TECHNICAL BRIEF

The Future of Data Infrastructure
Western Digital's Software Composable Infrastructure

Prepared by:
Western Digital

Contents

Introduction 2

Today's Data Center Challenges 2

Diversity of Applications and Workloads 3

Hyper-Converged Infrastructures 3

Need for Composable Infrastructures 4

- Open Standards Are Essential for CDIs 5
- How a CDI Works 5
- Benefits of a CDI 5
- Components of a CDI 6

Western Digital's Leadership 7

- Open Composable API 7
- Driving the Ecosystem 8
- Initial TCO Results 8

Looking Ahead 9

About Western Digital 10

Key Definitions 11

Introduction

Our data-centric world is evolving. Big data is getting bigger and faster. Fast data is getting faster and bigger. High-performance applications are increasing in diversity and being fueled by artificial intelligence, machine learning, autonomous vehicles, smart cities and cars, virtual and augmented reality, and hyper-analytics. This has also created an increase in workflow complexities. As such, the task of orchestrating and managing the data infrastructure at scale has never been more challenging. The need to 'spin up' or 'spin down' application instances quickly, without disrupting concurrent applications, is required. Western Digital provides insights into the architectural shifts underway to ensure that performance, scale and efficiency can be quickly and easily provisioned and optimized for the most demanding data environments. Through the convergence of emerging technologies, open standards, innovation and a growing ecosystem, the future is in composable infrastructures.

Today's Data Center Challenges

Data is no longer generated from just applications. It comes from diverse sources (e.g., mobile devices, production machines and sensors, video surveillance systems, IoT and IIoT devices, healthcare monitors and wearables, to name a few). It is being generated in both large-scale data centers at the core, as well as from remote and mobile sources at the edge of the network. To process and analyze this data for immediate or future use, fast data and big data environments are rising quickly. Big data are very large datasets that need to be analyzed, and through computations and algorithms, unmask trends, patterns and associations between disparate data groups. The ability to respond to events as they happen is also required, paving the way for fast data environments that process and transform data as it is captured. These applications leverage the algorithms derived from big data to provide real-time decisions and results.

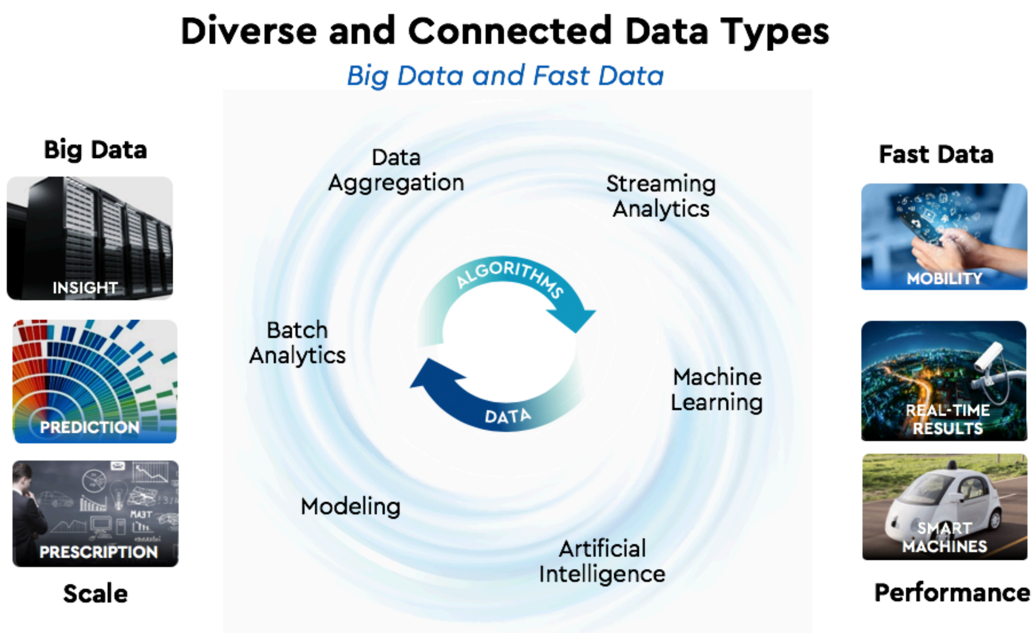
Data analysis leads to actionable insights

Applications within an enterprise change constantly throughout the day, requiring a flexibility to provision IT resources quickly to meet application demands when and where they are required. The ability to react to the expanding diversity of big and fast data applications is vital and requires the allocation of more compute, storage and network resources at scale, and in real time across the enterprise.

Diversity of Applications and Workloads

As big data and fast data environments proliferate, they are breaking the boundaries of traditional infrastructures and system architectures. When data-intensive workloads are supported by general-purpose architectures, a uniform ratio of resources are typically used to address all compute processing, storage and network bandwidth requirements. This 'one size fits all' approach is not effective for these diverse data-intensive workloads.

What is required for today's data-centric architectures are capabilities that enable more control over the blend of resources that each application needs so that optimized levels of processing, storage and network bandwidth can be scaled independently of one another. These extreme data-intensive workloads that drive analytics, machine learning, artificial intelligence and smart systems demand an infrastructure that is both flexible and composable.



Hyper-Converged Infrastructures

Given the challenges associated with general-purpose architectures (fixed resource ratios, underutilization and overprovisioning), converged infrastructures (CIs) emerged delivering a preconfigured package of hardware resources in a single system. The compute, storage and networking components are discrete and managed through software. CIs have progressed into **hyper-converged infrastructures (HCIs)** where all of the hardware resources are virtualized, delivering software-defined computing, storage and networking.

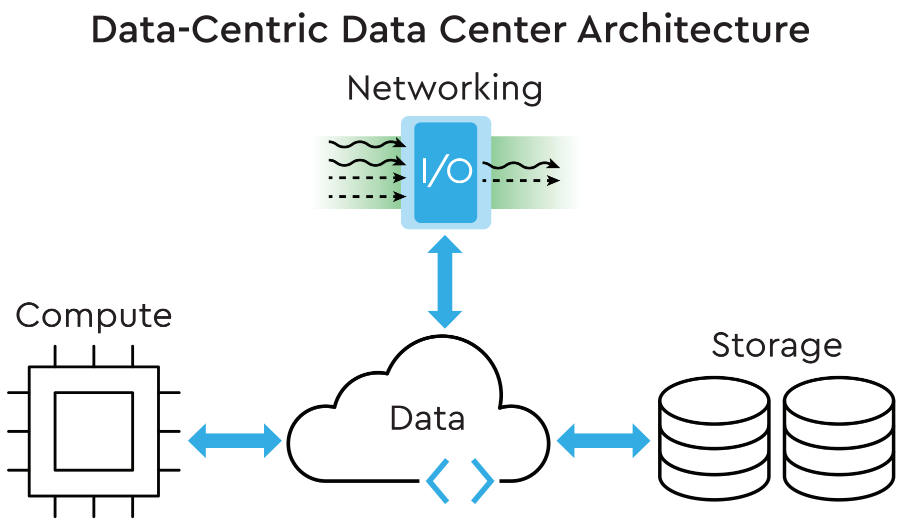
Though HCIs combine compute, storage and network resources into a single virtualized system, it is not without inefficiencies. For example, the scalability limits are defined by the processor, and access to resources are made through the processor. So, to add more resources, such as storage, memory or networking, HCI architectures deliver additional processors even if compute capacity is not required.

Additionally, when data is moved to a public cloud for archival in an HCI model, there is a lack of **performance predictability** due to **noisy neighbors**. With the need to optimize infrastructure **provisioning** to meet specific application requirements, data center architects are trying to build flexible infrastructures using inflexible building blocks.

The fixed building block approach cannot achieve the level of flexibility and predictable performance needed in today's data center as the disaggregated HCI model needs to be enabled and become easily composable through software tools based on an open **application programming interface (API)**.

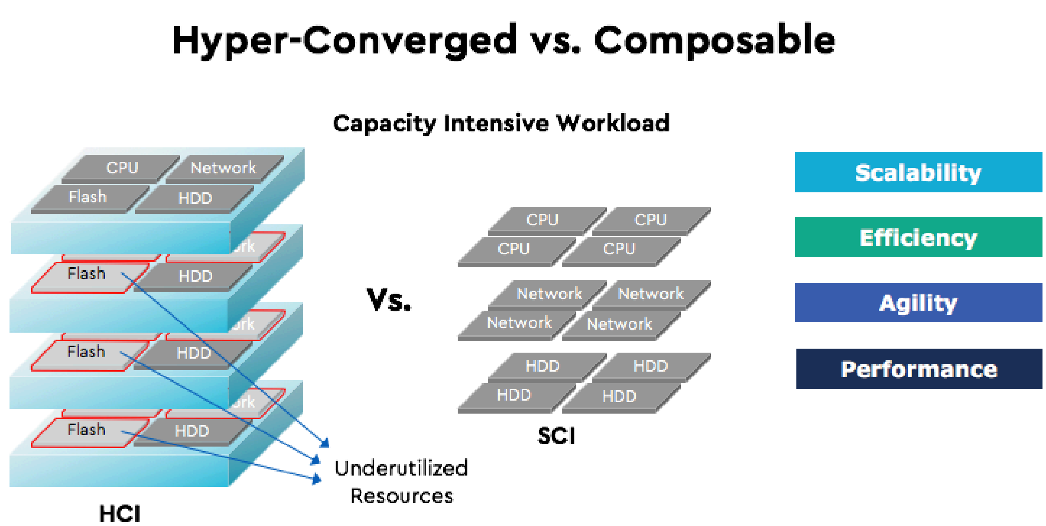
Need for Composable Infrastructures

A **composable disaggregated infrastructure (CDI)** is a data center architectural framework whose physical compute, storage and network fabric resources are treated as services. As such, high-density compute, storage and network racks use software to create a virtual application environment that provides whatever resources the application needs in real-time to achieve the optimum performance required to meet workload demands.



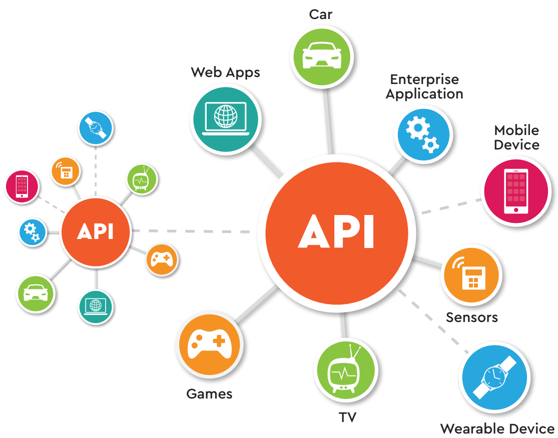
Today's data-centric data center architectures

In a CDI, virtual servers are created out of independent resource pools comprised of compute, storage and network devices, rather than partitioned resources that are hardwired as HCI servers. Thus, the servers can be provisioned and re-provisioned as needed, under software control, to suit the demands of particular workloads as the software and hardware components are tightly integrated.



Open Standards Are Essential for CDIs

With an API on the composing software, an application could request whatever resources are needed, delivering real-time server reconfigurations on-the-fly, without human intervention, and a step toward the self-managing data center.



APIs enable workloads to be provisioned quickly and efficiently

Through a single management API, a CDI can integrate operational silos, logically pool resources and automatically compose and recompose blocks of disaggregated IT services to effortlessly meet the changing needs of an application. This enables IT personnel to provision workloads just as quickly and efficiently as public cloud providers, but in a secure, scalable on-premises private cloud environment. When more compute, storage or networking resources are needed, one simply changes the virtual server configuration for that application from a computer. If a virtual server dies, the boot device can be moved to a new virtual server. The infrastructure is not only flexible, but composable, reducing underutilization and overprovisioning while creating a more agile, cost-effective data center environment.

Therefore, in order for the CDI initiative to be successful, it must be backed by an ecosystem to create those differentiated products for it to succeed and evolve. An open management API is essential for success as it provides developers with public access to software and code for development, as well as the set of requirements that govern how an application communicates and interacts with another.

How a CDI Works

A CDI resides on premises in the data center. Since compute, storage and network are treated as services, the 'composable' part refers to its ability to make those resources available on-the-fly, depending on the needs of each physical or virtual application. Therefore, a management layer within the API is required to not only discover and access the resource pool, but also to ensure that the right resources are in the right place when they are needed. In this infrastructure, if the workload needs additional compute or storage resources, a few keyboard clicks adds more to the server. Once the workload has completed its tasks, the resources can then be returned to the shared pool for use by other applications. This composition and re-composition under software control is fast and without physical involvement or reconfiguring the equipment. Resources are no longer trapped in separate silos, eliminating 'stranded storage' and underutilized CPUs.

Benefits of a CDI

The main benefit of CDIs is their ability to disaggregate compute, storage and network fabric resources into shared resource pools that can be available for on-demand allocation with an end result that vastly improves server utilization, performance, flexibility and agility in the data center. And, because composable infrastructures are compatible with existing physical infrastructures, they are very cost-effective since neither resources or the time spent deploying new applications are wasted. CDIs are also ideal for enabling SLA-driven architectures for compute and storage where various workloads are automatically assigned resources that have sufficient performance and capacities to meet application specific SLAs.

The key benefits of CDIs can be summated into three categories and include its ability to: (1) Automate Agility; (2) Improve Efficiency; and (3) Improve Performance Predictability.

1. **Automate Agility**
   CDIs enable the access and management of a resource pool that can be routed dynamically on demand to where they are needed the most and are logically pooled so that administrators do not have to physically configure hardware to support a specific application. As such, CDIs can adapt more quickly to changing workloads and provision them more purposefully in support of a range of big and fast data applications.

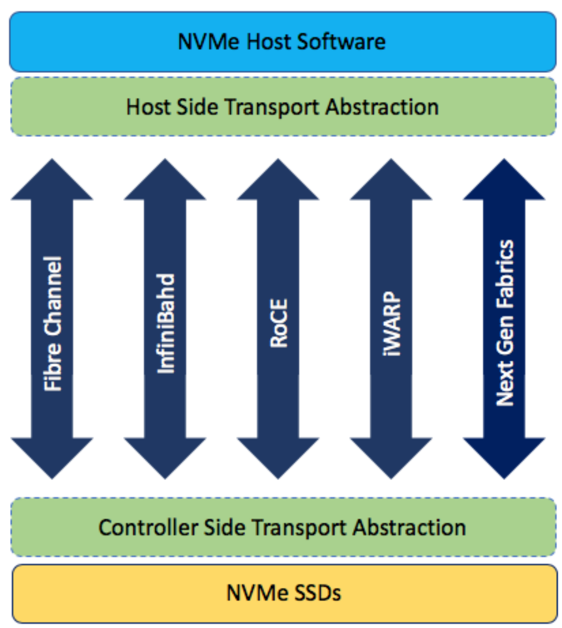
2. **Improve Efficiency**
   CDIs improve economic efficiency by enabling better utilization of compute and storage resources. By reducing overprovisioning of resources, it improves power consumption, cooling, associated costs, etc. It also enables resources to be expanded as needed so organizations can buy smaller amounts of storage upfront, for example, and expand the footprint gradually and evenly over time, based on need.

3. **Improve Performance Predictability**
   With the ability to allocate resources properly to meet application demand, less wasted or unused IOPS (input/output operations per second) occur as resources perform exactly as expected and are available when required. This reduces the variability in business and/or operating processes and execution as the lower the variance, the more consistent the outcome. Unpredictability in performance makes it difficult to establish normal operating parameters, costs become difficult to forecast, and profits and long-term operating expenses become difficult to calculate.

In a nutshell, CDIs can automatically and effortlessly compose and recompose blocks of disaggregated IT services to meet the changing needs of an application.

Components of a CDI:

There are two main components that enable a CDI. The first is software to dynamically provision and manage the resources, while the second is a network protocol to disaggregate the resources and make them shareable amongst multiple applications and servers. In the case of software, the IT resources can be dynamically provisioned and managed through an open API without physically touching hardware. A networked protocol is equally important to a CDI enabling compute and storage resources to be disaggregated from the server and available to other applications. Connecting compute or storage nodes over a fabric is important as it enables multiple paths to the resource. Emerging as the frontrunner network protocol for CDI implementations is **NVMe™-over-Fabric (NVMe-oF™)**. Regardless of the acronym, NVMe-oF promises to deliver the lowest end-to-end latency from application to storage, enabling CDIs to provide the data locality benefits of direct-attached storage (low latency and high performance), while delivering agility and flexibility by sharing resources throughout the enterprise. **Non-Volatile Memory Express™ (NVMe)** technology is a streamlined, high-performance, low-latency interface that utilizes an architecture and set of protocols developed specifically for persistent flash memory technologies. It increases non-volatile memory storage performance in locally-attached PCIe-based servers and SSDs by eliminating the DAS bottlenecks associated with legacy HDD interfaces and SCSI command stacks. SSDs based on PCIe NVMe are well-suited for applications that require extremely fast data transfer rates, a significant number of I/O operations performed per second, and short access times when delivering data.



NVMe-oF extends NVMe for use on network fabrics

The NVMe protocol has been extended beyond local attached server applications by providing the same performance benefits across a network through the NVMe-over-Fabric specification. This specification enables flash devices (such as PCIe SSDs and storage arrays) to communicate over networks, (such as InfiniBand™ or Converged Ethernet), delivering the same high-performance, low-latency benefits as local attached NVMe. As NVMe was designed for PCIe as a server attachment for storage, NVMe-oF extends NVMe for use on network fabrics where the primary advantage is scale. While PCIe supports a limited number of servers and devices, there is virtually no limit to the number of servers that can share NVMe-oF storage or the number of NVMe-oF storage devices that can be shared. As networks can be slow, narrow in focus or too expensive, the most cost-effective bandwidth is within servers, which is why so much storage locally resides inside of them. PCIe-based storage devices are helping to drive this migration because the interface protocol enables more data lanes to be added, which in turn, delivers fast I/O performance with very low latency built into every server. And, since almost every device supports a PCIe interface, there are no drivers to install, making it an effective interface for composable infrastructures.

Western Digital's Leadership

As one of the market leaders in developing innovative products, systems and solutions based on NVMe technology, Western Digital is further investing in the NVMe-oF standard to enable a composable SLA-driven architecture for compute and storage covering flash-based media, hard drives and eventually memory. Western Digital has unveiled its new OpenFlex™ architecture that delivers breakthrough levels of scalability, efficiency and performance for big data and fast data applications, enabling a new era in NVMe-oF based CDIs. Western Digital brings the essential building blocks, technological advancements, tools, resources and contacts to enable and help accelerate market adoption and will provide an open management API and storage abstraction layer to accelerate time-to value.

OpenFlex F3200 Series Fabric Device and 3U enclosure for mission critical apps and data

For flash-based, high-performance, mission-critical applications and data, the company offers its OpenFlex F3200 Series fabric device which is a hot-swappable NVMe flash device that supports capacities ranging from 12.8TB[1] to 61.4TB. The initial implementation is RDMA over Converged Ethernet (RoCE) which is ubiquitous in the data center and cost-effective to deploy. Up to ten (10) OpenFlex F3200 Series fabric devices can be housed in the OpenFlex E3000 fabric storage platform which is a separate, but required 3U10 enclosure that can deliver capacity points ranging from 128TB up to 614TB based on the number and type of devices installed.

Open Composable API

As part of the initial OpenFlex launch, Western Digital has developed and contributed to www.opencompute.org an **Open Composable API** to dynamically provision and manage the resources allowing flash storage, for example, to be shared by multiple servers in the CDI architecture. Resources can be scaled independently of each other. Capacity can be expanded at will to adapt to changing workloads. The result is the ability to enable heterogeneous compute environments, improve capacity utilization and serve multiple applications with a common pool of high-speed storage.

[1] Projected specifications, subject to change. As used for storage capacity, one petabyte (PB) = one quadrillion bytes, one terabyte (TB) = one trillion bytes and one gigabyte (GB) = one billion bytes. Total accessible capacity varies depending on operating environment.

Driving the Ecosystem

The OpenFlex CDI initiative will require a complete ecosystem to support it in order to thrive. Included will be infrastructure and system architects, HW/SW/FW developers, contributors, test and validation groups, solutions and application developers, compute/storage/ network fabric vendors, system integrators, and users. The company is committed to advancing CDI architectures in data centers globally for use in both mission-critical applications and for archival and other data protection schemes. Western Digital is committed to supporting the broad ecosystem of data center software and hardware vendors.

Initial TCO Results

To justify the investment required to migrate over to an NVMe-oF based CDI platform, the ecosystem will want assurances that this infrastructure strategy will ultimately provide significant business value, a savings in total cost of ownership (TCO) and a return on investment (ROI). In a head-to-head comparison of equipment, components and associated costs required in a traditional HCI versus Western Digital's new OpenFlex NVMe-oF-based CDI, a 34 percent cost savings can be achieved if all of the equipment and components are purchased upfront[2].

However, as the OpenFlex architecture delivers the proper blend of IT resources for on-demand workload allocation, not all of the equipment and components will need to be purchased in one fell swoop. In fact, by implementing a 'buy-as-you-grow' strategy, a 41 percent savings can be achieved over a traditional HCI, assuming the following:

- 50 to 60 percent of the hardware is purchased upfront
- One server is purchased every quarter for three years
- One OpenFlex F3200 Series fabric device is purchased annually. Network fabrics based on NVMe-oF enable separate resources to be added over time and scale them independent of each other.
- As such, the 'buy-as-you-grow' model enables expenses to be delayed over time, and with that, users can take advantage of lower dollar per gigabyte storage pricing over time, as well as the costs associated with running equipment and components that are underutilized, or worse, not being used.
- Other areas of cost-savings when comparing an OpenFlex architecture to a traditional HCI can include:
- Reducing the number of storage racks from two to one
- Utilizing more scaled up servers within the rack to saturate the ports
- Utilizing multi-node servers versus individual servers
- Reducing the number of servers by almost 50 percent
- Upgrading the CPUs and DRAM with more cores and capacity
- Reducing expenses in power and cooling as a result of the CPUs running faster and more efficiently
- Reducing expenses associated with networking given the efficiencies of NVMe-oF

With Western Digital's OpenFlex architecture, resources can more easily be added as needed, enabling a 'buy-as-you-grow' model that delivers an extra seven percent in savings over three-years.

---

[2] Assumptions include: (1) the HCI resources are not fully utilized; (2) CPU cores have 22 percent unused capacity; (3) DRAM has at least 13 percent of unused capacity; (4) storage capacity is at least 20 percent underutilized. The TCO values are based on internal analysis using those assumptions and public component pricing as of July 2018.

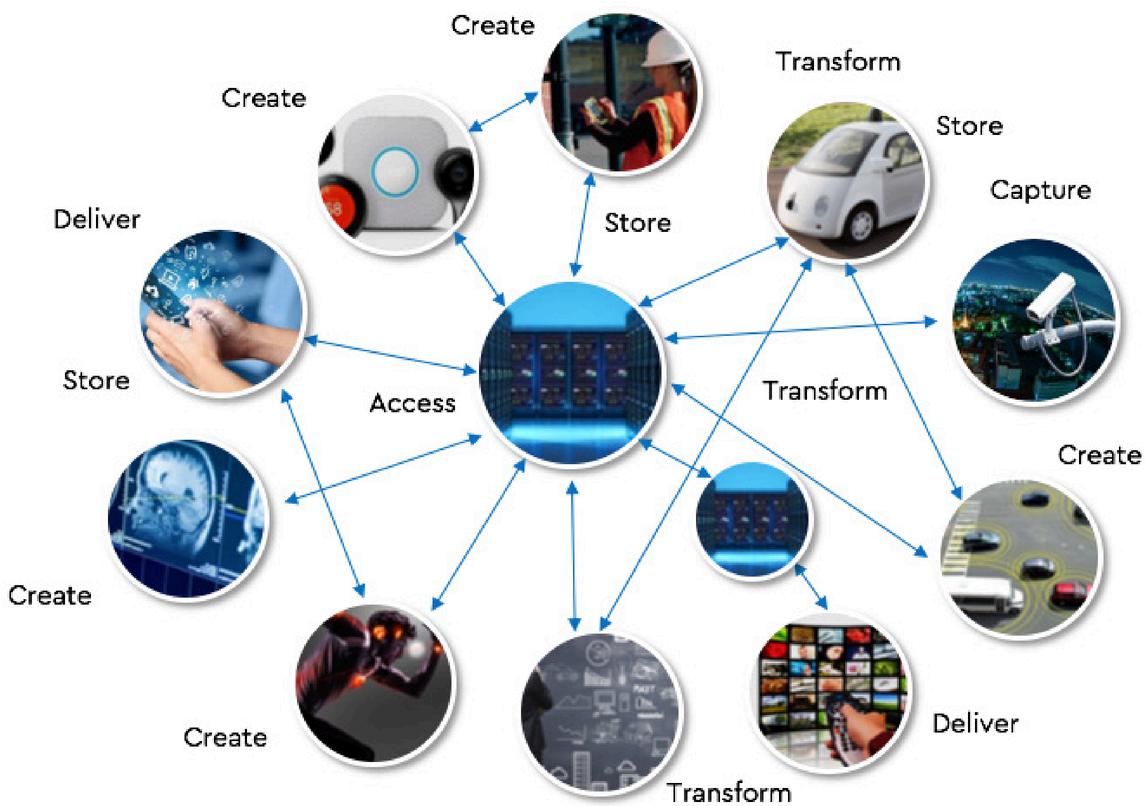
Looking Ahead

Big data in the core and fast data at the edge have exceeded the capabilities of traditional infrastructures and architectures especially relating to scalability, performance and efficiency. As general-purpose infrastructures are supported by a uniform ratio of resources to address all compute processing, storage and network bandwidth requirements, they are no longer effective for diverse data-intensive workloads. Western Digital's OpenFlex architecture is truly an open composable NVMe-oF platform that disaggregates compute, storage and network resources. Data center architects, cloud service providers, systems integrators, software-defined storage partners, OEMs, etc. can now deliver storage and compute services with greater economics, agility, efficiency and simplicity at scale, while enabling dynamic SLAs across workloads. Starting at the rack level, with federated management across multiple fabrics and sites, the OpenFlex architecture is based on the company's composable vision of:

- Leveraging an industry leader in both disk and flash storage to deliver a scalable, modular set of NVMe-oF storage devices based on a common interface
- Offering an open API and storage abstraction layer to accelerate time to value
- Reducing the complexities, inefficiencies and performance bottlenecks of current storage offerings to deliver optimum cost and performance
- Enabling CDIs from next-generation container technologies

Today's data-centric architectures require more control over the blend of resources that each application needs so that optimized levels of processing, storage and network bandwidth can be delivered and scaled independent of one another. These extreme workloads demand an infrastructure that is both flexible and composable. Western Digital is in a leading position to drive NVMe-oF-based CDI technology and market adoption forward and, together with an ecosystem, create environments for data to thrive – capturing, preserving, accessing and transforming it to its full potential. When data thrives, our people, communities and planet thrive as well, all through the power, potential and possibilities of data.

Current and Future
Data Pooled and Shared by Multiple Applications



About Western Digital

Western Digital creates environments for data to thrive. The company is driving the innovation needed to help customers capture, preserve, access and transform an ever-increasing diversity of data. Everywhere data lives, from advanced data centers to mobile sensors to personal devices, our industry-leading solutions deliver the possibilities of data. Western Digital data-centric solutions are marketed under the G-Technology™, SanDisk® and WD® brands.

Key Definitions

API	Application Programming Interface that features a set of subroutine definitions, protocols, tools and clearly defined methods for communicating with various components. In the case of SCIs, storage and other IT resources can be dynamically provisioned and managed through an open API without physically touching hardware. Open means publicly available to developers.
Container	A controlling element for an application instance that runs within a type of virtualization scheme called container-based virtualization. The container incorporates the application software, as well as instructions on how to self-provision the required resources for the application using virtualization or an SCI.
HCI	Hyper-Converged Infrastructure is a data center architectural framework that combines storage, compute and networking into a single system to reduce data center complexities and increase scalability. These platforms include a hypervisor for virtualized computing, software-defined storage and virtualized networking, and run on standard off-the-shelf servers. Multiple nodes can be clustered together to create pools of shared compute and storage resources.
Noisy Neighbors	When data is moved into a public cloud, performance issues may occur from other data being on the same cloud. Public clouds provide multitenant storage, so the sharing of resources also includes competing for the same dedicated servers that reside in the cloud provider's data center.
NVMe	Non-Volatile Memory Express is a streamlined, high-performance, low-latency interface that utilizes an architecture and set of protocols developed specifically for persistent flash memory technologies.
NVMe-of	The NVMe-oF protocol has been extended beyond local-attached server applications and across a network through the NVMe-over-Fabrics specification, and enables flash devices (such as PCIe SSDs and storage arrays) to communicate over networks, such as InfiniBand and Converged Ethernet, delivering the same high-performance, low-latency benefits as local-attached NVMe.
Performance Predictability	In a traditional storage infrastructure, applications compete for resources from a fixed number of IOPS and the devices perform exactly as expected. In an SCI, there is no wasted or unused IOPS as they are available when needed, guaranteeing performance without intervention.
Provisioning	The process of assigning storage capacity to servers, computers, virtual machines or any other computing device.
RDMA	RDMA Remote Direct Memory Access is when memory from one system or computer is provided to another without involving either operating system, enabling high-throughput, low-latency networking. Common RDMA implementations include the Virtual Interface Architecture, RDMA over Converged Ethernet (RoCE), InfiniBand, Omni-Path and internet Wide-Area RDMA Protocol (iWARP).

Western Digital.

5601 Great Oaks Parkway
San Jose, CA 95119, USA
www.westerndigital.com