

# Lanka Education and Research Network

---

## Measuring Network Performance with perfSONAR

19<sup>th</sup> to 22<sup>nd</sup> July 2022

*Network Management and Measurement Workshop*

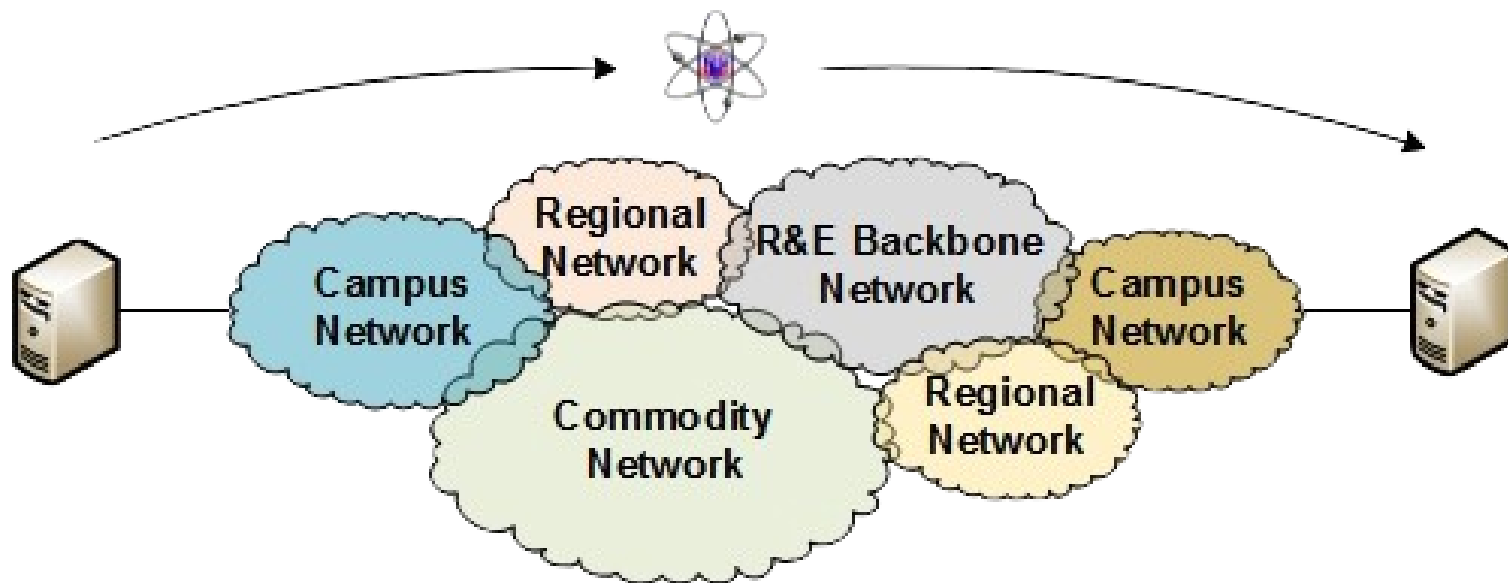
Dhammika Lalantha / LEARN



# Problem statement

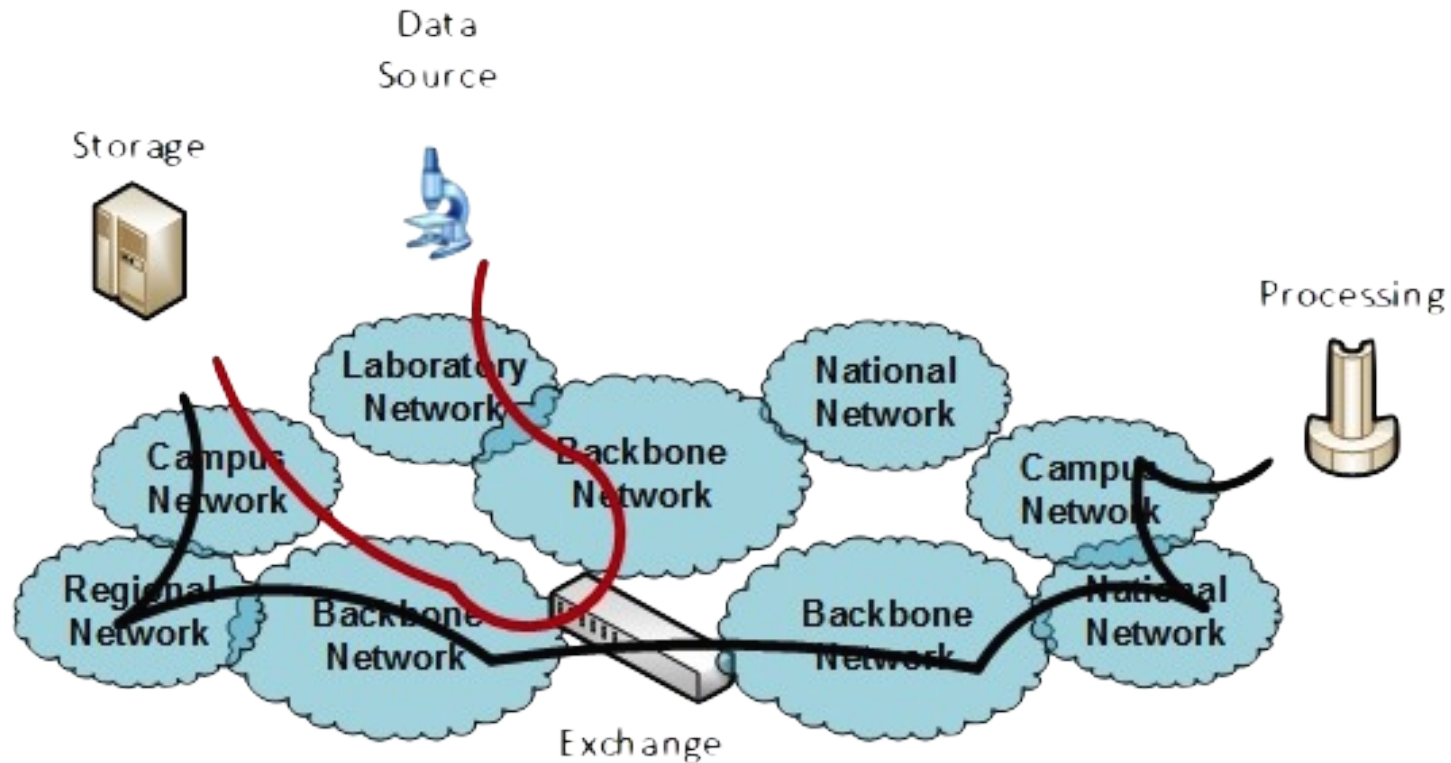
---

- This complex, heterogeneous set of networks must operate seamlessly from end to end to support science and research collaborations that are distributed globally.



# Problem statement

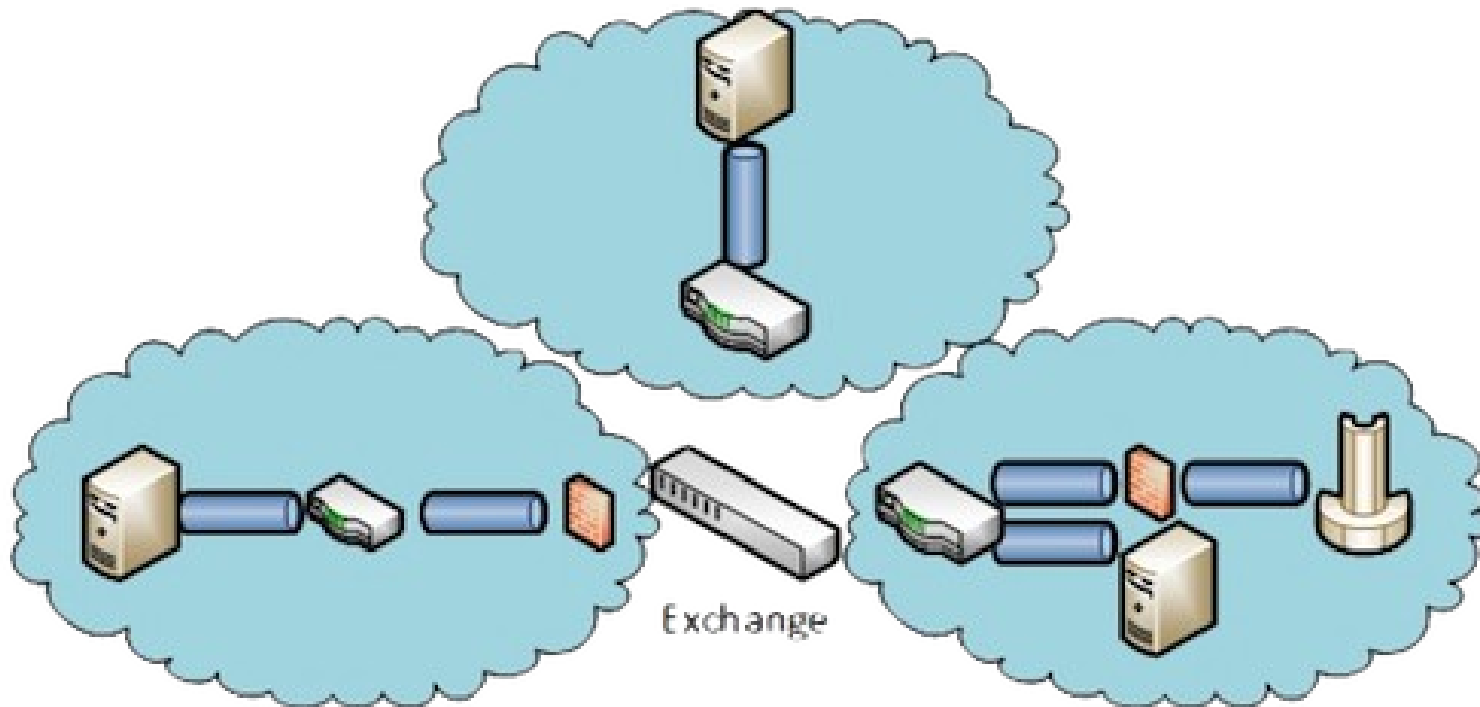
- Higher data rate applications and services of academic and research entity at a remote end are accessed by another similar entity at another location of a world is a must these days.



# Problem statement

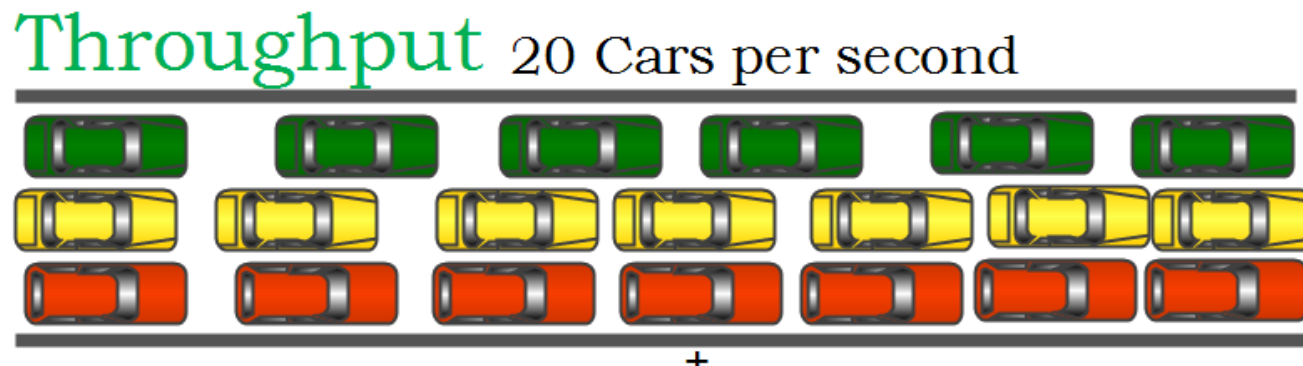
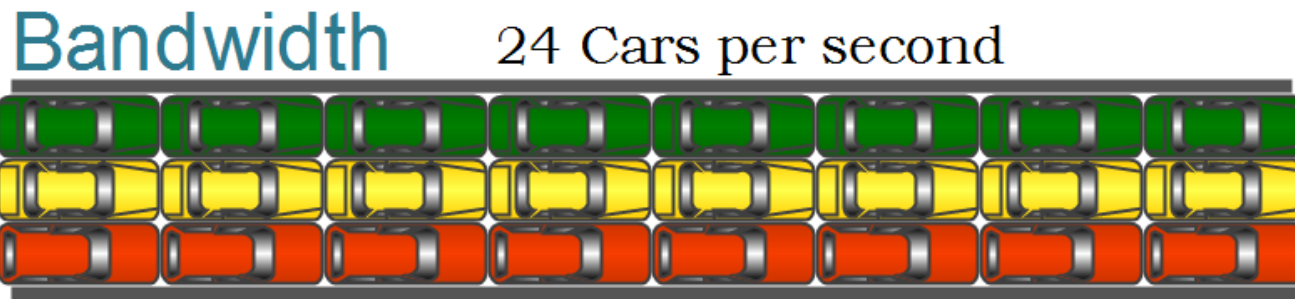
---

- While these networks all interconnect, each network is owned and operated by separate organizations (called domains) with different policies, customers, funding models, hardware, bandwidth and configurations.



# Throughput

- Bandwidth vs Throughput
  - Bandwidth – theoretical data rate, channel capacity, maximum capacity
  - Throughput – Actual data transfer rate



# Throughput

---

- Throughput limiting factors
  - Network Latency
  - Packet Loss and errors
  - Network Congestion
  - Enforced Limitation (Soft limit by ISP)
  - Protocol overhead (Transmission medium)
    - Frame Overhead (24 byte) = Preamble (7 byte) + Start frame delimiter (1 byte) + Inter-frame gap (12 byte) + VLAN tag (4 byte)

# Throughput

---

- Throughput calculation of a Network Interface
  - $\text{Throughput} = \left[ \frac{\text{Frame Size}}{\text{Frame Size} + \text{Frame Overhead}} \right] * \text{Switch Port Bit Rate}$
- Eg:
  - 1 Gigabit Ethernet Interface
  - Used frame size – 1518 byte
  - $\text{Throughput} = \left[ \frac{1518}{1518 + 24} \right] * 1000 \text{ Mbps}$   
 $= 984.43 \text{ Mbps}$



# TCP Problem

---

- TCP is mostly used for bulk data transfer which can fail due to packet loss and higher latency
- TCP throughput is reduced with the latency and packet loss
- Mathis Equation

Mathis Equation to calculate the maximum TCP throughput

$$\frac{MSS}{RTT} * \frac{1}{\sqrt{p}}$$

Where

- **MSS:** Maximum Segment Size
- **RTT:** Round-Trip Time
- **p:** Packet Loss

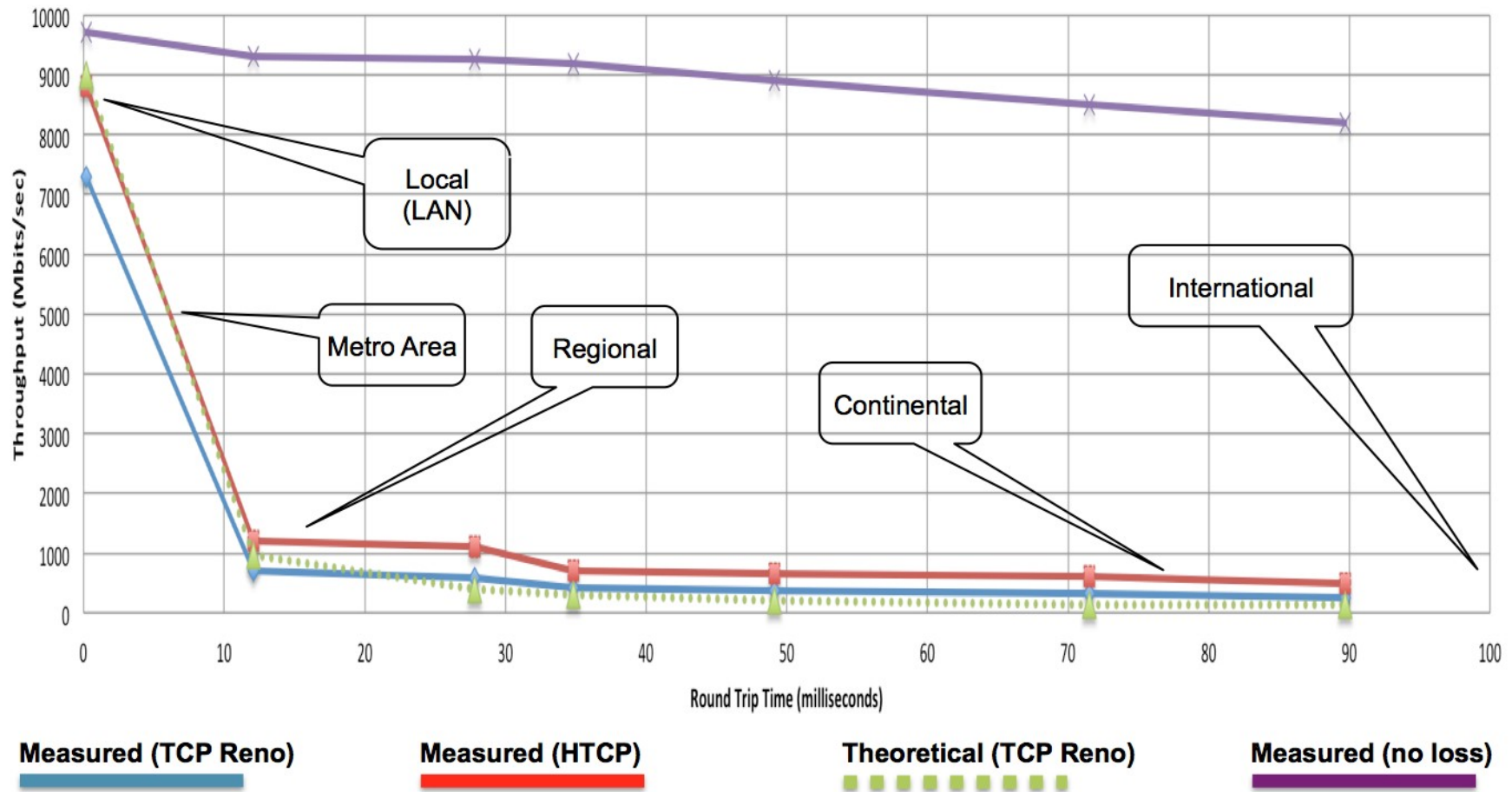
# TCP Problem

---

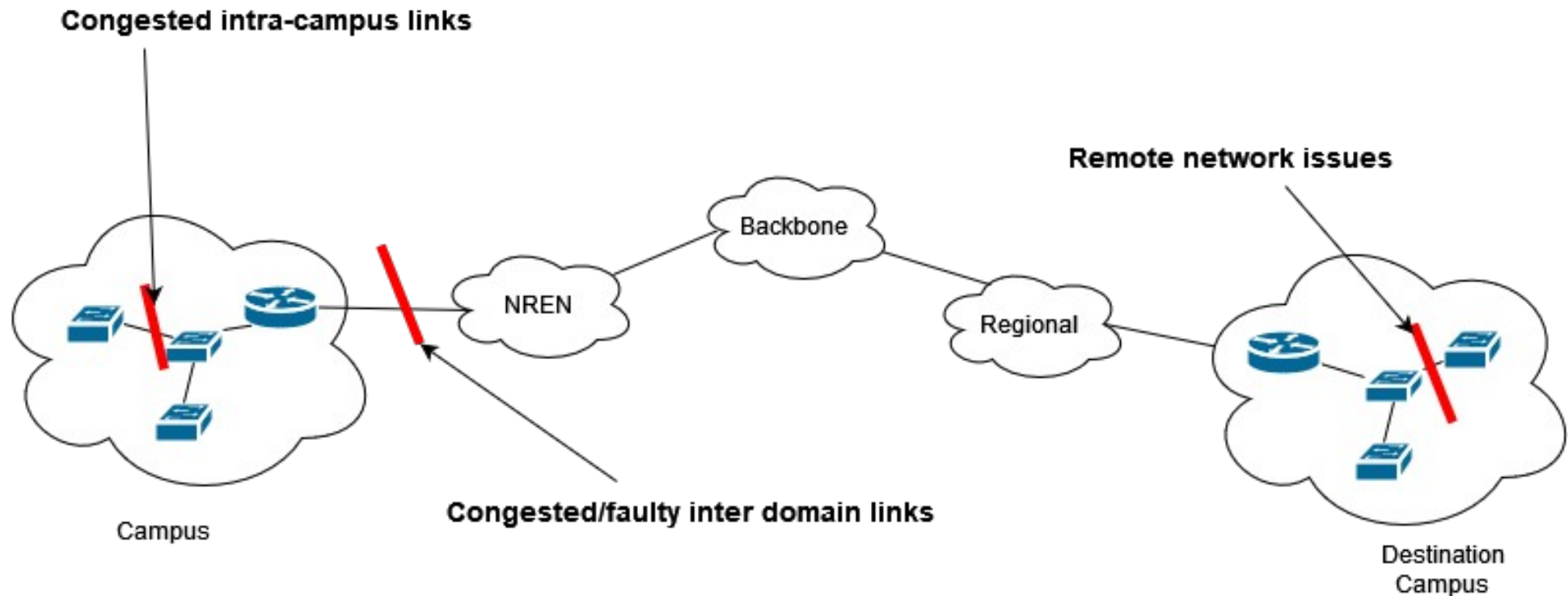
Round trip latency	TCP throughput with no packet loss Round trip latency	TCP throughput with 2% packet loss
0 ms	93.5 Mbps	3.72 Mbps
30 ms	16.2 Mbps	1.63 Mbps
60 ms	8.7 Mbps	1.33 Mbps
90 ms	5.32 Mbps	0.85 Mbps

# TCP Problem

Throughput vs. increasing latency on a 10Gb/s link with **0.0046%** packet loss



# Where the issues can happen ?



# Soft Failures vs Hard Failures

---

- Hard failures
  - Fiber cut
  - Power failure takes down routers
  - Hardware ceases to function
  - Easy to detect through classical monitoring systems
- Soft failures
  - Basic connectivity works
  - Performance is poor
  - Sometimes go undetected

# What is perfSONAR ?

---

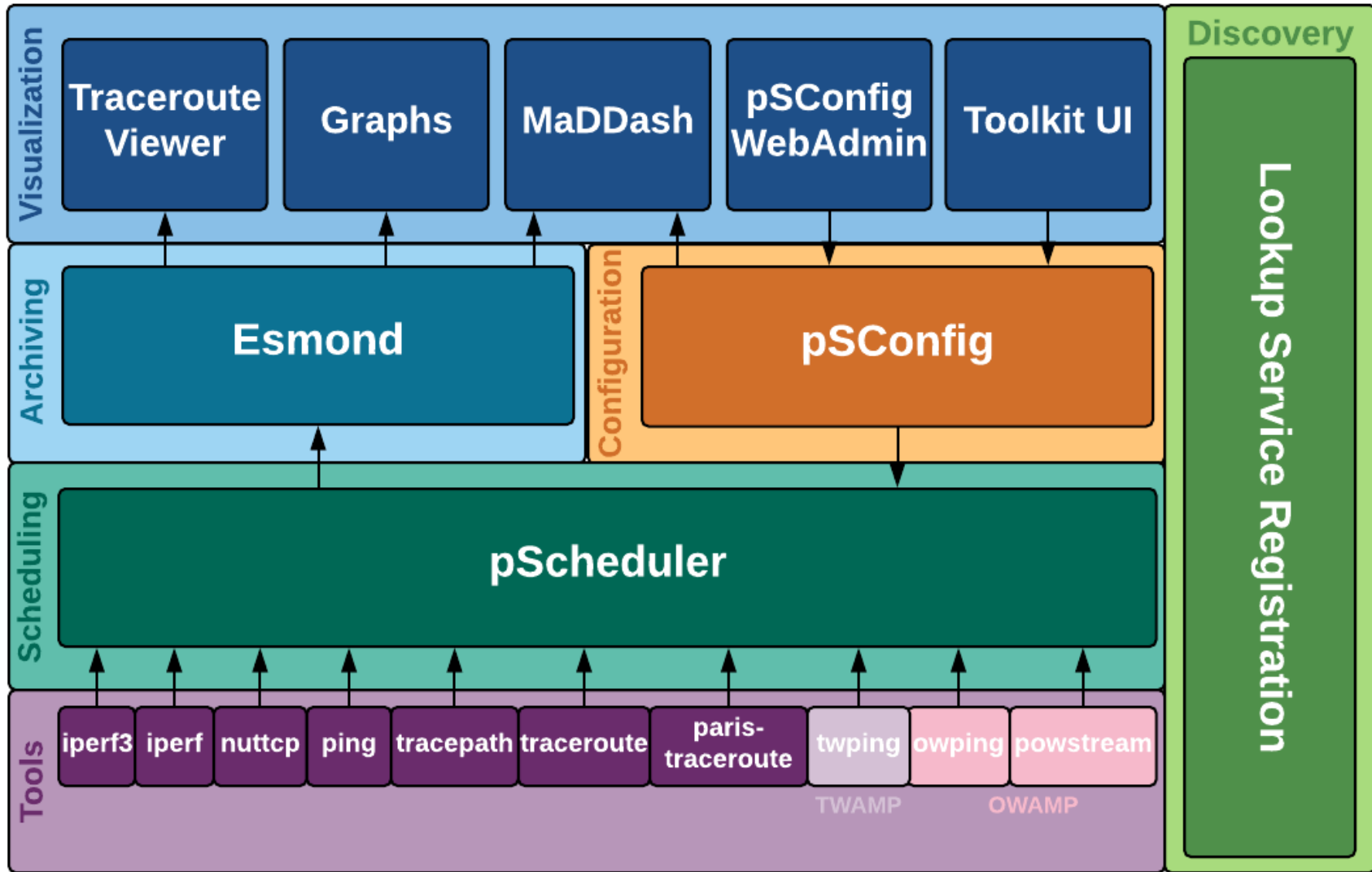
- perfSONAR is a collection of software for performing and sharing end-to-end network measurements.
- Package a collection of tools to do perform network measurements.
- Useful for detecting soft failures
- Widespread deployment of all around the world (2000 nodes)
- Schedules the measurements tests
- Archive the test results
- Visualize test results in useful way
- Joint effort lead by
  - ESnet, GEANT2 JRA1, Internet2, RNP (Brazil)
  - With other universities and institutes

# Measured Metrics

---

- **Throughput** (sometimes called achievable bandwidth) is the amount data that is transferred from a specific source to a specific destination at a given time.
- **One-way latency** is a measure of how long data takes to travel from one host to another.
- **Round-trip time** (or two-way latency) is the measure of how long data takes to travel from one host to another and back to the first host.
- **Packet loss** is a measurement of how many packets are dropped for any reason on a network segment or path.
- **Packet duplication** is a measure of how many packets are duplicated for any reason on a network segment or path.
- **Jitter** is the variation in arrival times for packets between two participating endpoints.

# What is perfSONAR ?





# perfSONAR tools

---

- **owamp** - A tool primarily used for measuring packet loss and one-way delay. It includes the command `owping` for single short-lived tests and the `powstream` command for long-running background tests.
- **twamp** - A tool primarily used for measuring packet loss and two-way delay.
- **iperf3** - A rewrite of the classic `iperf` tool used to measure network throughput and associated metrics.
- **iperf2** - Also known as just `iperf`, a common tool used to measure network throughput that has been around for many years.
- **nuttcp** - Another throughput tool with some useful options not found in other tools.
- **traceroute** - The classic packet trace tool used in identifying network paths

# perfSONAR tools

---

- **tracpath** - Another path trace tool that also measures path MTU
- **paris-traceroute** - A packet trace tool that attempts to identify paths in the presence of load balancers/multiple paths
- **ping** - The classic utility for determining reachability, round-trip time (RTT) and basic packet loss.

# Scheduling Tasks

---

- PerfSONAR user a scheduler to run different tasks in a scheduled manner.
- Finding time-slots to run the tools while avoiding scheduling conflicts that would negatively impact results
- Executing the tools and gathering results
- Sending to the results to the archiving layer (if needed)

# Archiving Measurement Results

---

- Archive measurement results time-series databased called esmond.
- Two ways to setup measurement archive
  1. It can be installed on each measurement host (perfSONAR node)
    - Comes with perfSONAR toolkit by default
  2. Installing a in central server at multiple measurement host environment
    - Need additional configurations
    - Can be used when measurement host has limited resources

# Configuration

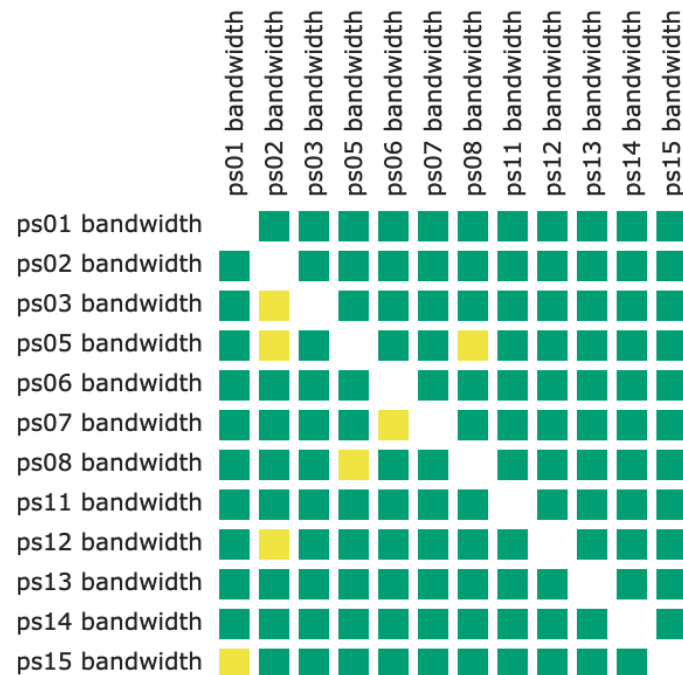
---

- In perfSONAR measurements tasks can be configured in configuration templates.
- This is done by pSConfig template framework
- PSConfig helps,
  - Scheduling tasks at multiple nodes
  - Maintaining visualization components to display results from multiple nodes
- PSConfig has two agents
  - Psconfig pScheduler Agent (Run on perfSONAR node)
    - read configuration templates and send task to the scheduler for execution
  - PSConfig MaDDash Agent (Run on MaDDash server)
    - Read configuration templates and helps MaDDash to display the results on a Dashboard

# Visualization Measurement Data

- MaDDash visualize measurement results in the measurement archive.
- View results of measurements of multiple nodes in a two dimensional way as a set of grids

✔ No problems found in grid



# Discovering Federated Nodes

- <https://stats.es.net/ServicesDirectory>

**perfSONAR**      **Lookup Service Directory**

---

**Search**

Filter results by searching for specific terms:  
🔍

**Search**    **Show All**

**Browser**

- ▶ pScheduler Server 1811
- ▶ BWCTL Server 202
- ▶ OWAMP Server 2143
- ▶ NDT Server 46
- ▶ NPAD Server 14
- ▶ Ping Responder 64
- ▶ Traceroute Responder 63
- ▶ MA 2005
- ▶ BWCTL MP 158
- ▶ OWAMP MP 142
- ▶ twamp 1883

**Service Information**


Service Name	Addresses	Geographic Location	Communities	Version	Custom

**Host Information**

Host Name	Hardware	System Info	Toolkit Version	Communities

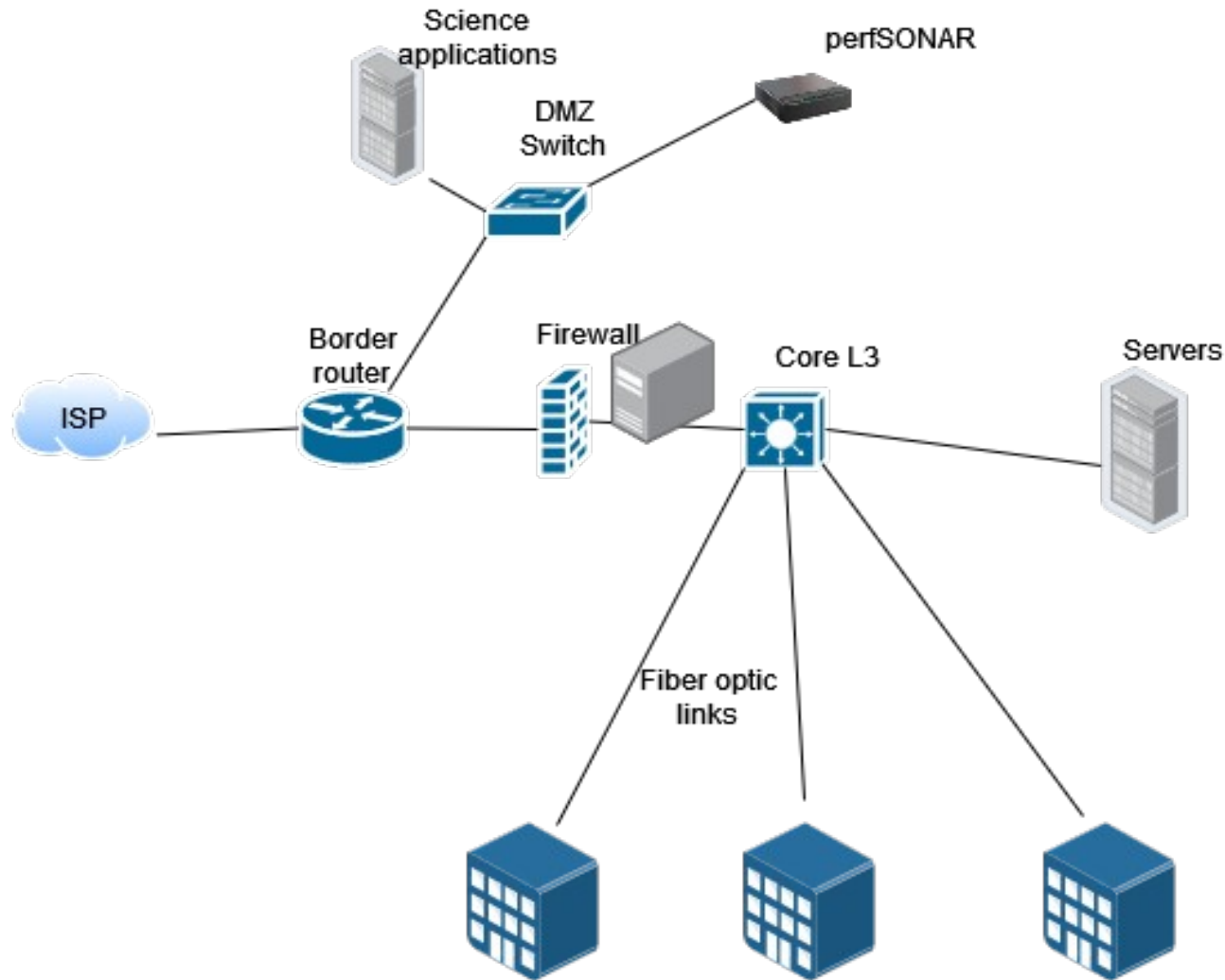
**Service Map**

Map    Satellite



# Deployment

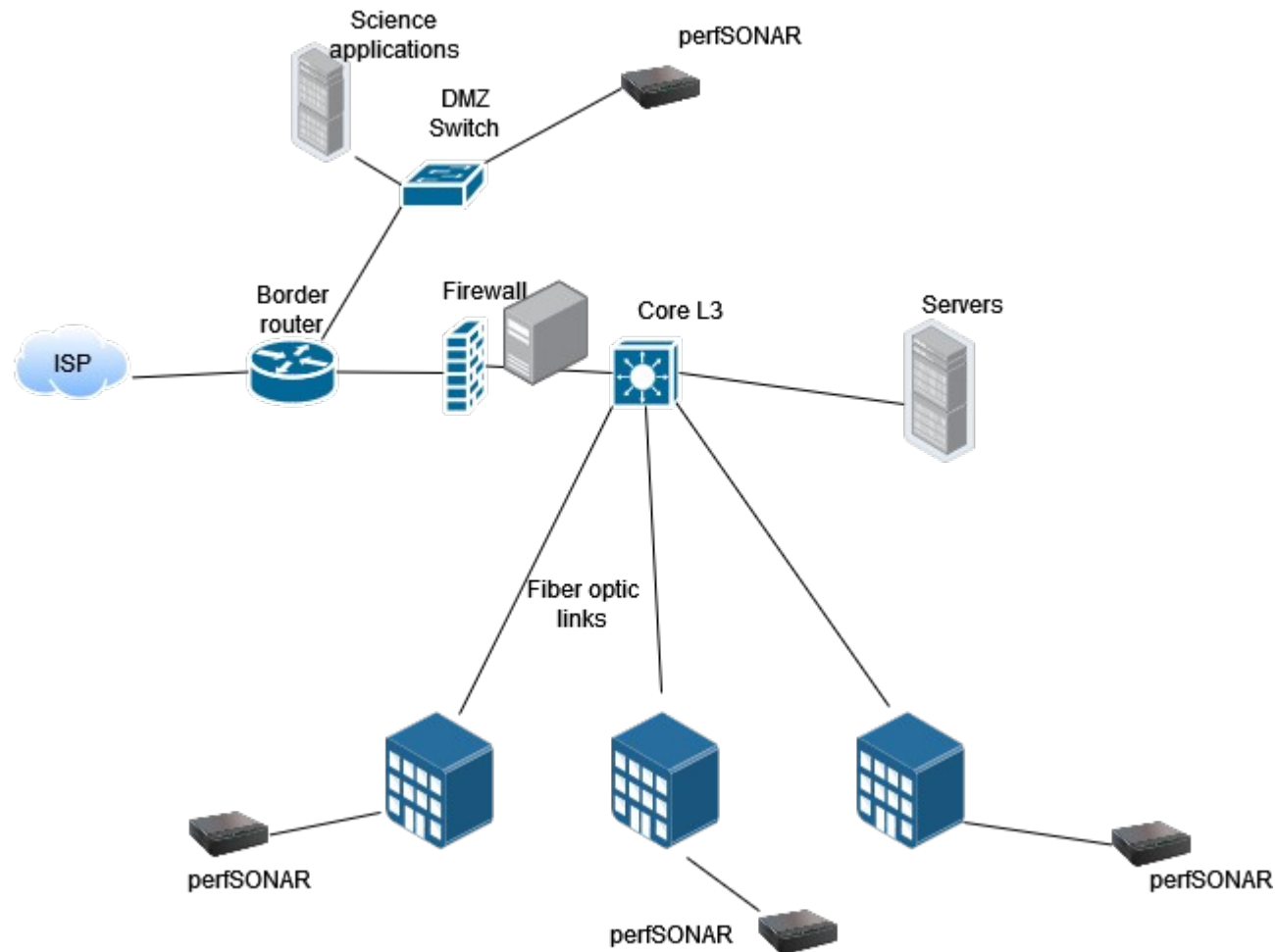
- Recommended to deploy in Science DMZ





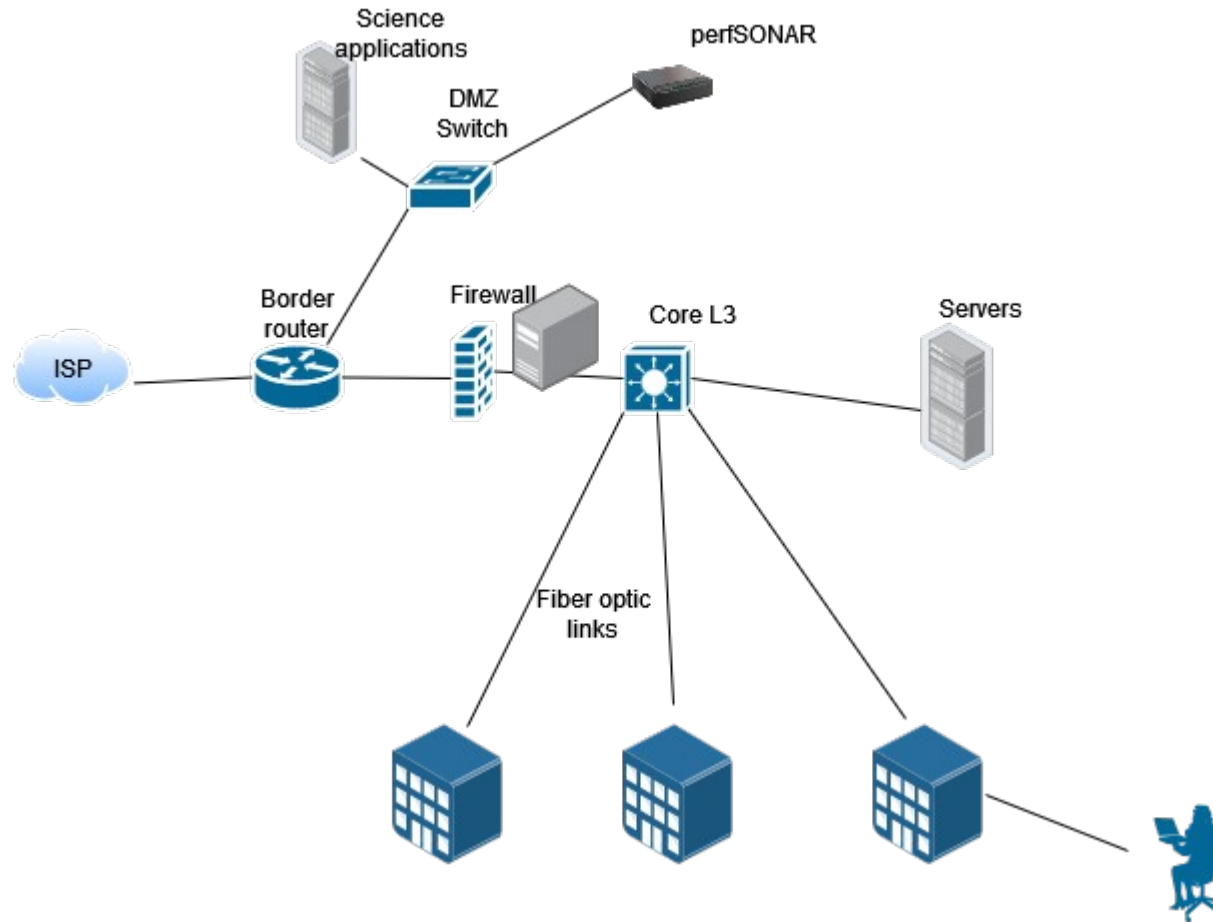
# Deployment – Measurements within campus

- To troubleshoot within campus network you can have multiple nodes between branches/sites/faculties



# On-demand testing

- When you face a issue on your links



# Develop a Test-plan

---

- What are you going to measure?
  - Throughput
    - Important collaborators
    - Mostly used destinations
    - Between sites
    - 4 times per day to each destination
    - Run for shorter periods (20 seconds)
    - Consider overnight/off peak time scheduling
  - Packet loss/Availability/Latency
    - Important collaborators
    - Mostly used destinations
    - Between sites
    - 6 times a day on a regular interval

# Lanka Education and Research Network

---

Thank You

Dhammika Lalantha/LEARN

Email: [lalantha@learn.ac.lk](mailto:lalantha@learn.ac.lk)