
OPEN
Compute Project
OCP U.S. SUMMIT 2017
Santa Clara, CA

Intel® Rack Scale Design:

A Deeper Perspective on Software Manageability for the Open Compute Project Community

Mohan J. Kumar

Intel Fellow

OPEN HARDWARE. OPEN SOFTWARE. OPEN FUTURE.

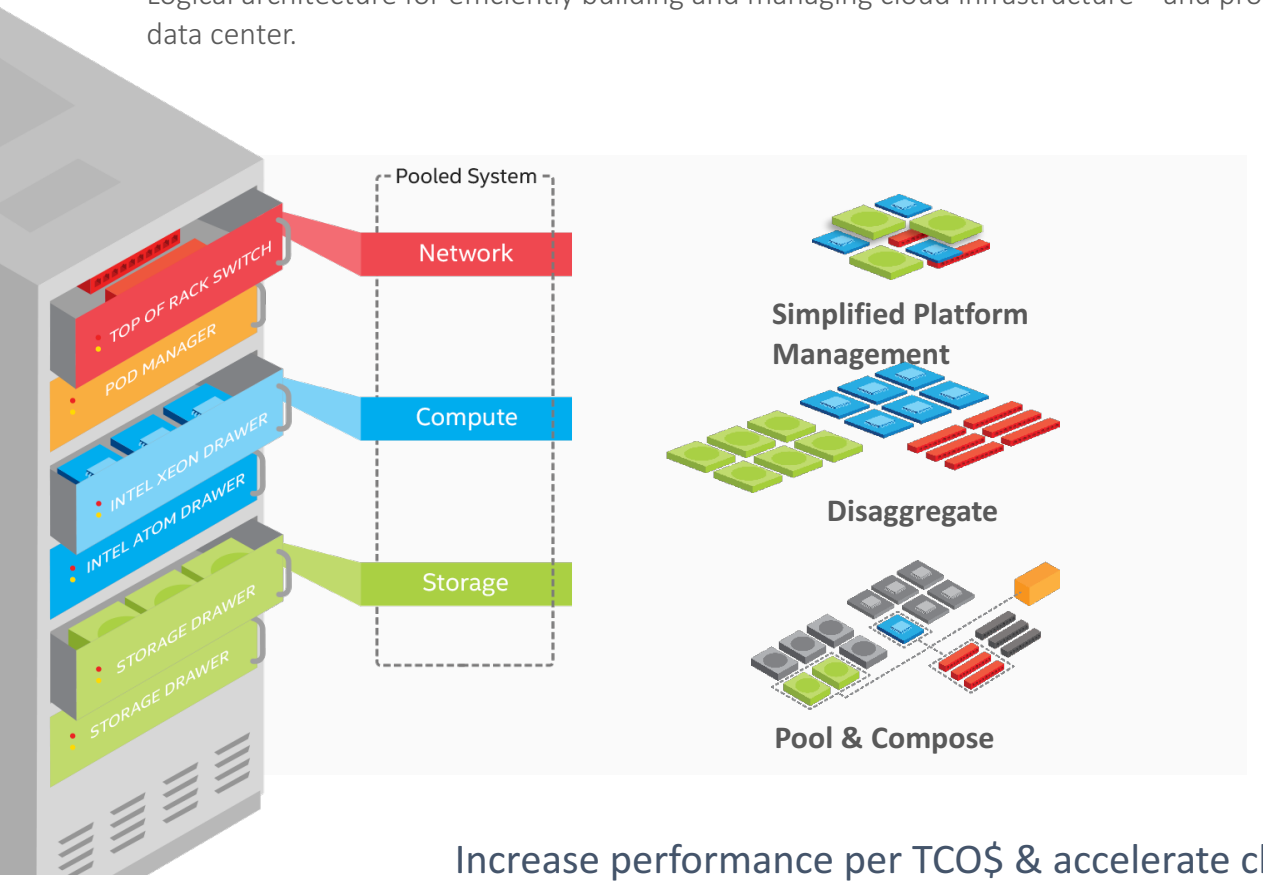
OPEN Compute Project

Agenda

- Rack Scale Design (RSD) Overview
- Manageability for RSD
- Rack Scale Design Future Considerations
- Summary

Intel® Rack Scale Design

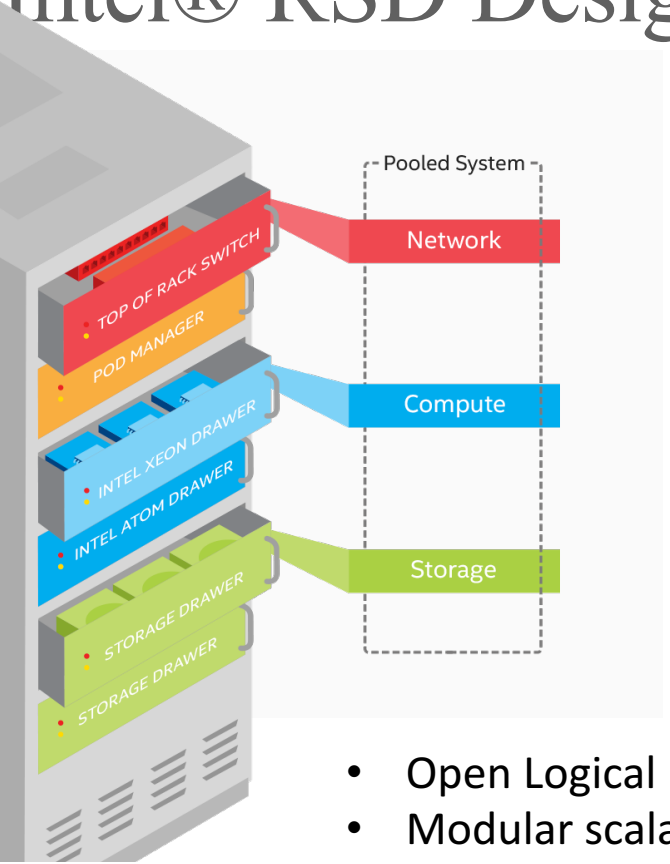
Logical architecture for efficiently building and managing cloud infrastructure—and providing the simplest path to a software defined data center.

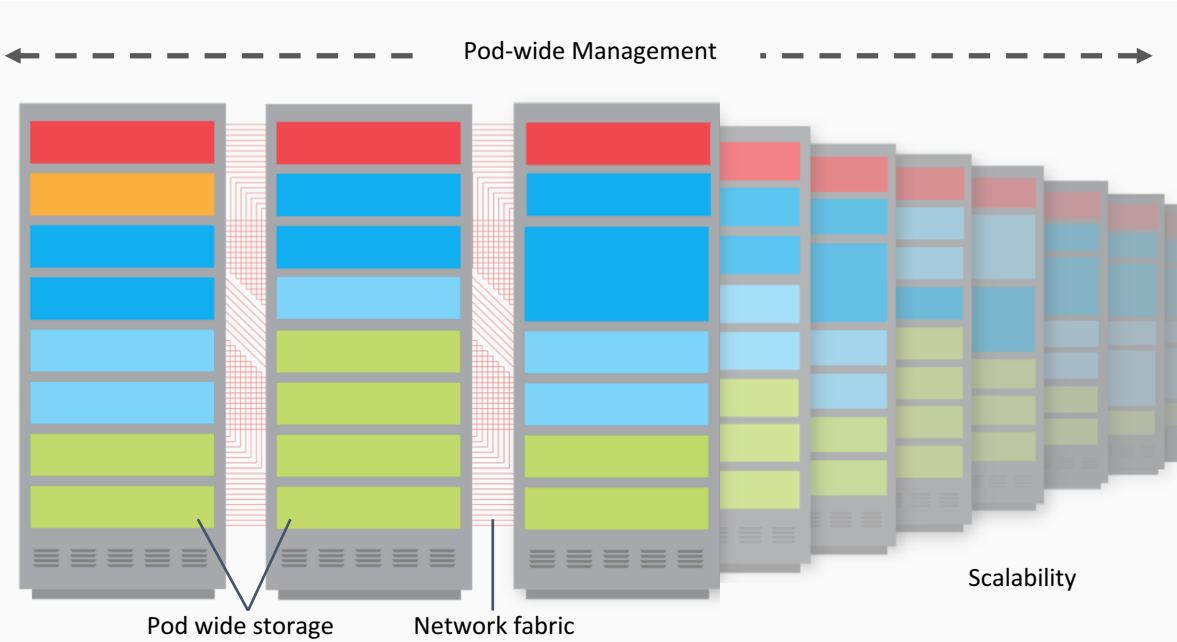


- ✓ **User-Defined Performance**
- ✓ **Maximum Utilization**
- ✓ **Interoperable Solutions**

Increase performance per TCO$ & accelerate cloud adoption

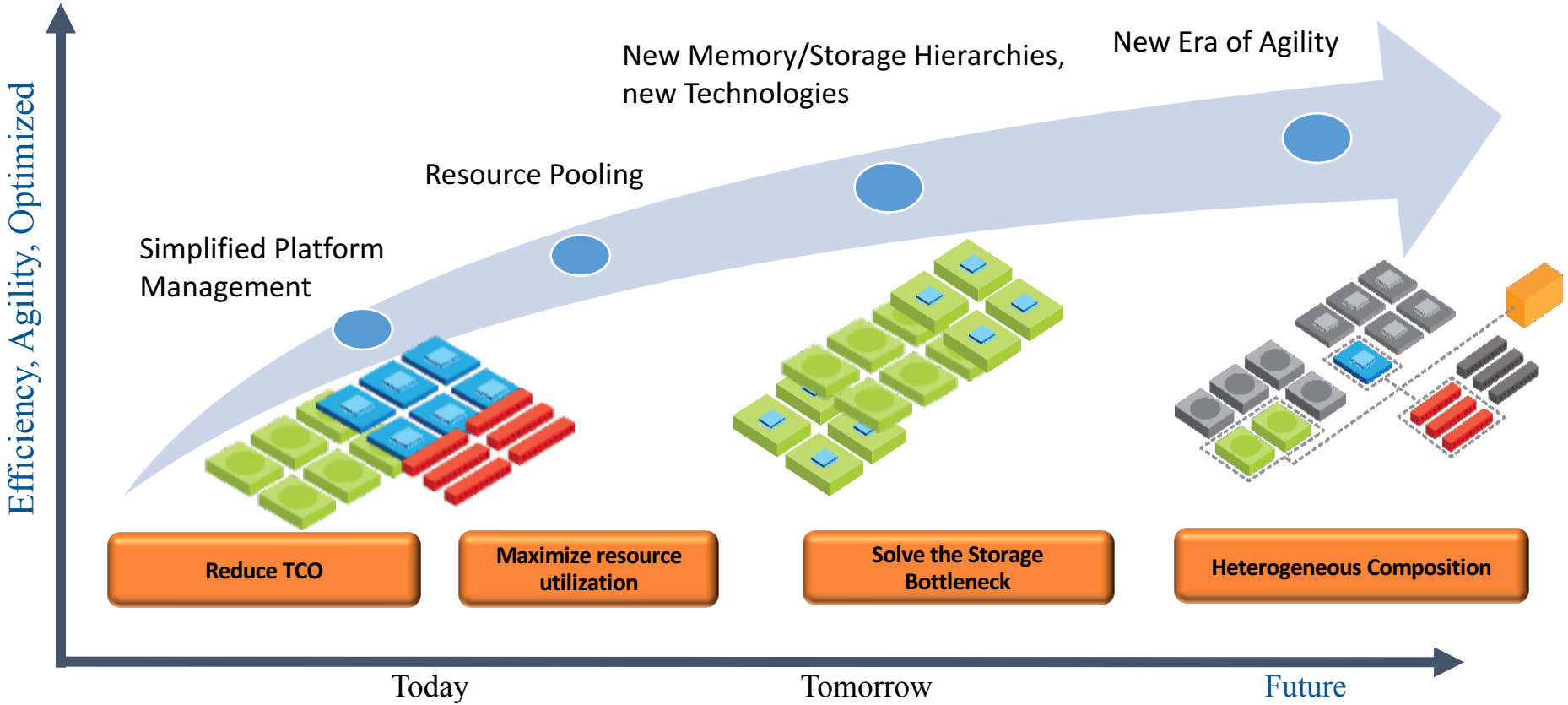
Intel® RSD Design Framework





- Open Logical Reference Architecture
- Modular scalable management architecture
- Comprehends Hardware, Firmware and Management Software
- Compliance and Interop program in the works

Evolution of Rack Scale Design



Manageability for RSD

OPEN HARDWARE. OPEN SOFTWARE. OPEN FUTURE.

OPEN Compute Project

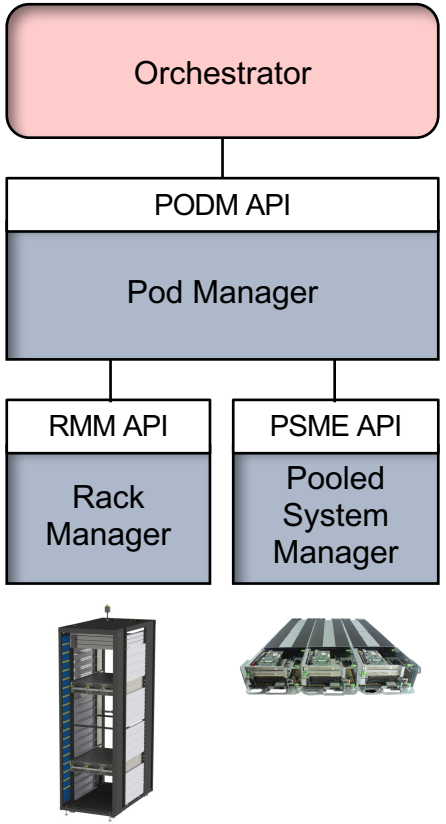
RSD Management Software Framework

- Asset & location discovery
- Disaggregated resource management
- Composable system support
- Support compute, network, and storage
- Built using DMTF† Redfish™

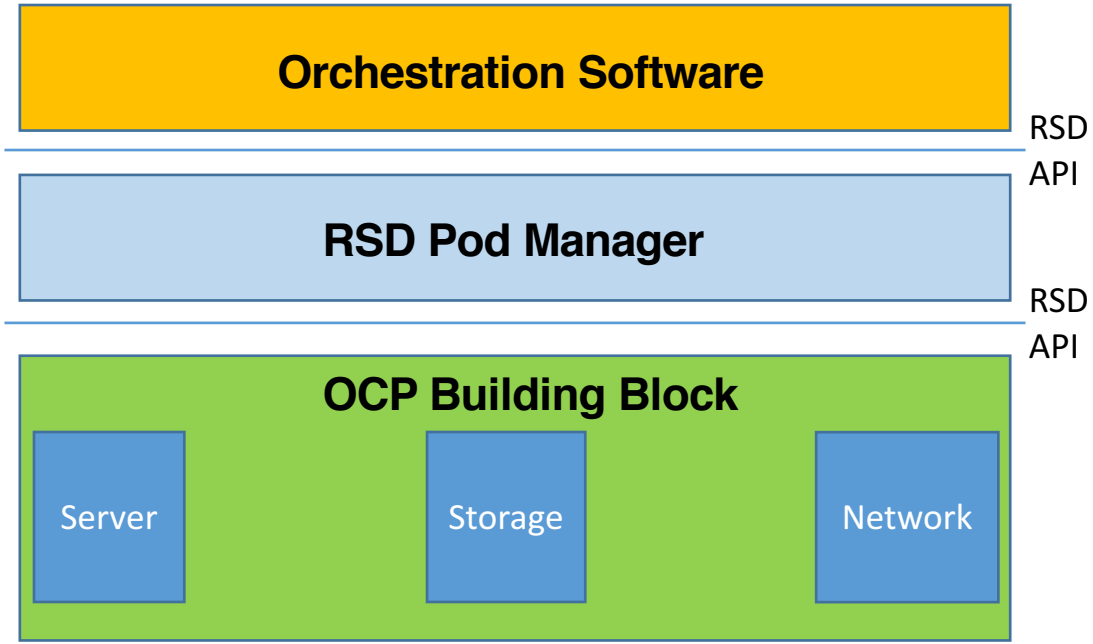
Comprehensive management architecture

RSD Manageability Standards

- Intel® Rack Scale Design manageability interfaces are based on Redfish™
  - Pod Manager (PODM) API
  - Rack Manager (RMM) API
  - Pooled System Manager (PSME) API
- Redfish™ has two parts
  - Interface specification (HTTP, JSON, OData)
  - Resource models for manageability
- Manageability Models
  - DMTF – physical platform, compute
  - SNIA – networked storage and Storage Service (Swordfish)
  - IETF – network devices and services (YANG-to-Redfish)



Intel® RSD and OCP



- Rack Scale and OCP directions are well aligned
- OCP hardware can be managed using Rack Scale Management API
- OCP hardware with RSD support available from RSD partners

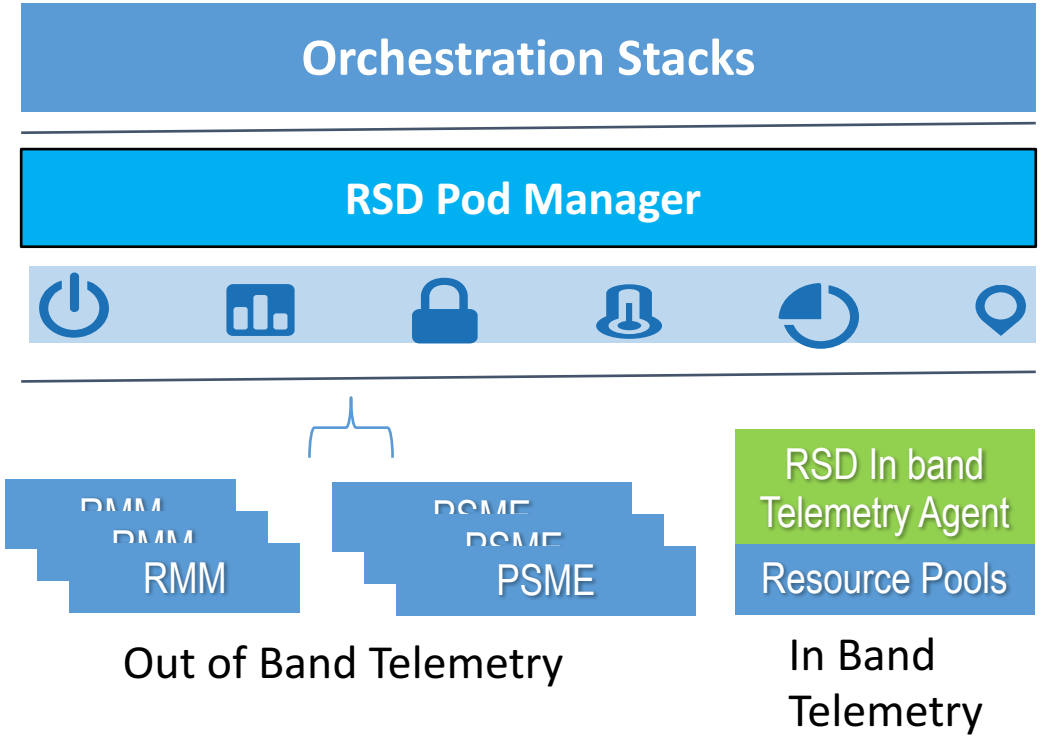
RSD Future Considerations

OPEN HARDWARE. OPEN SOFTWARE. OPEN FUTURE.

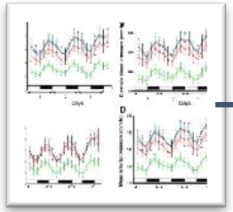
OPEN Compute Project

Intel® RSD Telemetry Overview

- Support Inband and Out-of-band (OOB) telemetry
- Built using Redfish extensions for Telemetry
- Feed relevant RSD telemetry to analytics at Orchestration layer
- Comprehends power, performance, security, thermal, utilization and location
- Hierarchical metric collection and representation in events and APIs



Intel® RSD Telemetry Flow



RSD Inband Telemetry

RSD Out-of-band Telemetry

RSD Pod Manager

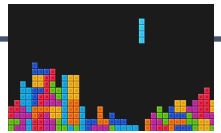
REPORTING

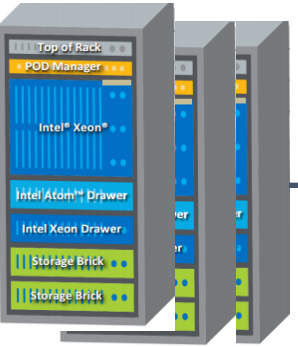
Analytics

Orchestrator Workload Placement and Resource allocation

CONTROL

RSD Pod Manager re-provisions HW and system resources



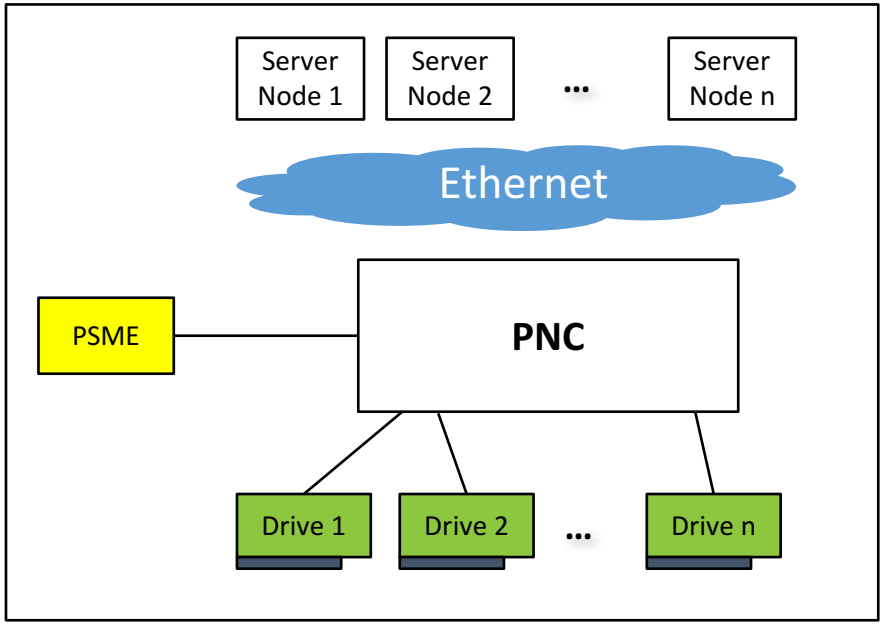


RSD Pod

openstack

vmware

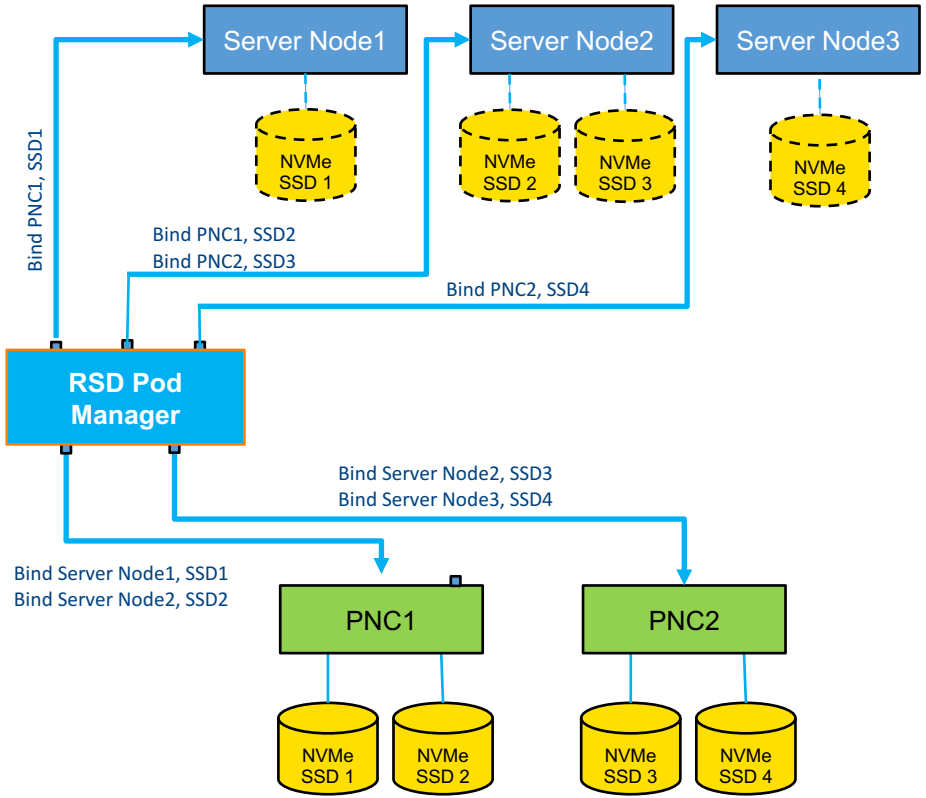
Intel® RSD Pooled NVM Express* Controller



PSME – Pooled System Management Engine
PNC – Pooled NVM Express Controller

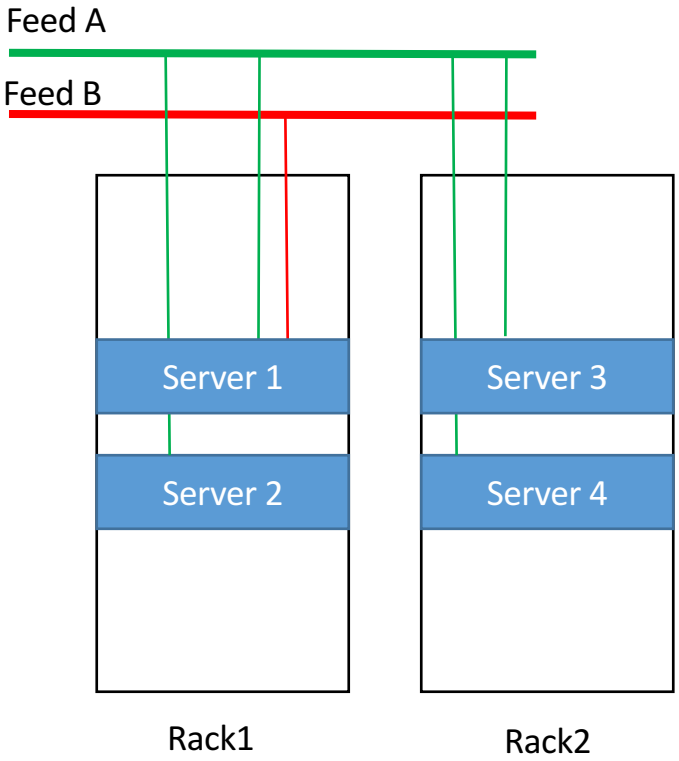
- Builds on the momentum of NVM Express over Fabric* (NVMe-oF*)
- Enable disaggregation of PCIe NVMe* devices
- Support higher radix pooling
- Assign storage to Compute or Storage nodes based on workload demand
- Intel® RSD complements NVMe-oF* by
  - Supporting discovery and management of manage NVMe-oF* devices.
  - Enables add/removal of NVMe* devices/subsystems to a host
  - Improved security by separation of in-band and management networks
  - Allows QoS policies across composed NVMe* devices

Intel® Rack Scale Pooled NVM Express* Controller - Example

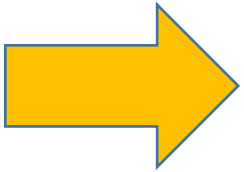


- Intel® Rack Scale Design pooled systems management engine configures NVMe over Ethernet Targets
- Intel® Rack Scale Design Pod manager composes the required platform by binding NVMe targets to hosts
- Pod manager handles QOS to the NVMe over Ethernet traffic
- Pod manger monitors drive health

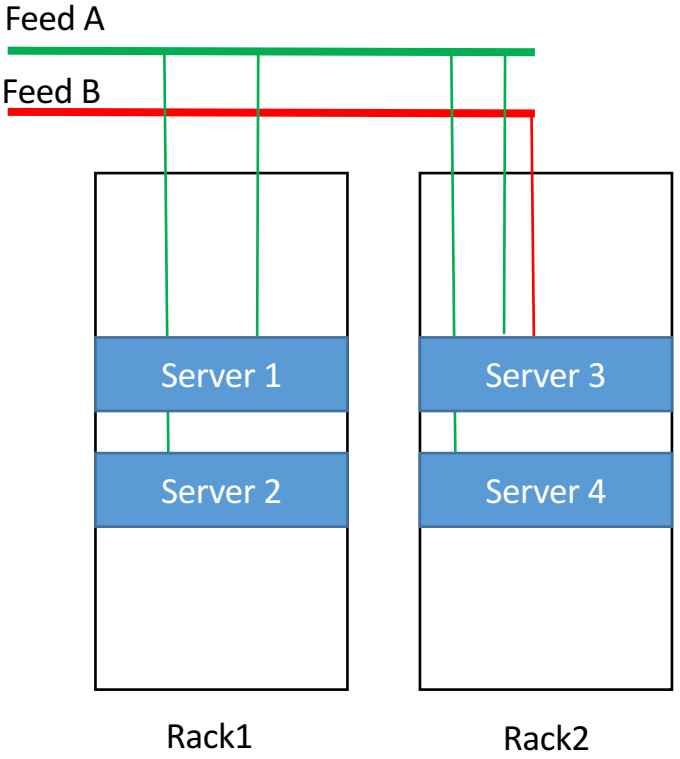
Composable Power using RSD



Server 1 – Dual Feed
Server 2-4 – Single Feed



RSD Composition



Server 1, 2, 4 – Single Feed
Server 3 – Dual Feed

Summary

OPEN HARDWARE. OPEN SOFTWARE. OPEN FUTURE.

OPEN Compute Project

Summary

- Rack Scale and OCP directions are well aligned
- OCP hardware can be managed using Rack Scale Management API
- Rack Scale Firmware / Software reference available online - http://www.intel.com/IntelRSD

Visit the Intel booth (A5) to learn more and see the Intel® Rack Scale Design demo.

Intel® Rack Scale Design Aligned Ecosystem

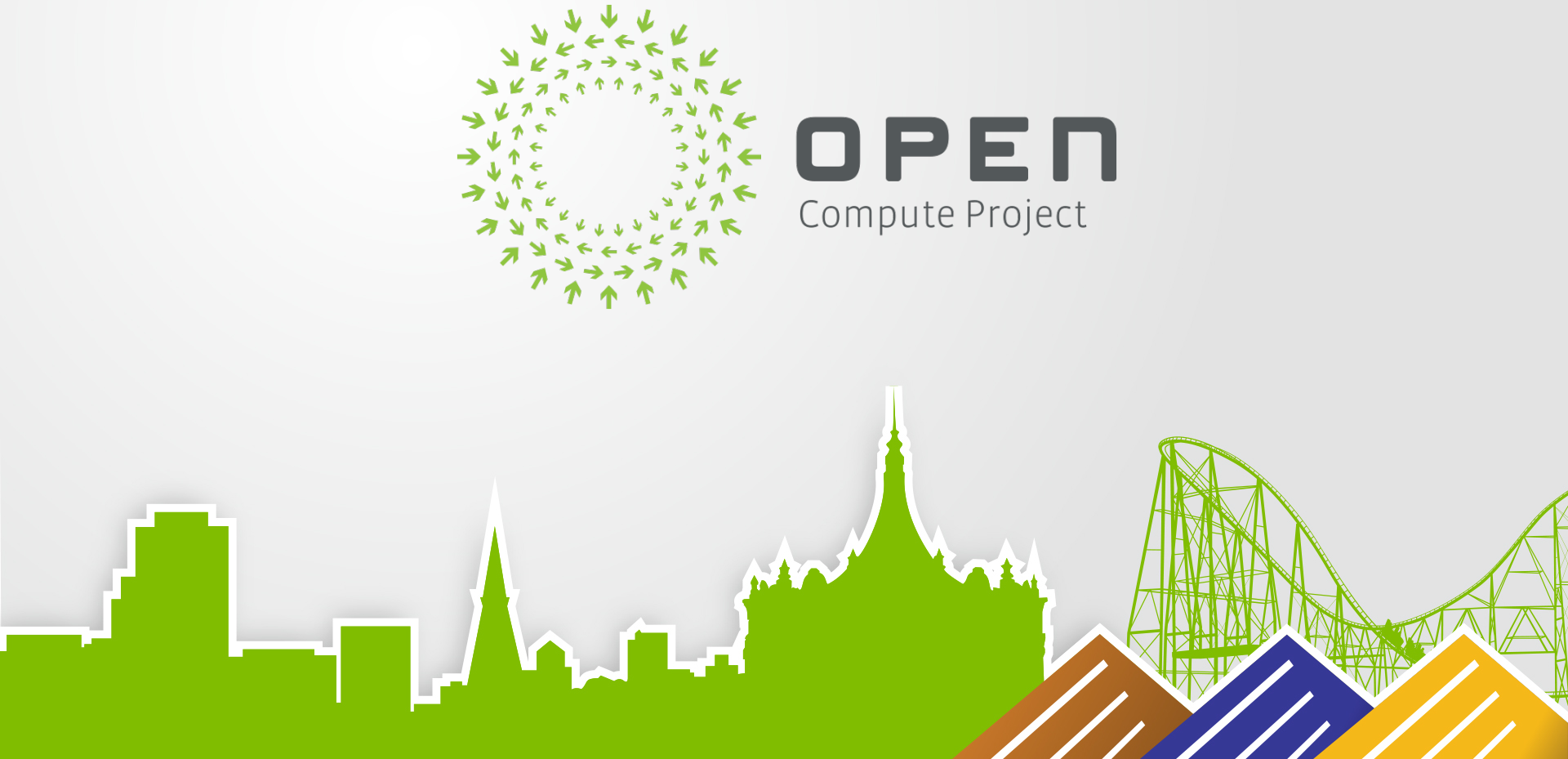
Common management framework and telemetry supports ecosystem partner requirements to develop a range of platforms and solutions

OEMs/ ODMs/TEMs*	Hewlett Packard Enterprise, inspur 浪潮, 英業達集團 Inventec, Microsemi, QCT, Supermicro, DELL EMC, Ericsson, Huawei, Lenovo, Orchestrating a brighter world NEC, radisys, wiwynn
ISVs/OSVs*	American Megatrends, Canonical, vmware, 九州云 Cloud
Industry initiatives/ Standards*	openstack, Open Compute Project, DMTF Redfish, 开放数据中心委员会 Open Data Center Committee
End Users/POCs*^	Baidu 百度, SK telecom, SoftBank, Tencent 腾讯, YAHOO!

*Other names and brands may be claimed as property of others. ^Contact your local Intel representative for POC information.

Legal Notices and Disclaimers

- Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com.
- Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.
- Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.
- Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.
- This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.
- No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.
- Statements in this document that refer to Intel's plans and expectations for the quarter, the year, and the future, are forward-looking statements that involve a number of risks and uncertainties. A detailed discussion of the factors that could affect Intel's results and plans is included in Intel's SEC filings, including the annual report on Form 10-K.
- All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice. The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.
- Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.
-
- *Other names and brands may be claimed as the property of others.


OPEN
Compute Project