



Accelerating Data Compression on Cisco Data Intelligence Platform with Xilinx FPGAs

Increase storage capacity through data compression with Xilinx accelerator cards and NoLoad software in Cloudera Data Platform with no performance loss

Existing data landscape and architectural evolution

The exploding volume of data through myriad devices generates a vast amount and variety of data. Extracting value from newer emerging workloads through data analytics poses increasing challenges to build systems that can cater to burgeoning data storage, run computation on fewer systems, reduce power usage and the overall data-center footprint so enterprises can gain and improve meaningful business insights.

Data-compression techniques have been widely used in data centers to reduce data storage and network transmission overhead. However, algorithms with a good compression ratio are also very CPU-intensive, which adds increased processing time. Data Lake and Hadoop, by default, use Snappy for compression because it is designed to be less CPU intensive for both compression and decompression. While Snappy is a good choice to run on CPUs for compression, the downside is that its compression ratio is limited. On the other hand, we can use a different compression algorithm, such as Gzip, which has a compression ratio that is 50 percent greater than Snappy's. However, using Gzip on a CPU is very compute intensive. This is where Field-Programmable Gate Arrays (FPGAs) come into the picture.

An FPGA can offload a CPU from specific tasks, such as compression, encryption, and other functionalities, and the FPGA can be programmed

or purpose-built to do specific tasks (compression, in this case) more efficiently and in less time. This way, we can achieve higher compression of data than Snappy, for example using Gzip, in the same or a shorter amount of time by offloading Gzip to an FPGA, thus providing higher compression and freeing up the CPU for other tasks, thereby accelerating the overall workload.

Next-generation bigdata systems will be data-driven heterogeneous architectures that leverage integration of CPU/GPU/FPGA-accelerated compute. By offloading compute-heavy compression tasks to the FPGA, the CPU is freed to perform other tasks, and the IT organization is free to take advantage of the significant savings in performance, power and cooling, and space that result from reducing the total number of systems they have to support.

This study highlights offloading enhanced data compression on Cisco® Data Intelligence Platform, a private cloud designed for exabyte scale data, with Cisco UCS® rack servers running Cloudera Data Platform Private Cloud Base (CDP PvC Base) to Xilinx FPGA accelerated with Eideticom's NoLoad computational storage solution.

Solution highlights

Improve TCO for Data Lake with accelerated compression

Cisco, Cloudera, Xilinx, and Eideticom bring together a high performance and scalable system architecture that delivers improved storage efficiency by achieving better compression to gain performance and reduce TCO.

Higher compression (3x) through Xilinx Alveo U50 cards

Cisco UCS C-series rack servers powered by Xilinx Alveo U50 data-center-accelerator cards and Eideticom's NoLoad transparent compression technology deliver three times the storage savings gained through improved data compression as measured using standard Hadoop sort benchmarks.

Cloud-scale architecture with diverse computing resources

The Cisco Data Intelligence Platform enables data to be operated by different computing constructs – whether using a CPU, a Graphics Processing Unit (GPU), or a Field Programmable Gate Array (FPGA) – based on application demand.

Independently scale storage and compute resources based on demand

With the Cisco Data Intelligence Platform, customers can start small and expand non-disruptively to facilitate in-time deployments while independently scaling storage and/or compute resources to thousands of nodes.

High performance and power efficient Xilinx Alveo U50 cards

Xilinx Alveo U50 cards bring flexibility, parallelism, and power efficiency, and ensure effective and efficient acceleration of data-processing workloads.

Eideticom NoLoad – bringing in high-performance offloading

Eideticom NoLoad delivers higher performance by transparently offloading the host CPU from compute-intensive compression operations, providing vastly improved application performance.

Cisco Data Intelligence Platform

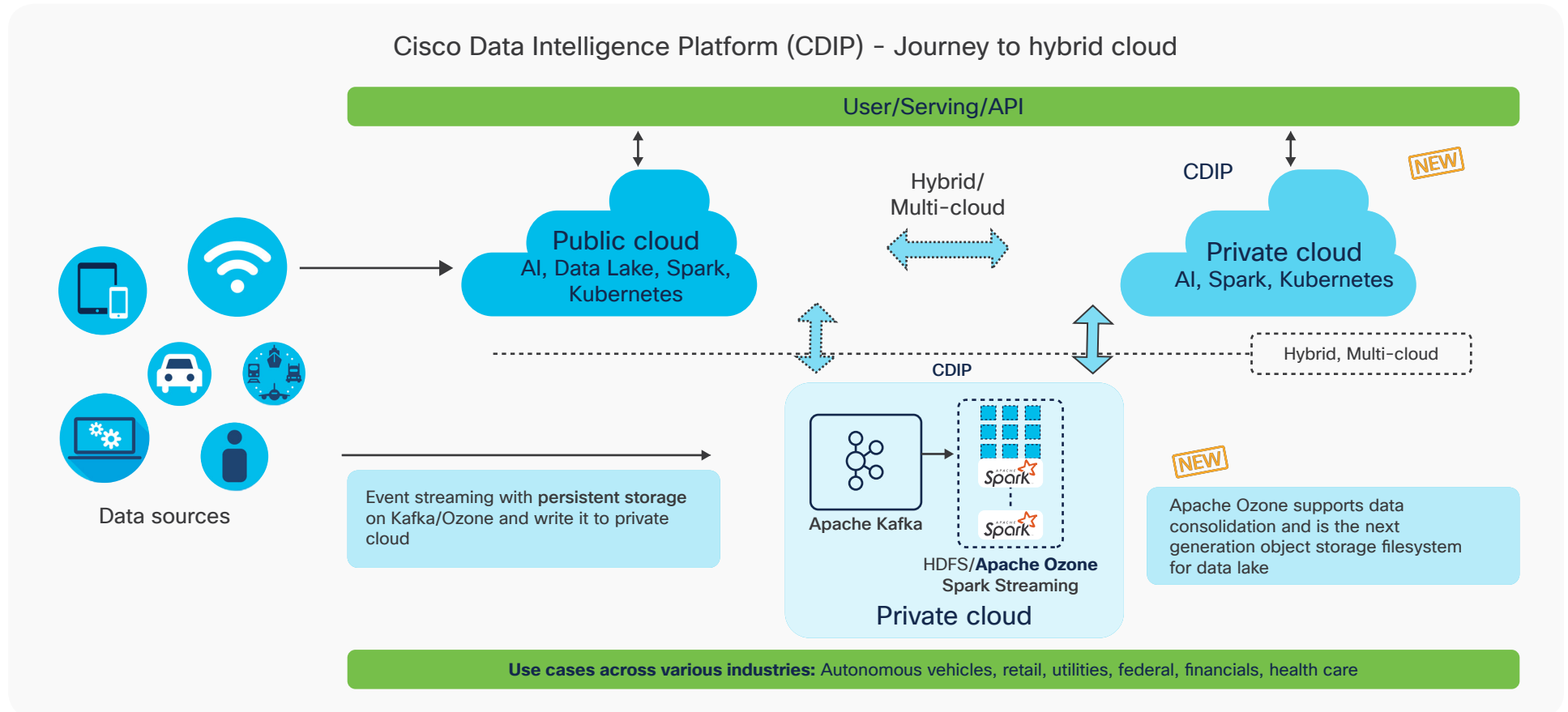
The Cisco Data Intelligence Platform (CDIP) is a cloud-scale architecture and a private cloud primarily for a data lake that brings together bigdata, an AI computing farm, and storage tiers to work together as a single entity, but also be able to scale independently to address the IT issues in the modern data center.

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 compute resources (GPUs, CPUs, and FPGAs) to work on this data for additional analytics processing
- A storage tier, allowing gradual retirement of data that has been worked on to a storage-dense system with a lower cost per terabyte, thus providing a better TCO. Next-generation Apache Ozone file system for storage in a data lake.

The Cisco Data Intelligence Platform supports today's evolving architecture (Figure 1), bringing 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, including the data lake, AI/ML, and object storage.

Figure 1. Cisco Data Intelligence Platform high-level

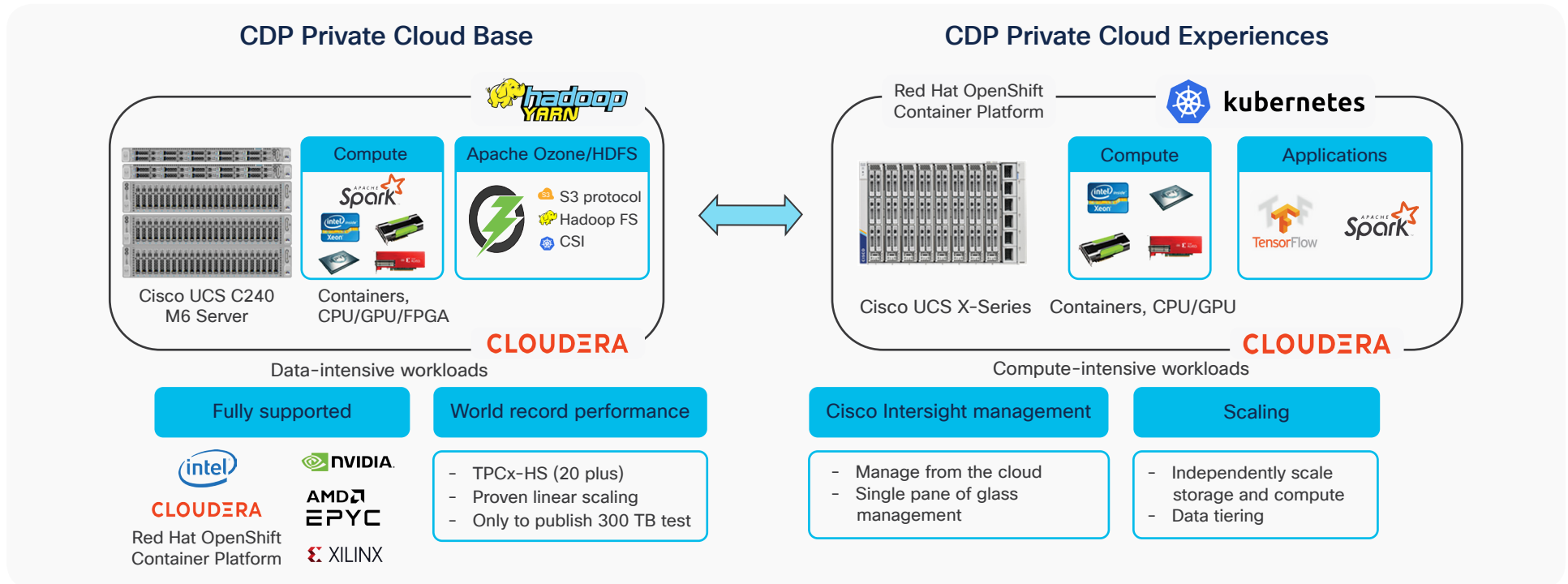


Cisco developed numerous industry leading Cisco Validated Designs (reference architectures) in the area of bigdata, compute farm with Kubernetes (CVD with Red Hat OpenShift Container Platform) and object storage.

A CDIP architecture as a private cloud can be fully enabled by the Cloudera Data Platform with the following components:

- Data lake enabled through CDP PVC Base
- Private Cloud with compute on Kubernetes can be enabled through CDP Private Cloud Experiences
- Exabyte storage enabled through Apache Ozone

Figure 2. Cisco Data Intelligence Platform – Private cloud on Cisco UCS M6

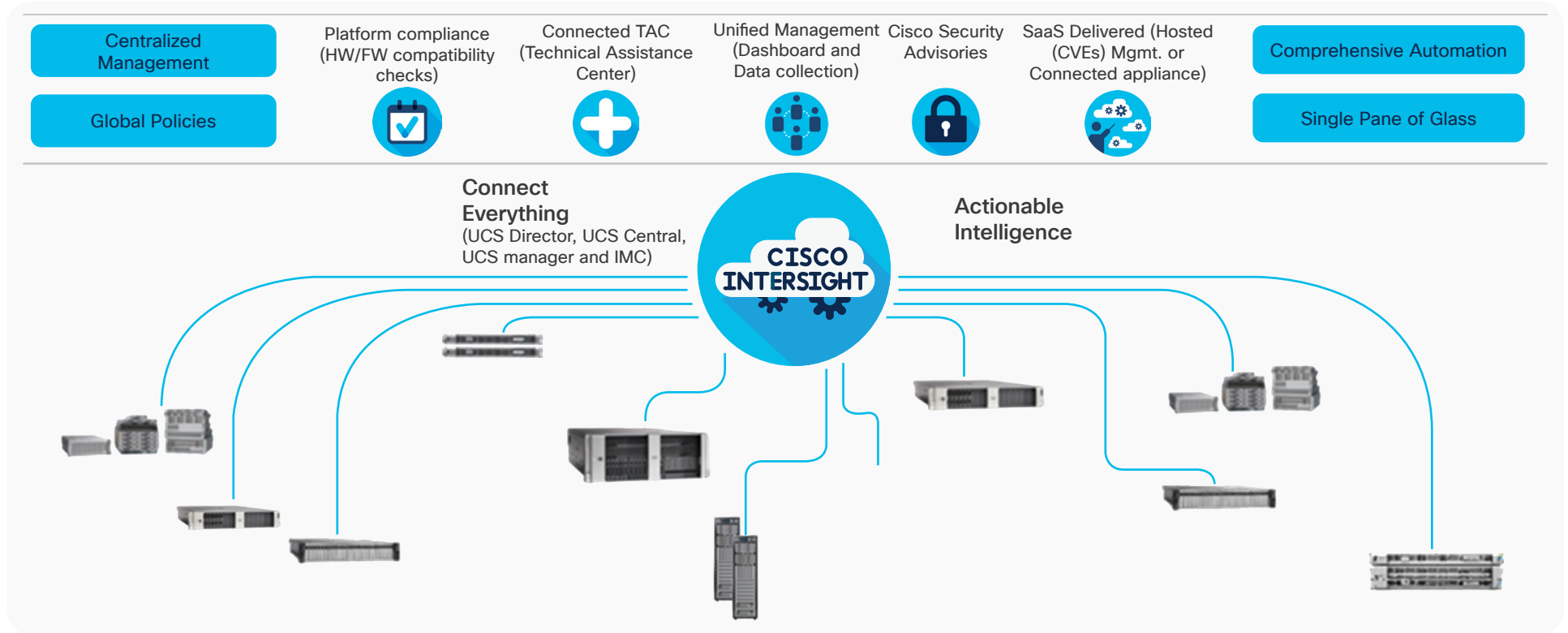


Cisco Intersight cloud-based centralized management

With Cisco Intersight, management is moved from the network into the cloud so that you can respond at the speed and scale of your business and manage all of your infrastructure.

Cisco Intersight is a Software-as-a-Service (SaaS) infrastructure management that provides single pane of glass management. Cisco Intersight enables IT operations managers to claim devices across different sites, presenting these devices in a unified dashboard. 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. The Intersight recommendation engine provides actionable intelligence for IT operations management. The insights are driven by expert systems and best practices from Cisco. For more details go to <https://www.cisco.com/c/en/us/products/cloud-systems-management/intersight/index.html>.

Figure 3. Cisco Intersight



Xilinx Alveo Data Center Accelerator Card

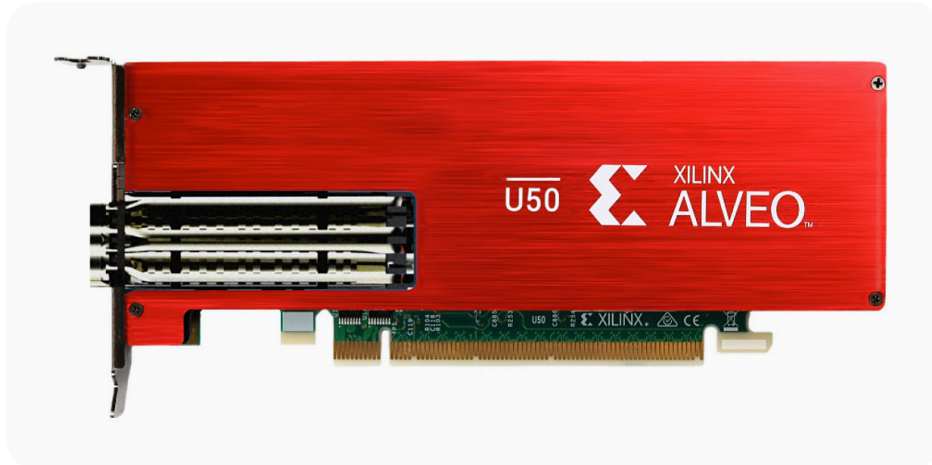
Xilinx Alveo data center accelerator cards are PCI Express compliant and designed to meet the constantly changing needs of the modern data center, providing up to 90X higher performance than CPUs for a variety of workloads, including machine learning inference, video transcoding, and database search and analytics. Built on the Xilinx 16nm UltraScale+ architecture, Alveo accelerator cards adapt to changing acceleration requirements and algorithm standards, capable of accelerating any workload type without changing hardware, and reducing overall cost of ownership.

Enabling Alveo accelerator cards is an ecosystem of Xilinx and partner applications for common data center workloads. The Vitis core development kit provides verified platforms defining all the required hardware and software interfaces, allowing developers to design custom acceleration applications that are easily integrated into the Vitis programming model.

Xilinx Alveo U50

Xilinx Alveo U50 data center accelerator cards provide optimized acceleration for workloads in financial computing, machine learning, computational storage, and data search and analytics. Built on the Xilinx UltraScale+ architecture and packaged in an efficient 75-watt, low-profile form factor, the Alveo U50 card includes 8GB HBM2, 100GbE networking, and PCI Express 4.0 and is designed for deployment in any server. For more details, go to <https://www.xilinx.com/products/boards-and-kits/alveo/u50.html>.

Figure 4. Xilinx Alveo U50 Data Center Accelerator Card



Xilinx Alveo U50 Card specifications

Hardware and software installation procedures for the half-height, half-length Alveo U50 data center accelerator card detailed steps are outlined in the [Alveo U50 Data Center Accelerator Card Installation Guide](#).

Table 1. Xilinx Alveo U50 card specification

| Features | Alveo U50 Card |
|-------------------|--|
| Architecture | UltraScale+ |
| Form factor | Half-height, half-length, single-slot, low-profile |
| Lookup tables | 872,000 |
| HBM2 memory | 8GB |
| HBM2 bandwidth | 316GB/s |
| Network interface | 1 x QSFP28 (100GbE) |
| Clock precision | IEEE 1588 |
| PCI Express | PCI Gen3 x16, dual PCIe Gen 4 x8, CCIX |
| Thermal solution | Passive |
| Power (TDP) | 75W |

How to Buy Alveo U50

To procure Alveo U50 please follow the instructions provided at this link: <https://www.xilinx.com/products/boards-and-kits/alveo/u50/noloadcsp-sales.html>

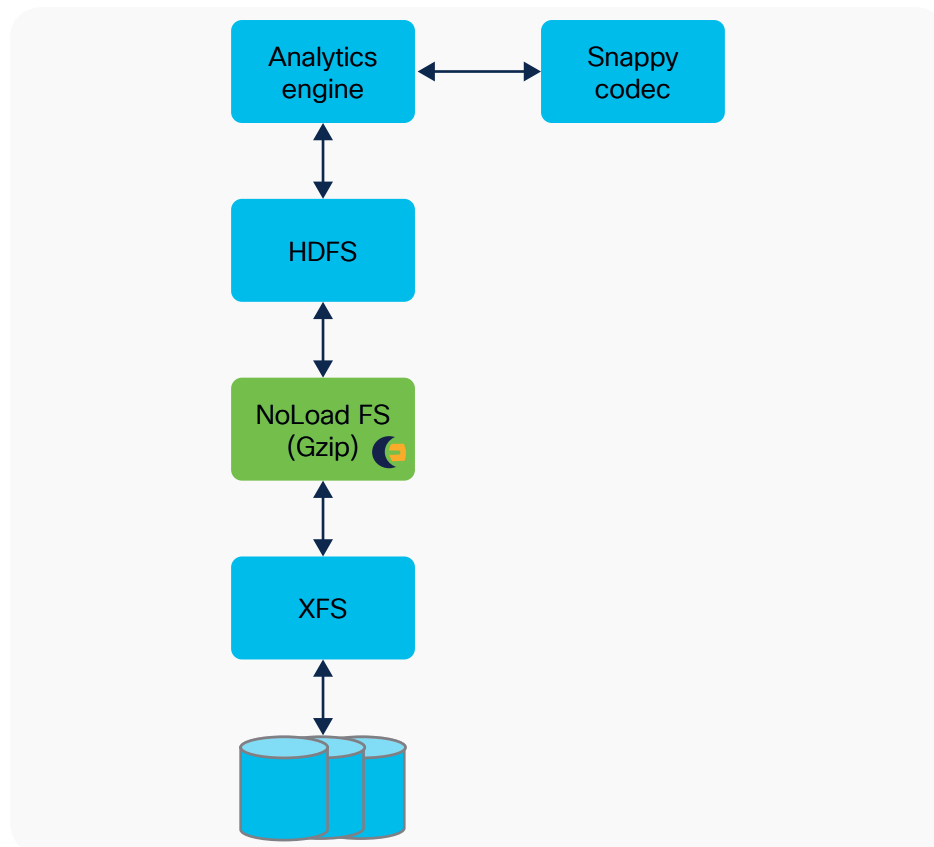
Eideticom NoLoad computational storage processor

Eideticom's NoLoad is a computational storage processor purpose-built for the acceleration of storage and compute-intensive workloads, including compression. Eideticom's NoLoad Transparent Compression for Hadoop requires no application changes, ties directly into file systems such as XFS/EXT4 and achieves line-rate compression while using 70 percent less CPU. NoLoad's FPGA-based solution accelerates efficient compression and fast

decompression enabling reduction in data size of the input, shuffle, and output; thus, speeding up overall processing time, which results in improved system throughput, higher resource utilization, and less hardware.

The Xilinx Alveo U50 card is a PCIe Gen3x16, or dual Gen4x8 FPGA accelerator card and is loaded with the Eideticom NoLoad computational storage solution. This solution provides computational services in an Add-In-Card (AIC) form factor and is considered a Computational Storage Processor (CSP).

Figure 5. NoLoad file system logical diagram with Hadoop - transparent compression with zero application changes



The NoLoad transparent compression kernel module was developed to allow applications such as Hadoop or other databases to offload critical storage tasks to the Alveo U50 card without any modifications. This offloading leads to improved performance and efficiency and reduced costs for the storage system. NoLoad supports several acceleration functions, including compression, encryption, erasure coding, deduplication, and data analytics.

As shown in Figure 5, the NoLoad File System resides on top of the operating system file system which is transparent to HDFS and requires no changes in the existing Hadoop configuration. Cloudera Manager provides the capability to apply Snappy codec as the default to compress data temporarily by enabling a compression-specific configuration. For more details on configuring data compression in Cloudera Manager, go to <https://docs.cloudera.com/cdp-private-cloud-base/7.1.7/managing-clusters/topics/cm-choosing-configuring-data-compression.html>.

Reference architecture

The reference architecture configuration details for the data lake, AI/ML components of the data lake, and object storage is listed in detail for [Cisco Data Intelligent Platform](#).

Table 2, below, lists the reference architecture for this study. You can deploy these configurations as is or use them as templates for building custom configurations. You can scale your solution as your workloads demand, including expansion to thousands of servers with Cisco Nexus® 9000 Series Switches. The configurations vary in disk capacity, bandwidth, price, and performance characteristics.

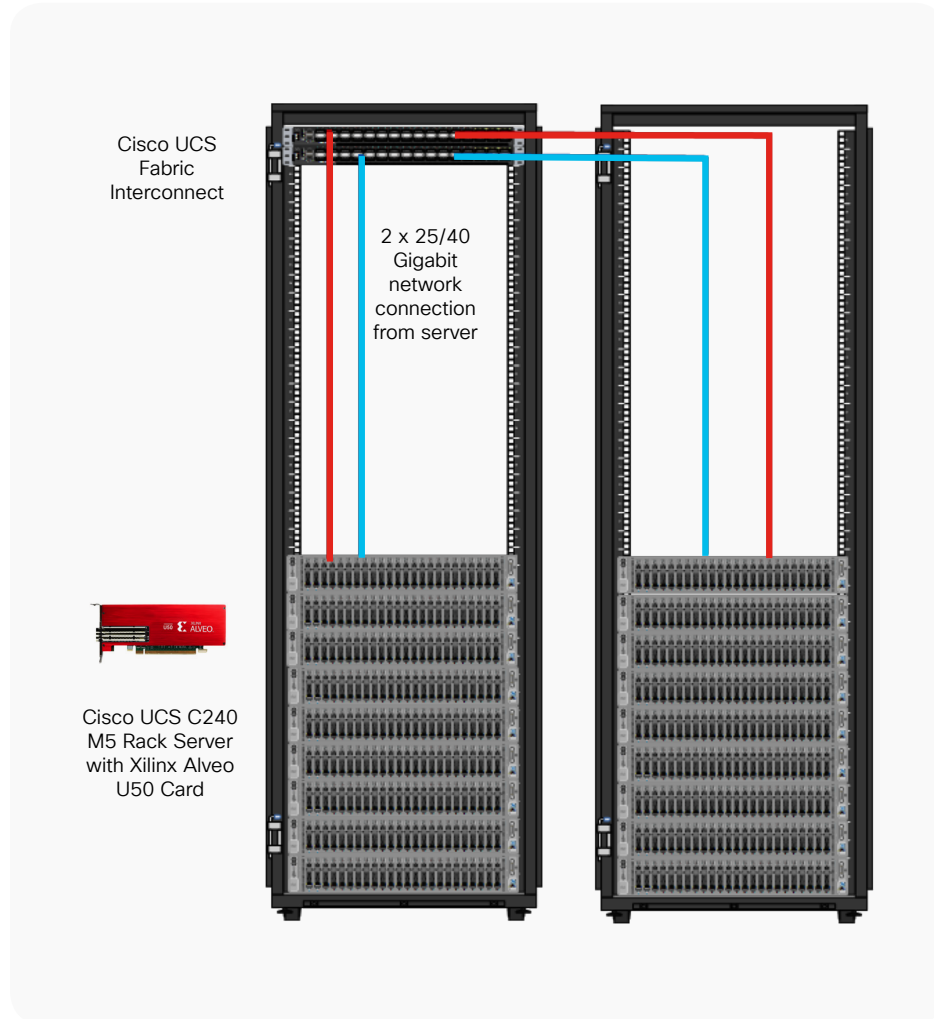
Table 2. Cisco Data Intelligence Platform on Cisco UCS M5 with Xilinx and Eideticom

| Server | Cisco UCS C240 M5 2RU Rack Server |
|---------------------------|--|
| CPU | 2 x 2nd Gen Intel Xeon Scalable Processors 6230R (2 x 26cores @ 2.1GHz) |
| Memory | 12 x 32GB 2933 MHz DDR4 (384 GB) |
| Boot | Cisco Boot-Optimized M.2 RAID Controller with 2 x 240GB SSDs |
| Storage | 26 x 2.4TB 10K rpm SFF SAS HDDs or 12 x 1.6TB Enterprise Value SATA SSDs |
| Network | 25 Gigabit Ethernet (Cisco UCS VIC 1457) or 40/100Gigabit Ethernet (Cisco UCS VIC 1497) |
| Storage controller | 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) |
| Xilinx FPGA card | 1 x Xilinx Alveo U50 Acceleration Card |
| Eideticom NoLoad | V6.2 |
| Software | Cloudera Private Cloud Base 7.1.7 Red Hat Enterprise Linux 8.2 |

Table 3. Cisco Data Intelligence Platform on Cisco UCS M6 with Xilinx and Eideticom

| Server | Cisco UCS C240 M6 2RU Rack Server |
|---------------------------|---|
| CPU | 2 x 3rd Gen Intel® Xeon® Scalable Processors 6330 processors (2 x 28 cores, at 2.0 GHz) |
| Memory | 16 x 32 GB RDIMM DRx4 3200 (8Gb) (512 GB) |
| Boot | Cisco Boot-Optimized M.2 RAID Controller with 2 x 960GB SSDs |
| Storage | 24 x 2.4TB 10K rpm SFF SAS HDDs or 24 x 15.3TB Enterprise Value SATA SSDs |
| Network | quad port 10/25 Gigabit Ethernet (Cisco UCS VIC 1467) or dual port 40/100 Gigabit Ethernet (Cisco UCS VIC 1477) |
| Storage controller | Cisco M6 12G SAS RAID Controller with 4GB FBWC or Cisco 12G SAS HBA |
| Xilinx FPGA card | 2 x Xilinx Alveo U50 Acceleration Card |
| Eideticom NoLoad | V6.2 |
| Software | Cloudera Private Cloud Base 7.1.7 Red Hat Enterprise Linux 8.2 |

Figure 6. Reference Architecture with Cisco UCS C240 M5 Rack Servers



Performance benchmark

As highlighted in Table 2, seventeen Cisco UCS C240 M5 Rack Servers were configured with one Xilinx Alveo U50 card per server. The benchmark used was Hadoop sort for different data sizes (3TB and 10TB) with both Spark and MapReduce and was compared with how much storage space is saved with and without hardware acceleration, as highlighted in Table 4 below.

Table 4. Storage savings by compressing data using Xilinx Alveo U50 Cards on Cisco UCS C240 M5

| Parameter | Default compression ¹ | Xilinx + NoLoad compression |
|-----------------------------------|----------------------------------|-----------------------------|
| Size on disk (3TB Spark) | 16.5 | 5.6 |
| Hadoop sort/hour (3TB Spark) | 15.9 | 16.4 |
| Size on disk (10TB Spark) | 55 | 18.7 |
| Hadoop sort/hour (10TB Spark) | 13.2 | 16.1 |
| Size on disk (3TB Map Reduce) | 16.5 | 5.7 |
| Hadoop sort/hour (3TB Map Reduce) | 9.8 | 8.9 |

¹ Default compression uses Snappy codec, which only compresses intermediate results without application changes.

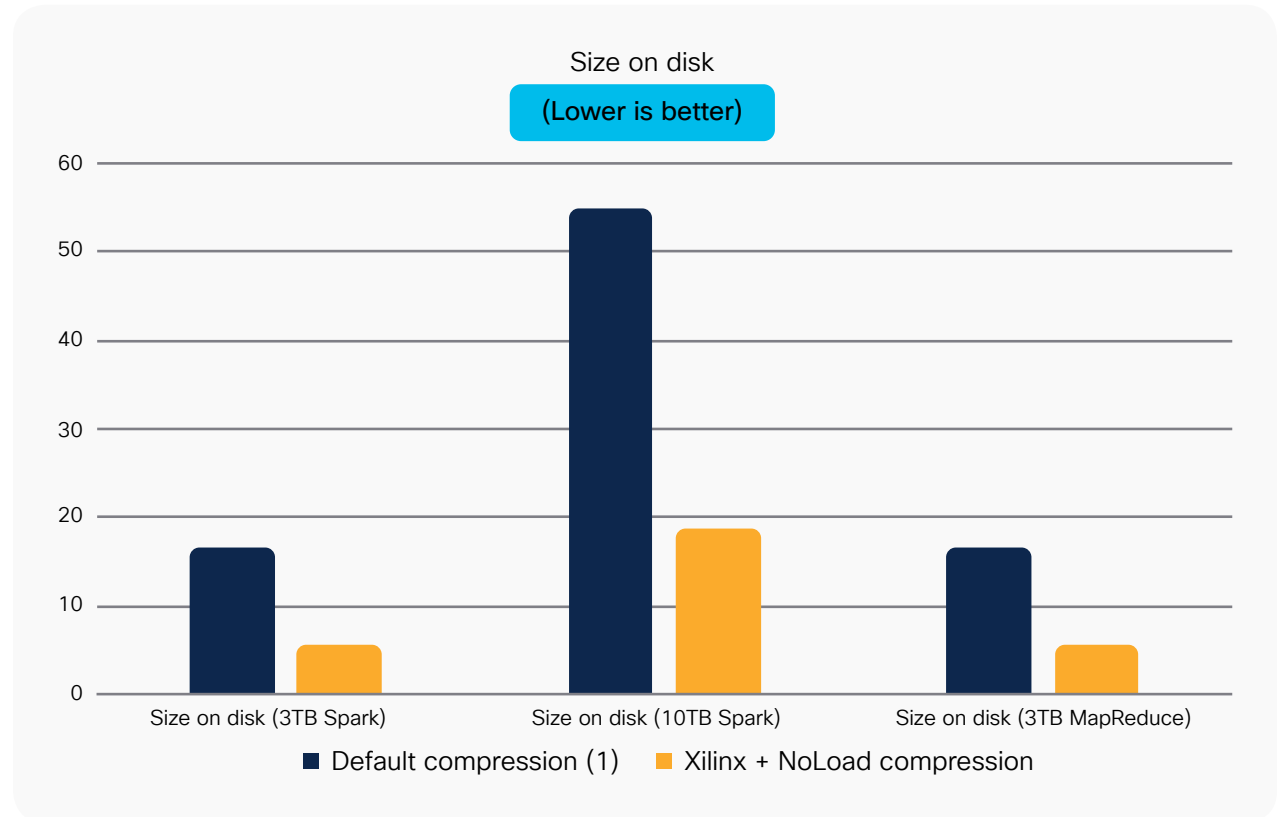
The results of this system showed that hardware-accelerated compression through a Xilinx Alveo U50 card will allow you to increase your current storage capacity by 44 percent by changing from a Snappy CPU compression of 2.05 to a Gzip hardware compression of 3.05. In addition, you are also saving CPU cycles for other tasks.

Conclusion

Unprecedented growth of data leads to the importance of data compression to reduce network congestion and improve storage and energy efficiency.

The configuration detailed in this document can be extended to clusters of various sizes, depending on application demands. Scaling beyond a single rack can be implemented by interconnecting multiple Cisco UCS domains using Cisco ACI® technology. This architecture is scalable to thousands of servers and to hundreds of petabytes of storage and can be managed from a single pane using Cisco Intersight.

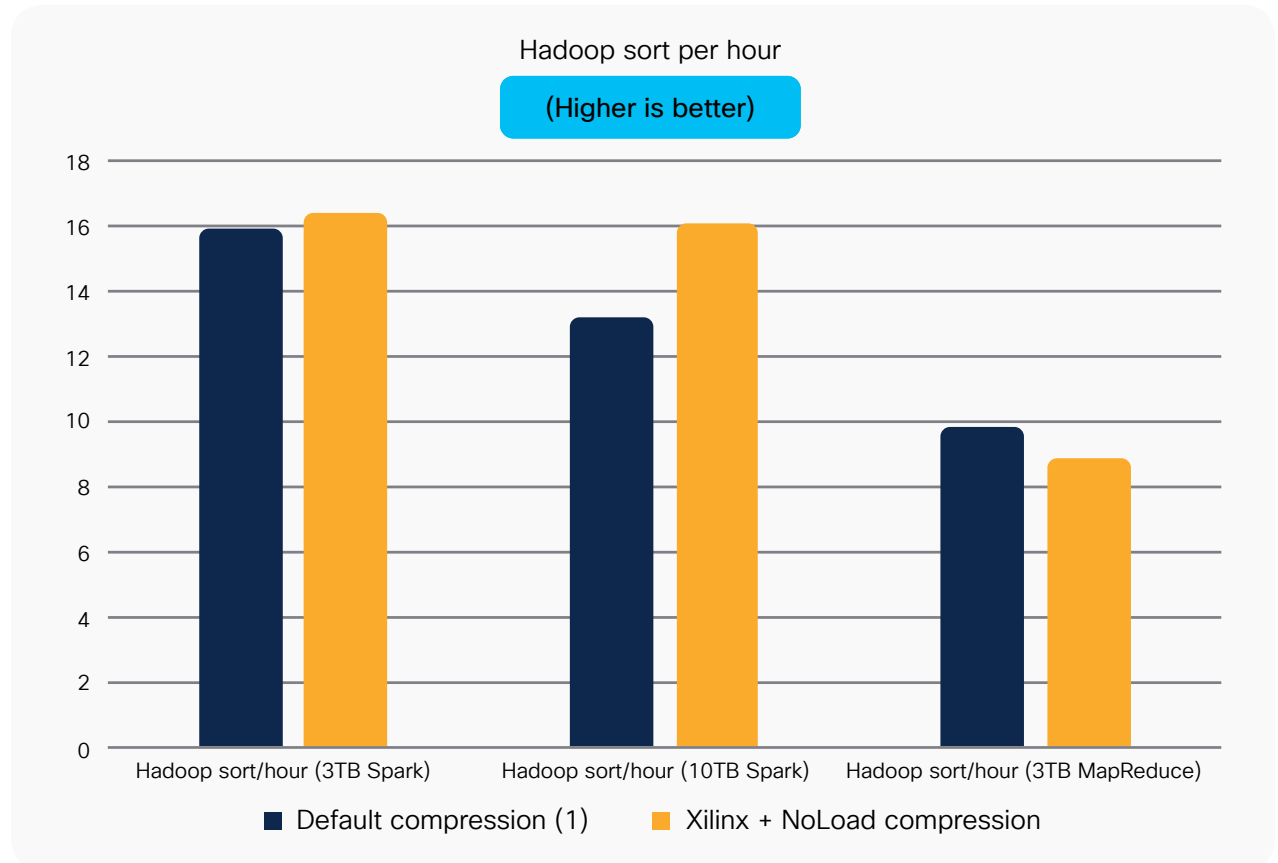
Figure 7. Storage utilization comparison with default Vs Xilinx + NoLoad compression



For more information

- To learn more about Cisco UCS big-data solutions, visit: <https://www.cisco.com/c/en/us/solutions/data-center-virtualization/big-data/index.html>
- To learn more about Cisco Data Intelligence Platform, visit: <https://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-c-series-rack-servers/data-intelligence-platform-cdip-so.html>
- To find out more about Cisco ACI solutions, visit: <http://www.cisco.com/go/aci>
- To access the Alveo Data Center Accelerator Card Platforms User Guide, visit https://www.xilinx.com/support/documentation/boards_and_kits/accelerator-cards/ug1120-alveo-platforms.pdf
- To access the Alveo U50 Data Center Accelerator Card Installation Guide, visit https://www.xilinx.com/support/documentation/boards_and_kits/accelerator-cards/1_7/ug1370-u50-installation.pdf

Figure 8. Hadoop sort per hour comparison with default Vs Xilinx + NoLoad compression



The Services and all content and other materials on the Services, including, but not limited to, the Cludera effects, our logo and all designs, text, graphics, logos, icons, images, photographs, audio clips, digital downloads, data compilations, software and the selection and arrangement thereof (collectively, the "Materials") are our property or are licensed to us and are protected by United States and/or international copyright laws. Except as otherwise provided in these Terms, you are not permitted to copy, modify, display, license, create derivative works of, or otherwise exploit our Materials without written permission of the copyright holder.

© Copyright 2021 Xilinx, Inc. Xilinx, the Xilinx logo, Alveo, Artix, Kintex, Spartan, Versal, Virtex, Vivado, Zynq, and other designated brands included herein are trademarks of Xilinx in the United States and other countries. All other trademarks are the property of their respective owners.

© 2021 Cisco and/or its affiliates. All rights reserved. Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: <https://www.cisco.com/go/trademarks>. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)