



Flash Memory Summit



Scalable Management of Ethernet- Attached Drives and EBOFs with SNIA Swordfish

Richelle Ahlvers, Intel

About the Presenter



Richelle Ahlvers

Storage Technology
Enablement Architect, Intel

Richelle Ahlvers is a Storage Technology Enablement Architect at Intel, where she promotes and drives enablement of new technologies and standards strategies. Richelle has spent over 25 years in Enterprise R&D teams in a variety of technical roles, leading the architecture, design and development of storage array software, storage management software user experience projects including mobility, developing new storage industry categories including SAN management, storage grid and cloud, and storage technology portfolio solutions.

Richelle has been engaged with industry standards initiatives for many years and is actively engaged with many groups supporting manageability including SNIA, DMTF, NVMe, OFA and UCle. She is Vice-Chair of the SNIA Board of Directors, Chair of the Storage Management Initiative, leads the SSM Technical Work Group developing the Swordfish Scalable Storage Management API, and has also served as the SNIA Technical Council Chair and been engaged across a breadth of technologies ranging from storage management, to solid state storage, to cloud, to green storage. She also serves on the DMTF Board of Directors as the VP of Finance and Treasurer.

Abstract

- The enterprise storage market is rapidly expanding to include NVMe and NVMe-oF products pervasively. This provides a challenge: how do you manage these as part of your enterprise datacenter?
 - As the NVM Express family of specifications continue to develop, the corresponding Swordfish management capabilities are also evolving: the SNIA Swordfish specification has expanded to include full NVMe and NVMe-oF enablement and alignment across DMTF, NVMe, and SNIA for NVMe and NVMe-oF use cases.

In conjunction with Redfish management of servers, Swordfish's capabilities to manage Ethernet-attached drives and EBOFs in the enterprise provide a seamless management ecosystem.

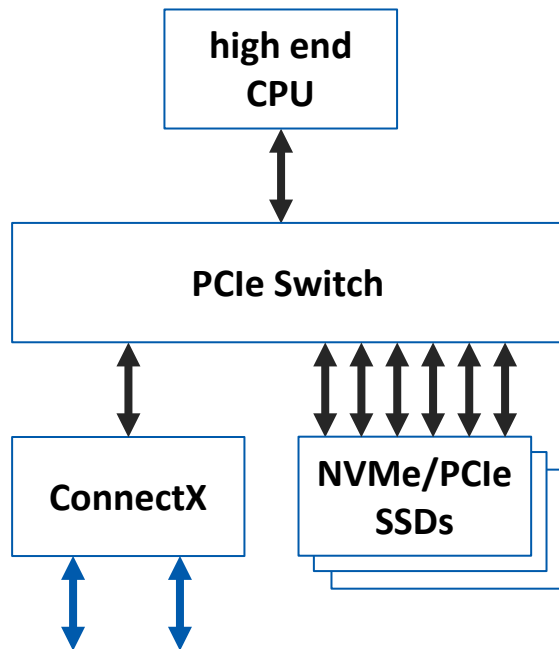
- This presentation will provide an overview of the capabilities available in SNIA Swordfish® to manage Ethernet-attached drives and EBOFs .

Overview

- Introduction to Ethernet attached drives
 - What are Ethernet attached drives?
 - What is driving emergence?
 - Use cases
- Standards-based management for Ethernet-attached drives (and EBOFs)
 - Layering Redfish and Swordfish with NVMe / NVMe-MI
 - Overview of standard NVMe and storage management models
 - Ethernet-attached drive and EBOF examples
 - Available materials
- SNIA Native NVMe-oF™ Drive Specification

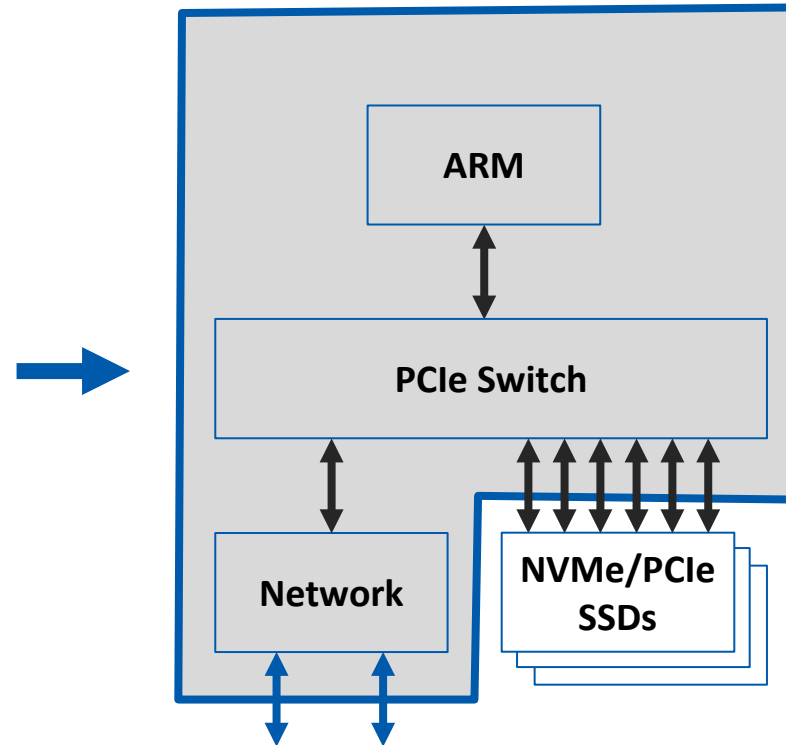
NVMe JBOF architecture options

Current JBOF Architecture



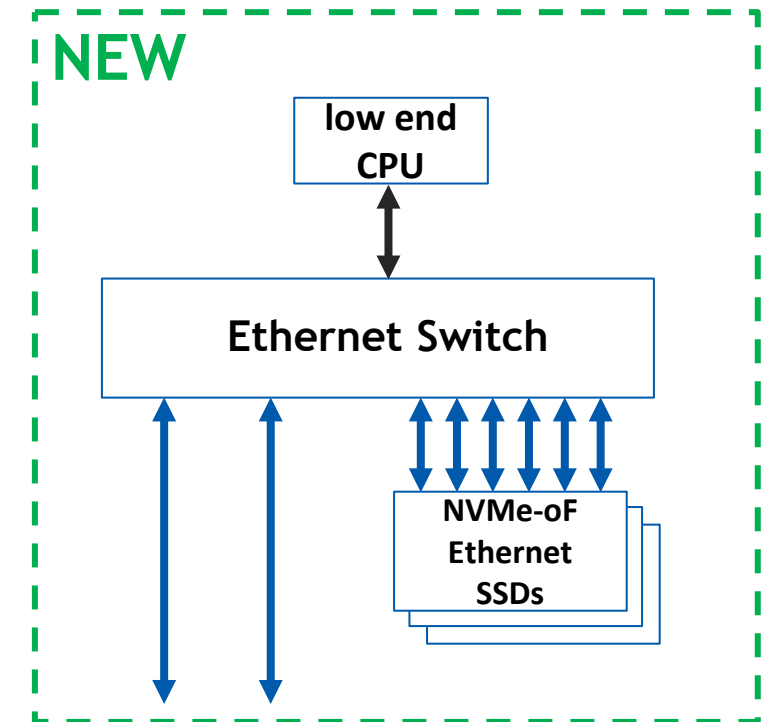
Shipping Since 2017

DPU Based JBOF



Shipping 2020

EBOF Architecture



Sampling now

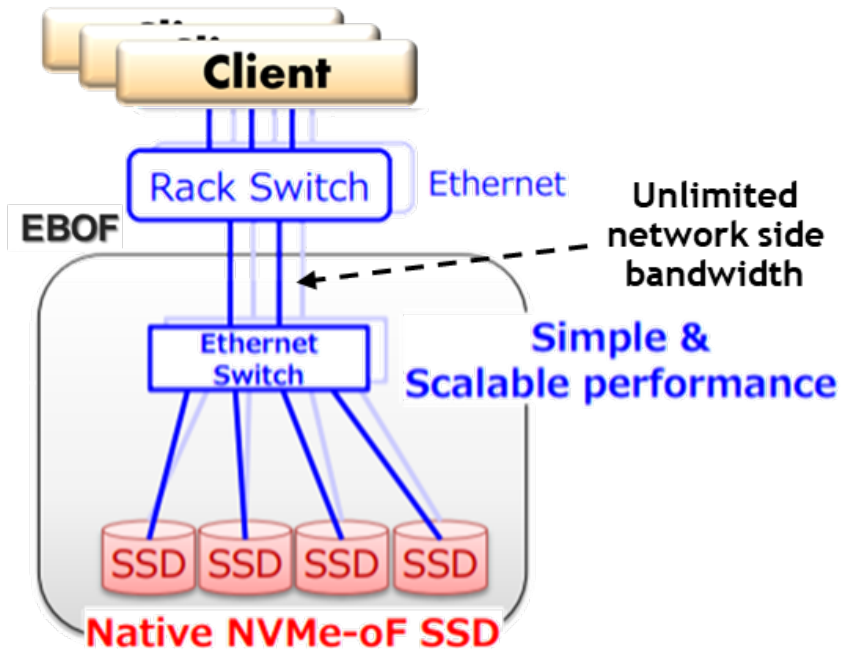
Why ESSDs?



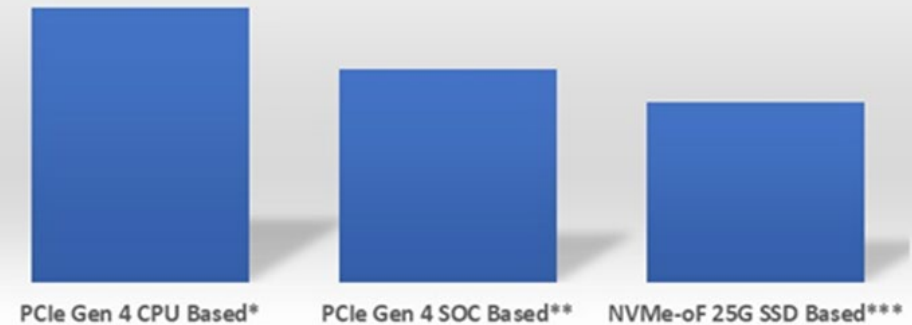
Flash Memory Summit

Two reasons this new architecture is attractive

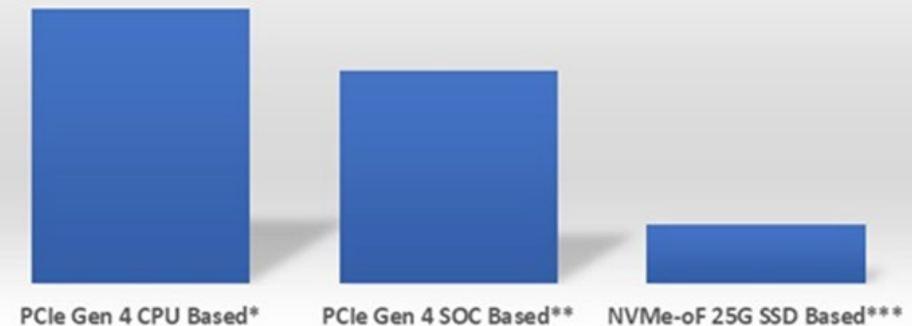
- 1) Performance
- 2) Cost



JBOF price comparison
(Excluding SSD cost)
NetApp 2019 FMS



JBOF Price per Gbit of performance

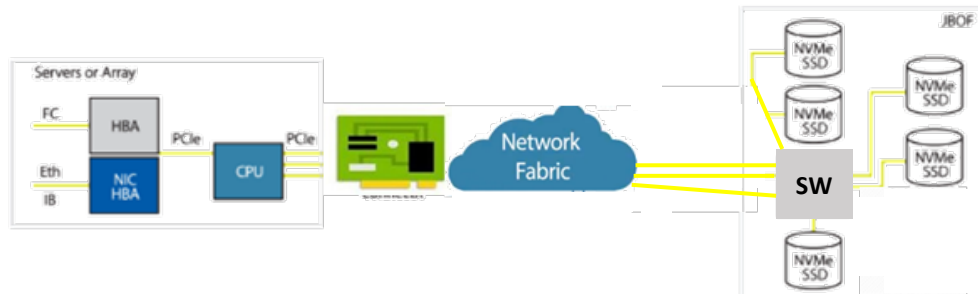


* Supports one 2x200G RNIC connected with x16 PCIe Gen4

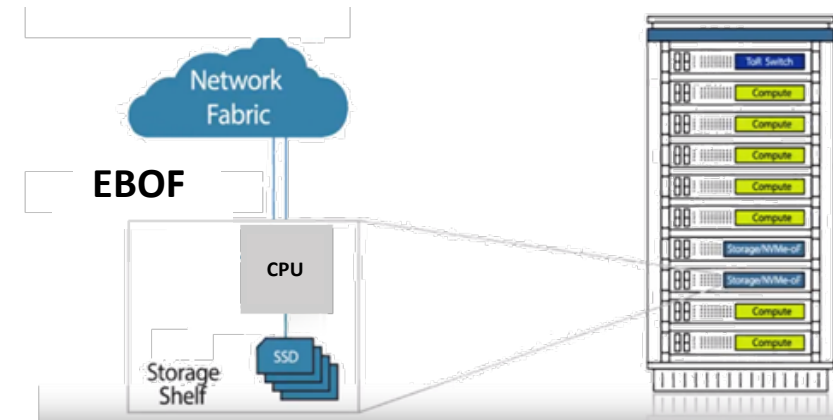
** Supports one 2x200G SOC RNIC connected with x16 PCIe Gen4

*** Supports three 200G Host connected Ethernet ports

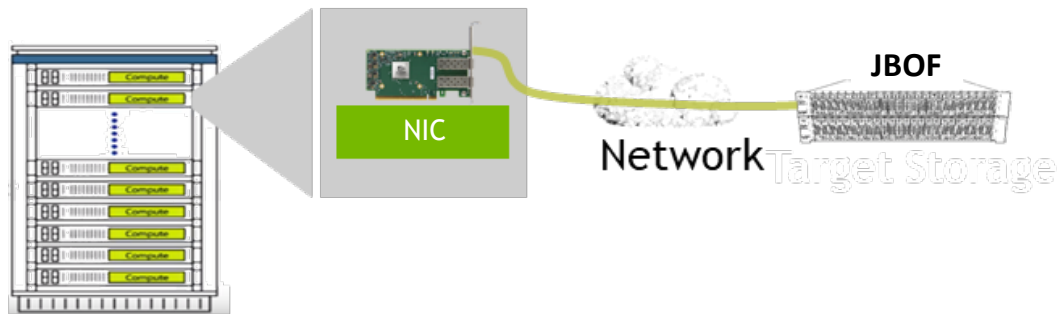
Use Cases for Ethernet SSDs



Backend NVMe-oF Cluster



Compute and Storage Disaggregation



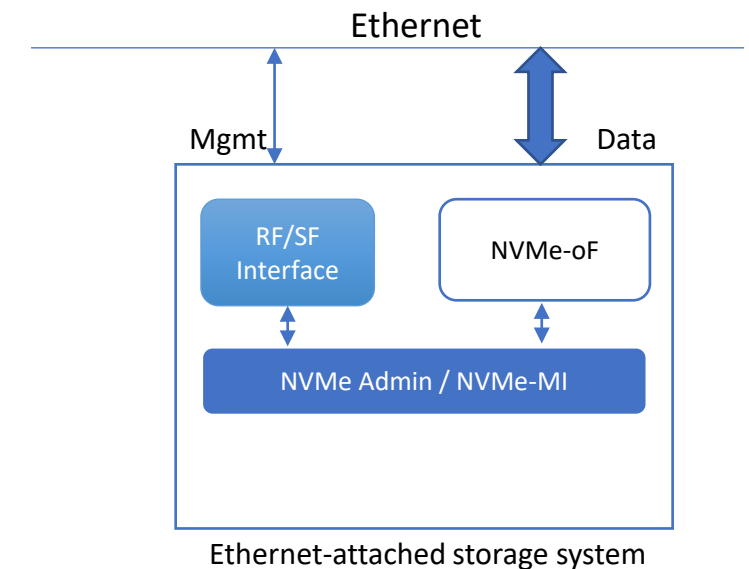
Server Based Storage / SDS

Standards-based Management for Ethernet-Attached Drives and EBOFs

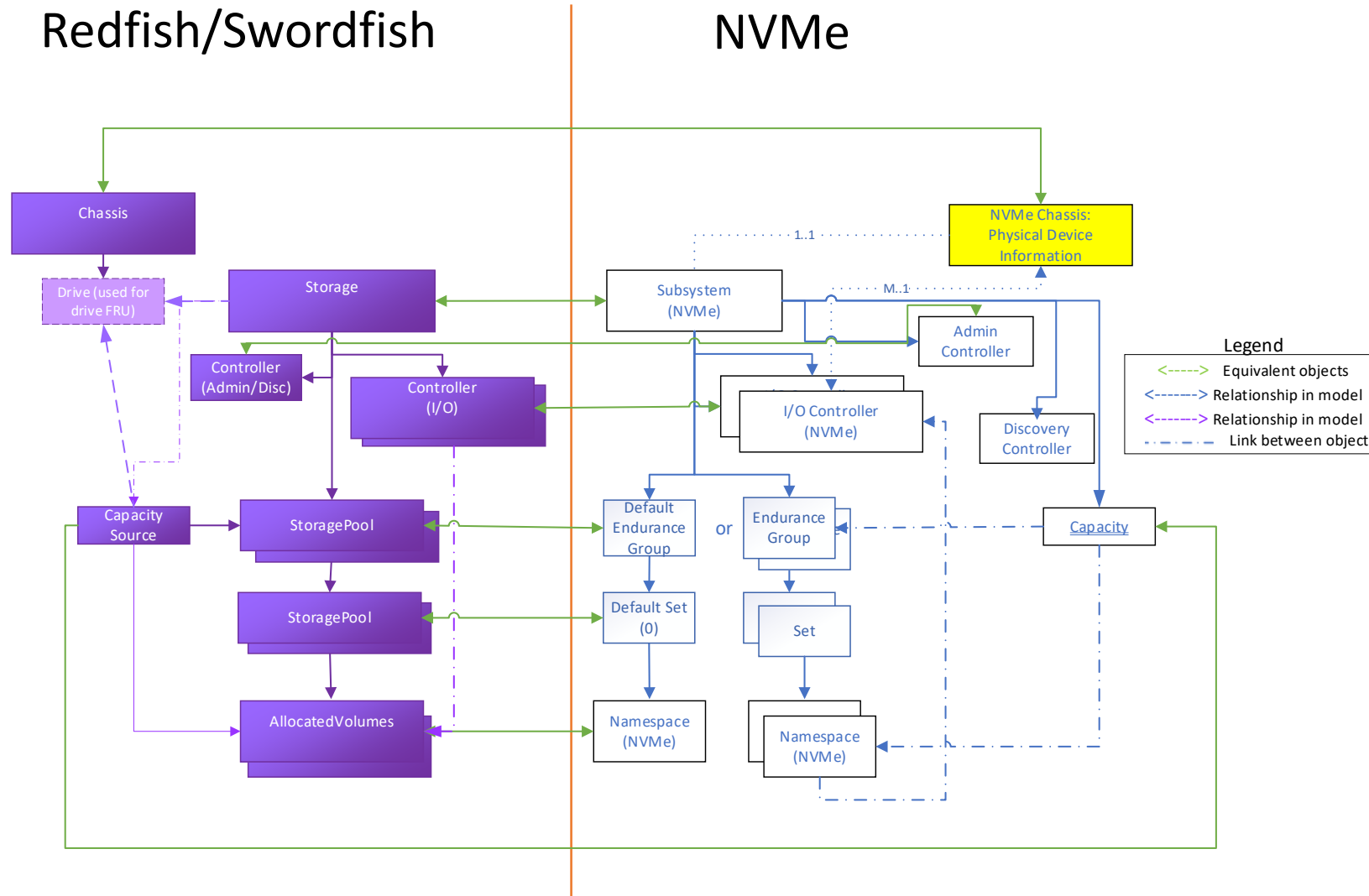


Management

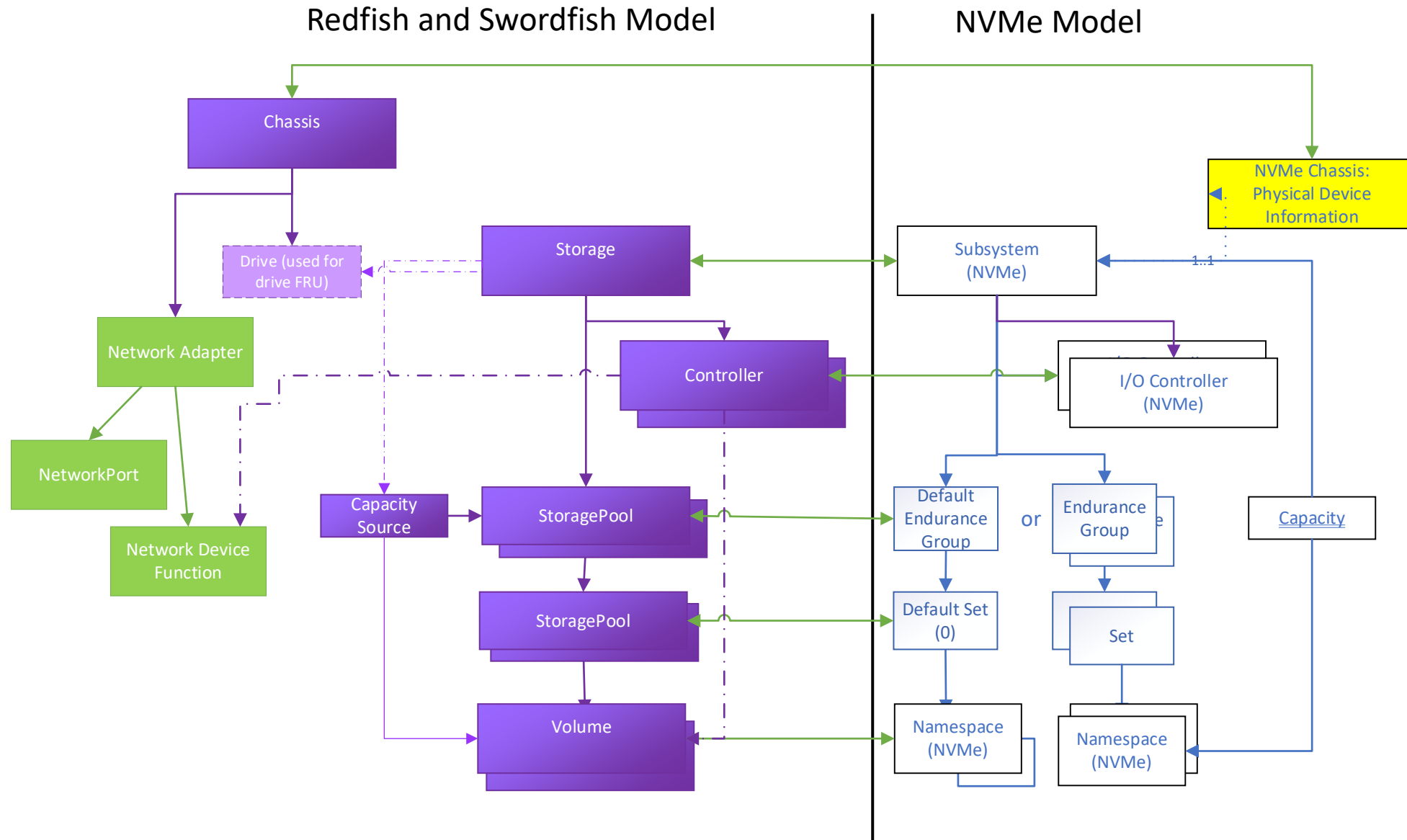
- Scale out orchestration of 10's of thousands of drives possible by using a RESTful API such as DTMF Redfish™ and SNIA Swordfish®
 - Redfish provides physical component models
 - Swordfish components add storage-specific functionality
- Each device can report its own management information directly
 - HTTP/TCP/Ethernet based management interface
 - Each system component provides a RF/SF interface
- Provides the same management capabilities as other NVMe devices
 - RF/SF use the available low-level transports to get device / transport specific information into the common models
 - RF/SF uses the commands that are provided in the NVMe/NVMe-oF/NVMe-MI specs
 - NVMe-MI can also be used as the low-level I/F to get the information into the high-level management environment as OOB access mechanism when appropriate
- Requirements included in Native NVMe-oF™ Drive Specification



Starting Point: Basic NVMe Subsystem Model



Ethernet Attached Drive: Management Model



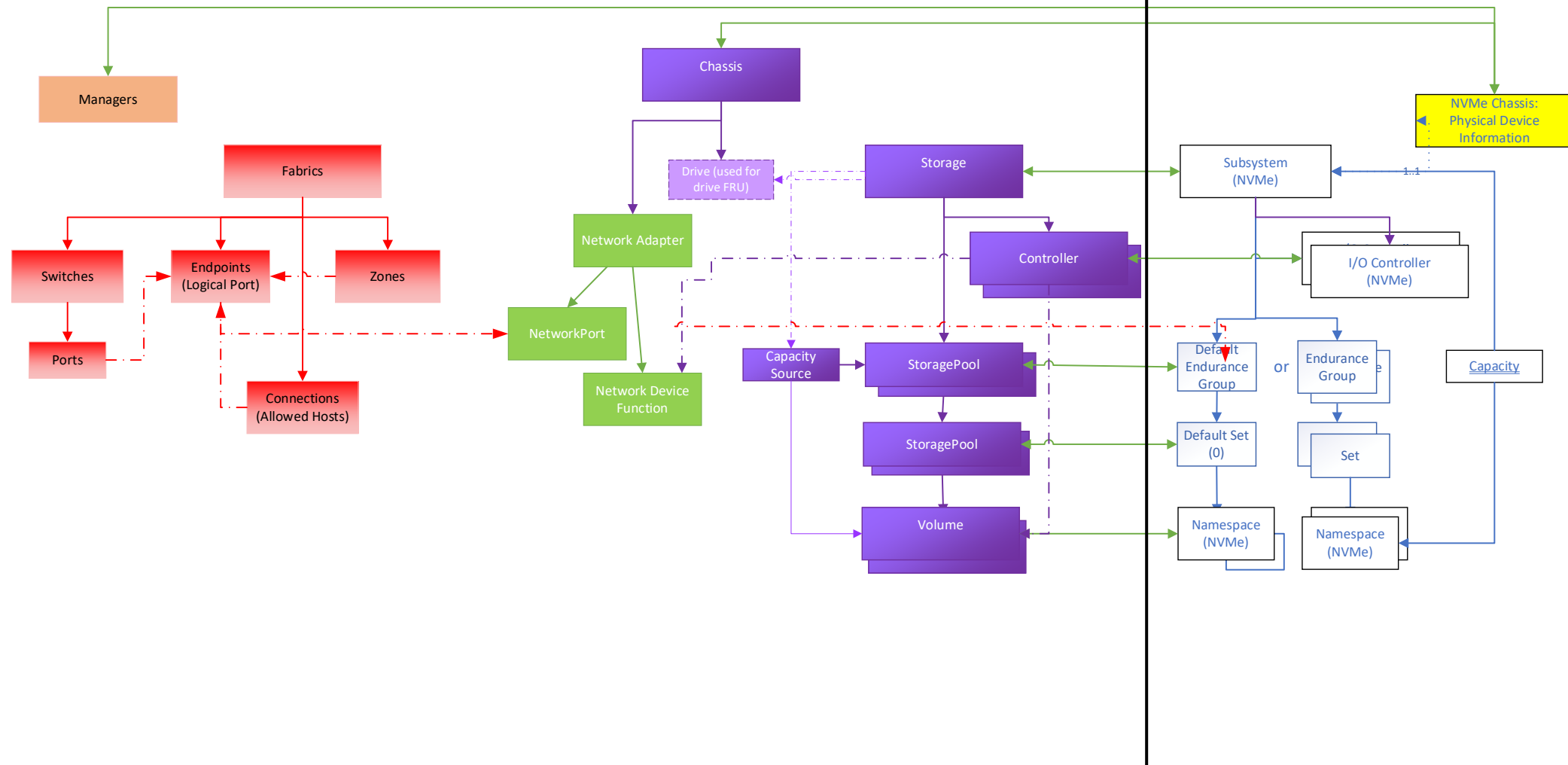
Adding the Enclosure and Switch...



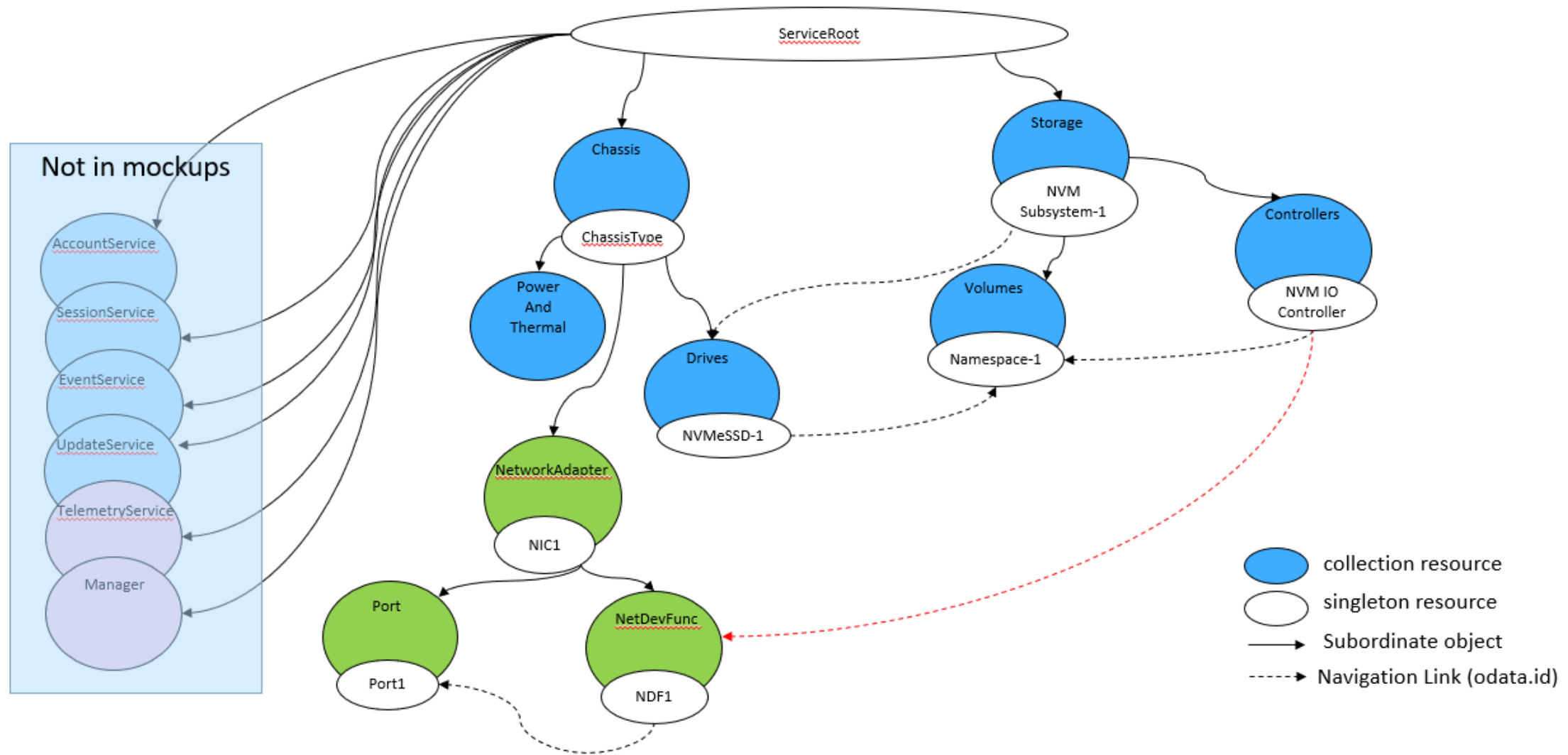
Flash Memory Summit

Redfish and Swordfish Model

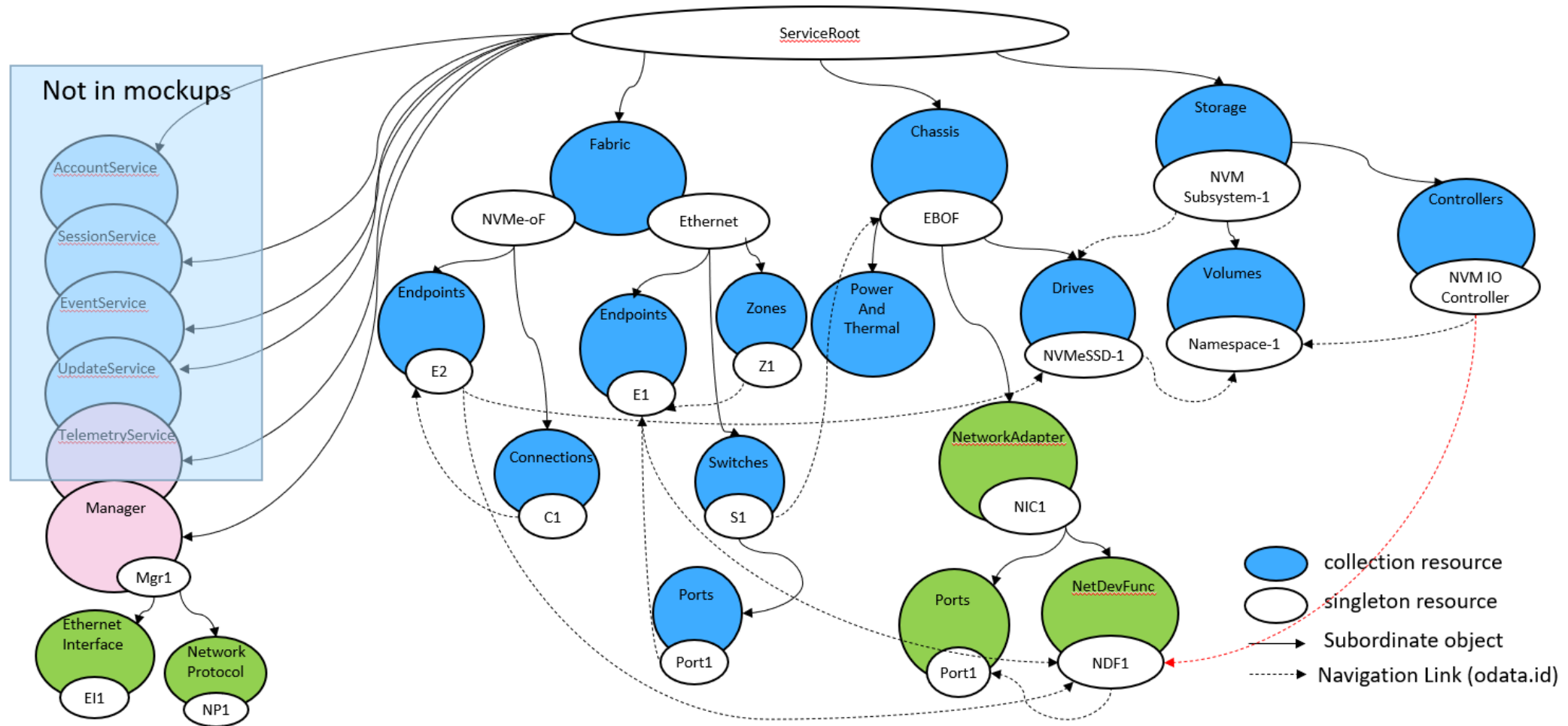
NVMe Model



Ethernet Attached Drive: Management Components



EBOF: Combining Enclosure and Drives



Where are the Details?

- Swordfish NVMe Model Overview and Mapping Guide
 - Defines the model to manage NVMe and NVMe-oF storage systems with Redfish and Swordfish, and provides the detailed mapping information between the NVMe, NVMe-oF specifications and the Redfish and Swordfish specifications.
- Swordfish profiles
 - Profiles define the detailed required and recommended functionality to implement
 - Clients use these as shorthand to specify requirements
 - Also used in vendor-neutral conformance testing – “Swordfish CTP”



Complete the Picture: From Requirements...

- Example: Space used in a volume
 - From the SwordfishNVMeFrontEnd.json profile the AllocatedBytes and ConsumedBytes values are required for “Volume” (NVMe namespace)

```
"Volume": {
  "PropertyRequirements": {
    "BlockSizeBytes": {},
    "Capacity": {
      "PropertyRequirements": {
        "Data": {
          "PropertyRequirements": {
            "AllocatedBytes": {},
            "ConsumedBytes": {}
          }
        }
      }
    },
    "DisplayName": {},
    "Description": {
      "Comparison": "Equal",
      "Values": ["A Namespace is a quantity of non-volatile memory that may be formatted in
to logical blocks. When formatted, a namespace of size n is a collection of logical blocks with
logical block addresses from 0 to (n-1). NVMe systems can support multiple namespaces."],
      "ReadRequirement": "Mandatory"
    }
  },
}
```


To Mapping from RF/SF to NVMe....

Get implementation guidance:

- How to I map the number of bytes used?

| | | | |
|---------|-----|--|-----|
| • RF/SF | 126 | BlockSizeBytes mapping | 210 |
| • From | 127 | Capacity.Data.ConsumedBytes mapping | 211 |
| | 128 | Capacity.Data.ProvisionedBytes mapping | 212 |

- From table 127 – this is in section 6.5.2.2, part of Namespace
 - Description – The number of bytes consumed in this data store for this data type
 - Mandatory - Yes

* Multiple NVMe objects have a concept of capacity used so this shows up several places

Show the End-to-End View

Property

Capacity.Da

ta.ConsumedBytes

NVM Spec Property /

Field: Namespace

Utilization (NUSE) **NVM**

Spec: Section:Figure

NVMe 1.4a: Section 5.15.2.1

(Identify Namespace),

Figure 247

- Find the guidance for a specific property:
ConsumedBytes
 - The mapping guide points you to the exact NVMe construct
 - Read the notes too, for important information

| | Redfish/Swordfish | NVMe / NVMe-oF |
|-------|---|--|
| Notes | Reporting capacity in bytes is the Redfish and Swordfish standard mechanism. Clients expect the capacity information to be reported consistently for these devices, so the calculation here is to convert the NVMe properties (in blocks) to bytes. | Returned in bytes 23:16 of the Identify Namespace Data Structure (NVM Command Set Specific). Reference NVMe Base Specification section n 5.15.2.1 and figure 247). |

SNIA Native NVMe-oF™ Drive Specification

- Discover and Configure: the drives, their interfaces, the speeds, the management capabilities
- Connectors
 - Some connectors may need to configure the PHY signals based on the type of drive interface
 - Survivability and mutual detection is important
- Pin-outs
 - For common connectors and form factors
- NVMe-oF integration
 - Discovery controllers / Admin controllers
- Management
 - Standard capabilities available for monitoring, control via RF/SF
 - Through Ethernet/TCP for Datacenter-wide management

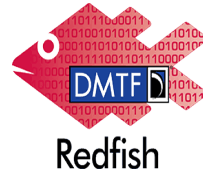
Where to Find More Info...

SNIA Swordfish™

- Swordfish Standards
 - Schemas, Specs, Mockups, User and Practical Guide's, ...
<https://www.snia.org/swordfish>
- Swordfish Specification Forum
 - Ask and answer questions about Swordfish
 - <http://swordfishforum.com/>
- Scalable Storage Management (SSM) TWG
 - Technical Work Group that defines Swordfish
 - Influence the next generation of the Swordfish standard
 - Join SNIA & participate: https://www.snia.org/member_com/join-SNIA
- Join the SNIA Storage Management Initiative
 - Unifies the storage industry to develop and standardize interoperable storage management technologies
 - <https://www.snia.org/forums/smi/about/join>

DMTF Redfish™

- Redfish Standards
 - Specifications, whitepapers, guides,...
<https://www.dmtf.org/standards/redfish>



SNIA Native NVMe-oF™ Drive Specification

SNIA Object Drive TWG

- Description & Links
<https://www.snia.org/nvme>

NVM Express



- Specifications
<https://nvmexpress.org/developers/>
- Join: <https://nvmexpress.org/join-nvme/>