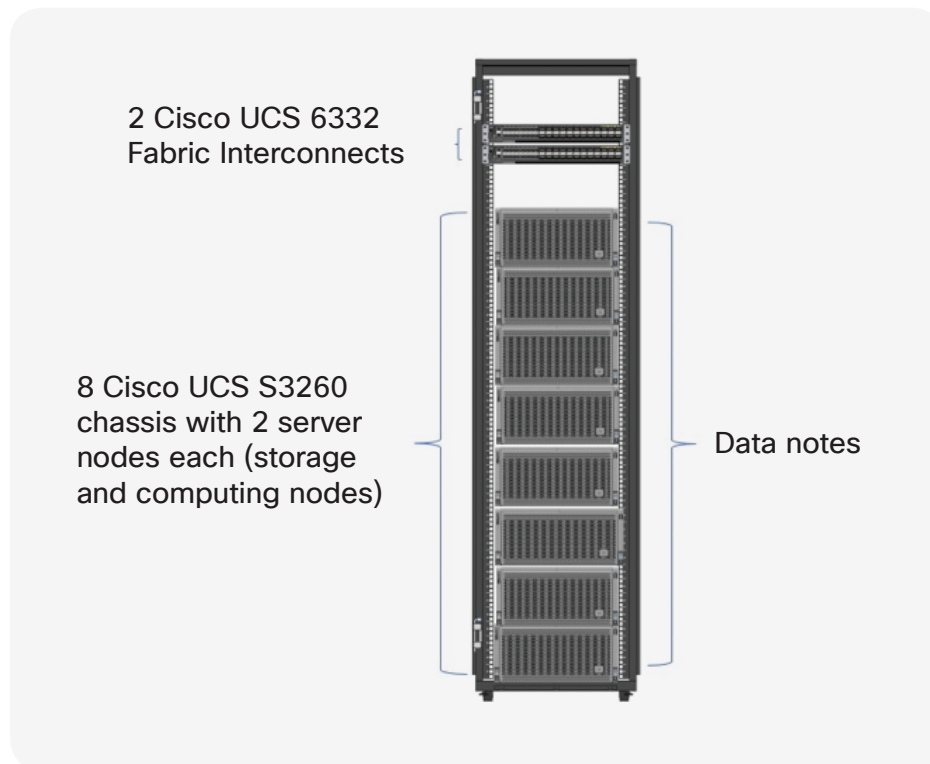
Tiered storage with Hadoop (warm data)

Tiered storage with Hadoop maintains three copies of the data; however, much data is consolidated in a single rack. This approach still retains three copies, but because of consolidation the warm tier stores almost three times as much data on a single rack as is the case on the hot tier. Thus, the amount of I/O bandwidth, network bandwidth, computing cores, and memory working on this data is reduced to almost one third (Figure 7). This approach still follows the Hadoop principle of moving the computing to the data, and processing is performed where the data is located, avoiding additional network load.

ılıılı
CISCO

The bridge to possible

Tiered storage offers a number of advantages:

- **Cost effectiveness:** Tiered storage promotes huge reductions in storage costs. Use of a single type of storage for all data is a waste of money for most businesses.
- **Operational efficiency:** With tiered storage, only the data that serves important business functions ends up in a performance-optimized layer, and archival data ends up in a capacity-optimized layer.
- **Flexibility:** You have the flexibility, particularly with automated tiering, to move data between different storage media as its value changes.

Figure 7. 2.4 petabytes of Hadoop data consolidated in a single rack with eight Cisco UCS S3260 M5 servers

2 Cisco UCS 6332 Fabric Interconnects

8 Cisco UCS S3260 chassis with 2 server nodes each (storage and computing nodes)

Data notes

Object-storage solution on Cisco UCS S3260 (cold tier)

With cold-tier storage, instead of HDFS managing data, a large amount of data is consolidated onto a single rack for the object store. Hadoop has a limitation of 100 TB per node; however, with the object store the entire 720 TB of storage can be used. Furthermore, most object stores have erasure encoding for data resiliency, which means that only 1.7 times the data is stored as opposed to 3 times the data with Hadoop.

The data in the object store is considered cold because processing on this data is performed outside these nodes, and only one copy of the data is available to the server for data read operations. Also, the data must be copied to a YARN-only Hadoop computing node or to another computing framework to be processed, leading to additional network load. Primarily, the data in the object store is data that is expected to be accessed in fewer than 5 percent of the queries, but is required to be stored.

> Cisco has partnered with industry-leading object-storage vendors and has published multiple Cisco Validated Designs with Scality, and Cloudian on Cisco UCS S3260 servers.

Today, many enterprises are facing the challenge of growing data sets. The rapid increase in the amount of structured and unstructured data being created requires enterprises to find cost-effective ways to store and manage their data. Because of its data protection features using erasure coding and its capability to scale to multiple-petabytes ranges at much lower costs, object storage typically is an excellent solution for storing massive amounts of data.

Object-storage systems are based on HTTP and representational state transfer (REST) protocol. Some object-storage vendors have added support for enterprise file-sharing protocols such as NFS. In addition, some object-storage systems support other important HTTP-based protocols, such as Simple Storage Service (S3) APIs, which is a standard interface for object storage. Object storage can scale out to thousands of petabytes and costs less to manage.

Data transfer between tiers

The Cisco Data Intelligence Platform allows data to be transferred between the various tiers (Figure 8).

Figure 8.   Data transfer between different tiers of Cisco Data Intelligence Platform
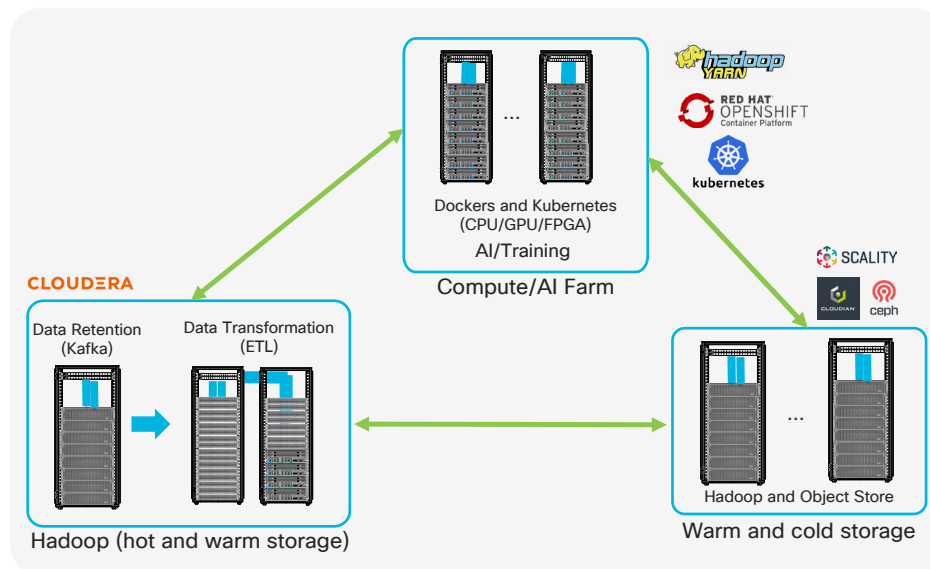
Data access between the Hadoop (data lake) and computing farm

Data can be accessed in several ways between the Hadoop data lake and computing farm.

YARN computing nodes

One way to enable a computing farm in Hadoop is by deploying YARN-only nodes. The computing resource scheduler in Hadoop is YARN, and a YARN-only node can schedule CPU, GPU, memory, and even Docker containers for applications. YARN nodes are by default part of a Hadoop cluster and have access to data in HDFS. If used to primarily schedule AI workloads on Docker containers and GPU, these nodes could be scaled independently from the data lake nodes.

Cloudera Data Science Workbench

Another way to enable an AI computing farm in Hadoop is through Data Science Workbench. This platform can be deployed with both Cloudera and Hortonworks. It is primarily a Kubernetes cluster that manages and schedules the CPU, GPU, and memory as resources attached to a Docker container to form a computing farm. It delegates the orchestration of these containers to Kubernetes. At session launch, project files, libraries, and client configurations are mounted to a newly started Docker container preconfigured with secured access to the Hadoop cluster, because all containers are clients of the Hadoop cluster.

Cisco UCS S3260 Storage Servers provide a high-capacity cold storage tier. Software-defined storage (SDS) with Cisco UCS S3260 brings together the simplicity and agility of the cloud and the cost benefits of industry-standard servers. It offers an excellent S3-compatible object-storage platform that is highly scalable and optimized for capacity and I/O performance.

The bridge to possible

This architecture is transparent to end users. They do not need to know about these components and will interact with the platform only through the web user interface.

This solution is suitable for those frameworks that have APIs in programming languages supported by Data Science Workbench: Python, R, and Scala.

Kubernetes

The computing farm accessing data on Hadoop can be a standalone Kubernetes cluster (such as Red Hat OpenShift OCP). The containers launched by Kubernetes in this cluster can access data on Hadoop either through the Kubernetes persistent volume framework support brought in by Hadoop or by accessing data from Hadoop as an NFS mount point.

Data access between computing farm and Hadoop tiered storage (warm data)

The YARN computing nodes, Data Science Workbench, or Kubernetes clusters accessing data on Hadoop tiered storage (warm data) are the same as discussed in the previous section.

Data access between computing farm and object store (cold storage)

S3-compatible object storage provides an HTTP and REST-based API for reading and writing objects. All operations, such as HTTP PUT, GET, POST, DELETE, and HEAD, can be performed using the S3 API. For example, applications can get and put objects into object storage using the object API.

Data between the computing farm (YARN computing nodes, Data Science Workbench, or Kubernetes clusters) and object storage can also be accessed by setting up a client application that provides simplified access to containers and buckets in the object store.

Data access between HDFS cluster (hot tier) and object store (cold tier)

Cloudera Manager enables data replication across data centers for backup and disaster-recovery scenarios. It also allows replication of HDFS data to and from S3, and Hive data and metadata to and from S3 also can be replicated.

Similarly, Hortonworks DataPlane Service (DPS) is a common set of services for managing, securing, and governing data assets across multiple tiers and types. It does this for data at rest and in multiple clusters and tiers.

Hortonworks Data Lifecycle Manager (DLM) is a user interface service that is enabled through DPS. It supports replication of HDFS and Hive data between underlying HDFS and S3 storage. From DLM, replication and disaster-recovery policies and jobs can be created. With autotiering, customers can create dynamic policies suited to their needs and can move data between different tiers, such as from expensive Solid-State Disk (SSD) to hard-disk drive (HDD) media or to the S3 bucket to reduce TCO.
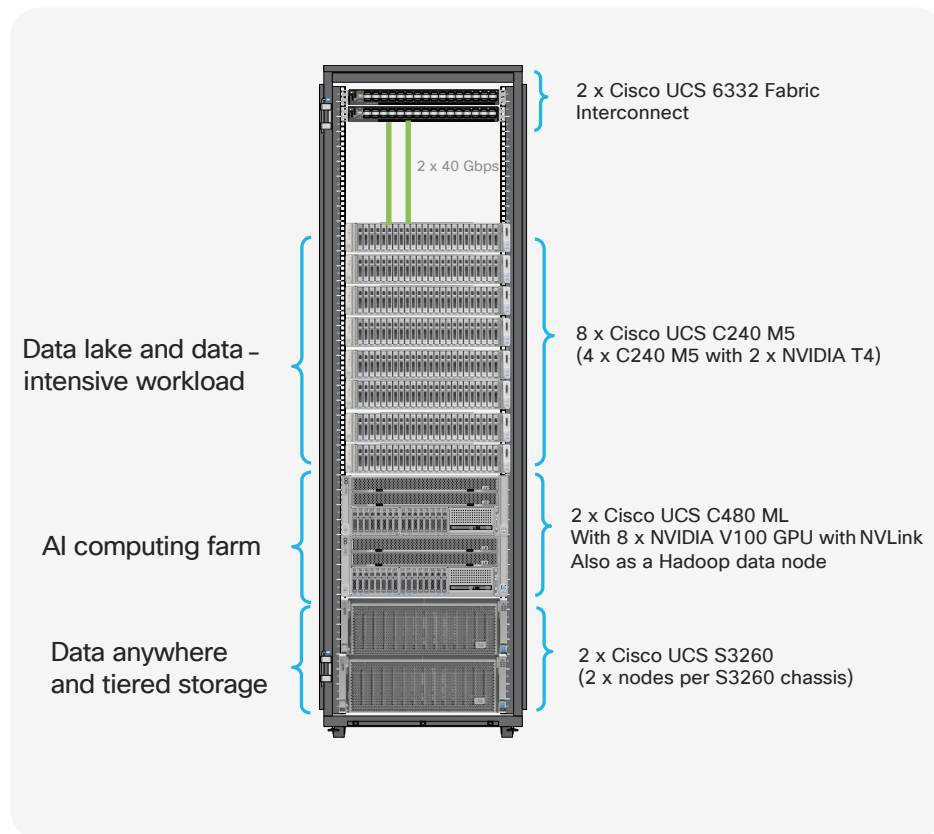
Data access between hot and warm tiers in Hadoop cluster

The application stack discussed in the preceding sections is also relevant to the movement of data between different tiers of storage within a Hadoop cluster, such as between hot and warm storage.

Scaling the architecture

This solution can be deployed in a single rack and can be scaled to thousands of nodes. Figure 9 shows the reference architecture for a single rack. This design can be scaled to more than a thousand nodes.

Figure 9. Cisco Data Intelligence Platform deployed in a single rack

2 x Cisco UCS 6332 Fabric Interconnect

2 x 40 Gbps

Data lake and data-intensive workload

8 x Cisco UCS C240 M5
(4 x C240 M5 with 2 x NVIDIA T4)

AI computing farm

2 x Cisco UCS C480 ML
With 8 x NVIDIA V100 GPU with NVLink
Also as a Hadoop data node

Data anywhere and tiered storage

2 x Cisco UCS S3260
(2 x nodes per S3260 chassis)

In the reference architectures discussed here, each of the components is scaled separately, and for the purposes of this example, scaling is uniform. Two scale scenarios are discussed here:

- Scaled architecture with 3:1 oversubscription with Cisco fabric interconnects and Cisco ACI
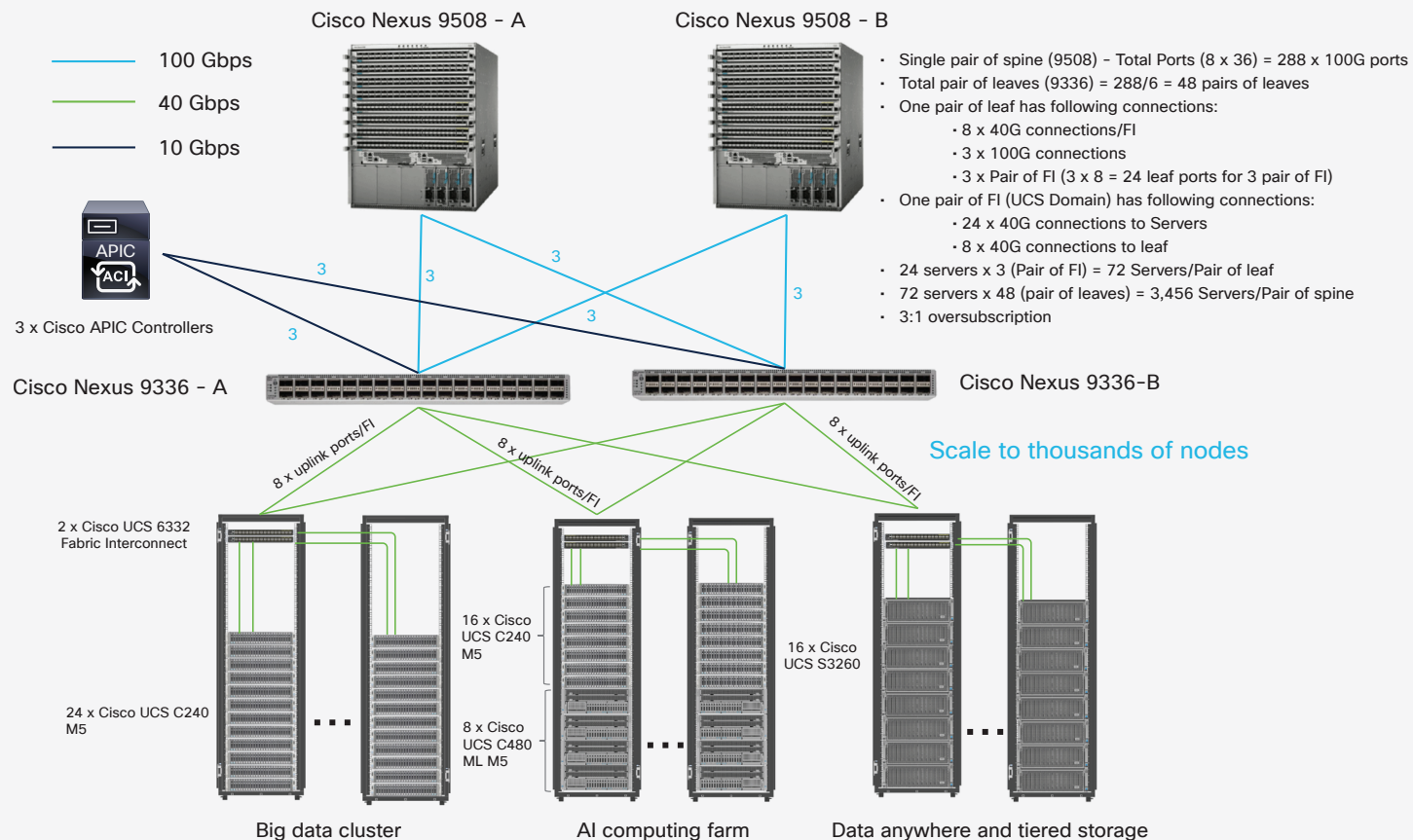- Scaled architecture with 2:1 oversubscription with Cisco ACI

In the following scenarios, the goal is to populate up to a maximum of 200 leaf nodes in a Cisco ACI domain. Not all cases reach that number because they use the Cisco Nexus® 9508 Switch for this sizing and not the Cisco Nexus 9516 Switch.

Scaled architecture with 3:1 oversubscription with Cisco fabric interconnects and Cisco ACI

The architecture discussed here and shown in Figure 10 supports 3:1 network oversubscription from every node to every other node across a multidomain cluster (nodes in a single domain within a pair of Cisco fabric interconnects are locally switched and not oversubscribed).

From the viewpoint of the data lake, 24 Cisco UCS C240 M5 Rack Servers are connected to a pair of Cisco UCS 6332 Fabric Interconnects (with 32 x 40-Gbps throughput). From each fabric interconnect, 8 x 40-Gbps links connect to a pair of Cisco Nexus 9336 Switches. Two pairs of fabric interconnects can connect to a single pair of Cisco Nexus 9336 Switches (8 x 2 40-Gbps links). Each of these Cisco Nexus 9336 Switches connects to a pair of Cisco Nexus 9508 Cisco ACI switches with 6 x 100-Gbps uplinks (connecting to a Cisco N9K-X9736C-FX line card).

Figure 10.   Scaled architecture with 3:1 oversubscription with Cisco fabric interconnects and Cisco ACI

Scaled architecture with 2:1 oversubscription with Cisco ACI

In the scenario discussed here and shown in Figure 11, the Cisco Nexus 9508 Switch with the Cisco N9K-X9736C-FX line card can support up to 36 x 100-Gbps ports, each and 8 such line cards.

Here, for the 2:1 oversubscription, 30 Cisco UCS C240 M5 Rack Servers are connected to a pair of Cisco Nexus 9336 Switches, and each Cisco Nexus 9336 connects to a pair of Cisco Nexus 9508 Switches with three uplinks each. A pair of Cisco Nexus 9336 Switches can support 30 servers and connect to a spine with 6 x 100-Gbps links on each spine. This single pod (pair of Cisco Nexus 9336 Switches connecting to 30 Cisco UCS C240 M5 servers and 6 uplinks to each spine) can be repeated 48 times (288/6) for a given Cisco Nexus 9508 Switch and can support up to1440 servers.

ılıılı
CISCO

The bridge to possible

To reduce the oversubscription ratio (to get 1:1 network subscription from any node to any node), you can use just 15 servers under a pair of Cisco Nexus 9336 Switches and then move to Cisco Nexus 9516 Switches (the number of leaf nodes would double).

To scale beyond this number, multiple spines can be aggregated.

Reference architecture

Tables 1, 2, and 3 summarize the reference architecture configuration details for the data lake, AI/ML components of the data lake, and tiered storage.

Data lake reference architecture

Table 1 lists the data lake reference architecture configuration details for Cisco UCS Integrated Infrastructure for Big Data and Analytics.

Figure 11.   Scaled architecture with 2:1 oversubscription with Cisco ACI
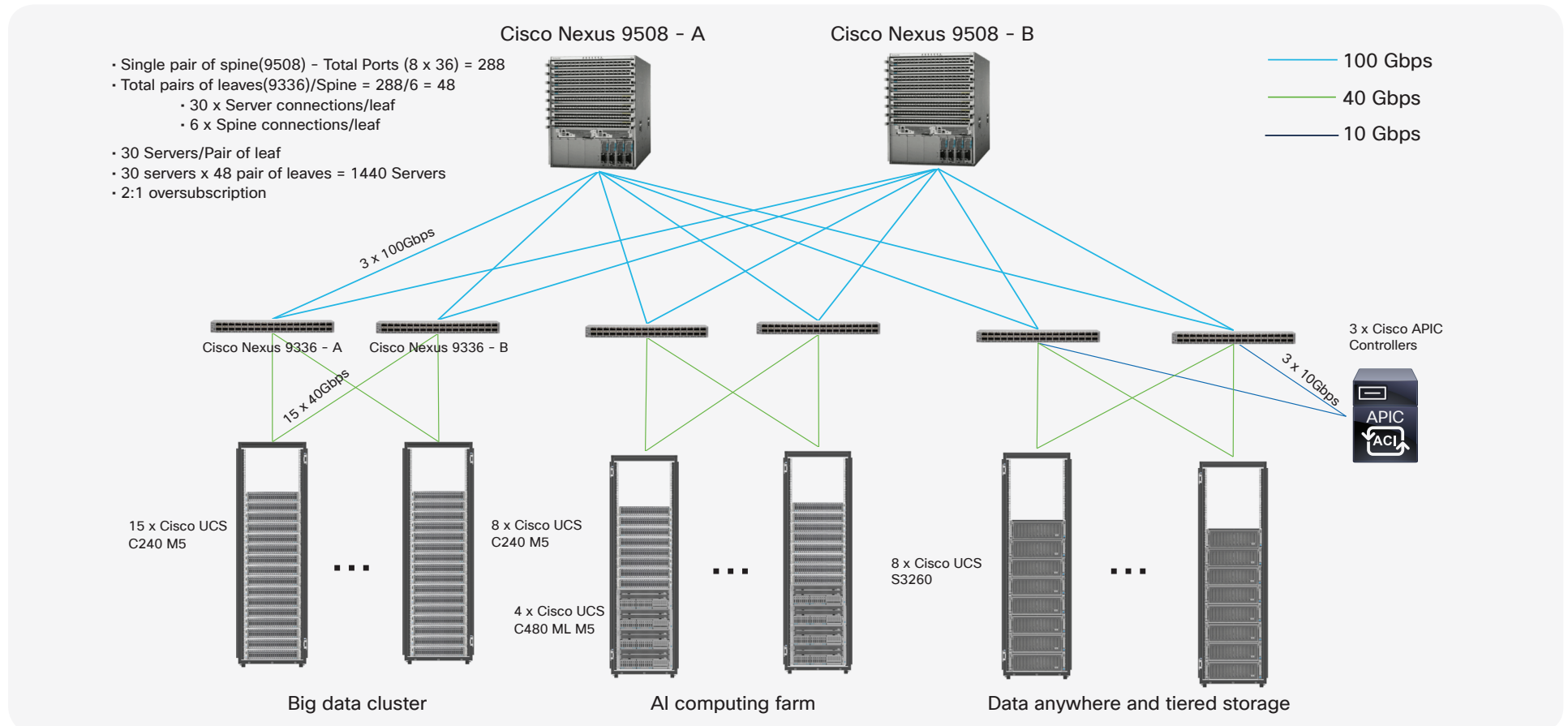
ııılıılı
CISCO

The bridge to possible

Table 1.   Cisco UCS Integrated Infrastructure for Big Data and Analytics configuration options for data lakes

	NVMe performance	Flash performance	Performance	Capacity	High capacity
Servers	16 x Cisco UCS C220 M5SN Rack Servers with Small- Form-Factor (SFF) drives (UCSC-C220-M5SN)	8 x Cisco UCS C4200 Series Rack Servers with 4 x Cisco UCS C125 M5 Rack Server Nodes	16 x Cisco UCS C240 M5 Rack Servers with Small- Form-Factor (SFF) drives	16 x Cisco UCS C240 M5 Rack Servers with Large- Form-Factor (LFF) drives	8 x Cisco UCS S3260 Storage Servers with two S3260 M5 server nodes
CPU	2 x 2nd Gen Intel® Xeon® Scalable Processor 6230R (2 x 26 cores, at 2.1 GHz)	1 x AMD 7352 Processor (24 cores, at 2.3 GHz)	2 x 2nd Gen Intel® Xeon® Scalable Processor 5218R (2 x 20 cores, at 2.1 GHz)	2 x 2nd Gen Intel Xeon Scalable Processor 4210R (2 x 10 cores, at 2.4 GHz)	2 x 2nd Gen Intel Xeon Scalable Processor 6230R (2 x 26 cores, 2.1 GHz)
Memory	12 x 32-GB DDR4 (384 GB)	16 x 32-GB DDR4 (512 GB)	12 x 32-GB DDR4 (384 GB)	12 x 32-GB DDR4 (384 GB)	12 x 32-GB 2666 MHz (384 GB)
Boot	Cisco Boot-Optimized M.2 RAID Controller with 2 x 240-GB SSDs	M.2 with 2 x 240-GB SATA SSDs	Cisco Boot-Optimized M.2 RAID Controller with 2 x 240-GB SSDs	Cisco Boot-Optimized M.2 RAID Controller with 2 x 240-GB SSDs	2 x 240-GB SATA Boot SSDs
Storage	10 x 8-TB 2.5-in U.2 Intel P4510 NVMe high-performance value endurance	6 x 7.6-TB enterprise-value SATA SSDs	26 x 2.4-TB 10K RPM SFF SAS HDDs or 12 x 1.6-TB enterprise-value SATA SSDs	12 x 8-TB 7.2K RPM LFF SAS HDDs	14 x 8-TB 7.2K RPM LFF SAS HDDs
Virtual Interface Card (VIC)	25 Gigabit Ethernet (Cisco UCS VIC 1457) or 40/100 Gigabit Ethernet (Cisco UCS VIC 1497)	25 Gigabit Ethernet (Cisco UCS VIC 1455)	25 Gigabit Ethernet (Cisco UCS VIC 1457) or 40/100 Gigabit Ethernet (Cisco UCS VIC 1497)	25 Gigabit Ethernet (Cisco UCS VIC 1457) or 40/100 Gigabit Ethernet (Cisco UCS VIC 1497)	25 Gigabit Ethernet (Cisco UCS VIC 1455) or 40/100 Gigabit Ethernet (Cisco UCS VIC 1495)
Storage controller	NVMe switch included in the optimized server	Cisco 12-Gbps SAS 9460-8i RAID controller with 2-GB FBWC	Cisco 12-Gbps SAS modular RAID controller with 4-GB Flash-Based Write Cache (FBWC) or Cisco 12-Gbps modular SAS Host Bus Adapter (HBA)	Cisco 12-Gbps SAS modular RAID controller with 2-GB FBWC or Cisco 12-Gbps modular SAS Host Bus Adapter (HBA)	Cisco UCS S3260 dual RAID controller
Network connectivity	Cisco UCS 6332 Fabric Interconnect or Cisco UCS 6454/64108 Fabric Interconnect	Cisco UCS 6454/64108 Fabric Interconnect	Cisco UCS 6332 Fabric Interconnect or Cisco UCS 6454/64108 Fabric Interconnect	Cisco UCS 6332 Fabric Interconnect or Cisco UCS 6454/64108 Fabric Interconnect	Cisco UCS 6332 Fabric Interconnect or Cisco UCS 6454/64108 Fabric Interconnect
GPU (optional)	Up to 2 x NVIDIA Tesla T4 with 16 GB of memory each		Up to 2 x NVIDIA Tesla V100 with 32 GB of memory each or up to 6 x NVIDIA Tesla T4 with 16 GB of memory each	2 x NVIDIA Tesla V100 with 32 GB memory of each or up to 6 x NVIDIA Tesla T4 with 16 GB of memory each	

Note: The Cisco UCS C240 M5 SFF Hybrid with NVMe reference architecture is described in the UCS C240 M5 section.

cisco
CISCO
The bridge to possible

Table 2.   Cisco UCS Integrated Infrastructure for Big Data and Analytics configuration options for high-density CPU cores and GPU nodes

	Select stack	Elite stack	Premier stack
Servers	8 x Cisco UCS C240 M5 Rack Servers 4 x Cisco UCS C480 M5 Rack Servers	8 x Cisco UCS C240 M5 Rack Servers 4 x Cisco UCS C480 ML M5 Rack Servers	8 x Cisco UCS C4200 Rack Server Chassis, each with 4 x Cisco UCS C125 M5 Rack Server Nodes
CPU	2 x 2nd Gen Intel Xeon Scalable Processor 6230R (2 x 26 cores at 2.1 GHz)	2 x 2nd Gen Intel Xeon Scalable Processor 6230R (2 x 26 cores at 2.1 GHz)	2 x AMD 7552 processor (2 x 48 cores at 2.2 GHz)
Memory	12 x 32-GB DDR4 (384 GB)	12 x 32-GB DDR4 (384 GB)	16 x 32-GB DDR4 (512 GB)
Boot	M.2 with 2 x 960-GB SSDs	M.2 with 2 x 960-GB SSDs	M.2 with 2 x 240-GB SATA SSDs
Storage	24 x 1.8-TB 10K rpm SFF SAS HDDs or 12 x 1.6-TB enterprise-value SATA SSDs	24 x 1.8-TB 10K rpm SFF SAS HDDs or 12 x 1.6-TB enterprise-value SATA SSDs	6 x 3.8-TB enterprise-value SATA SSDs
Virtual Interface Card (VIC)	25 Gigabit Ethernet (Cisco UCS VIC 1457) or 40/100 Gigabit Ethernet (Cisco UCS VIC 1497)	25 Gigabit Ethernet (Cisco UCS VIC 1457) or 40/100 Gigabit Ethernet (Cisco UCS VIC 1497)	25 Gigabit Ethernet (Cisco UCS VIC 1455)
Storage controller	Cisco 12-Gbps SAS modular RAID controller with 4-GB FBWC or Cisco 12-Gbps modular SAS HBA	Cisco 12-Gbps SAS modular RAID controller with 4-GB FBWC or Cisco 12-Gbps modular SAS HBA	Cisco 12-Gbps SAS 9460-8i RAID controller with 2-GB FBWC
Network connectivity	Cisco UCS 6332 Fabric Interconnect or Cisco UCS 6454/64108 Fabric Interconnect	Cisco UCS 6332 Fabric Interconnect or Cisco UCS 6454/64108 Fabric Interconnect	Cisco UCS 6454/64108 Fabric Interconnect
GPU	For C240 M5: 2 x NVIDIA TESLA V100 with 32 GB of memory each or up to 6 x NVIDIA T4 For C480 M5: 4 x NVIDIA TESLA v100 with 32 GB of memory each or 4 x NVIDIA T4	For C240 M5: 2 x NVIDIA TESLA V100 with 32 GB of memory each or up to 6 x NVIDIA T4 For C480 M5 ML: 8 x NVIDIA TESLA V100 with 32 GB of memory each and with NVLink	

Table 3.   Cisco UCS Integrated Infrastructure for Big Data and Analytics configuration options for object storage

Servers	Cisco UCS S3260 with Single Node	Cisco UCS S3260 with Dual Node
CPU	2 x 2nd Gen Intel Xeon Scalable Processor 6230R (2 x 26 cores, 2.1 GHz)	2 x 2nd Gen Intel Xeon Scalable Processor 6230R (2 x 26 cores, 2.1 GHz)
Memory	12 x 32GB 2666 MHz (384 GB)	6 x 32GB 2666 MHz (192 GB) per node
Boot	2 x 1.6TB SATA Boot SSDs	4 x 1.6TB SATA Boot SSDs
Storage	UCS S3260, 4 rows of drives – 56 x 14 TB per node	UCS S3260, 2 rows of drives – 28 x 14 TB per node
Virtual Interface Card (VIC)	25 Gigabit Ethernet (Cisco UCS VIC 1455) or 40/100 Gigabit Ethernet (Cisco UCS VIC 1495)	25 Gigabit Ethernet (Cisco UCS VIC 1455) or 40/100 Gigabit Ethernet (Cisco UCS VIC 1495)
Storage controller	Cisco UCS S3260 dual RAID controller	Cisco UCS S3260 dual RAID controller
Network connectivity	Cisco UCS 6332 Fabric Interconnect or Cisco UCS 6454/64108 Fabric Interconnect	Cisco UCS 6332 Fabric Interconnect or Cisco UCS 6454/64108 Fabric Interconnect
Cache	2x UCS S3260 M5 SIOC 2TB NVMe	1x UCS S3260 M5 SIOC 2TB NVMe/node

Cisco UCS and Cisco ACI platforms

This section describes the Cisco UCS and Cisco ACI products and their features.

Cisco Unified Computing System

Cisco UCS is a next-generation data center platform that unites computing, networking, storage access, and virtualization resources into a cohesive system designed to reduce TCO and increase business agility. The system integrates a low-latency, lossless 10 and 40 Gigabit Ethernet unified network fabric with enterprise-class, x86-architecture servers. The system is an integrated, scalable, multichassis platform in which all resources participate in a unified management domain.

Cisco UCS 6300 and 6400 Series Fabric Interconnects

The Cisco UCS 6300 Series Fabric Interconnects are a core part of Cisco UCS, providing both network connectivity and management capabilities for the system. The Cisco UCS 6300 Series offers line-rate, low-latency, lossless 40-Gigabit Ethernet, Fibre Channel over Ethernet (FCoE), and Fibre Channel functions, as well as unified ports capable of either Ethernet or Fibre Channel operation.

The Cisco UCS 6454 Fabric Interconnect offers line-rate, low-latency, lossless 10, 25, 40, and 100 Gigabit Ethernet, FCoE, and Fibre Channel functions.

The Cisco UCS 6300 Series and 6454 Fabric Interconnects provide the management and communication backbone for the Cisco UCS B-Series Blade Servers and C-Series Rack Servers. All servers attached to the Cisco UCS fabric interconnects become part of a single, highly available management domain.

Cisco UCS C240 M5 Rack Server

The Cisco UCS C240 M5 Rack Server (Figure 12) is a dual-socket, 2-rack-unit (2RU) server with the latest second-generation Intel Xeon and Intel Xeon Scalable processors, 24 DIMM slots for DDR4 DIMMs, one dedicated internal slot for a 12-Gbps SAS storage controller card, and up to 26 internal SFF drives or up to 12 front-facing internal LFF drives. It offers industry-leading performance and expandability for a wide range of storage and I/O-intensive infrastructure workloads, such as big data, analytics, and collaboration. In addition, the server has two modular M.2 cards that can be configured for boot. A modular LAN-on-motherboard (mLOM) slot supports dual 40 Gigabit Ethernet network connectivity with the Cisco UCS VIC 1387.

Figure 12.   Cisco UCS C240 M5 Rack Server

Cisco UCS C4200 M5 Rack Server

The Cisco UCS C4200 Rack Server (Figure 13) is Cisco's densest computing solution, with up to four Cisco UCS C125 M5 Rack Server nodes in 2RU of rack space. This density makes it well suited for the network edge or in the data center for scale-out applications. The C125 M5 node contains AMD EPYC processors, up to 2 TB of memory, and up to six SAS/SATA drives or four plus two Non-Volatile Memory Express (NVMe) drives. Additional Secure Digital (SD) or M.2 storage modules can be used as boot devices or additional storage. Fourth-generation VICs and an OCP 2.0 mezzanine slot offer exceptional levels of performance, flexibility, and I/O throughput to run your applications.

The bridge to possible

Figure 13. Cisco UCS C4200 Rack Server

Cisco UCS S3260 Storage Server

The Cisco UCS S3260 Storage Server (Figure 14) is a modular, storage server with dual server nodes. It is optimized for large data sets used in scenarios such as big data, cloud, object-storage, video surveillance, and content delivery environments.

The S3260 server helps achieve the highest levels of data availability and performance. With a dual-node capability that is based on the latest second-generation Intel Xeon and Intel Xeon Scalable processors, it offers up to 840 TB of local storage in a compact 4RU form factor. Network connectivity is provided with dual-port 40-Gbps nodes in each server.

Figure 14. Cisco UCS S3260 Storage Server

Cisco UCS C480 ML M5 Rack Server

The Cisco UCS C480 ML M5 Rack Server (Figure 15), is a purpose-built 4RU server for deep-learning environments. It supports the latest second-generation Intel Xeon and Intel Xeon Scalable processors and 8 NVIDIA Tesla V100 32-GB Tensor Core GPUs with NVLink interconnects. It supports up to 3 TB of DDR4 memory in 24 slots, up to 24 SFF hot-swappable SAS/SATA SSDs and HDDs, up to 6 PCIe NVMe disk drives, and up to 2 internal M.2 drives.

Figure 15. Cisco UCS C480 ML M5 purpose-built deep-learning server

Cisco Intersight cloud-based management

Cisco Intersight software enables IT operations managers to claim devices across different sites, presenting these devices in a unified dashboard (Figure 16). The adaptive management of the Cisco Intersight software provides visibility and alerts to firmware management, showing compliance across managed Cisco UCS domains as well as proactive alerts for upgrade recommendations. Integration with the Cisco Technical Assistance Center (TAC) allows the automated generation and upload of tech support files from the customer.

Figure 16. Features of Cisco Intersight and how it fits in the infrastructure

ılıılı
CISCO

The bridge to possible
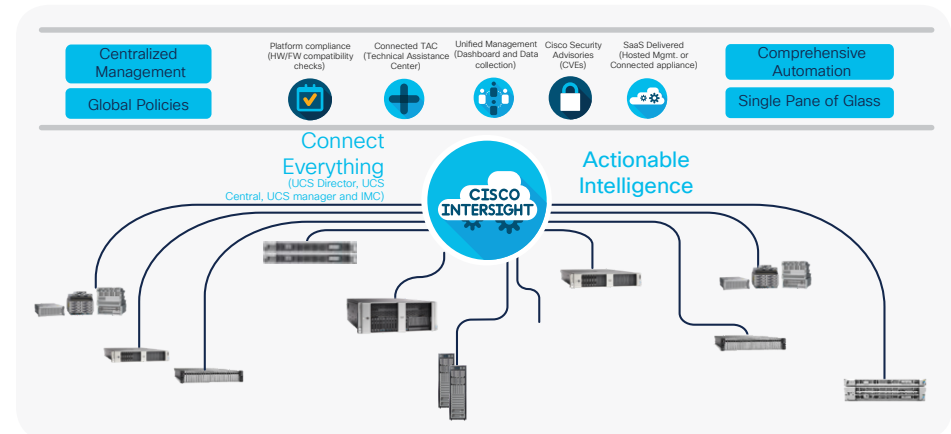
Cisco Application Centric Infrastructure

The Cisco ACI fabric uses Cisco Nexus 9000 Series Switches with the Cisco Application Policy Infrastructure Controller (APIC) running in the Cisco ACI leaf-and-spine fabric mode. These switches form a "fat tree" network by connecting each leaf node to each spine node; all other devices connect to the leaf nodes. The APIC manages the Cisco ACI fabric. Figure 17 provides an overview of the Cisco ACI leaf-and-spine fabric.

Figure 17.   Cisco ACI spine-and-leaf fabric
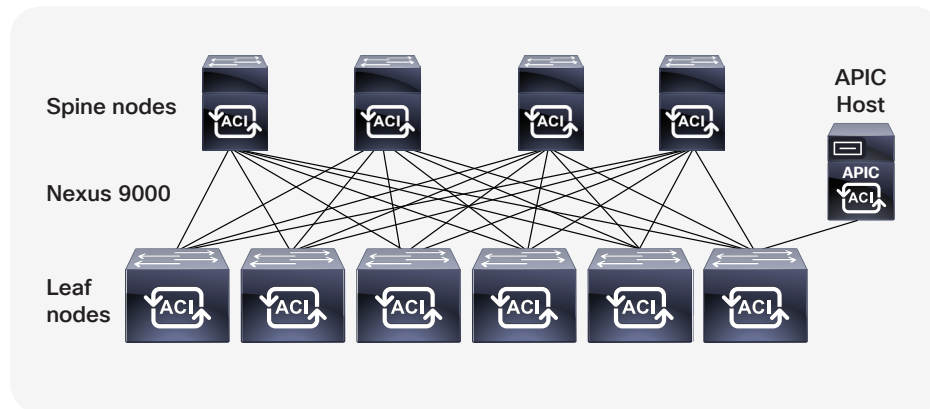
The Cisco ACI fabric provides consistent low-latency forwarding across high-bandwidth links (40 and 100 Gbps). Traffic with the source and destination on the same leaf switch is handled locally, and all other traffic travels from the ingress leaf to the egress leaf through a spine switch. Although this architecture appears as two hops from a physical perspective, it is actually a single Layer 3 hop because the fabric operates as a single Layer 3 switch.

Cisco Application Centric Infrastructure (ACI) and Cisco Unified Computing System (Cisco UCS), working together, can cost-effectively scale capacity, and deliver exceptional performance for the growing demands of big data processing, AI, and storage workflows. For larger clusters and mixed workloads, Cisco ACI uses intelligent, policy-based flowlet switching and packet prioritization to deliver:

- Centralized Management for the entire Network
- Dynamic load balancing
- Dynamic Packet Prioritization
- Multi-Tenant and Mixed Workload Support
- Deep Telemetry

Centralized Management for the Entire Network

Cisco ACI treats the network as a single entity rather than a collection of switches. It uses a central controller to implicitly automate common practices such as Cisco ACI fabric startup, upgrades, and individual element configuration. The Cisco Application Policy Infrastructure Controller (Cisco APIC) is the unifying point of automation and management for the Cisco Application Centric Infrastructure (ACI) fabric. This architectural approach dramatically increases the operational efficiency of networks, by reducing the time and effort needed to make modifications to the network and, also, for root cause analysis and issue resolution.

Dynamic Load Balancing

Cisco's Application Centric Infrastructure is not only aware of the congestion points but is able to make dynamic decisions on how the traffic is switched/routed. This could be new flows that are about to start or existing long flows which could benefit from moving to a less congested route. Dynamic load balancing takes care of these decisions at run time automatically and helps utilize the links optimally – both the healthy and the congested links. This is useful in both congested link scenarios and scenarios where there are link failures. Even when there is no congestion this will maintain close to optimal distribution of traffic across the spines.

Dynamic Packet Prioritization

Dynamic Packet Prioritization (DPP), prioritizes short flows higher than long flows; a short flow is less than approximately 15 packets. Short flows are more sensitive to latency than long ones. Small and urgent data workloads, such as database queries, may suffer processing latency delays because larger data sets are being sent across the fabric ahead of them. This approach presents a challenge for instances in which database queries require near-real-time results.

ılıılı
CISCO

The bridge to possible

Dynamic Packet Prioritization can improve overall application performance. Together these technologies enable performance enhancements to applications, including Big Data workloads.

Multi-Tenant and Mixed Workload Support

Cisco ACI is built to incorporate secure multi-tenancy capabilities. The fabric enables customers to host multiple concurrent clusters (Big Data/AI/ Object Store) on a shared infrastructure. Cisco ACI provides the capability to enforce proper isolation and SLA's for workloads of different tenants. These benefits extend beyond just Big Data, AI or Object store workloads – Cisco ACI allows the same cluster to run a variety of different application workloads including containers, micro-services, with the right level of security and SLA for each workload.

Deep Telemetry of Tenant and Application Network

One of the core design principles behind Cisco ACI is to provide complete visibility into the infrastructure – physical and virtual. Cisco APIC is designed to provide application and tenant health at a system level by using real-time metrics, latency details, atomic counters, and detailed resource consumption statistics.

If your application is experiencing performance issues, you can drill down easily into the lowest possible granularity – be it at a switch level, line card level, or port level.

The holistic approach to correlate virtual and physical and tie that intelligence to an application or tenant level ensures that troubleshooting becomes extremely simple across your infrastructure, through a single pane of glass.

Cisco Nexus 9508 Switch

The Cisco Nexus 9508 Switch (Figure 18) offers a comprehensive feature set, high resiliency, and a broad range of high-density Ethernet line cards to meet the most demanding requirements of enterprise, service provider, and cloud data centers.

Organizations can improve performance and efficiency with the Cisco Nexus 9508 Switch. It provides high-density 1, 10, 25, 40, 50, and 100 Gigabit Ethernet in a compact, 13RU modular chassis. This versatile modular switch is designed for high-density, end-of-row, and high-performance aggregation-layer deployments in both traditional and Cisco ACI enabled data centers.

Figure 18.   Cisco Nexus 9508 Switch

Cisco Nexus 9336 Switch

Powered by cloud-scale technology, the Cisco Nexus 9336 Switch (Figure 19) offers flexible port speeds supporting 36 ports with 1, 10, 25, 40, and 100 Gbps in a compact 1RU form factor. Designed to meet the changing needs of data centers, big data applications, and automated cloud environments, this powerful switch supports both Cisco ACI and standard Cisco Nexus environments (NX-OS mode). This capability provides access to industry-leading programmability (Cisco NX-OS) and the most comprehensive automated, policy-based, systems-management approach (Cisco ACI).

Figure 19.   Cisco Nexus 9336 Switch

ılıılı
CISCO

The bridge to possible

For more information

For additional information, see the following resources:

- To find out more about Cisco UCS big data solutions, see https://www.cisco.com/go/bigdata.

- To find out more about Cisco UCS big data validated designs, see https://www.cisco.com/go/bigdata_design

- To find out more about Cisco UCS AI/ML solutions, see https://www.cisco.com/go/ai-compute

- To find out more about Cisco ACI solutions, see https://www.cisco.com/go/aci

- To find out more about Cisco validated solutions based on Software Defined Storage at https://www. cisco.com/c/en/us/solutions/data-center-virtualization/software-defined-storage-solutions/index.html.

Conclusion

When building an infrastructure to enable this modernized architecture which could scale to thousands of nodes, operational efficiency can't be an afterthought.

To bring in seamless operation of the application at this scale, one needs:

- Infrastructure automation of Cisco UCS servers with service profiles and Cisco Data Center network automation with application profiles with Cisco ACI

- Centralized Management and Deep telemetry and Simplified granular trouble-shooting capabilities and Multi-tenancy allowing application workloads including containers, micro-services, with the right level of security and SLA for each workload

- Cisco UCS with Intersight and Cisco ACI can enable this cloud scale architecture deployed and managed with ease