Swordfish implementation in Rack Scale Design NVMe over Fabrics

Mariusz Krzywienski
Rafal Sztejna
Intel
Agenda

- Code availability
- Rack Scale Design overview
- NVMe over Fabrics using Swordfish/Redfish
- RSD Proposed enhancements
Code availability

- All discussed code available on GitHub: https://github.com/intel/intelRSD
- It will be presented on SDC Workshop
RSD Overview
Data Center Agility, Built on Open Standards

Today’s data center Challenges

Current Infrastructure

• Fixed ratio of compute, storage, and accelerator resources
• Expensive refresh & scale out
• Outdated software interface
• Cumbersome hardware provisioning process

“an industry-aligned architecture for composable, disaggregated infrastructure built on modern, open standards.”

Disaggregated

Composable

Interoperable

Decrease Costs

Increase Agilit

50% Efficiency IT Operations

45% People Hours Per Update

3 People Per Rack

5% Utilization Equipment

4 Equipment

Intel® Rack Scale Design

2. Source: Disaggregated Server Architecture Drives Data Center Efficiency and Innovation, Shesha Krishnapura, Intel Fellow and Intel IT CTO, 2017
Intel® RSD Key Attributes

Disaggregated

Composable

Interoperable

Intel® Rack Scale Design

Spend less up front and save $$ over time

Compose hardware resources “on the fly”

Choose the best now without vendor lock-in

Intel Pod Manager

Composed Node 1

Composed Node 2

App 1

App 2

App 3

Orchestration

Single-Pane-of-Glass Management

Vendor A Hardware

Vendor B Hardware

Vendor C Hardware

Vendor D Hardware

Compose hardware resources “on the fly”
Intel Rack Scale SW Stack

Cloud or Data Center Manager
- VMware
- OpenStack
- 3rd Party Datacenter Mgr
- DIY

Rack Scale POD Management API
- POD Wide Asset Management
- POD Wide Composed Nodes Management
- POD Wide Fabric Management
- POD Wide Storage Management

Rack Scale POD Manager
- POD Manager Functionality
- POD Manager Functionality

Rack Scale Pooled Systems API
- Pooled Systems
- Compute Module
- Network Module
- FPGA Module
- Memory Module
- Accelerator Module
- Fabric Module
- Storage Module

Network Services API
- Network Services
- Network Infrastructure
- Network Protocols
- Networking Services

Storage Services API
- Storage Services
- Storage Pools
- Logical Volumes
- Storage Endpoints

Chassis Management API
- Chassis Management
- Power
- Thermal

Rack Manager

2018 Storage Developer Conference. © Intel Corp. All Rights Reserved.
Intel Rack Scale Storage Services

Core Management Application

- Asset Manager
- CMDB

Redfish/Swordfish API

Storage Management

- Targets
- Logical Devices
- Physical Devices

Generic Assets Management Interface (JSON-RPC)

- NVMe over Fabrics
- iSCSI
- LVM Native Linux
- CEPH
- Native Linux Disk Mgmt

Storage Assets (Physical and Logical)
NVMe over Fabrics using Swordfish
RSD components for NVMe-oF
Swordfish resources in RSD

- StorageService
- DriveCollection
- EndpointCollection
- StoragePool (+Collection)
- Volume (+Collection)
- Capacity
- StorageReplicaInfo
NVMe over Fabrics using Swordfish

Create Endpoint

Volume

"ConnectedEntities": [
  { "EntityLink": {
    "@odata.id": "/redfish/v1/StorageServices/1/Volumes/1" },
    "EntityRole": "Target" },
"Identifiers": [
  { "DurableName": "nqn.2014-08.org.nvmexpress:uuid:397f9b78-(...)",
    "DurableNameFormat": "NQN" },
"IPTransportDetails": [
  { "TransportProtocol": "RoCEv2",
    "IPv4Address": { "Address": "192.168.0.10" },
    "Port": 1023 },
"Links": {
  "Oem": {
    "Intel_RackScale": {
      "Zones": [ { "@odata.id": "/redfish/v1/Fabrics/NVMeoE/Zones/1" } ],
      "Interfaces": [ { "@odata.id": "/redfish/v1/Systems/Target/EthernetInterfaces/1" } ]
    }
  }
}]

* Target host only
* Discovery Service & Target Host
Create Endpoint

```
"ConnectedEntities": [ 
  { "EntityLink": null, 
    "EntityRole": "Initiator"  } ],
"Identifiers": [ 
  { "DurableName": "nqn.2014-08.org.nvmexpress:uuid:12345678-(...)",
    "DurableNameFormat": "NQN"  },
  "IPTransportDetails": [ ],
  "Links": {},
  "Oem": { 
    "Intel_RackScale": { 
      "Zones": [ {"@odata.id": "/redfish/v1/Fabrics/NVMeoE/Zones/1"} ],
      "Interfaces": [ ]
    }
  }
} }
```
NVMe over Fabrics using Swordfish

Create Zone

```
"Links": {
  "Endpoints": [
    {
      "@odata.id": "/redfish/v1/Fabrics/NVMeoE/Endpoints/1"
    },
    {
      "@odata.id": "/redfish/v1/Fabrics/NVMeoE/Endpoints/2"
    }
  ]
}``

Initiator Endpoint

Target Endpoint
RSD Enhancements
RSD Enhancements to Swordfish/Redfish

- **Endpoint**
  - **Authentication** – indicating what authentication is required by this endpoint
    - Currently CHAP model supported
    - Planned to extend for key-based authentication
  - **Zones** – array of all zones where this endpoint exist
  - **Interfaces** – array of available NICs
RSD Enhancements to Swordfish/Redfish

- Volume
  - Endpoints – array of links to Endpoints that connect to this volume
  - Bootable – flag
  - Metrics – volume metrics (telemetry)

- StorageService
  - ManagedBy – link to service manager
RSD Enhancements to Swordfish/Redfish

- Drive:
  - FwVersion – useful for inventory
  - UsedBy - link to StoragePool
- Metric – subresource containing metrics:
  - TemperatureKelvin, UnitsRead, UnitsWritten, HostRead(Write)Commands, PowerCycles
Can’t Miss: SNIA Swordfish™ Events

- Hands-on Workshop, Tuesday, 9/25
  - Mezzanine, 2:50 p.m. and 5:00 – 7:00 p.m.
  - Visit interactive workstations and see actual implementations
  - Work with open source development tools

- BoF, Monday, 9/24
  - Winchester, 7:00 p.m. – 9:00 p.m.
  - Discussions from adoption to integration
  - Have a beer and get your questions ready!
Questions?