

# Gnodal



## **Lowest Latency, Highest Density, Scalable Ethernet Networks**

150ns Latency - 1.5Tbps Throughput - 72 ports in 1U - 1.6W / port

Scalable, Ultra-low Latency 10/40 Gigabit Ethernet Switching  
Solutions for the High Performance Data Center

**Presented to UKNOF  
Dr. John Taylor**

# Overview of Presentation



- **High Performance Data Centre Ethernet**
  - **Problem**
  - **Solution**
  
- **Gnodal Products and Technology**
  - **GS-Series**
  
- **Performance Characteristics**
  
- **Case Study – Distributed Fairness**



# Fundamental Problems

- **Ethernet switches congest, leading to...**
  - Reduced network bandwidth
  - High unpredictable latency
  - Limited scalability and over provisioning
  - Poor quality of service and lost business
  - Increased power consumption
  - Inefficiency and increased total cost of ownership
  
- **Who cares: customer examples & verticals**
  - Financial Services: HFT, Equities, Algo, CoLo, DCs
  - Telco, SP: DC Hosting
  - Academic & research: HPC Clusters, converged networks
  - Oil & Gas: Seismic analysis, DCs
  - Media: Post-production, Back-end delivery

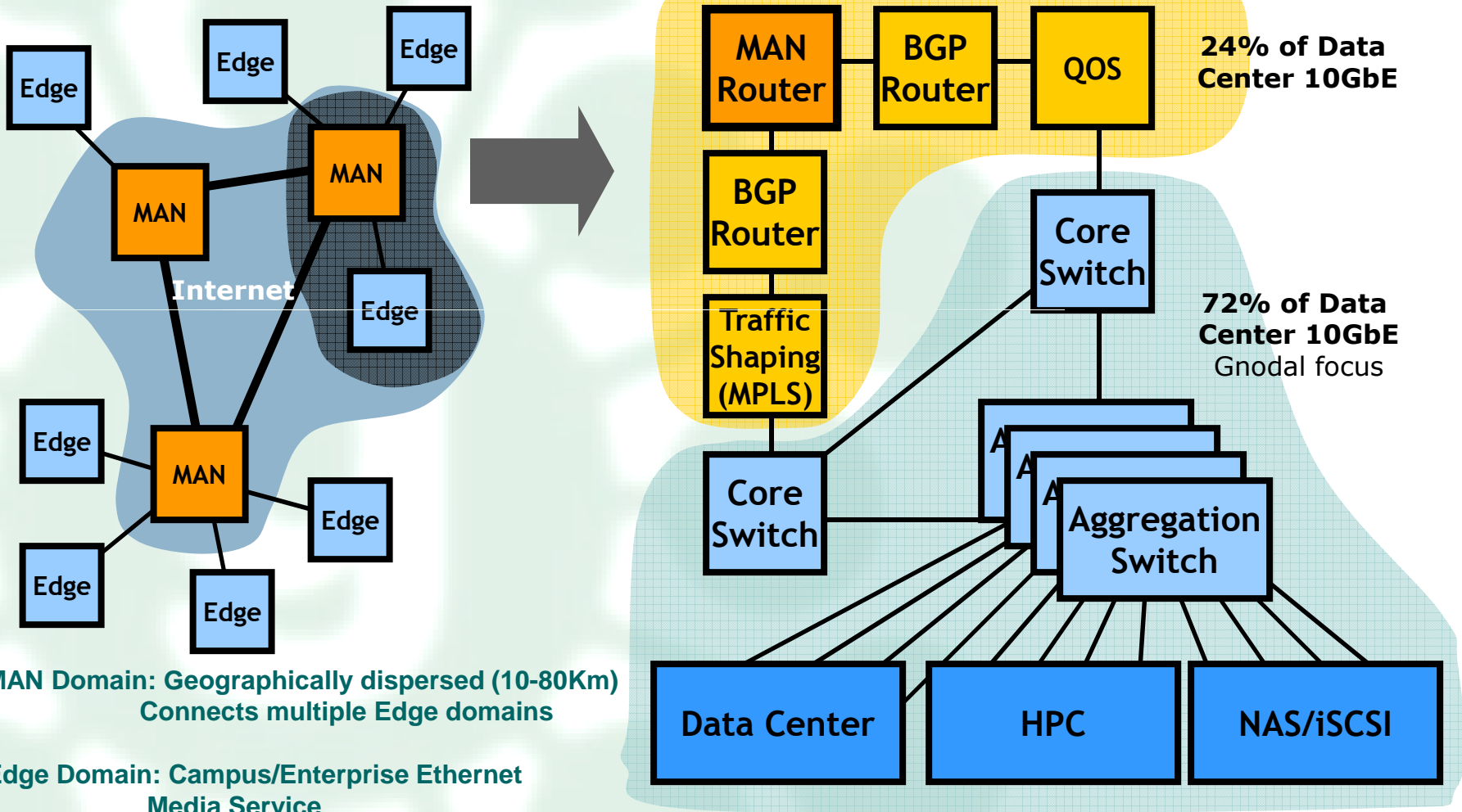
# Revolutionary Solution



- **Apply Techniques Honed From High Performance Computing to the Ubiquity of Ethernet**
  - Develop ASIC encapsulating Scalability
  - Support Ethernet Edge with Gnodal Fabric
  - Mitigate Ethernet Standards where insufficient
  - Support highest line-rate with state-of-the-art
  - Support L2 and L2+ features
  - Design-in High Reliability
  - Provided flexibility in product packaging
  - Ensure ease of administration

**World's fastest and only fully adaptive, load balancing Ethernet Fabric**

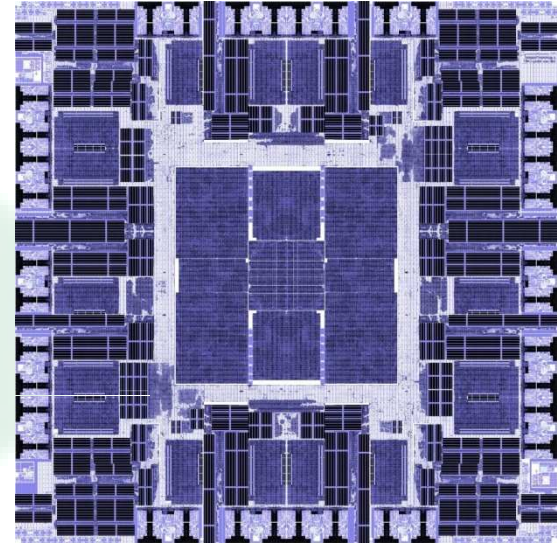
# Current HPDC Space



# Technology Peta ASIC Features



- Fab'ed by IBM
  - 65nm Cu Process
  - Proven SERDES technology
- Low latency
  - 64ns Pin to Pin latency
  - 150ns First hop
  - Full crossbar/Non blocking
- 72 Port 10GHz Integrated PHYs
  - 72 x 10GbE MAC
  - 18 x 40GbE MAC
  - 18 x 40Gb Gnodal MAC



# Product Description

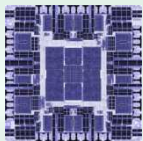


## Gnodal's System:

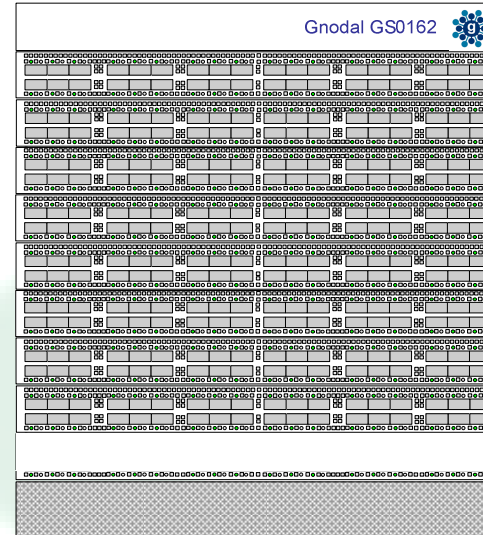
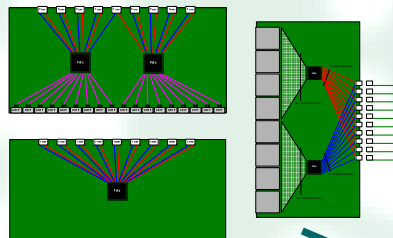
- Scales from 72 to 64000 ports
- Largest, most cost effective data center switch
- Highest per port bandwidth
- Best Performance, Power and Utilisation
- Highly Flexible

## Gnodal designed chip:

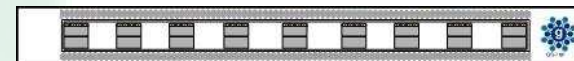
- **Solves fundamental congestion problem**
- Lowest network latency
- Lowest power consumption



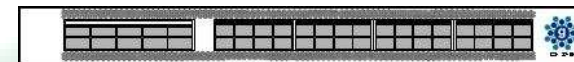
**Gnodal Peta**  
10/40Gb Ethernet ASIC



**Large Modular Fabric Switches**

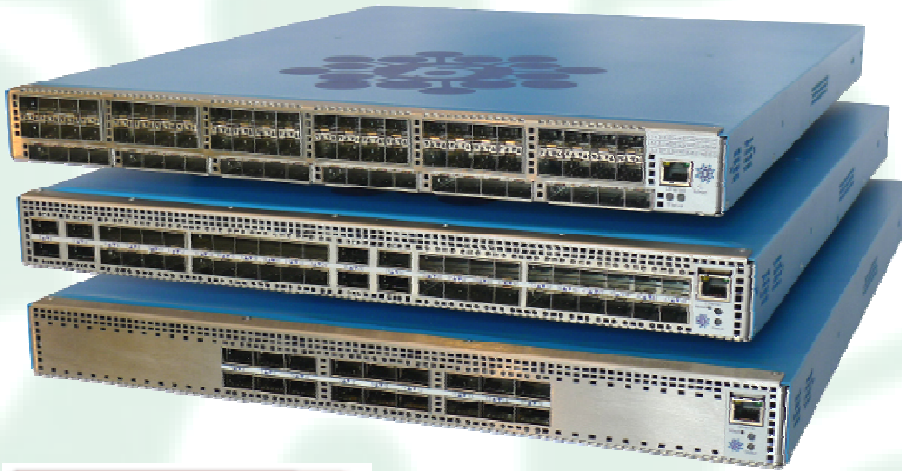


**Small Fixed Fabric Switches**



**Top of Rack Access Switches**

# GS Series Switches

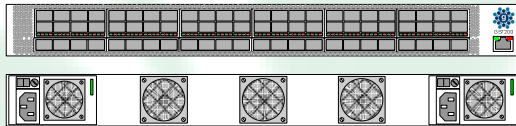




# Initial Products - Fixed



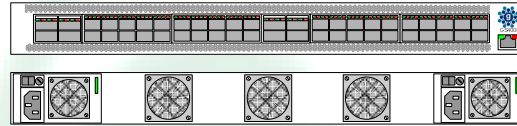
GS7200



**72 x 10GbE SFP+**

**112 - 262W**

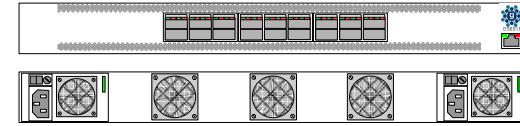
GS4008



**40 x 10GbE SFP+**  
**8 x 40Gb QSFP**

**112 - 242W**

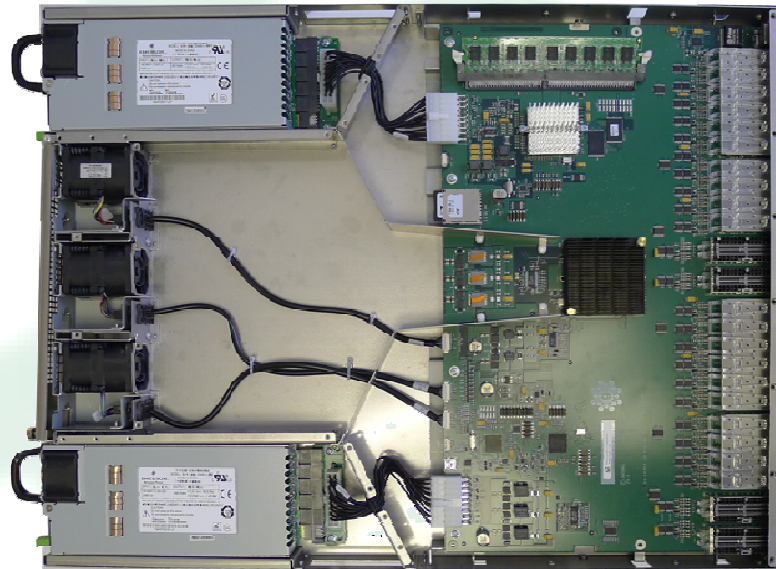
GS0018



**18 x 40Gb QSFP**

**112 - 202W**

**1.5Tb/s Throughput**  
**Hot-swap PSU / Fans**  
**1RU**





# Performance

- Latency
- Scalability – Distributed Fairness
- Throughput – Putting it Altogether

# Performance – Latency



## Gnodal hardware switching performed by switch ASIC

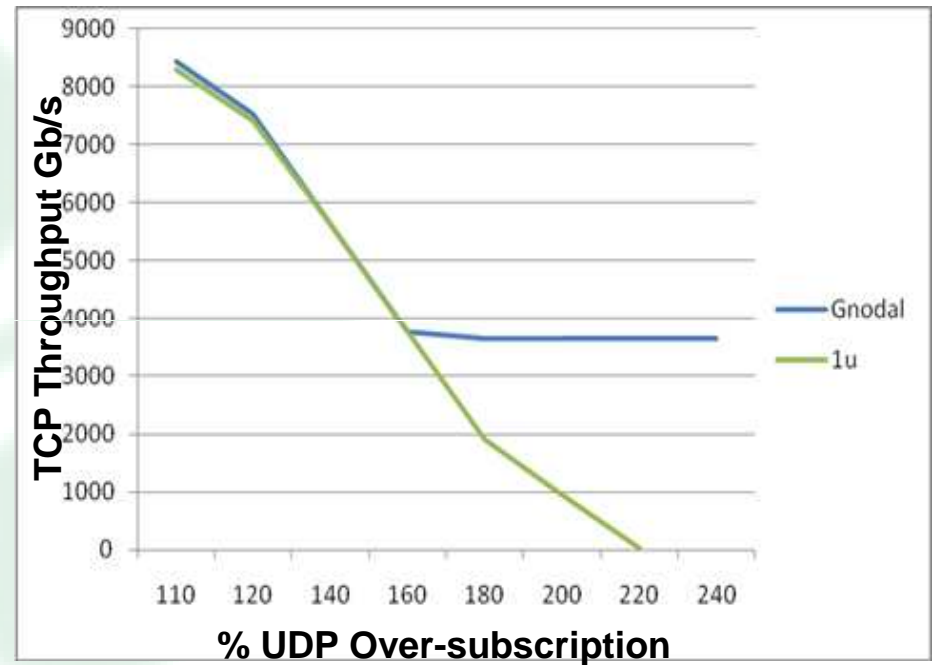
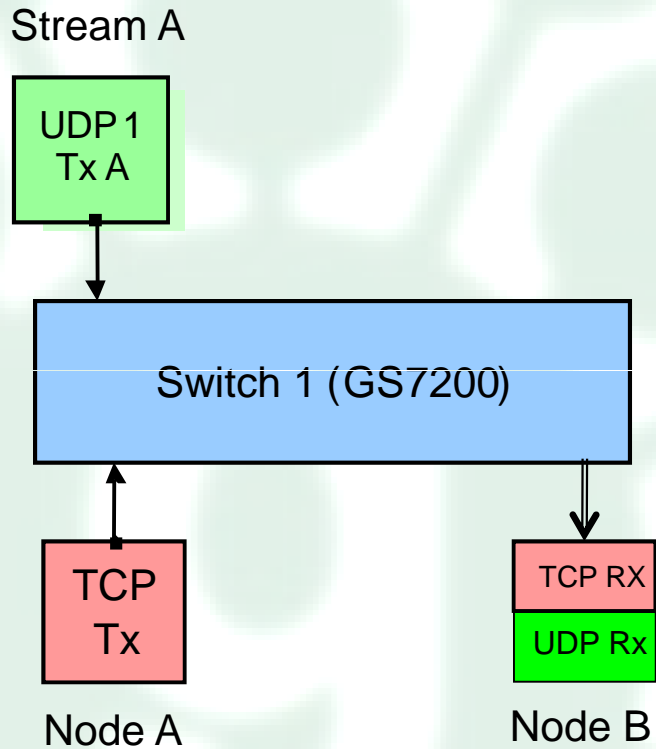
- Minimum Latency (Measured per IETF RFC1242)

| Ports                     | Hops | Gnodal<br>10GbE | Gnodal<br>40GbE |
|---------------------------|------|-----------------|-----------------|
| 72                        | 1    | 150ns           | 150ns           |
| Inter switch <sup>¥</sup> | +1   | 66ns            | 66ns            |
| 720 <sup>¥</sup>          | 3    | 282ns           | 282ns           |
| 6480 <sup>¥</sup>         | 5    | 414ns           | 414ns           |

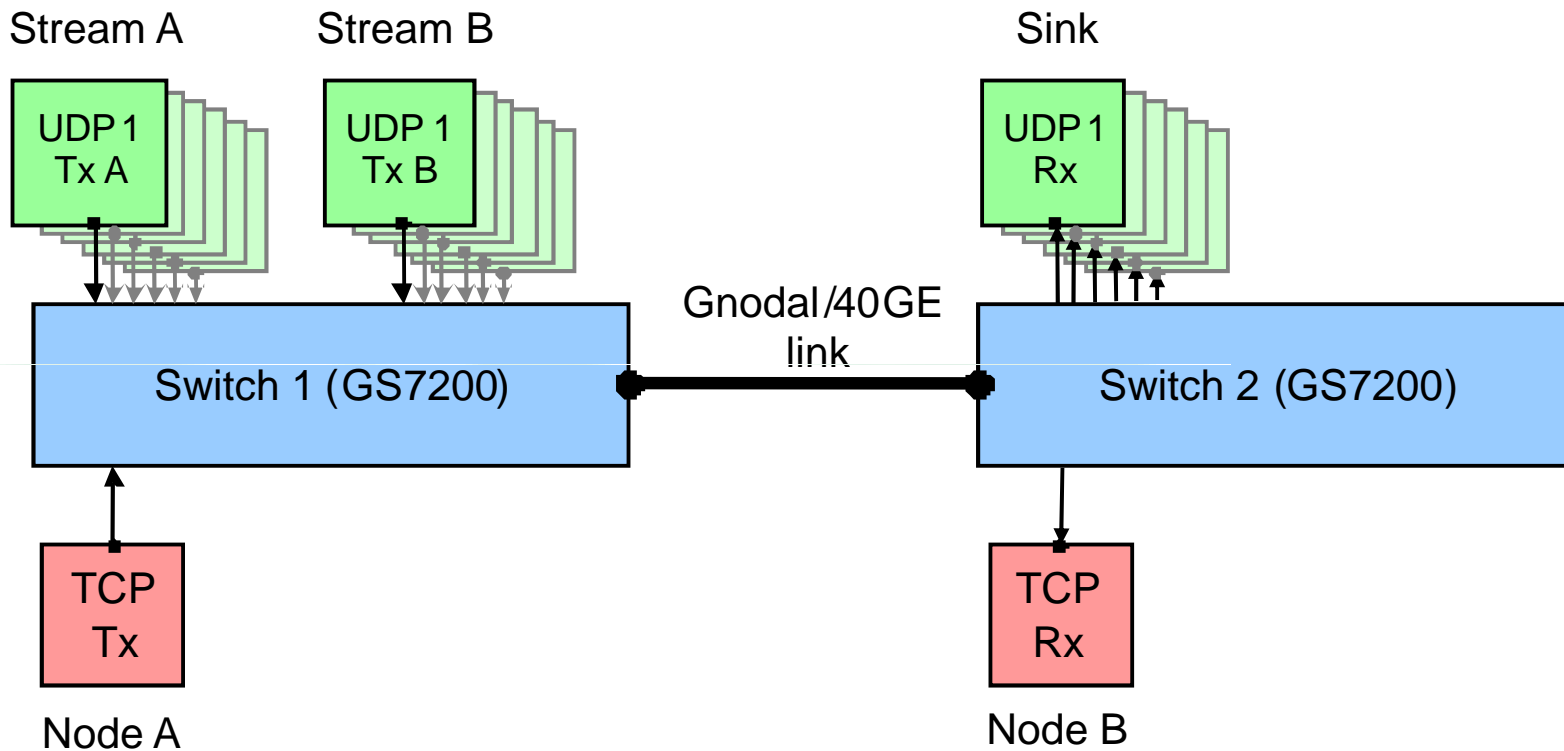
**150ns First Hop with 64nS cut-through latency on any additional hop**

<sup>¥</sup> 0.25m cable

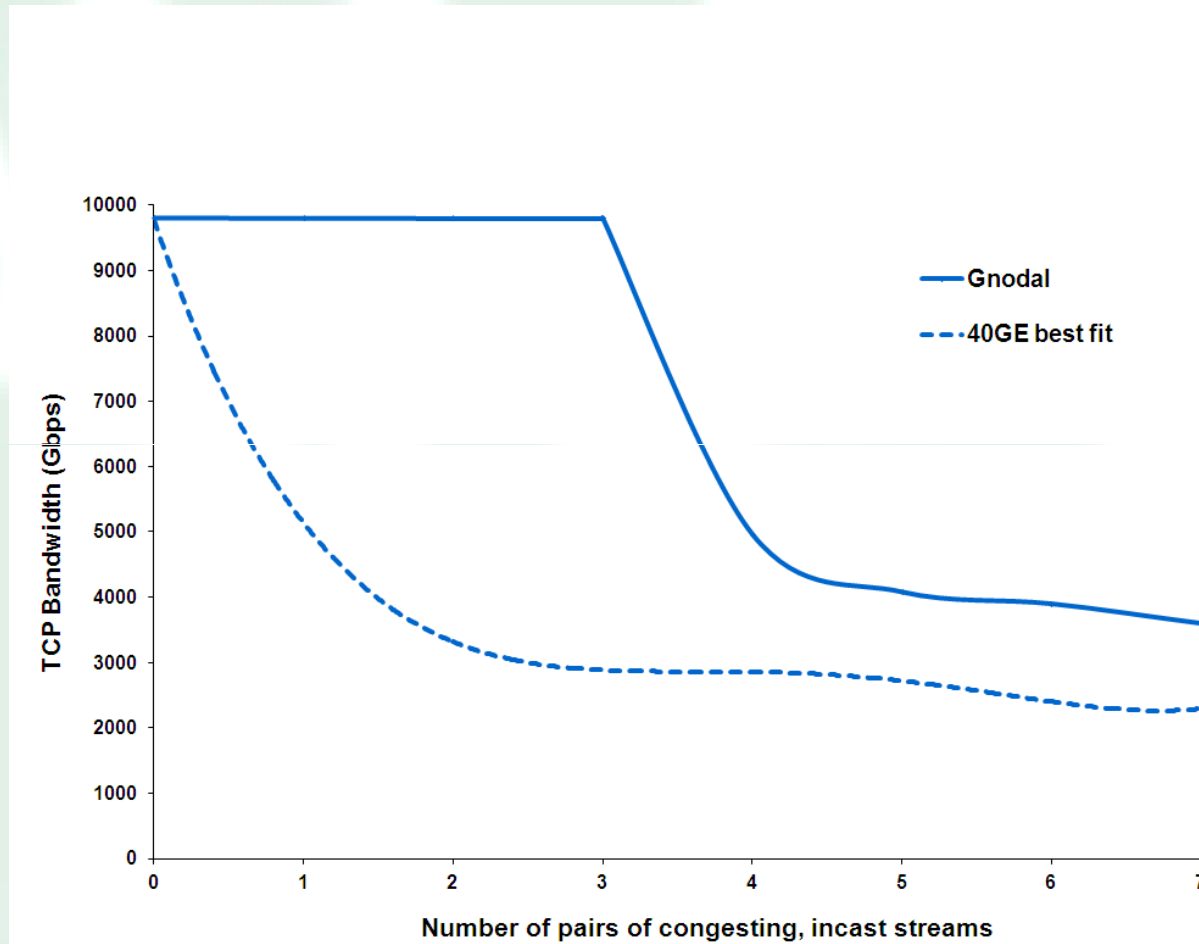
# Single Switch Test



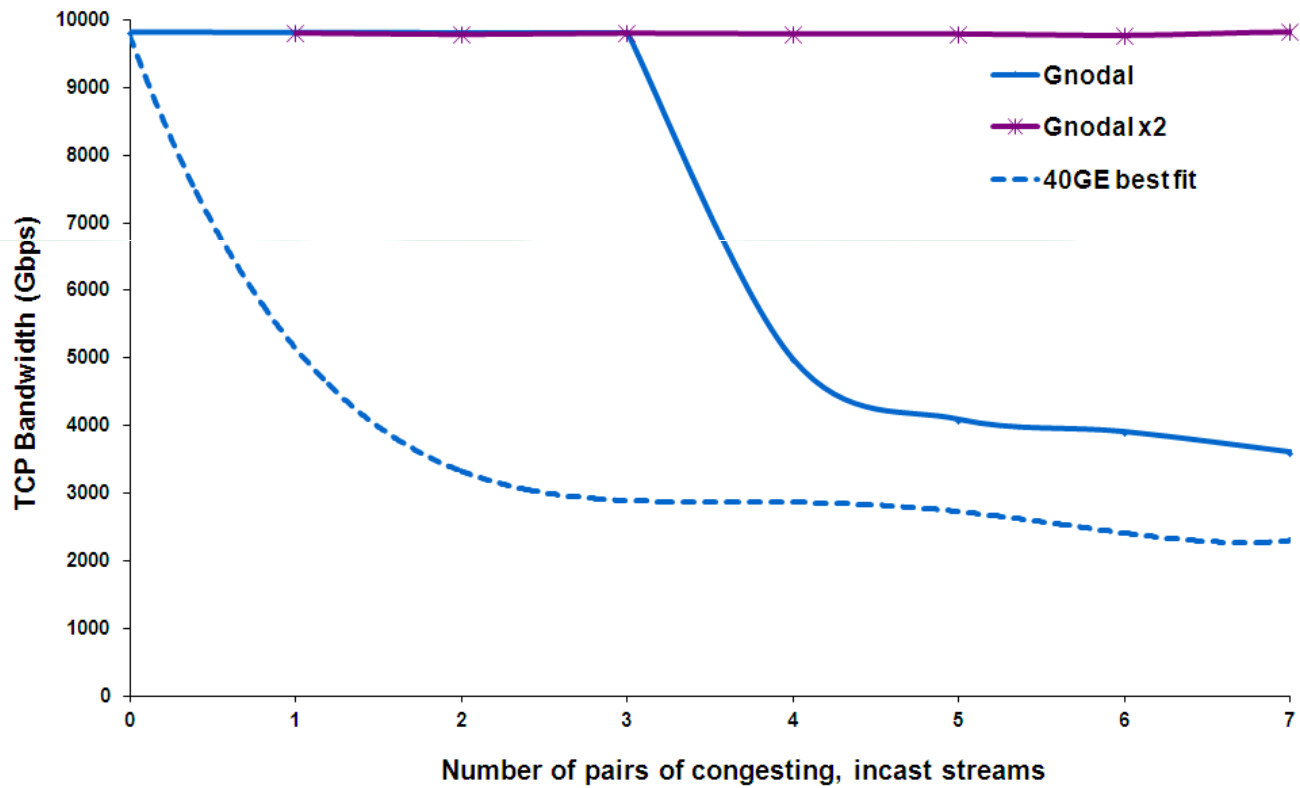
# Fabric Test - 1



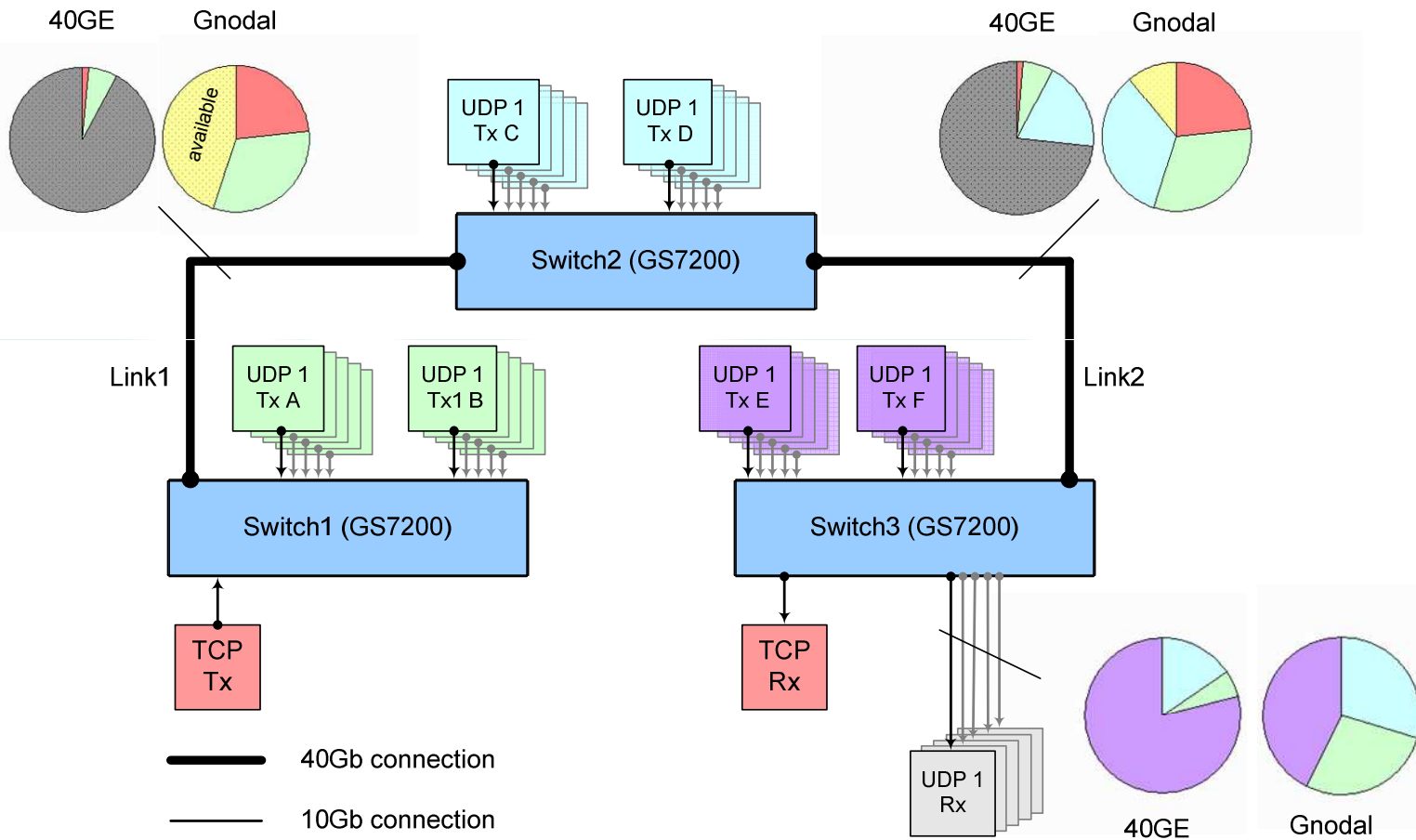
# Fairness



# Adding Another Fabric Link



# Fabric Test - 2





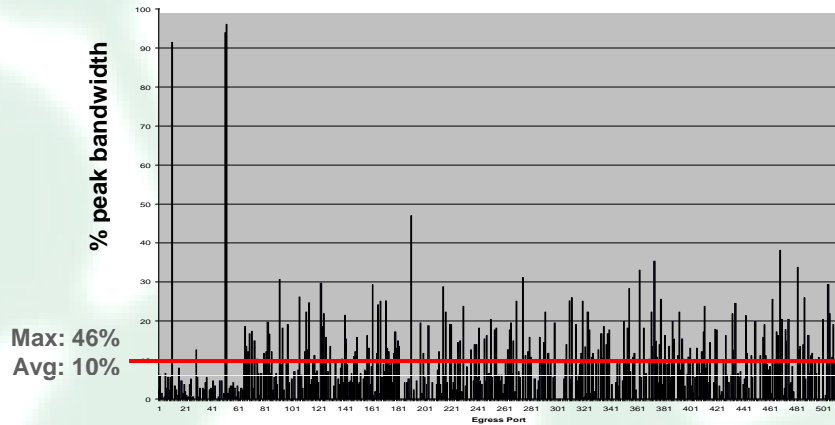
# Throughput Performance



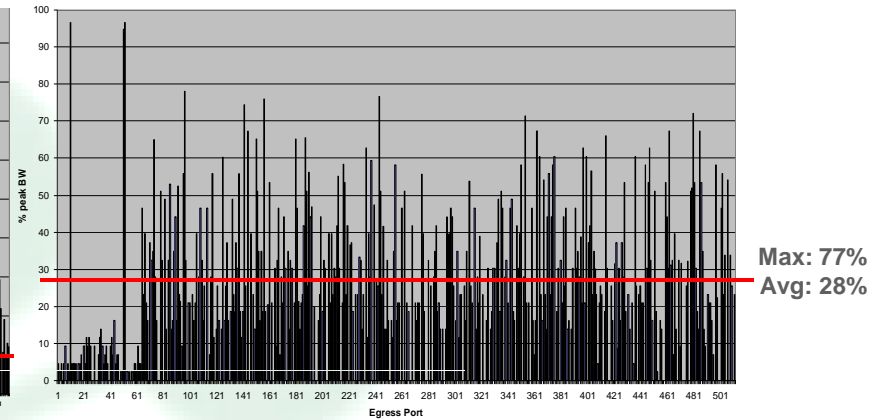
- Network
  - 512 port fat tree
  - 24 switches (16 leaf, 8 spine)
- Traffic pattern (typical HPDC)
  - Congesting incast
    - 106 ports (chosen at random) sending to 3 ports on two (left hand) leaf switches
  - Non-contending background traffic
    - 406 ports, randomly pairing, sending number frames of varying size
    - Occasionally re-pairing (random)

# Performance - Bandwidth

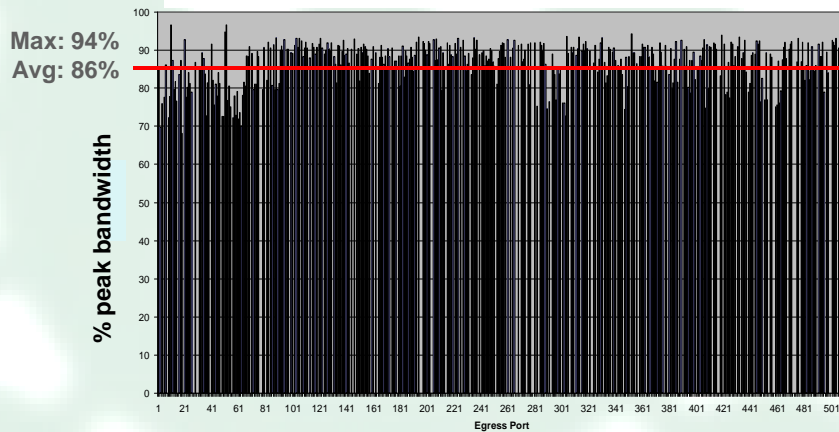
## Performance Characteristics During End-point Congestion



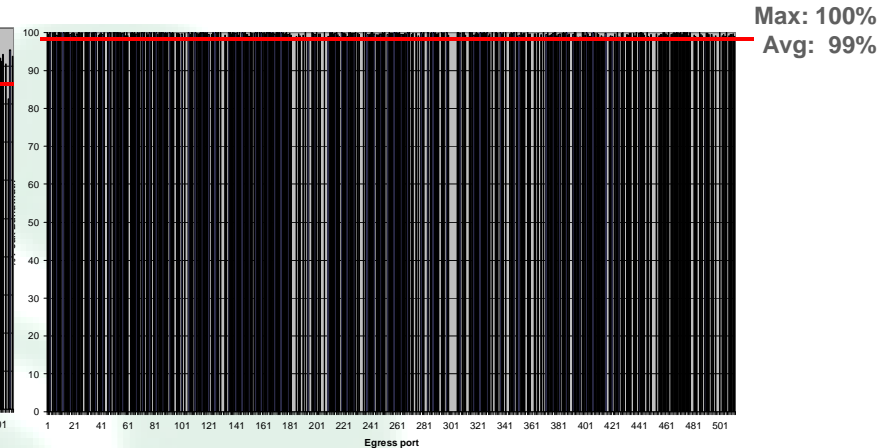
Typical Ethernet Switching



Proprietary Switching



Gnodal Ethernet Switching



Gnodal Ethernet Switching with Anti-congestion

# Competitive Advantages



| <b>Feature</b>   | <b>Benefit</b>   |
|--|--|
| <b>Gnodal ASIC (14 filed patents; 8 in draft)</b>                        | Smarts supported directly in hardware and not a “software integration” |
| <b>L2+ with lowest latency</b>   | Supports optimal device level interconnection for applications         |
| <b>Highest density 10 and 40 GbE</b>                                     | Saves on rack space and power budget                                   |
| <b>Native 40 GbE capability</b>  | Future proof for storage aggregation and direct 40 GbE NIC support     |
| <b>Scalability + Low power</b>   | Up to 64K Ports in single bridge                                       |
| <b>Multi-path capability in multi-unit installations</b>                 | High reliability and high utilization                                  |
| <b>Congestion free architecture with dynamic re-routing</b>              | Maximizes performance and utilization                                  |
| <b>Topology Invariant</b>  | Flexibility of deployment  |
| <b>Loss-less transport</b>   | Supports network convergence and RDMA                                  |
| <b>Up to million MAC Address per port and 1000s of Multi-Cast Groups</b> | Highly scalable  |
| <b>De facto like CLI</b>   | Ease of Administration   |
| <b>Open source AVL for optics and cables</b>                             | Optimal TCO and protects current investments                           |

# Summary



- L2+ with lowest latency and power
- Highest density in 1U
  - 10GbE and 40GbE edge access
- Multi-path capability in multi-chassis installations
  - 40Gb Load balancing fabric
  - Support for millions of MAC Addresses
  - Multi-stage Scalability to 64k edge ports
- Congestion free architecture with dynamic re-routing
  - Always In Order Frame delivery
- Effective Arbitrary Size Crossbar Performance