Testing AI Datacenter Fabric with Emulated Collective Communication Library Workload

Spirent Communications, Akihiro NAKAMURA

Agenda

- 自己紹介
- ネットワークにAIを適用する展望
- AI/ML向けDCネットワークの現状と検証課題
- 検証例の紹介

自己紹介

- 2001年3月　大学卒業
- 2001年4月~2017年3月

  測定器専門商社にて光通信・情報通信関連の製品・ソリューションを販売
  - 新製品の立ち上げや市場開拓、コミュニティ・学会にも積極的に参加
- 2017年4月~　現職
- 好きなスポーツ：野球、ゴルフ
- MPLS Japan 2010ごろ？から参加中

Agenda

- 自己紹介
- **ネットワークにAIを適用する展望**
- AI/ML向けDCネットワークの現状と検証課題
- 検証例の紹介

Proprietary and Confidential

AI's potential impact on **Telco business cases**

$450B - $680B

Value to the global telecommunications industry

- McKinsey | 2024 -

Customer Service	Marketing & Sales	**Network**	IT	Operation

REDUCED OPEX

AI streamlines operations and cuts costs by proactively identifying and addressing inefficiencies, faults, and areas for improvement in real time, leading to more efficient resource utilization and reduced downtime.

REDUCED CAPEX

By enhancing operational efficiency and optimizing resource utilization, AI allows for more strategic and need-based management of equipment upgrades and energy costs, thus minimizing ongoing capital expenditures.

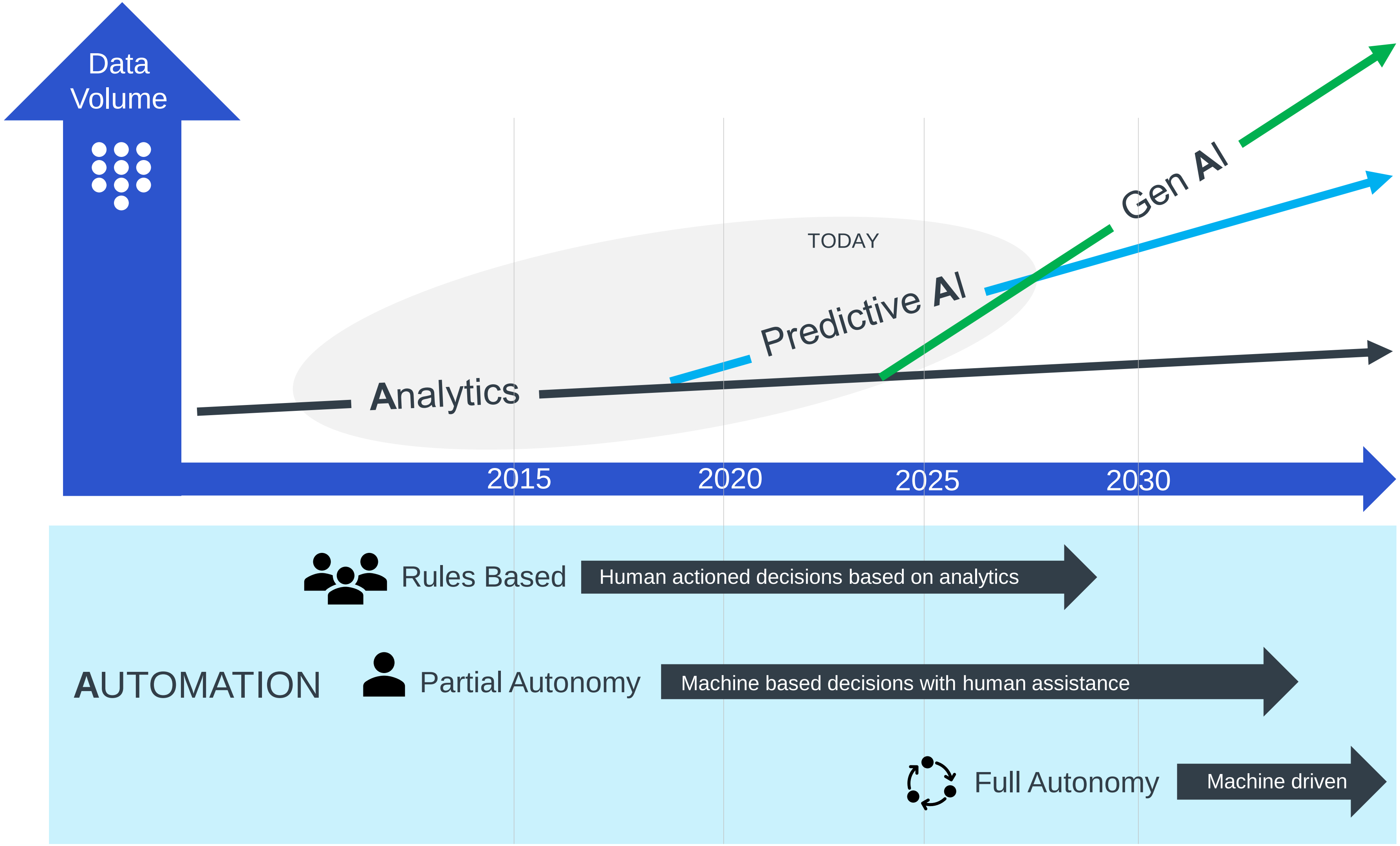
INCREASED CUSTOMER EXPERIENCE

AI enhances customer satisfaction and service quality by analyzing engagement and sentiment, offering proactive recommendations, thereby reducing churn and boosting revenue.

INCREASED REVENUE

AI leverages customer engagement and behavioral trends to uncover new revenue opportunities, enhancing existing services and creating personalized offerings.

The power of 3 (**A**nalytics, **A**utomation & **AI**) over time



Credit: Diagram concept based on Appledore Research "Progression of AI and Computational Resources Demanded"

Use cases of AI on **Network** & **Infrastructure**

Seen across our global telecom operators

in NETWORK management

NETWORK DESIGN	OPERATIONS & MGT.	SECURITY
Traffic prediction	Network optimization	Threat prediction
Capacity planning	Predictive maintenance	Fraud detection
Radio map planning	Fault prediction	Resilience planning
	Anomaly detection	
	Root cause analysis	

in INFRASTRCUTURE

RAN	DEVICES (UE)
Energy optimization	GenAI on device
Spectral efficiency	AI/ML in NR Air Interface
Traffic steering	
Load balancing	
Mobility optimizations	

How **GenAI** will **complement PredAI** in Network Management

Today

Analytics & Predictive AI/ML

Top 5 network mgt. use cases (today):

1. Network management
2. Predictive maintenance
3. Traffic prediction (network planning)
4. Anomaly detection
5. Performance Assurance

+

Evolution

Generative AI

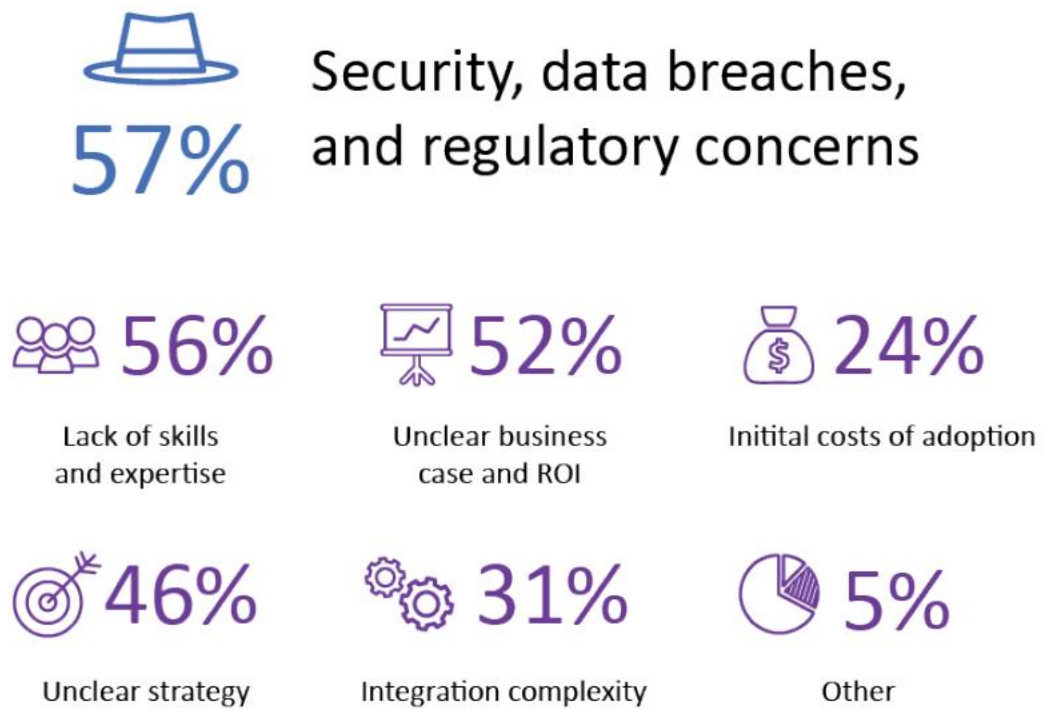
For:	Content creation:	Querying (human-machine):	Supporting predictive models:
Use cases:	• Troubleshooting guides • Incident reports • Network topologies & coverage maps • Configuration scripts • Test traffic generation (for adversarial scenarios)	• Document interrogation (field and support teams) • Consulting historical data (similar issues / resolutions)	• Anomaly detection • Augmenting data sets • Preventing overfitting

Save $30B - $45B / Year

Incremental value to the global telecommunications industry

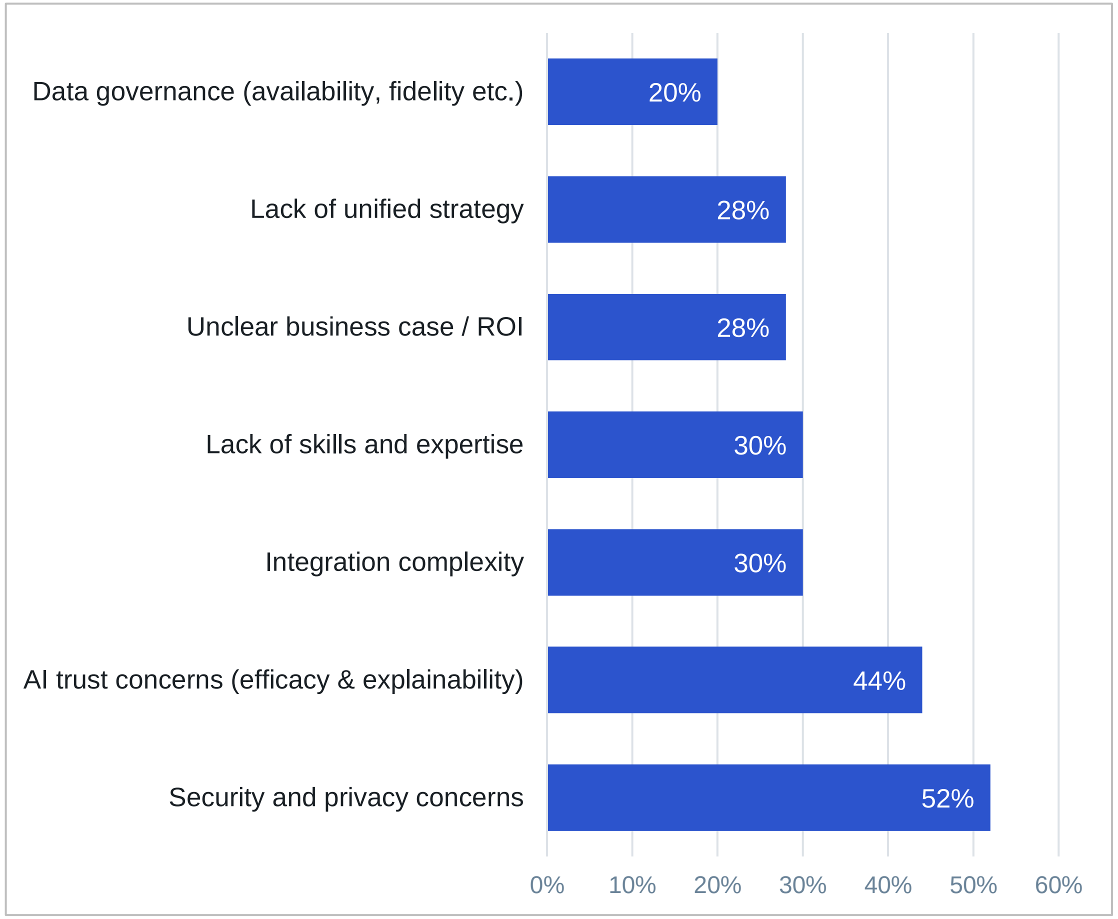
Barriers to adoption

Figure 3: Biggest barriers to widespread AI adoption in telecom



Note: Respondents could select multiple options.
Source: Omdia, Telecoms.com *Annual Industry Survey 2023*, n=87

What key challenges are slowing down the adoption of AI in your organization's network? (Select Top 3)



Spirent / RCR poll | N=50

Case study – Validating Security AI assistants

~99%

Firewall breaches are due to misconfigurations compounded by thousands of overlapping rules & policies

AI

dependency aware assistants

Can optimize & troubleshoot overlapping rules while adding new rules more efficiently = enhanced solution performance & trust.

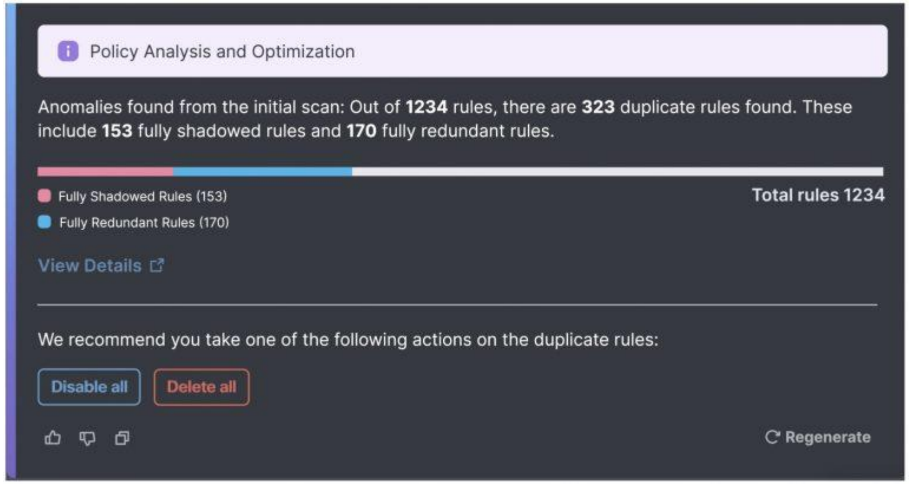
How to validate AI assistant optimized rules & policies do not impact the security or performance of the network

Automated & continuous Security testing

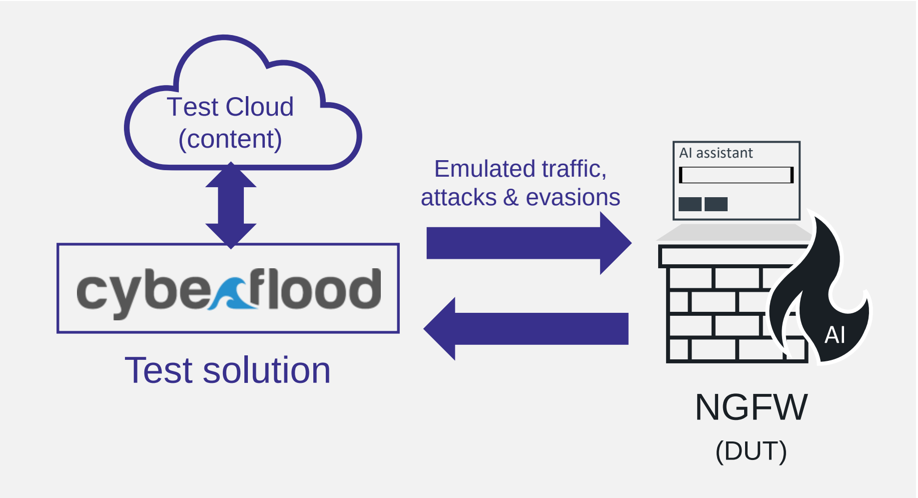
Realistic traffic & attack generation

Validates

- AI changes are implemented correctly.
- Efficacy of AI predictive blocking.
- Traffic performance impact
- Hallucinations or laziness



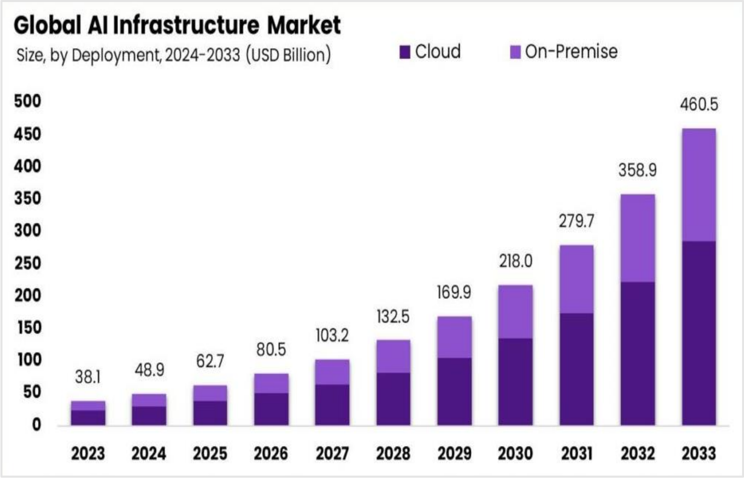
Example: Cisco Firewall AI Assistant



Agenda

- 自己紹介
- ネットワークにAIを適用する展望
- **AI/ML向けDCネットワークの現状と検証課題**
- 検証例の紹介

AI Infra Market Trends



Market Scoop April '24

- Ten-year overall AI infrastructure market expected to grow from $38B in '23 to $460B in 2033.
- Ten-year CAGR of 28.3%

AI Adoption in Finance '24

- 46% are already using large language models (LLMs)

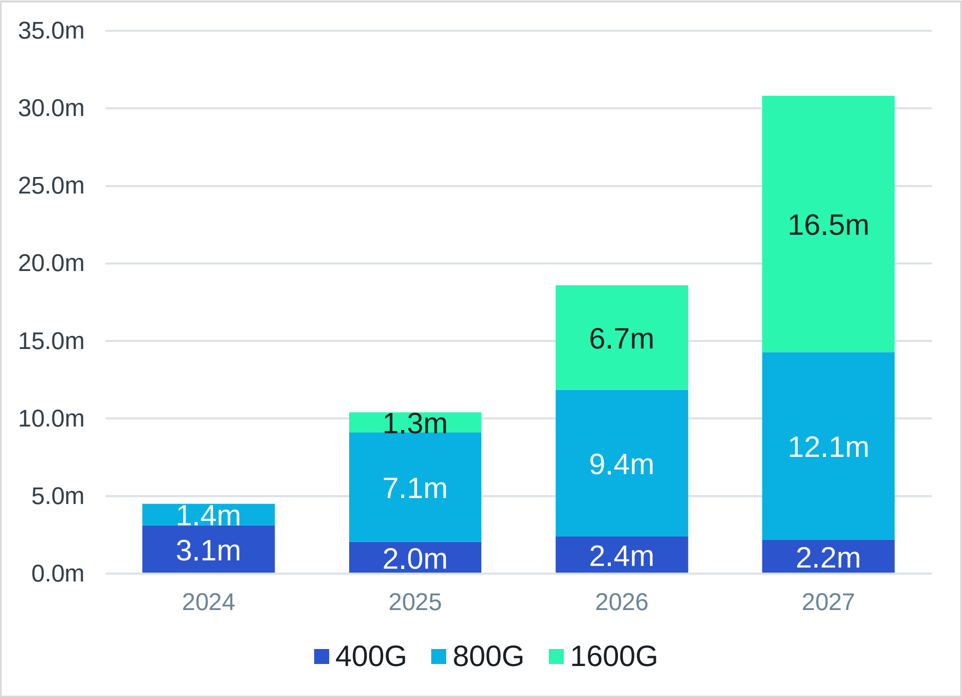
  - NVIDIA 4/24 Survey

- Standards bodies defining next generation network specifications for evolving AI demands

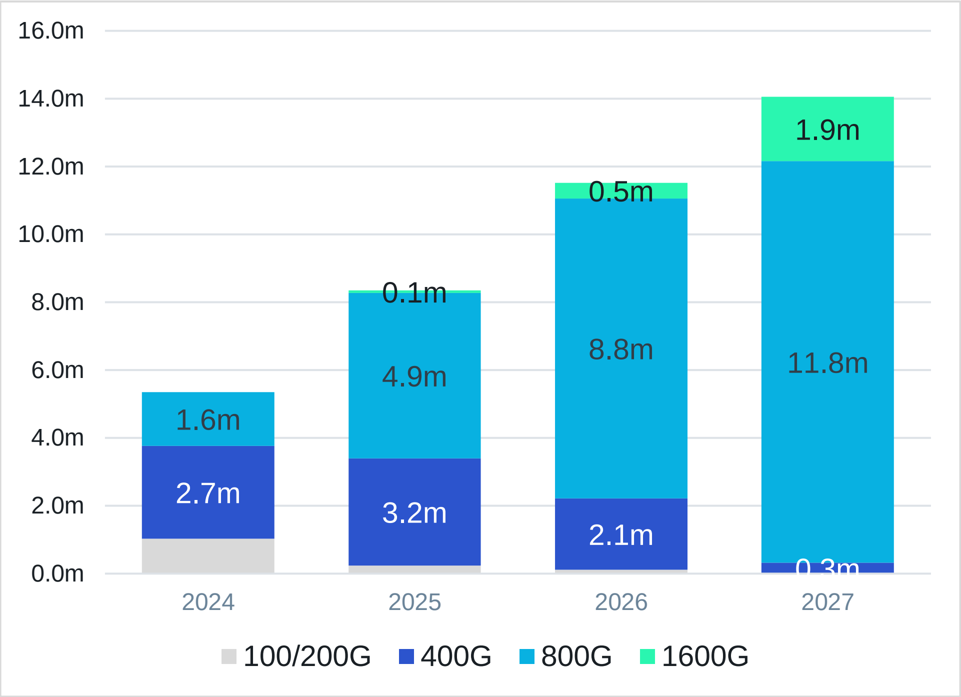
Dell'Oro Datacenter Report Q4 '23

Data Center Networking Link Speed update

AI Back-End Networks
Ethernet speed growth (ports)



Front-End Networks
Ethernet speed growth (ports)



Source: Dell'Oro Q3 2024

- Ethernet revenue share vs InfinBand will grow from 17% in 2023 to **51% in 2028**

Bottlenecks For Training Performance

Model name	Model Size (billion params)	Dataset size (billion tokens)	Training ZFLOPS (10^21)	GPU	GPU FLOPS	GPU utilization	Training time (in weeks)	Number of GPUs
OPT	175	300	430	H100	3,000	0.5	1	474
LLaMA	65	1,400	600	H100	3,000	0.5	1	662
LLaMA 2	34	2,000	400	H100	3,000	0.5	1	441
LLaMA2	70	2,000	800	H100	3,000	0.5	1	882
GPT-3	175	300	420	H100	3,000	0.5	1	463
GPT-4* (est.)	1,500	2,600	31,200	H100	3,000	0.5	1	34,392

Table 3 — GPU cluster size for training various LLM models with H100 GPUs, FP8 arithmetic, and 50% utilization.

GPU fabrics for GenAI workloads By Sharada Yeluri on 8 Dec 2023

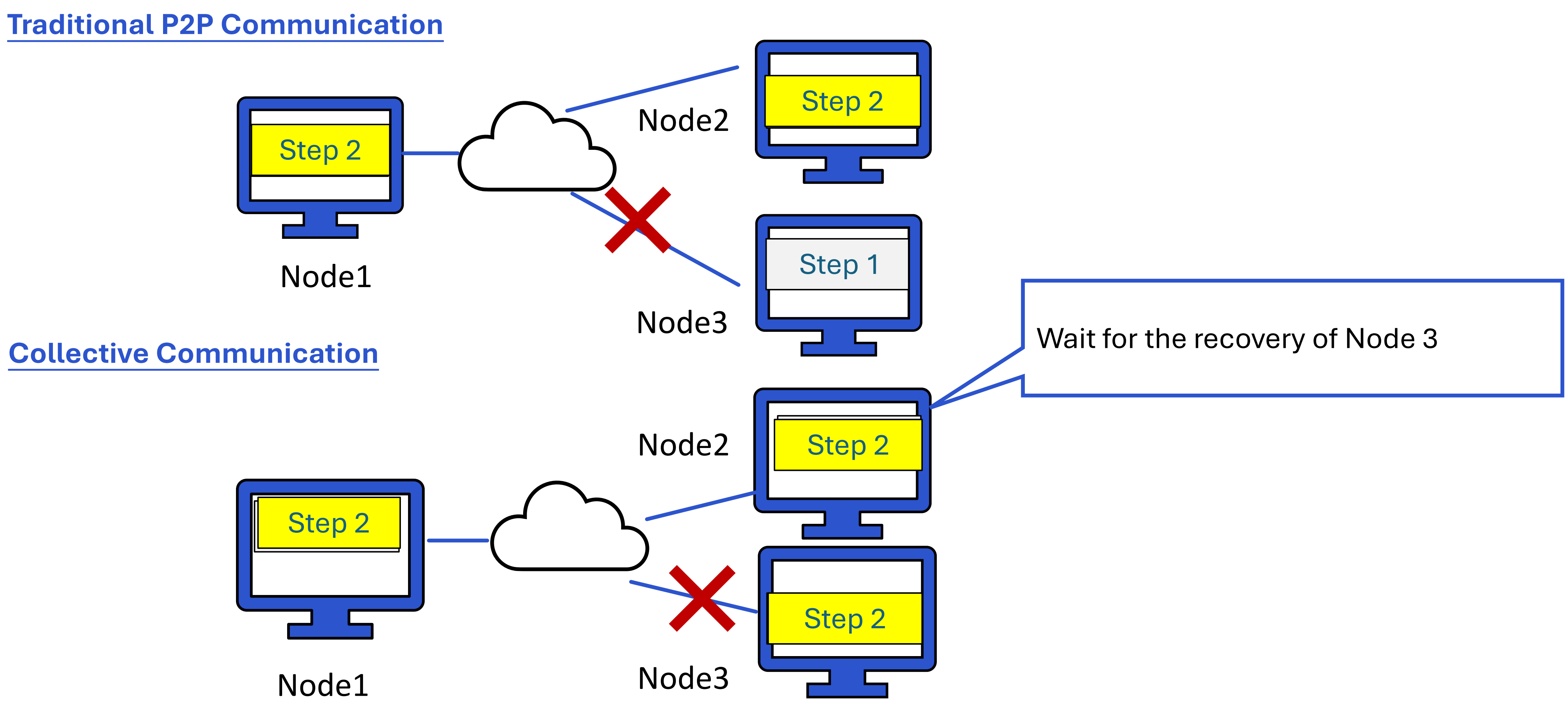
Training Performance Optimization

Memory Access

Network Performance

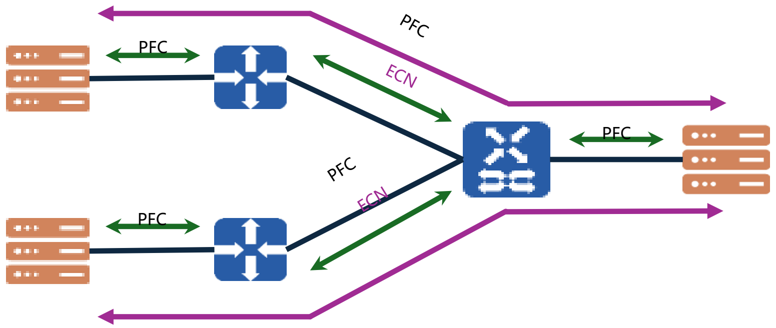
1% packet loss = 30%+ training performance down

Collective Communication Amplifies The Network Bottleneck

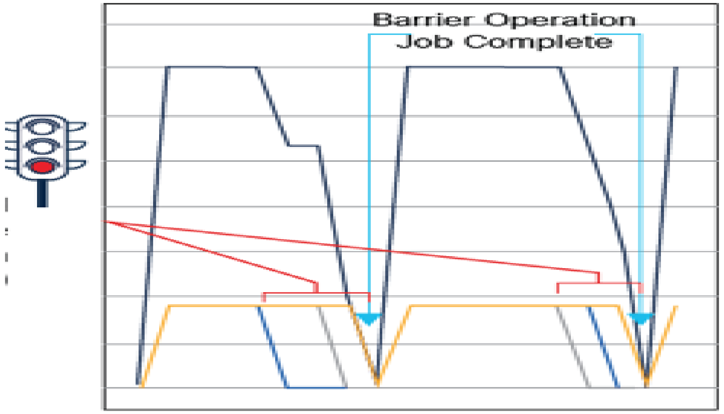


Challenges Of Network For LLM Training

Challenges – Network Congestions

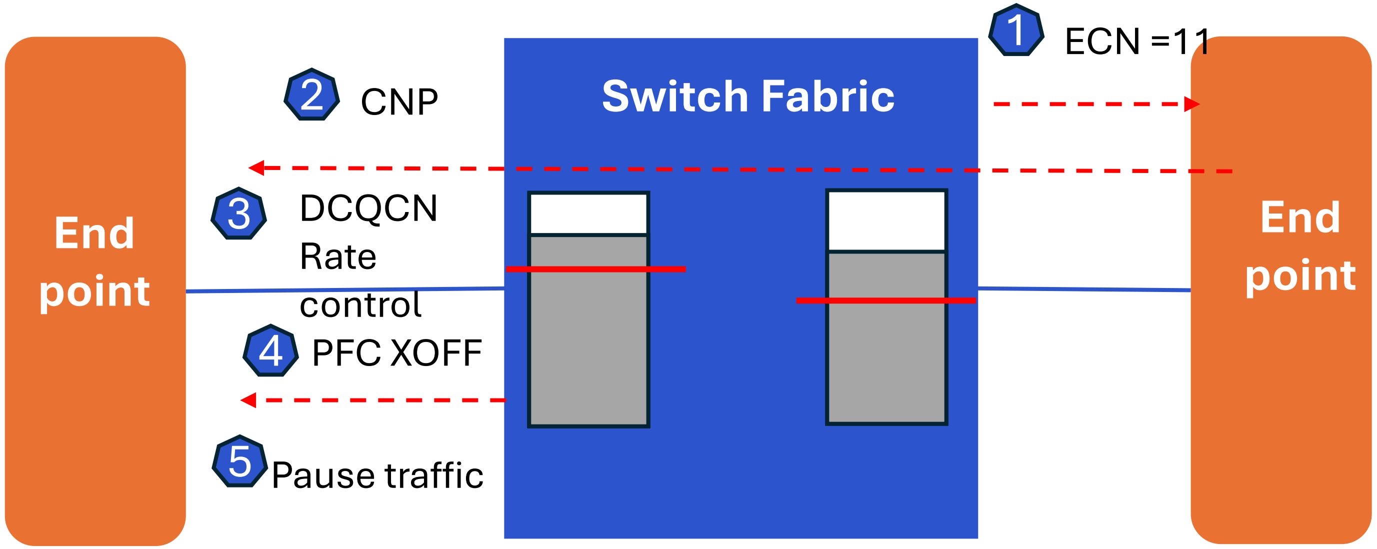


Frequent incast traffic



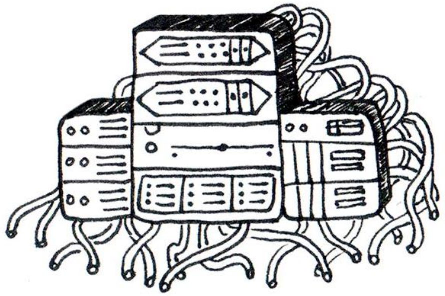
Bursty but elephant flows

Congestion control – DCQCN w/ ECN & PFC



Testing AI Fabric with Real xPU-Based Systems

Takes a datacenter to test a datacenter



Burden is on the end customer at all levels of solution

Requires staff with specific knowledge of AI network landscape, traffic and protocols

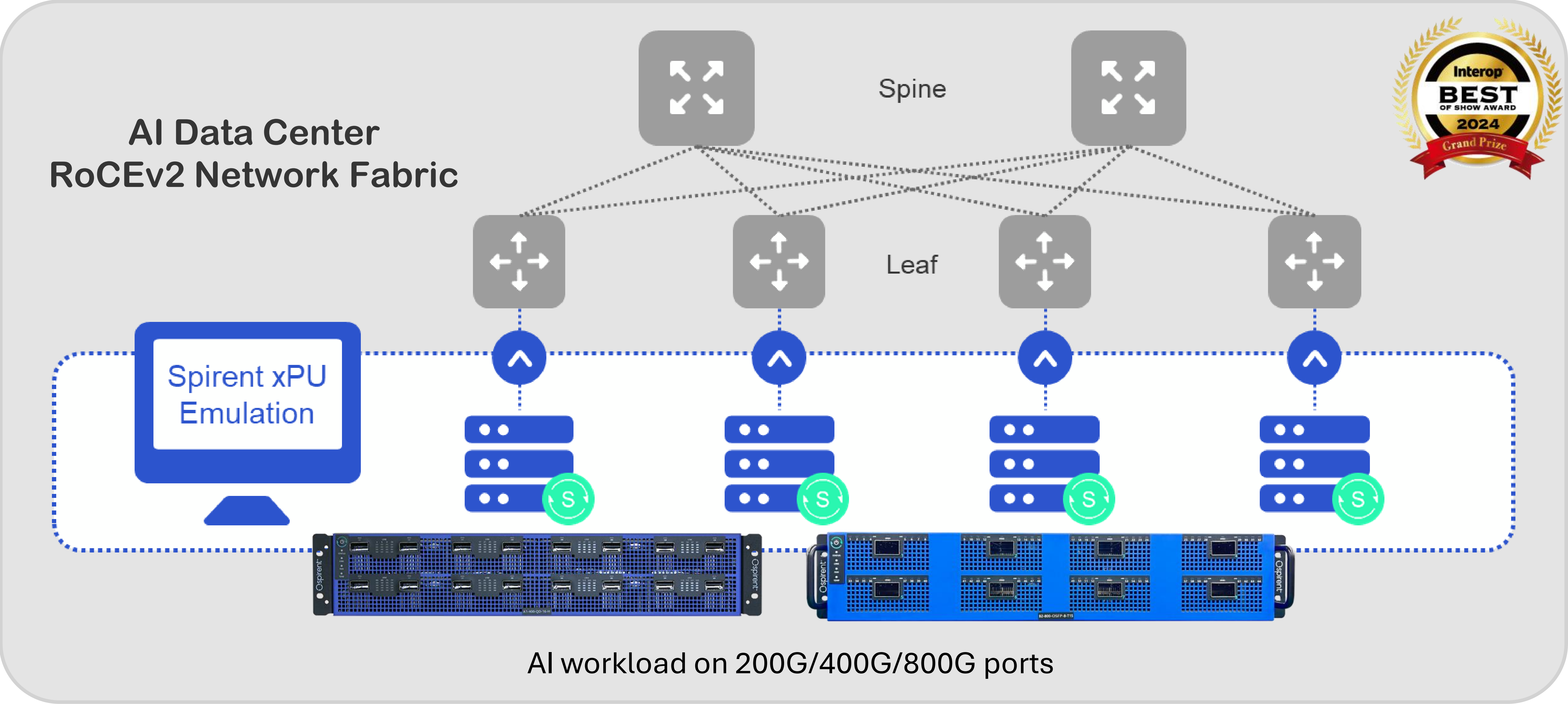
High Cost – Limited Repeatable Testing:

Procurement, Deployment, Management, Test Creation/Execution, Power Consumption, Results/Analytics and Ongoing Test Development

Agenda

- 自己紹介
- ネットワークにAIを適用する展望
- AI/ML向けDCネットワークの現状と検証課題
- **検証例の紹介**

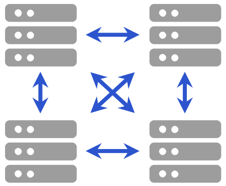
Spirent A1: Hyper-Realistic AI Workload Emulation



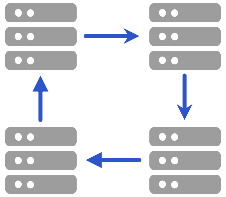
Comprehensive Test Coverage

- Collective Communications Library (CCL) Traffic Patterns
- Congestion Scenarios
- Network Performance
- Service Performance
- Resiliency (Negative Traffic Testing)
- Multi-Speed Roadmap from 100G to 1.6Tb

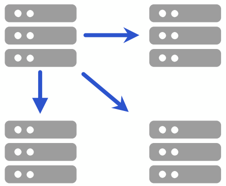
AI Traffic Patterns



All to All



Ring Reduce



Broadcast

and more….

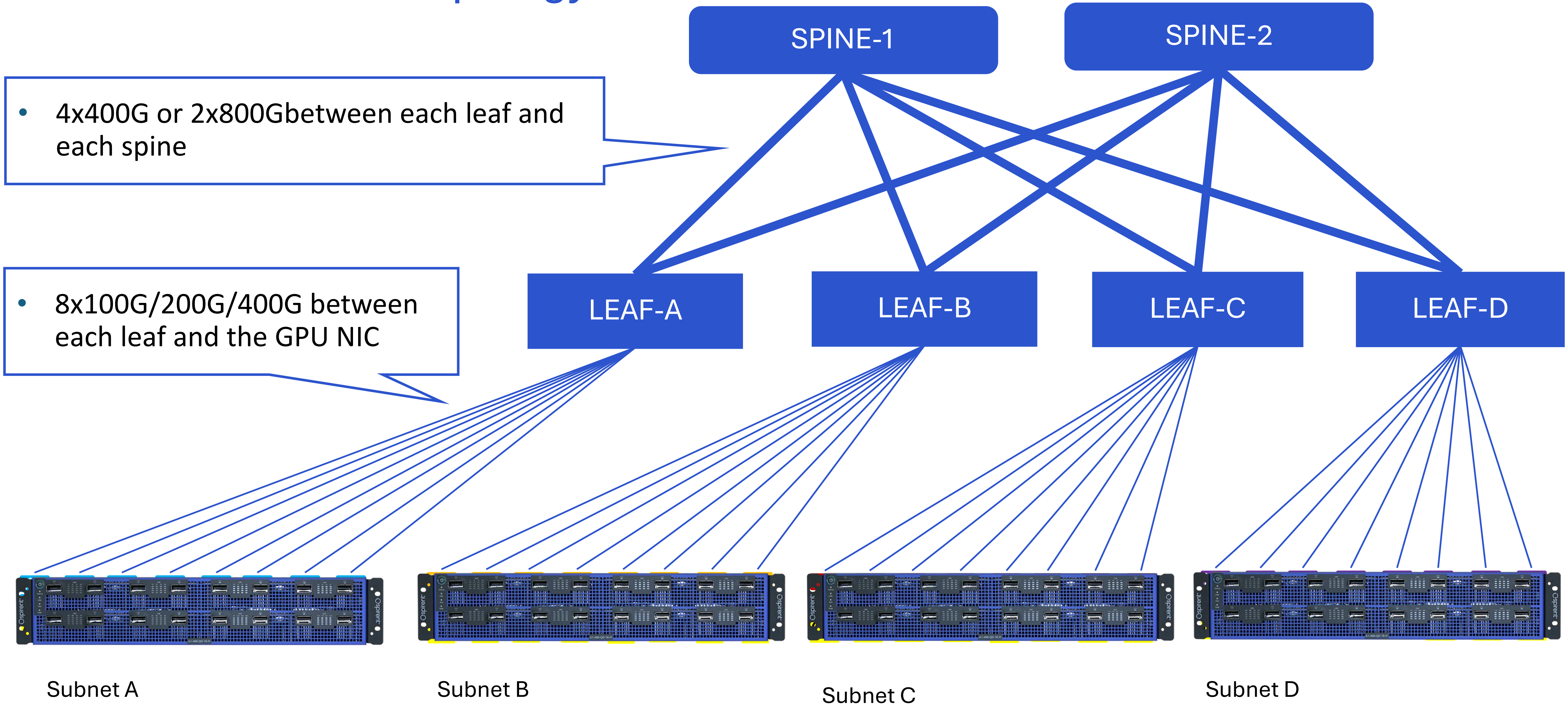
Spirent vs “Do it Yourself”

	Spirent	DIY
Equipment	One Spirent A1 with 16 x 400G	Two Eight-Card xPU Server with 400G NIC Four Eight-card xPU Server with 200G NIC
Material Cost	• Less than a real 16 x port xPU system from NVIDIA. • One time cost.	• Varies per region; as high as $500K+ /server in some regions. • Double the cost every time a new generation GPU releases
Power	4,500W	10,000W+ per server
Expertise/Skills	No additional cost. Current Spirent test experts can use our solution to test the AI fabric. Complexity of in building traffic patterns masked by the RocEv2 & Traffic Wizard.	Cost to train network engineer to have AI knowledge or hire new AI engineers. Complexity in creation of test scenarios.
Maintenance & Development Cost	Minimum maintenance cost.	Maintenance of the xPU test bed (incl. Servers).
Test Cases	**Consistent and Repeatable.**	Inconsistent across multiple runs.
Additional values	General test platform which can be used for AI tests and Routing/Switching tests **greatly improves ROI.**	Single purpose test bed ONLY for AI fabric testing.
Availability	Short lead times. Less than 4 weeks.	Long lead times due to xPU shortages.

A Test Practice – Topology

- 4x400G or 2x800Gbetween each leaf and each spine

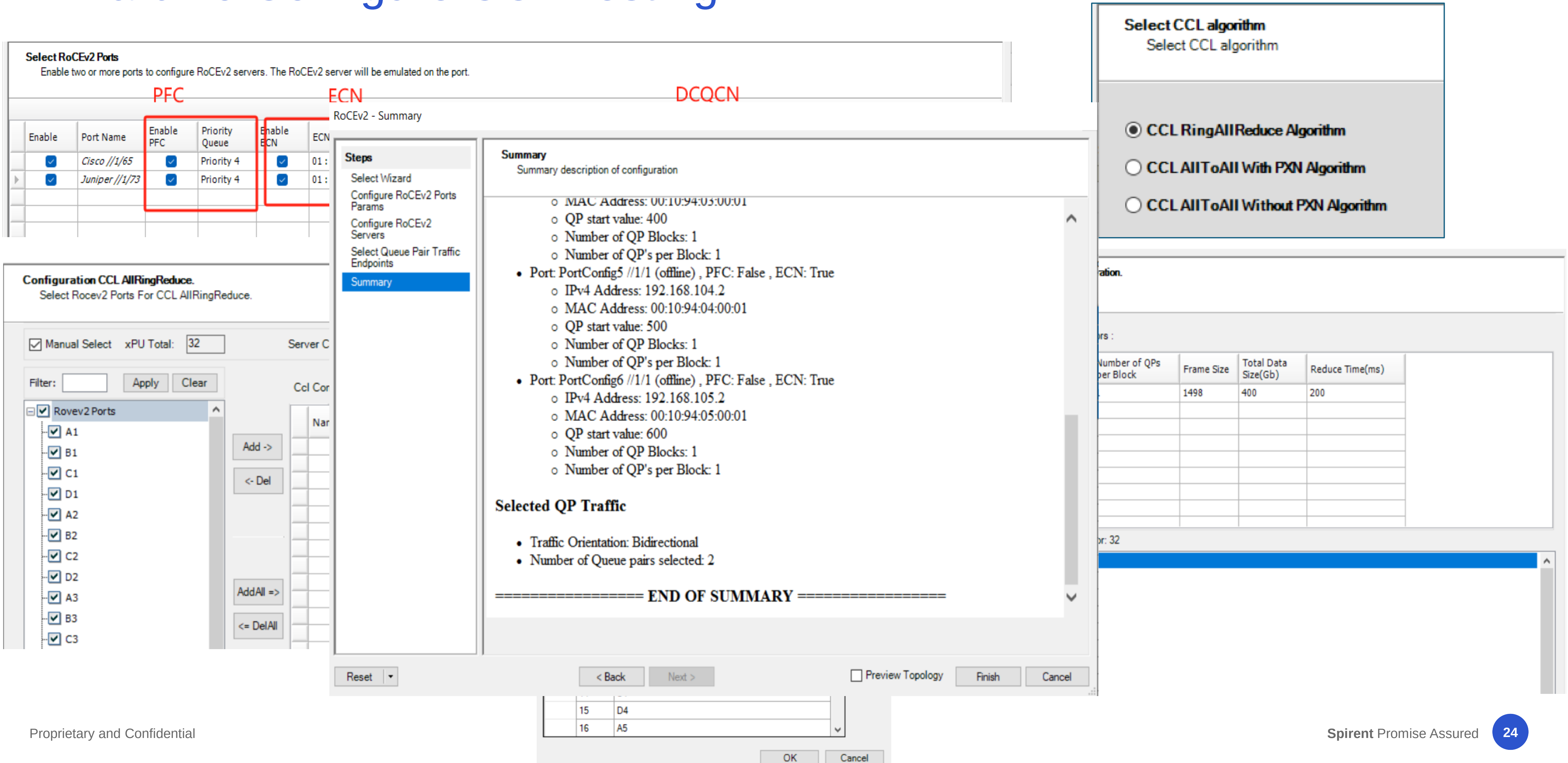
- 8x100G/200G/400G between each leaf and the GPU NIC



A Test Practice – Test Cases

CCL Pattern – 32-port		Scenarios		Parameters
1 x 32-port Ring 8 x 4-port Ring 1 x 32 Alltoall with PXN 4 x 8 Alltoall w/o PXN	X	Benchmarking No congestion, ideal load balancing 80% rate-limiting Imbalance Break fibers	X	Packet size：514B,1582B, 8194B Data size: 128MB-64GB DUT: per flow LB or Packet Spray LB

Wizard To Configure CCL Testing



Results Analysis

Primary Collective KPIs

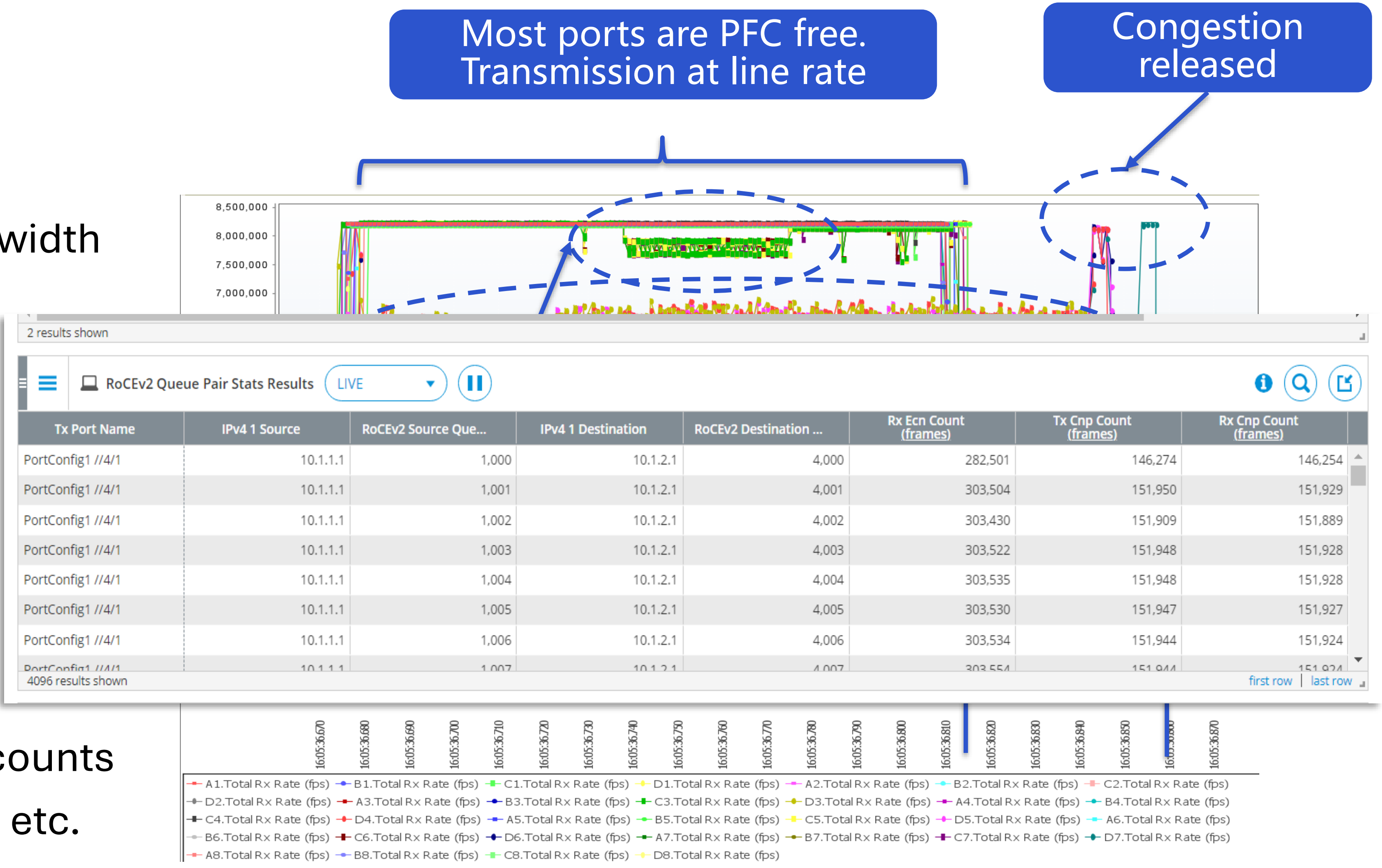
- Algorithm bandwidth, Bus bandwidth

Collective related Stats:

- Transmission Time
- Job Completion Time
- Tail Latency

Network Stats:

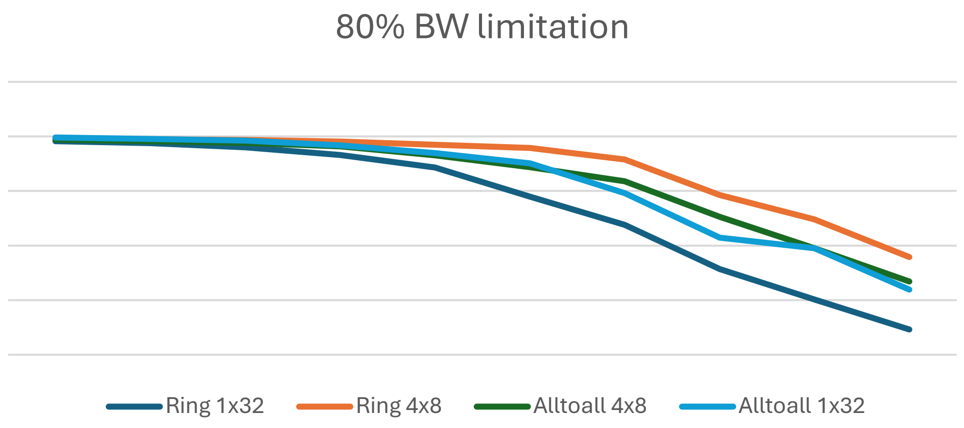
- PFC counts, ECN counts, CNP counts
- Packet loss, Latency, reorder, etc.

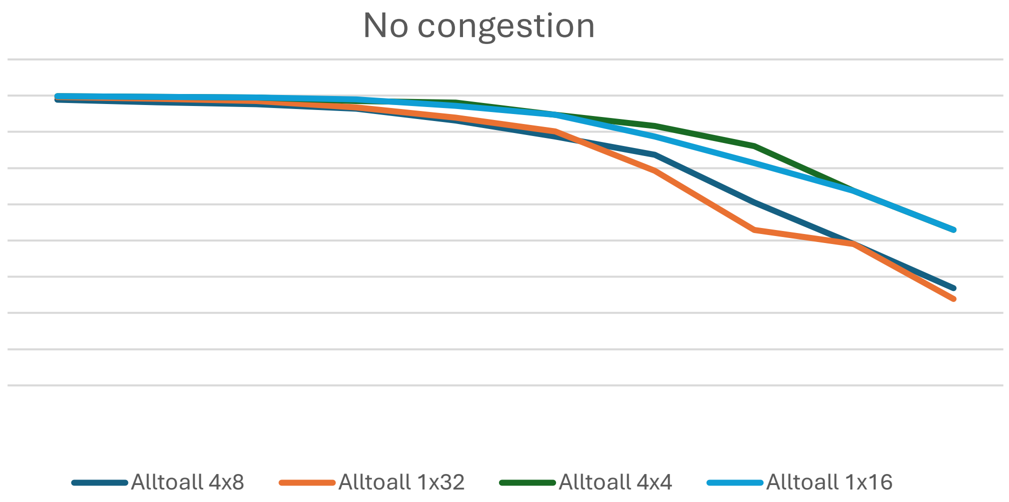


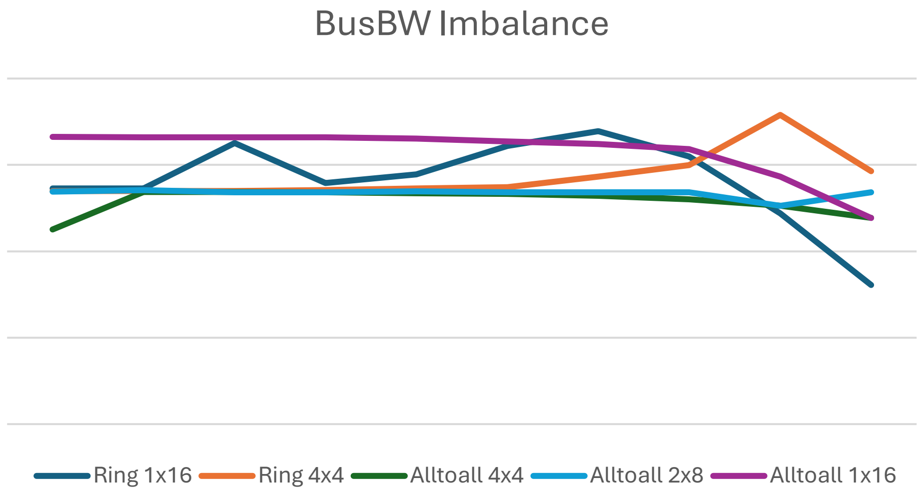
Tx Port Name	IPv4 1 Source	RoCEv2 Source Que...	IPv4 1 Destination	RoCEv2 Destination ...	Rx Ecn Count (frames)	Tx Cnp Count (frames)	Rx Cnp Count (frames)
PortConfig1 //4/1	10.1.1.1	1,000	10.1.2.1	4,000	282,501	146,274	146,254
PortConfig1 //4/1	10.1.1.1	1,001	10.1.2.1	4,001	303,504	151,950	151,929
PortConfig1 //4/1	10.1.1.1	1,002	10.1.2.1	4,002	303,430	151,909	151,889
PortConfig1 //4/1	10.1.1.1	1,003	10.1.2.1	4,003	303,522	151,948	151,928
PortConfig1 //4/1	10.1.1.1	1,004	10.1.2.1	4,004	303,535	151,948	151,928
PortConfig1 //4/1	10.1.1.1	1,005	10.1.2.1	4,005	303,530	151,947	151,927
PortConfig1 //4/1	10.1.1.1	1,006	10.1.2.1	4,006	303,534	151,944	151,924
PortConfig1 //4/1	10.1.1.1	1,007	10.1.2.1	4,007	303,554	151,944	151,924

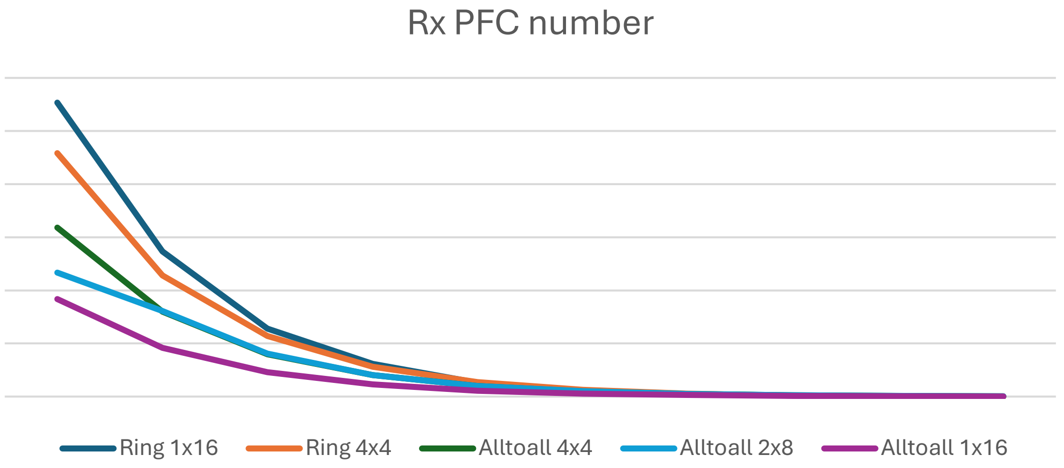
1ms high resolution sampling

A Test Practice – Sample Of Results









Validate the readiness of new Ethernet infrastructures for next-gen AI applications

Thank you.

Akihiro.Nakamura@spirent.com