

# 100 Gigabit Ethernet

Developments and Operational Experience

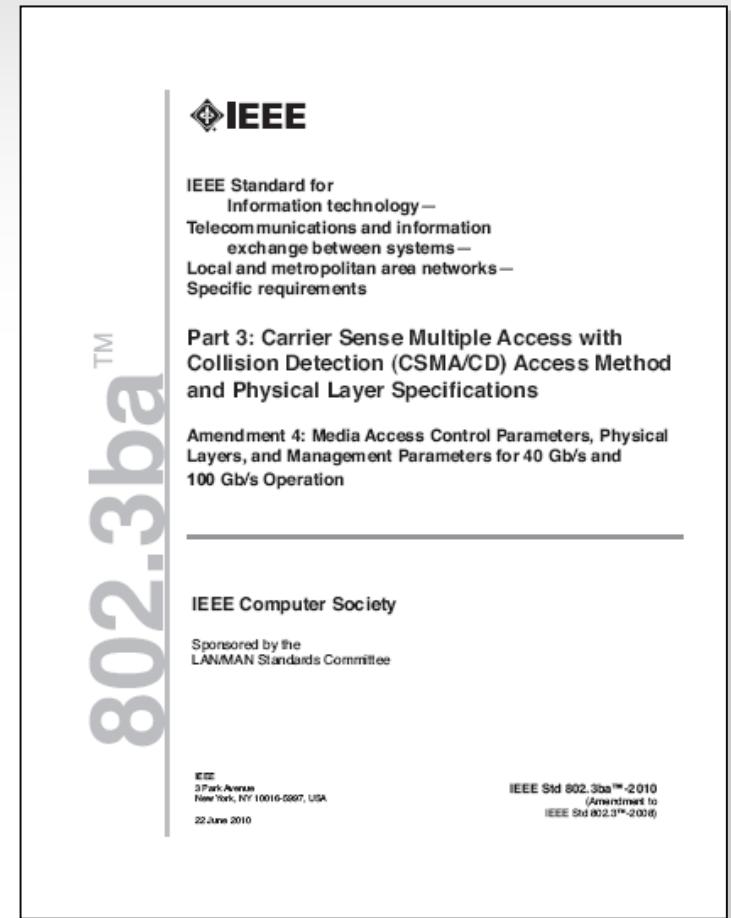
Greg Hankins  
Brocade  
[ghankins@brocade.com](mailto:ghankins@brocade.com)

Martin Pels  
AMS-IX  
[martin.pels@ams-ix.net](mailto:martin.pels@ams-ix.net)

RIPE63, Vienna  
Nov 1<sup>st</sup>, 2011

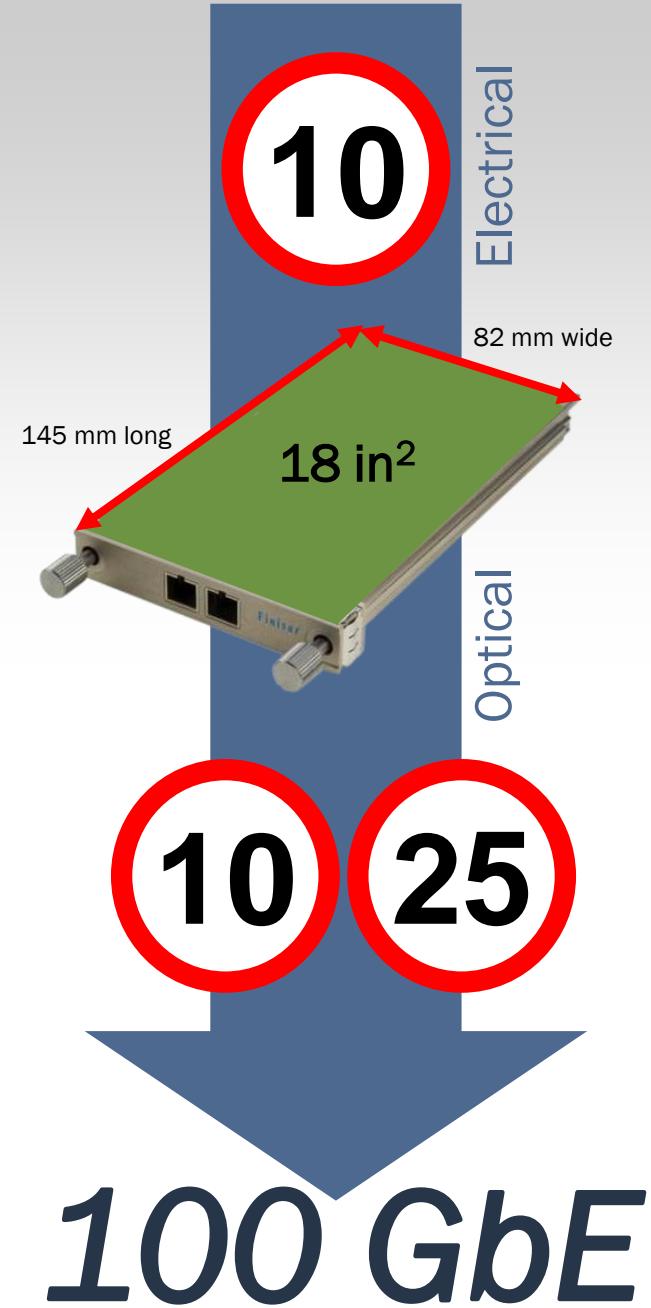
# 40 Gigabit and 100 Gigabit Ethernet Developments

- IEEE 802.3ba standard approved June 17, 2010
  - 457 pages added to IEEE 802.3-2008
- Shipping 1<sup>st</sup> generation media, test equipment, router interfaces, and optical transport gear in 2010/2011
- 2<sup>nd</sup> generation technology projects for both 40 and 100 GbE have started
  - Expected on the market in 2012 - 2013



# Current State of the Industry

- Fundamental 1<sup>st</sup> generation technology constraints limits higher 100 GbE density and lower cost
- Electrical signaling inside the box
  - 100 Gbps Attachment Unit Interface (CAUI) uses 10 x 10 Gbps
- Optical signaling outside the box
  - 10x10 MSA: 10 x 10 Gbps
  - 100GBASE-LR4 and 100GBASE-ER4: 4 x 25 Gbps
- CFP module size and power consumption



# Recent 100 GbE Developments

- 10x10 MSA finished several projects
  - Up to 26 members including AMS-IX, Facebook and Google
  - Initial 10x10-2km standard published in March, 2011
  - Additional 10x10-10km and 10x10-40km standards finished in August, 2011
- 2nd generation projects based on 4 x 25 Gbps electrical signaling have started
- New IEEE P802.3bj 100 Gb/s Backplane and Copper Cable Task Force was started in September, 2011
  - 100GBASE-KR4: 4 x 25 Gbps over >1 m backplane
  - 100GBASE-CR4: 4 x 25 Gbps over >5 m copper twinax cable
  - <http://www.ieee802.org/3/bj/>

# Recent 100 GbE Developments

- New IEEE Next Generation 100Gb/s Optical Ethernet Study Group was started in July, 2011
  - 100GBASE-SR4: 4 x 25 Gbps over 100 m OM3/OM4 MMF
  - 100GBASE-FR4: 4 x 25 Gbps over <2? km SMF
  - CAUI-4: electrical signaling to the CFP2
  - CPPI-4: electrical signaling to the QSFP28 and CFP4
  - QSFP28 and CFP2/4 will be competing for the highest front panel density in 2013+
  - <http://www.ieee802.org/3/100GNGOPTX/index.html>

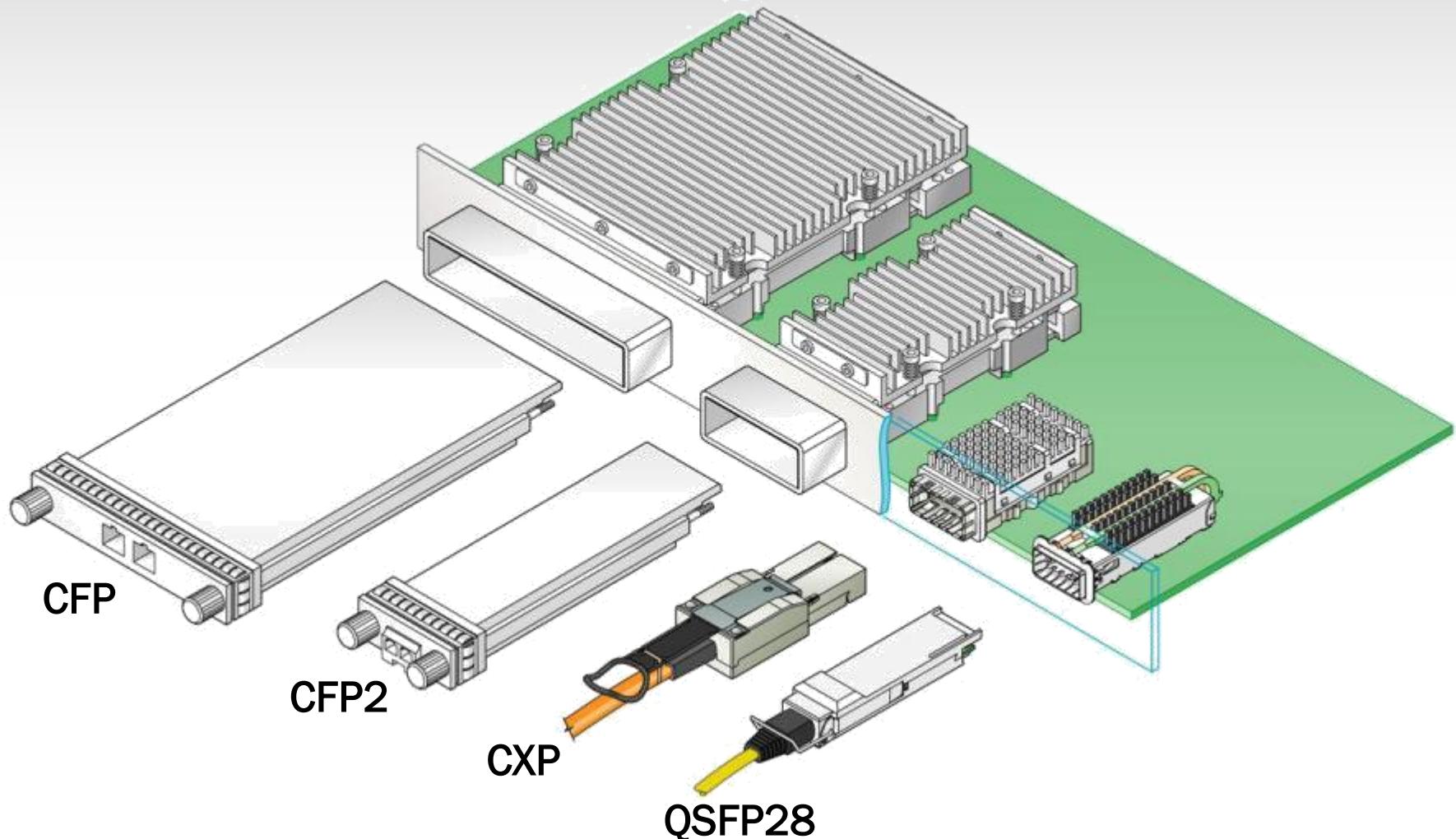
# IEEE Ethernet Bandwidth Assessment

## Ad Hoc

- Laying groundwork and investigating industry interest for the next Ethernet speed
  - Evaluate Ethernet wireline bandwidth requirements
  - Provide data and reference material to the IEEE
  - Gather information only, will not make a recommendation
  - [http://www.ieee802.org/3/ad\\_hoc/bwa/index.html](http://www.ieee802.org/3/ad_hoc/bwa/index.html)
- Network operator input is needed on future requirements
  - Speed, density, distance, cost, topology, anything really
- Please get involved... this means ***you!!***
- Ad Hoc Chair contact:  
John D'Ambrosia, <[jdambrosia@ieee.org](mailto:jdambrosia@ieee.org)>

# 100 Gbps Module Evolution

Graphical View of Module Form Factors





LOW COST 100Gb/s PLUGGABLE OPTICAL TRANSCEIVER



- Current IEEE standards have a gap
  - 100GBASE-SR10 supports up to 150 m on OM4 MMF
  - 100GBASE-LR4 supports up to 10 km on SMF
- 100GBASE-LR4 100 GbE optics are very complex and expensive
- 10x10 MSA bridges the gap
  - Support for 2 km, 10 km and 40 km on SMF
  - Considerably more economical
  - Eliminate expensive components
  - Consume lesser power
- Network operator members!

## Members



[www.10x10msa.org](http://www.10x10msa.org)

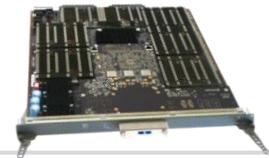
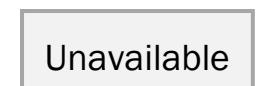
# 100 GbE Market Overview

## CFP Modules

Physical Layer Reach	100 m OM3/ 150 m OM4 MMF	2 km SMF	10 km SMF	40 km SMF		
CFP Module						
Media Type						
Standard	2010 IEEE 802.3ba	March 2011 10x10 MSA	August 2011 10x10 MSA	2010 IEEE 802.3ba	August 2011 10x10 MSA	2010 IEEE 802.3ba
Electrical Signaling (Gbps)	10 x 10	10 x 10	10 x 10	10 x 10	10 x 10	10 x 10
Optical Signaling (Gbps)	10 x 10	10 x 10	10 x 10	4 x 25	10 x 10	4 x 25
Power (W)	~8	~19	~19	~24	TBD	TBD
Target Availability	2012	Now	Now	Now	2012	2012
Relative List Price	€	€€€€€	€€€€€€€€€€	€€€€€€€€€€€€	TBA	TBA

# 100 GbE Market Overview

## Diverse Multivendor Router Interfaces and Media

Vendor	Feature Set	Product Line	CFP Media
Alcatel-Lucent 	L2, IP, MPLS	7450 ESS, 7750 SR	 10x10-10km, 100GBASE-LR4
BROCADE 	L2, IP, MPLS	MLX/XMR, MLXe	 10x10-2km, 10x10-10km, 100GBASE-LR4
CISCO 	IP, MPLS	CRS-3	 100GBASE-LR4
	L2, IP, MPLS	ASR 9000	 100GBASE-SR10, 100GBASE-LR4
HUAWEI 	IP, MPLS	NE5000E	 Unavailable
JUNIPER NETWORKS 	IP, MPLS	T1600, TX Matrix Plus	 100GBASE-LR4

1<sup>st</sup> Generation IEEE  
1<sup>st</sup> Generation 10x10 MSA  
2<sup>nd</sup> Generation IEEE

# 100 GbE Technology Reference

Physical Layer Reach	>1 m Back-plane	>5 m Copper Cable	7 m Copper Cable	100 m OM3/OM4 MMF	100 m OM3, 150 m OM4 MMF	<2? km SMF	2 km SMF	10 km SMF	40 km SMF		
Name	100GBASE-KR4	100GBASE-CR4	100GBASE-CR10	100GBASE-SR4	100GBASE-SR10	100GBASE-FR4	10x10-2km	10x10-10km	100GBASE-LR4	10x10-40km	100GBASE-ER4
Standard Status	Future IEEE 802.3bj	Future IEEE 802.3bj	2010 IEEE 802.3ba	Possible Future IEEE	2010 IEEE 802.3ba	Possible Future IEEE	March 2011 10x10 MSA	August 2011 10x10 MSA	2010 IEEE 802.3ba	August 2011 10x10 MSA	2010 IEEE 802.3ba
Generation	2 <sup>nd</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	1 <sup>st</sup>	1 <sup>st</sup>	1 <sup>st</sup>	1 <sup>st</sup>
Electrical Signaling (Gbps)	4 x 25	4 x 25	10 x 10	4 x 25	10 x 10	4 x 25	10 x 10	10 x 10	10 x 10	10 x 10	10 x 10
Media Signaling (Gbps)	4 x 25	4 x 25	10 x 10	4 x 25	10 x 10	4 x 25	10 x 10	10 x 10	4 x 25	10 x 10	4 x 25
Media Type	Backplane	Twinax	Twinax	MPO MMF	MPO MMF	Duplex SMF	Duplex SMF	Duplex SMF	Duplex SMF	Duplex SMF	Duplex SMF
Media Module	Backplane	QSFP28, CFP2, CFP4	CXP	QSFP28, CFP2, CFP4	CXP, CFP	QSFP28, CFP2, CFP4	CFP	CFP	CFP	CFP	CFP
Availability	2014	2014	2010	2013+	2010	2013+	Q1 2011	Q3 2011	2010 (CFP2 in 2013+)	Q3 2011	2012

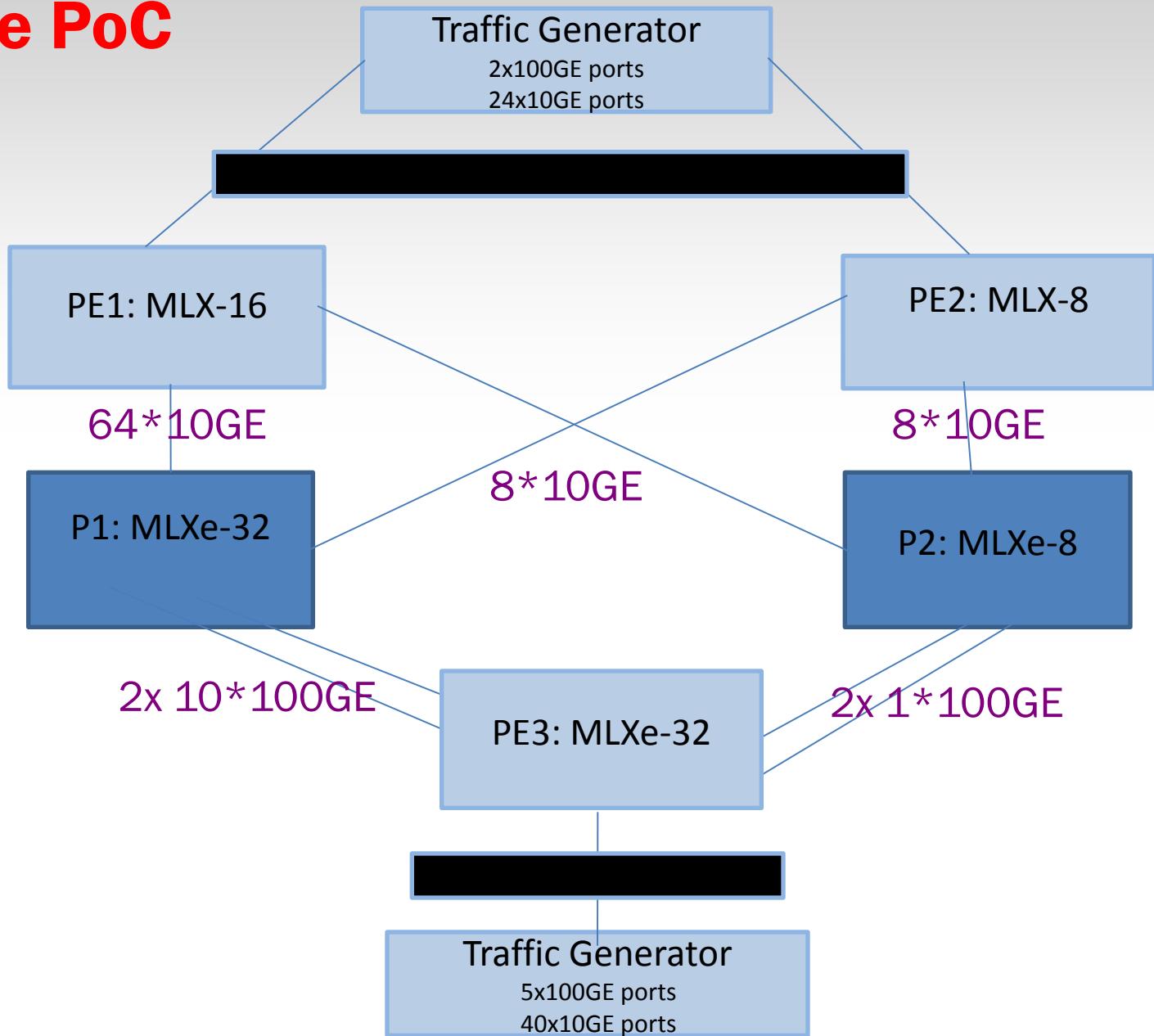
# Operational experiences

- Brocade PoC
- Customer trial
- Metro-area tests
- Long distance trial



# Brocade PoC

July 2011



# Brocade PoC

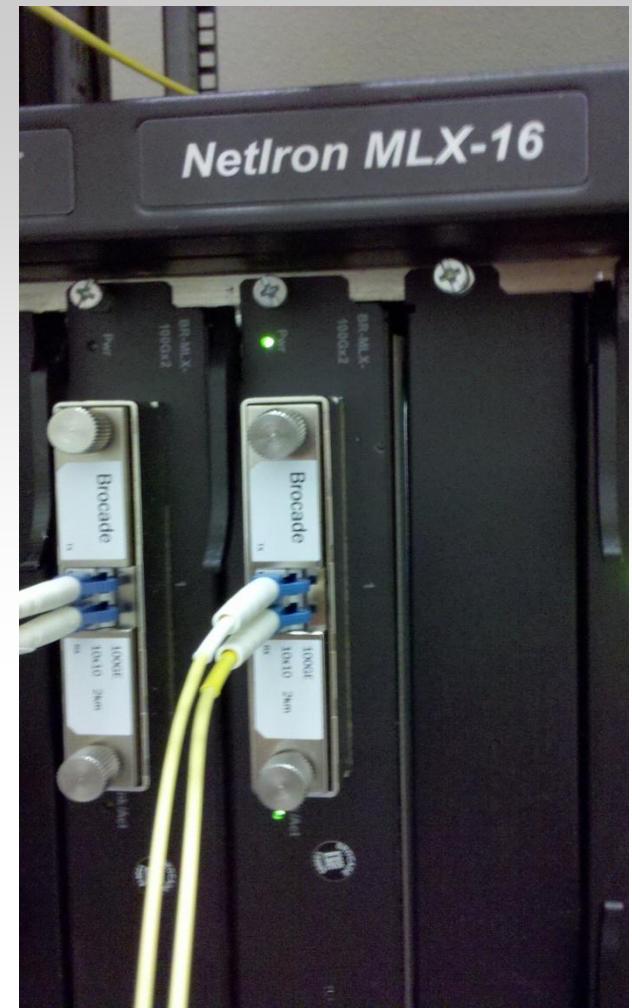
July 2011

- High card density (fully loaded MLXe-32)
- Large aggregates (64 \* 10GE, 10 \* 100GE)
- High traffic volume (4.6Tbit/chassis)
- Several bugs found
- Verified fixes in PoC setup and AMS-IX lab

# Customer trial (Limelight)

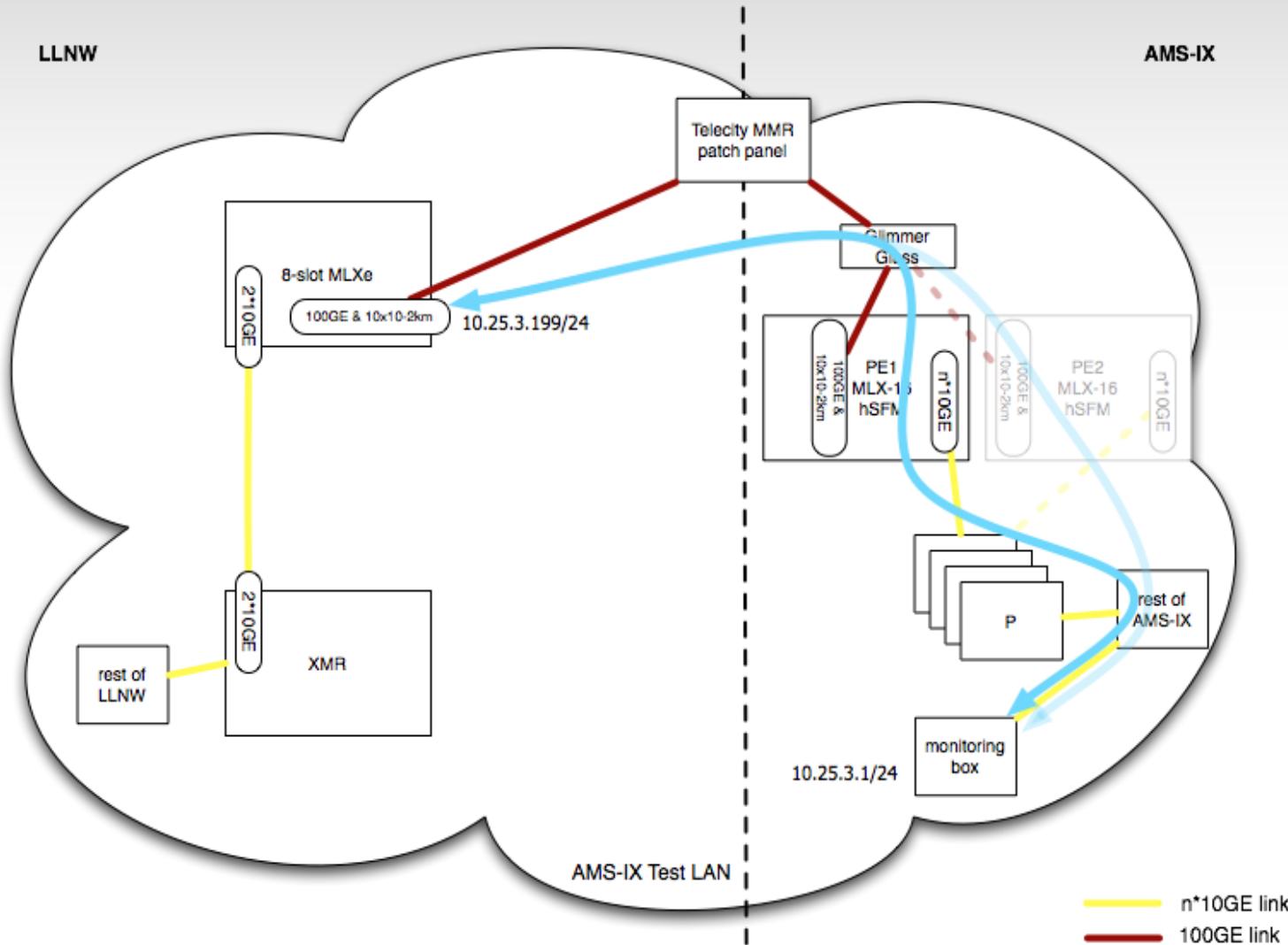
May 2011 - current

- Brocade MLXe 8-slot at LLNW
  - High-speed Switch Fabric Modules
  - 2 \* 10GE to LLNW backbone
- 2 \* MLX-16 at AMS-IX
  - High-speed Switch Fabric Modules
  - 16 \* 10GE to AMS-IX backbone
- 3 \* 2-port 100GE-modules
- 4 \* 10x10-2km CFP optics
- 2 \* 100GE Anritsu traffic generator
- 1 \* 10GE Anritsu traffic generator



# Customer trial (Limelight)

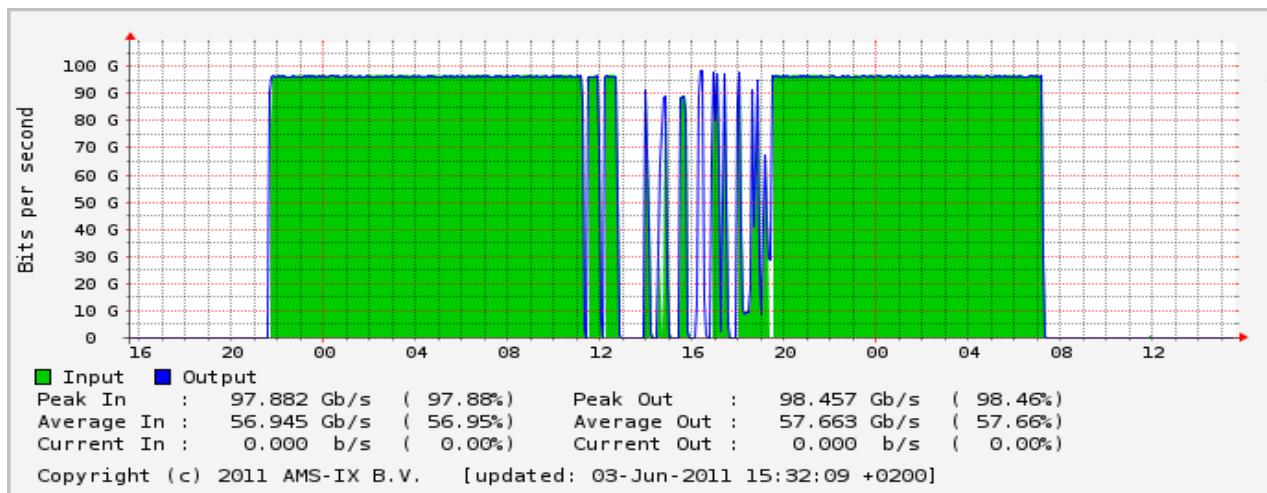
May 2011 - current



# Customer trial (Limelight)

## Tests

- L2 forwarding
- Routing IPv4
- AMS-IX topology failovers
- Routing IPv6
- Several revisions of soft- and hardware



# **Customer trial (Limelight)**

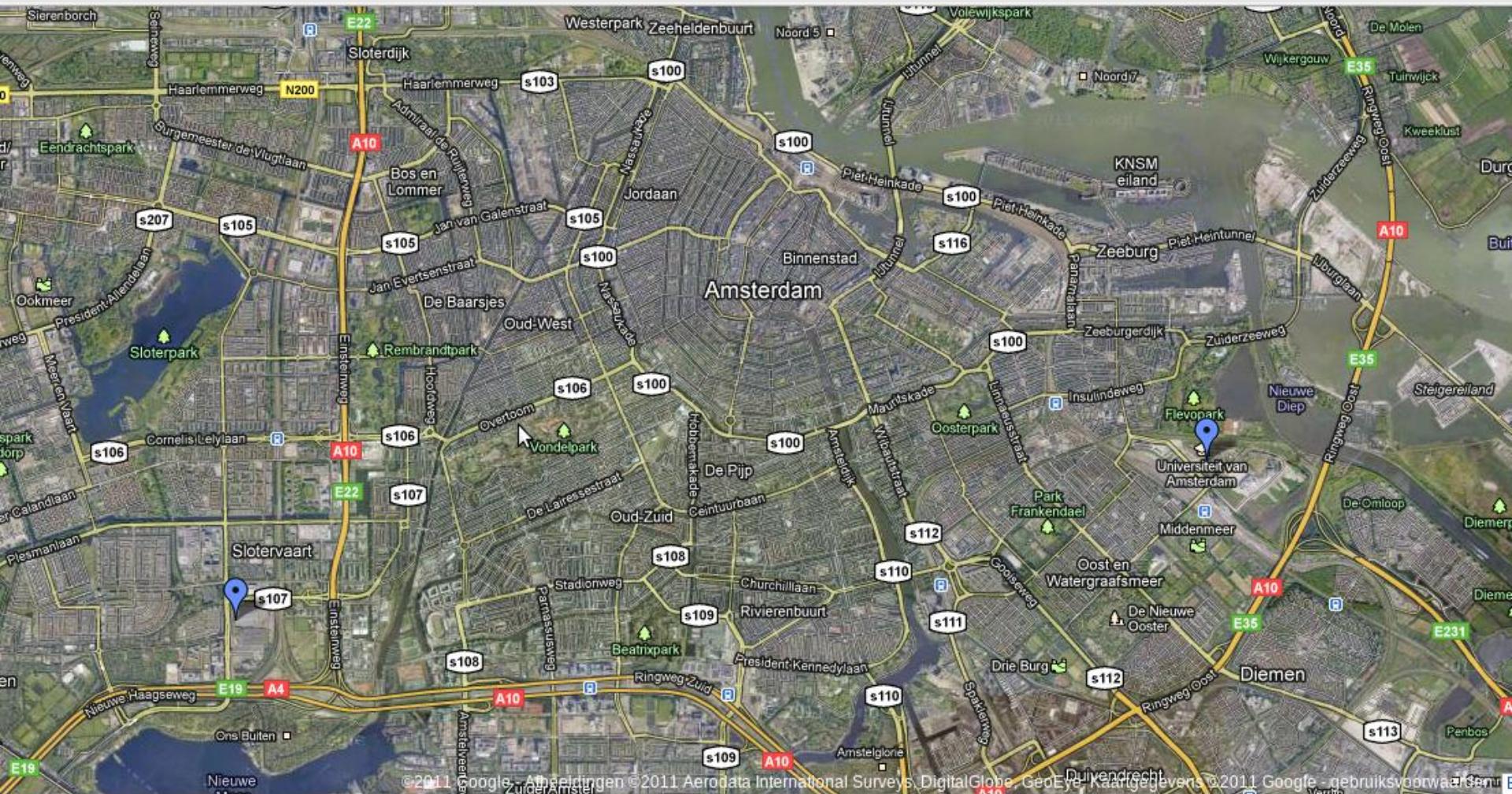
## Results

- Line-rate forwarding for Internet traffic mixes
- Line-rate routing for Internet traffic mixes
- AMS-IX topology failovers: ~190ms average service interruption
- sFlow sampling (up to 2048)
- Jumbo frames (only briefly tested)

**In service!**

# Metro-area tests

10 <> 40km



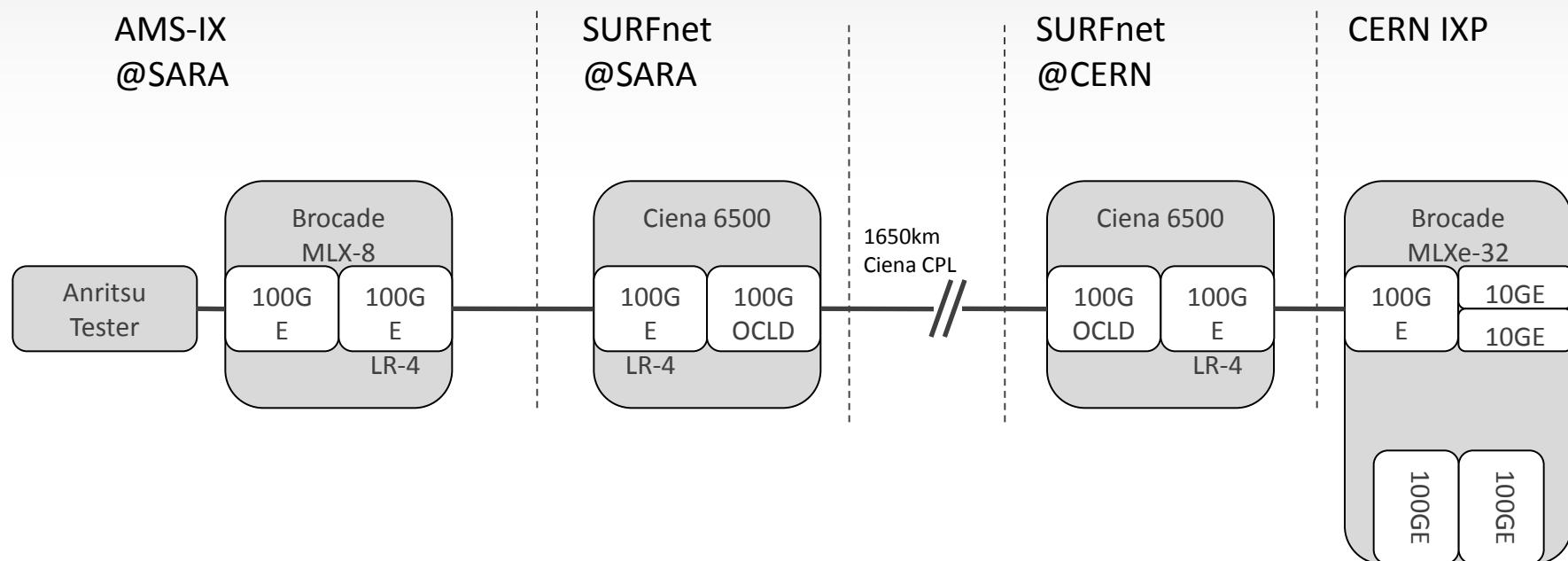
# Metro-area tests

10 <> 40km

- Needed for backbone connections
- 100GBASE-ER4 not ready yet
- LR4 (1300nm range)
  - Bridged 28km (16dB attenuation) using 75mA SOA
  - Looking for production-quality solutions
- 10x10 (1550nm range)
  - SOA out of range (higher channels)
  - EDFA out of range (lower channels)
  - Open for suggestions

# Long distance trial

August 2011



# Long distance trial

August 2011

- Joint trial with SURFnet, CERN, Ciena, Brocade
- 1650km DWDM circuit (Amsterdam - Geneva)
- Brocade MLX / MLXe
- Ciena OME6500
- 100GBASE-LR4 optics



# Long distance trial

## Results

- 10GE loop at CERN, 10Gbit/s traffic load
  - 24 hour no frame loss
- 100GE loop at CERN, 100Gbit/s traffic load
  - 24 hour no frame loss on long distance part
  - MLXe throughput within IEEE-defined BER ( $10^{-12}$ )

# **What's next?**

- Deployment in backbone (Q1 2012)
- Multi-vendor customer trials

**Ready for customers!**

# Questions?