SNIA®

Redfish® and Swordfish®:

The Data Source for Predictive Models and Decision Making

Chris Lionetti

HPE Senior Tech Marketing Engineer

Sept, 2024

Abstract

In today's rapidly changing datacenters, it is very difficult to visualize your current equipment utilization, much less be able to predict when and where to expect bottlenecks, issues and failures to occur. Particularly in a multi-vendor solution, DMTF Redfish® and SNIA Swordfish® can bridge the gap to provide the key instrumentation needed.

This presentation will provide an overview of techniques and examples to use Redfish/Swordfish instrumentation, such as metrics and counters, to populate graphs and create predictive models of Asset utilization, Power and Temp metrics, and Capacity and Performance statistics. The models can facilitate decision making, chargeback opportunities, infrastructure optimization, and refresh/upgrade cycles.

Agenda

- Basics of Redfish/Swordfish and Native Rest APIs
- Asset Utilization: finding orphans and the underutilized
- Asset Retirement/Replacement: Populating TCO calculations to make smart business decisions
- Active Optimization Automation

Disclaimer

- These processes and methods illustrated here increase in value as the size of your infrastructure increases. It represents a good amount of work and automation required to see these reductions.

SNIA®

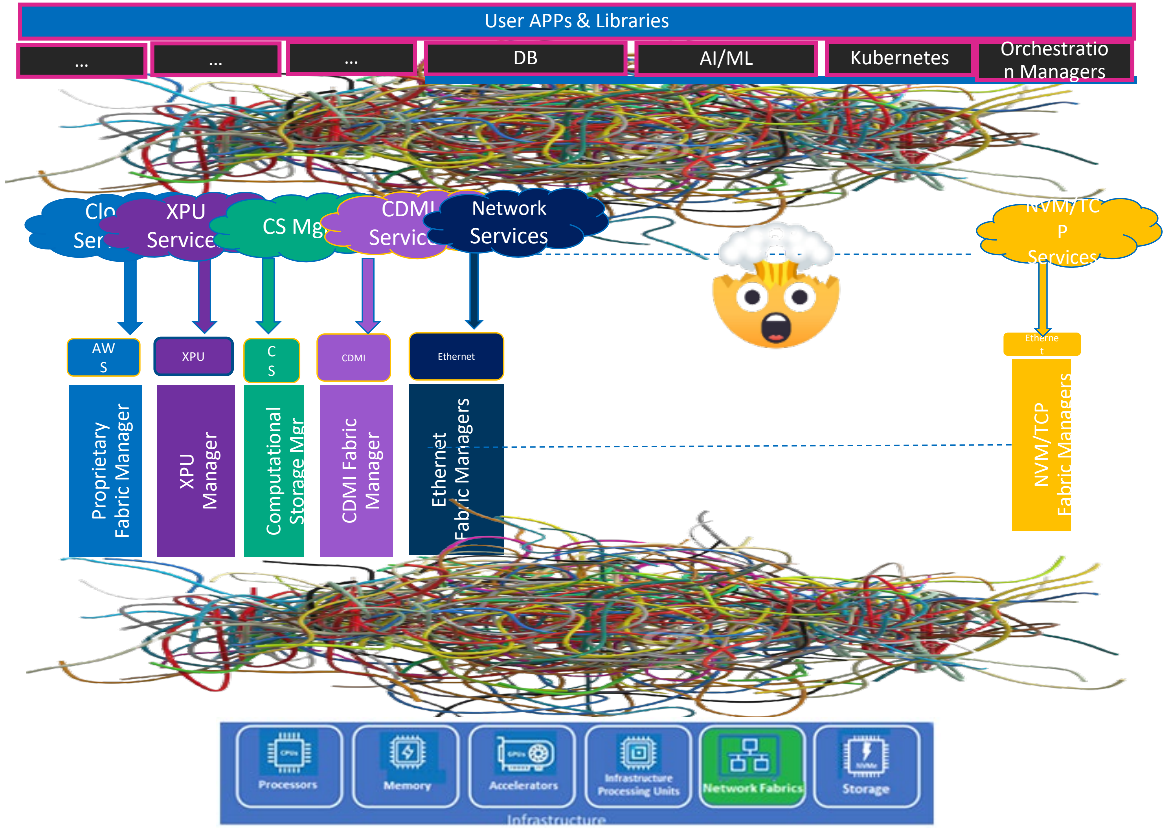
What came Before Redfish?

- IPMI (Intelligent Platform Management Agent)
  - First Published 1998
  - Last updated 2015
- SNMP (Simple Network Management Protocol)
  - Initial version 1990
  - Last Updated 2004 (v3)
- CIM (Common Information Model)
  - First Published 1999
  - Discontinued 2019
- SMI-S (Storage Mgmt Initiative – Specification)
  - Extension to CIM to cover Enterprise Storage
  - First Published 2002
  - Discontinued

SNIA

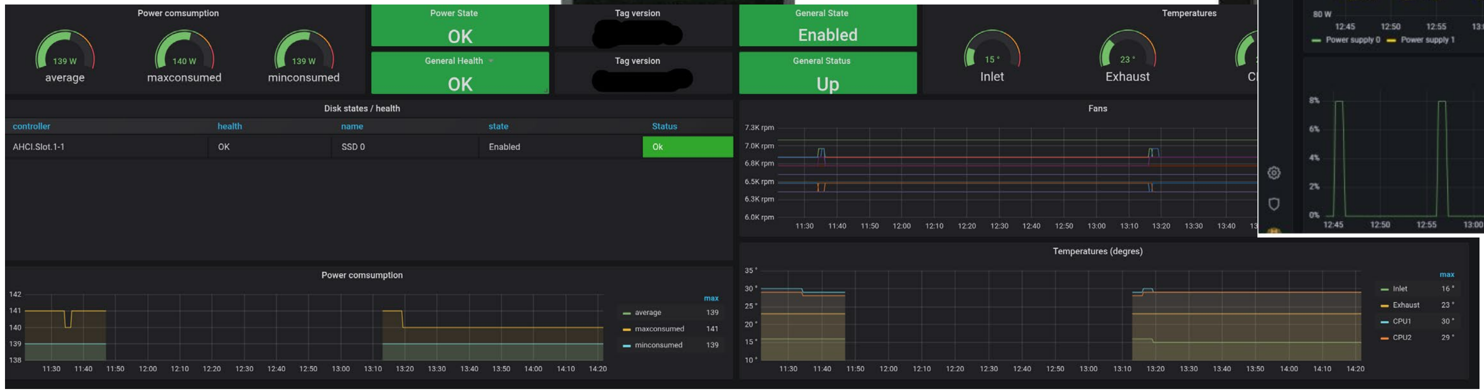
These Trends Lead to … a Manageability Headache

- Adding additional virtualization and acceleration technologies presents more flexibility and options to build a custom configuration but doing so increases the management challenges exponentially.
- Workload management and optimization is different for each type of technology, device, and vendor. Sometimes, even per version.
- Administrators are being asked to manage an increasingly heterogenous device and network infrastructure, each with its own management standard and model.
- Doing DevOps, instead of specialized management domains.
- With the proliferation of different technologies and multiple vendors, standards-based management is more critical than ever to successfully develop, integrate, deploy, and manage systems and storage at scale



See all the Things: Visibility of your Infrastructure

- Hardware: Servers → Redfish RestAPI
  - Close to 100% of Industry Standard Server
    - If server has a BMC/iDrac/ILO it likely supp
    - If your server is newer than 4 Years old, it li
  - Can support ALL classes of servers
    - smallest MicroServers to the largest Superc



See https://www.hpe.com/psnow/doc/a00134351enw : Datacenter monitoring using Redfish

SNIA

See all the Things: Visibility of your Infrastructure

- Hardware: Storage → Swordfish RestAPI
  - Some adoption; can write common codebase among multiple vendors
    - In the case of the HPE MSA, since this device supports Swordfish natively, you can extract information directly either via the RestAPI, or a simplified CLI via powershell
      - https://github.com/SNIA/Swordfish-Powershell-Toolkit
    - You can also see an example of a Swordfish Implementation where the interface was offered as a emulator as well
      - https://github.com/Cray-HPE/csm-redfish-interface-emulator/blob/master/setup.sh
  - When Swordfish Implementation doesn't exist
    - consider writing a converter based on the opensource existing Swordfish API Emulator (Python)
      - https://github.com/SNIA/Swordfish-API-Emulator
    - Example code base to convert HPE Nimble Storage / Alletra 6000 native RestAPI to Swordfish (PowerShell)
      - https://github.com/chris-lionetti/SwordfishMockup
  - As last resort; you can write against the native vendor API
    - Example of a well-defined API for which data can be normalized as inserted into a DB

      This example is used to retrieve data directly from HPE Alletra 9000 and Primera type arrays
      - https://developer.hpe.com/blog/new-powershell-toolkit-available-for-managing-hpe-data-services-cloud-console/

SNIA

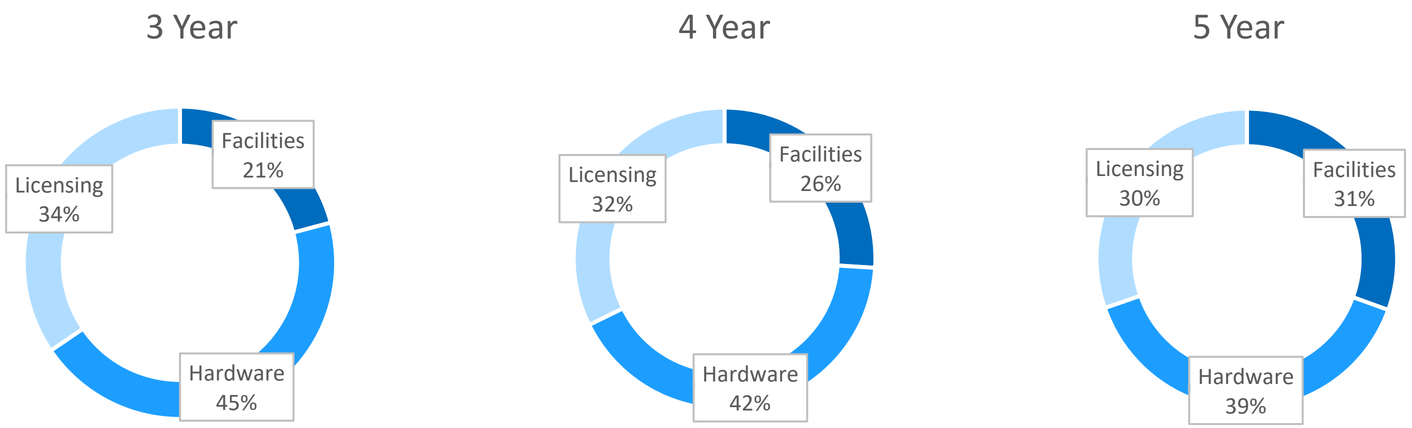
See all the Things.

- This first step is visibility of your infrastructure
  - Power Distribution
    - A number of major PDU vendors support Redfish
      - https://help.raritan.com/json-rpc/4.2.10/Redfish_8idl_source.html
  - Operating Systems
    - Windows Failover Clustering : Inbox Module
    - Performance Counters: Inbox Module
  - Hypervisors
    - VMware PowerCLI : https://www.powershel
    - Hyper-V: Inbox Module (Hyper-V 2.0.0.

SNIA

Where does the money go?

- Cost of the hardware is the obvious answer
  - Below examples take sample 32 Core System from CDW
    - Adding significant NVMe and/or GPUs can drive this price up drastically.
- Cost of Licensing for this System is
  - Below examples assume either RHEL or Windows Server perpetual license
    - Adding either Databases, Mail, other licensed application software drives this price up drastically
    - Adding an additional Hypervisor may also drive this price up drastically
- Cost of Power + UPS Services + RackU Facilities
  - Below example assumes lowest commercial power costs in nation ($0.10/KwH),
    - Deploying to higher cost locations can increase these costs drastically ($0.40/KwH)



SNIA

Running Benchmark

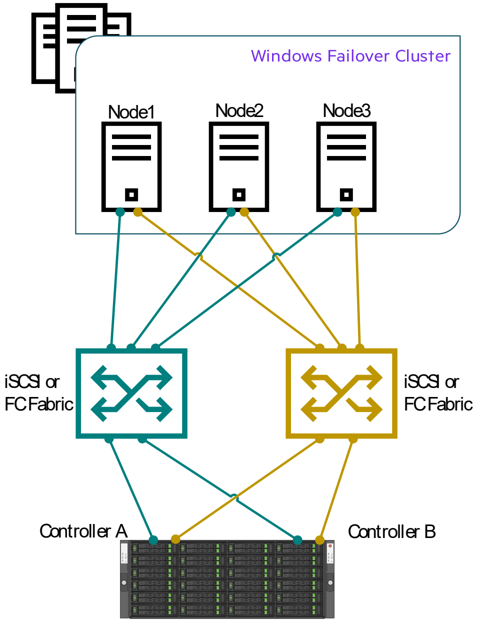
- Can help you determine your Workload
  - To simulate a VM Workload; consider tools
    - Max out the Network, CPU, Disk, and Memory
      https://techcommunity.microsoft.com/t5/azure-stack-blog/vmflee
    - Measure Power/Cooling/etc during Idle, and at max load.
    - Should be able to generate the following metrics
      - IOPS per Core or Watt, GB/s per Core or Watt
    - Ensure that every core is engaged for best ratings
      - A server running idle gets bad efficiency.
    - Can evaluate Previous Generation vs Current Generation
      - Surprising effect. 2 Similar systems a generation apart
        - Newer Gen = less port at idle, and less power under sa

SNIA

Using what we have learned to Deploy a VM Cluster

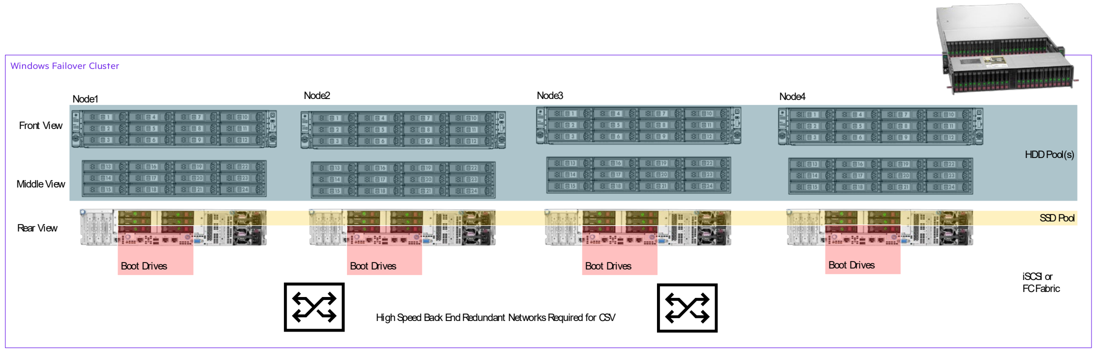
You can choose SAN

- Can be implemented as Triple Parity. i.e. 24 Drives = 20 Drives worth of Capacity
- Can Tier to Flash
- Lower Wattage as servers don't house Storage. Different Scale Pattern



You can choose SDS (Software Defined Storage)

- Commonly uses 3-way Mirroring, i.e. 60 Drives = 20 Drives worth of capacity
- Can Tier to Flash or be all flash
- Doesn't Need Storage Array



Optimizing a SAN Based Failover Cluster

Saving Power

- Use Anti-Affinity rules to ensure separation of VMs that need to exist on different nodes
- Heavily Load a few cluster Servers, leave some nodes evacuated. Allows you to shut down unneeded loads
- (N/2)+1 nodes running in cluster to maintain redundancy and quorum, Off nodes can be booted on demand.
- Optimal regular needs should equate to N/2 Nodes, peak demand should equate to N Nodes.
  - If peek exceeds N Nodes add more nodes live.
  - If optimal needs lower than N/2, remove nodes live

Saving License Costs

- Purchase enough perpetual Licenses for the Optimal Load
- Purchase Pay-as-you-go licenses for on-demand nodes.

Saving Hardware Costs

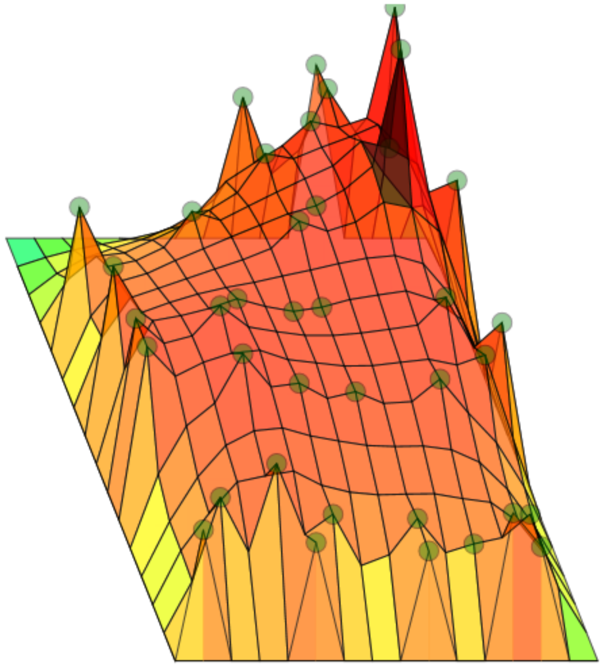
- Consider Purchasing replacement Nodes when TCO calculations justify it
- Consider Purchase for Optimal Nodes
- Consider Lease for Demand Nodes

SNIA

Optimize Environmental

Every Watt towards a System becomes a Watt of Heat

- Devices have MANY tem
  - You should be able to mo
- Devices report current al
  - If workload remains const
  - Most devices have compl
  - Ensure that each device i
- Enabler for air-side econ
  - Consider monitoring exte
  - Can predict future weathe



Front of server

04-P1 DIMM 1-6	Memory	8	4	OK	99F	Caution: 189F; Critical: N/A
05-P1 DIMM 7-12	Memory	13	5	OK	102F	Caution: 189F; Critical: N/A
06-P2 DIMM 1-6	Memory	1	5	OK	106F	Caution: 189F; Critical: N/A
07-P2 DIMM 7-12	Memory	6	4	OK	102F	Caution: 189F; Critical: N/A
08-P1 Mem Zone	Memory	8	7	OK	99F	Caution: 158F; Critical: 167F
09-P1 Mem Zone	Memory	14	6	OK	104F	Caution: 158F; Critical: 167F
10-P2 Mem Zone	Memory	1	6	OK	106F	Caution: 158F; Critical: 167F
11-P2 Mem Zone	Memory	7	7	OK	97F	Caution: 158F; Critical: 167F
12-HD Max	System	12	0	OK	95F	Caution: 140F; Critical: N/A
13-Chipset 1	System	8	9	OK	111F	Caution: 221F; Critical: N/A
14-Chipset1 Zone	System	9	10	OK	106F	Caution: 158F; Critical: 167F
15-P/S 1 Inlet	Power Supply	1	11	OK	93F	Caution: N/A; Critical: N/A
16-P/S 1 Zone	Power Supply	1	8	OK	104F	Caution: 158F; Critical: 167F
17-P/S 2 Inlet	Power Supply	5	11	OK	91F	Caution: N/A; Critical: N/A
18-P/S 2 Zone	Power Supply	5	7	OK	102F	Caution: 149F; Critical: 158F
19-PCI #1	I/O Board	11	13	Not installed	N/A	N/A
20-PCI #2	I/O Board	8	13	Not installed	N/A	N/A
21-VR P1	Power Supply	11	1	OK	97F	Caution: 239F; Critical: 248F
22-VR P2	Power Supply	4	1	OK	104F	Caution: 239F; Critical: 248F
23-VR P1 Mem	Power Supply	9	1	OK	81F	Caution: 239F; Critical: 248F
24-VR P1 Mem	Power Supply	13	1	OK	82F	Caution: 239F; Critical: 248F
25-VR P2 Mem	Power Supply	2	1	OK	90F	Caution: 239F; Critical: 248F
26-VR P2 Mem	Power Supply	6	1	OK	82F	Caution: 239F; Critical: 248F
27-VR P1Mem Zone	Power Supply	9	0	OK	79F	Caution: 158F; Critical: 167F
28-VR P1Mem Zone	Power Supply	13	0	OK	81F	Caution: 158F; Critical: 167F
29-VR P2Mem Zone	Power Supply	1	0	OK	88F	Caution: 158F; Critical: 167F
30-VR P2Mem Zone	Power Supply	5	0	OK	82F	Caution: 158F; Critical: 167F
31-HD Controller	System	12	10	OK	172F	Caution: 221F; Critical: N/A
32-HD Cntlr Zone	System	12	11	OK	118F	Caution: 149F; Critical: 158F
33-PCI 1 Zone	System	8	12	OK	104F	Caution: 158F; Critical: 167F
34-PCI 1 Zone	System	10	13	OK	106F	Caution: 151F; Critical: 160F
35-LOM Card	I/O Board	13	13	OK	154F	Caution: 212F; Critical: N/A

Great secret: You don't need to ensu
exceed its critical threshold that force

SNIA

So how do I do this

- We have already covered how to obtain Hardware info and control that hardware
  1. To automate/script control of VMWare: See ‘Move-VM –VM %VmName%’
  2. To automate/script control of Hyper-V: See ‘Suspend-ClusterNode –name ‘Node’ –drain’
  3. To physically shut off or reboot a server, us the SNIA/Swordfish Toolkit and the following commands

     Invoke-RedfishSystemReset –ResetType { GracefulShutdown | On }

What Next

- This modeled data can be used to generate TCO models
- Can also be used to create chargeback models for internal consumers
- This same model information can be used to populate your carbon footprint values
- This also allows customers to run meaningful bake-off evaluations of new hardware

SNIA

Chris Lionetti

SNIA Board Member & SNIA SMI Vice-Chair

Chris.Lionetti@hpe.com

Why does this matter to me? Because I LIVE IT!

- I built a Micro-Da e.
  - ed Floor,
- for Fr
  - old air fro
  - d Fall: Pu
  - Pull air fr
  - sac

This slide is just extra credit…. ☺

SNIA