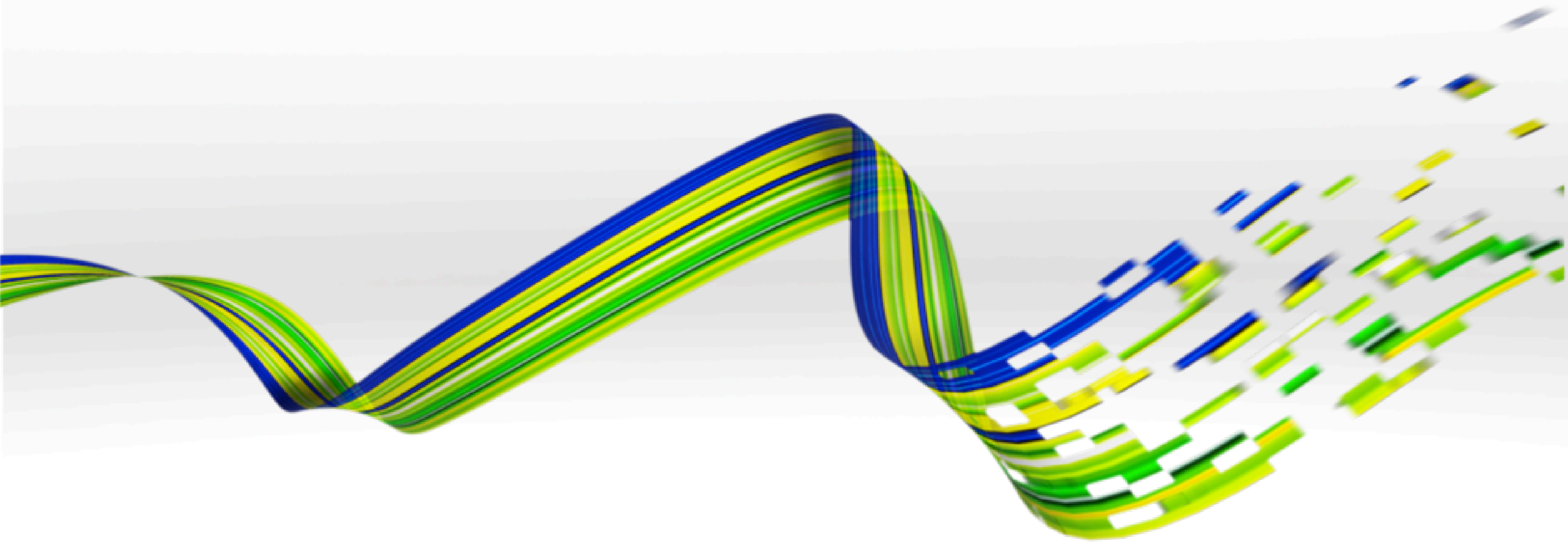
Testing First Generation 100GE Hardware

Elisa Jasinska
elisa@llnw.com

Limelight NETWORKS

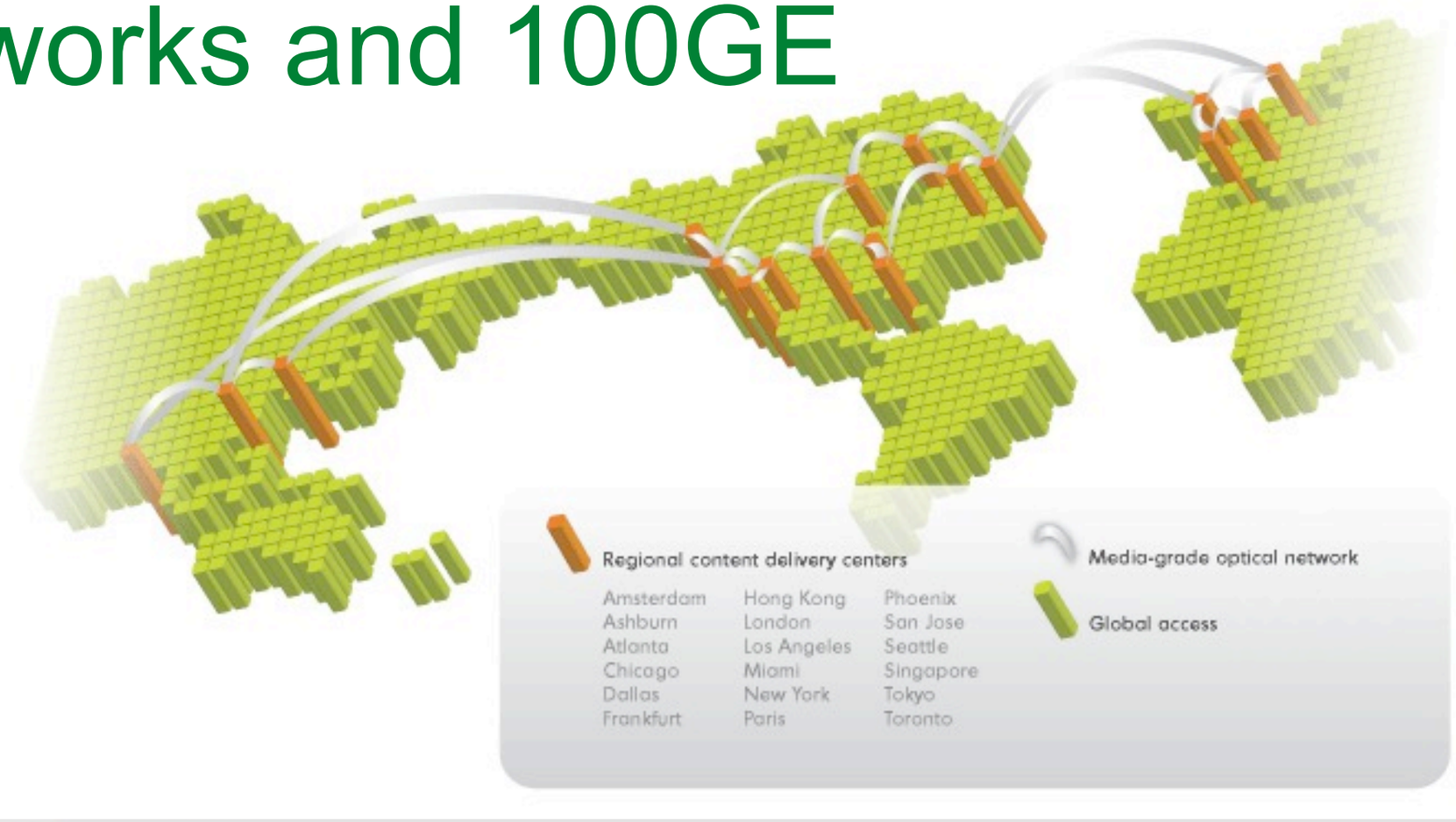
The Need for 100GE



Limelight NETWORKS

Limelight Networks and 100GE

- Global Network
- Present at 50 locations
- Peering at 47 Internet Exchanges



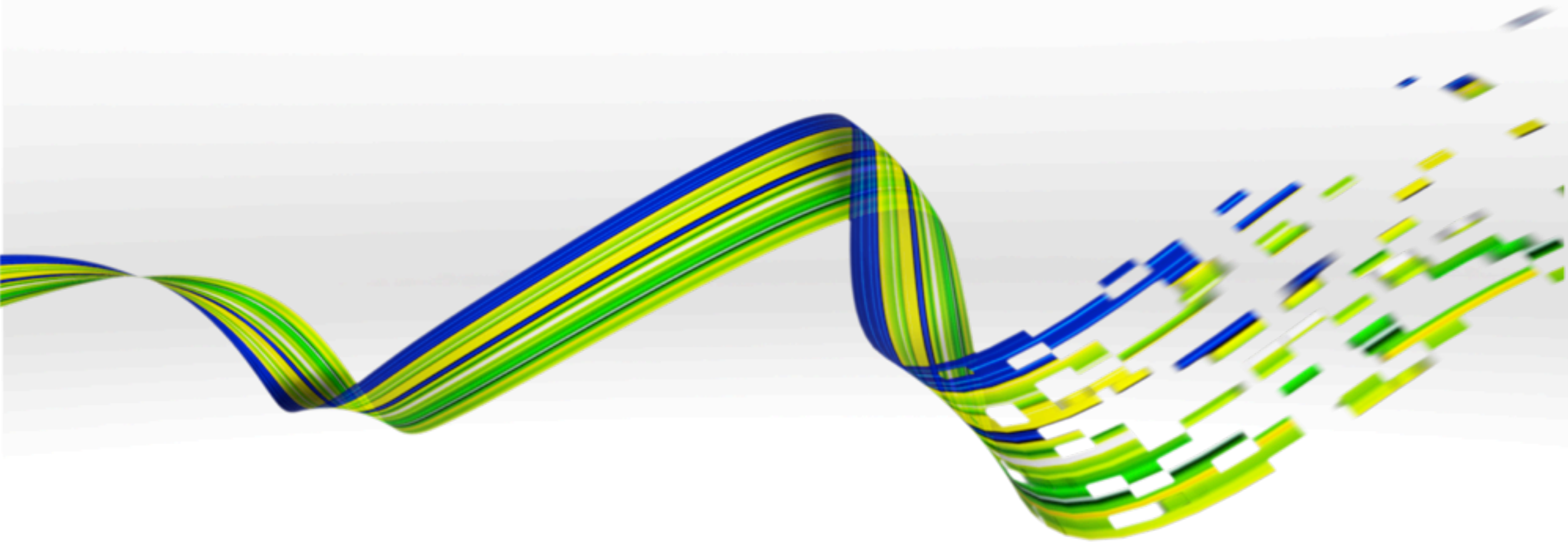
- Higher capacity interfaces beyond 10 GbE needed
  - Limitation on number of links in LAGs
  - Port density (limitation on number of LAGs)
  - Hashing algorithms
  - Large flow distribution

Limelight NETWORKS

Use of 100GE

- Connectivity between various facilities in a city
  - Using optical gear and dark fiber
- Connectivity between routers in a particular facility
  - To maintain core capacity
- Peering exchanges
- PNIs with larger networks
- Backbone between city pairs

A Brief History of 100GE



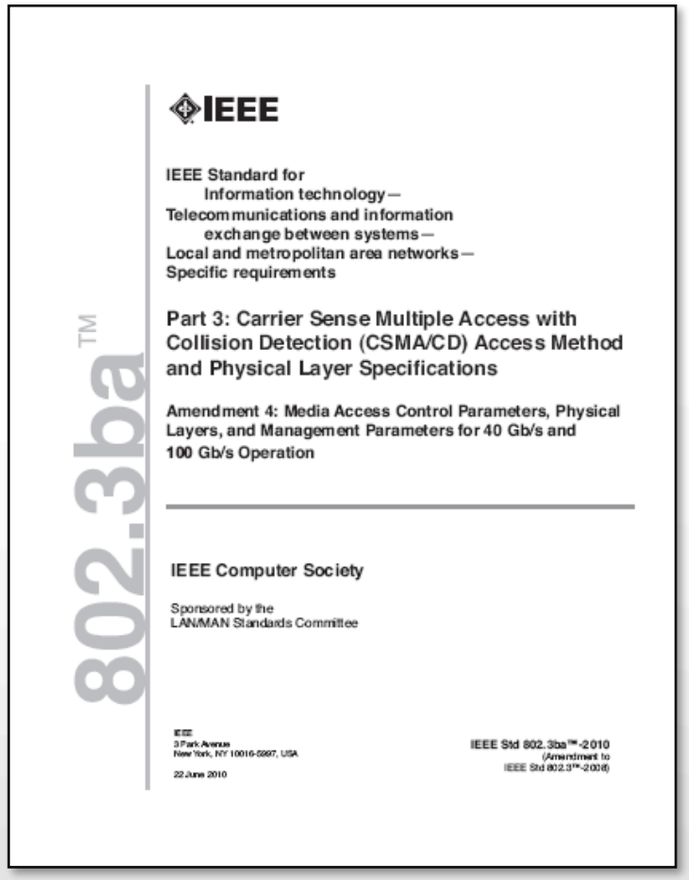
Limelight NETWORKS

IEEE 802.3ba

- Higher Speed Study Group (HSSG) in 2006
- Closed in December 2007
- Transition into IEEE 802.3ba Task Force

- IEEE 802.3ba (approved June 17th, 2010)
- 1st generation equipment 2010/2011
- 2nd generation equipment 2012/2013

- Current IEEE standards
  - 100GBase-CR10 supports up to 7m on Copper
  - 100GBase-SR10 supports up to 150m on OM4 MMF
  - 100GBase-LR4 supports up to 10km on SMF
  - 100GBase-ER4 supports up to 40km on SMF


IEEE

IEEE Standard for
Information technology—
Telecommunications and information
exchange between systems—
Local and metropolitan area networks—
Specific requirements

Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications

Amendment 4: Media Access Control Parameters, Physical Layers, and Management Parameters for 40 Gb/s and 100 Gb/s Operation

802.3ba™

IEEE Computer Society

Sponsored by the
LAN/MAN Standards Committee

IEEE
3 Park Avenue
New York, NY 10016-5997, USA

22 June 2010

IEEE Std 802.3ba™-2010
(Amendment to
IEEE Std 802.3™-2008)

Limelight NETWORKS

10x10 MSA

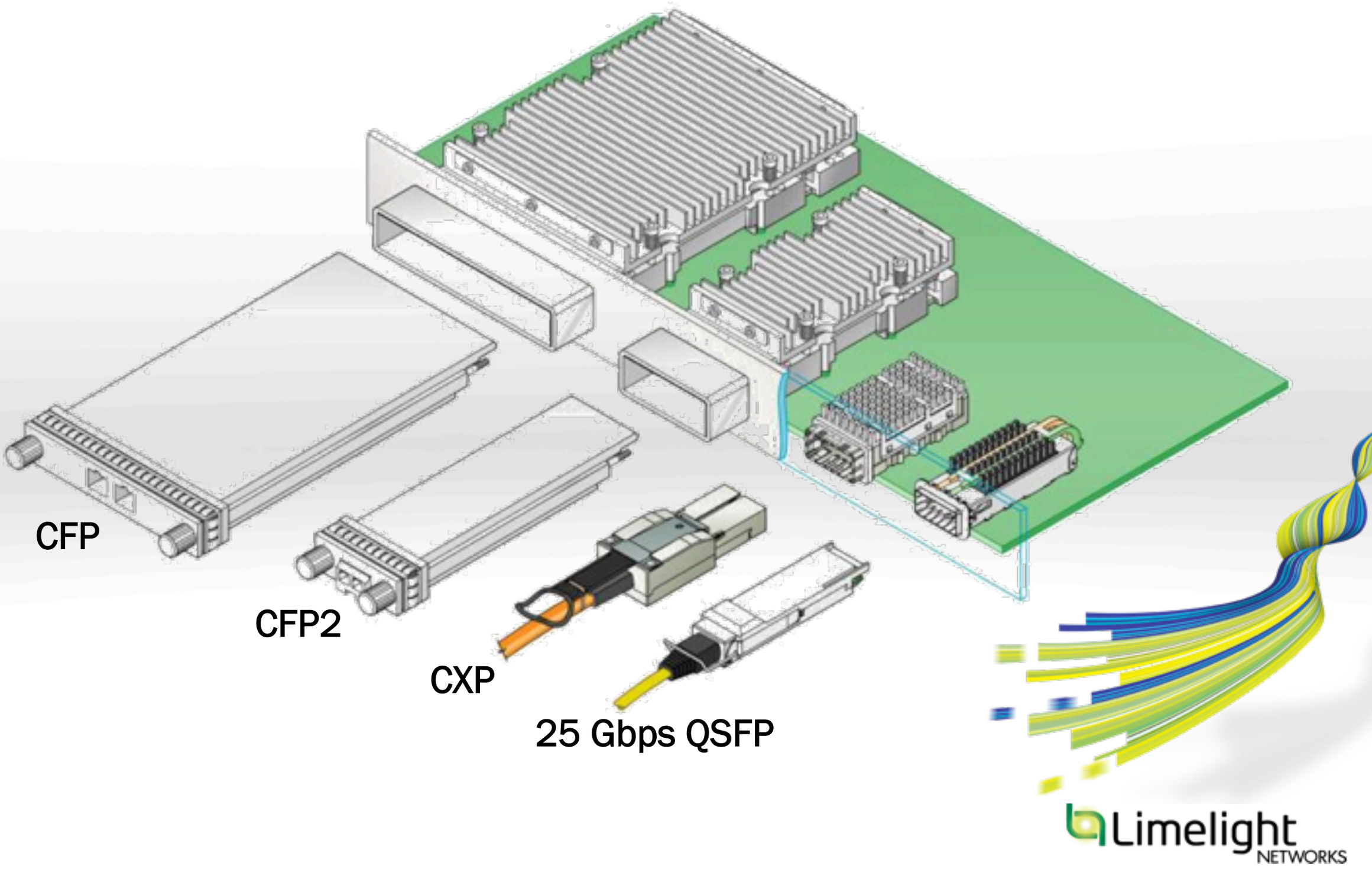
- 10x10 MSA (Multi Source Agreement)
- 26 members including AMS-IX, Brocade, Facebook, Google, ...
- Lower cost, lower power

- Support for 2 km, 10 km and 40 km on SMF
  - Initial 10x10-2km standard published in March, 2011
  - Additional 10x10-10km and 10x10-40km standards finished in August, 2011

www.10x10msa.org

Limelight NETWORKS

1st and 2nd Generation Form Factors



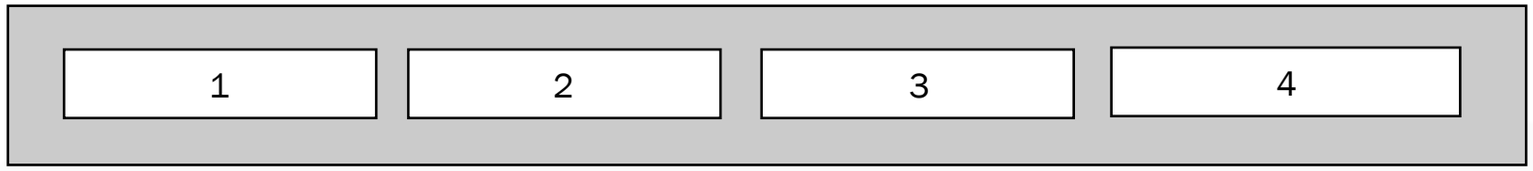
Limelight NETWORKS

First Generation Form Factor Brocade

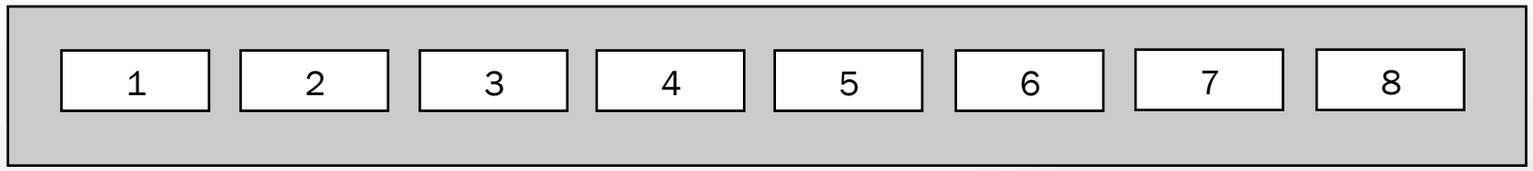
Limelight NETWORKS

Second Generation Form Factors

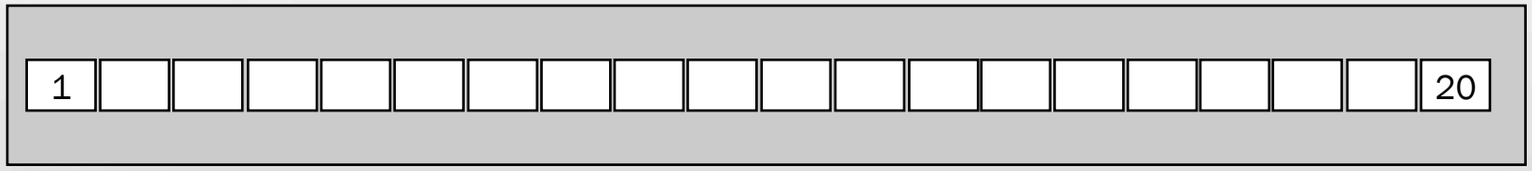
- CFP (2010)



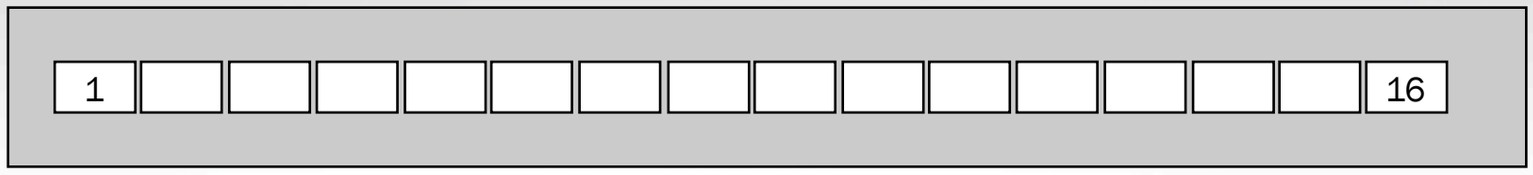
- CFP2 (2013)



- QSFP (2013)



- CFP4 (2014)



Limelight NETWORKS

Overview

100GBase-CR10	100GBASE-SR10	100GBase-LR4	100GBase-ER4	10x10-2km	10x10-10km	10x10-40km
2010 IEEE 802.3ba	2010 IEEE 802.3ba	2010 IEEE 802.3ba	2010 IEEE 802.3ba	March 2011 10x10 MSA	August 2011 10x10 MSA	August 2011 10x10 MSA
CXP	CXP, CFP	CFP	CFP	CFP	CFP	CFP
7m	150m	10km	40km	2km	10km	40km
Copper	MMF	SMF	SMF	SMF	SMF	SMF
2010	2010	2010	2012	2011	2011	2011

Limelight NETWORKS

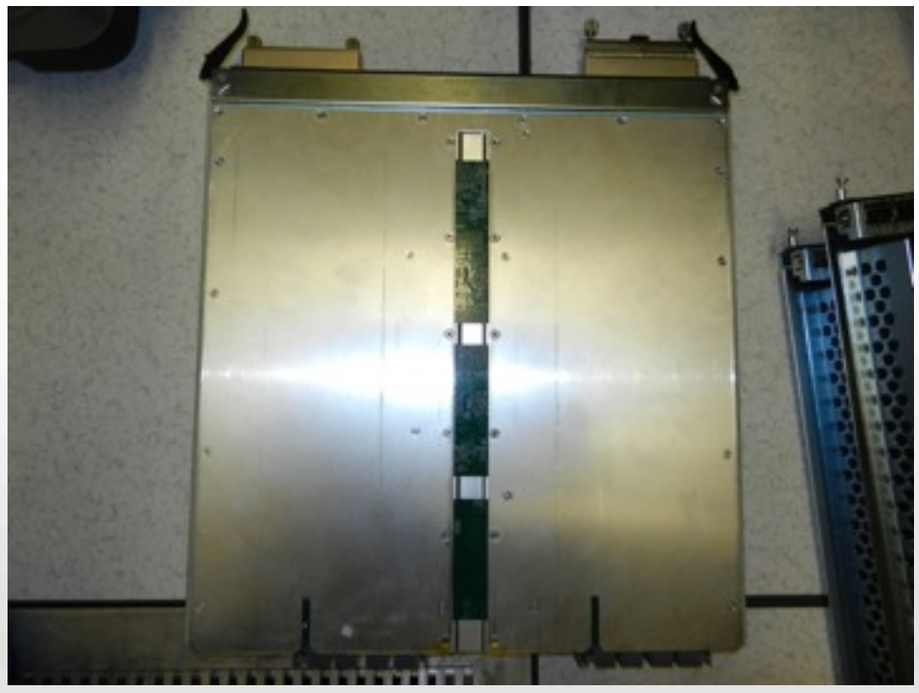
Testing Brocade Hardware

Limelight NETWORKS

Brocade Hardware

BROCADE

- MLXe announced in September 2010
  - Increased backplane capacity to allow for 100GE usage
- Initial 100GE modules and optics shipped in May 2011
  - BR-MLX-100Gx2-X 2-port 100GbE
- New High-speed Switch Fabric Modules



Limelight NETWORKS

Brocade Firmware Releases

- Initial test on BETA releases
  - Version 5.2.0B4T163 (Apr 19 2011)
  - Version 5.2.0B5T177 (May 20 2011)

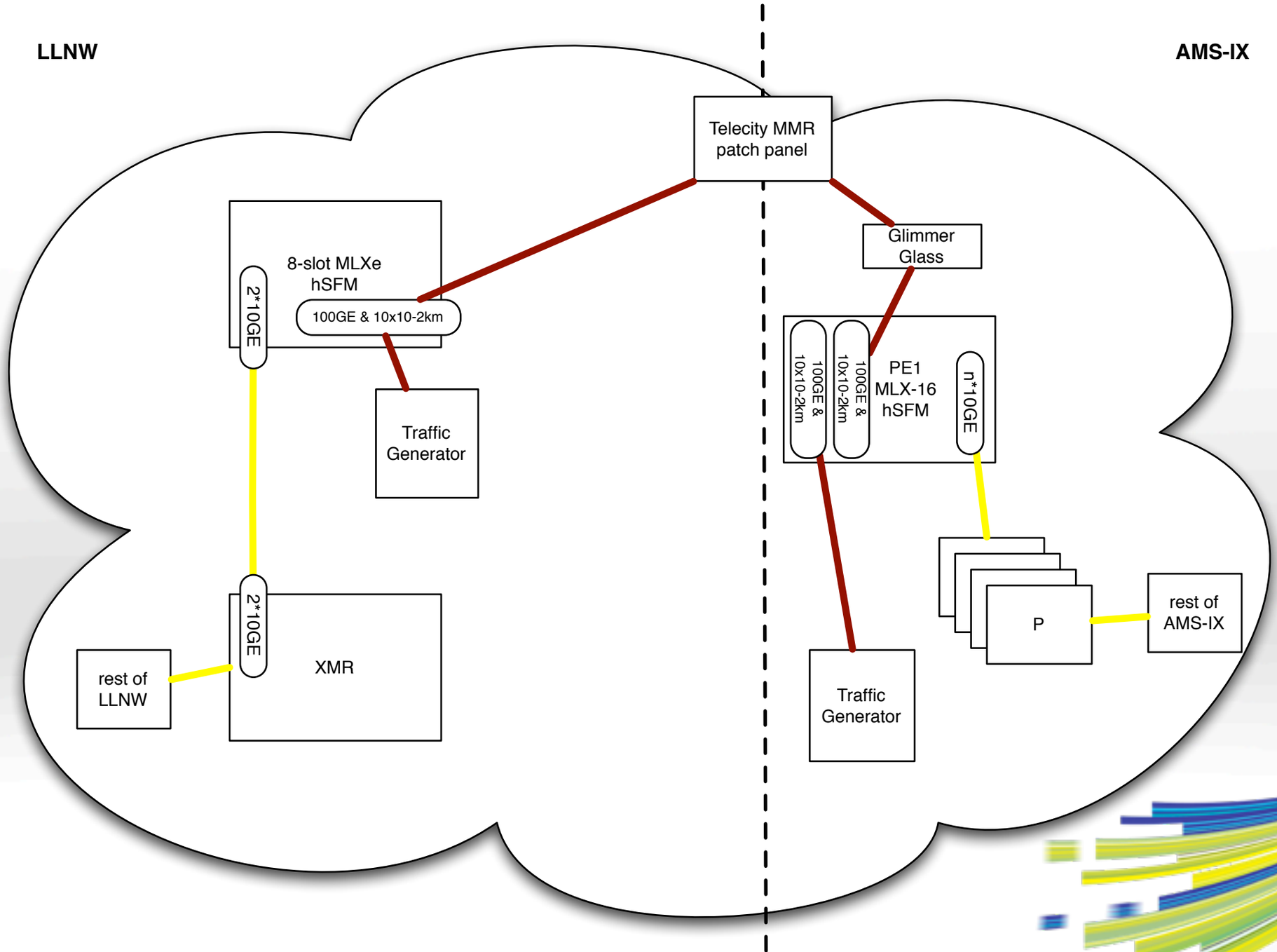
- More testing on 5.2 release
  - Version 5.2.0T177 (Jun 12 2011)

Hardware in Test Setup

- LLNW
  - MLXe-8 chassis
  - BR-MLX-100Gx2-X 2-port 100GbE Modules
  - High-speed Switch Fabric Modules
  - 2 * 10GE to LLNW backbone
- AMS-IX
  - MLX-16 chassis
  - BR-MLX-100Gx2-X 2-port 100GbE Modules
  - High-speed Switch Fabric Modules
  - 16 * 10GE to AMS-IX backbone

- 4 * 10x10-2km CFP optics
- 2 * 100GE Anritsu traffic generator
- 1 * 10GE Anritsu traffic generator

Test Setup

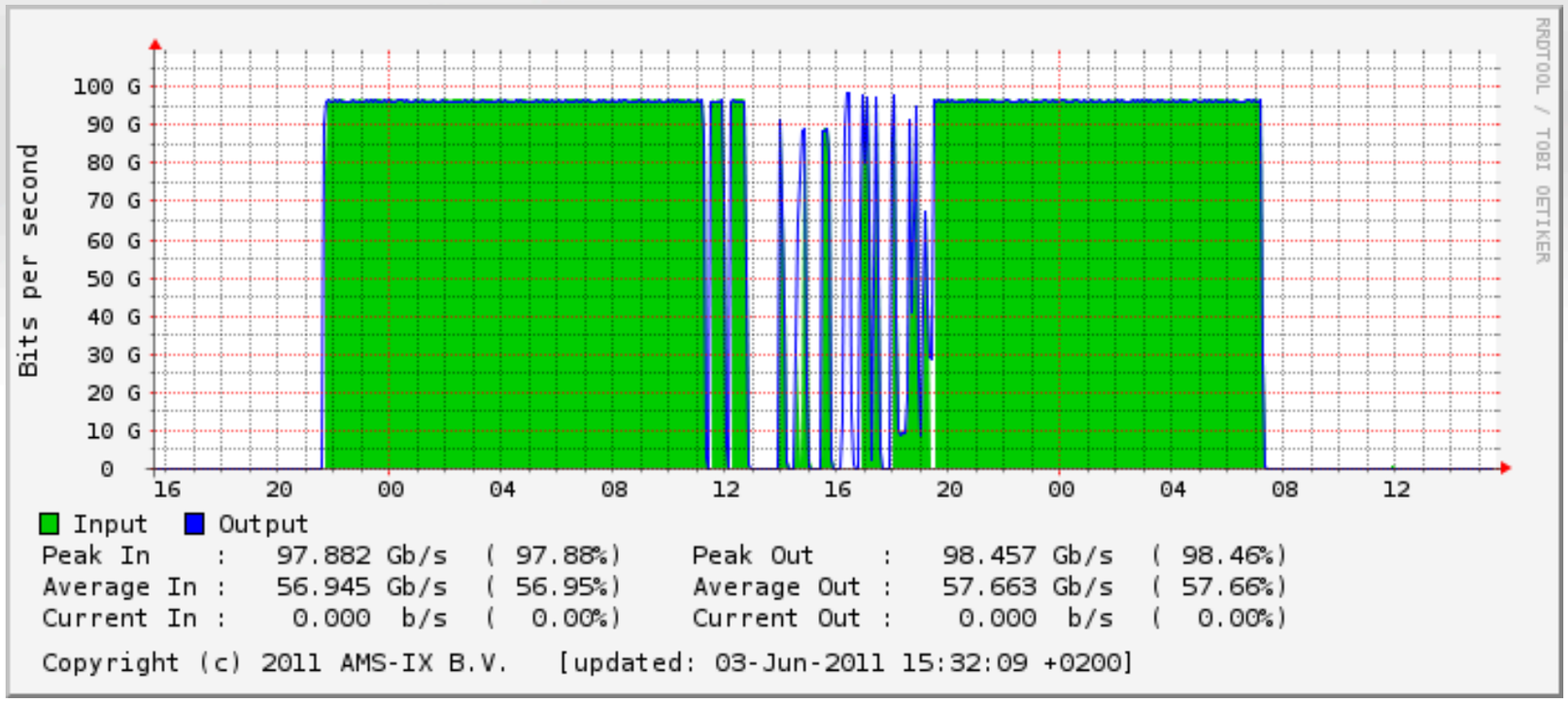


Traffic Flow
n*10GE link
100GE link

Limelight NETWORKS

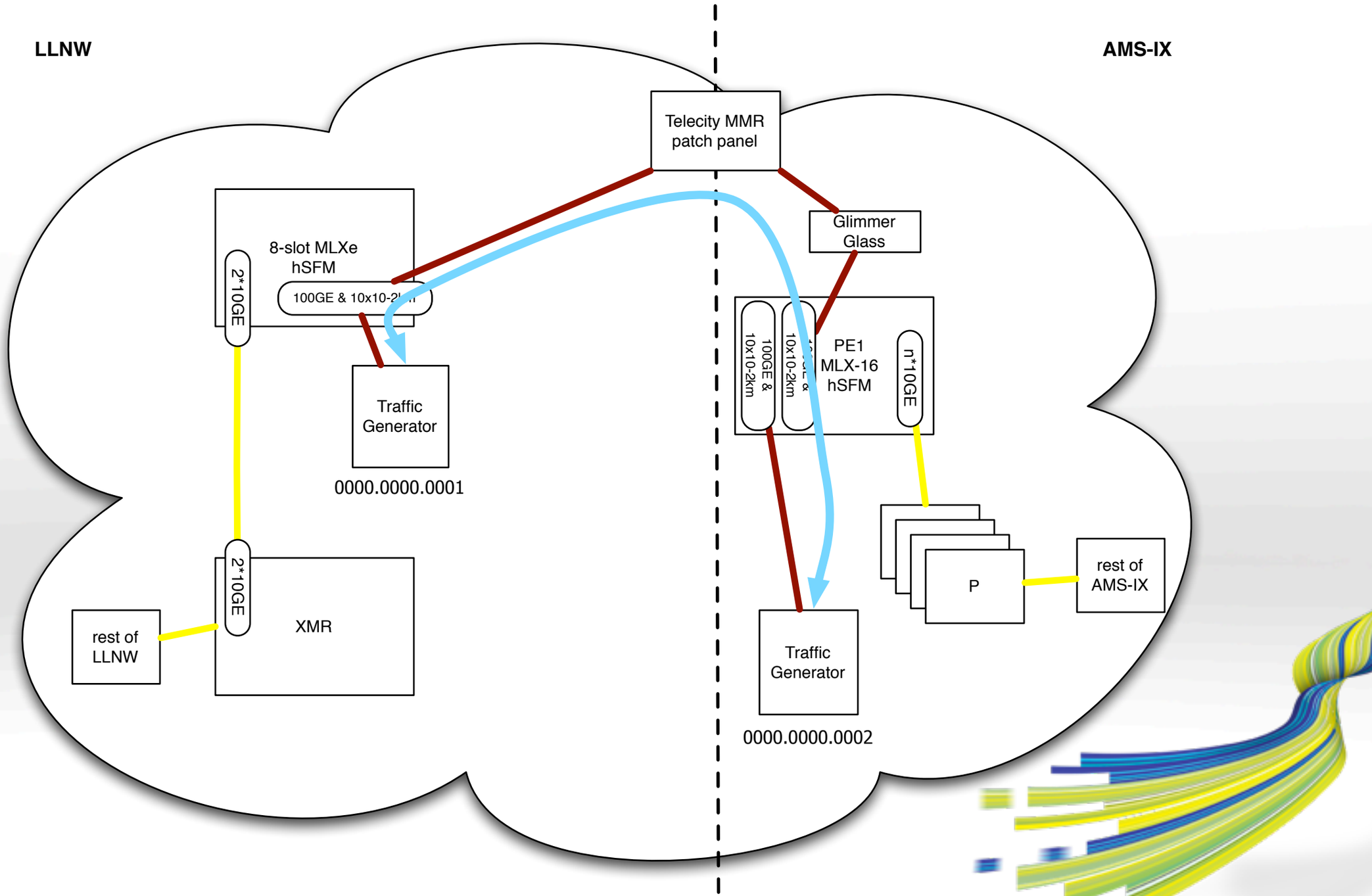
Testing

- L2 Forwarding
- Routing IPv4
- AMS-IX Topology Failovers
- Routing IPv6
- More Routing



Limelight NETWORKS

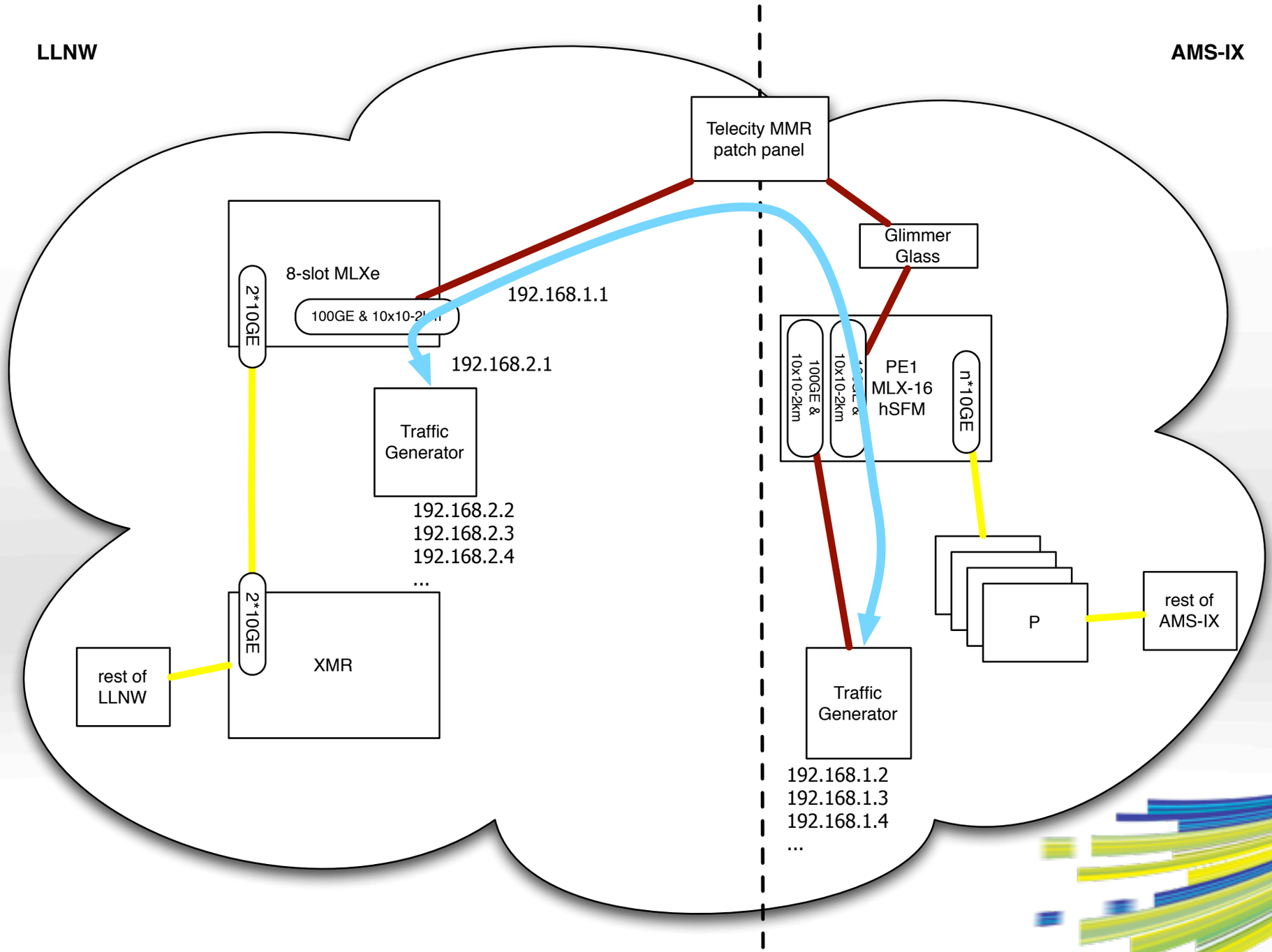
Forwarding Tests



Traffic Flow
n*10GE link
100GE link

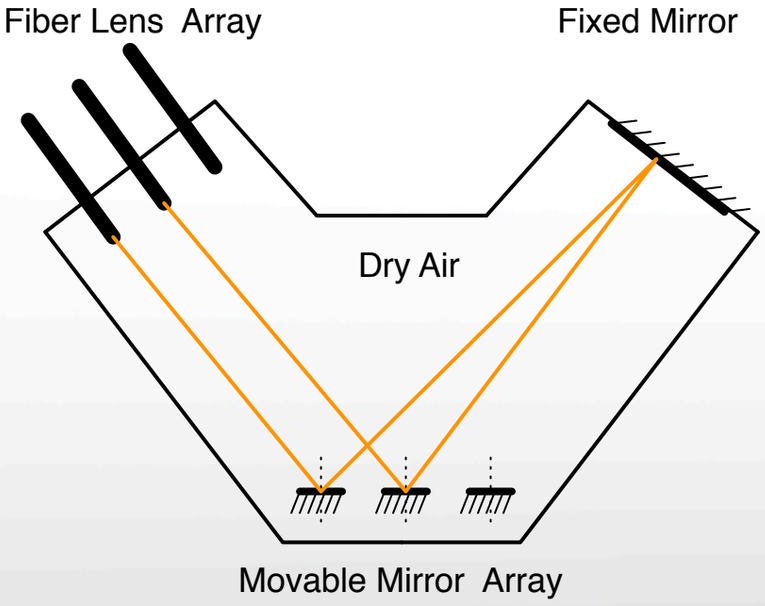
Limelight NETWORKS

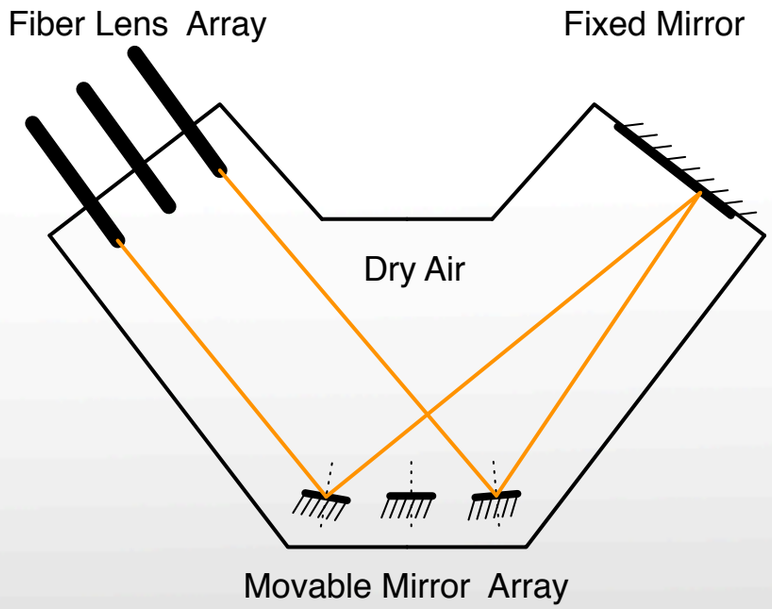
Routing Tests



Limelight NETWORKS

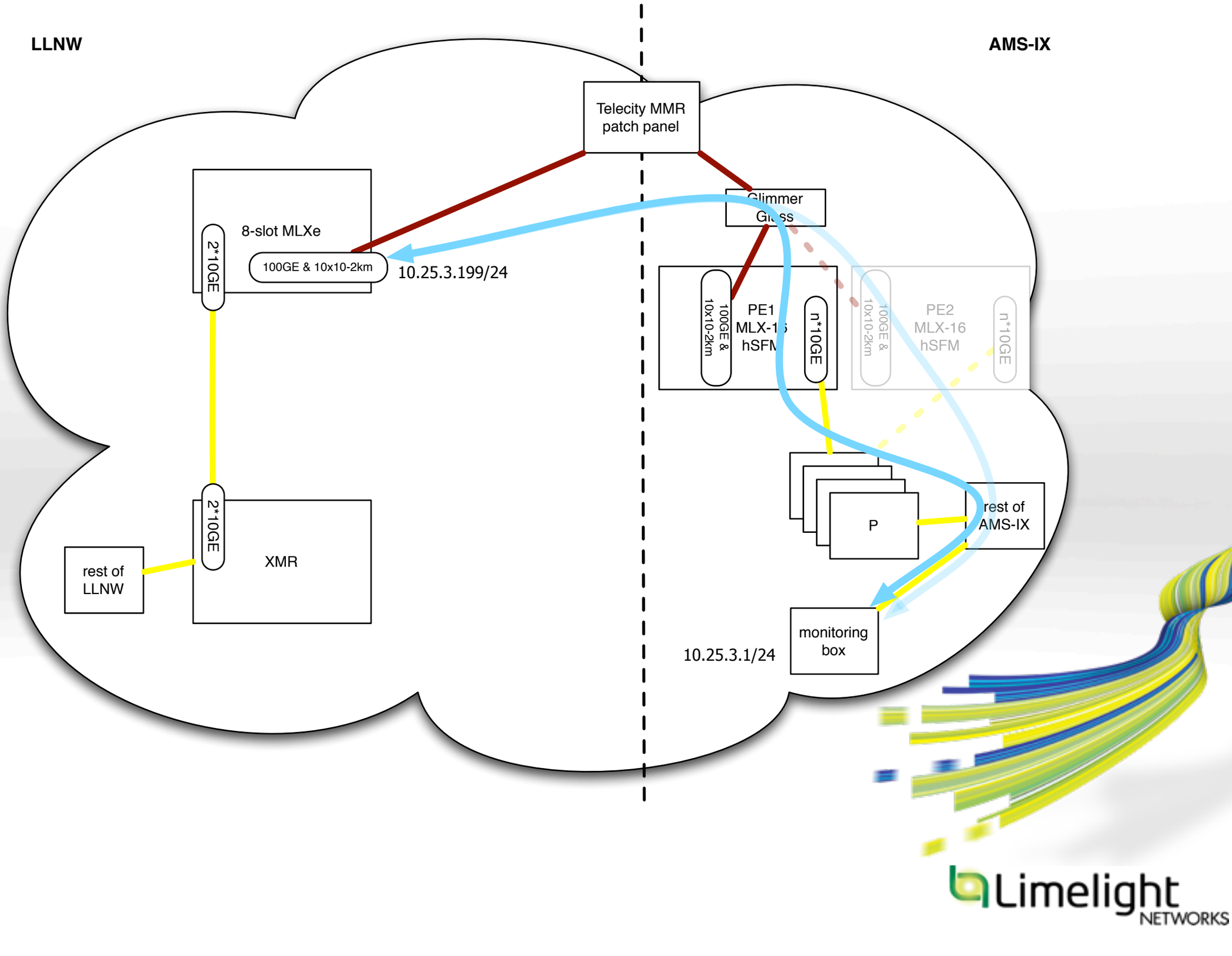
Glimmerglass at AMS-IX





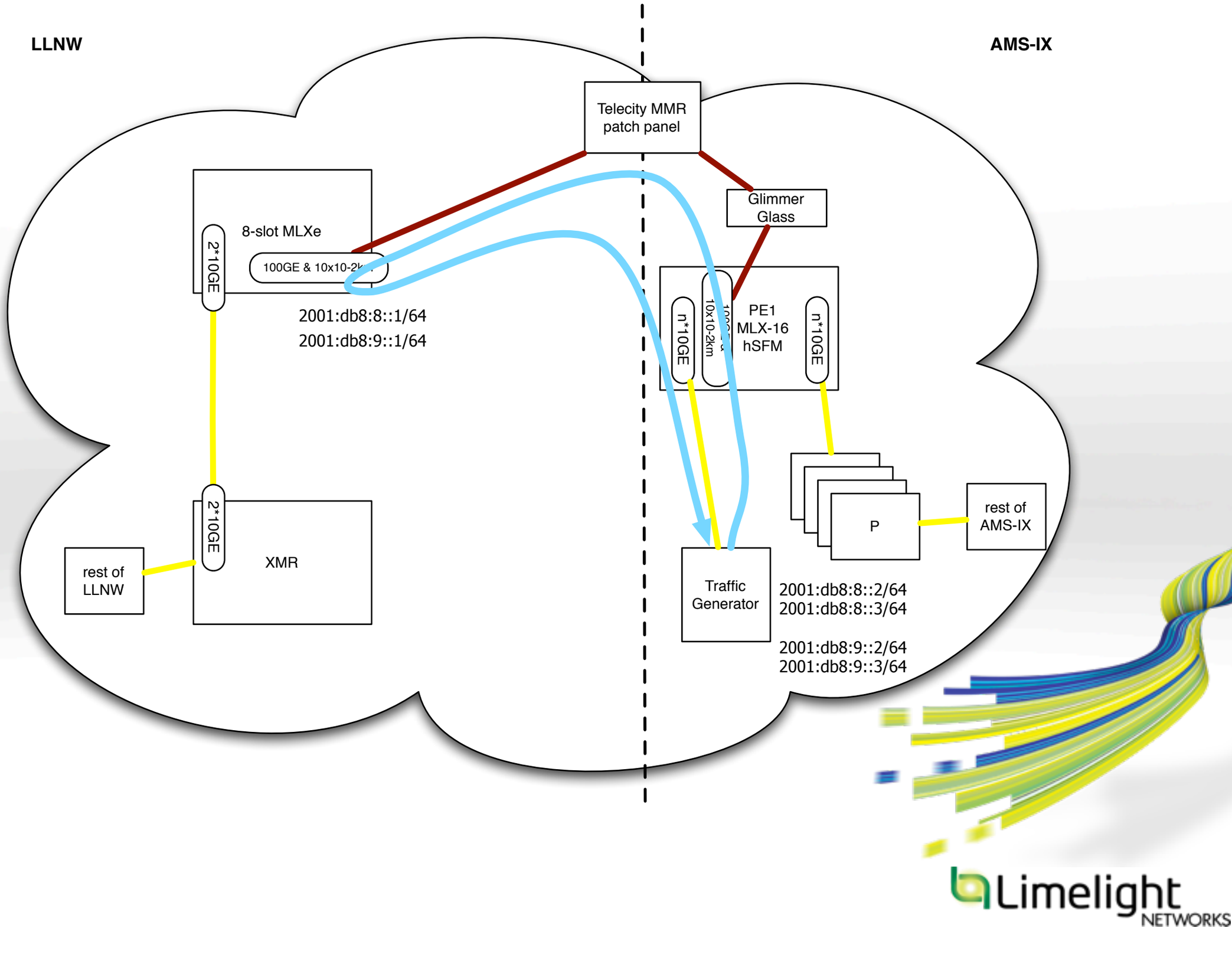
- Photonic cross connect
- Remote patch panel

AMS-IX Swap Tests

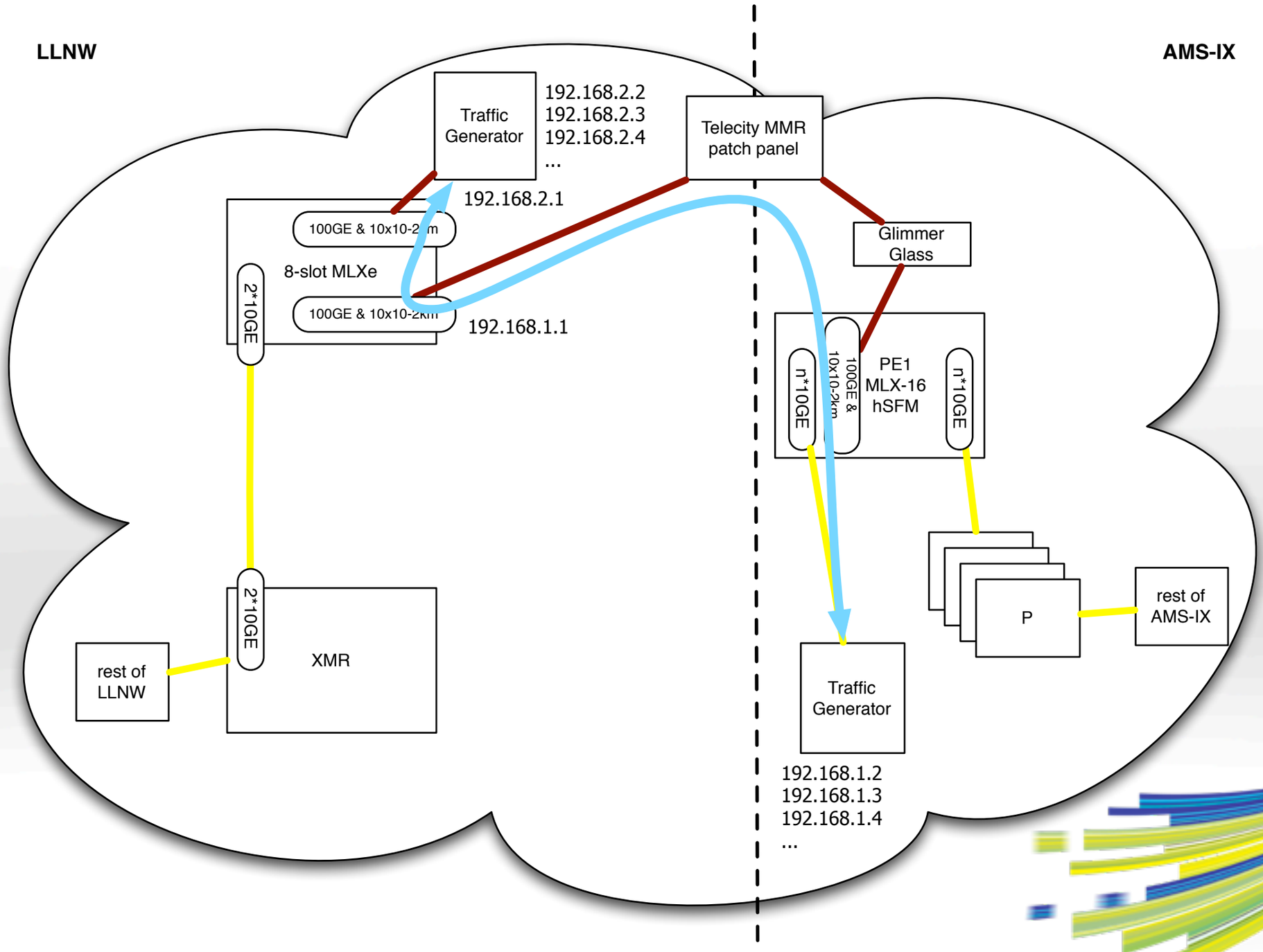


Limelight NETWORKS

IPv6 Tests


LLNW
AMS-IX
Telecity MMR
patch panel
Glimmer
Glass
8-slot MLXe
2*10GE
100GE & 10x10-2k
2001:db8:8::1/64
2001:db8:9::1/64
PE1
MLX-16
hSFM
n*10GE
10x10-2km
n*10GE
P
rest of
AMS-IX
2*10GE
XMR
rest of
LLNW
Traffic
Generator
2001:db8:8::2/64
2001:db8:8::3/64
2001:db8:9::2/64
2001:db8:9::3/64
Traffic Flow
n*10GE link
100GE link

More Routing Tests



Traffic Flow
n*10GE link
100GE link

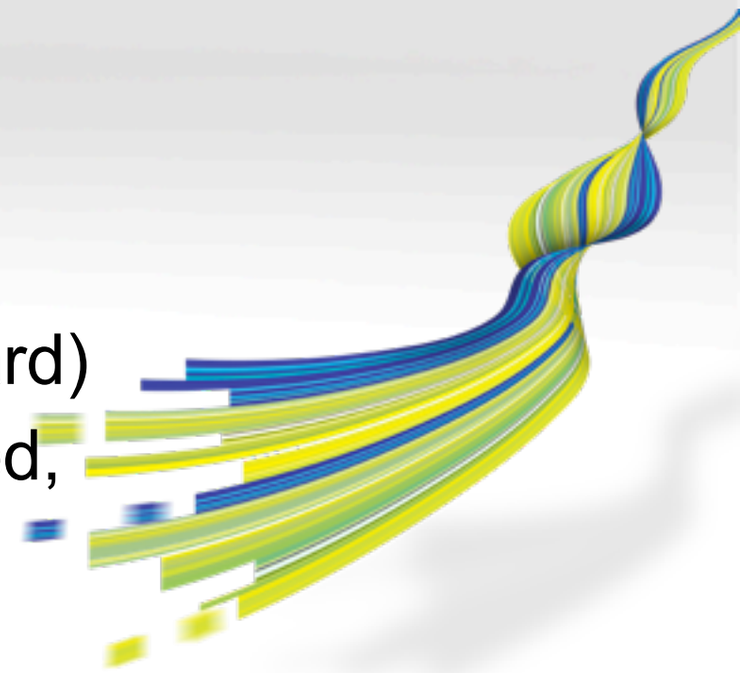
Limelight NETWORKS

Results - Summary

- Line rate forwarding for Internet traffic mixes works
- Line rate routing for Internet traffic mixes works
- AMS-IX topology failovers work (190 ms average service interruption)
- IPv6 works

- sFlow sampling rate (up to 2048)
- Jumbo frames (only briefly tested)

- LP crash and subsequent traffic drop (known issue)
- Bogus MAC learning (single occurrence, reload line card)
- Small frame drops across 100GE card (tuning improved, results within IEEE standard of 10^-12)
- Various minor inconveniences in 5200 release



Limelight NETWORKS

Thanks!
Questions?

Elisa Jasinska
elisa@llnw.com

Limelight NETWORKS