perfSONAR

Embedded Network Systems

Juniper Research & Education Networks Symposium (J-RENS)

October 10, 2018

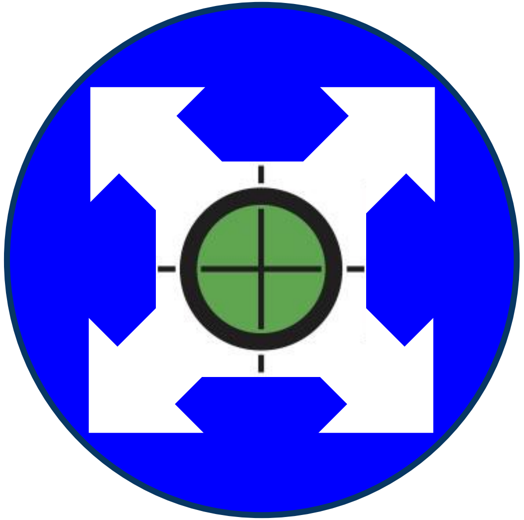
Eric Boyd <ericboyd@umich.edu>, Ed Colone <epcjr@umich.edu>

INFORMATION AND TECHNOLOGY SERVICES

UNIVERSITY OF MICHIGAN

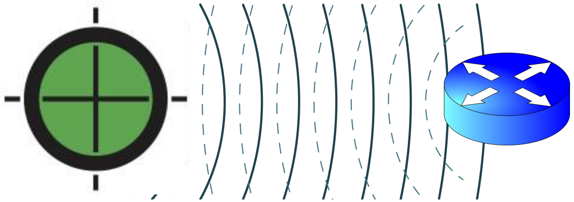
perfSONAR Embedded Systems

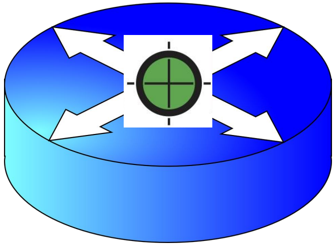
1. Background
2. Standards·Principles·Requirements
3. Vision and End Goal
4. Juniper Collaboration
5. Advice for Next Generation Appliances



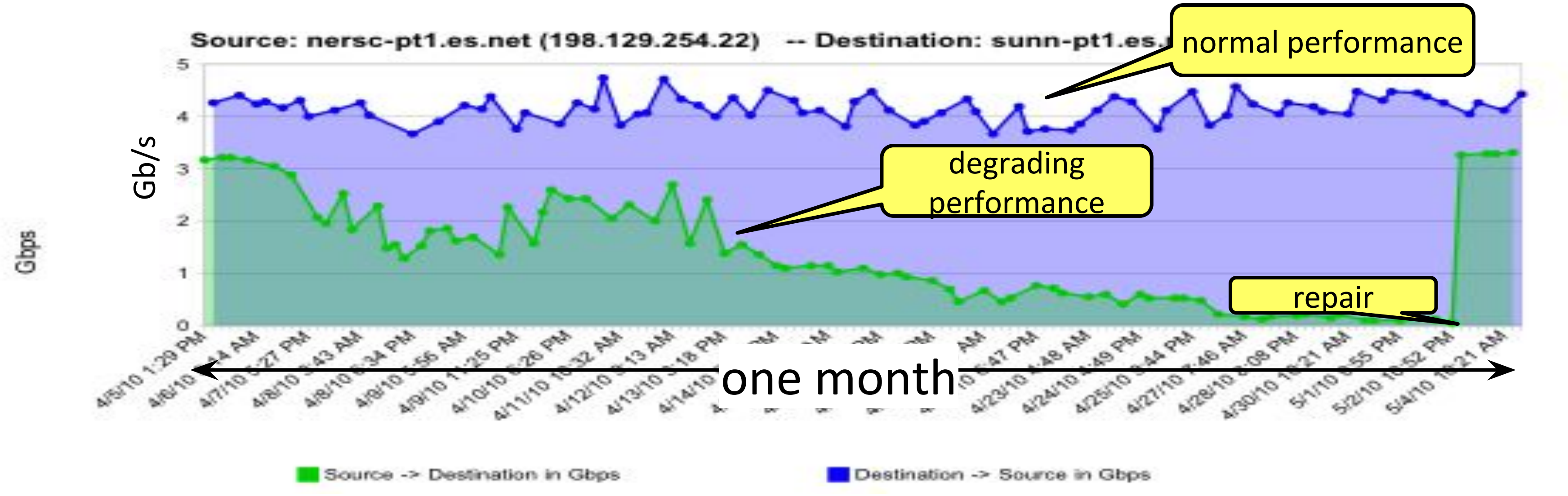
Background: Emerging Technology

- perfSONAR: over 2,000 deployments
- embedded systems: control many devices in common use today
- *NEW: next generation of the network gear - perfSONAR embedded in major network vendor hardware*





Sample Soft Failure: failing optics



Credit: Jason Zurawski, ESNet

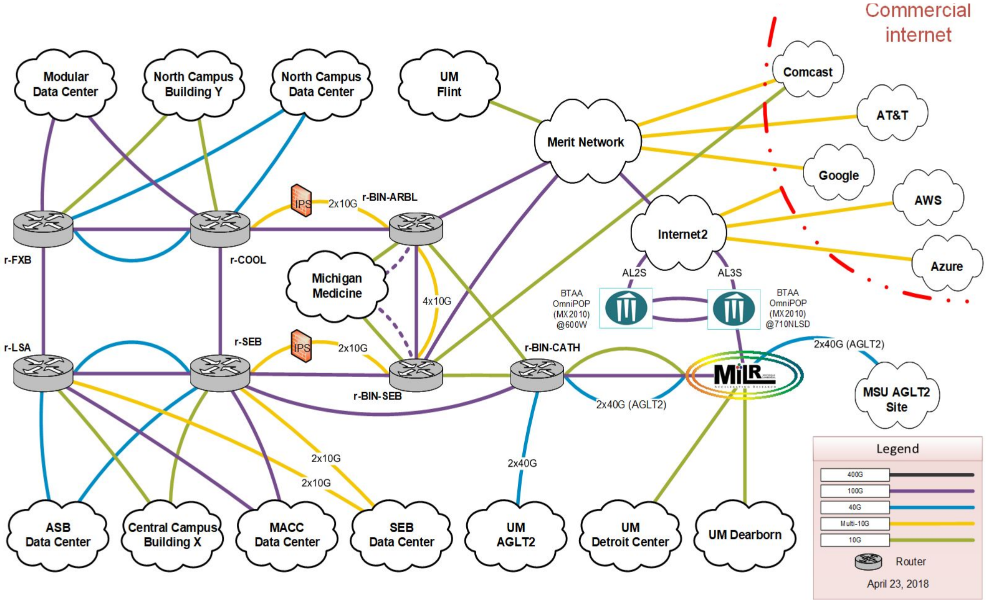
perfSONAR Embedded Systems

Background: Problem of Scale

The ability to benchmark and troubleshoot the “last mile” of large distributed networks.

U-M current scope:

- Number of AL switches: 101 models, 3,825 devices
- Number of DL switches: 22 models, 239 devices



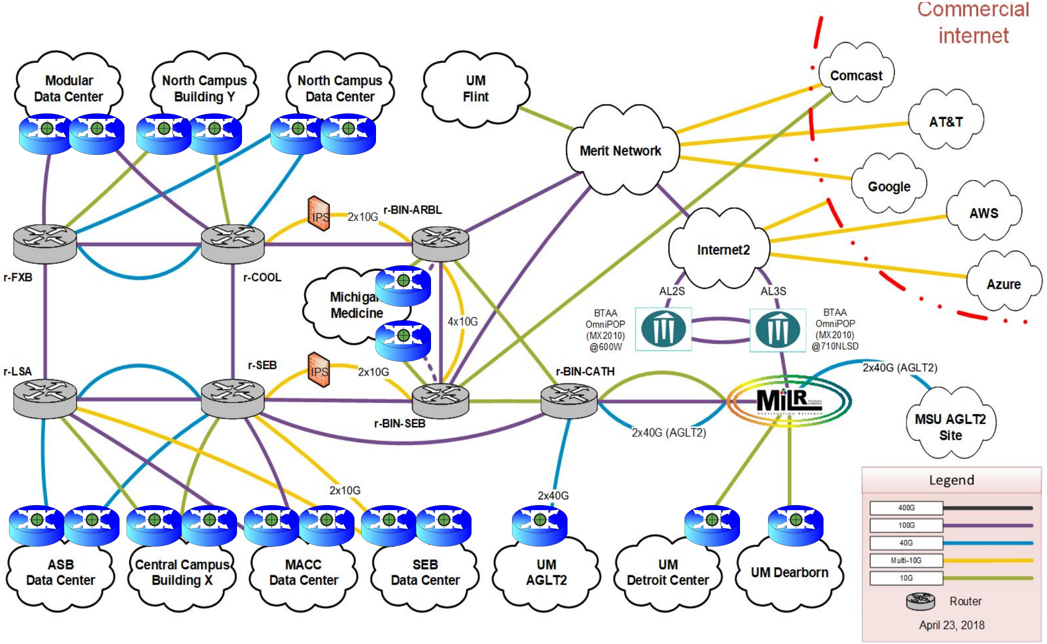
INFORMATION AND TECHNOLOGY SERVICES UNIVERSITY OF MICHIGAN · ESnet ENERGY SCIENCES NETWORK · GÉANT · INDIANA UNIVERSITY · INTERNET2 · UNIVERSITY OF MICHIGAN

perfSONAR Embedded Systems

Emerging Technology Vision

Provision perfSONAR on the network gear itself.

- Scalable
  - Deploy
  - Support
  - Grow
- Fewer costs/ fewer machines
- Fewer dependencies
- Becomes a part of the in-building network replacement lifecycle



INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN

ESnet ENERGY SCIENCES NETWORK

GÉANT

INDIANA UNIVERSITY

INTERNET2

UNIVERSITY OF MICHIGAN

perfSONAR Embedded Systems

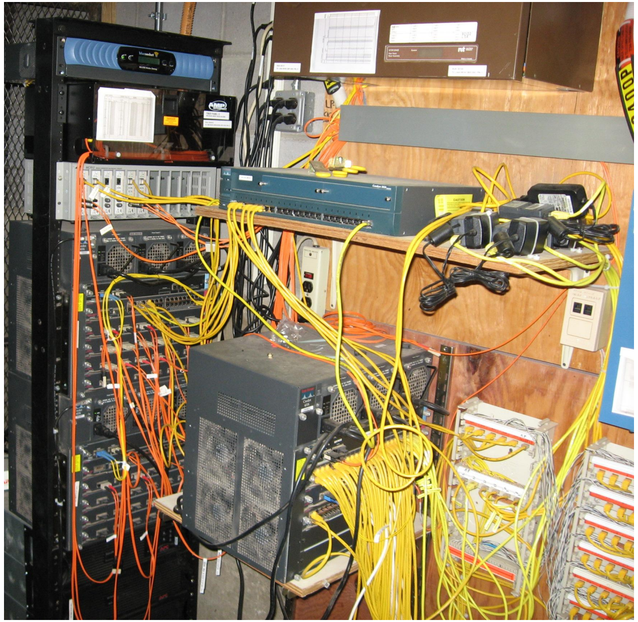
Background: Challenge

Legacy location issues:

- Space
- Power
- Climate control

On-site support challenges:

- Too many locations (>1,000)
- Impossible to provide 24x7 operations
- Servers with management planes more expensive



INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN

ESnet
ENERGY SCIENCES NETWORK

GÉANT

INDIANA UNIVERSITY

INTERNET2

UNIVERSITY OF MICHIGAN

Standards, Principles, Requirements

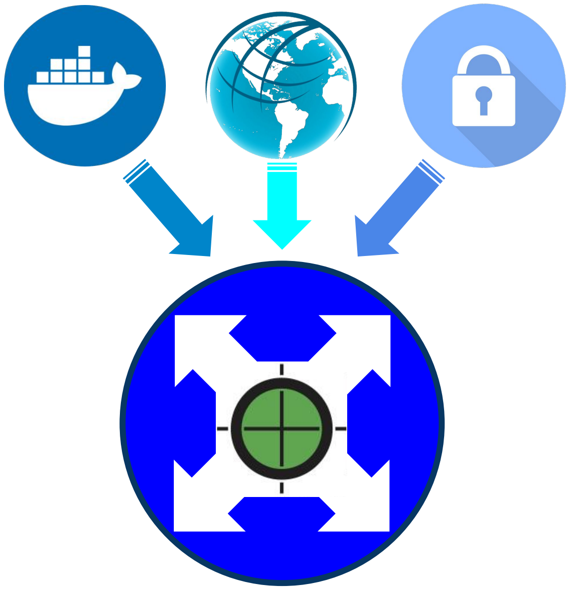
Embedded Solution vs. a stand-alone perfSONAR node:

- **Don't gouge:** cost of the Embedded solution must not be greater than a discreet solution
- **Don't suck:** Network hardware must perform switching and routing duties unimpeded.
- **Don't lie**: performance measurements must be as accurate as discreet servers.

INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN
ESnet ENERGY SCIENCES NETWORK
GÉANT
INDIANA UNIVERSITY
INTERNET2
UNIVERSITY OF MICHIGAN

Obstacles: Net Admin Perspective

- Multiple proficiencies required to configure
- Complicated container creation
- Network Architecture
  - Container requires public IP address + small subnet per building
  - Often rebooting is not an option
  - “Switch First” philosophy
- Security Architecture
  - Non-network admins require administrative access to the switch
  - How well is the container separated from control plane
  - What other security issues does this present...



Embedded Platform Options

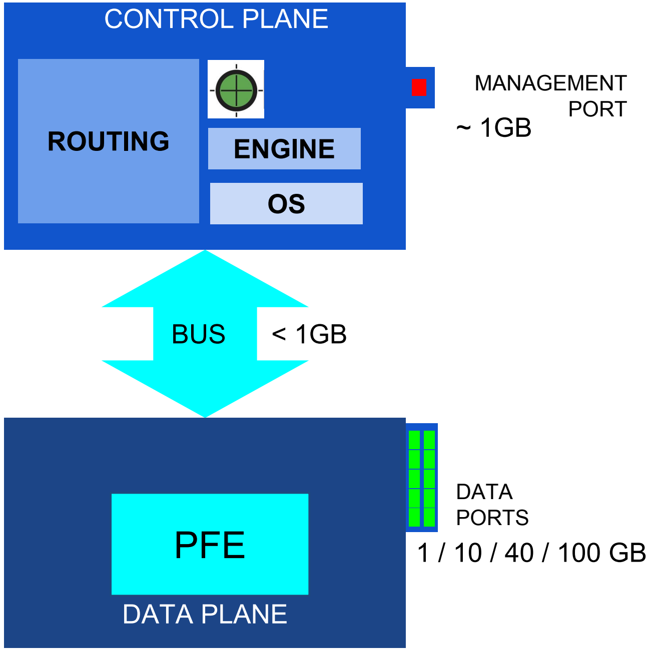
- "Bare Metal" port (Yocto / Wind River)
  - High effort - Embedded Linux not a dev environment
  - Unsupportable
- KVM
  - Poor performance
  - Larger containers
- Docker, LXC
  - Better performance
  - Standards for creation
  - Some existing support (perfSONAR on Docker)
- Proprietary container process
  - Start with base image/container (Docker, LXC, etc.)
  - More complex provisioning process
  - Security concerns may require perfSONAR redesign / development

perfSONAR Embedded Systems

Emerging Technology Architecture

Current throughput constraints:

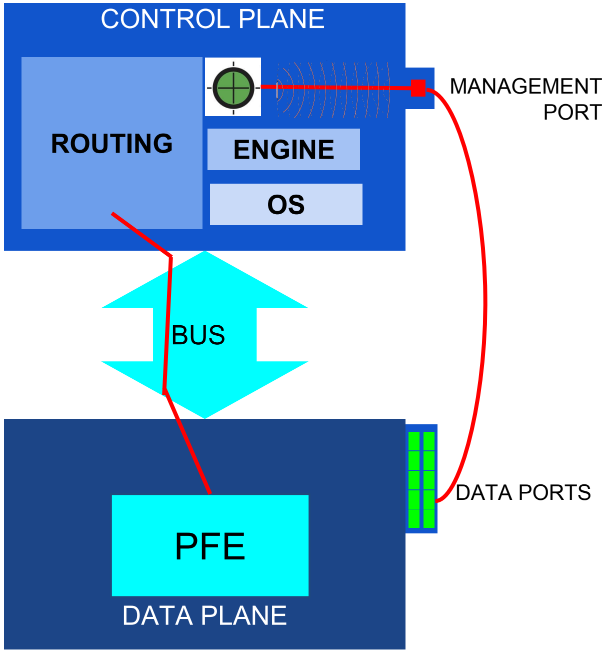
- Desired performance
  - Max data port rate (10GB / 40GB / 100GB / etc.)
- Management port constraints
  - ~ 1 GB (Sometimes limited to copper interfaces)
- Bus performance
  - Currently < 1 GB



INFORMATION AND TECHNOLOGY SERVICES UNIVERSITY OF MICHIGAN · ESnet ENERGY SCIENCES NETWORK · GÉANT · INDIANA UNIVERSITY · INTERNET2 · UNIVERSITY OF MICHIGAN

Container Networking: Management Port

- Currently has most throughput performance (~1GBE)
- If Management port limited to copper, could be incompatible with the rest of the switch, bad for “leaf nodes”
- Least desirable from Net Admin perspective
  - Precludes Out-Of-Band Management
- Cable required
  - Outage risk from possible misidentification as a network loop
  - Adds installation complexity

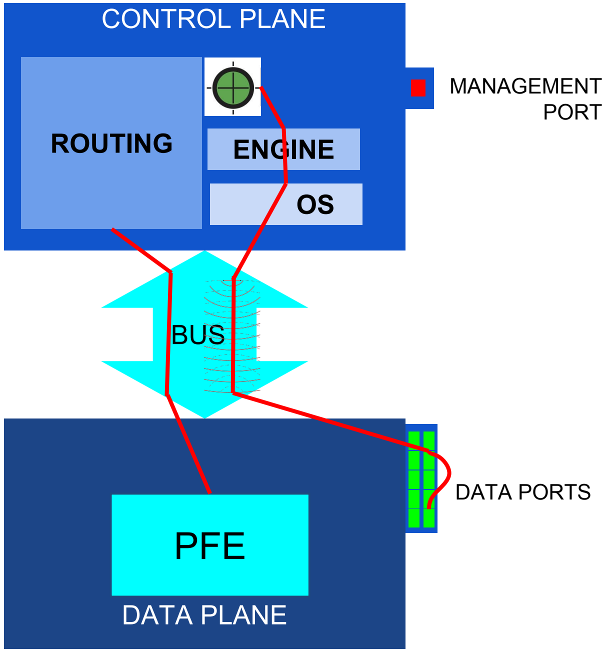


INFORMATION AND TECHNOLOGY SERVICES UNIVERSITY OF MICHIGAN
ESnet ENERGY SCIENCES NETWORK
GÉANT
INDIANA UNIVERSITY
INTERNET2
UNIVERSITY OF MICHIGAN

perfSONAR Embedded Systems

Container Networking: Single Data Port

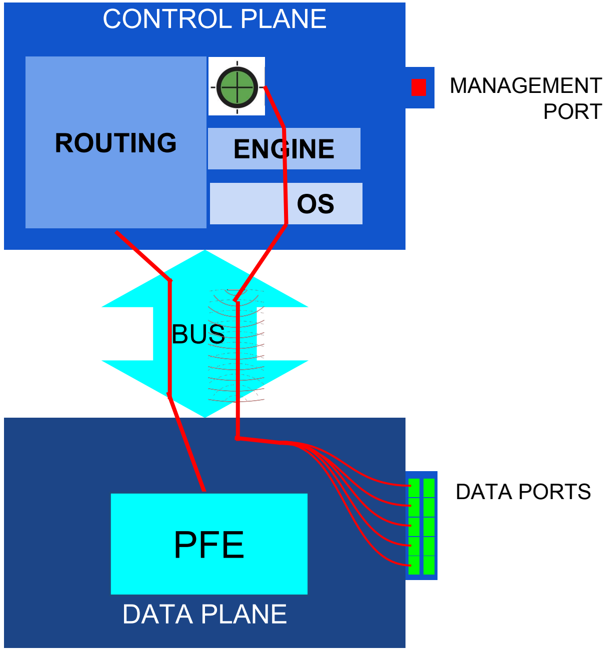
- Limited throughput performance (<1GBE)
- Bus is shared with routing duties
  - Unknown failure modes
- Cable required
  - Outage risk from possible misidentification as a network loop
  - Adds installation complexity



INFORMATION AND TECHNOLOGY SERVICES UNIVERSITY OF MICHIGAN · ESnet ENERGY SCIENCES NETWORK · GÉANT · INDIANA UNIVERSITY · INTERNET2 · UNIVERSITY OF MICHIGAN

Container Networking: Any Port

- Any data port "cableless"
  - Most desirable
  - Still goes over internal bus
- Bus is shared with routing duties
  - Unknown failure modes
- Ideally, pS container would be bridged to SVI
  - zero touch deployment
  - does not have to consume any ports on the front



INFORMATION AND TECHNOLOGY SERVICES UNIVERSITY OF MICHIGAN · ESnet ENERGY SCIENCES NETWORK · GÉANT · INDIANA UNIVERSITY · INTERNET2 · UNIVERSITY OF MICHIGAN

Obstacles: Hardware Limitations

- High performance platforms, no budget solution yet
- Discreet solution currently has some performance advantages
- Historically switches didn’t require a lot of storage:
    - Storage quickly becomes an issue with the larger images required to run
    - Typical flash might not have the space
    - External USB storage a possibility
        - Speed / bandwidth / performance concerns
- Shared bus = shared bandwidth

INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN
ESnet ENERGY SCIENCES NETWORK
GÉANT
INDIANA UNIVERSITY
INTERNET2
UNIVERSITY OF MICHIGAN

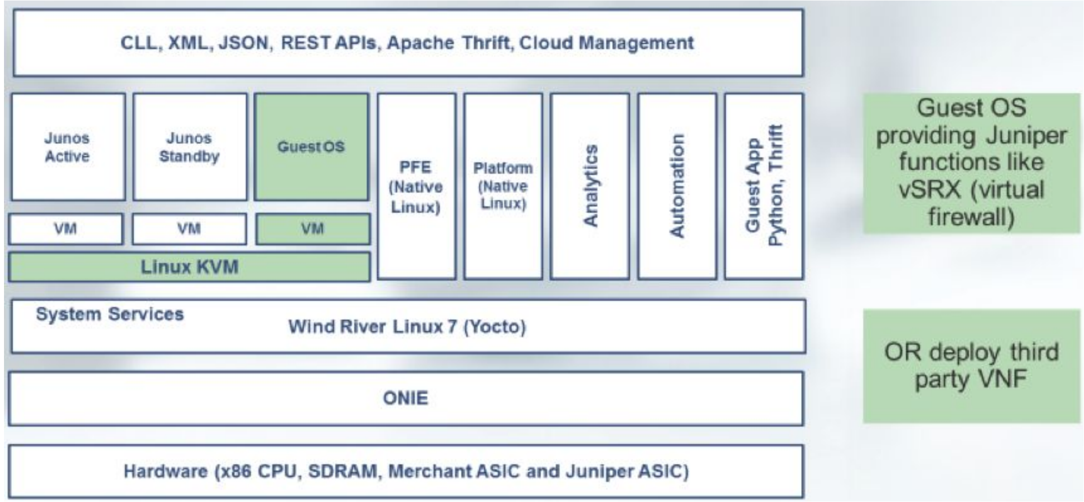
perfSONAR Embedded Systems

Juniper History

- perfSONAR demo at SC17
  - Ran on KVM
  - Management Port Networking
- Non-EVO experiments (2018)
  - Learning experience
  - Container Networking

    Limitations prevented

    Successful pilot program



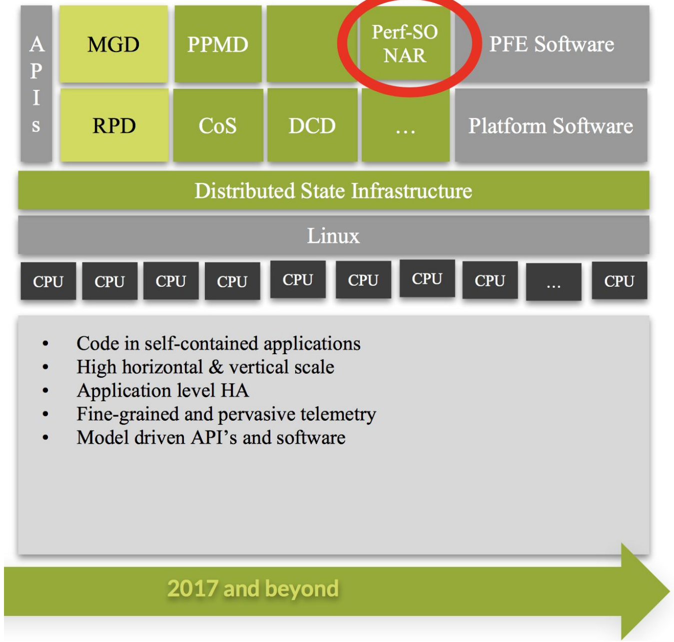
INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN
ESnet ENERGY SCIENCES NETWORK
GÉANT
INDIANA UNIVERSITY
INTERNET2
UNIVERSITY OF MICHIGAN

perfSONAR Embedded Systems

Juniper EVO + perfSONAR

Current Issues

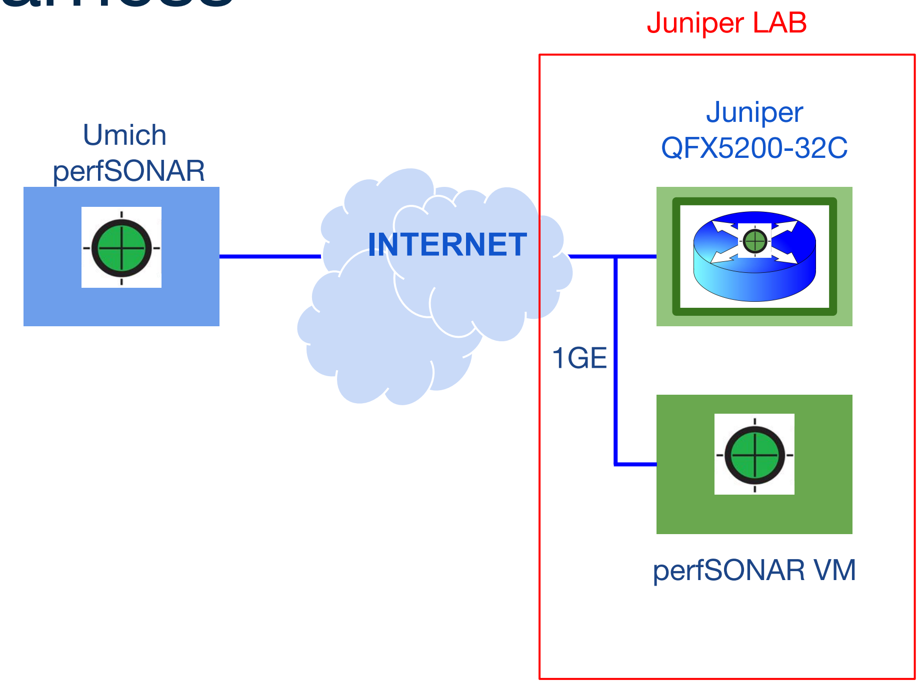
- Complicated container process
  - LXC with distrobuilder, chroot, no systemd, etc.
  - Currently Ubuntu only
  - Privileged vs unprivileged container security model
  - Requires perfSONAR REST interface re-development
- EVO on 5200
  - Experimental Release, Only in Juniper's Lab



INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN
ESnet ENERGY SCIENCES NETWORK
GÉANT
INDIANA UNIVERSITY
INTERNET2
UNIVERSITY OF MICHIGAN

Test Harness

- perfSONAR pscheduler was able to communicate from 5200 to UM over port 80
- Firewall issues prevented further perfSONAR testing to UM
- Proof-of-concept throughput and latency testing done in Juniper's lab



INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN
ESnet ENERGY SCIENCES NETWORK
GÉANT
INDIANA UNIVERSITY
INTERNET2
UNIVERSITY OF MICHIGAN

perfSONAR Embedded Systems

Performance Results: RTT

- 1GE Lab connection
- Expected low latency

1	CentOS-TP.juniper.net	64 Bytes	TTL 64	RTT	0.3580 ms
2	CentOS-TP.juniper.net	64 Bytes	TTL 64	RTT	0.3720 ms
3	CentOS-TP.juniper.net	64 Bytes	TTL 64	RTT	0.4640 ms
4	CentOS-TP.juniper.net	64 Bytes	TTL 64	RTT	0.3770 ms
5	CentOS-TP.juniper.net	64 Bytes	TTL 64	RTT	0.4740 ms

Performance Results: Latency

- Useful Latency statistics
  - 1GE Lab connection
  - Expected low latency scenario

```
One-way Latency Statistics
--------------------------
Delay Median ......... -1.97 ms
Delay Minimum ........ -2.07 ms
Delay Maximum ........ -1.84 ms
Delay Mean ........... -1.97 ms
Delay Mode ........... -1.96 ms -1.95 ms
Delay 25th Percentile ... -2.00 ms
Delay 75th Percentile ... -1.95 ms
Delay 95th Percentile ... -1.89 ms
Max Clock Error ...... 4.84 ms
Common Jitter Measurements:
    P95 - P50 ........ 0.08 ms
    P75 - P25 ........ 0.05 ms
    Variance ......... 0.00 ms
    Std Deviation .... 0.04 ms
```

perfSONAR Embedded Systems

Performance Results: Throughput

- Useful 1GE performance
  - 943.54 Mbps
  - 1GE copper interface
  - Potentially useable for 1GE link throughput analysis

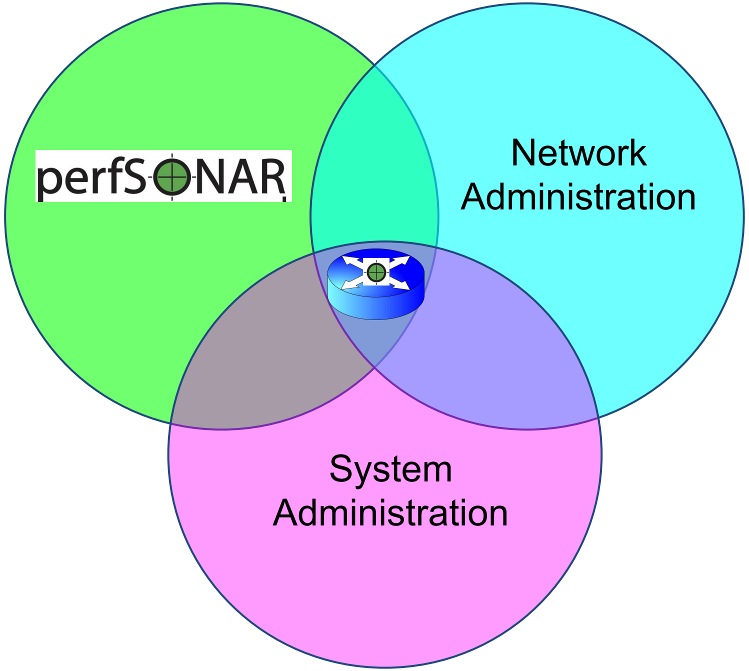
```
* Stream ID 6
Interval       Throughput     Retransmits    Current Window
0.0 - 1.0      953.64 Mbps    0              377.93 KBytes
1.0 - 2.0      942.48 Mbps    0              395.30 KBytes
2.0 - 3.0      941.44 Mbps    0              414.13 KBytes
3.0 - 4.0      950.62 Mbps    0              474.94 KBytes
4.0 - 5.0      940.56 Mbps    0              498.11 KBytes
5.0 - 6.0      941.93 Mbps    0              498.11 KBytes
6.0 - 7.0      935.20 Mbps    0              524.18 KBytes
7.0 - 8.0      943.51 Mbps    0              524.18 KBytes
8.0 - 9.0      944.59 Mbps    0              524.18 KBytes
9.0 - 10.0     941.39 Mbps    0              524.18 KBytes

Summary
Interval       Throughput     Retransmits
0.0 - 10.0     943.54 Mbps    0
```

perfSONAR Embedded Systems

Lessons Learned: Staff Proficiency

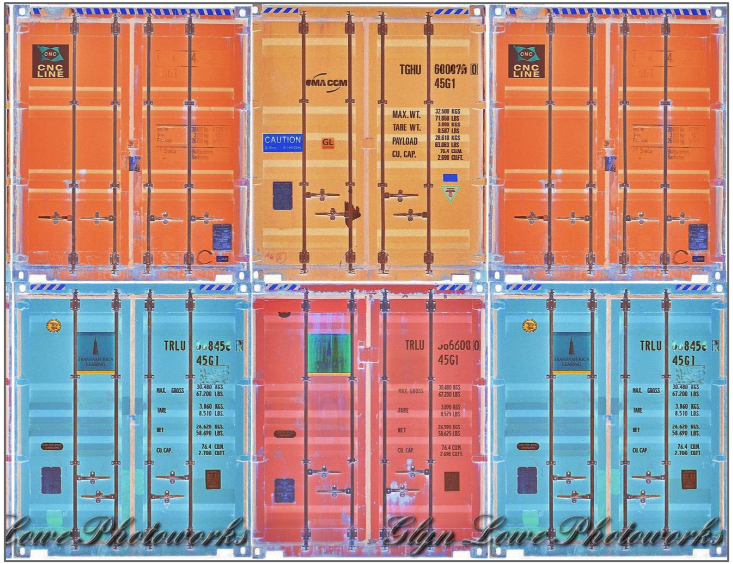
Necessary Breadth and Depth of Skills

- perfSONAR Administrator
  - deploy / maintain perfSONAR nodes & meshes
- Network Administrator
  - Switch configuration
  - Network Architecture
- System Administrator
  - Operating System Expertise: UNIX
  - Containers: Docker, LXC, proprietary, etc.
  - NTP / Security / System Tuning



Advice for Next Generation Appliances

- Throughput speeds = max line speed
  - Faster / parallel bus?
  - Application specific ports?
- Implement container architecture
  - Docker, LXC, etc.
- Simple & powerful container networking
  - Reduce configuration complexity
  - Route through any data port ("wireless")

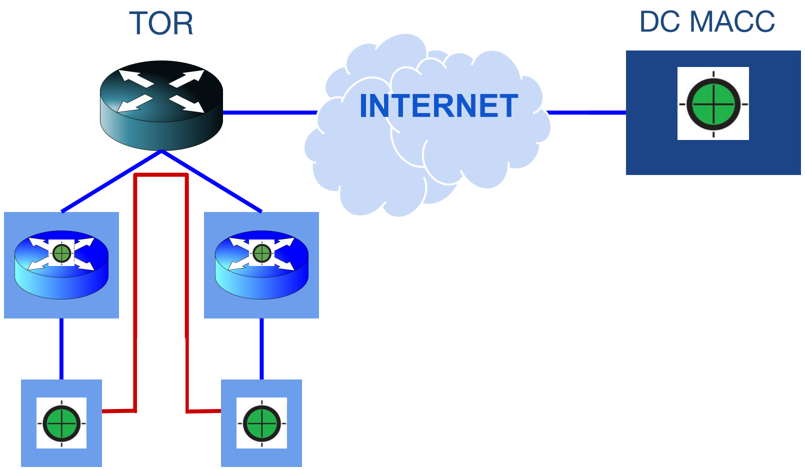


INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN
ESnet ENERGY SCIENCES NETWORK
GÉANT
INDIANA UNIVERSITY
INTERNET2
UNIVERSITY OF MICHIGAN

perfSONAR Embedded Systems

Next Steps

- Deploy perfSONAR in Umich Lab
  - EVO on 5200
- Enhance documentation
- Automate container creation
- Explore container security architecture
- Further testing
  - Add traffic, more hosts
  - Inclusion in test meshes



INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN
ESnet
ENERGY SCIENCES NETWORK
GÉANT
INDIANA UNIVERSITY
INTERNET2
UNIVERSITY OF MICHIGAN

perfSONAR Embedded Systems

Credits

University of Michigan

Eric Boyd, Phil Camp, Ed Colone, Dan Eklund, Brady Farver, Ryan Goniwiecha, Amy Liebowitz, John Simpkins, Katarina Thomas

Juniper

Erik DeHaas, Mrinal Dhillon, JJ Jamison, Robin McClear, Shane Praay

Internet2

Mark Feit

INFORMATION AND TECHNOLOGY SERVICES
UNIVERSITY OF MICHIGAN

ESnet
ENERGY SCIENCES NETWORK

GÉANT

INDIANA UNIVERSITY

INTERNET2

UNIVERSITY OF MICHIGAN