Intel Xeon Scalable 2nd Generation Processor Recommendations for Cisco UCS M5 Servers

April 2021

Contents

Choose recommended Intel Xeon Cascade Lake and Cascade Lake Refresh processors on Cisco UCS M5 servers for industry-standard enterprise solutions.

Executive summary

This document provides guidance to help you select the best 2nd generation Intel® Xeon® Scalable processors for Cisco UCS® M5 servers for enterprise solutions. The recommendations in this document use a variety of available metrics, such as CPU core count, clock speed, industry-standard SPEC CPU2017 benchmark results, and software licensing details.

Scope

Choosing the CPU for enterprise applications is not an easy task, whether you use a simple back-of-the envelope calculation to determine your needs or perform exhaustive capacity planning with analytical model simulations for workloads and their future growth. Often, lack of details (or inaccurate information) about the input data and the workload type leads to uncertain estimations about the rightsizing of the hardware (CPU type, core count and clock speed, memory capacity, etc.), leading to suboptimal resource utilization.

This document provides default processor and memory recommendations when only a limited set of inputs is available. Processor and memory selections can be fine-tuned later when more detailed workload information is available. This document provides guidance only for the selection of Intel Cascade Lake CPUs / Cascade Lake Refresh CPUs and the associated default memory configurations. For other metrics, such as storage requirements etc., please consult the vendor documentation for the relevant enterprise software.

2nd generation Intel Xeon Scalable processors – Cascade Lake (CLX-SP)

Intel Xeon Scalable processors provide a foundation for powerful data center platforms with an evolutionary leap in agility and scalability. Disruptive by design, this innovative processor family supports new levels of platform convergence and capabilities across computing, storage, memory, network, and security resources.

Cascade Lake (CLX-SP) is the code name for the next-generation Intel Xeon Scalable processor family that is supported on the Purley platform serving as the successor to Skylake SP. These chips support up to eight-way multiprocessing, use up to 28 cores, incorporate a new AVX512 x86 extension for neural-network and deep-learning workloads, and introduce persistent memory support. Cascade Lake SP based chips are manufactured in an enhanced 14-nanometer (14-nm++) process and use the Lewisburg chip set. Cascade Lake SP based models are branded as the Intel Xeon Bronze, Silver, Gold, and Platinum processor families (Figure 1).

Cascade Lake is set to run at higher frequencies than the current and older generations of the Intel Xeon Scalable products. Additionally, it supports Intel Optane™ DC Persistent Memory. The chip is a derivative of Intel's existing 14-nm technology (first released in 2016 in server processors). It offers 26 percent performance improvement compared to the earlier technology while maintaining the same level of power consumption.

Figure 1.
2nd generation Intel Xeon Scalable processors

Main features of Intel Cascade Lake processors

The new Cascade Lake processors incorporate a performance-optimized multichip package to deliver up to 28 cores per CPU and up to 6 DDR4 memory channels per socket. They also support Intel Optane DC Persistent Memory and are especially valuable for in-memory computing SAP workloads.

Cascade Lake delivers additional features, capabilities, and performance to our customers:

- Compatible with the Purley platform through a six-channel drop-in CPU
- Improved core frequency through speed-path and processing improvements
- Support for DDR4-2933 with two DIMMs Per Channel (DPCs) on selected SKUs and 16-Gbps devices
- Scheduler improvements to reduce load latency
- Additional capabilities such as Intel Optane DC Persistent Memory Module (DCPMM) support
- Intel Deep Learning Boost with Vector Neural Network Instructions

Cascade Lake Refresh (CLXR-SP) processors

Cascade Lake Refresh Intel Xeon Gold 6200 Series CPUs are optimized to deliver higher performance per core plus larger cache for latency sensitive workloads while the Cascade Lake Refresh Intel Xeon Gold 5200 Series CPUs provide balanced performance with optimized cores, frequency and price point and the Cascade Lake Refresh Intel Xeon Silver 4200 Series CPUs provide a 6% - 8% performance (SIR) improvement. **Cascade Lake Refresh Intel Xeon Gold 6200 Series CPUs also provide enhanced digital experiences without requiring software optimization or enablement**.

Recommended Intel Cascade Lake and Cascade Lake Refresh processors for data center solutions on Cisco UCS M5 servers

Table 1 provides processor recommendations for various workloads and memory configurations on Cisco UCS M5 servers.

Table 1. Recommended Cascade Lake and Cascade Lake Refresh processors for various workloads and associated memory configurations (A, B, C, D, E – see note below) for dual-socket servers

Part number	Cores	Speed (GHz)	Power (Watts)	Memory (MHz)1	SPEC CPU2017 Integer Rate[2]	VSI	VDI	Oracle	Microsoft SQL Server	SAP	DP	Hadoop	Splunk	Scale Out Storage	Microsoft Azure Stack
UCS-CPU-I4210R	10	2.4	100	2400[3]	114	A					E	A	A		
UCS-CPU-I4214R	12	2.4	100	2400[3]	148						E		A	A	E
UCS-CPU-I4215R[4]	8	3.2	130	2400[3]	109				B						
UCS-CPU-I5218R	20	2.1	125	2666	223	B	B	C			E	A	A	A	E
UCS-CPU-I5220R[4]	24	2.2	150	2666	272	B	C								E
UCS-CPU-I5222	4	3.8	105	2666	65			B							
UCS-CPU-I6230R	26	2.1	150	2933	287	B					E	B	A		E
UCS-CPU-I6234	8	3.3	130	2933	111			C							
UCS-CPU-I6238R[4]	28	2.2	165	2933	306										E
UCS-CPU-I6240R[4]	24	2.4	165	2933	287		D			E	E				
UCS-CPU-I6242R[4]	20	3.1	205	2933	282		B	C	C		E				
UCS-CPU-I6246R[4]	16	3.4	205	2933	248			C	C						
UCS-CPU-I6248R[4]	24	3.0	205	2933	311		B	D							

Part number	Cores	Speed (GHz)	Power (Watts)	Memory (MHz)1	SPEC CPU2017 Integer Rate[2]	VSI	VDI	Oracle	Microsoft SQL Server	SAP	DP	Hadoop	Splunk	Scale Out Storage	Microsoft Azure Stack
UCS-CPU-I6258R[4]	28	2.7	205	2933	339					E					
UCS-CPU-I8276	28	2.2	165	2933	301					E					

Note:

- 1: Apache Pass (AEP) ready, unless specifically called out
- 2: SPEC CPU2017 Integer Rate for two-socket system
- 3: No AEP
- 4: Not Supported on the S3260 platform (Supported on B200, C220, C240)
- A = 192 GB (12 x 16 GB), B = 384 GB (12 x 32 GB), C = 768 GB (12 x 64 GB), and D = 1536 GB (12 x 128 GB); for E, see the discussions of the specific workloads later in this document.

Recommended default memory configuration for data center solutions on Cisco UCS M5 servers

Table 2 provides recommendations for default memory configurations on Cisco UCS M5 servers.

Table 2. Recommended default memory configurations for Cascade Lake and Cascade Lake Refresh processors on Cisco UCS M5 servers

Intel Xeon processors	Intel Xeon Platinum 8200	Intel Xeon Gold 6200	Intel Xeon Gold 5200	Intel Xeon Silver 4200
Default memory configurations for dual-socket servers	768 GB (12 x 64 GB)	384 GB (12 x 32 GB)	384 GB (12 x 32 GB)	192 GB (6 x 32 GB)

Note: Memory specifications in Table 1 are based on the default memory configurations specified in Table 2, unless specified differently for a workload in the sections that follow.

Intel Cascade Lake and Cascade Lake Refresh processor recommendations by workload

This section discusses the processor recommendations for particular types of workload.

Virtual Server Infrastructure (VSI)

For virtual server infrastructure (VSI), virtual machines can be of different types: general purpose, compute optimized, high performance, etc., based on use cases. For example, testing and development, smaller databases, and lower-traffic web servers on **general-purpose** virtual machines; medium-size databases and application servers on **compute-optimized** virtual machines; and batch processing and analytics on **high-performance** virtual machines. To address these various use cases, five CPUs are recommended: UCS-CPU-I4210R (Intel Xeon Silver 4210R) with 10 cores, UCS-CPU-I5218R (Intel Xeon Gold 5218R) with 20 cores, UCS-CPU-I5220R (Intel Xeon Gold 5220R) with 24 cores, UCS-CPU-I6230R (Intel Xeon Gold 6230R) with 26 cores and UCS-CPU-I6238R (intel Xeon Gold 6238R) with 28 cores.

Consider an example to see how the recommended CPUs can be used to provide a range of virtualization options. Four virtual CPUs (vCPUs) per core are assumed, and the utilization of CPUs and RAMs are capped at 80% to provide headroom of 20% for safety. Four types of virtual machines—small.VM, medium.VM, large.VM, and xlarge.VM—with the configurations in Table 3 are considered. Table 4 shows the number of virtual machines available for the default memory capacity.

Table 3. Example showing virtual machine types

Virtual machine type	small.VM	medium.VM	large.VM	xlarge.VM	Total	mean.VM
% of VMs in Environment	40%	30%	20%	10%	100%	
# of VMs in Environment	400	300	200	100	1000	
vCPU per virtual machine	2	4	8	16		5.2
# of Virtual CPUs	800	1200	1600	1600	5200	
Memory per virtual machine (GB)	8	16	32	64		20.8
Total Memory (GB)	3200	4800	6400	6400	20,800	

Table 4. Example showing the number of virtual machines with 80% utilization for a 2-socket CPU server with the minimum recommended memory

Processor part number and quantity	Cores	Memory (GB)	small.VM Number of virtual machines[1]	medium.VM Number of virtual machines[1]	large.VM Number of virtual machines[1]	xlarge.VM Number of virtual machines[1]	mean.VM Number of virtual machines[1]
UCS-CPU-I4210R x 2	20	192 (12 x 16 GB)	19	9	4	2	7
UCS-CPU-I5218R x 2	40	384 (12 x 32 GB)	38	19	9	4	14
UCS-CPU-I5220R x 2	48	384 (12 x 32 GB)	38	19	9	4	14
UCS-CPU-I6230R x 2	52	768 (12 x 64 GB)	76	38	19	9	29
UCS-CPU-I6238R x 2	56	768 (12 x 64 GB)	76	38	19	9	29

[1] The number of virtual machines is limited by memory capacity.

Virtual Desktop Infrastructure (VDI)

Choosing the right CPU for Virtual Desktop Infrastructure (VDI), which includes Microsoft Windows 10 Virtual Desktop Infrastructure (VDI), virtual applications, and Microsoft Remote Desktop Server Host (RDSH) user sessions, requires you to consider several factors, including feature sets and hardware requirements. Proper processor selection is crucial. For VDI solutions, given adequate memory, storage performance, and network bandwidth, the CPU is the element that determines user density and performance. Different user types benefit from different processor and memory configurations.

During the design phase of a VDI deployment, you should carefully consider the selection of the CPU. Ideally, you should deploy a Proof-of-Concept (PoC) system before you order the hardware, but this may not always be possible. For such cases, the VDI team has conducted testing using Login Virtual Session Indexer (Login VSI) to provide benchmark data for baseline sizing. These Cisco® Validated Designs can be found on the Design Navigator page at Cisco.com. Sizing tools for VDI deployed on the Cisco Unified Computing System™ (Cisco UCS) and Cisco HyperFlex™ platforms are available to Cisco and partner account teams to identify starting-point configurations based on customer input.

Based on Login VSI results and VDI solution engineering team experience, we recommend three Intel Xeon processors as the starting point CPU for task worker, knowledge worker, and power user systems for Cisco UCS M5 and Cisco HyperFlex M5 servers. Additional factors relevant to the choice of the processor include the following:

- The current build of Microsoft Windows 10 is 50 percent more graphics intensive than Windows 7.
- Graphics elements perform better with higher-frequency processors.
- More graphics content translates to higher CPU use per desktop.

The base memory for VDI knowledge worker configurations is 768 GB per server. To accommodate growing demand for memory per virtual machine, a 12 x 64 GB DIMM configuration is recommended to allow expansion. Tables 5 and 6 show starting-point virtual machine and server configurations for the different user types.

Table 5. Example showing virtual machine types

Virtual machine type	Task worker	Knowledge worker	Power user
vCPU per virtual machine	1	2	4
Memory per virtual machine	2 GB	4 to 8 GB	8 to 16 GB

Table 6. Example showing the number of virtual machines with 75 to 80 percent utilization for a 2-socket CPU server with the minimum recommended memory

Processor part number and quantity	Cores	Memory (GB)	Number of task workers	Number of knowledge workers	Number of power users
UCS-CPU-I5218R x 2	36	384 (12 x 32 GB)	175 to 180		
UCS-CPU-I5220R x 2	40	768 (12 x 64 GB)		90 to 190	
UCS-CPU-I6240R x 2	40	1.5 TB (24 x 64 GB)			95 to 185

Note: For VDI virtual machines for professional graphics application virtualization, other Intel Xeon Scalable processors will be the processors of choice based on user and application requirements. For light and medium class users, start with UCS-CPU-I6242R. For heavy/rendering class users, start with UCS-CPU-I6248R. For these use cases, GPUs are required along with the high frequency CPUs.

Oracle

Oracle Database normally is licensed by customers on a per-processor-core basis. The list price per core for Oracle is already substantial for a single-instance database and can double if Oracle RAC is deployed. In addition, 90 percent of customers also license Oracle Partitioning for such use cases. Therefore, the majority of customers demand core-optimized processors with low core counts but high clock rates for high performance. For most Oracle deployments, you should lead with the previous generation of Cascade lake processors which offer 4-core Intel Xeon Gold 5222, or the 8-core Intel Xeon Gold 6234 processor which provides the best performance for the lowest core count. The Oracle licensing savings will make up for the higher cost of these processors. These processors are a particularly good choice for the Cisco UCS B200 M5 Rack Servers, which historically account for 80 percent or more of the Cisco servers and blades used for Oracle solutions.

If the customer however insists on using the Cascade Lake "R" processors, then the choices are limited to the list in Table 7 below.

Alternatively, larger enterprise customers may have purchased an enterprise or sitewide license from Oracle. This license allows the customer to run as many instances of Oracle Database and Oracle RAC across any number of servers and processor cores, so the number of processor cores is not a concern in these configurations. Instead, raw performance will be a stronger factor in the CPU decision. In such cases, you should lead with the UCS-CPU-I6248R, a 24-core processor. The higher number of cores helps achieve higher performance levels and the 28 core alternatives really do not buy the customer much benefit relative to the potential additional cost of Oracle licenses. Again, UCS-CPU-I6248R (or higher core count CPU) makes sense only in the presence of a sitewide license; otherwise, the per-core license cost can be overwhelming.

Approximately 32 GB memory per core can be considered as a default for online transaction processing (OLTP), data warehouse, and business analytics use cases. Table 7 shows the calculated memory and the closest recommended memory. The actual memory configured may be different if customer-specific requirements are known.

Table 7. Recommended memory for Oracle use cases

Processor part number and quantity	Number of cores	Calculated memory at 32 GB per core (GB)	Recommended memory (GB)
UCS-CPU-5222 x 2	8	256	12 x 32 GB = 384 GB
UCS-CPU-6234 x 2	16	512	12 x 64 GB = 768 GB
UCS-CPU-I6242R x 2	40	1280	12 X 64 GB = 768GB
UCS-CPU-I6246R x 2	32	1024	12 x 64 GB = 768 GB
UCS-CPU-I6248R x 2	48	1536	12- x 128 GB = 1.5TB

Microsoft SQL Server

Microsoft SQL Server is more widely deployed than Oracle and on average tends to have a smaller-size database than the Oracle solution. Larger databases on SQL Server running on Linux may be a consideration for customers who may be trying to improve their economics through migration from alternate enterprise databases. However, given that the pricing of a SQL Server license in most cases is still per core, you should lead with the UCS-CPU-I4215R (Intel Xeon Silver 4215R) 8-core processor. The UCS-CPU-I6242R (Intel Xeon Gold 6242R) 20-core processor can be an alternative if the customer is seeking higher total performance and is not deterred by the higher licensing cost required by the higher core counts. Sitewide licenses from Microsoft are less common. Table 8 shows the recommended memory for various SQL Server use cases.

Table 8. Recommended memory for Microsoft SQL Server use cases

Processor part number and quantity	Number of cores	Calculated memory at 32 GB per core (GB)	Recommended memory (GB)
UCS-CPU-I4215R x 2	16	512	12 x 32 GB = 384 GB
UCS-CPU-I6242R x 2	40	1280	12 x 64 GB = 768 GB
UCS-CPU-I6246R x 2	32	1024	12 x 64 GB = 768 GB

SAP

SAP HANA is an in-memory, column-oriented, relational database management system developed and marketed by SAP SE. Its dual purpose is to function as a standard transactional database and as an advanced analytics database that includes extract, transform, and load (ETL) capabilities.

SAP has partnered with hardware vendors such as Cisco to provide the hardware infrastructure to support SAP HANA analytics and transactional workloads. In that partnership, SAP has also defined two design and delivery models: one based on a preconfigured certified appliance and the other based on SAP Tailored Datacenter Integration (TDI).

The appliance model is based on a preconfigured, tested, and certified design with fixed delivery and break/fix services provided by the hardware vendor. SAP has mandated the use of the two top Intel CPUs for all appliance configurations. Cisco recommends UCS-CPU-I8276 (base and large [L] versions) with the Intel Xeon Platinum 8276 CPU to meet the appliance requirements, as well as to ensure performance for customers using the TDI alternative. For customers with significantly heavy

workloads, the UCS-CPU-I8276 can be replaced with the UCS-CPU-I8280L (base and large [L] versions) with the Intel Xeon Platinum 8280L CPU for more cores. This higher-level CPU has heavy power requirements.

In the TDI delivery model, the customer can select a lower-level CPU based on an analysis of each customers use case and performance needs. The UCS-CPU-I8276 processor group is the primary choice for TDI to avoid performance issues, and the alternative UCS-CPU-I6240R (Intel Xeon Gold 6240R base and L versions) has an appropriate number of cores while providing sufficient performance overhead to meet the demands of most 4 socket server SAP HANA workloads. For 2 socket server implementations including Cisco Hyperflex systems the alternative UCS-CPU-I6258R (Intel Xeon Gold 6258R base and L versions) is recommended as there is sufficient frequency to maximize the vHANA capacity of the Cisco Hyperflex and UCS systems.

Ultimately, the final CPU selection in an SAP HANA TDI environment depends on the results of the SAP HANA sizing exercise. The preceding recommended CPUs are good candidates for most SAP HANA workloads. Lighter SAP HANA workloads may take advantage of even lower-level CPU for additional cost savings.

Memory selection for the SAP HANA server is based exclusively on evaluation of customer database expectations and sizing and is bounded by SAP requirements for appliance and TDI implementations. The customer database will reside in memory in its entirety and will almost always require either 64-GB DDR4 LRDIMMs and RDIMMs or 128-GB DDR4 RDIMMs. The memory slots can be fully populated, or half populated. Mixed memory configurations are not encouraged, and unbalanced channel configurations are not supported.

CPU and memory options for appliances are listed in the SAP HANA Product Availability Matrix (PAM) located at https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/appliances.html.

Data protection

Cohesity

Cisco and Cohesity jointly recommends the CPU and memory configuration based on thorough performance characterization that is needed to perform optimal backup, recovery and various scale-out NAS workloads. These recommendations are hard limits and cannot be changed and are summarized in Table 9.

Table 9. Cohesity: CPU and Memory selection summary

Configurations	Processor part number and quantity	Total cores	Memory (GB)
Cohesity on Cisco UCS C220 M5 (24 TB)	UCS-CPU-I4210R x 2	20	64 (2x 32 GB)
Cohesity on Cisco UCS C220 M5 (36 TB)	UCS-CPU-I4210R x 2	20	128 (4 x 32 GB)
Cohesity on Cisco UCS C240 M5 (48 TB)	UCS-CPU-I6242R x 2	40	128 (4 x 32 GB)
Cohesity on Cisco UCS C240 M5 (120 TB)	UCS-CPU-I6242R x 2	40	128 (4 x 32 GB)
Cohesity on Cisco UCS S3260 M5 (210 TB)	UCS-CPU-I6240 x 2	48	256 (8x 32 GB)
Cohesity on Cisco UCS S3260 M5 (420 TB)	UCS-CPU-I6240 x 2	48	256 (8x 32 GB)

Commvault

Commvault ScaleProtect solutions use the scale-out Commvault HyperScale, which achieves performance and capacity growth by adding UCS compute and storage resources through appropriately configured building blocks (nodes), published in the Cisco Validated Designs (CVDs). This approach allows the gradual and incremental addition of performance (CPU, memory, and network bandwidth) and capacity (hard-disk drives) as the need arises.

For Commvault ScaleProtect and Commvault MediaAgent solutions UCS-CPU-I4210R is recommended. Typical configurations and memory requirements are listed in Table 10 below.

Table 10. Commvault: CPU and Memory selection summary

Configurations	Processor part number and quantity	Total cores	Memory (GB)	Capacity (TB)
CVLT ScaleProtect UCS C220 M5 4TB (6TB, 8TB, 10TB, 12TB)	UCS-CPU-I4210R x 2	20	128 GB (8 x 16GB)	4TB x4 (6TB x4, 8TB x4, 10TB x4, 12TB x4)
CVLT ScaleProtect UCS C240 M5 6TB (8TB, 10TB, 12TB, 14TB)	UCS-CPU-I4210R x 2	20	256 GB (8 x 32 GB)	6TB x12 (8TB x12, 10TB x12, 12TB x12, 14TB x12)
CVLT ScaleProtect UCS S3260 M5 6TB (8TB, 10TB, 12TB, 14TB, 16TB, 18TB)	UCS-CPU-I4210R x 2	20	384 GB (12 x 32 GB)	6TB x12 (8TB x12, 10TB x12, 12TB x12, 14TB x12, 16TB x12, 18TB x 12)
CVLT ScaleProtect UCS S3260 M5 6TB 2N (8TB, 10TB, 12TB, 14TB, 16TB, 18TB)	2 x (UCS-CPU-I4210R x 2)	2 x 20	768 GB [2x (12 x 32GB)]	6TB x12 (8TB x12, 10TB x12, 12TB x12, 14TB x12, 16TB x12, 18TB x12)
CVLT MediaAgent C240 M5 32TB X	UCS-CPU-I4210R x 2	20	32 GB (2 x 16 GB)	4TB x12
CVLT MediaAgent C240 M5 49TB S	UCS-CPU-I4210R x 2	20	64 GB (4 x 16 GB)	6TB x12
CVLT MediaAgent S3260 M5 80TB M	UCS-CPU-I4210R x 2	20	128 GB (8 x 16 GB)	8TB x14
CVLT MediaAgent S3260 M5 174TB L	UCS-CPU-I4210R x 2	20	128 GB (8 x 16 GB)	8TB x28
CVLT MediaAgent S3260 M5 283TB XL	UCS-CPU-I4210R x 2	20	128 GB (8 x 16 GB)	8TB x42

Veeam

For Veeam Data Protection (DP) solutions using Cisco UCS C240 M5 servers, UCS-CPU-I4214R (Intel Xeon Silver 4214R) with 12 cores and a speed of 2.4 GHz is recommended. This CPU is appropriate for smaller to medium-size backup targets with storage up to 144 TB. For Veeam solutions on Cisco UCS S3260 M5 servers, UCS-CPU-I5218R (Intel Xeon Gold 5218R) with 20 cores and a speed of 2.1 GHz is recommended. For the high capacity provided by Cisco UCS S3260 M5 servers (max. capacity - 874 TB) and fast restore performance, CPU with higher performance, UCS-CPI-I6240 (Intel Xeon Gold 6240) may be required as Veeam software is CPU intensive. Typical configurations are summarized in Table 11 below.

For the Veeam Availability on Cisco HyperFlex solution (VACH), the UCS-CPU-I5218R (Intel Xeon Gold 5218R) processor with 20 cores and a speed of 2.1 GHz is recommended. The VACH solution is CPU intensive, given that it has multiple software layers (Veeam, Microsoft Windows, Red Hat Linux, Cisco HyperFlex HX Data Platform, and VMware ESXi) stacked to provide an integrated hyperconverged solution. Also, additional CPU cores are required to handle the end-user virtual machines for various data lab use cases that use backup copies for testing and development and disaster-recovery scenarios.

Table 11. Veeam: CPU and Memory selection summary

Configurations	Processor part number and quantity	Total cores	Memory (GB)	Capacity (TB)
Veeam Data Protection C240 M5 48TB	UCS-CPU-I4214R x 2	24	128 GB (8 x 16 GB)	4TB x12
Veeam Data Protection C240 M5 72TB	UCS-CPU-I4214R x 2	24	128 GB (8 x 16 GB)	6TB x12
Veeam Data Protection C240 M5 96TB	UCS-CPU-I4214R x 2	24	128 GB (8 x 16 GB)	8TB x12
Veeam Data Protection C240 M5 120TB	UCS-CPU-I4214R x 2	24	128 GB (8 x 16 GB)	10TB x12
Veeam Data Protection C240 M5 144TB	UCS-CPU-I4214R x 2	24	128 GB (8 x 16 GB)	12TB x12
Veeam Data Protection S3260 M5 8TB	UCS-CPU-I6240 x 2	36	256 GB (8 x 32 GB)	8TB x14, 8TB x28, 8TB x56
Veeam Data Protection S3260 M5 12TB	UCS-CPU-I6240 x 2	36	256 GB (8 x 32 GB)	12TB x14, 12TB x28, 12TB x56

Big data: Hadoop

Hadoop and most big data solutions are I/O intensive. Hadoop (including support for Hadoop by our Hadoop independent software vendor [ISV] partners) requires the ratio of CPU cores to HDDs to be greater than 1. i.e, the server needs to have at least as many physical cores as the number of HDDs. The CPUs recommended for Hadoop and other big data applications support these requirements and provide overall higher performance. This requirement is captured as part of our reference architecture and in Cisco Validated Designs for Hadoop. UCS-CPU-I4210R (Intel Xeon Silver 4210R) is the recommended CPU for the Cisco UCS C240 Large Form Factor (LFF), with up to 12 LFF drives; UCS-CPU-I5218R (Intel Xeon Gold 5218R) is the recommended CPU for Cisco UCS C240 M5 Small Form Factor (SFF) with 26 drives. UCS-CPU-I6230R (Intel Xeon Gold 6230R) is the recommended CPU for both Cisco UCS S3260 M5 with 28 LFF drives and all NVMe configuration with Cisco UCS C220 M5.

The ever-increasing demand for both high performances along with a lower lab footprint is driving the need for NVMe solutions. NVMe drives, which can store 8TB in an SFF is ideal to achieve better TCO with high performance and low lab footprint.

Depending on customer requirements, the cores and speed for the CPU selected may be higher than the recommendation.

Table 12 summarizes the recommendations.

Table 12. Recommended memory for Hadoop use cases

Processor part number and quantity	Number of cores	Memory (GB)
UCS-CPU-I4210R x 2	20	12 x 16 GB = 192 GB
UCS-CPU-I5218R x2	40	12 x 16 GB = 192 GB
UCS-CPU-I6230R x 2	40	12 x 32 GB = 384 GB

Big data: Splunk

A basic Splunk enterprise solution requires 12 or more CPU cores, which function at 2 GHz or greater, and 64 GB of RAM in the indexing and search tiers. However, high-performance Splunk servers equipped with 32 or more cores per server and 192 GB (or more) of RAM would be beneficial for use cases in which the daily data ingest rate is 300 GB and multiple simultaneous online and schedule searches need to be allowed. The Splunk workload requires 1200 or more random read and write I/O Operations Per Second (IOPS) with a read:write mix of 70:30 from the storage subsystem of indexing servers.

Tables 13, 14, and 15 provide guidelines for Splunk servers.

Table 13. Splunk Indexers

Configuration	CPU (dual)	Memory	Notes
Minimum requirement	UCS-CPU-I4214R	64 GB to 192 GB (32 GB DIMMs)	Small and medium deployments
Performance (recommended)	UCS-CPU-I5218R	192 GB (6 X 32 GB)	Most common
Overprovisioned	UCS-CPU-I6230R	192 GB (6 X 32 GB)	Consider this option only when using all-flash storage. In other cases, add a server in the indexer tier.

Table 14. Splunk Search Heads

Configuration	CPU (dual)	Memory	Notes
Minimum requirement	UCS-CPU-I4214R	64 GB to 192 GB (32 GB DIMMs)	Small and medium deployments
Performance (recommended)	UCS-CPU-I5218R	192 GB (6 X 32 GB)	Most common
Overprovisioned	UCS-CPU-I6230R	192 GB (6 X 32 GB)	Consider this CPU or a higher option only if this is a single (non-clustered) enterprise security search head.

Table 15. Splunk administration nodes (cluster master, license master, deployer, monitoring console, etc.)

Configuration	CPU (dual)	Memory	Notes
Minimum requirement	UCS-CPU-I4210R	64 to 192 GB (32 GB DIMMs)	Splunk services such as License Master, Cluster Master, Deployment Server, Monitoring Console, Deployer.

Scale-out object and file storage

Object-storage use cases are not performance centric. Most such deployments are focused on archival and longer-term retention. Therefore, based on the guidance from object-storage vendors and from our solution engineers, after testing various object-storage solutions in the lab, we recommend UCS-CPU-I4214R (Intel Xeon Silver 4214R) with 12 cores because it can provide sufficient processing power for an object-storage architecture. UCS-CPU-I5218R (Intel Xeon Gold 5218R) with 20 cores is recommended for IBM Cloud Object Storage (COS) because IBM COS embeds the accessor layer on the storage node. The accessor layer encrypts objects before storing them and acts as a front end for applications.

Table 16 summarizes the recommendations.

Table 16. Scale-out object and file storage: CPU and memory selection by chassis

Configuration	CPU (dual socket)	Memory
Cisco UCS C240 LFF	UCS-CPU-I5218R (IBM COS)	192 GB (6 x 32GB)
Cisco UCS S3260: Single Server Node	UCS-CPU-I4214R (Cloudian and Scality)	384 GB (12 x 32GB)
Cisco UCS S3260: Dual Server Node	UCS-CPU-I4214R	192 GB (6 x 32 GB) per server node

Microsoft Azure Stack

Azure Stack is very different from Oracle and SQL Server databases because the Azure Stack infrastructure software is provided as a free license to customers. Although some Azure services that run on top of Azure Stack are priced based on CPU use, the price is not calculated directly based on the number of physical cores in the system. Hence, the number of processor cores does not affect the cost to the customer except for the initial purchase price of the system.

In most initial-sizing scenarios, analysis will not yet have been performed to determine the optimal CPU choice. In such cases, you should start with the UCS-CPU-I5218R (Intel Xeon Gold 5218R), a 20-core processor, because this system offers a good balance of core to clock speed for configurations with less than 1024 GB per server. Most opportunities will fall in this range. If the customer requires more than 1024 GB per server, the UCS-CPU-I5220R (24-core processor) or the UCS-CPU-I6230R (26-core processor) or the UCS-CPU-I6238R (28-core processor), must be used. Furthermore, the customer may choose to invest in these processors based on either the workloads forecast to be placed on the system or the results of the Microsoft Azure Stack sizing tool, which may indicate the need for these processors. This tool provides a recommendation based on the size and number of Azure virtual machines to be run on the system. The Cisco and Azure Stack presales design team (called Jedi) can perform this calculation or the Cisco Global Virtual Engineering (GVE) team can make this recommendation.

For more information

For more information about Cisco UCS M5 servers, see https://www.cisco.com/c/dam/en/us/products/collateral/servers-unified-computing/ucs-family-poster-m5.pdf.

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 **https://www.cisco.com/go/offices**.

Printed in USA C11-742358-01 04/21