# LINX
# The World's First Juniper IX

"Why we did it…and what happened!"

Derek Cobb, LINX CTO

NANOG54 – March 2012

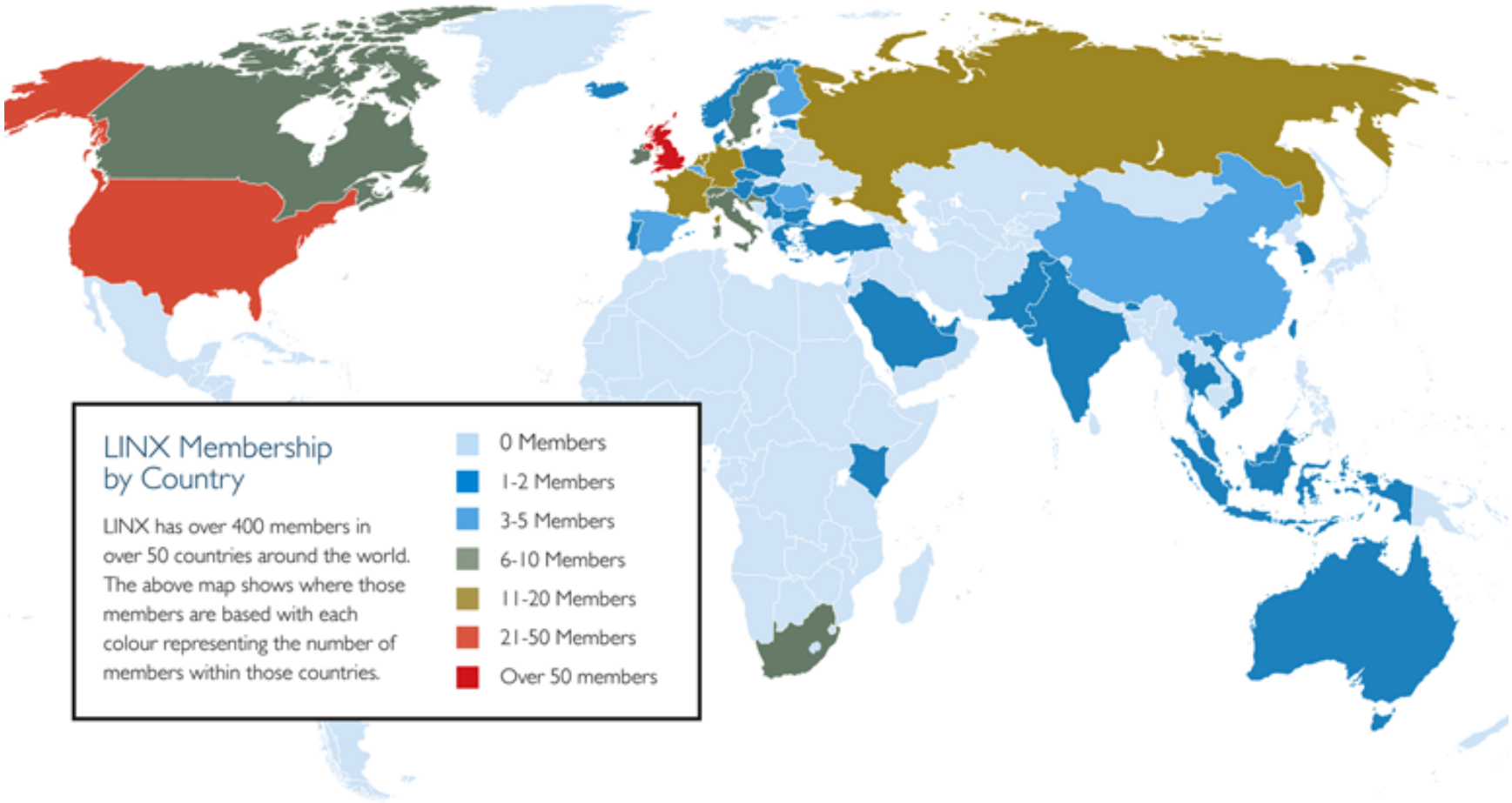
# Content

1. About LINX
2. Pre-Juniper
3. New architecture
4. RFQ and contract agreement
5. Design
6. Programme
7. Lessons learned
8. War stories

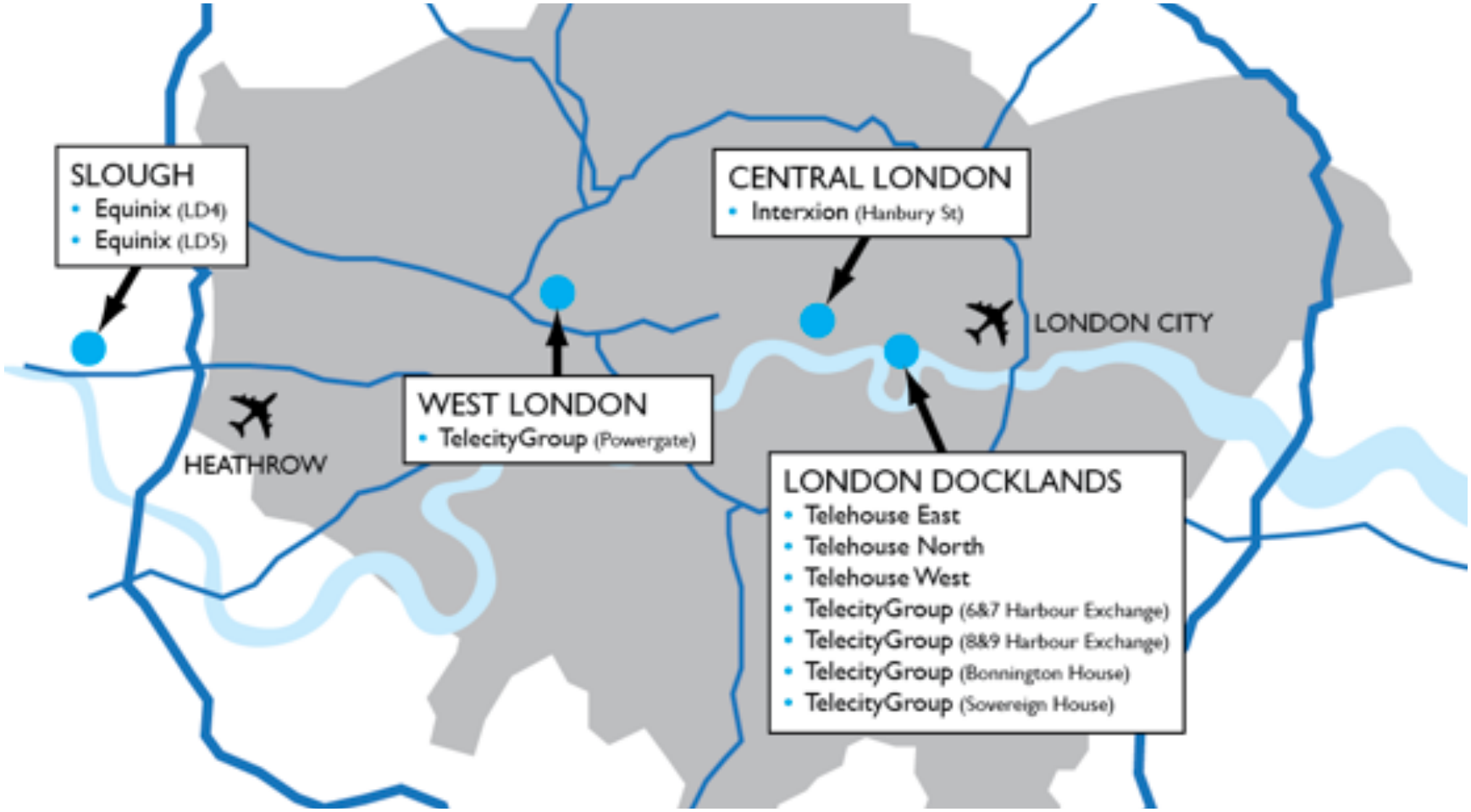
# About LINX: Mission

- Dual mission:
  - To facilitate Internet interconnection in the UK, especially through public peering
  - To represent our members interests in matters of public policy
- As a neutral, mutually-owned membership association
  - With over 400 members from 50 countries around the world

# About LINX: Members

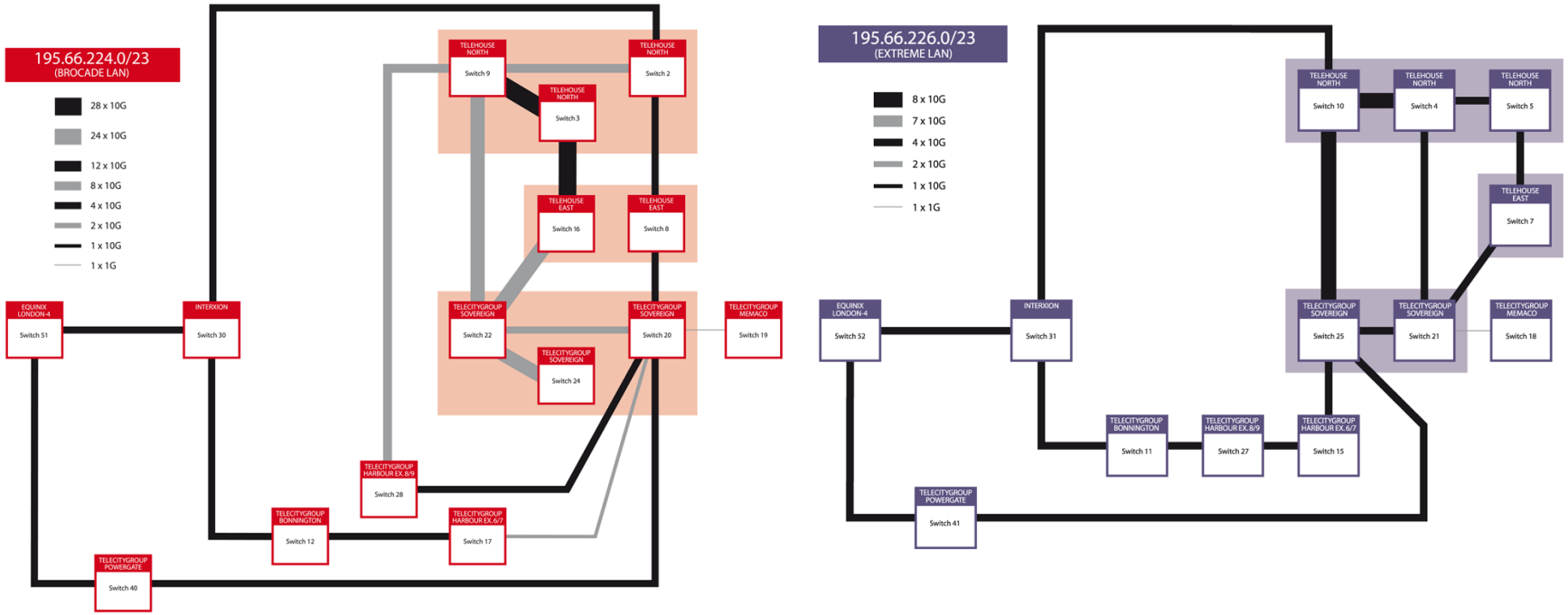


# About LINX: Geography



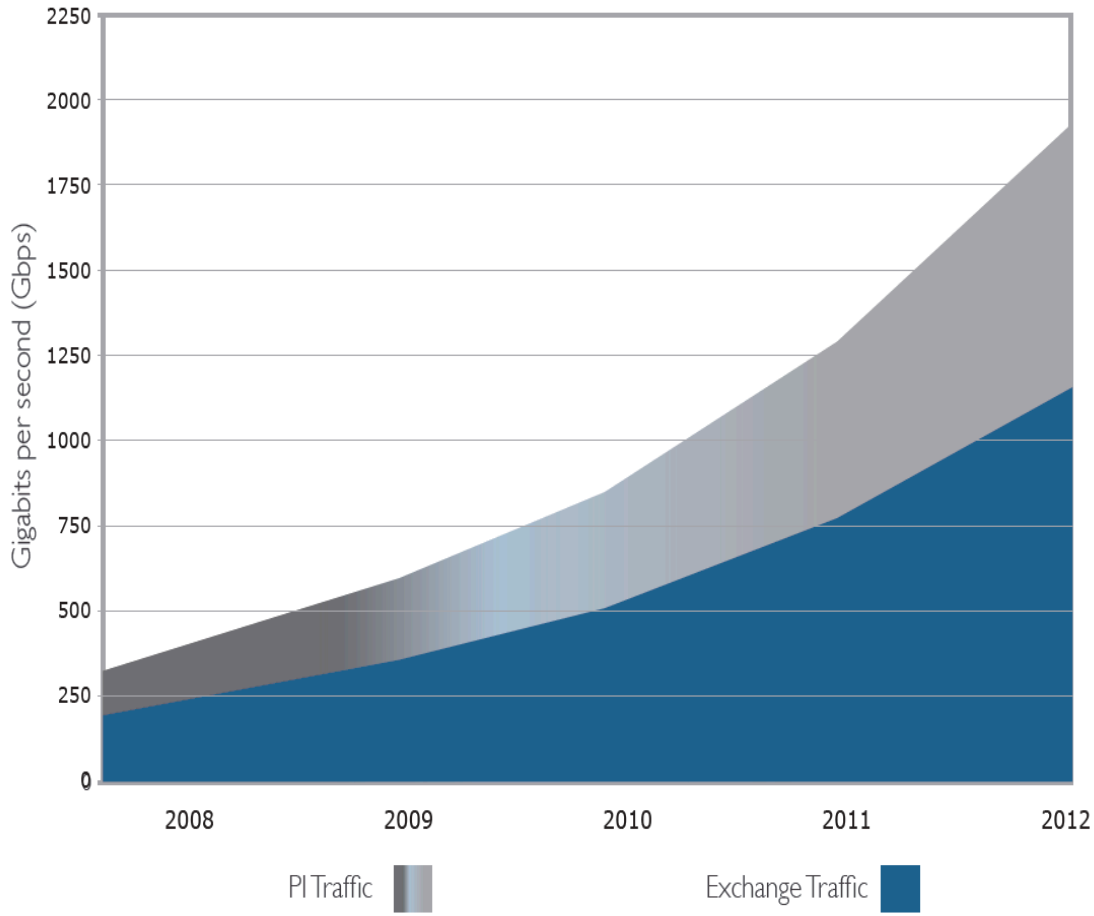
# About LINX: Networks (Jan 11)

- 2x independent n*10G Layer 2 Ethernet rings (Brocade & Extreme switches), IPv4/6 dual stack
- Private Interconnect over LINX managed fibre or systems at many sites and between many sites



# About LINX: Statistics

- ~400 members
- Total traffic >1.8+ Tb/sec (includes private peering)
- 80% of global routing table
- 876 member ports
- 415x10G member ports



# New architecture

- We had realised that our ring-based architecture wasn't scaling well
- We had done the design work on a new architecture in 2010
- We had chosen VPLS as the new design for our primary peering LAN
- We had gained membership agreement to the new architecture in 2010
- Proof of concept work began in 2010 ahead of RFQ in 2011

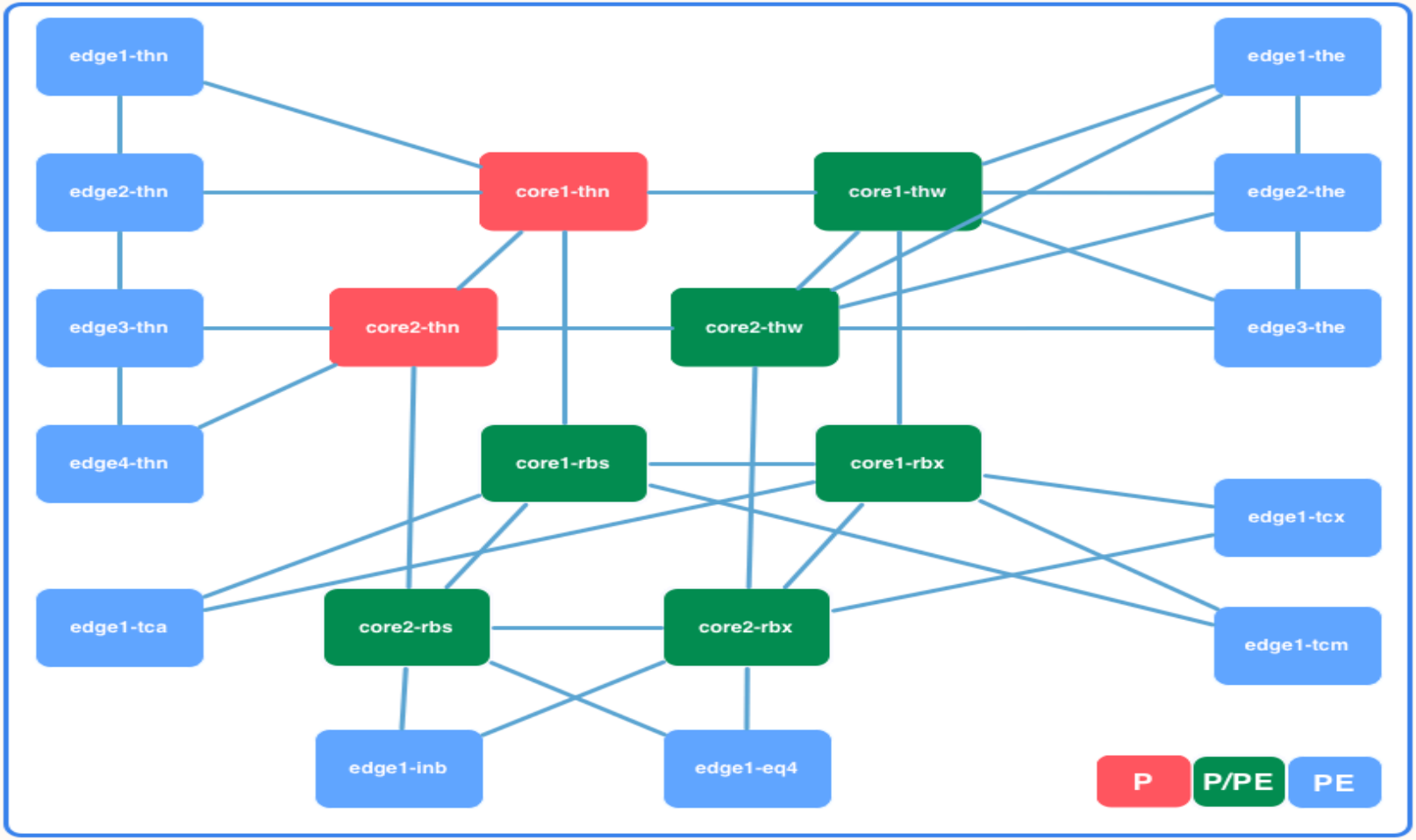
# RFQ

- Juniper and Brocade selected as candidates for VPLS LAN replacement
  - POC testing in Q1
  - RFQ in Q2
- LINX Board award contract to Juniper on 15 June
  - 2 year design agreed to support 100% growth by end 2012
  - Initially MX based with PTX introduction to core in 2012

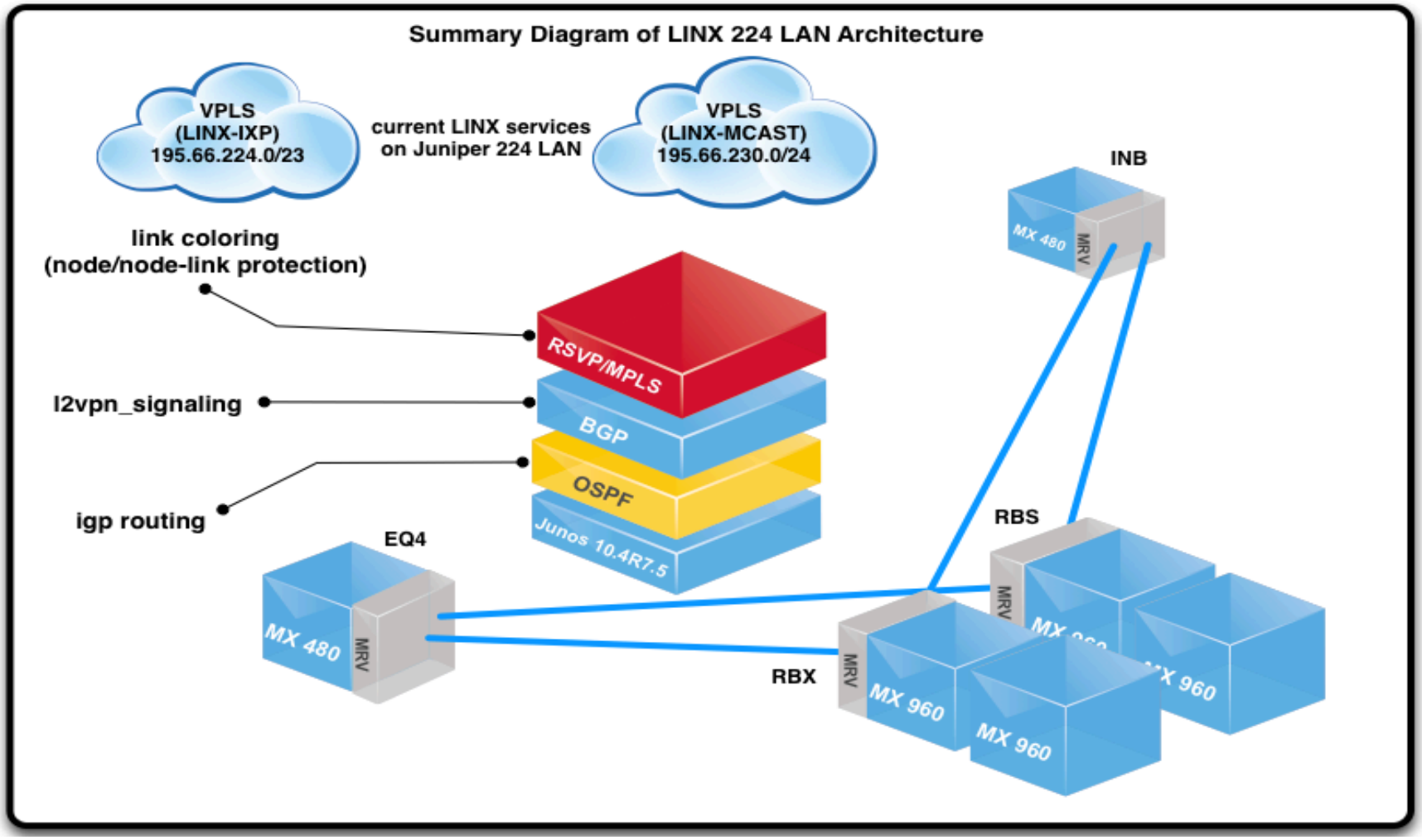
# Contract agreement

- Joint LINX/Juniper sign-off on HLD, LLD & migration plan
  - Juniper professional services and resident engineer contracted
  - Staging of network prior to deployment
  - Juniper resident engineer and enhanced TAC support during migration
  - 30 day retention after final site cut-over

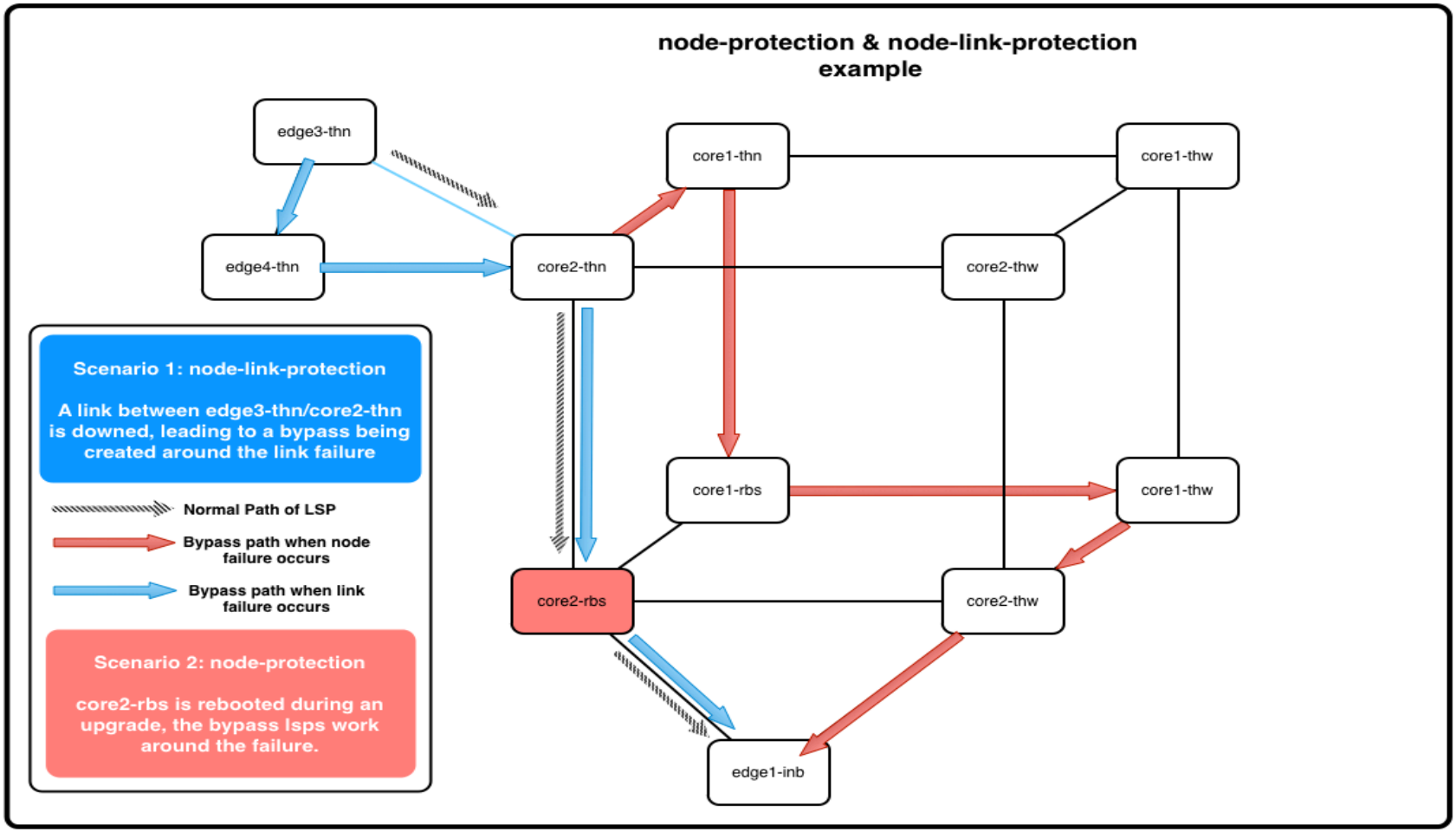
# Juniper LAN Design



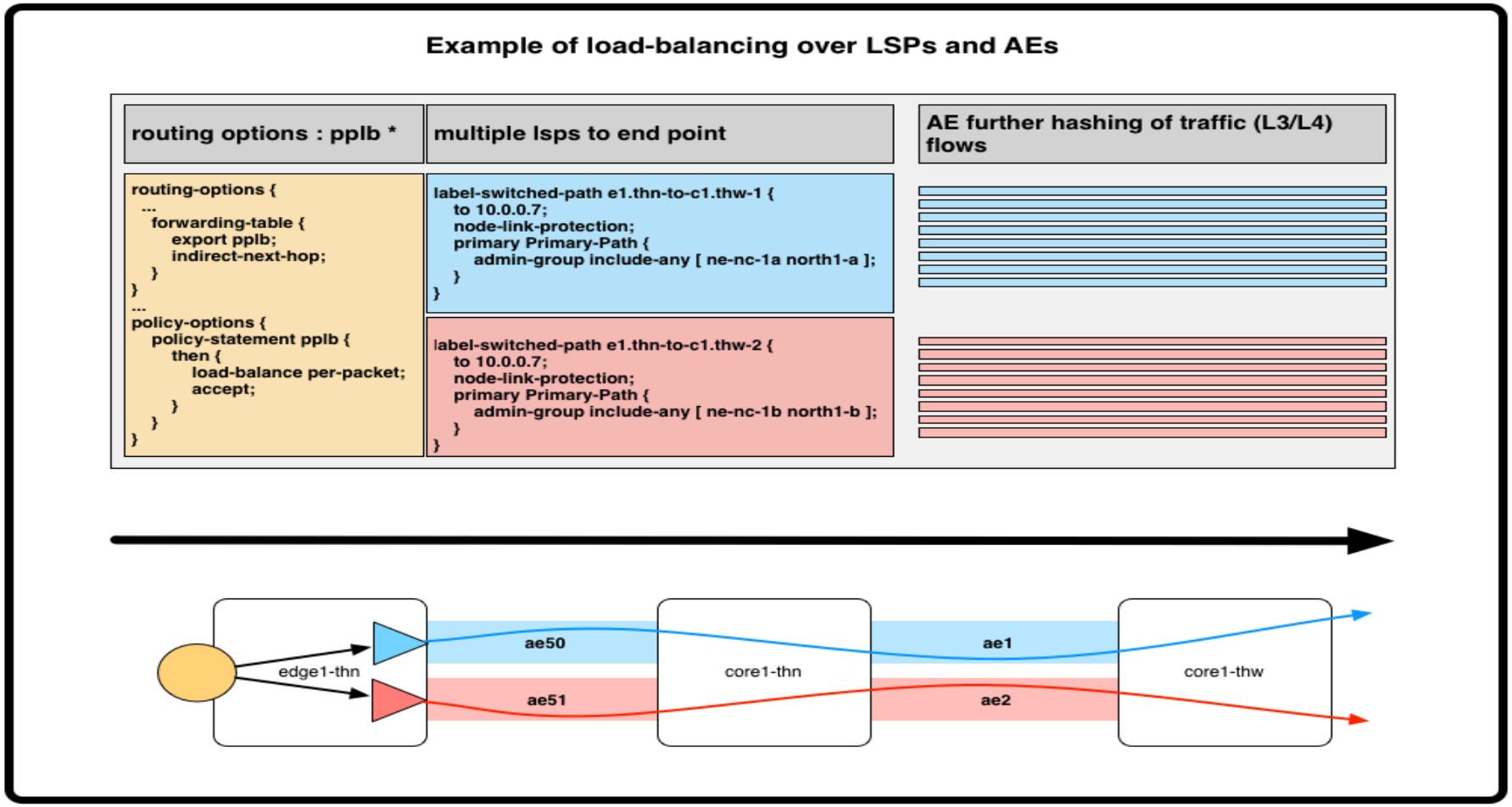
# Design Outline



# Design - Link and Node Protection



# Design – Load Balancing



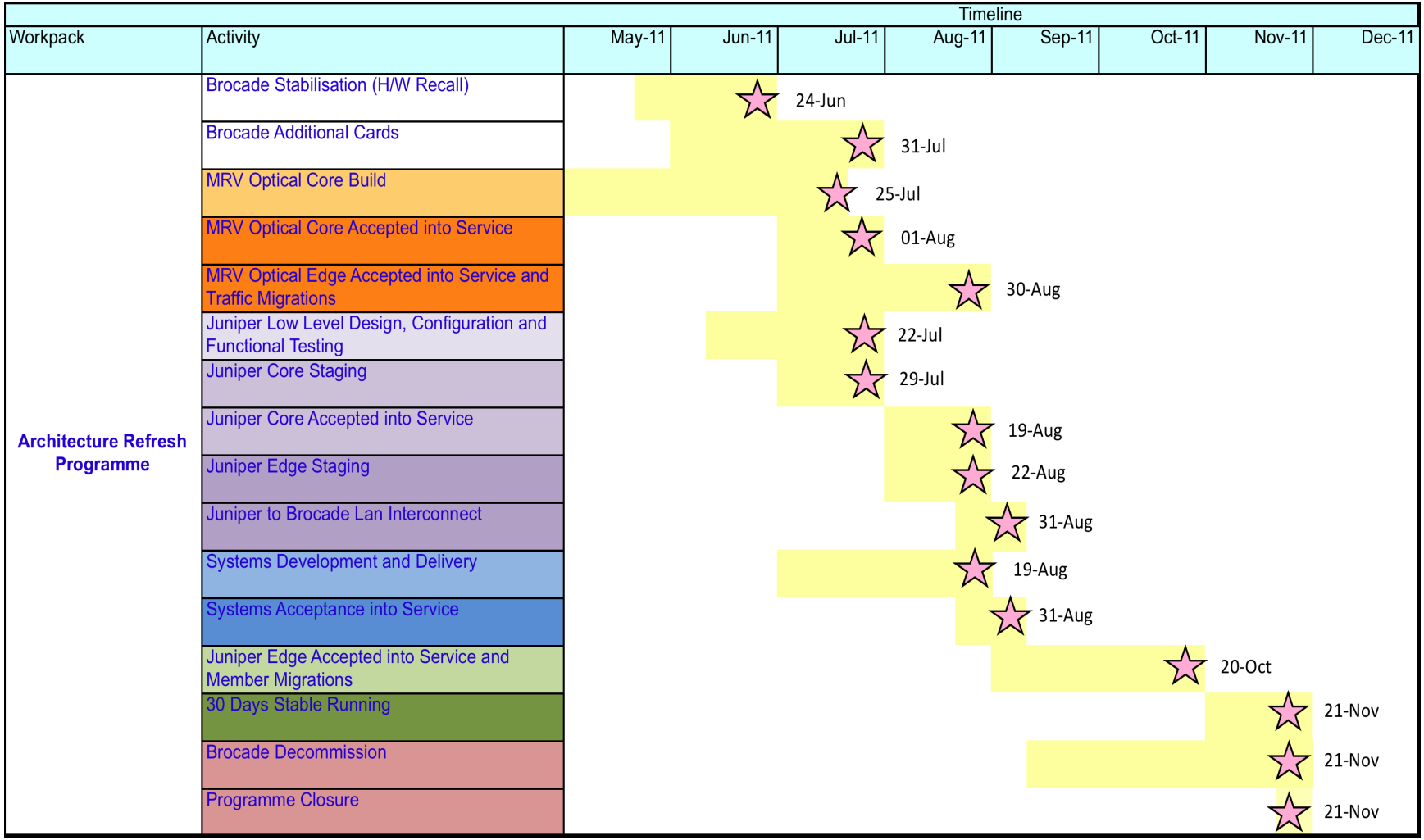

* per flow, not per-packet

# Programme

We're a 35 person company with 13 engineers not used to doing stuff this big, so we…

- Hired a proper programme manager with a track record
- Formed a steering board with clearly defined responsibilities
- Built a clearly defined programme with milestones for:
  - High level design and proof of concept
  - Low level design and detailed configuration
  - Operational Readiness Testing (the network works)
  - Acceptance into Service (network, systems and processes are ready)
  - Migration
  - Decommissioning
- Held weekly programme reviews
- Set up collaboration and sign-off with Juniper at each stage
- Built a clear communication process with LINX members

# Programme Plan



# So What Did We Do Well?

- We came in on time!
- No unplanned downtime!
- We showed the value of close collaboration with Juniper
- We showed we have some good engineers
- We showed a high level of commitment
- We supported each other and worked as one team
- We learned from an early near-miss!
- Our preparation for the major sites was meticulous and paid off
- We communicated well – but sometimes we had to think about it!

# What Weren't We So Good At?

- We didn't manage space and power well enough
  - 2 core sites needed late changes
  - drove additional cost and risk into programme
- We weren't managing our relationships with Data Centre owners
  - day-to-day is fine but anything else seemed to need CEO escalation
- Our technology roadmaps need to convert to real plans
  - programme was too time constrained, increasing risk and stress to team

# What Weren't We So Good At?

- We need to be better at forecasting
  - be better prepared and less reactive
  - better at budget forecasting
  - we need to manage the order pipeline better
- Our communications can be improved
  - migration web site and weekly updates appreciated by members
  - maintenance announcements were good, but clearly didn't reach all the right people!
  - migration ticket updates were appreciated by members

# And the Final Summary is?

- Programme approved - 14 June
- Low level design and ORT - 22 July
- Build complete and Acceptance into Service - 22 September
- Migration started 27 September
- Migration completed 29 October
  - 10 sites
  - 110 x 100 Mb/s ports
  - 196 x 1GB/s ports
  - 277 x 10 Gb/s ports
  - 700GB/s peak traffic

# Migration War Stories

# Staging

# Dark Fibre

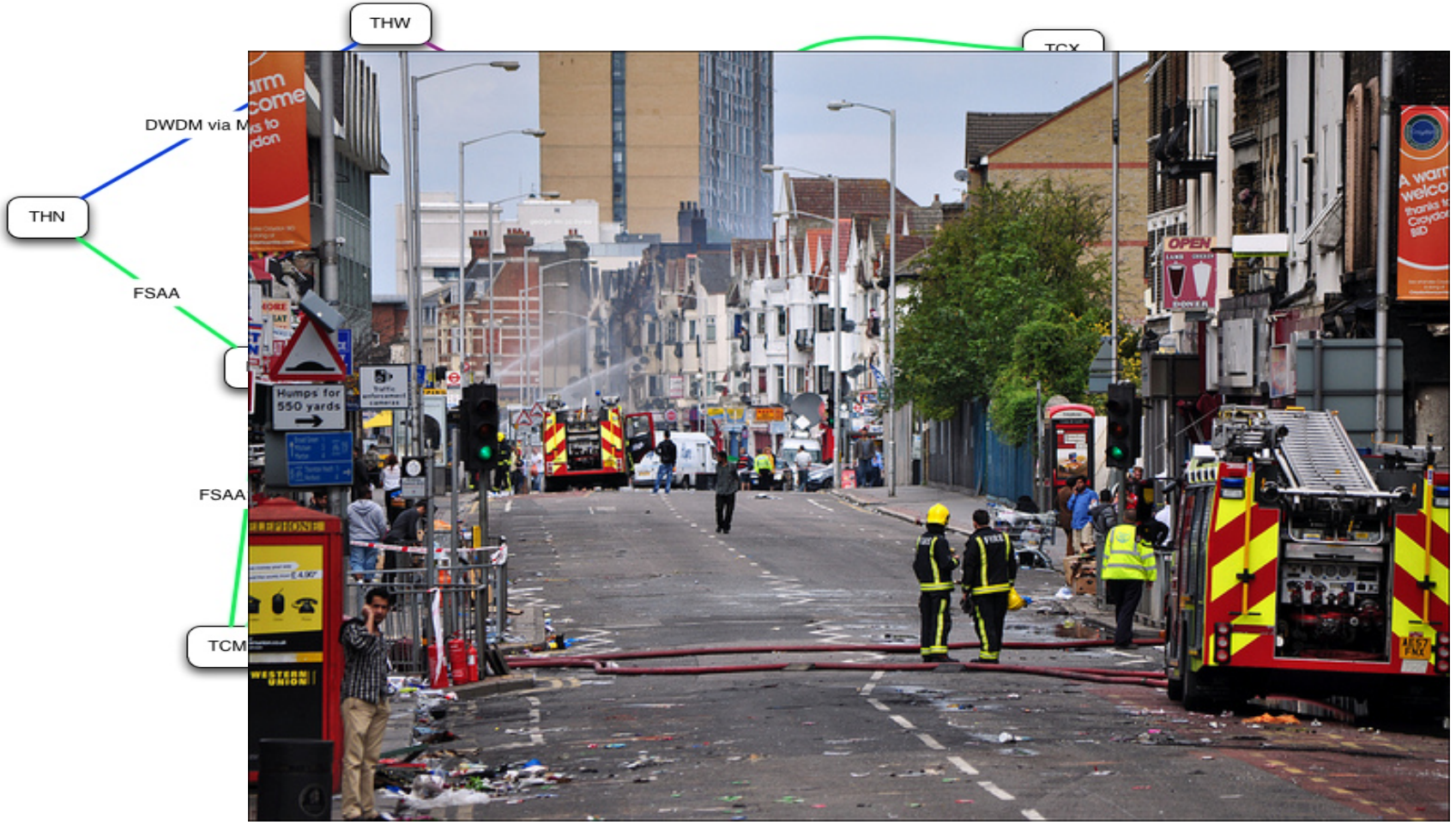
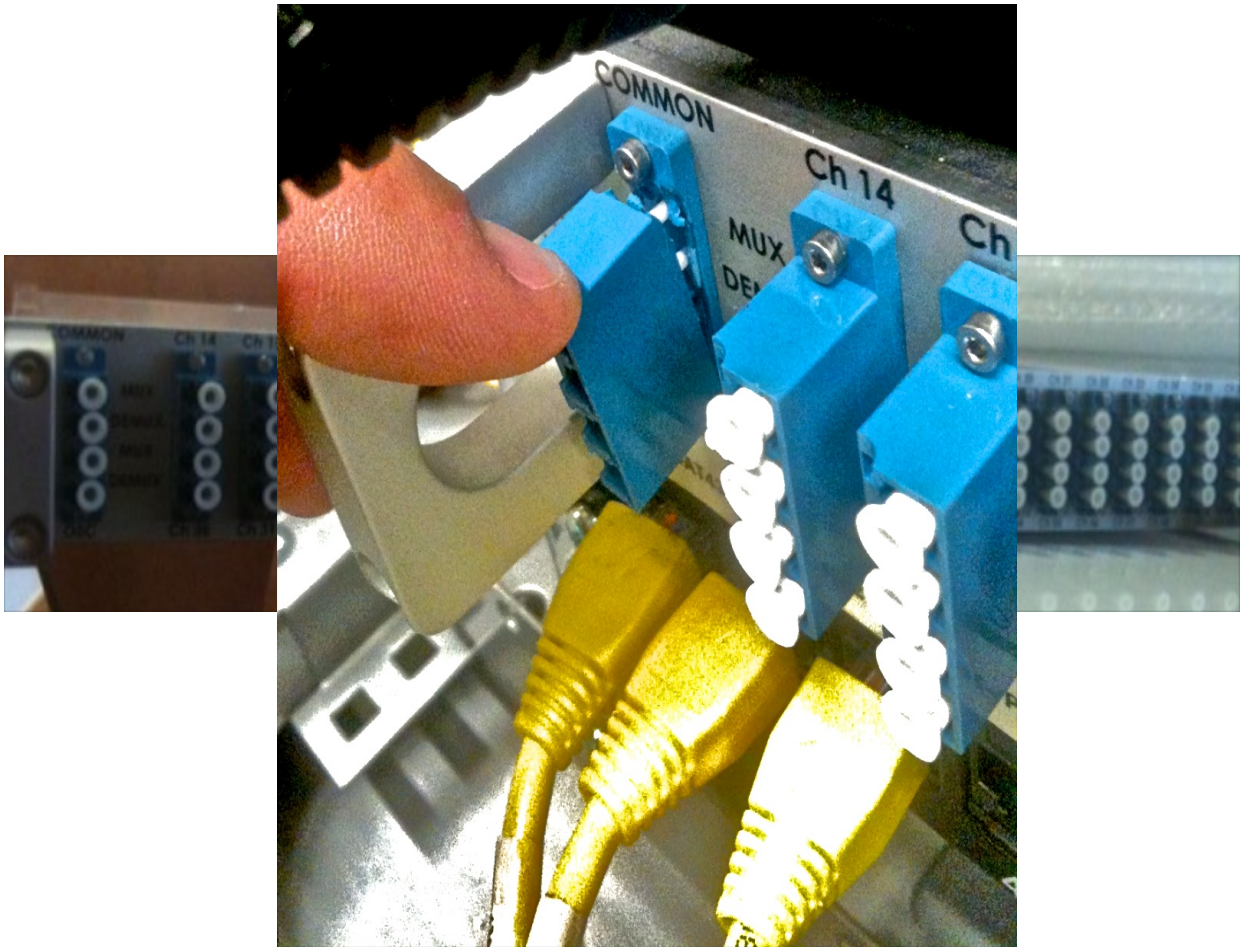
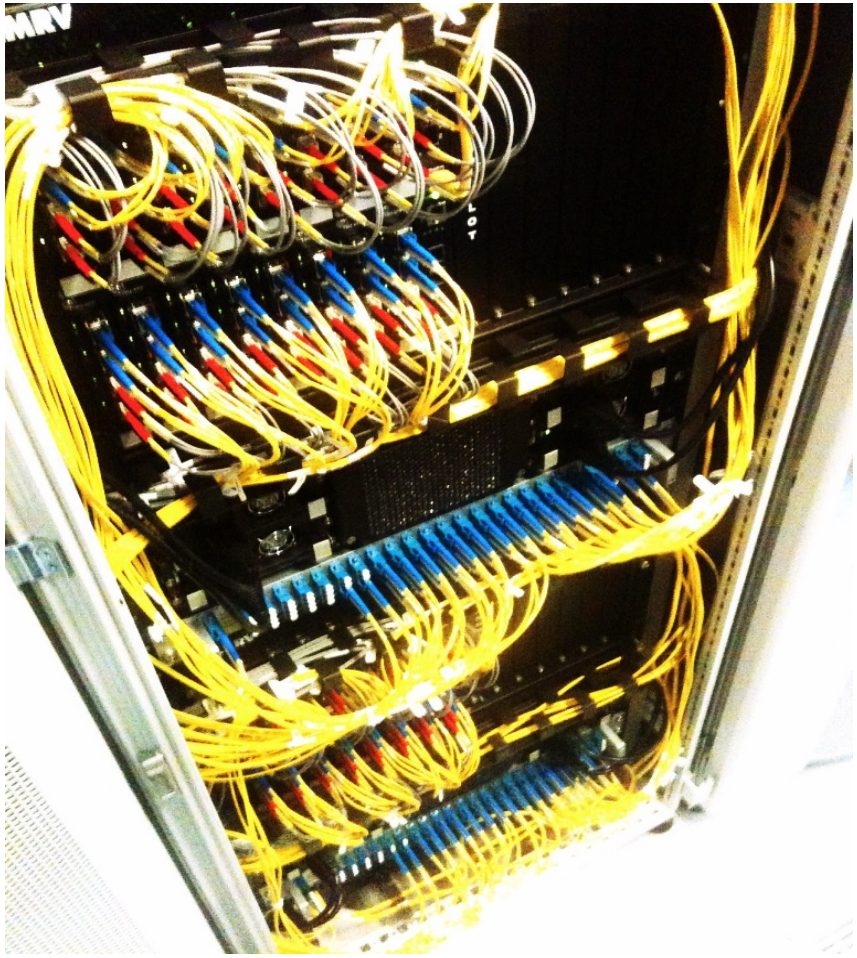



Image: George Rex /Flickr

# Connectors



# More Patching



Some numbers:

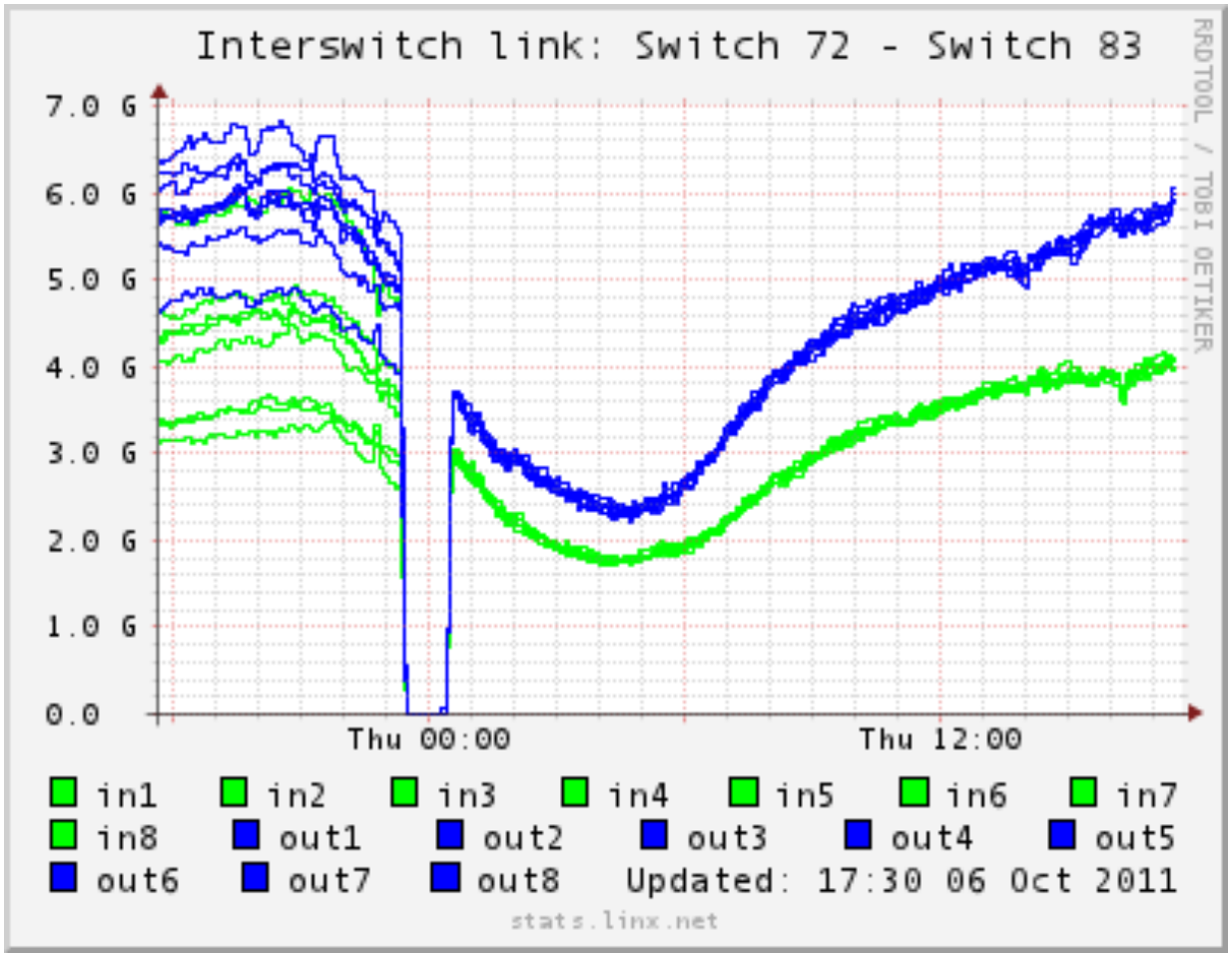
- 1576 ports connected on Juniper devices
- around 900 are ISL ports
- around 500 are connected to DWDM links

# Junos

- PoC testing done on 10.4R1 (early 2011)
- Final testing done on 10.4R5 (pre-staging)
- During staging, upgrade to 10.4R6 following BGP related defect
- Upgrade to 10.4R7.5 after first migration (Load-balancing related issues)

# Load Balancing

Before:
10.4R6




After:
10.4R7

# Current Juniper Capacity

- ~400 members connected
- Current traffic peak ~1Tb/s
- End 2012 design 1.4Tb/s
- Current potential maximum MX edge capacity 9.6Tb/s
- Current potential maximum MX core capacity 3.2Tb/s
  - PTX upgrade increases potential to approx 9Tb/s in 2013
- No architectural limits to scaling beyond these figures

# What's next?

## 2012

- Evolution of Phase 1 design and protection of investment
- PTX5000 implemented at core Telehouse sites (THN and THW)
- MX fabric upgrades allow full line rate use of 16x 10 Gb/s cards
- Controlled introduction of 100Gb/s member access on MX
- Review of transport layer to support 100Gb/s
- sFlow early release on MX in 2H 2012
- Committed for release in Q1 2013

## 2013

- Review of PTX 3000 at core sites at RBX and RBS
- Review of Qfabric for Telehouse sites
- sFlow fully supported on MX

Thank you…

and any questions!