Open Composable Architecture & OpenFlex™

Sridhar Sabesan
Director, Platform Engineering

06.15.2019

Consume. Collaborate. Contribute.

Increasingly Dynamic Workloads

A survey of mid-sized and large-enterprise IT users found...

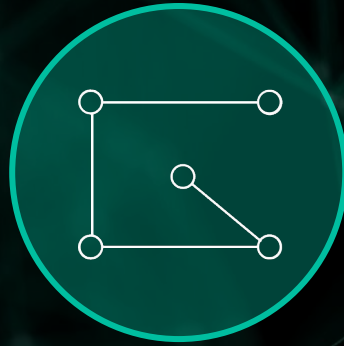
45%

of compute hours and
storage capacity are
utilized

70%

report inefficiencies
in the time required to provision
compute and storage resources

Western Digital

6/15/19   2

Driving New Demands on Data Infrastructure

Scalability

Efficiency

Agility

Performance

Composable Infrastructure

The What	• Data center infrastructure that **disaggregates** compute, storage, and network resources into shared pools that can be composed for on-demand allocation
The How	• Hardware Disaggregation – **OpenFlex hardware** • Infrastructure Composability – **Open Composable API** & orchestration software
The Why	• Greater productivity, agility, performance and faster time-to-market

Western Digital's Open Composable API

- Designed for data center composability
  - Logical composability of resources abstracted from the underlying physical hardware
  - Discovers, assembles, and composes self-virtualized resources via peer-to-peer communication
- Enables virtual system composition in existing HCI and next-generation SCI environments
  - Futureproofs the transition from hyper-converged to disaggregated architectures
  - Complements existing Redfish®/Swordfish™ usage
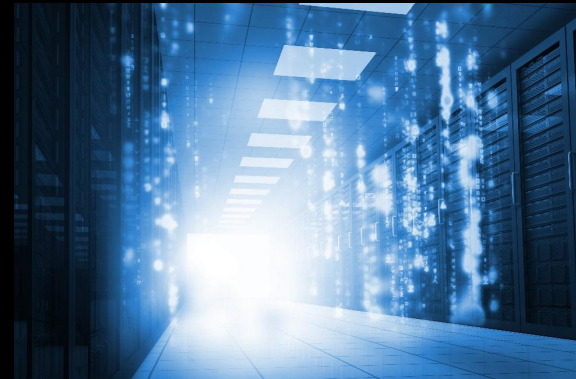
Our Composable Infrastructure Vision

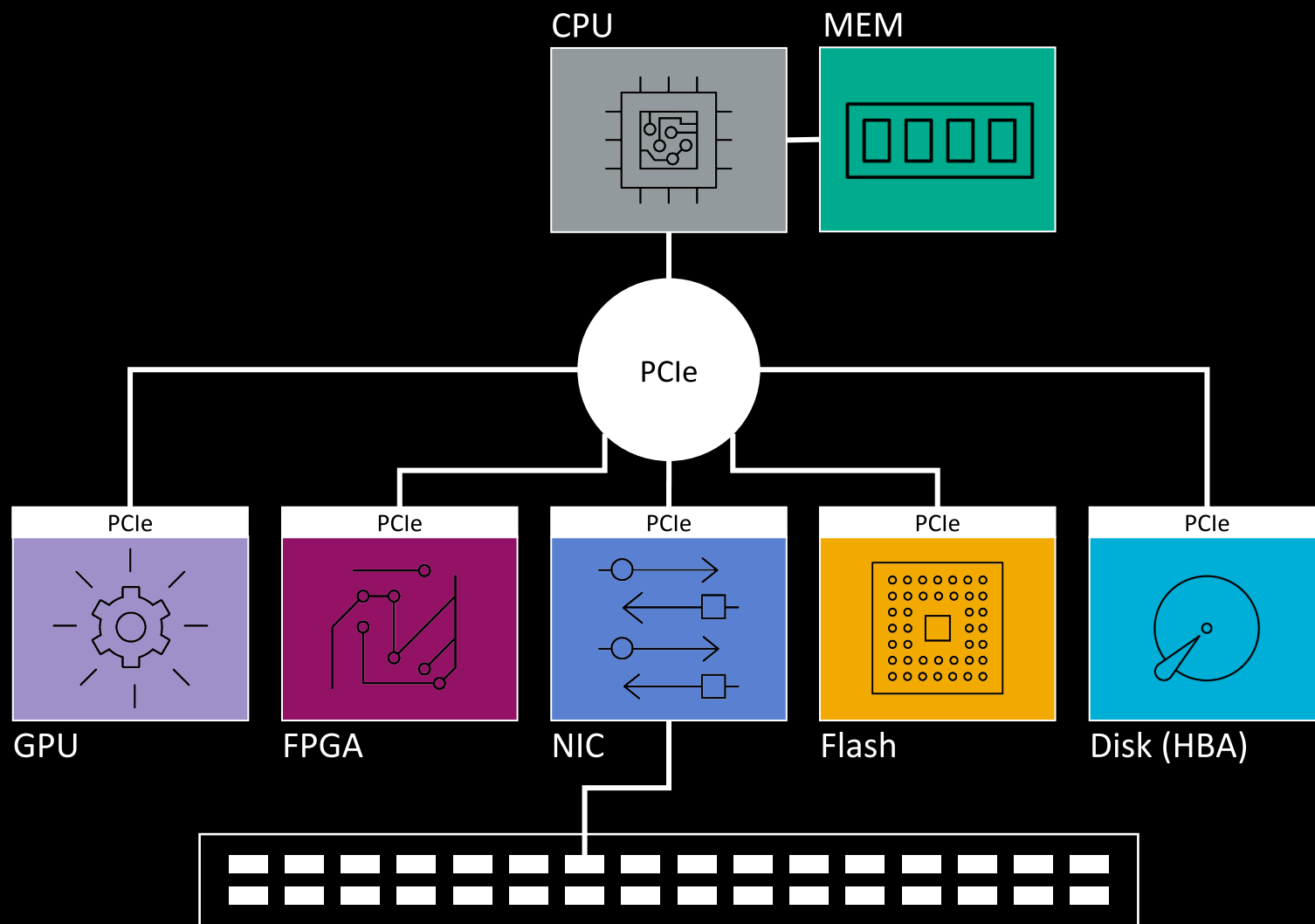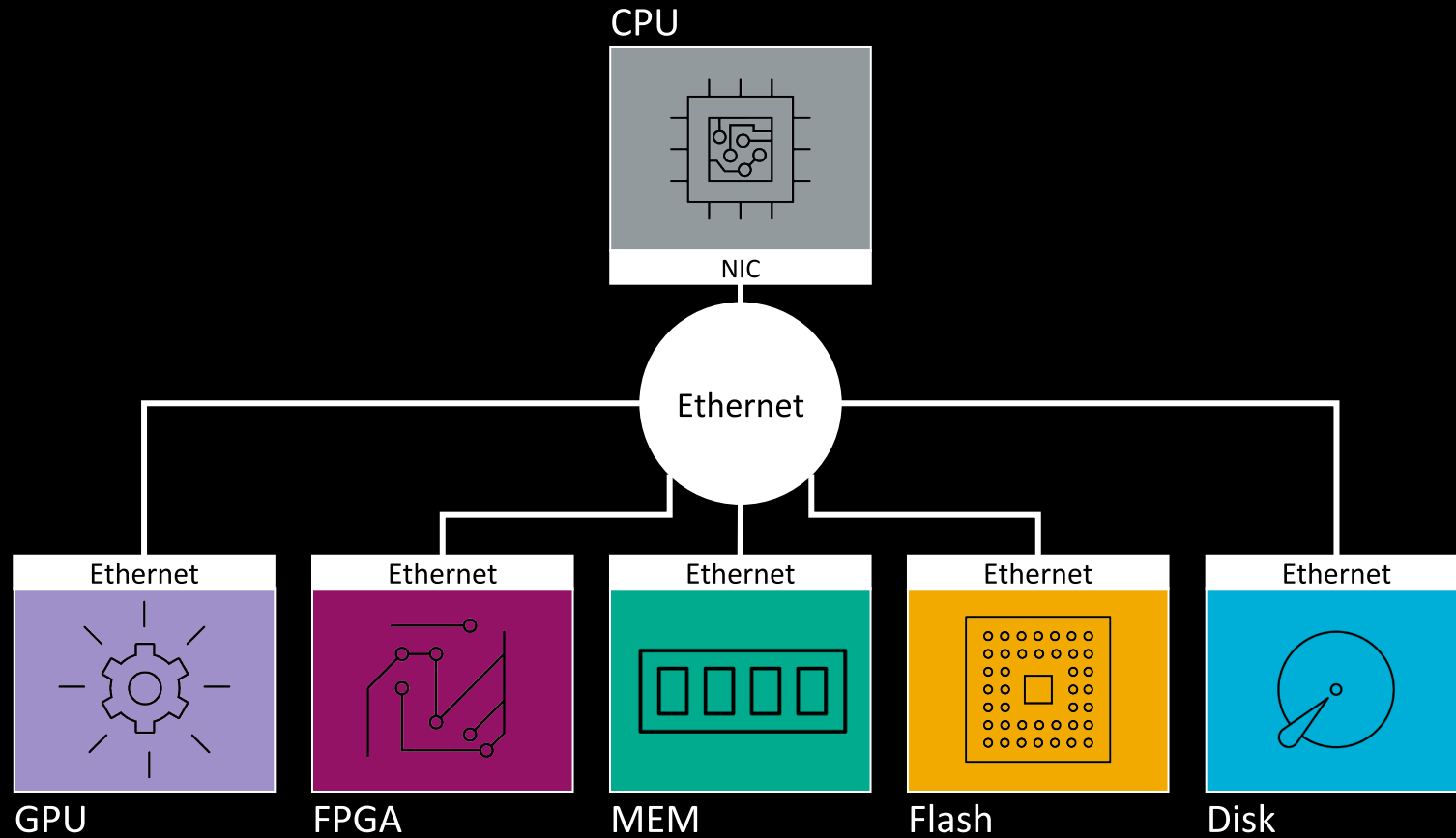| Open	Scalable	Disaggregated	Extensible
Open in both form factor and API for management and orchestration of composable resources	Independent performance and capacity scaling from rack-level to multi-rack	True disaggregation of storage and compute for independent scaling to maximize efficiency, agility and to reduce TCO	Flash, disk and future composable entities can be independently scaled, managed and shared over the same fabric

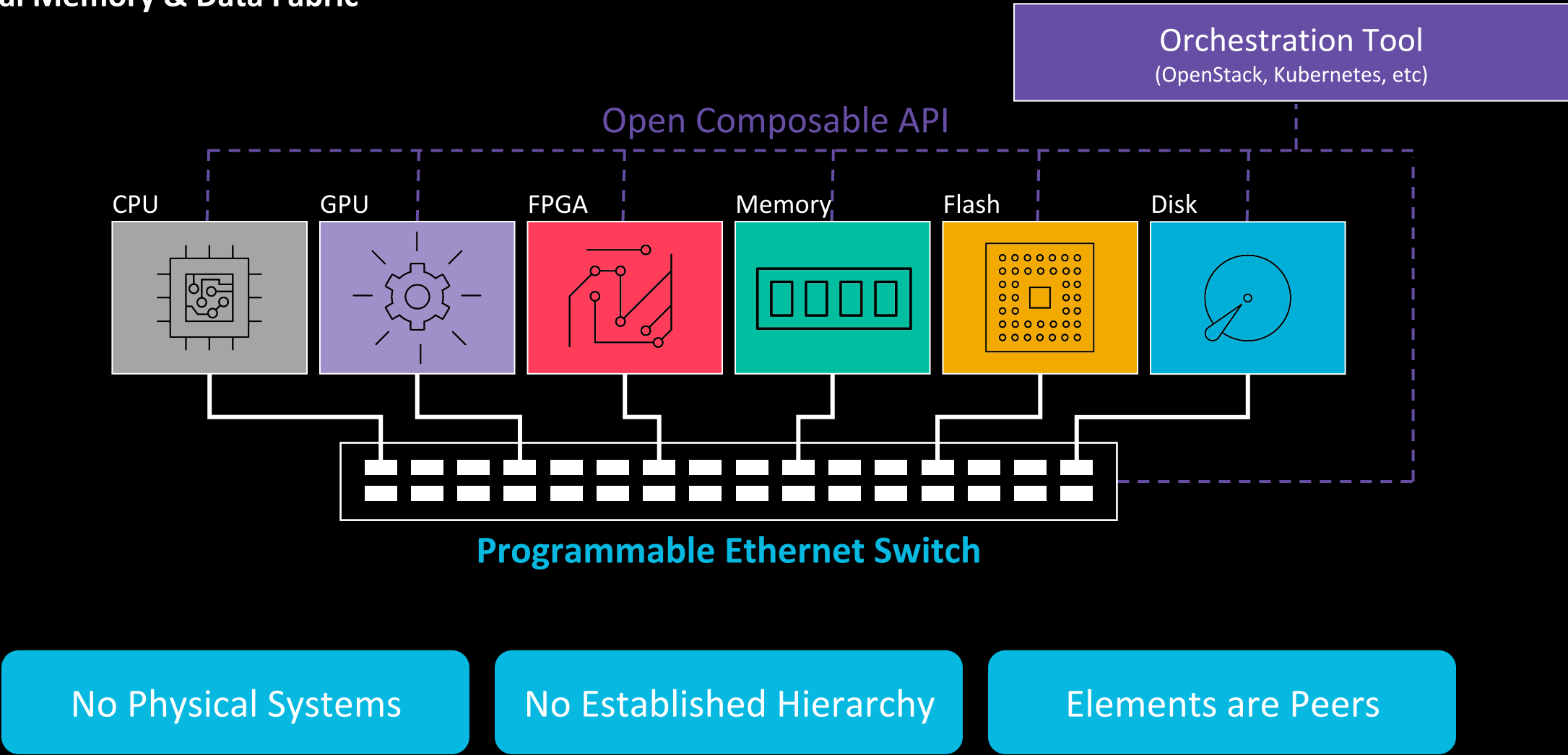# System Transformation

From Direct Attached...

System Transformation
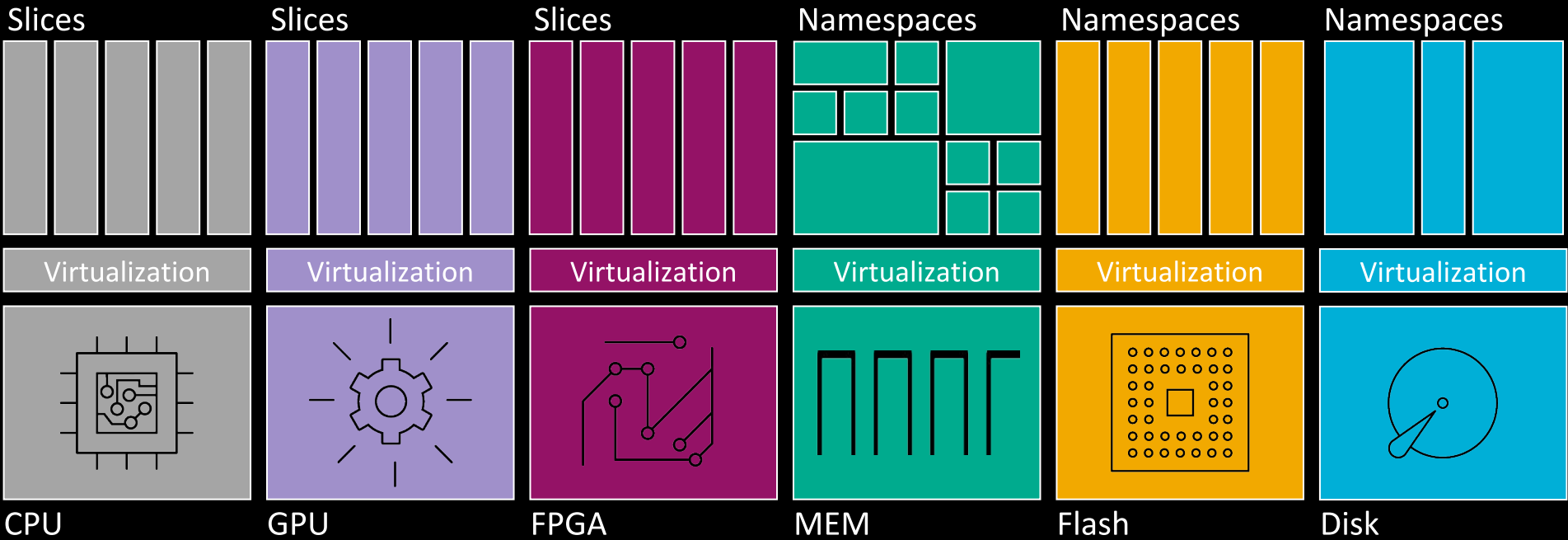
To Shareable, Using NVMe-over-Fabric

CPU

NIC

Ethernet

Ethernet | Ethernet | Ethernet | Ethernet | Ethernet

GPU | FPGA | MEM | Flash | Disk

Shared Accelerated Storage

Enabling Composable Infrastructure

Universal Memory & Data Fabric

Orchestration Tool
(OpenStack, Kubernetes, etc)

Open Composable API

| CPU | GPU | FPGA | Memory | Flash | Disk |

Programmable Ethernet Switch

No Physical Systems

No Established Hierarchy

Elements are Peers
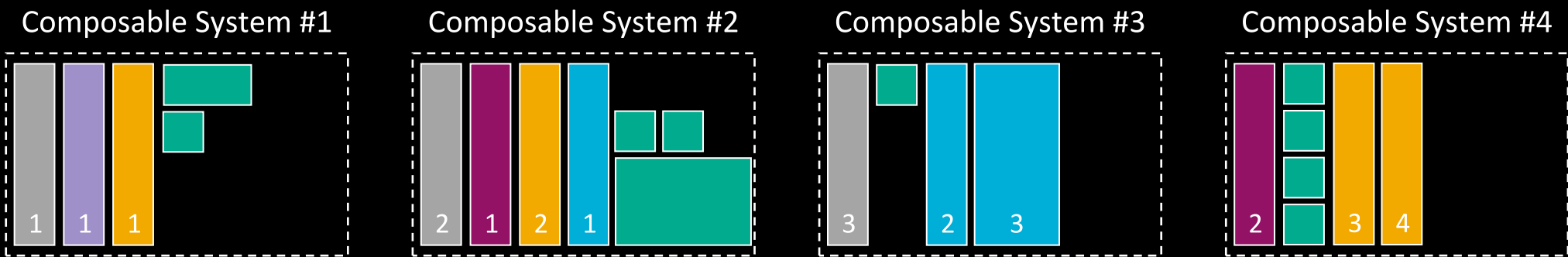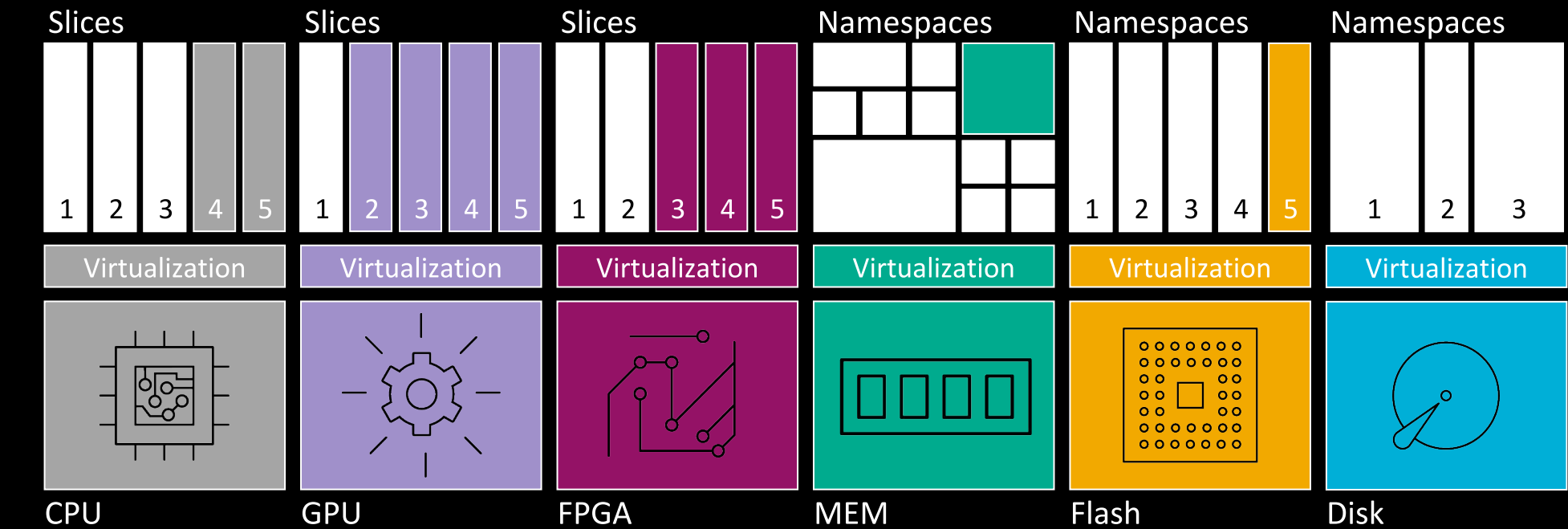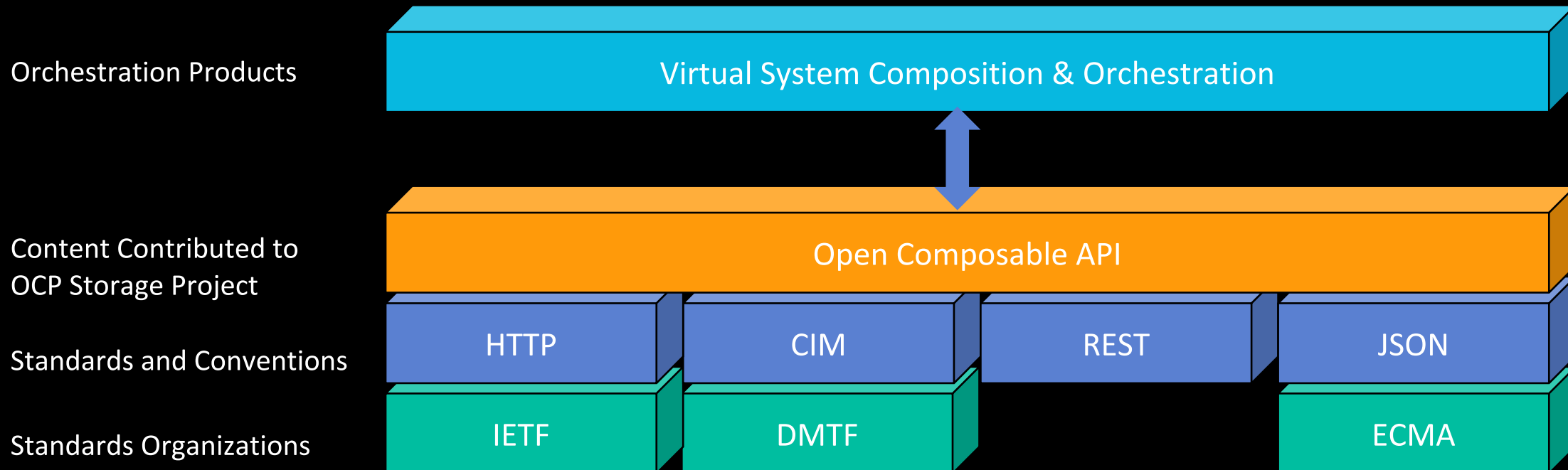
Delivering Real-Time Resource Allocation

Delivering Real-Time Resource Allocation

Open Composable API Ecosystem

Existing Open Standards are the foundation of the Open Composable API

Orchestration Products

Virtual System Composition & Orchestration

Content Contributed to
OCP Storage Project

Open Composable API

Standards and Conventions

| HTTP | CIM | REST | JSON |

Standards Organizations

| IETF | DMTF | | ECMA |

HTTP = HyperText Transfer Protocol
CIM = Common Information Model
REST = REpresentational State Transfer
JSON = JavaScript Object Notation
IETF = Internet Engineering Task Force
DMTF = Distributed Management Task Force
ECMA = European Computer Manufacturers Association or European association for standardizing information and communication systems

Open Composable API

REST based commands to discover and compose virtual systems

Device Discovery
GET http://<ip>/Query
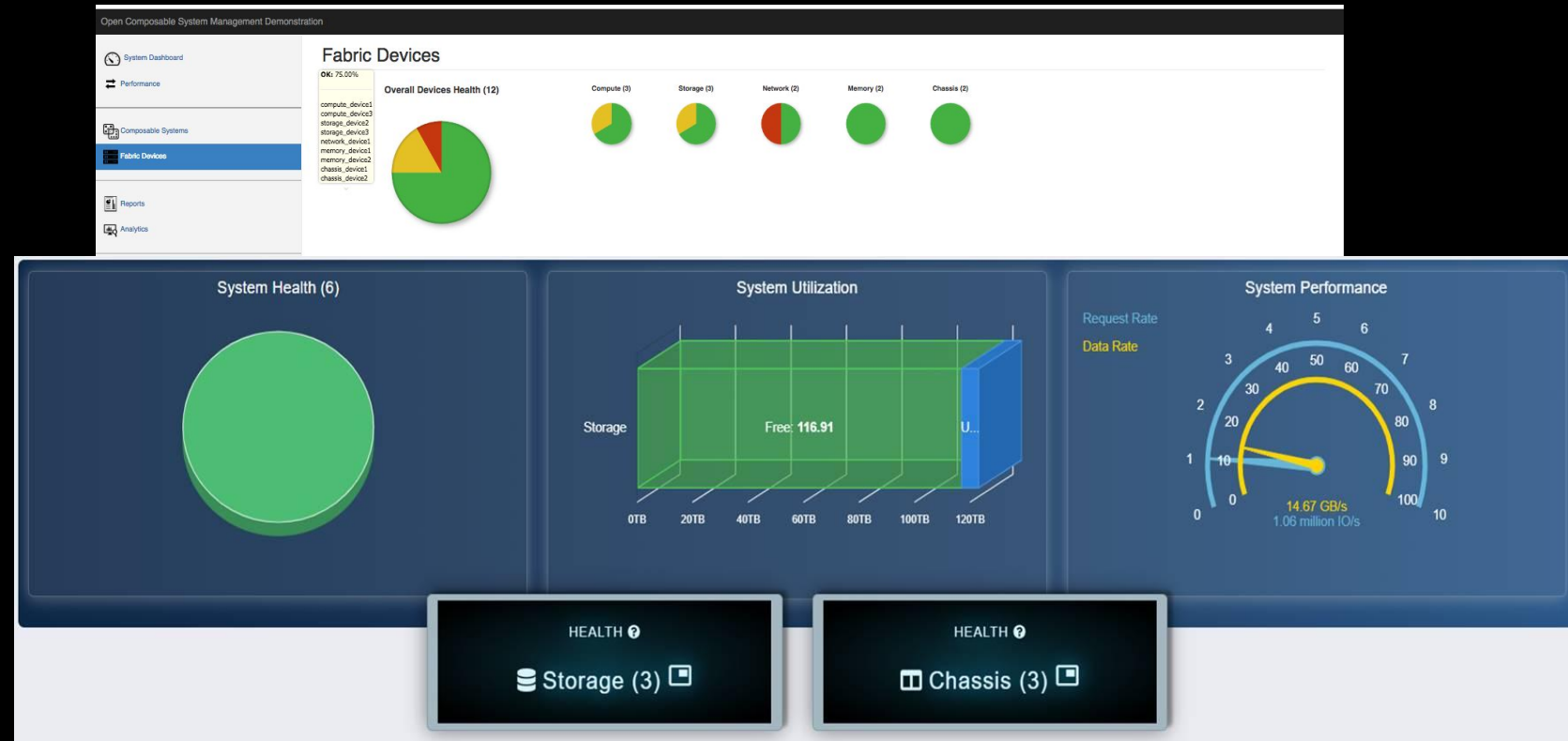
System Discovery
/System/Query

Compose Systems
GET /System/Composites

Create Storage Volumes
GET /Storage/Devices/{id}/Volumes/{id}

Compose your virtual systems with one API

Email inquiries to OpenComposableAPI@wdc.com

Introducing OpenFlex™ Composable Infrastructure

All-Flash Fabric Device and Enclosure
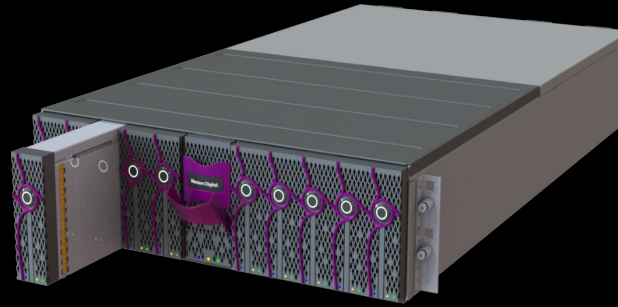
TRUE HARDWARE DISAGGREGATION @ SCALE ON -DEMAND

◄— **OpenFlex E3000** —►

High performance, low latency for fast data

OpenFlex | **NVMe-over-Fabric | Infrastructure Disaggregation | Software Composable**

NVMf Fabric Devices

OpenFlex F3000 Fabric Device and E3000 Enclosure

Dual-port, high-performance, low-latency fabric-attached SSD

Self-virtualized device with up to 256 namespaces per Fabric device for dynamic provisioning

3U enclosure with 10 dual-port slots offering up to 614 TB

Multiple storage tiers over the same wire – Flash and Disk(future) accessed via NVMf
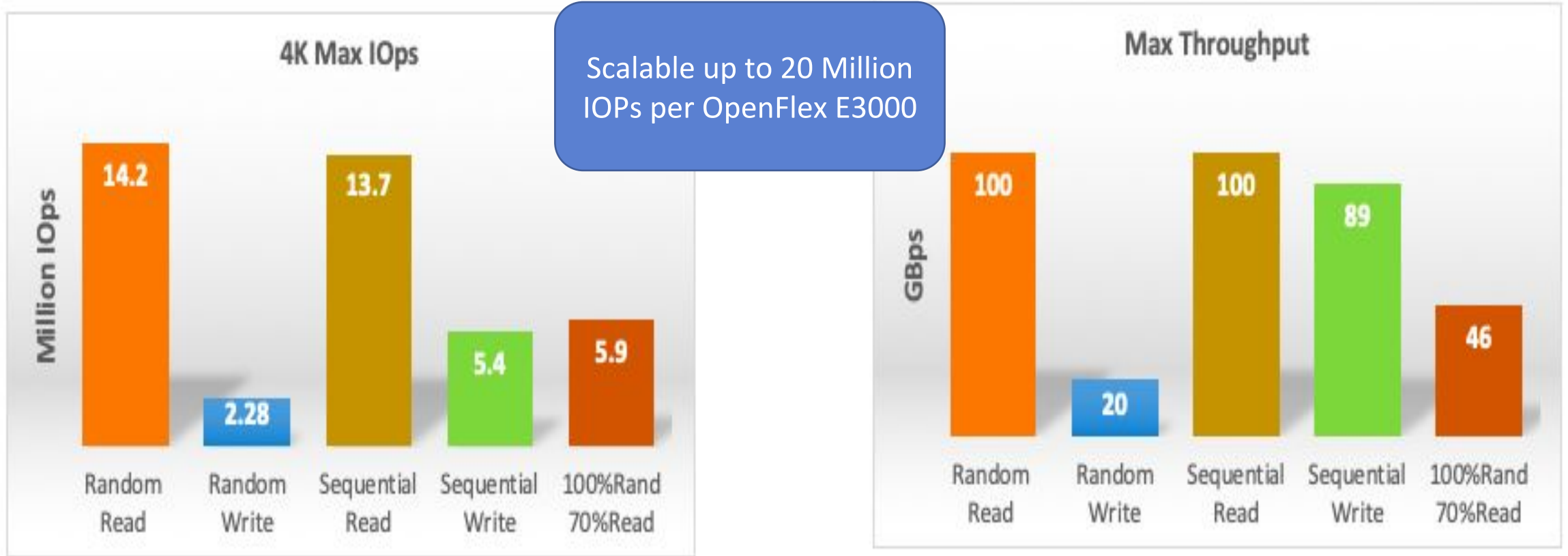
OpenFlex

NVMe-over-Fabrics | Infrastructure Disaggregation | Software Composable

OpenFlex *F3000 Fabric Device and E3000 Enclosure*

Optimizing Data Center Bridging features in high speed Ethernet Fabrics

✔**2 x 50G per Fabric Device scaling up to 20 X 50G per OpenFlex E3000**

✔**Reduced CPU cycles on the HOST server enabling better bandwidth usage and higher scalability.**

✔**OpenFlex^{TM} F3000 - Extensive Interop tests**

✔**Support Multiple Ethernet Switching & NIC Vendors**

✔**Ecosystem & Partner Enablement**

OpenFlex E3000 Performance

4K Max IOps

Million IOps

- 14.2 — Random Read
- 2.28 — Random Write
- 13.7 — Sequential Read
- 5.4 — Sequential Write
- 5.9 — 100%Rand 70%Read

Scalable up to 20 Million IOPs per OpenFlex E3000

Max Throughput

GBps

- 100 — Random Read
- 20 — Random Write
- 100 — Sequential Read
- 89 — Sequential Write
- 46 — 100%Rand 70%Read

Scalable Performance with incremental Fabric Devices.
OpenFlex E3000 capable of scaling up to 20 Million IOPs with High Performance SKUs

Fabric Comparison

Match the Fabric to the Required Service Level

	NVMe/RoCE	NVMe/TCP	NVMe/FC
Technology	RDMA over Ethernet	Standard TCP over Ethernet	Fibre Channel
Max Link Speed Today	100G	100G	32G
Link Aggregation	Yes	Yes	Limited MDIO based
Next Gen Speed	400G	400G	128G
End-to-End Latency	Lowest	Low	Med
Performance	Highest	High	Med
Encapsulation	UDP	TCP	FC
Routability	Routable UDP based	Routable TCP based	Limited
Scale	Multi Rack	Multi Rack	Limited
Convergence with other traffic	Yes	Yes	No
Challenges	HoL Blocking (PFC only) Fabric Configuration Fabric Tuning Fabric Debug	HoL Blocking Delayed Acks Increase Latency Incast Problems Lack of HW Acceleration*	Zoning Challenges Incast Problems DC Silos Network Mgmt Overhead Inferior Roadmap

OCP Storage Project

Forming Open Composable Architectures Sub-Project

- Seeking your feedback / interest from the OCP community on forming a sub-project

- WDC Contributions
  – F3000 Mechanical Design and Electrical Interface
  – Open Composable API Specification

- VPN to our lab
  – Virtual access to compute, fabric, and F3000 storage

- API Emulator Access for the community – *email OpenComposableAPI@wdc.com*

Join the conversation by subscribing to OCP-Storage@OCP-All.groups.io
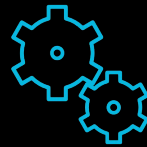
Western Digital OpenFlex

Positioned to accelerate market adoption

Open	Ecosystem	Vertical Innovation	Trust
Firm commitment to an open standards-based approach	Strategic position in the ecosystem to help accelerate market adoption	Vertical Innovation for Timely Productization of Technology	Trusted leader in data center products, technologies and infrastructure

Western Digital: Storage Technology and Product Leader

Sridhar.Sabesan@wdc.com

Vivek.Tyagi@wdc.com

Harsha.N@wdc.com

6/15/19