STORAGE DEVELOPER CONFERENCE
SDC 21
BY Developers FOR Developers
Virtual Conference
September 28-29, 2021
A SNIA Event

Managing Exported NVMe-oF™ Resources and Fabrics in Swordfish™ and Redfish™

Phil Cayton
Senior Staff Engineer
Intel Corporation

- What is NVMe-oF and how could you use it?
- Why all this fuss about managing NVMe-oF resources
- What are we doing about it?
- How you can help

What is NVMe-oF and How Could You Use It?

Maps and extends NVMe's™ performance and latency benefits

- Enables scaling NVMe™ out and across network fabrics
- Provides higher IOPS / reduced latency from the host software stack, through data fabric to backend storage

Recent sea-change - The content from the NVMe-oF specification has been absorbed into other NVMe specifications

- The NVMe-oF protocol is now found in a combination of the NVMe 2.0 Base Specification and the NVMe Transport Specifications

Today, NVMe-oF delivers a new level of performance for busines-critical applications

- Shared Storage at speeds that rival Direct Attach Storage
- Composable, pooled, disaggregated storage with increased utilization efficiency

Opens new possibilities for HPC & clouds, what to expect from workloads and use-cases which use very large datasets

- Big-data analytics, Predictive modelling, Large-volume OLTP, AI/ML, …

Why All This Fuss About Managing NVMe-oF

Current state of NVMe-oF configuration management:

- Configuration and management of each individual NVMe-oF resource
  - Enumeration of NVMe & fabric resources prior to configuration requires a-priori knowledge
  - Configuring (and reconfiguring) NVMe-oF Subsystems and Namespaces
  - Configuring (and reconfiguring) fabric transport(s) and paths
  - Configuring (and reconfiguring) access control
  - Provisioning Host access to NVMe-oF resources
  - Informing Hosts of changes to configured NVMe-oF resources

Configuration and management mechanisms

- Manual or static configuration: labor intensive, lack flexibility, decrease practical scalability
- Proprietary tools: non-standard, non-interoperable

Limits usability, flexibility, scale of dynamic installations of NVMe-oF

What Are We Doing About It?

To support large-scale, dynamic, deployments of NVMe-oF, more is needed

- Scalable/dynamic configuration management/provisioning at a datacenter level and scale
- Standard and much more efficient way to:
  - Abstract physical storage
  - Enable them to be used over one or more fabrics and/or fabrics-paths
- Provide a standardized, efficient, management method to virtualize NVMe Namespaces

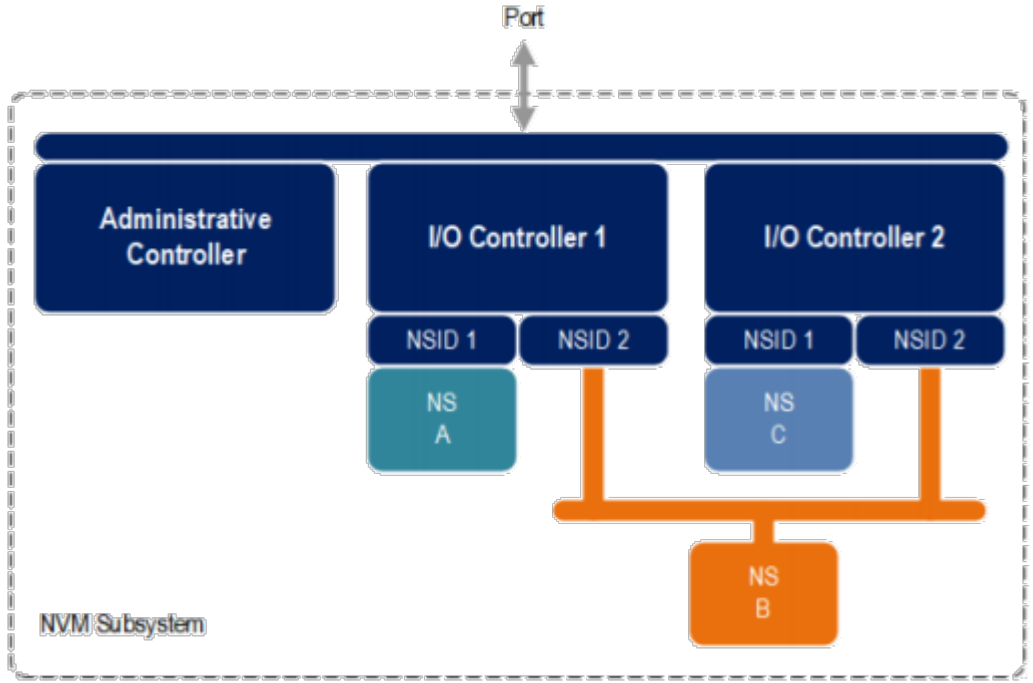
DMTF Redfish & SNIA Swordfish standards extensions for NVMe/NVMe-oF

- Already define, configure, manage datacenter resources
- DMTF, SNIA, and NVMe experts adding comprehensive scalable-storage profiles to Swordfish to manage NVMe and NVMe-oF resources

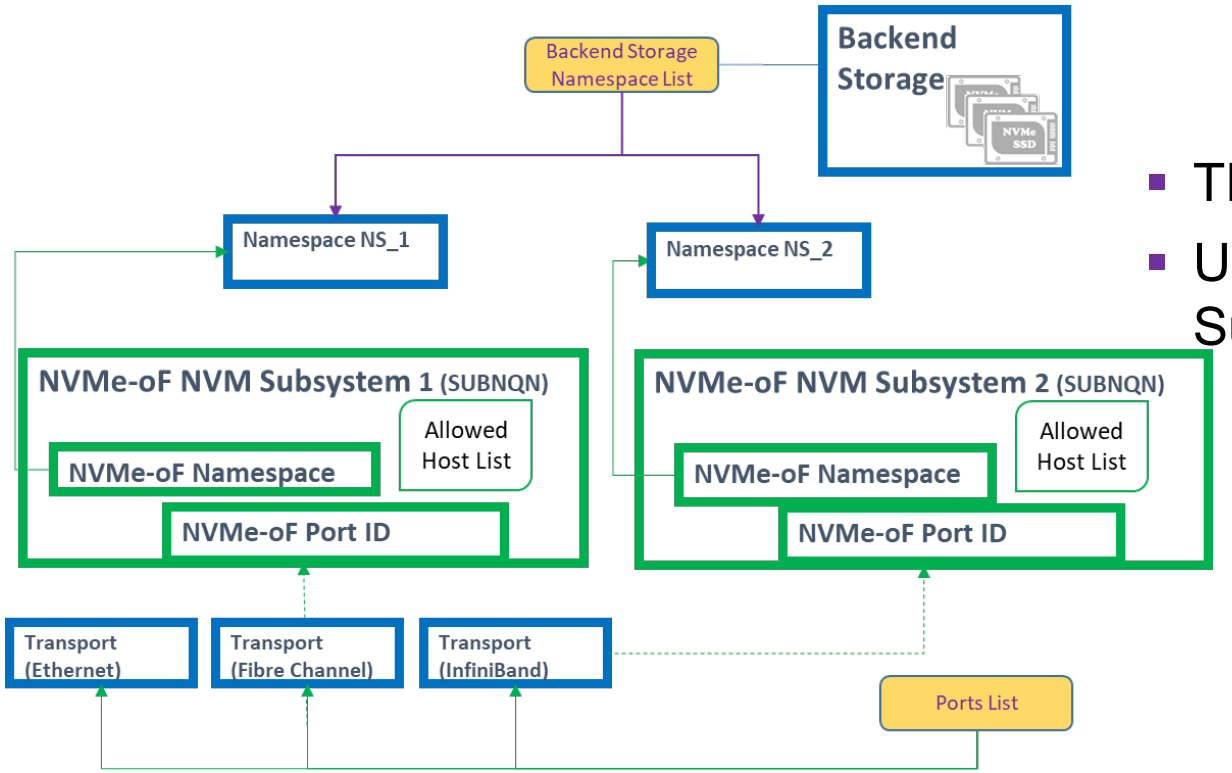
Starting with NVMe Objects

- **NVM Subsystem** presents a collection of one or more NVMe controllers which are used to access namespaces
- **NVMe Namespace (NS):** a storage volume – a collection of logical block addresses accessible to host software
- **Namespace ID (NSID):** an identifier used by a controller to provide access to an NVMe Namespace

Next up: Mapping NVMe Namespaces onto fabrics

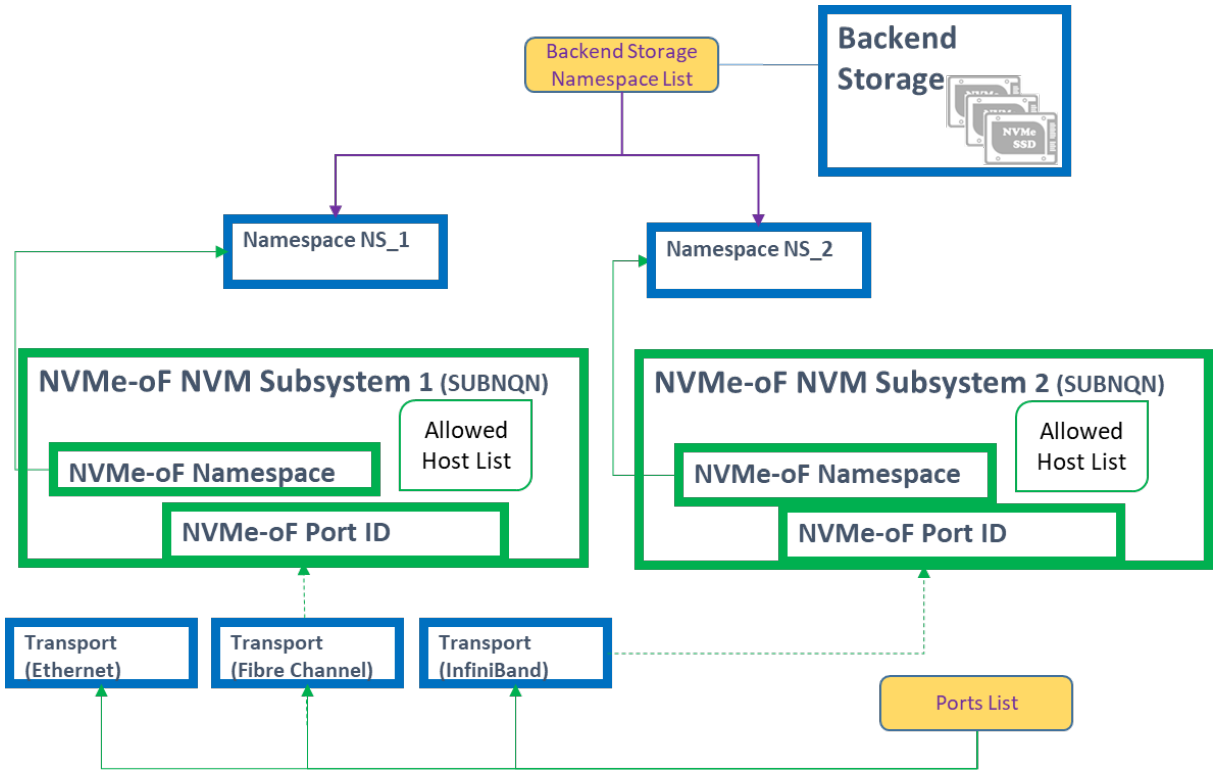


Mapping NVMe Objects Onto Fabrics



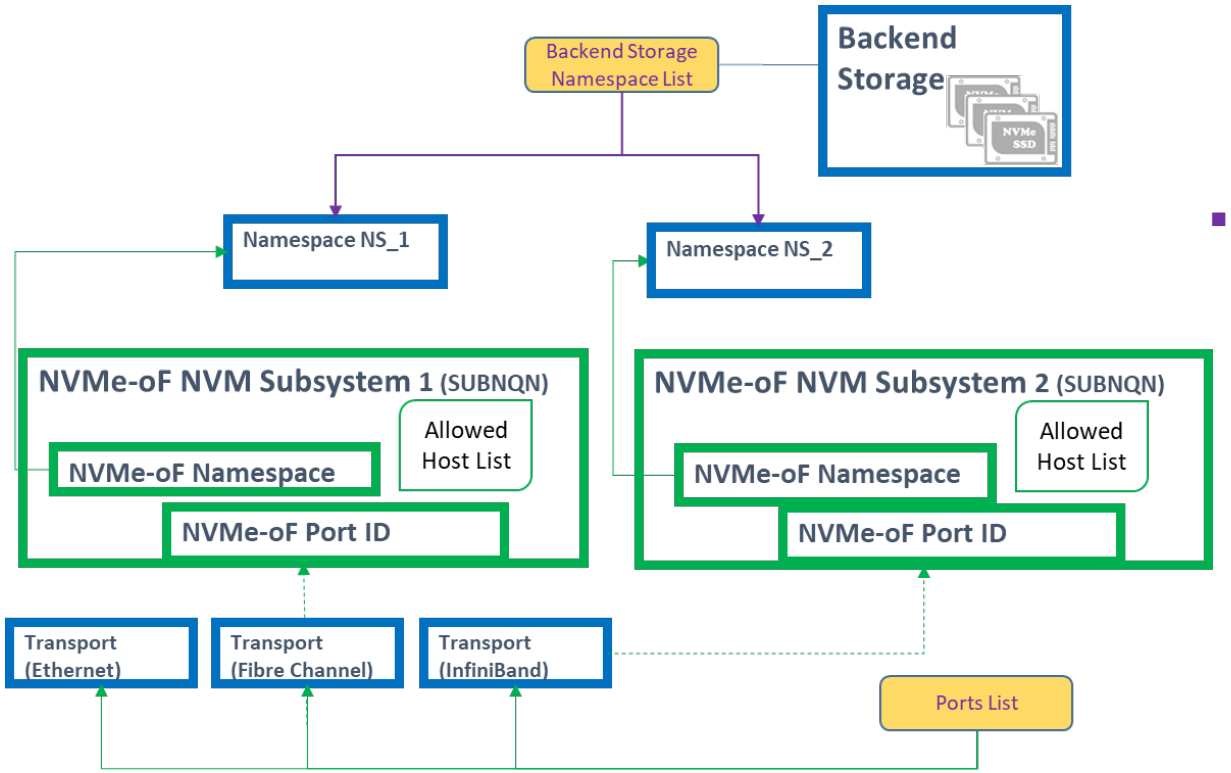
- This represents managing layered storage constructs
- Underlying NVM Subsystem and exported logical NVM Subsystems Maintain a ‘landlord / tenant relationship’
  - Exported NVM Subsystems use Underlying NVM resources
  - Exported Namespaces do manage backend physical resources (e.g., capacity)

Mapping NVMe Objects Onto Fabrics



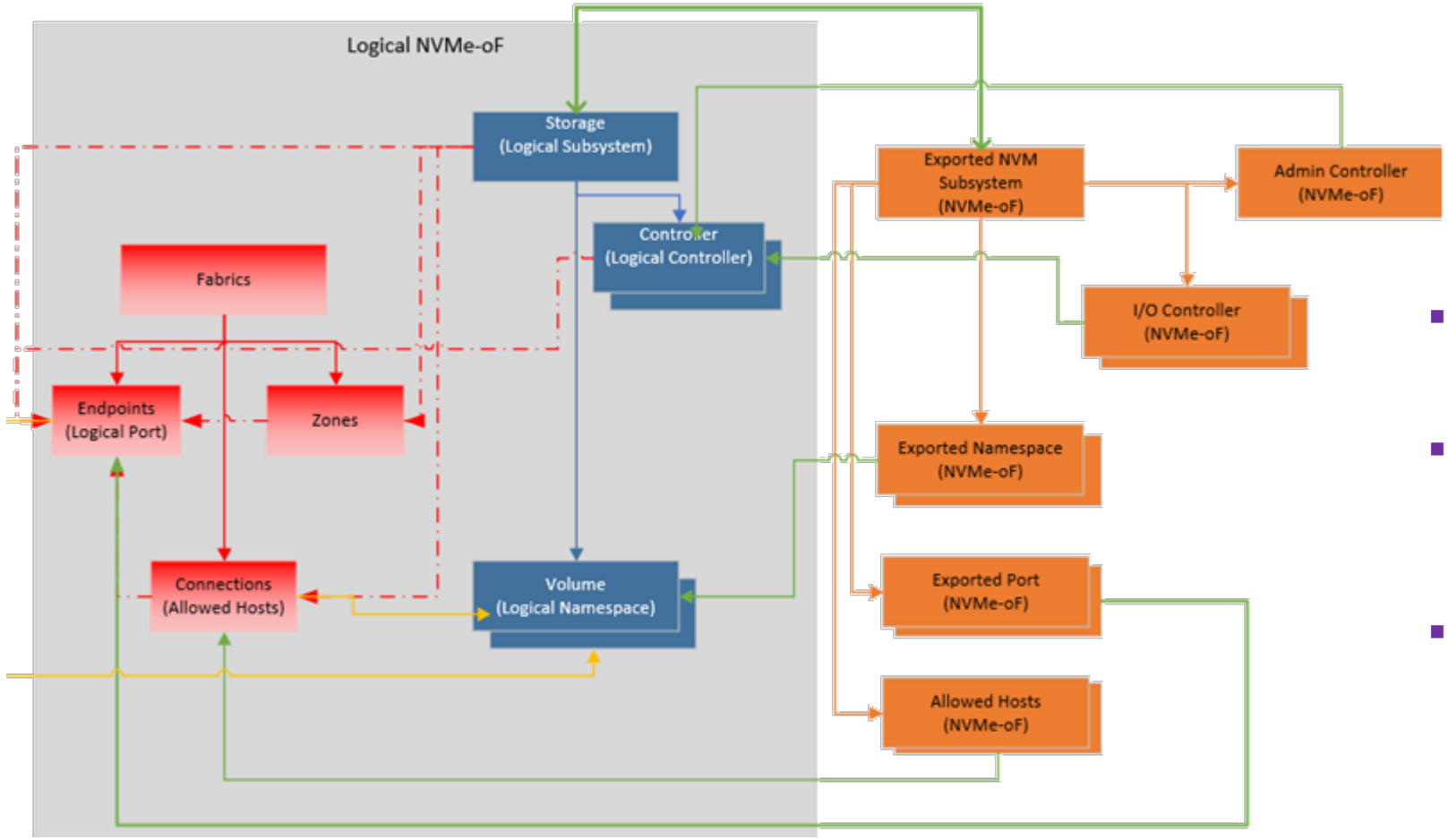
- These are virtual constructs on top of physical instances of storage
- Doesn't have to be NVMe SSD backed
  - This is what it looks like on top of NVMe
  - Could simply be an NVMe front-end onto ANY block storage

Interfaces For Standardized Management



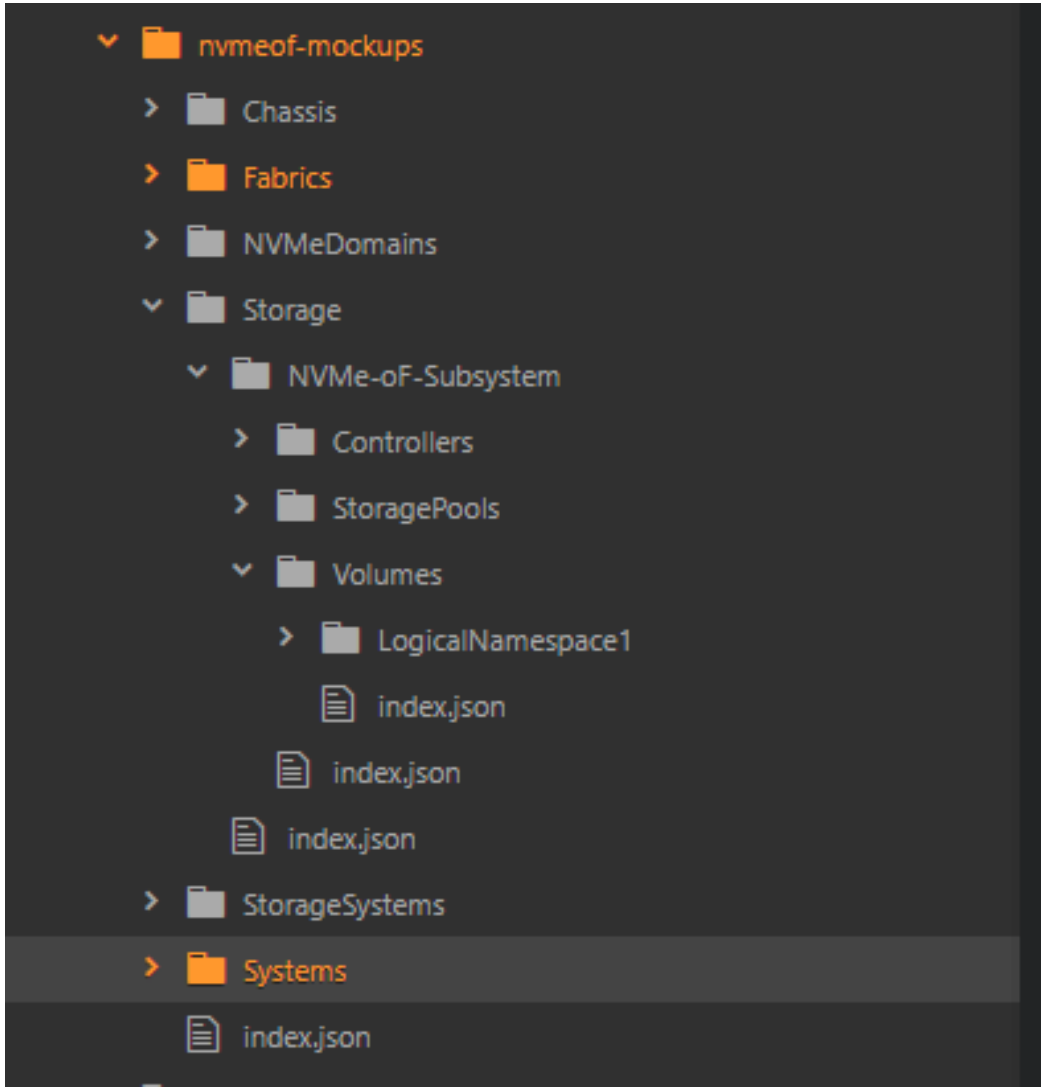
- The NVMe Consortium is developing standardized mechanisms to enable a framework for configuration and management for exporting NVMe resources
  - Get List of available storage resources
  - Get List of Ports available
  - Create/Delete exported NVM Subsystems
  - Link Namespaces to NVM Subsystems
  - Manage exported Transport Configuration
  - Manage Host access rights to Exported Subsystems
- SNIA is working in parallel with NVMe to use this framework to enable scalable storage management at datacenter scale

What is SNIA Swordfish Doing About It?



- Leveraging existing RF/SF models
- Map NVMe-oF objects onto RF/SF to replicate the NVMe model
- Enables configuration of logical representations of NVMe resources

Representing NVMe-oF in Redfish & Swordfish



Representing NVMe-oF in Redfish & Swordfish

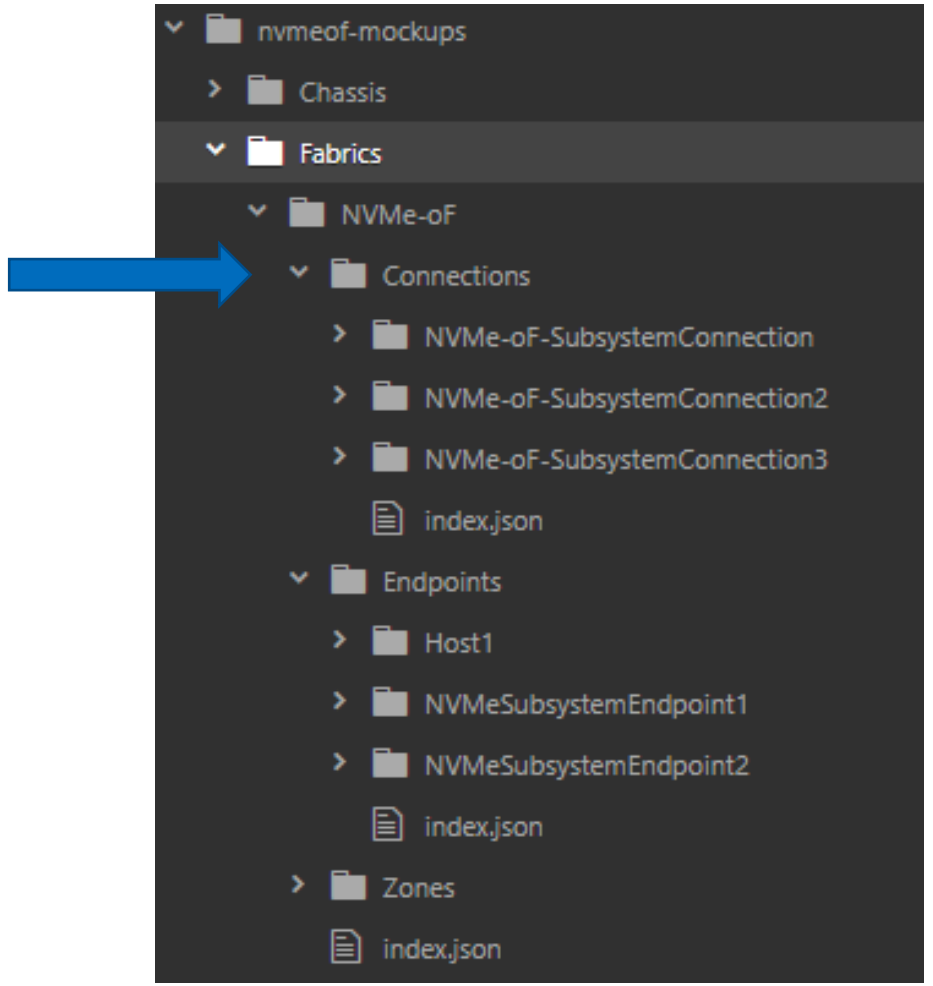

```
{
  "@Redfish.Copyright": "Copyright 2014-2020 SNIA. All rights reserved.",
  "@odata.id": "/redfish/v1/Storage/NVMe-oF-Subsystem",
  "@odata.type": "#Storage.v1_10_0.Storage",
  "Id": "1",
  "Name": "NVMe-oF Logical NVM Fabric System",
  "Description": "An NVM Express Subsystem is an NVMe device that contains one or more NVM Expr
  "Status": {
    "State": "Enabled",
    "Health": "OK",
    "HealthRollup": "OK"
  },
  "Identifiers": [{
    "DurableNameFormat": "NQN",
    "DurableName": "nqn.2014-08.org.nvmexpress:uuid:6c5fe566-10e6-4fb6-aad4-8b4159f50245"
  }],
  "Controllers": {
    "@odata.id": "/redfish/v1/Storage/NVMe-oF-Subsystem/Controllers"
  },
  "Volumes": {
    "@odata.id": "/redfish/v1/Storage/NVMe-oF-Subsystem/Volumes"
  },
```


Representing NVMe-oF in Redfish & Swordfish

```
"Links": {
    "Endpoints": [
        {
            "@odata.id": "/redfish/v1/Fabrics/NVMe-oF/Endpoints/NVMeSubsystemEndpoint1"
        },
        {
            "@odata.id": "/redfish/v1/Fabrics/NVMe-oF/Endpoints/NVMeSubsystemEndpoint2"
        }
    ]
    "Connections":[
      {
        "@odata.id": "/redfish/v1/Fabrics/NVMe-oF/Connections/NVMe-oF-SubsystemConenction2",
      },
      {
        "@odata.id": "/redfish/v1/Fabrics/NVMe-oF/Connections/NVMe-oF-SubsystemConenction3",
      }
    ]
}
}
```


Provisioning Logical Namespaces

Representing NVMe-oF in Redfish & Swordfish

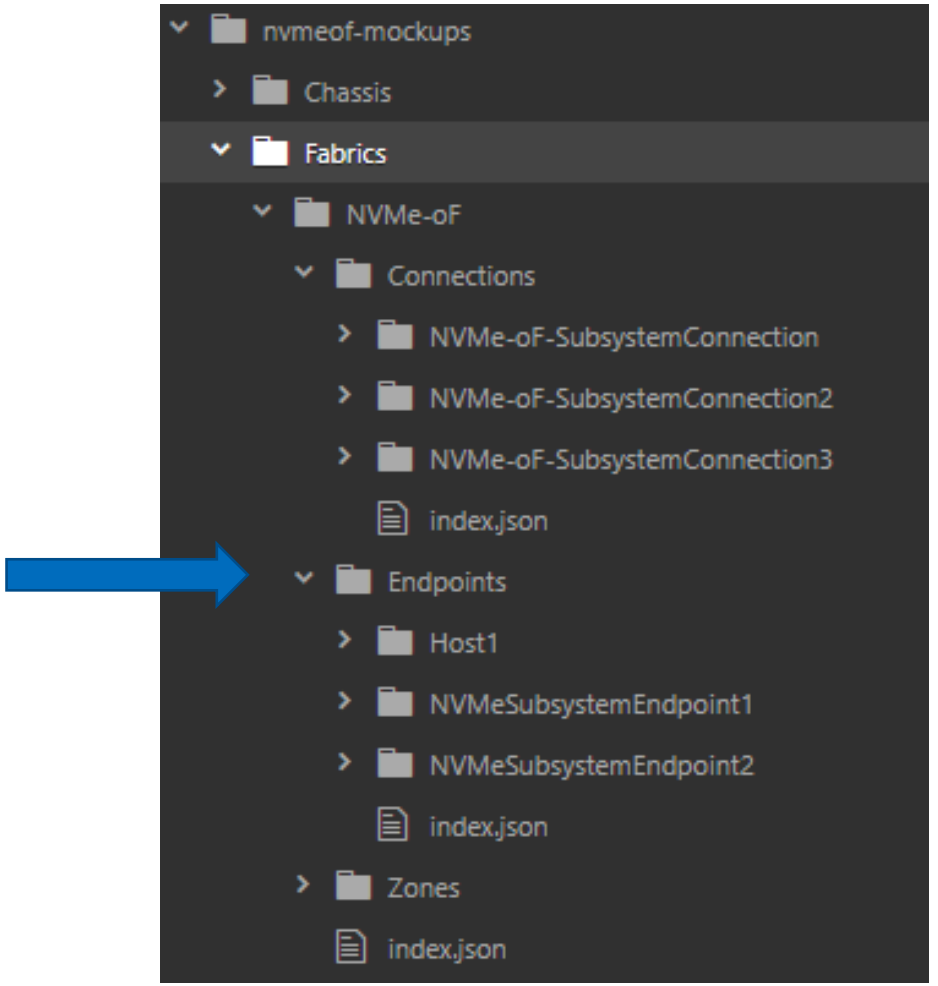


Configuring Access: Unrestricted Read/Write

Configuring Access: Restricted Read-Only

```
{
  "@odata.type": "#Connection.v1_0_0.Connection",
  "Id": "NVMe-oF-SubsystemConnection",
  "Name": "Connection info for NVMe-oFSubsy    m Restricted access",
  "ConnectionType": "Storage",
  "VolumeInfo": [{
    "AccessCapabilities": ["Read"],
    "Volume": {
      "@odata.id": "/redfish/v1/Storage/NVMe-oF-Subsystem/Volumes"
    }
  }],
  "Links": {
    "InitiatorEndpoints": [{
        "@odata.id": "/redfish/v1/Fabrics/NVMe-oF/Endpoints/Host2"
        },
        {
        "@odata.id": "/redfish/v1/Fabrics/NVMe-oF/Endpoints/Host3"
        }],
    "TargetEndpoints": [{
        "@odata.id": "/redfish/v1/Fabrics/NVMe-oF/Endpoints/NVMeSubsystemEndpoint1"
    }]
  },
  "@odata.id": "/redfish/v1/Fabrics/NVMe-oF/Connections/NVMe-oF-SubsystemConenction2",
  "@Redfish.Copyright": "Copyright 2014-2020 SNIA. All rights reserved."
}
```


Representing NVMe-oF in Redfish & Swordfish



Representing NVMe-oF in Redfish & Swordfish


```
{
  "@odata.type": "#Endpoint.v1_4_0.Endpoint",
  "Id": "1",
  "Name": "NVMeSubsystemEndpoint",
  "Description": "Endpoint connected NVMe-oF Subsystem (used in Zone)",
  "EndpointProtocol": "NVMeOverFabrics",

  "ConnectedEntities": [{
      "EntityType": "StorageSubsystem",
      "EntityRole": "Target",
      "EntityLink": {
         "@odata.id": "/redfish/v1/Systems/sys-1/Storage/NVMe-oF-Subsystem"
      },
    },
    {
      "EntityType": "NetworkController",
      "EntityRole": "Target",
      "EntityLink": {
        "@odata.id": "/redfish/v1/Chassis/Sys-1Chassis/NetworkAdapters/1/NetworkDeviceFunctions/1"
      }
    }
  ],
```


Representing NVMe-oF in Redfish & Swordfish


```
"Identifiers": [{
  "DurableNameFormat": "NQN",
  "DurableName": "nqn.2014-08.org.nvmexpress:uuid:6c5fe566-10e6-4fb6-aad4-8b4159f50246"
}],

"IPTransportDetails": [{
  "TransportProtocol": "RoCEv2",
  "IPv4Address": {
    "Address": "192.168.155.55"
  },
  "Port": 4420
}],

"@odata.id": "/redfish/v1/Fabrics/NVMe-oF/Endpoints/NVMeSubsystemEndpoint1",
"@Redfish.Copyright": "Copyright 2014-2021 SNIA. All rights reserved."
}
```

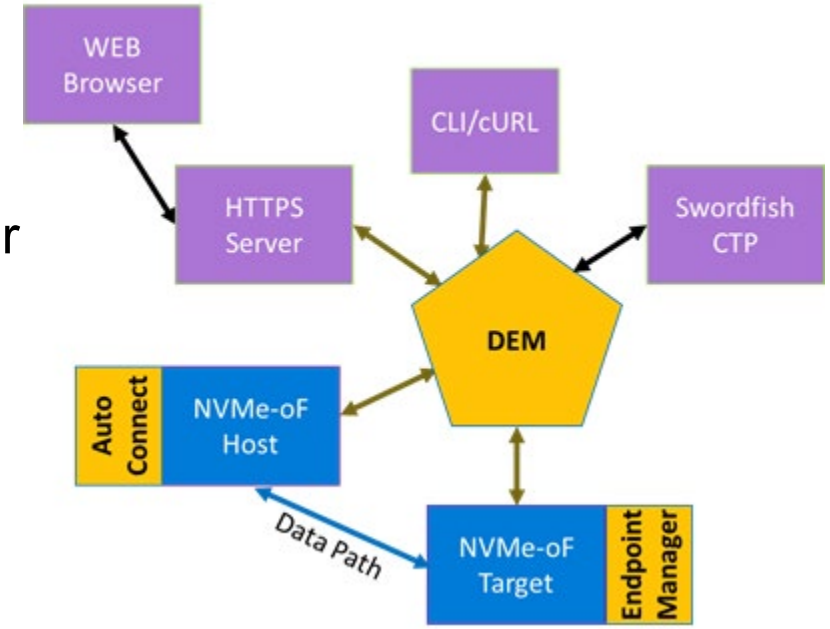

What Are We Doing About It?

- The NVMe Consortium is developing standardized mechanisms to enable a framework for configuration and management for exporting NVMe resources
- SNIA is working in parallel with NVMe to use this framework to enable scalable storage management in the datacenter
  - SNIA is verifying modeling of NVMe and fabric components is complete in RF/SF
  - producing schema, mockups, enhanced mapping guide(s)
  - Publishing CTP compliant open-source industry strawman demonstration / reference platform
- For more on using Swordfish to implement NVMe management, please watch Curtis Ballard's presentation: 'Completing the Picture for NVMe and NVMe-oF Management: Guidelines for Implementations'

Distributed Endpoint Manager (DEM)

- Open-source reference implementation of an NVMe-oF Management Suite
  - HTTP front-end implementing a Swordfish RESTful interface to NVMe-oF Targets
  - Remote configuration & provisioning of NVMe-oF resources via RESTful interface
  - Complies with SNIA Swordfish Conformance Test Program

- For more on this, please watch Rajalaxmi Angadi's presentation: Accelerating NVMe / NVMe-oF RF/SF Development Using DEM For
  - an in-depth discussion of the components and capabilities
  - A demonstration of the management suite
  - How to get it, use it, and how to help it grow



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

Open Fabric Management Framework (OFMF)

- OFMF Working Group (OFMFWG)
  - Description & Links https://www.openfabrics.org/working-groups/
- OFMFWG mailing list subscription
  - https://lists.openfabrics.org/mailman/listinfo/ofmfwg
- Join the Open Fabrics Alliance
  - https://www.openfabrics.org/membership-how-to-join/

NVM Express

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

Please take a moment to rate this session.

Your feedback is important to us.

