Evolution of Rack Scale Architecture Storage

Murugasamy (Sammy) Nachimuthu, Principal Engineer
Mohan J Kumar, Fellow
Intel Corporation

Agenda

- Introduction to Intel® Rack Scale Design
- Storage in Intel® Rack Scale Design (Today and Future)
- Intel® Rack Scale Design Storage Orchestration
- Summary

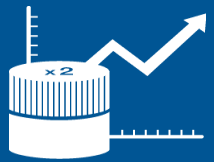
Agenda

- Introduction to Intel® Rack Scale Design
- Storage in Intel® Rack Scale Design (Today and Future)
- Intel® Rack Scale Design Storage Orchestration
- Summary

Data Center Challenges

Infrastructure has not kept up with increasing business demands

Inefficiency
Less than 50% server utilization[2]



Growth
Data growth doubles every 18 months[1]

Agility
New services can take a week or more to provision[1]

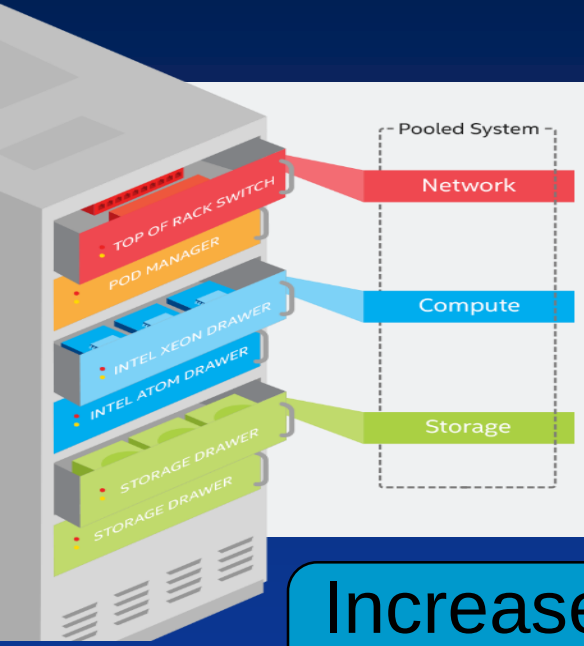
Business Needs

- **Reduce** operational and capital expenses.
- **Deliver** new services in minutes, not months.
- **Optimize** data center based on real-time analytics.
- **Address** application workload needs with agility.
- **Scale** capacity without interruption

1 *Worldwide and Regional Public IT Cloud Services 2013–2017 Forecast.* IDC (August 2013) idc.com/getdoc.jsp?containerId=242464
2 IDC's Digital Universe Study, sponsored by EMC, December 2012

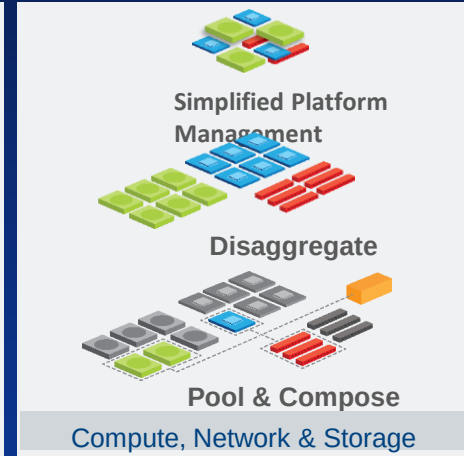
Intel® Rack Scale Design

Logical architecture for efficiently building and managing cloud infrastructure—and providing the simplest path to a software defined data center.



Benefits

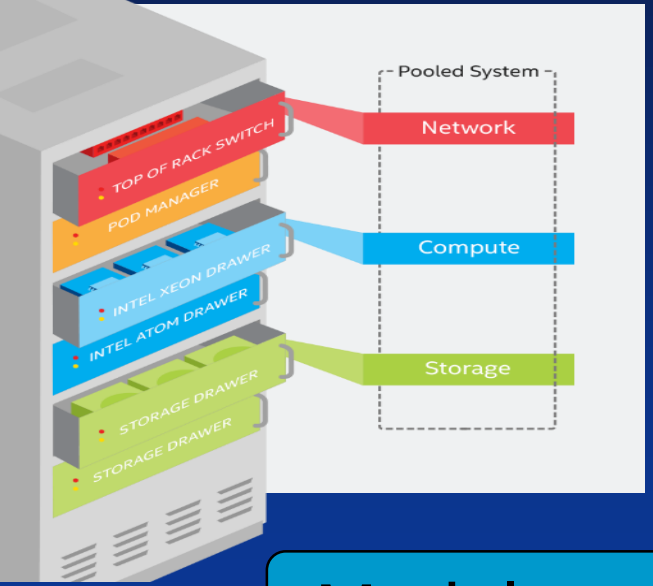
- >25% decrease in capital costs
- Increase capacity/IT $
- Reduce time to cloud deployment

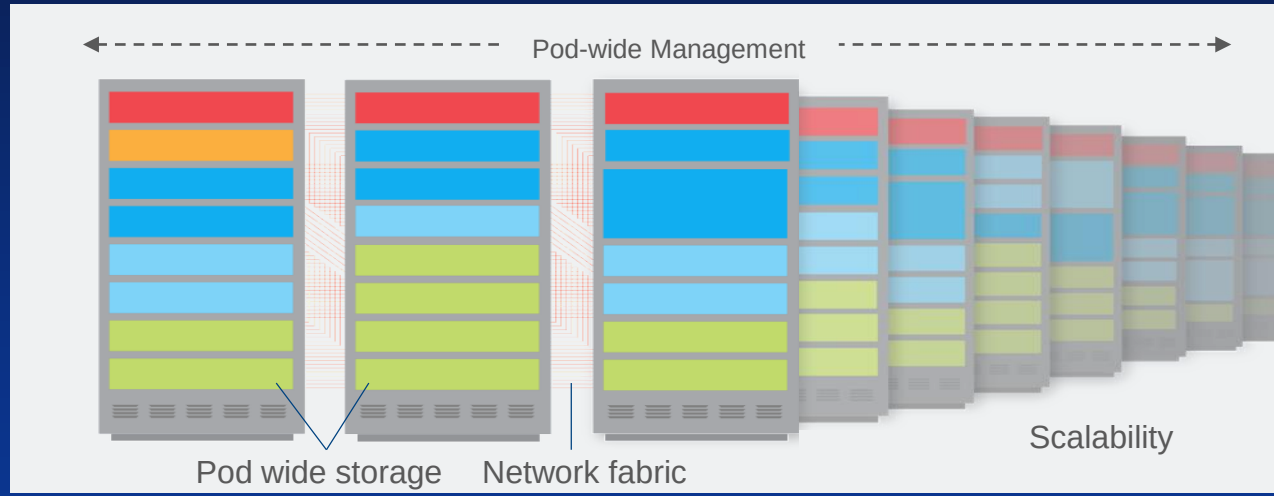


Increase performance per TCO$ & accelerate cloud adoption

Intel® Rack Scale Design Framework

1. Pooled systems 2. Pod management 3. Network fabric 4. Pod-wide Storage



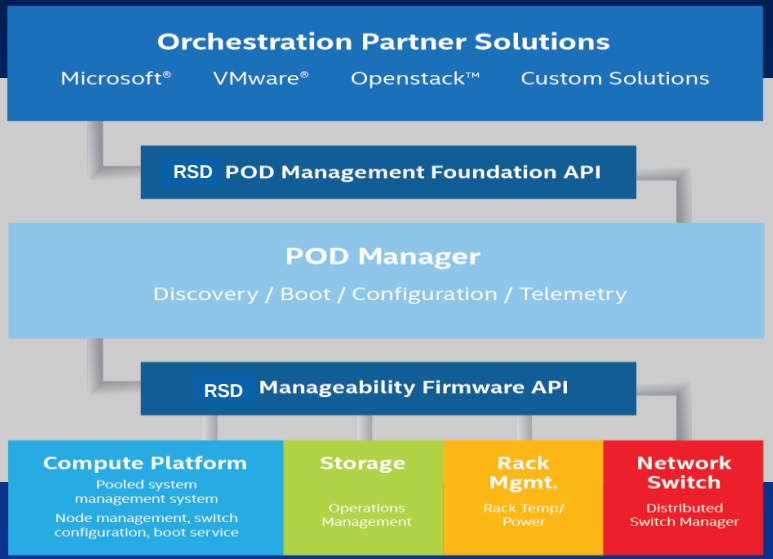


Modular scalable management architecture

Management Software Framework

Flexible management architecture allowing for range of implementation options

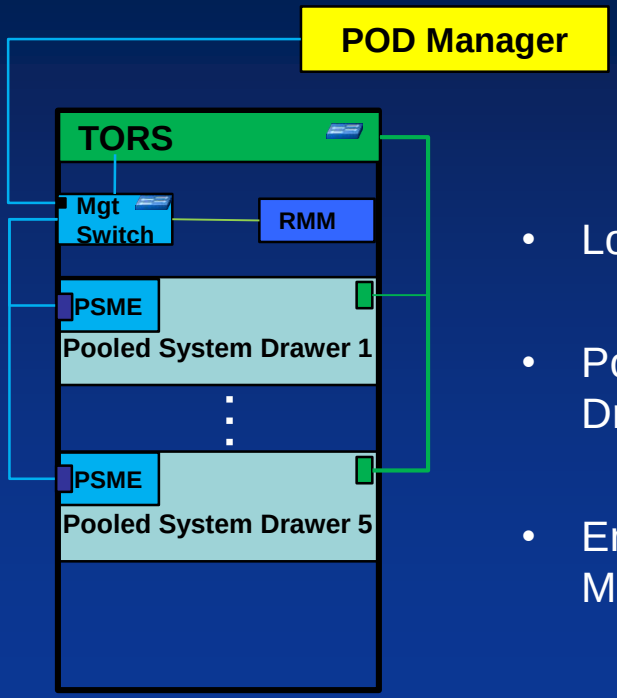
- Asset & location discovery
- Disaggregated resource management
- Composable system support
- Support compute, network, and storage,



Comprehensive management architecture

RSD Management - Example

Redundancy allowed,
Not Shown



RSD Rack

Rack Private Management Network

- NIC attached to external network
- NIC attached to private management network
- Switch
- Services outside the Rack Scale Rack

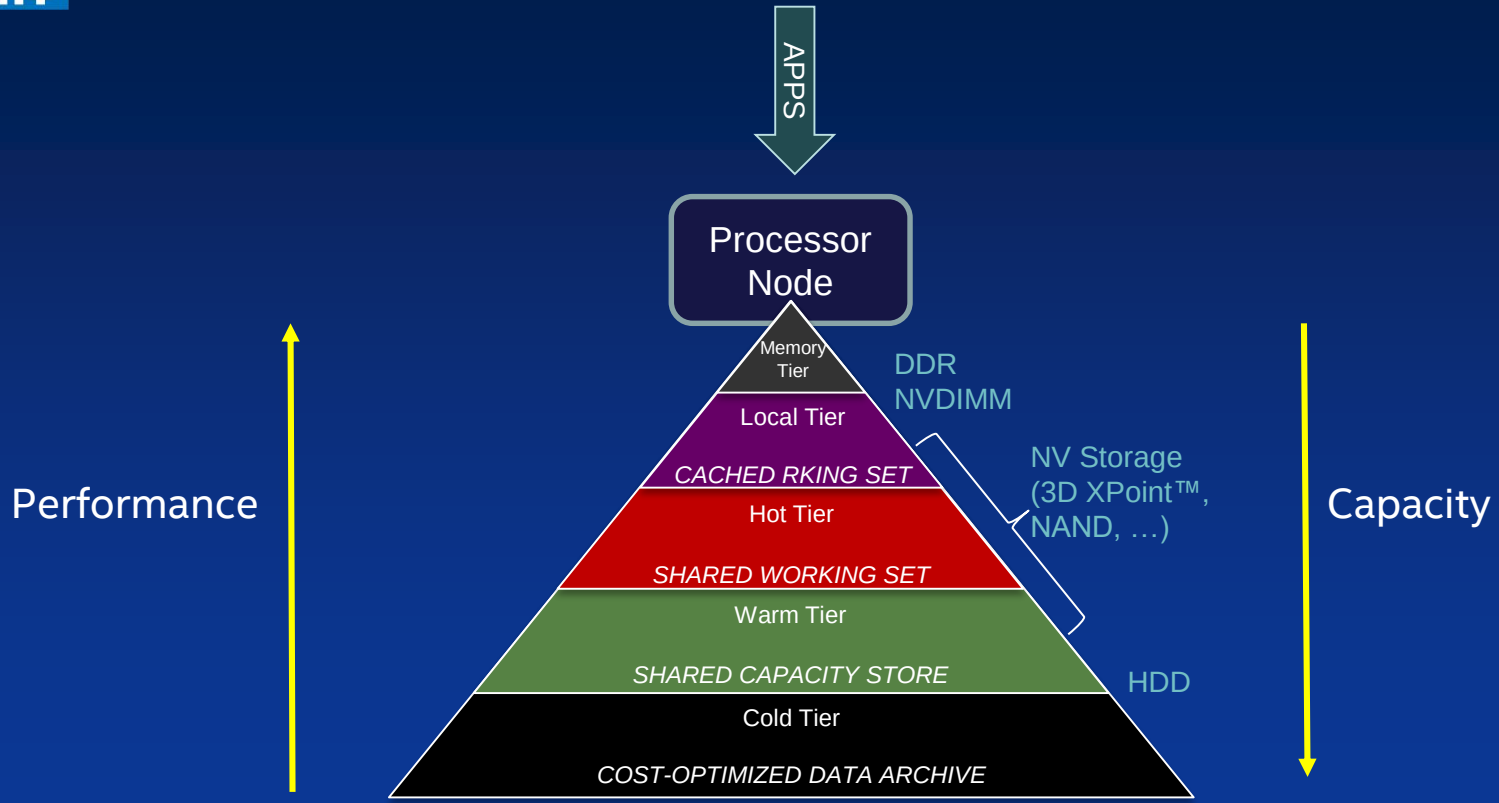
- Location aware compute, network, storage
- Pooled System Management Engine per Rack Scale Drawer (Multi-node aggregation)
- Environmental Management using Rack Management Module

PSME – Pooled System Management Engine
RMM – Rack Management Module

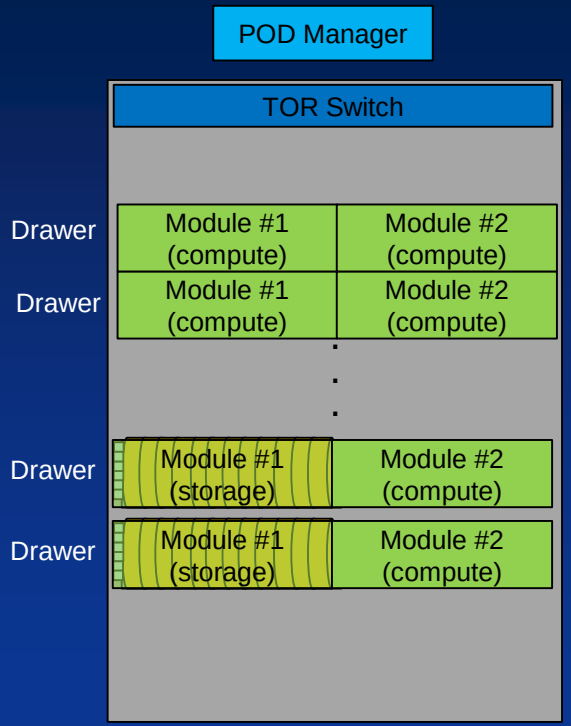
Agenda

- Introduction to Intel® Rack Scale Design
- Storage in Intel® Rack Scale Design (Today and Future)
- Intel® Rack Scale Design Storage Orchestration
- Summary

Storage Trends

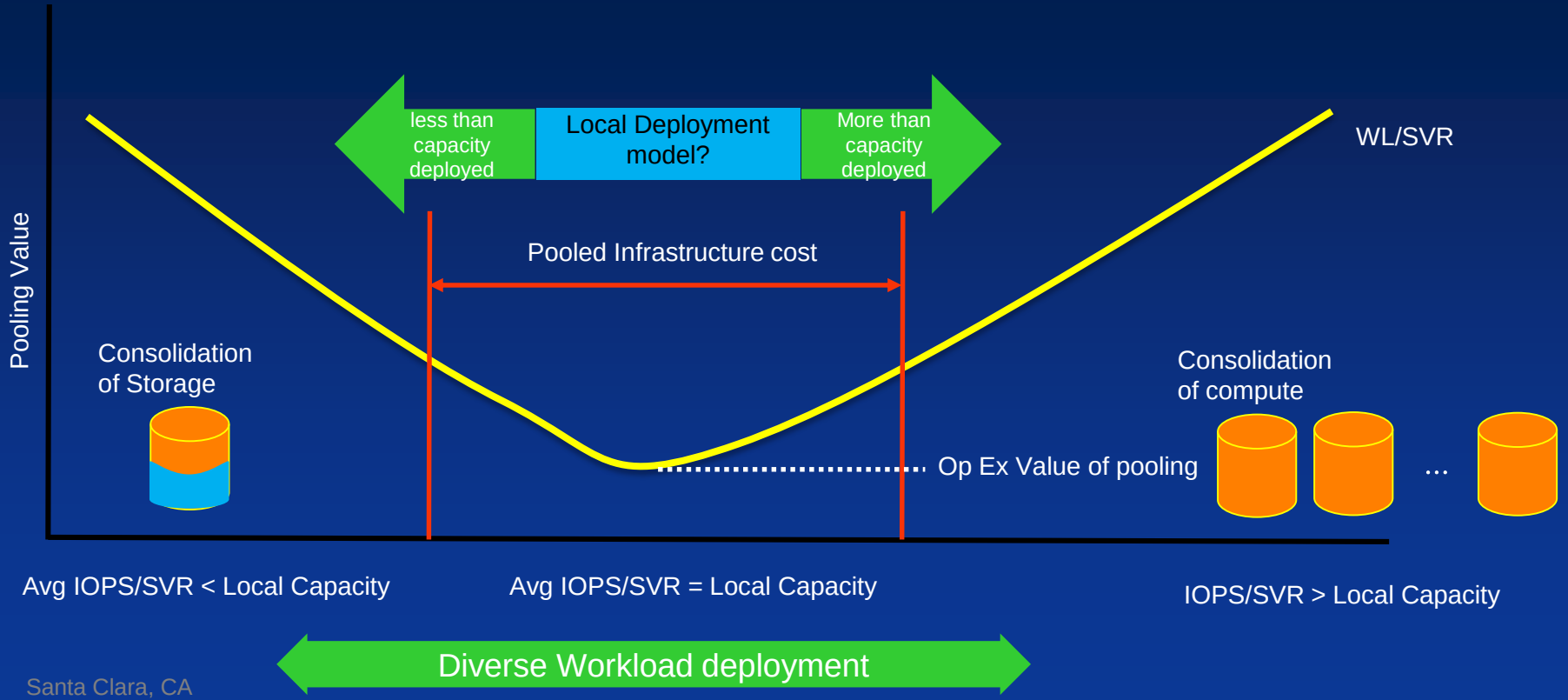


RSD Storage

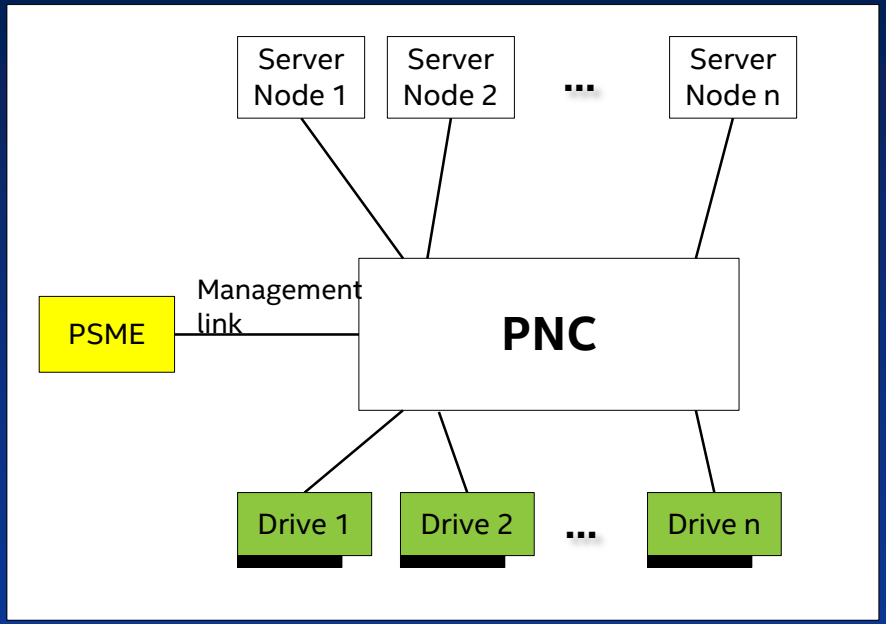


- Storage node is a collection of storage drives and storage controller
- Storage nodes are identified and exposed to Orchestration layer
- Comprehends storage pooling
- Today SAS/SATA drives are widely used
- Storage node based on Ethernet fabric

Local vs. Pooled Storage

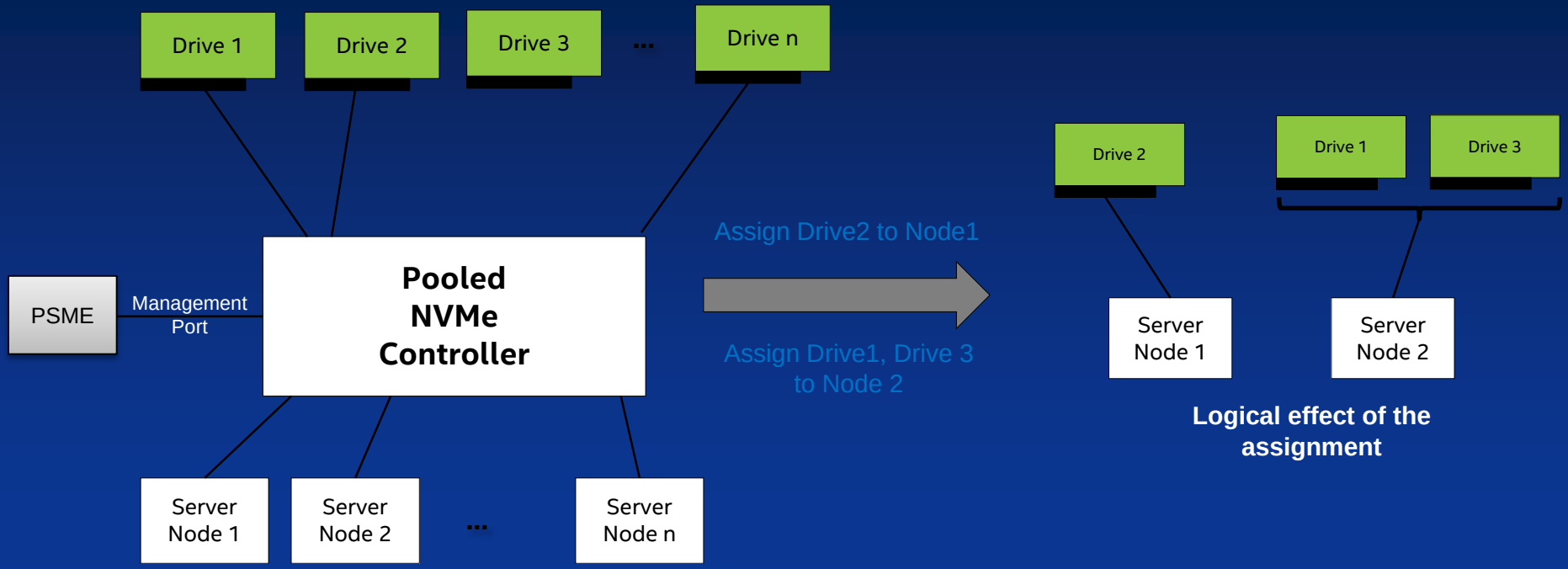


RSD Pooled NVMe Controller (PNC)

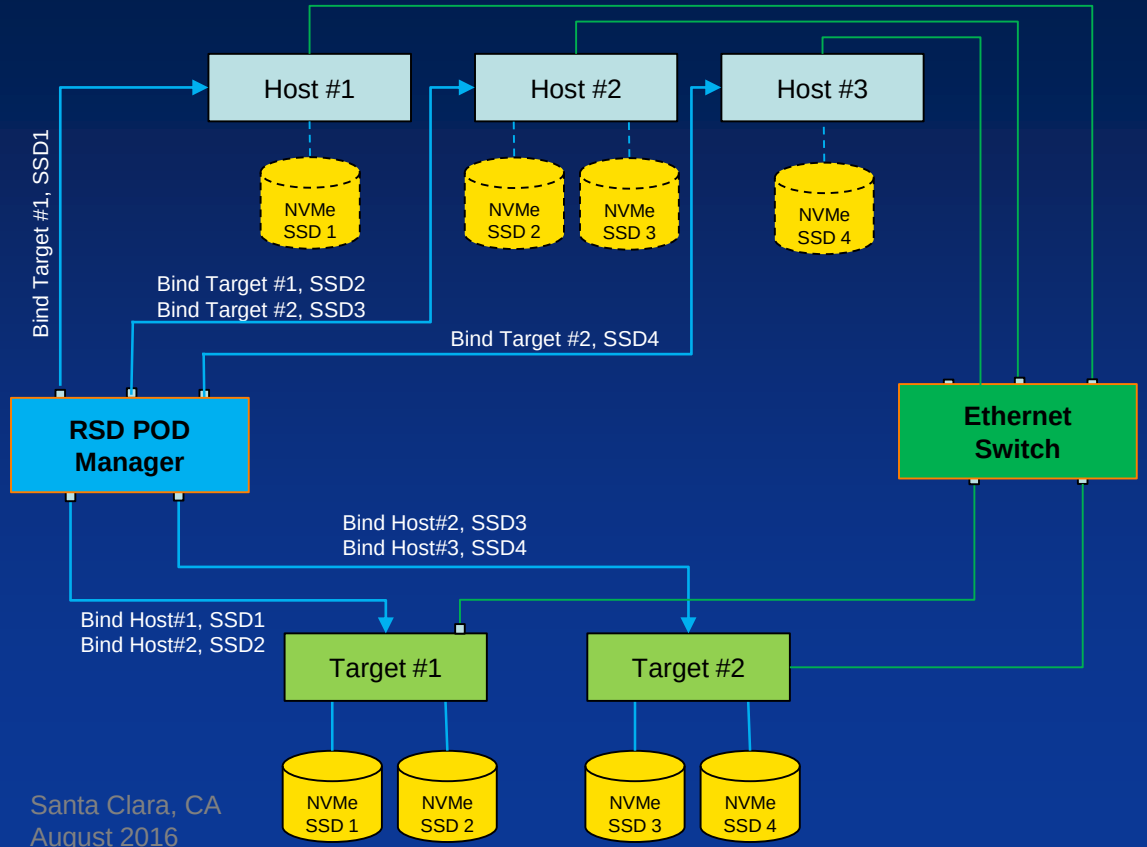


- Enable pooling of NVMe devices
- Assign high performance storage to nodes based on workload demand
- Prevent SPOF through host failover
- Enables ease of workload migration in hyperscale cloud environment
- Enables better utilization of DC resources by allowing composable high performance IO capacity

RSD Pooled NVMe Controller (PNC) (cont…)

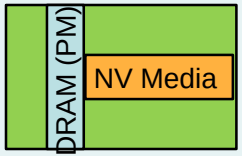
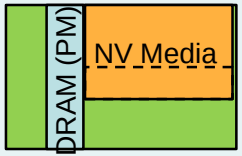


NVMe Over Ethernet

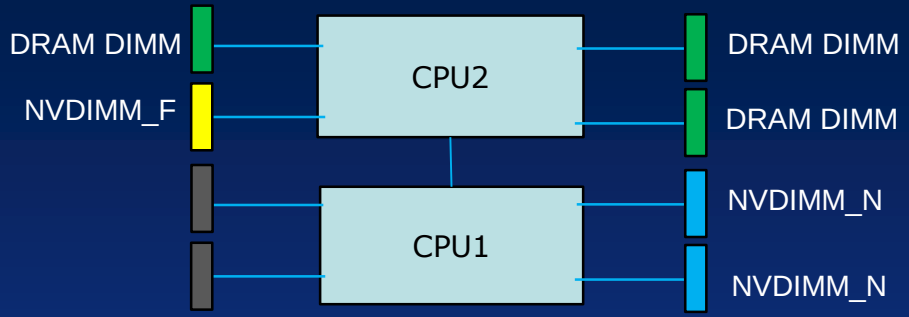


- Configures the NVMe Over Ethernet Targets and binds the host and targets
  - Network access path established
  - Storage NameSpace assigned
  - Security policies established
- Assign QOS to the NVMe over Ethernet traffic
- Monitors drive health

NVDIMM Types

NVDIMM-N	NVDIMM-F	NVDIMM-P
		
• Only DRAM is addressable by SW • NV Media acts as backup for DRAM • NV Media not addressable • At least 1:1 Capacity Ratio between DRAM & NV Media • Tracks DRAM latency & memory channel BW for Read and Write	• No DRAM • NV Media is directly addressable via Window mechanism • Tracks NV Media latency • Benefits from memory channel bandwidth	• Combination of NVDIMM-N and NVDIMM-F • Flash memory beyond that needed for persistence is accessible as block

RSD Memory - Example

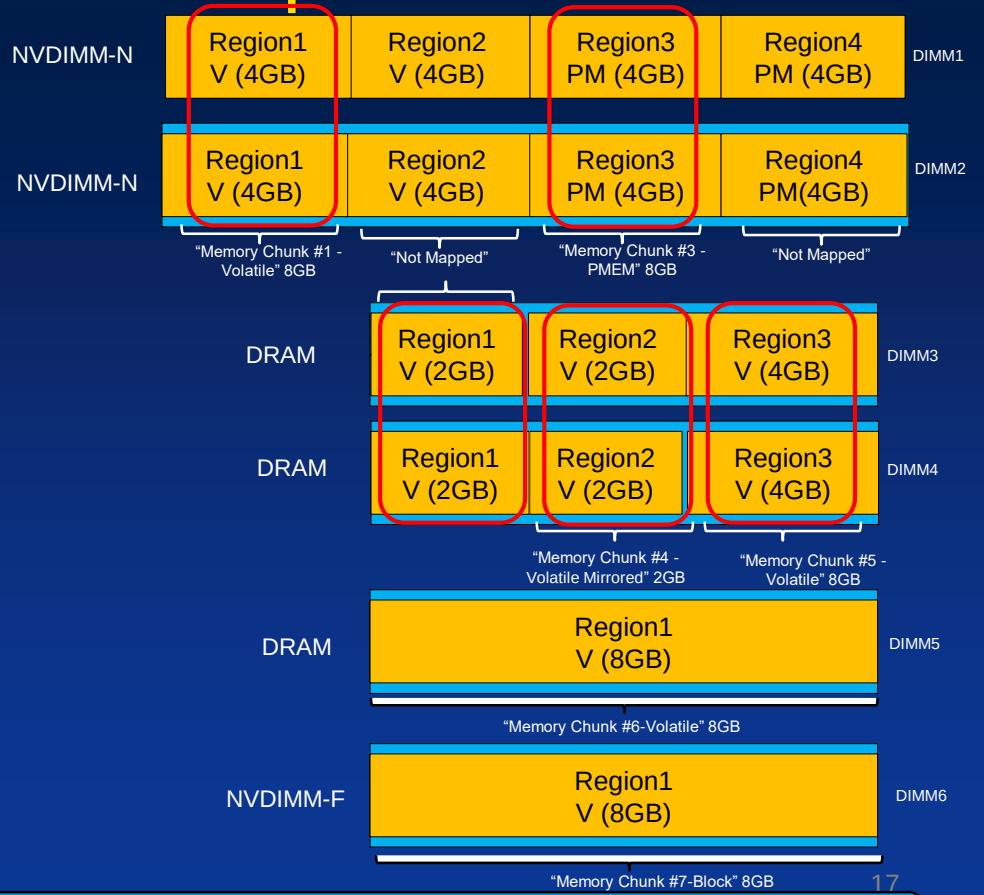


DIMM Type and Sizes

- **Two NVDIMM-N 16GB**
  - DIMM1
  - DIMM2
- **Three DDR4 DIMMs 8GB**
  - DIMM3
  - DIMM4
  - DIMM5
- **One NVDIMM-F 8GB**
  - DIMM6

Memory Chunks

- **#1 – 8GB Volatile**
- **#2 – Not present**
- **#3 – 8GB PMEM**
- **#4 – 2GB Mirrored Volatile**
- **#5 – 8GB Volatile**
- **#6 – 8GB Volatile**
- **#7 – 8GB Block**

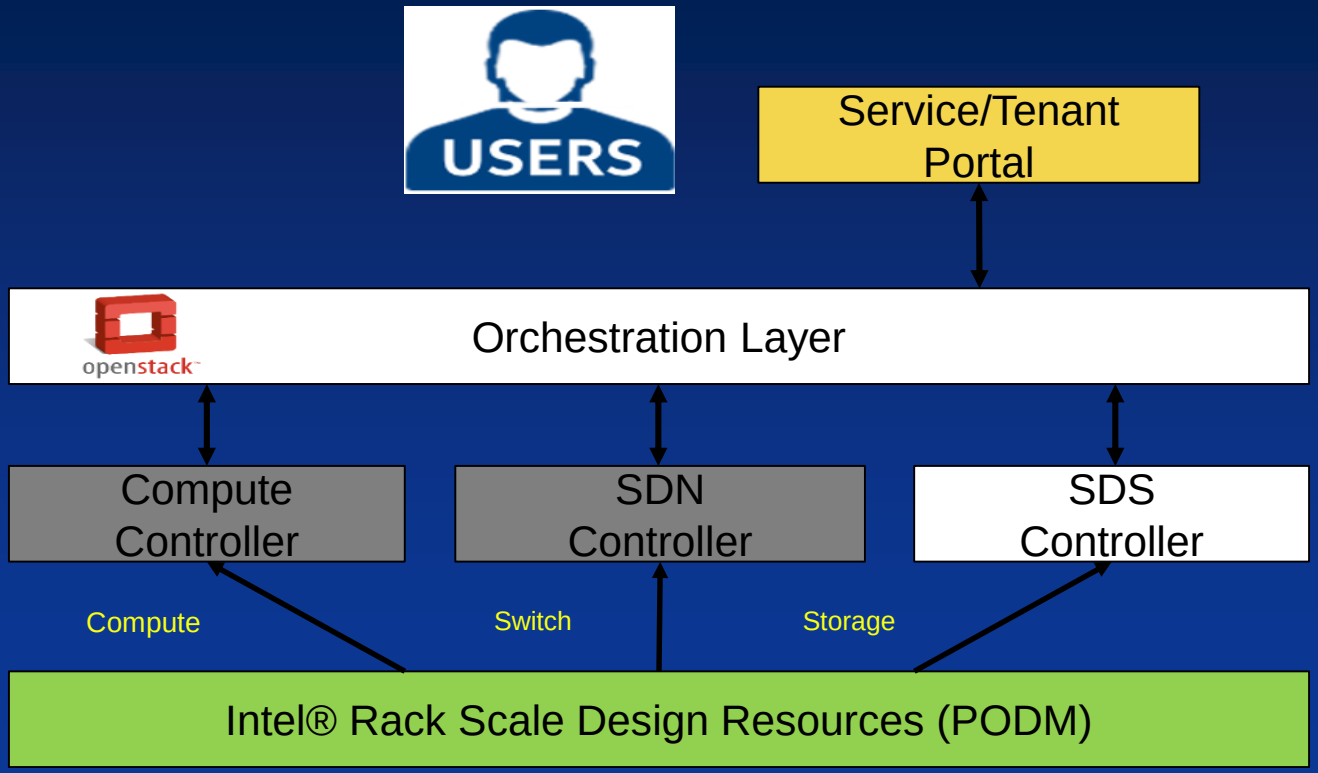


RSD Comprehends Memory Including NVDIMMs

Agenda

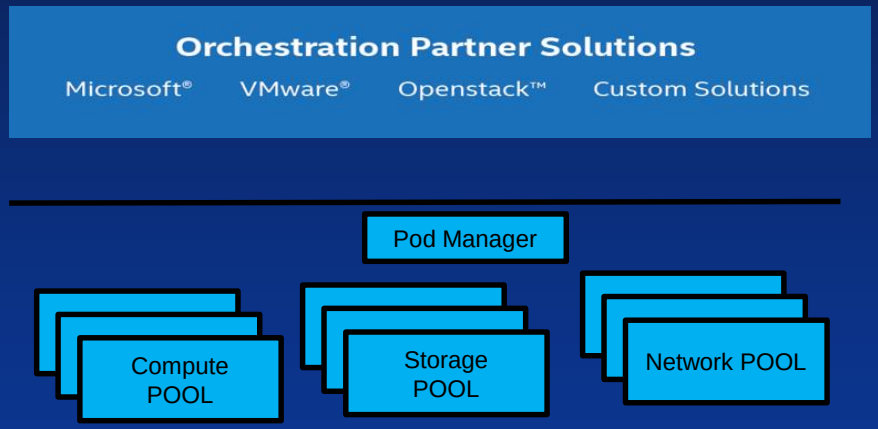
- Introduction to Intel® Rack Scale Design
- Storage in Intel® Rack Scale Design (Today and Future)
- Intel® Rack Scale Design Storage Orchestration
- Summary

Intel® Rack Scale Design Orchestration



Storage Orchestration

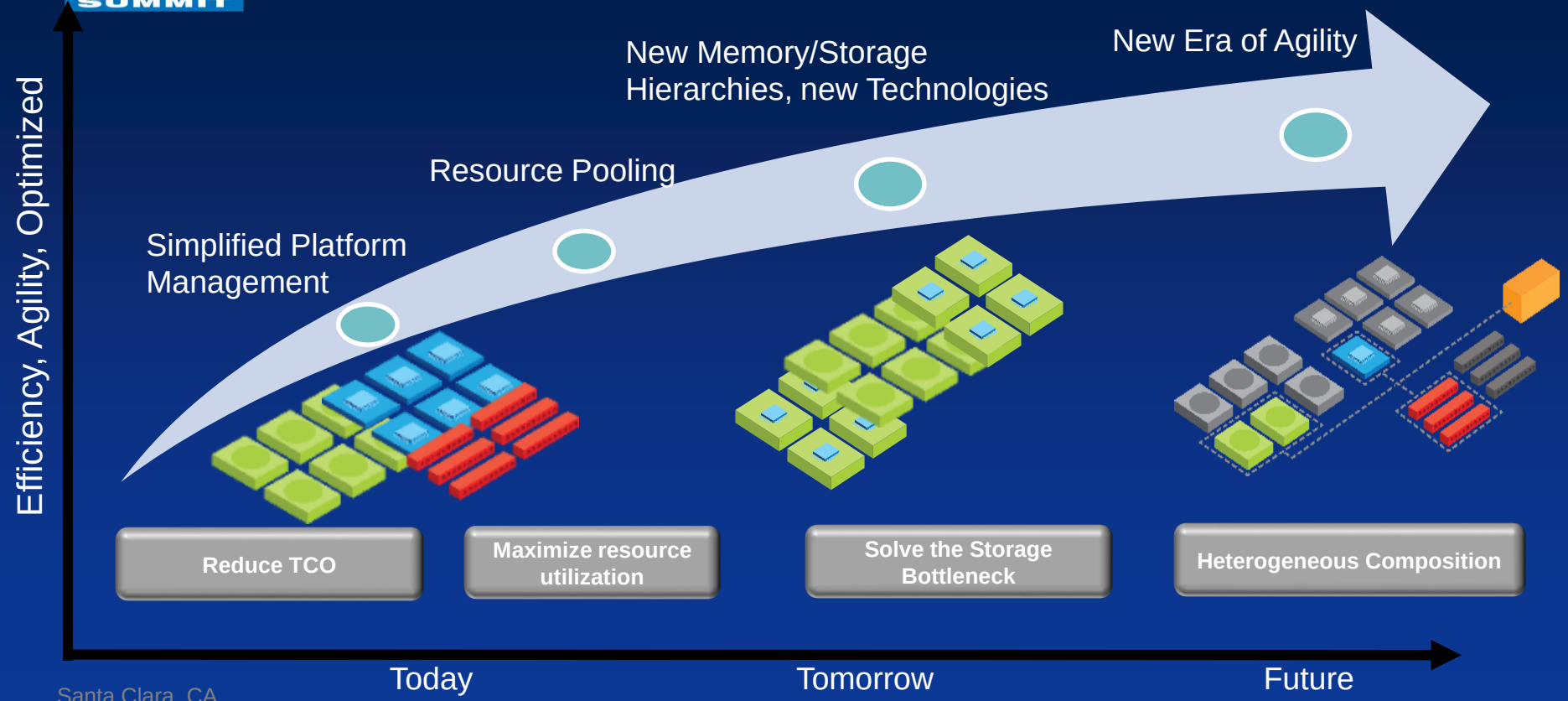
- Rack Scale Storage related elements are exposed to the Pod Manager through Intel® Rack Scale Design API
  - RSD API exposes storage availability zones to Orchestration layer
- Orchestration participates in configuring and managing the pooled storage configuration



Agenda

- Introduction to Intel® Rack Scale Design
- Storage in Intel® Rack Scale Design (Today and Future)
- Intel® Rack Scale Design Storage Orchestration
- Summary

Evaluation of Rack Scale Design



Summary

- Intel® Rack Scale Design delivers the Next Generation Data Center Architecture and comprehends advancements in storage

- Intel® Rack Scale Design references available online – http://www.intel.com/content/www/us/en/architecture-and-technology/rack-scale-design-overview.html

- Intel® Rack Scale Design overview, whitepaper and specifications http://www.intel.com/content/www/us/en/architecture-and-technology/rack-scale-design-overview.html