FLASH MEMORY SUMMIT 2018

The True Performance of Flash Storage

Flash Memory Summit

DEVELOPING LOW-LATENCY DATA SERVICES ON NVME-OF SHARED STORAGE

Presented by

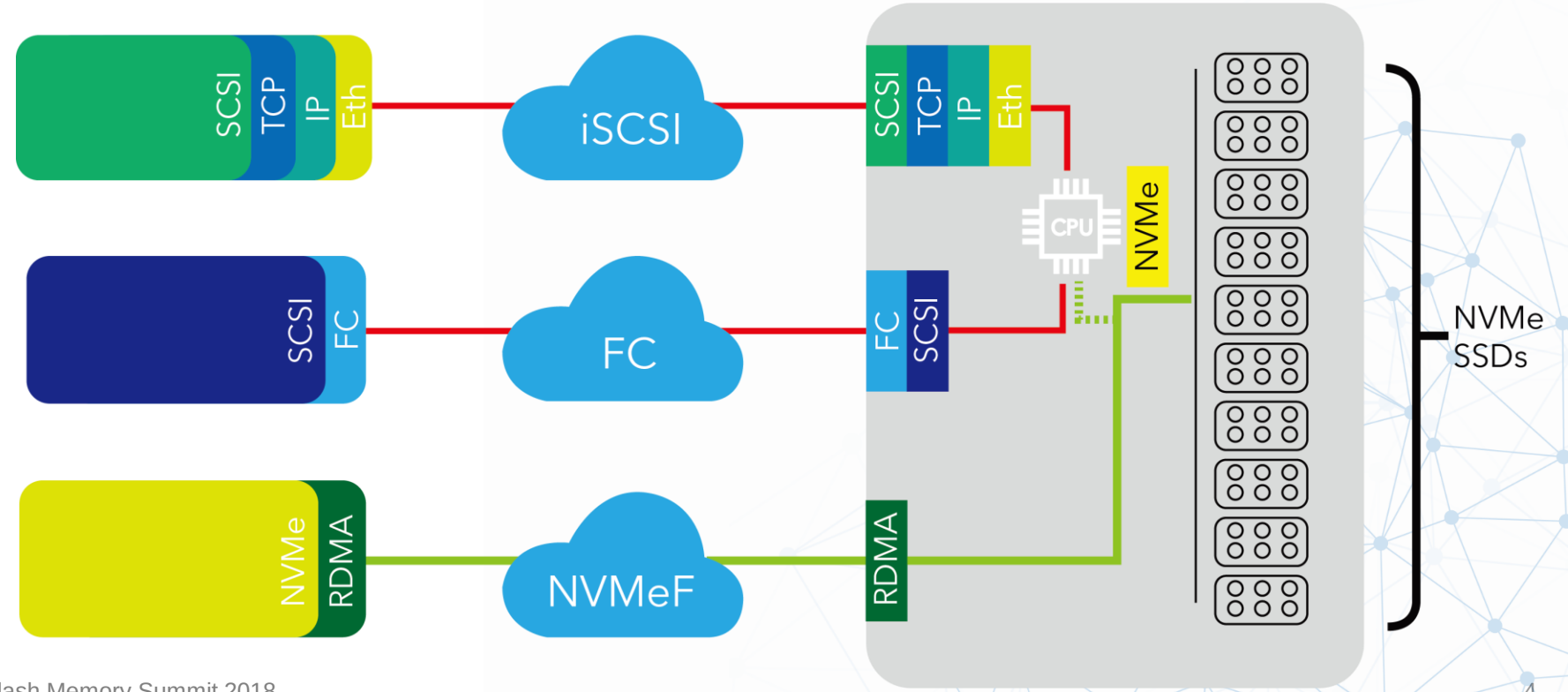
Chaan W Beard

TOSHIBA

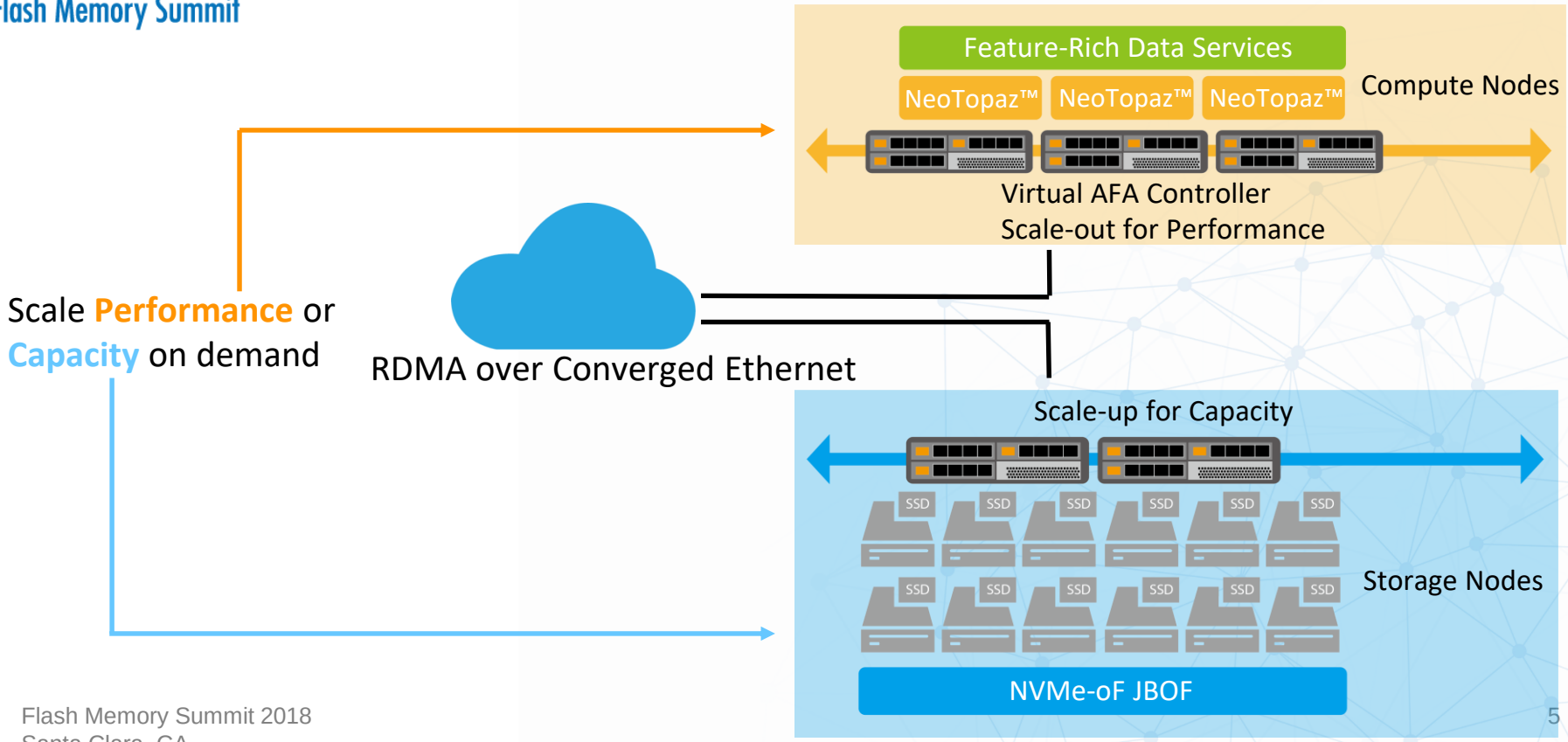
- Toshiba and AccelStor have been working jointly on NVMe-oF technology – we are grateful to Toshiba for the wonderful work and contributions they have made

- Our other SSD partners use Toshiba Memory in their products

NVMe-oF Changes the Rules for All-Flash Arrays



NVMe-oF Decouples Compute and Storage Nodes



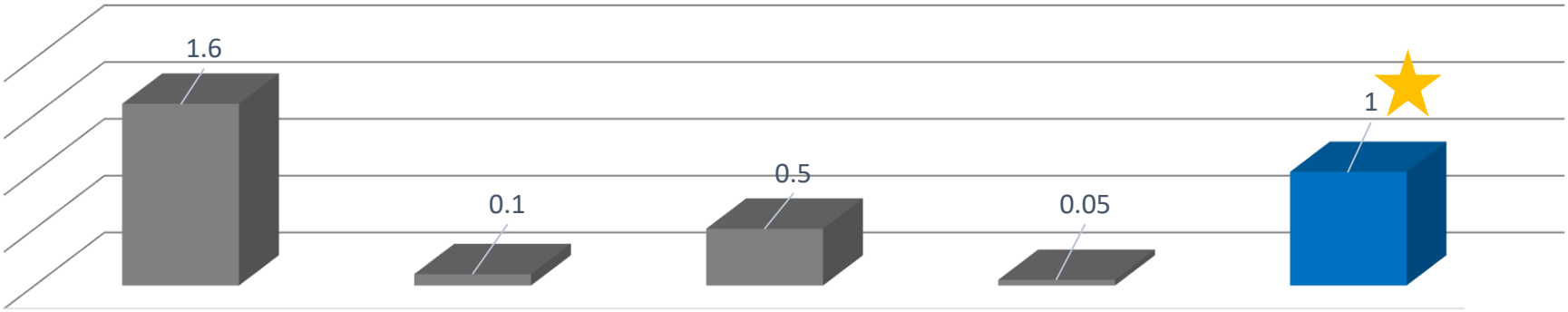
NVMe-oF Tools and Development Kit Overview

4KB Random Write IOPS (Million)/ Per Compute Node

Fast and Low Latency

Flexible Architecture

Feature-Rich Data Services



SPDK
without data services

SPDK
+ AIO
+ LVM
+ Linux NVMe Driver
in Physical Machine

FlexiRemap®
+ SPDK
+ AIO
+ Linux NVMe Driver
in Physical Machine

SPDK
+ AIO
+ LVM
+ Linux NVMe Driver
in Virtual Machine

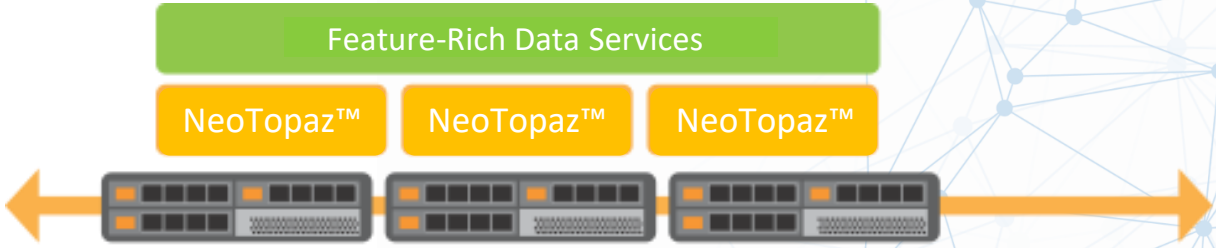
FlexiRemap®
+ SPDK
+ IF-AIO
in Virtual Machine

Virtual All-Flash Array, vFlexiArray™ Introduction

- NeoTopaz™ Virtual Controller
- Virtual NVMe SSD Namespace
- Virtual AFA creation
- Data Services same as physical Array
  (Snaps, Clones, thin/thick provision etc.)
- Fast, Flexible, and Feature-Rich
- Hyperscale now becomes a reality
- High Performance and Data Protection can now be a Guaranteed Service

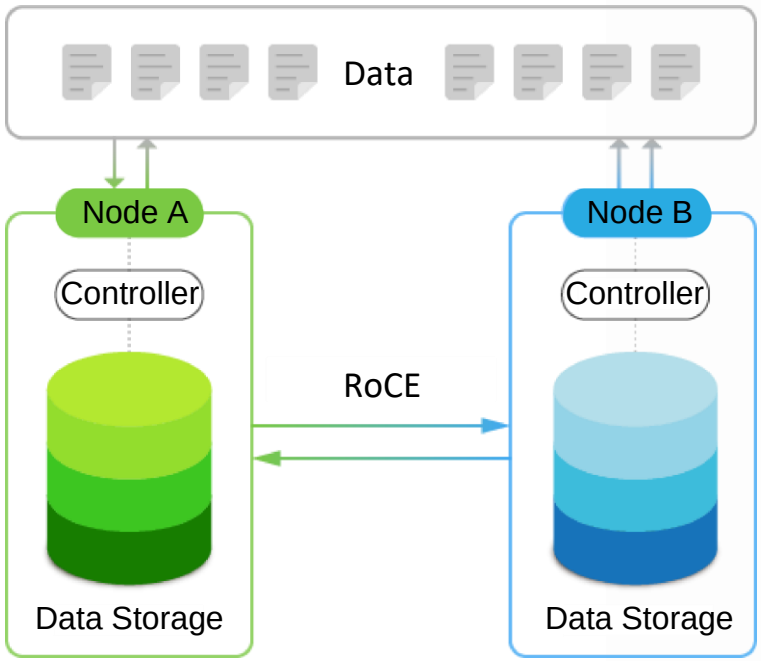
NeoTopaz™ Virtual Controllers

- Real virtualization not logical
- All the data services run in NeoTopaz™ Virtual Machine Storage OS
- Isolated, secure and reliable
- Communication via 100GB RoCE
- Delivers 1M IOPS @4K random writes, and scales up to 40M IOPS

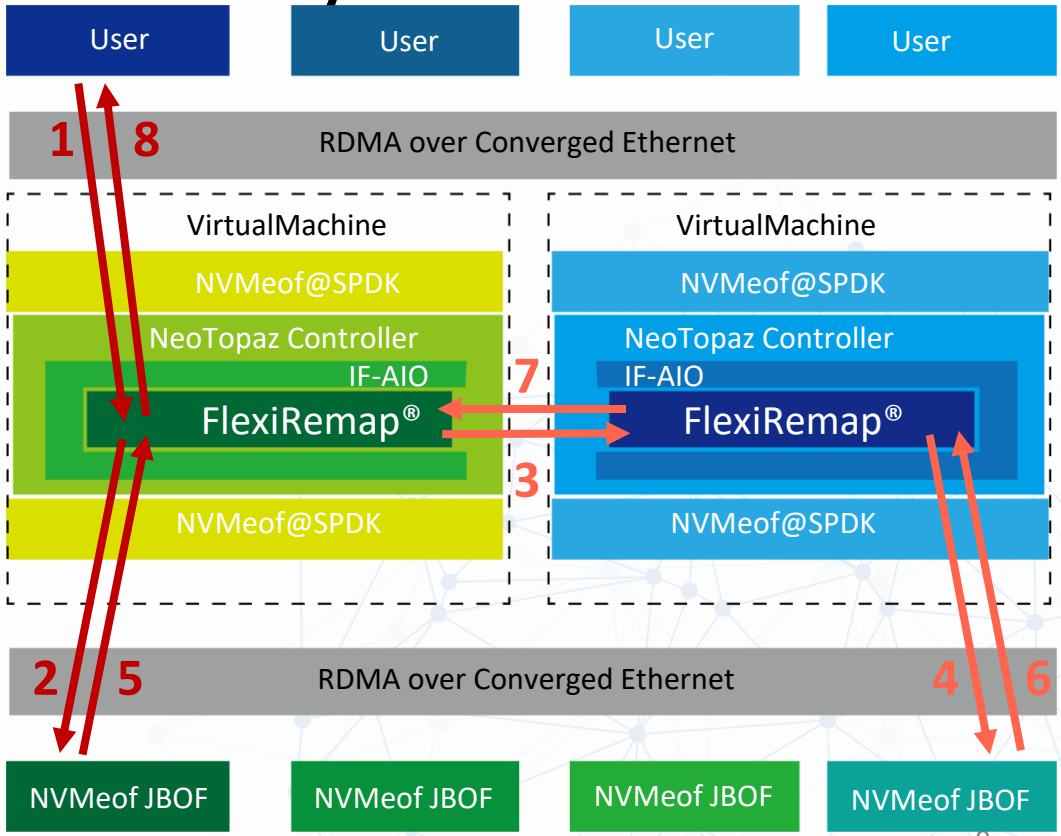


Virtual AFA Controller
Scale-out for Performance

NeoTopaz™ High Availability Solution

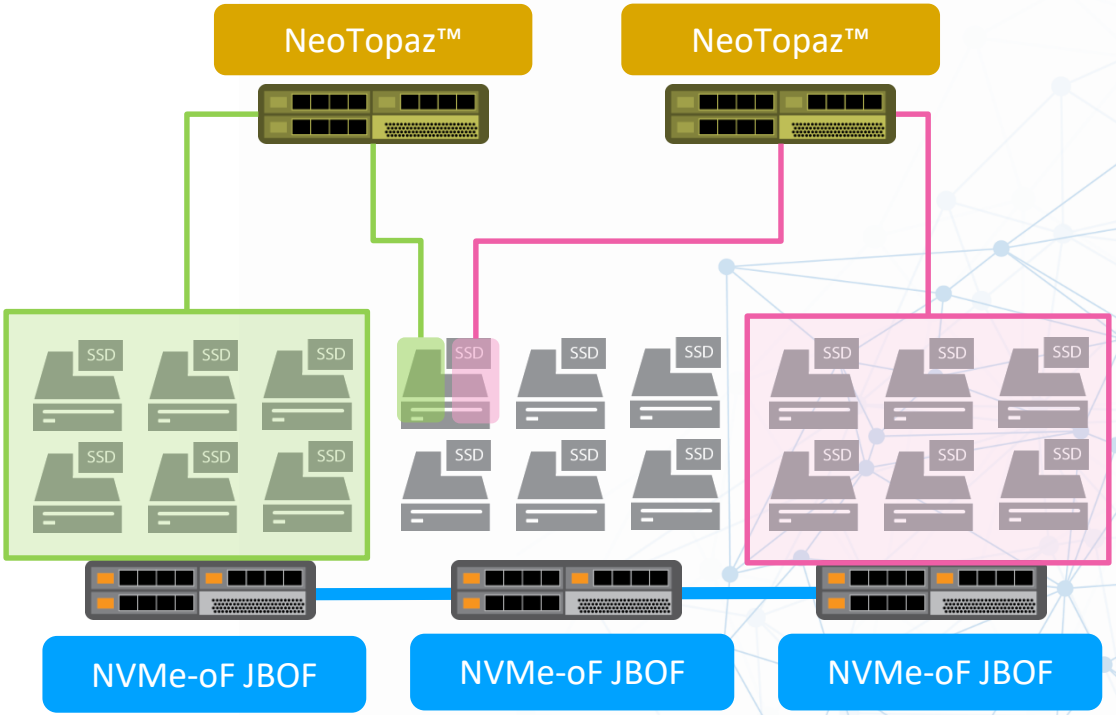


Shared-Nothing HA

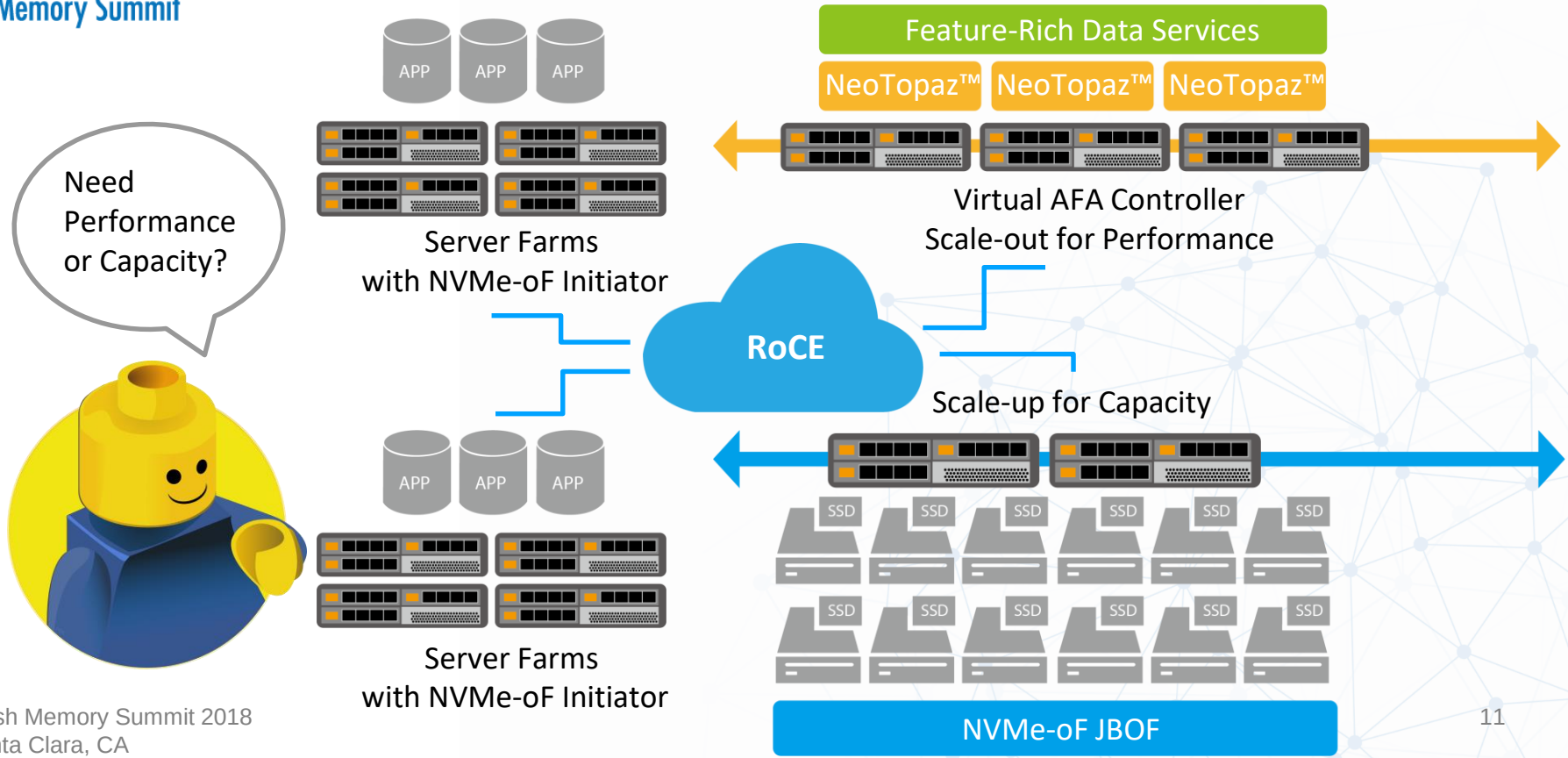


Data flow of vFlexiArray™

Virtual SSD Namespace



vFlexiArray™ Creation



Flash Memory Summit

Fast, Flexible, and Feature-Rich

Fast, Flexible, and Feature-Rich

Build a high capacity - low IOPS vFlexiArray™ using Virtual SSD Pools with required SSD characteristic
(Fewer SSD with more capacity)

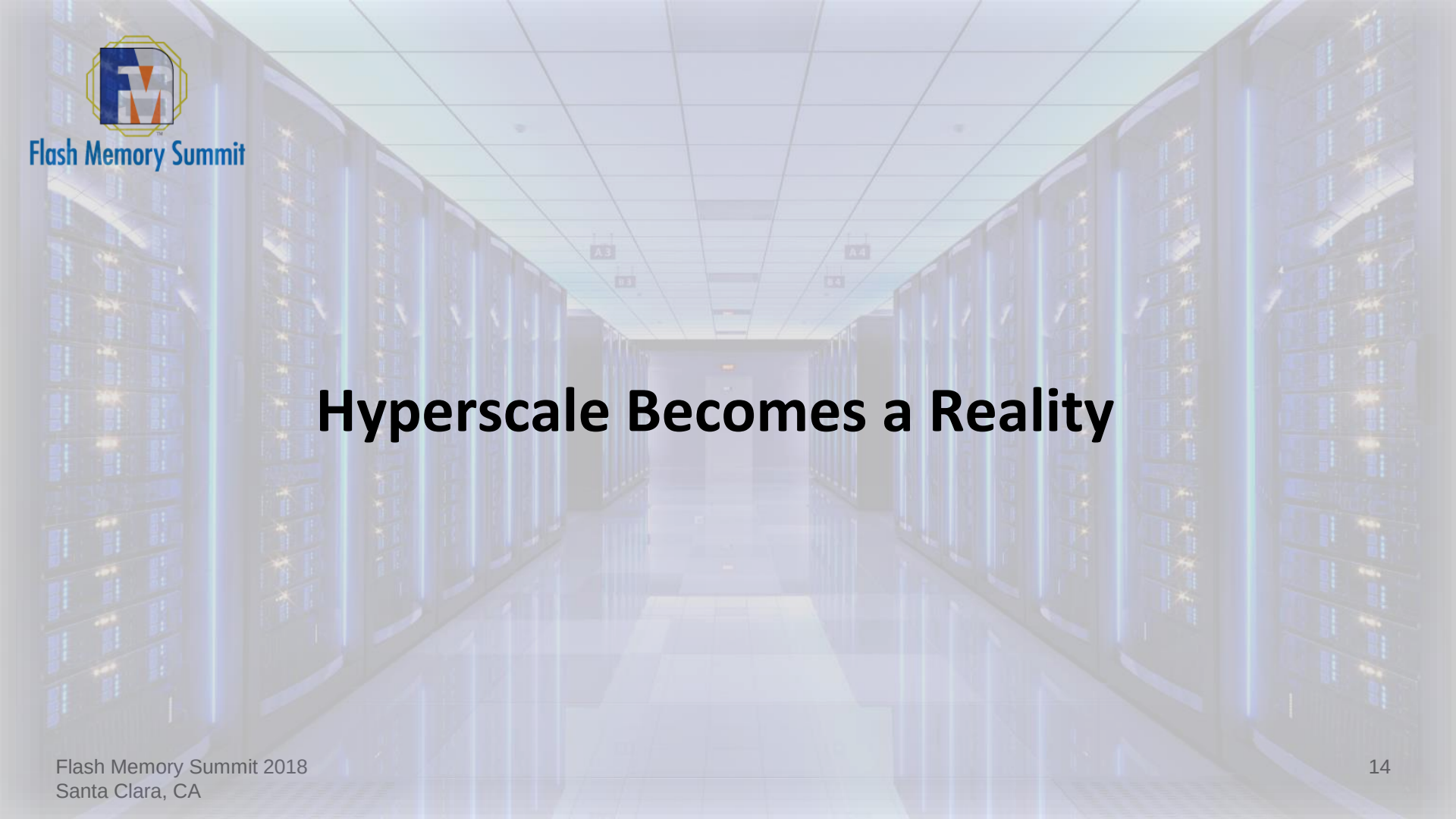
Add in HA via configuration

Build a high IOPS, small Capacity vFlexiArray™ with tailored, high performance SSD

So granular even a single large capacity and high IOPS SSD can be shared by multiple vFlexiArray™ via the Namespace feature

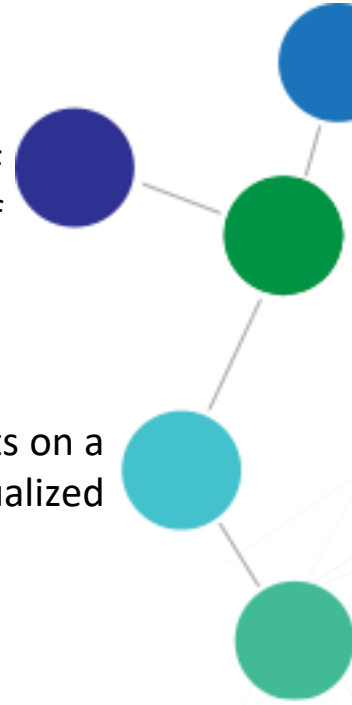
Tailor the vFlexiArray™ any way you desire with any SSD spec

Feature-rich data service: (FlexiRemap®,FlexiDedupe™, FlexiVirtualArray™, FlexiSnap™, FlexiClone™ etc.)



Hyperscale Becomes a Reality

Hyperscale Becomes a Reality



Unlimited extension of Capacity

Relationship between controller and SSD liberated by NVMe-oF network itself

A vFlexiArray™ can dynamically add any number of SSD in the pool

No limit of the number of JBOF ports on a vFlexiArray™ since it is virtualized

Expansion Simple and easy, just plug in the new JBOF to the RDMA bus, the management system detects and updates using REST API automatically!

Flash Memory Summit

High Performance & Data Protection

Guaranteed!

High Performance and Data Protection Guaranteed

Every vFlexiArray™ can provide up to 1 Million IOPS with either WRITE or READ random operations (@4K) with FULL Data Protection and HA built in

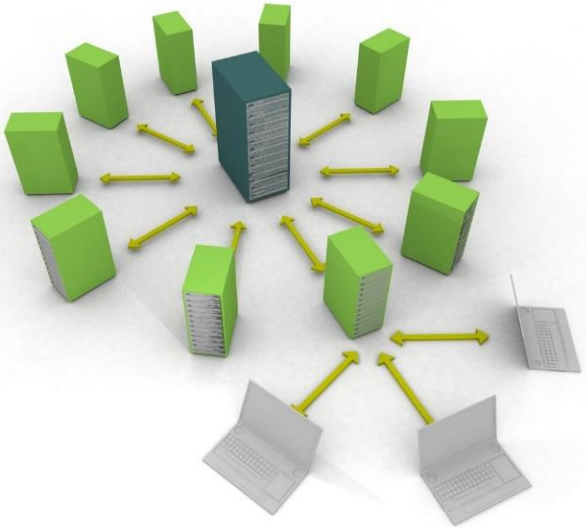
Our technology guarantees there will be no IOPS degradation using our vFlexiArray™

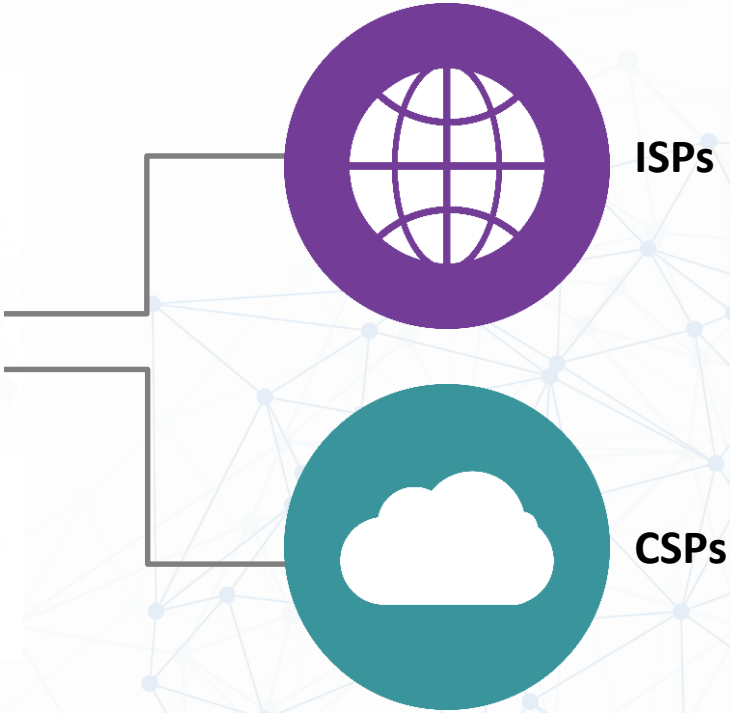
Performance also tailored by adding computing nodes as required

No longer need expensive storage hardware to service full HA needs

vFlexiArray™ Use Cases

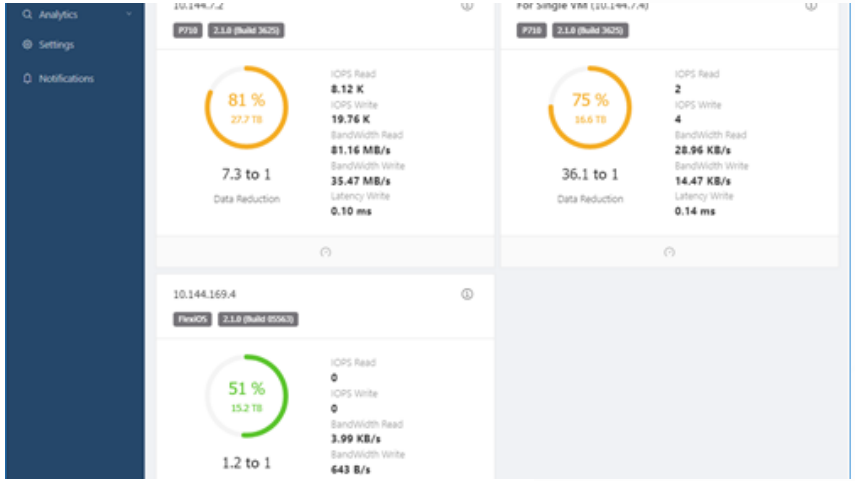
Multi-tenant data centers

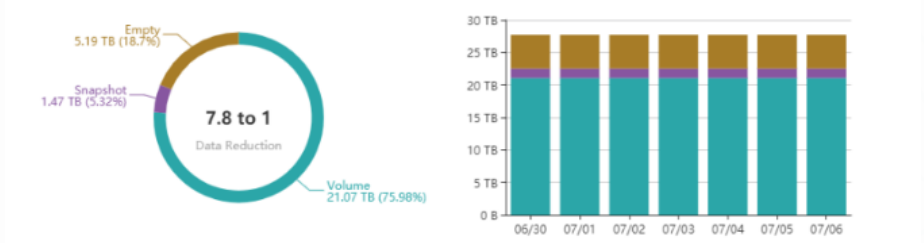


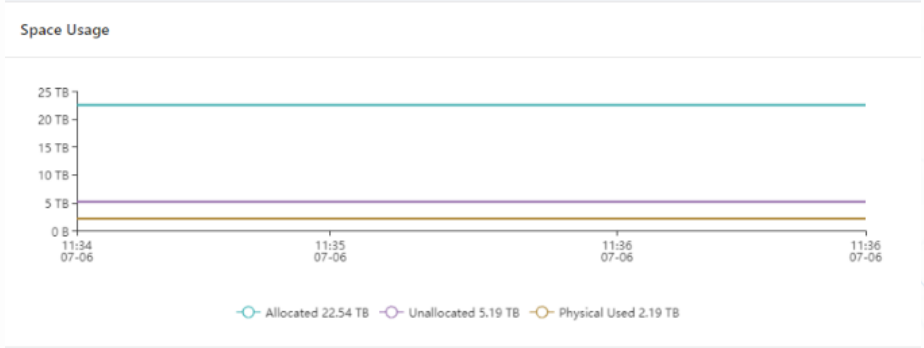


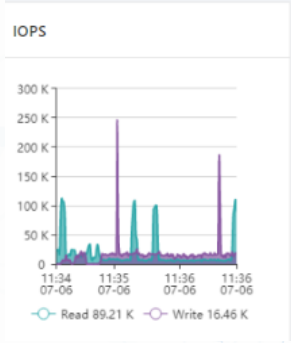
Flash Memory Summit

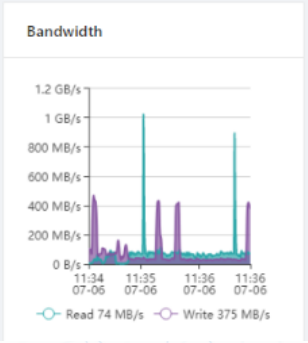
User Interface

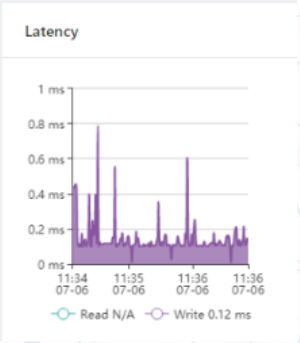












Flash Memory Summit

Thank You

Free Software Resources

- Install Linux Centos 7.2 or above for target-side
- Install Linux Centos 7.2 or above and upgrade kernel to 4.8 for initial-side
- Download SPDK: lock-free software
- Multiple modules in SPDK:
  nvme driver, nvmeof target, iscsi target.
- User-layer application
- Replace the role of the Linux driver for nvme

Missed Features