



OPENFABRICS  
ALLIANCE

15<sup>th</sup> ANNUAL WORKSHOP 2019

# DISTRIBUTED ENDPOINT MANAGEMENT

## AN NVME-OF™ SCALE-OUT MANAGEMENT SOLUTION

Phil Cayton

Intel Corporation

March, 2019

# AGENDA

- **NVMe over Fabrics (NVMe-oF) Overview**
- **Current State of NVMe-oF Management and Administration**
- **Distributed Endpoint Management (DEM) Project**
- **Brief Demonstration**
- **Development Opportunities and Wrap-up**

# NVME OVER FABRICS

## NVMe-oF Overview

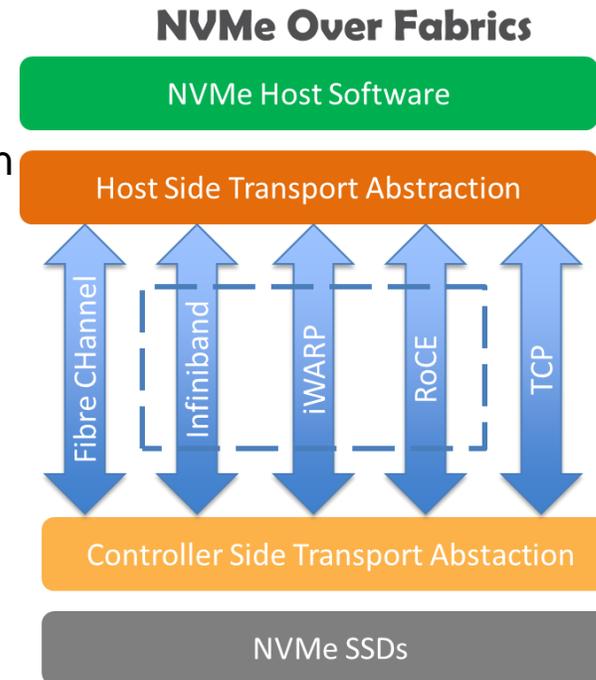
### **NVMe: industry standard interface and storage protocol for PCIe SSDs**

- High-performance, low-latency PCI SSD interface
- Eliminates unnecessary protocol translations (i.e., SCSI)
- Defines partitioning PCIe SSDs into one or more subsystems

### **NVMe-oF: extends NVMe efficiency over Fabrics**

- Builds on base NVMe architecture with thin encapsulation of base NVMe across a fabric
- Enables low-latency and high IOPS access to remote NVMe storage
- Defines end-to-end mechanisms to transfer NVMe commands and data structures

The NVMe-oF spec is not Fabric specific;  
separate Transport Bindings are defined for each Transport



# NVME OVER FABRICS

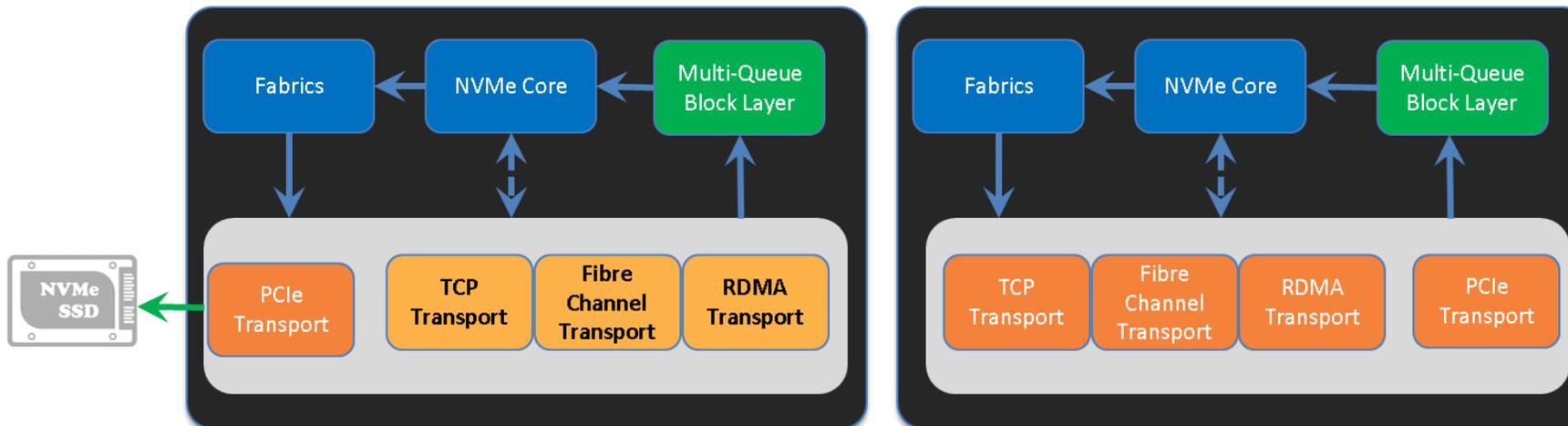
## NVMe-oF Overview

### Targets

- Create logical NVMe subsystems and Controllers that are presented to Hosts
- Logically map NVMe Namespaces to physical NVMe block devices
- Export NVMe subsystems virtualizing NVM Namespaces
- May be provisioned to allow individual Hosts access to specific resources

### Hosts

- Discover provisioned NVMe-oF resources from Targets
- Connect to provisioned resources



# DISTRIBUTED ENDPOINT MANAGEMENT

## Current State of NVMe-oF Management and Administration

**Linux In-kernel Implementation only supports local management**

**Each individual Target is either manually or statically configured**

- Configure each Fabric with Address, Port, ...
- Define each NVMe-oF Subsystem
- Assign NVMe resources to each NVMe-oF Subsystem
- Set up Individual Host Access rights to each NVMe-oF Subsystem

**Each individual Host either accesses NVMe-oF resources through:**

- Static predefined configurations
- Manual resource discovery process (as defined by NVMe-oF specification)
  - Connects to each individual Target
  - Requests resources on that Target that they may access

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



OPENFABRICS  
ALLIANCE

# DISTRIBUTED ENDPOINT MANAGEMENT (DEM) PROJECT

# DISTRIBUTED ENDPOINT MANAGEMENT

## DEM Project

### Distributed Endpoint Management (DEM): An Open-Source Project

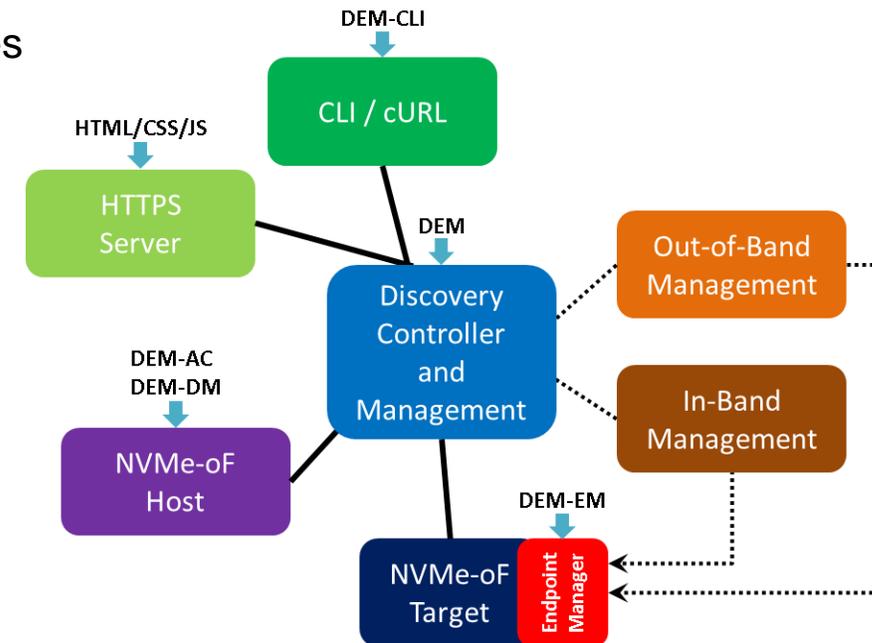
Enable efficient, dynamic configuration and provisioning of NVMe-oF Resources

Started out as a strawman driving specification changes into NVMe, NVMe-oF, and NVMe-MI

Management suite enabling:

- Remote configuration of NVMe-oF resources through RESTful interface
- Centralized enumeration of provisioned resources
- Single source for notification of changes to resources

Adopted by UNH-IOL with additional scripts written for Interoperability Test Suite



# DISTRIBUTED ENDPOINT MANAGEMENT

## DEM Project

### Distributed Endpoint Management (DEM): An Open-Source Project

Enable efficient, dynamic configuration and provisioning of NVMe-oF Resources

Started out as a strawman driving specification changes into NVMe, NVMe-oF, and NVMe-MI

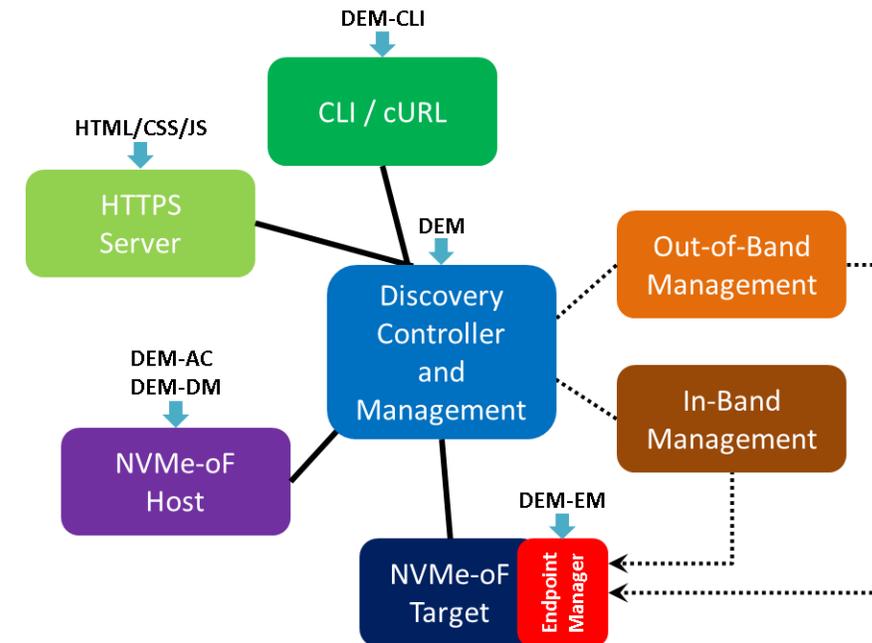
Management suite supporting:

#### NVMe-oF Transports

- RDMA - Validated on IB/iWARP/RoCE
- TCP

#### Configuration via

- In-band (i.e., using NVMe-oF protocol)
- Out-of-Band (i.e., using RESTful interface via JSON)



# DISTRIBUTED ENDPOINT MANAGEMENT

## DEM Project

### Project Components

#### Discovery controller + Management (DEM)

#### Endpoint Manager (DEM-EM) – customized for Target implementation

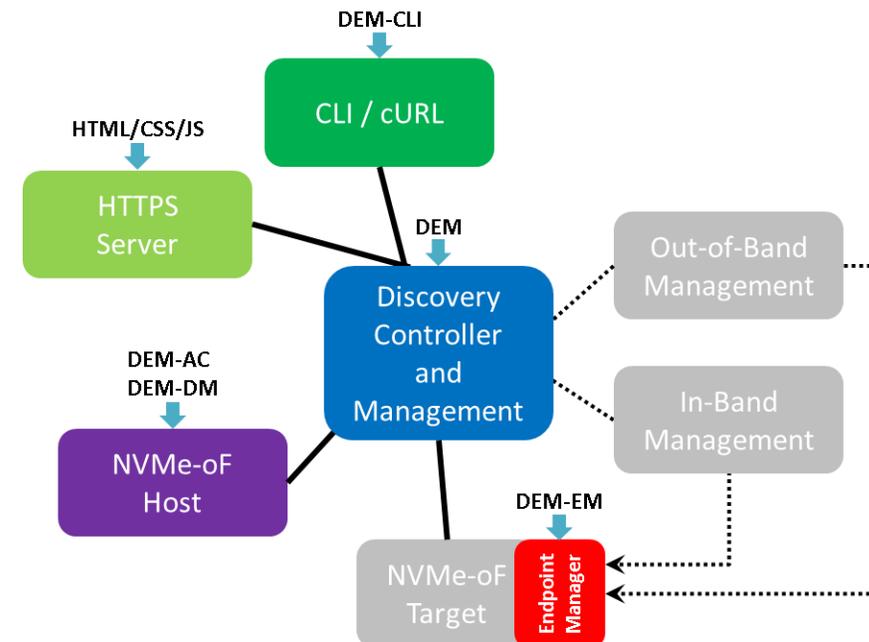
- In-Band Mode
- Out-of-Band Mode

#### Optional Host Tools

- Auto Connect (DEM-AC)
- Discovery Log Page Monitor (DEM-DM)

#### RESTful Management Interfaces

- Web Pages (HTML / CSS files)
- Command Line Interface (DEM-CLI)

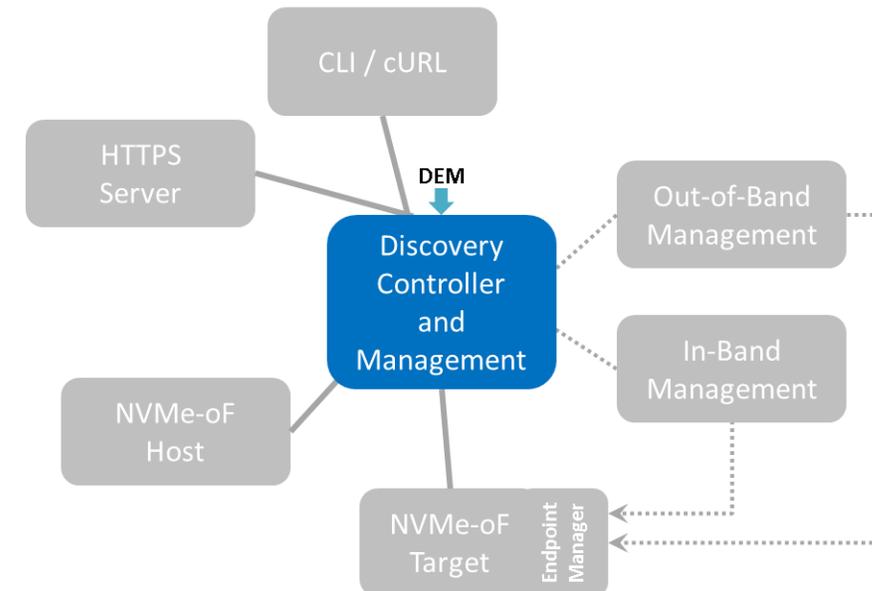


# DISTRIBUTED ENDPOINT MANAGEMENT

## Project Components

### Discovery controller + Management (DEM)

- Primary component for remote configuration and provisioning
- Plug-in module architecture for NVMe-oF supported Fabrics
- Configures remote NVMe resources via In-Band or Out-of-Band interfaces
- Collects & distributes tailored Discovery Log Pages to Hosts
- Receives notification of changes to NVMe-oF resources
- Reports changes to NVMe-oF resources to affected registered Hosts
- Enables additional access restrictions

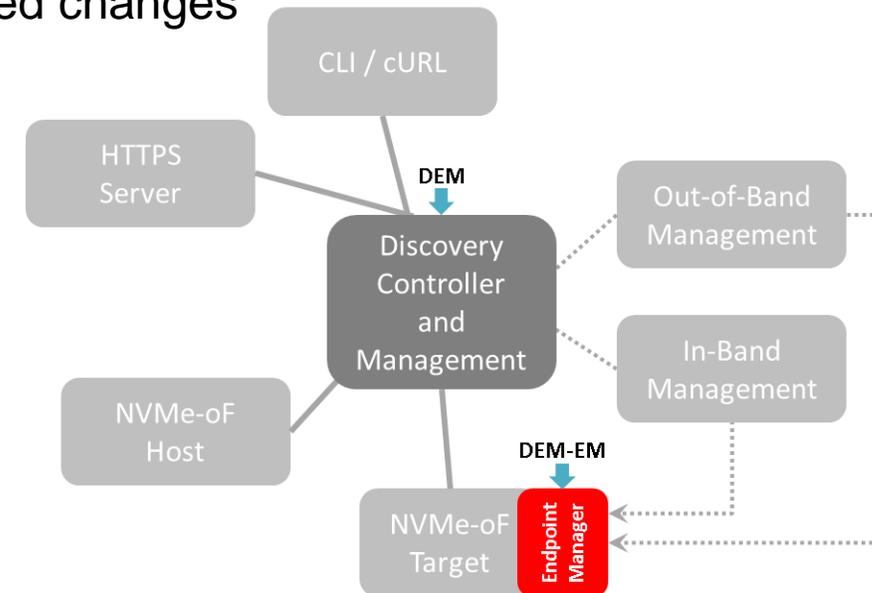


# DISTRIBUTED ENDPOINT MANAGEMENT

## Project Components

### Endpoint Manager (DEM-EM)

- Agent running on Target enabling remote configuration
- Reuses DEM plug-in module architecture for NVMe-oF supported Fabrics
- Plug-in configuration model enabling implementation-specific management of NVMe-oF resources
- In-Band configuration based on proposed changes to NVMe-MI specification
- Out-of-Band RESTful configuration based on proposed changes to RF/SF
- Used for Targets not managed by other means



# DISTRIBUTED ENDPOINT MANAGEMENT

## Project Components

### Optional Host Tools

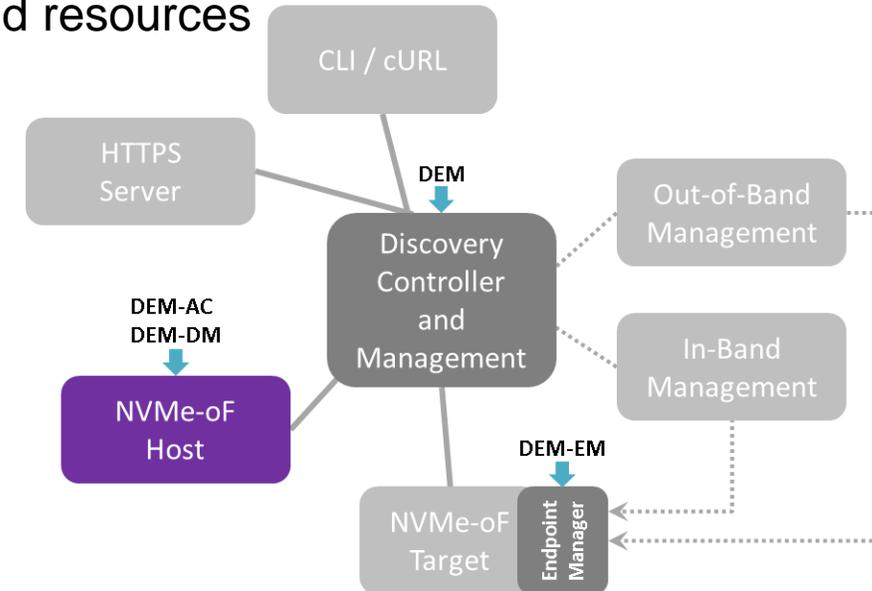
- Reuses DEM plug-in module architecture for NVMe-oF supported Fabrics

### Auto Connect (DEM-AC)

- Establishes persistent connection
- Collects Discovery Log Pages
- Automatically connects to its provisioned resources

### Discovery Log Page Monitor (DEM-DM)

- Establishes persistent connection
- Reports Log Change Events and displays updated Discovery Log Pages



# DISTRIBUTED ENDPOINT MANAGEMENT

## Project Components

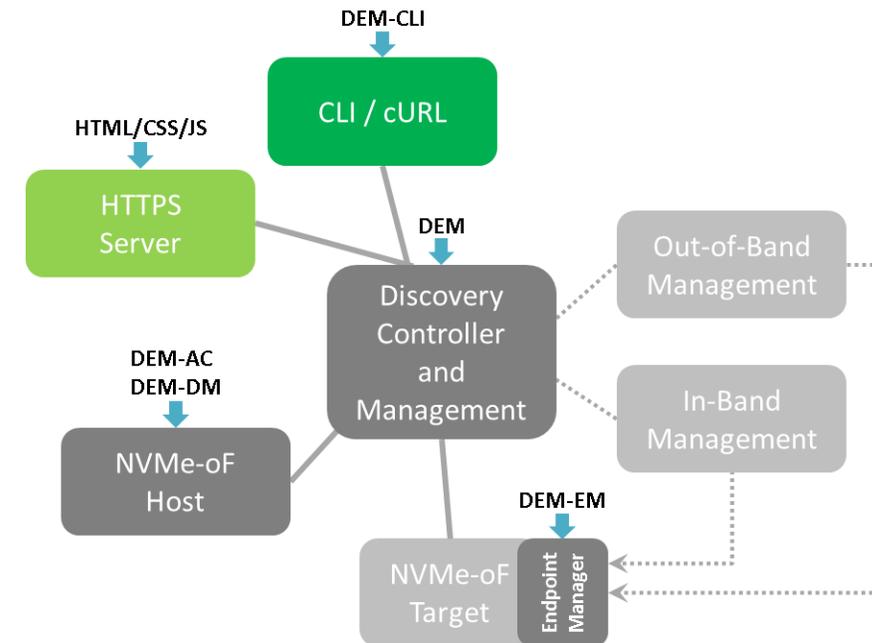
### RESTful Management Interfaces

#### DEM Command Line Interface (DEM-CLI)

- Local Interface to the DEM via console command line

#### Web interface

- Interface to the DEM via web interface
- Project contains complete set of HTML, CSS, and JS files





OPENFABRICS  
ALLIANCE

# BRIEF DEM-ONSTRATION

# DEMONSTRATION

NVMe-oF DEM - dem x +

← → ↻ ⚠ Not secure | 10.23.221.69/dem/#dem

## Distributed Endpoint Management

an implementation of a Centralized NVMe-oF Discovery Controller

Objects Info

### About

Use the menus to the right to view the defined objects managed by this **Distributed Endpoint Manager (DEM)**.

Currently active fabric interfaces for NVMe-oF Hosts to query are listed below.

### Interfaces

0: rdma ipv4 192.168.1.1:4422

version 0.1

Screen shows interfaces Hosts use to access Log Pages from the DEM

# DEMONSTRATION

NVMe-oF DEM - dem

Not secure | 10.23.221.69/dem/#dem

## Distributed Endpoint Management

an implementation of a Centralized NVMe-oF Discovery Controller

**Objects** **Info**

**Targets**

**Hosts**

**Groups**

### About

Use the menus to the right to view the defined objects managed by this **Distributed Endpoint Manager (DEM)**.

Currently active fabric interfaces for NVMe-oF Hosts to query are listed below.

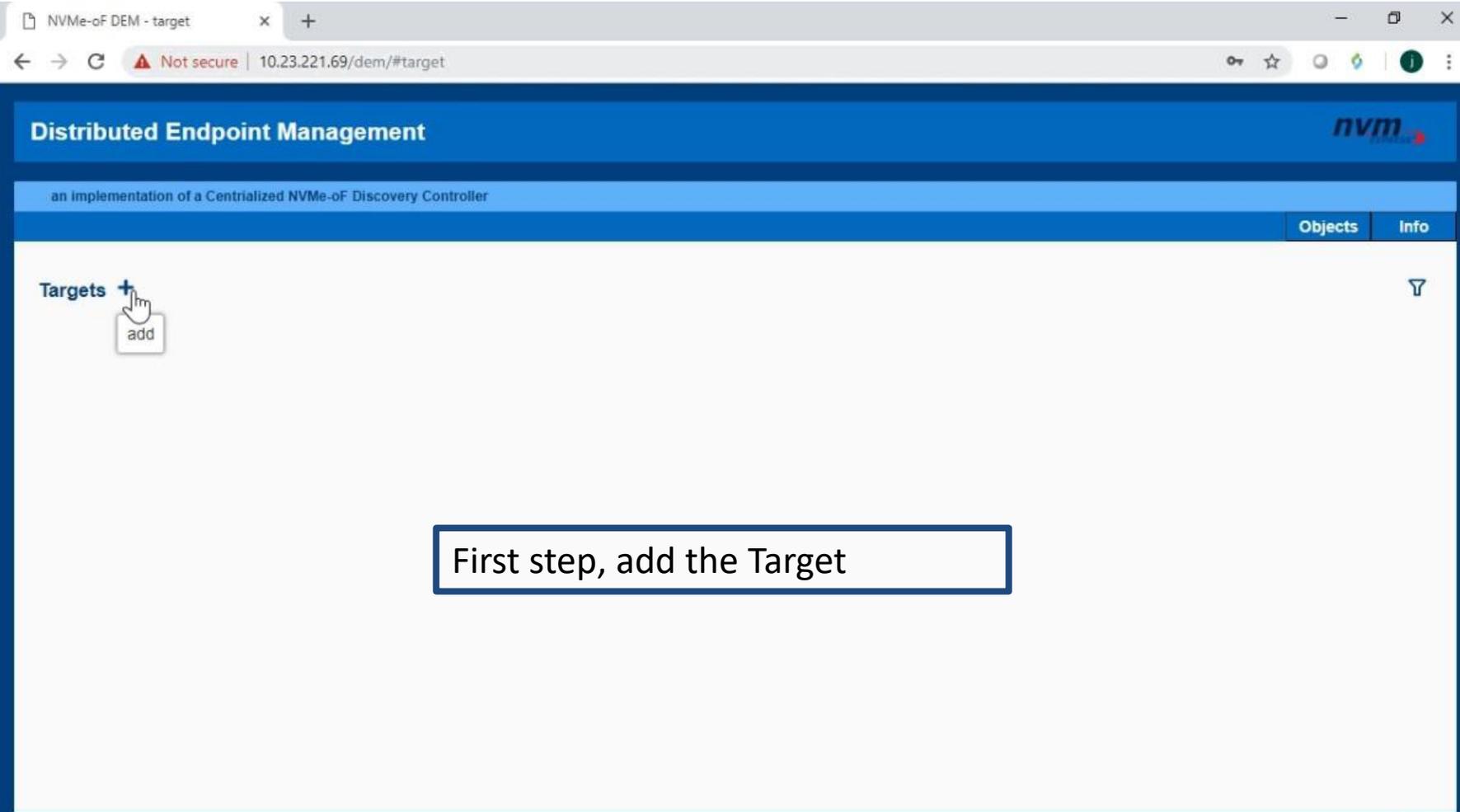
### Interfaces

0: rdma ipv4 192.168.1.1:4422

version 0.1

Configure remote Target to determine existing configuration

# DEMONSTRATION



The screenshot shows a web browser window with the title "NVMe-oF DEM - target". The address bar shows "10.23.221.69/dem/#target" with a "Not secure" warning. The page header is "Distributed Endpoint Management" with the "nvm" logo. Below the header, it says "an implementation of a Centralized NVMe-oF Discovery Controller". There are tabs for "Objects" and "Info". The main content area is titled "Targets" and contains a "+" icon with a hand cursor and an "add" button. A search icon is visible in the top right of the content area. The version "0.1" is displayed in the bottom right corner.

Targets +  
add

Objects Info

version 0.1

First step, add the Target

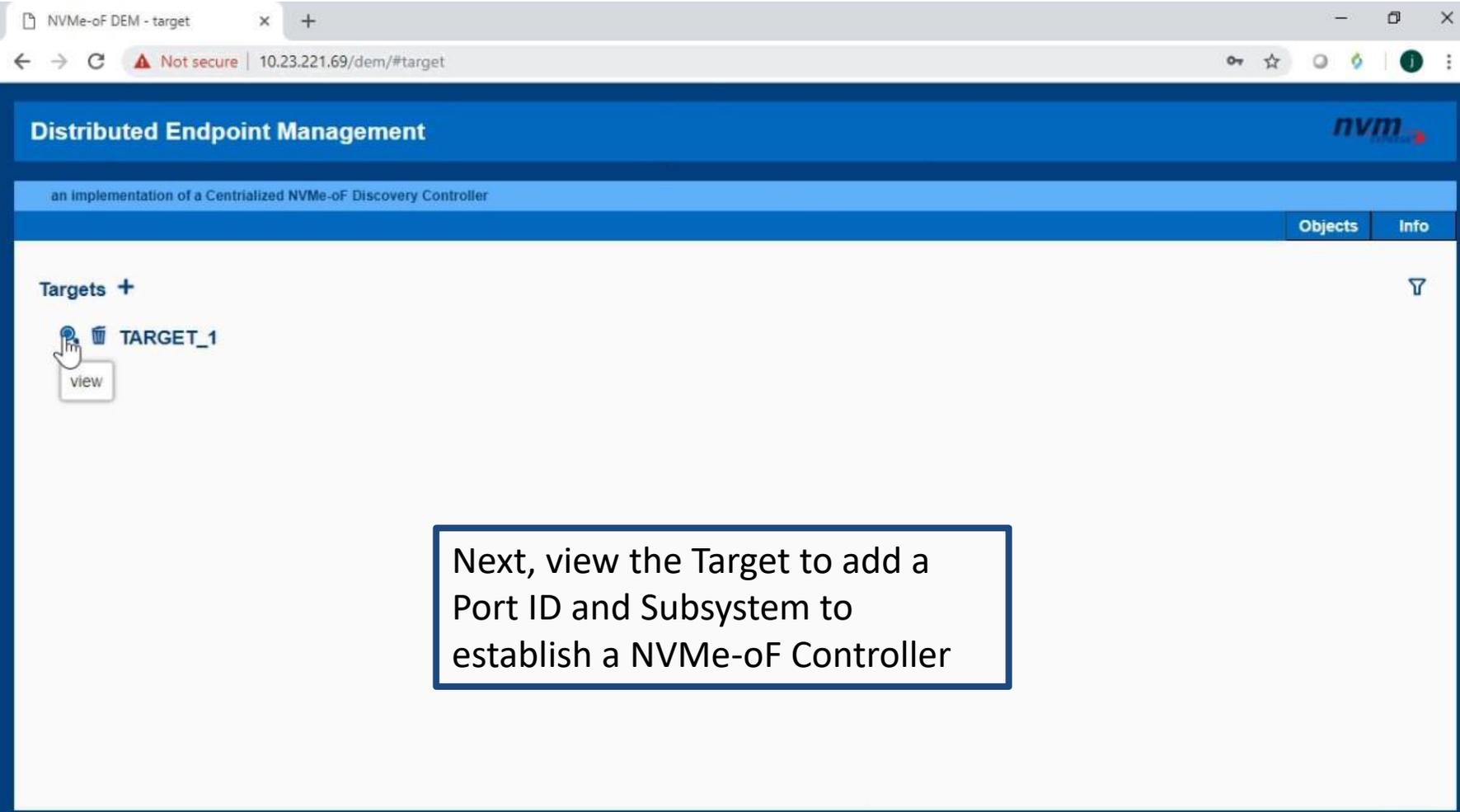
# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target`. The page title is "Distributed Endpoint Management". A modal dialog titled "Add a Target" is open, containing the following fields and options:

- Alias:**
- Management Mode:**  (dropdown menu)
- Locally Managed:** Targets need to poll logpages periodically for resource changes
- Periodic Resource Updates:** Refresh:  minutes - 0 disables timer Log Page refreshing

A callout box with a blue border contains the text: "Set the Target as Locally Managed. This will allow DEM to view a Target's configuration".

# DEMONSTRATION



The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target`. The page title is "Distributed Endpoint Management" and it includes the *nvm* logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". There are tabs for "Objects" and "Info". A section titled "Targets +" contains a single entry, "TARGET\_1", with a "view" button and a trash icon. A mouse cursor is hovering over the "view" button.

Next, view the Target to add a Port ID and Subsystem to establish a NVMe-oF Controller

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it is described as "an implementation of a Centralized NVMe-oF Discovery Controller". The interface features a blue header with the "nvm" logo and tabs for "Objects" and "Info". The main content area displays the following information:

- Alias: TARGET\_1 (with edit, copy, and back icons)
- Management Mode: Local
- Port IDs (with a plus icon and a hand cursor pointing to it)
- Subsystem (with an "add" button and a plus icon)

A callout box with a blue border and white background is positioned over the "Port IDs" text, containing the text "Next, add a Port ID".

# DEMONSTRATION

The screenshot shows a web browser window with the address bar displaying "10.23.221.69/dem/#target/TARGET\_1". The main content area is titled "Distributed Endpoint Management" and features a modal dialog box for adding a port to a target. The modal contains the following fields:

- Port ID: 1
- TRTYPE: rdma
- ADRFAM: ipv4
- TRADDR: 192.168.1.1
- TRSVCID: 4420

A text box with a blue border provides the following instruction:

Provide the existing information to connect to the Target. Info must be known a priory for connection to the existing configuration

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it includes the `nvm` logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". The main content area shows configuration for "Alias: TARGET\_1" with "Management Mode: Local". Under "Port IDs", there is one entry: "1: rdma ipv4 192.168.1.1:4420". The "Subsystems" section has a plus sign and an "add" button. A callout box with a blue border contains the text: "Next, add an arbitrary Subsystem for the creation of a NVMe-oF Controller".

# DEMONSTRATION

NVMe-oF DEM - target/TARGET\_1 x +

← → ↻ ⚠ Not secure | 10.23.221.69/dem/#target/TARGET\_1

Distributed Endpoint Management

**Add a Subsystem to Target 'TARGET\_1'**

Subsystem NQN:

Allow Any Host:

Provide the Subsystem NQN and set to Allow Any Host access

javascript:void(0)

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it includes the `nvm` logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". The interface has tabs for "Objects" and "Info". The main content area displays configuration for "Alias: TARGET\_1" with a "Management Mode: Local" and a "view log pages" button. Under "Port IDs", there is one entry: "1: rdma ipv4 192.168.1.1:4420". Under "Subsystems", there is one entry: "Subsystem NQN: SUBSYS1 (Allow Any Host)". A text box in the foreground explains that the controller can now query the target for preconfigured log pages.

Alias: TARGET\_1  
Management Mode: Local  
view log pages  
Port IDs +  
1: rdma ipv4 192.168.1.1:4420  
Subsystems +  
Subsystem NQN: SUBSYS1 (Allow Any Host)

Now that there is a Controller to connect to, DEM can query the Target for the set of preconfigured Log Pages

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1/logpage`. The page title is "Distributed Endpoint Management" and it includes the "nvm" logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". The main content area shows "Target: TARGET\_1" and "Log Pages" with a refresh icon. A red heading "Unattached Log Pages" is displayed above a JSON string: `subnqn="preconfig_subsys" subtype="nvme subsystem" portid=1 trtype="rdma" adrfam="ipv4" traddr=192.168.1.1 trsvcid=4420 req="not required" rdma: prtype="not specified" qptype="connected" cms="rdma-cm" pkey=0x0000`. A text box at the bottom of the screenshot explains that this is a preconfigured Log Page that does not match the DEM configuration, as indicated by the "Unattached" heading. The version "0.1" is visible in the bottom right corner of the interface.

This screen shows a preconfigured Log Page that does not match how the DEM has been configured as is evident from the Unattached heading

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it includes the `nvm` logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". The interface has tabs for "Objects" and "Info". The main content area displays configuration for "Alias: TARGET\_1". Under "Management Mode", there is an "edit" button. Under "Port IDs", there is one entry: "1: rdma ipv4 192.168.1.1:4420". Under "Subsystems", there is one entry: "Subsystem NQN: SUBSYS1 (Allow Any Host)". A text box in the foreground contains the text: "Next, the Target will be reconfigured to allow the DEM to configure the Target".

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management". A modal dialog titled "Update Target 'TARGET\_1'" is open, displaying the following configuration fields:

- Alias:
- Management Mode:
- Endpoint Manager configuration using RESTful interface to configure target
- Family:
- Address:
- RESTful Port:
- Periodic Resource Updates
- Refresh:

A text box with a blue border contains the following text: "The Target is configured for Out-of-Band Management through the DEM-EM residing on the Target. Admin must know how Target EM is configured/started".

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it includes the `nvm` logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". The interface has tabs for "Objects" and "Info". The main content area displays the following configuration for "Alias: TARGET\_1":

- Management Mode: Out-of-Band
- Interface: ipv4 192.168.1.1:22334
- Interfaces**
  - 0: rdma ipv4 192.168.1.1
- NSDevices**
  - 0: Device: ID -1
- Port IDs**
  - 1: rdma ipv4 192.168.1.1:4420
- Subsystems**
  - Subsystem NQN: SUBSYS1 (Allow Any Host)
  - NSIDs**

A text box with a blue border is overlaid on the right side of the interface, containing the text: "The DEM has now retrieved the Interfaces and NVMe Namespace Devices exported by the Target".

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it includes the `nvm` logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". The interface has tabs for "Objects" and "Info". The main content area displays configuration for "Alias: TARGET\_1".

Management Mode: Out-of-Band  
Interface: ipv4 192.168.1.1:22334

**Interfaces**  
0: rdma ipv4 192.168.1.1

**NSDevices**  
0: Device: ID -1

**Port IDs** +  
1: rdma ipv4 192.168.1.1:4420

**Subsystems** +  
Subsystem NQN: SUBSYS1 (Allow Any Host)

**NSIDs** +

An "edit" button is visible next to the Port ID entry. A text box with a blue border is overlaid on the interface, containing the text: "Once the Interfaces have be retrieved, setting of the Port ID will be different from the prior screen".

version 0.1

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management". A modal dialog titled "Update Port ID 1 on Target 'TARGET\_1'" is open, containing the following fields:

- Port ID:
- Interface:
- TRSVCID:

A text box with a blue border is overlaid on the interface field, containing the text: "Instead of providing the Transport Type, Address and Family, the Interface must be selected from the list".

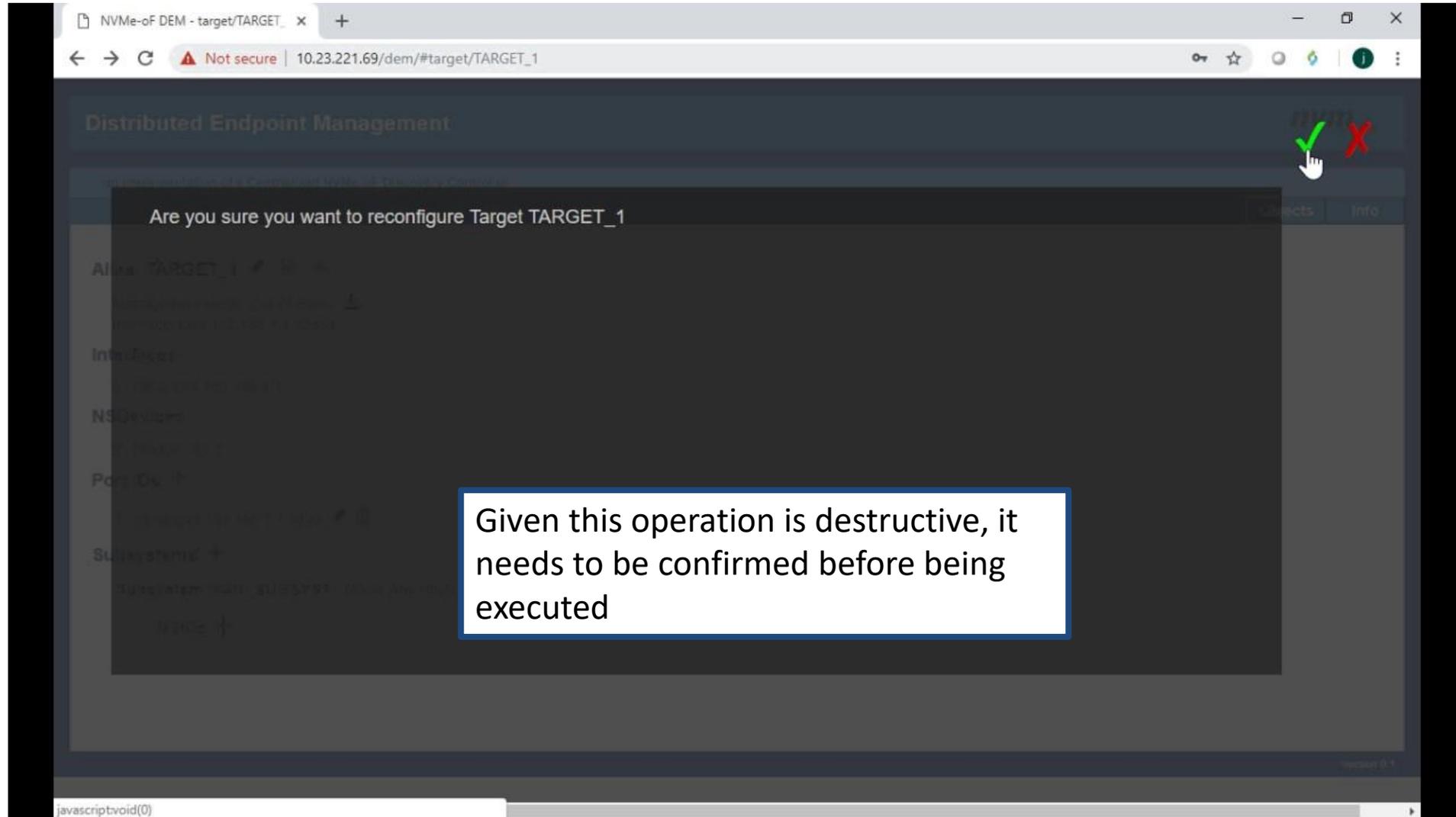
# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it includes the `nvm` logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". The main content area displays configuration details for "Alias: TARGET\_1". A mouse cursor is hovering over a "reconfigure" button next to the "Management Mode: Out-of-Band" field. The configuration includes:

- Management Mode: Out-of-Band
- Interface: ipv4 192.168.1.1:22334
- Interfaces: 0: rdma ipv4 192.168.1.1
- NSDevices: 0: Device: ID -1
- Port IDs: 1: rdma ipv4 192.168.1.1:4420
- Subsystems: Subsystem NQN: SUBSYS1 (Allow Any Host)
- NSIDs

A text box with a blue border contains the text: "Once a configuration is created, it can be pushed to the Target to replace any existing configuration".

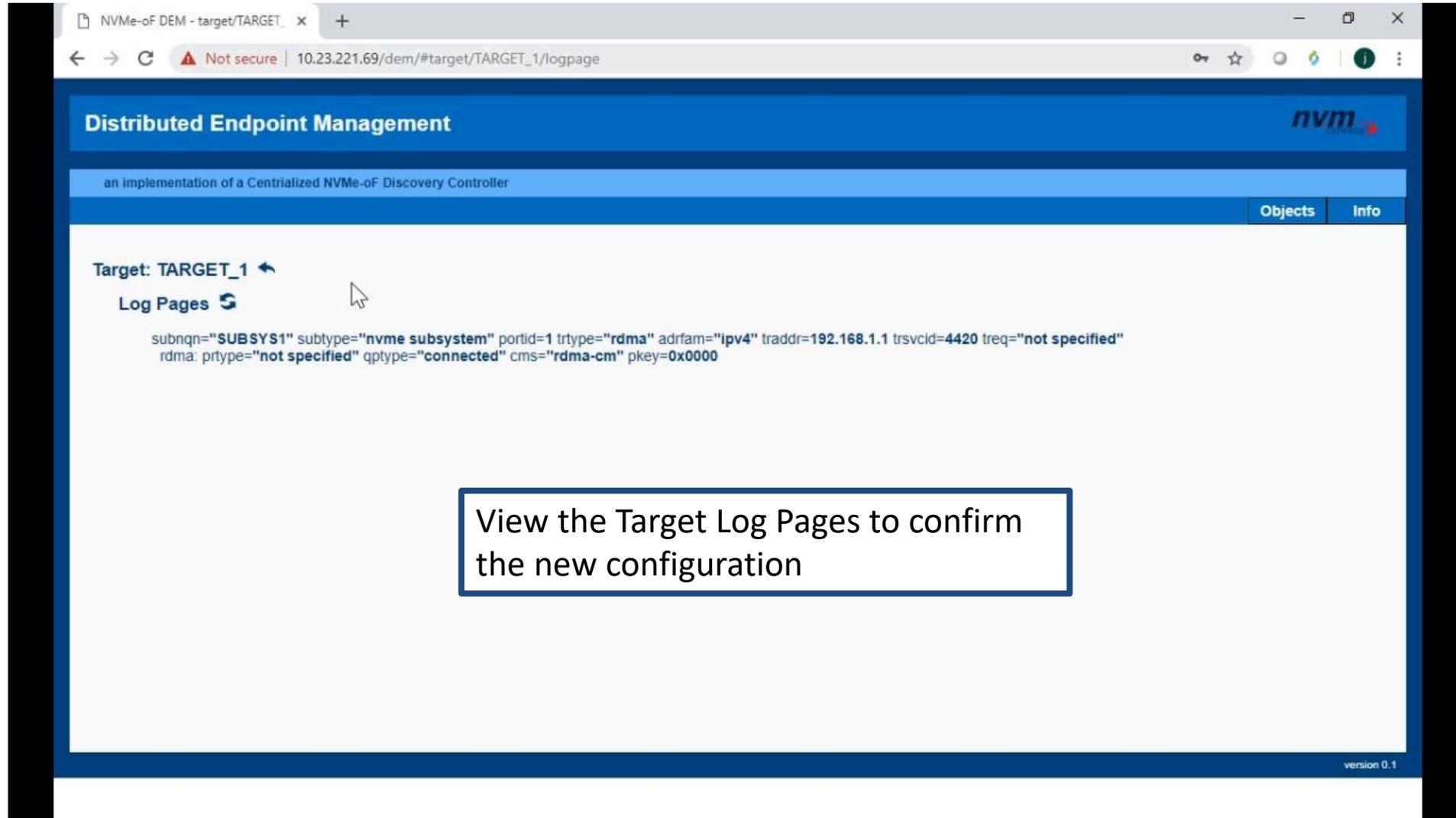
# DEMONSTRATION



# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it includes the `nvm` logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". The main content area displays configuration details for "Alias: TARGET\_1". A mouse cursor is hovering over a button labeled "view log pages". Other configuration items include "Management Mode: Out-", "Interface: ipv4 192.168.1...", "Interfaces" (0: rdma ipv4 192.168.1.1), "NSDevices" (0: Device: ID -1), "Port IDs" (1: rdma ipv4 192.168.1.1:4420), "Subsystems" (Subsystem NQN: SUBSYS1 (Allow Any Host)), and "NSIDs". A callout box with a blue border contains the text: "View the Target Log Pages to confirm the new configuration". The version number "version 0.1" is visible in the bottom right corner of the interface.

# DEMONSTRATION



The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1/logpage`. The page title is "Distributed Endpoint Management" and it includes the `nvm` logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". There are tabs for "Objects" and "Info". The main content area shows "Target: TARGET\_1" with a refresh icon and "Log Pages" with a refresh icon. Below this, the following configuration string is displayed: `subnqn="SUBSYS1" subtype="nvme subsystem" portid=1 trtype="rdma" adrfam="ipv4" traddr=192.168.1.1 trsvcid=4420 treq="not specified" rdma: prtype="not specified" qptype="connected" cms="rdma-cm" pkey=0x0000`. A callout box with a blue border contains the text: "View the Target Log Pages to confirm the new configuration". The version number "version 0.1" is visible in the bottom right corner of the interface.

# DEMONSTRATION

The screenshot shows a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it includes the `nvm` logo. Below the title, it states "an implementation of a Centralized NVMe-oF Discovery Controller". The main content area is divided into sections: "Alias: TARGET\_1", "Management Mode: Out-of-Band", "Interface: ipv4 192.168.1.1:22334", "Interfaces" (0: rdma ipv4 192.168.1.1), "NSDevices" (0: Device: ID -1), "Port IDs" (1: rdma ipv4 192.168.1.1:4420), and "Subsystems" (Subsystem NQN: SUBSYS1 (Allow Any Host)). Under the Subsystems section, there is an "NSIDs" section with an "add" button. A callout box with a blue border contains the text: "To complete the NVMe-oF configuration, the Subsystem needs NS Devices via NSIDs".

# DEMONSTRATION

10.23.221.69/dem/#target/TARGET\_1

Distributed Endpoint Management

**Add a Namespace to Target 'TARGET\_1' Subsystem 'SUBSYS1'**

NS ID:

NS Device:

To link the NS Device, an NS ID is required as well as selecting a NS Device from the available list

# DEMONSTRATION

The screenshot displays a web browser window with the URL `10.23.221.69/dem/#target/TARGET_1`. The page title is "Distributed Endpoint Management" and it includes a sub-header "an implementation of a Centralized NVMe-oF Discovery Controller". The interface features a blue header with the "nvm" logo and navigation tabs for "Objects" and "Info". The main content area shows configuration details for "Alias: TARGET\_1", including "Management Mode: Out-of-Band" and "Interface: ipv4 192.168.1.1:22334". It lists "Interfaces" (rdma ipv4 192.168.1.1), "NSDevices" (Device: ID -1), "Port IDs" (rdma ipv4 192.168.1.1:4420), "Subsystems" (Subsystem NQN: SUBSYS1 (Allow Any Host)), and "NSIDs" (Device: ID -1). A text box with a blue border is overlaid on the right side of the configuration, containing the text "Finally, a complete Target configuration". The version "0.1" is visible in the bottom right corner of the interface.



OPENFABRICS  
ALLIANCE

# DEVELOPMENT OPPORTUNITIES AND WRAP-UP

# DISTRIBUTED ENDPOINT MANAGEMENT

## Development Opportunities

### Discovery controller + Management (dem)

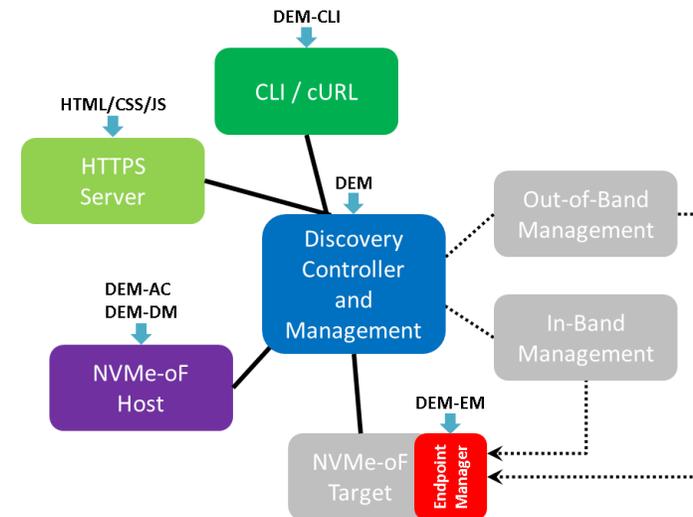
- Redundancy and Failover
- Generate Log Pages for Targets without a Discovery Controller
- Target usage monitoring
- Convert DEM to RedFish / SwordFish schema
- Convert DEM in-band configuration to current NVMe-MI Specification proposal

### Endpoint Manager (dem-em)

- Target usage monitoring
- Extend for other Targets

### Support:

- More NVMe-oF Transports (e.g., FC)
- Asymmetric Namespace Access Groups
- Namespace Subtypes/Partitions
- Transport Required (TREQ)
- Multipath



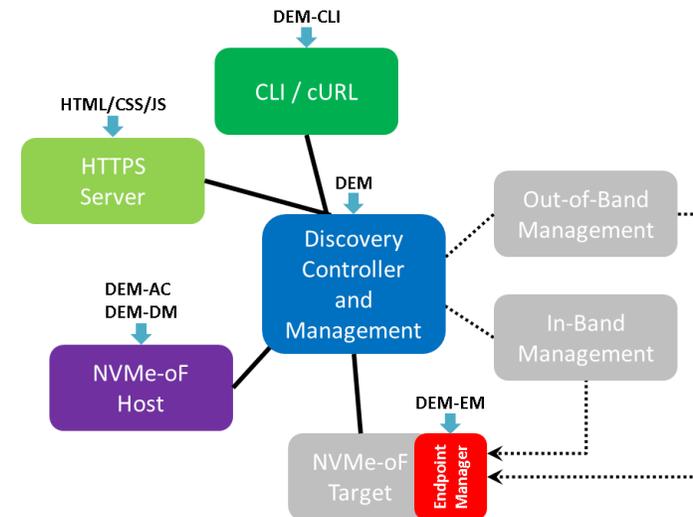
# DISTRIBUTED ENDPOINT MANAGEMENT

## Wrap-up

### Currently Available:

- As a Dual GPL / BSD licensed project
- On GitHub (<https://github.com/linux-nvme/nvme-dem>)
- Wiki (<https://github.com/linux-nvme/nvme-dem/wiki>)

**Clone  
Use  
Contribute**





OPENFABRICS  
ALLIANCE

15<sup>th</sup> ANNUAL WORKSHOP 2019

**THANK YOU**

Phil Cayton

Intel Corporation