Replacing Cisco 7600s

Better routing through capital expenditure

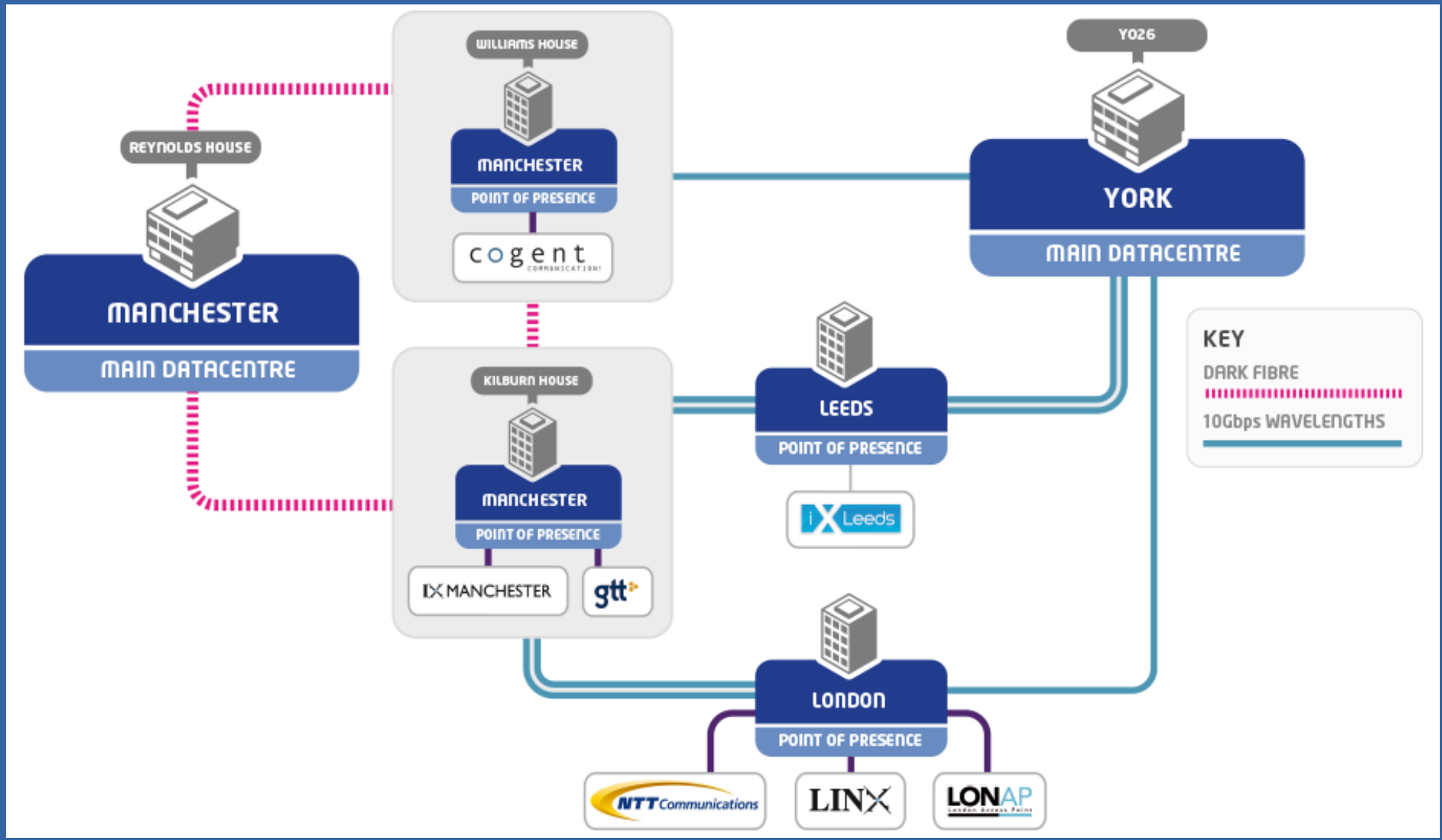
Tom Hill for UKNOF 35, September 2016

A little about Bytemark

- We're a hosting company
- Based in York, which is a bit different
- Try our best to be good, lots of sponsorship
- Built our own data centre in York (YO26)
- And we still have some space in Manchester
- We're growing, and aiming to improve on that

As seen from the ECML

(That's the train line that runs up the East of the Country)



As seen from our website

(There's a line missing. I'll let you work out where.)

As seen from space

(Behold, the majesty of our super-straight fibre paths)

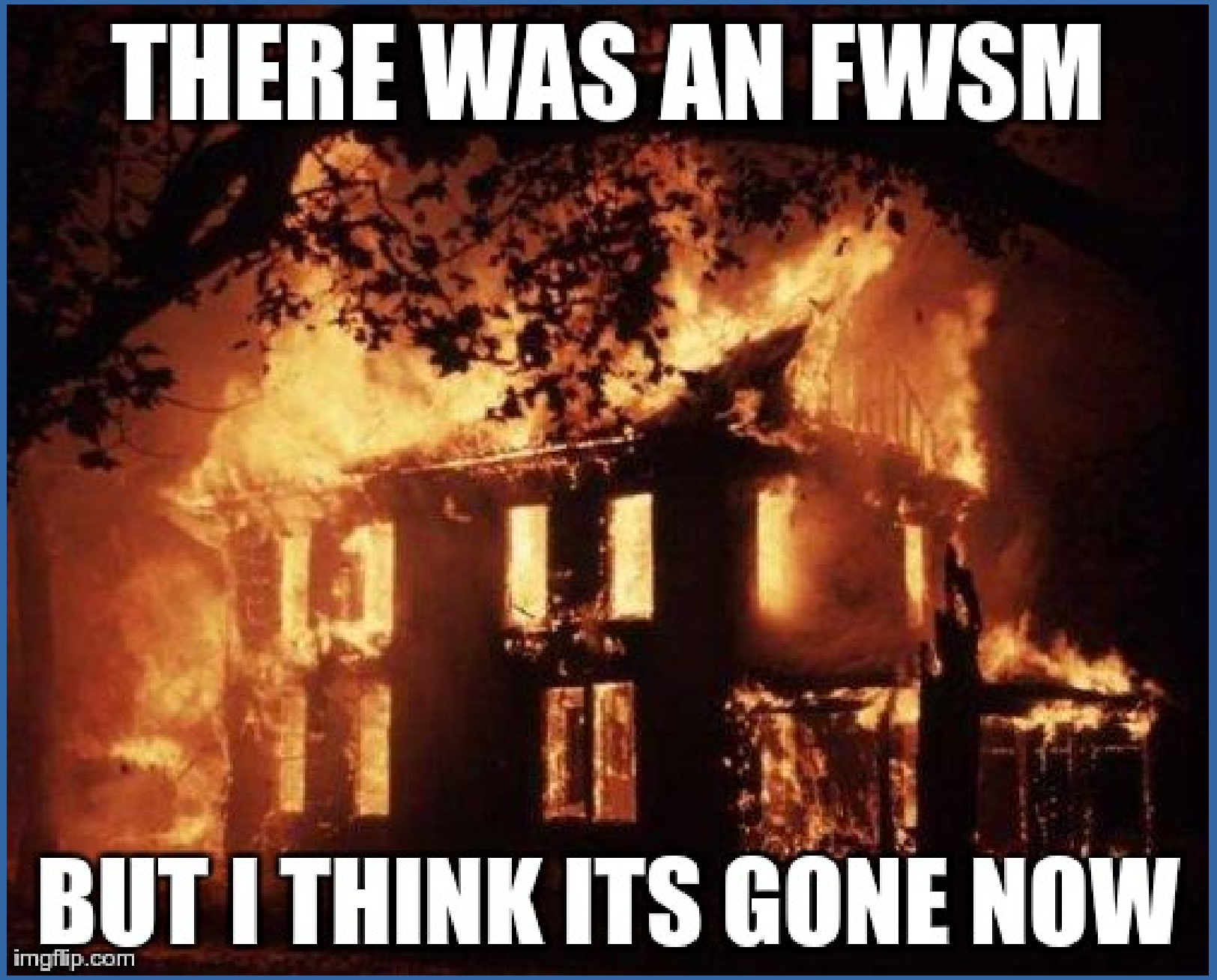
Why do we have all of this?

- Small, but well built – very few outages
- Lots of UK peering (4 IXPs, 3 Cities)
- Dual stack! IPv6-only!
- No SPoFs – makes for hassle-free maintenance
- Generously overprovisioned (helps with DDoSes)
- Make use of plenty of EoMPLS tunnels
- Somehow wound-up with a lot of 7600s...

Look around you

The 7600 wasn't that bad!

- It switches and routes at predictable speeds
- Flagship device, with a strong featureset: MPLS, EoMPLS, MP-BGP, etc.
- Modular & upgradeable
- Tonnes of interface options: ATM, SDH, Ethernet – and probably some I forget
- Long-term vendor support (even now)
- *Loads* of community support (e.g. c-nsp)
- Built-in firewalling!


THERE WAS AN FWSM
BUT I THINK ITS GONE NOW
imgflip.com

It's not aging well

- FIB limit at 1M v4 entries (no dynamic recarving)
- 1GB RAM limit in SUP720 (game over)
- Limitations on IPv6 (URPF, netflow, etc.)
- No multipoint L2VPN without WAN cards
- Backplane capacity is hilarious
- SUP2T was late, and too little an upgrade
- Watts per Gigabit is far too high
- It's. So. Slow. At. Everything.

The catalyst
(No, not that one)

- In 2012 Bytemark had some 'peering issues'
- A.K.A. The SUP720-3BXL melted
- An RSP720 was ordered to help out (a little)
- It managed to route everything in software
- At that point, we discovered the support costs
- The RSP720 was flung, the SUP720 went back in
- In early 2013, we bought an ASR 9001
- Discovered that it ate BGP sessions for breakfast

Likely Options

Cisco SUP2T/SUP6T

- Practically the same device, incremental upgrades
- Wouldn't gain us any real saving (new everything)
- Internally, BU focus has shifted – uncertain on XXL scale

Cisco ASR9k

- Known quantity: we bought one for peering
- Seems to be popular with others

Juniper MX

- Unknown quantity: no internal use previously
- Also seems to be popular with others

A special mention

(Because otherwise *someone* will ask the question)

- When we started this, Arista's R-series weren't things that existed
- It's not gone unnoticed, and we have spoken with them
- The software stack isn't **entirely** ready for us as yet, but beginning 2017 it could be
- Certainly haven't ruled them out for future expansion

The Decider

- Cisco pricing on ASR9k was very amicable
  - We had a *really* good account manager
  - Refurb pricing filled in the gaps
- Hardware support costs for the Juniper routers were just huge
- Some concerns over MX104 control plane
- Mostly: we had the one 9001, and we were quite taken with its performance
- The ASR9k inter-op'd perfectly with the 7600s

Picking what to buy is where it got complicated...

Software

- IOS-XR, not IOS – complete with crazy upgrades
- 'commit' system, familiar to JunOS users
- Route policy is much improved over route-maps
- There's a major reworking of the user permissions system – also 'admin' accounts
- Far bigger emphasis on CoPP for CP security
- Licenses for: 1) any L3VPN use, and 2) over 8 VRFs
- QNX based, moving to a Linux base with 6.1 (on selected platforms)

Chassis Lessons

- Two types:
  - The 9xxx, top-out at 880Gbit/slot
  - The 99xx, will scale somewhat further (~1.6Tbit/slot)
- 9006/9904 run with side-to-back airflow
  - Baffle kits are available to convert that
- Some models need v2 fan trays for denser line cards, when used with “high power optics”
- DC and AC “PEMs” available, multiple versions

Supervisors

- All have excellent control plane performance!
- Scale comes in two options: TR and SE
- RP1 or RP2 for the 9912 & 9922 (+ up to 7 SFCs)
- “9910 RSP” for the 9910 (+ optional 5 SFCs)
- RSP440 or RSP880 for the 9006, 9010, or 9904
- RSP440-LT and RSP880-RL are license-upgradeable
- You get half the advertised bandwidth with only one RSP

Line cards (Typhoon)

- 2nd generation of cards, with 60Gbit / 45Mpps per NPU
- Scale comes in TR or SE guise (doesn't affect FIB size)
- 'MOD' cards hold up to 2 MPAs, 8x10G or 2x40G at most
- 24x10G and 36x10G cards are available, but watch out for the 'ports per NPU' design choices
- 1x100G & 2x100G CFP cards available, but mostly pointless now
- The 40/56G cards were a late addition, but make a decent 1G aggregation option

Line cards (Tomahawk)

- 3rd generation of cards, with 240Gbit / 149Mpps per NPU
- MACSec support is engrained in the design
- 'MOD' cards hold up to 2 MPAs, 20x10G or 2x100G
  - And backwards compatible with **all** previous MPAs
- 4x100G & 8x100G CPAK cards available, and a 'CM' variant of the 8x100G that starts at 2x100G (licenses!)
- 12x100G QSFP28 card has really small buffers, and is aimed at core-facing only
  - Needs a license for full FIB, too

The 9001

- 9001 is the 2U, “hide it in a closet” peering router
- Does side-to-side airflow. And has a v2 fan tray option
- Essentially Typhoon-based MOD80-SE line card with something like an RSP440 sup welded to it
- MPA options are limited to 20x1G, 4x10G or 1x40G
- Onboard SFP+ utilises 'spare' capacity from Typhoon NPUs (~60G per NPU).
- The 9001-S is an introduction model, which comes with one NPU disabled (license upgradeable)

DO NOT BUY TRIDENT

The 1st generation ASR9k LCs, not the Nuclear Deterrent

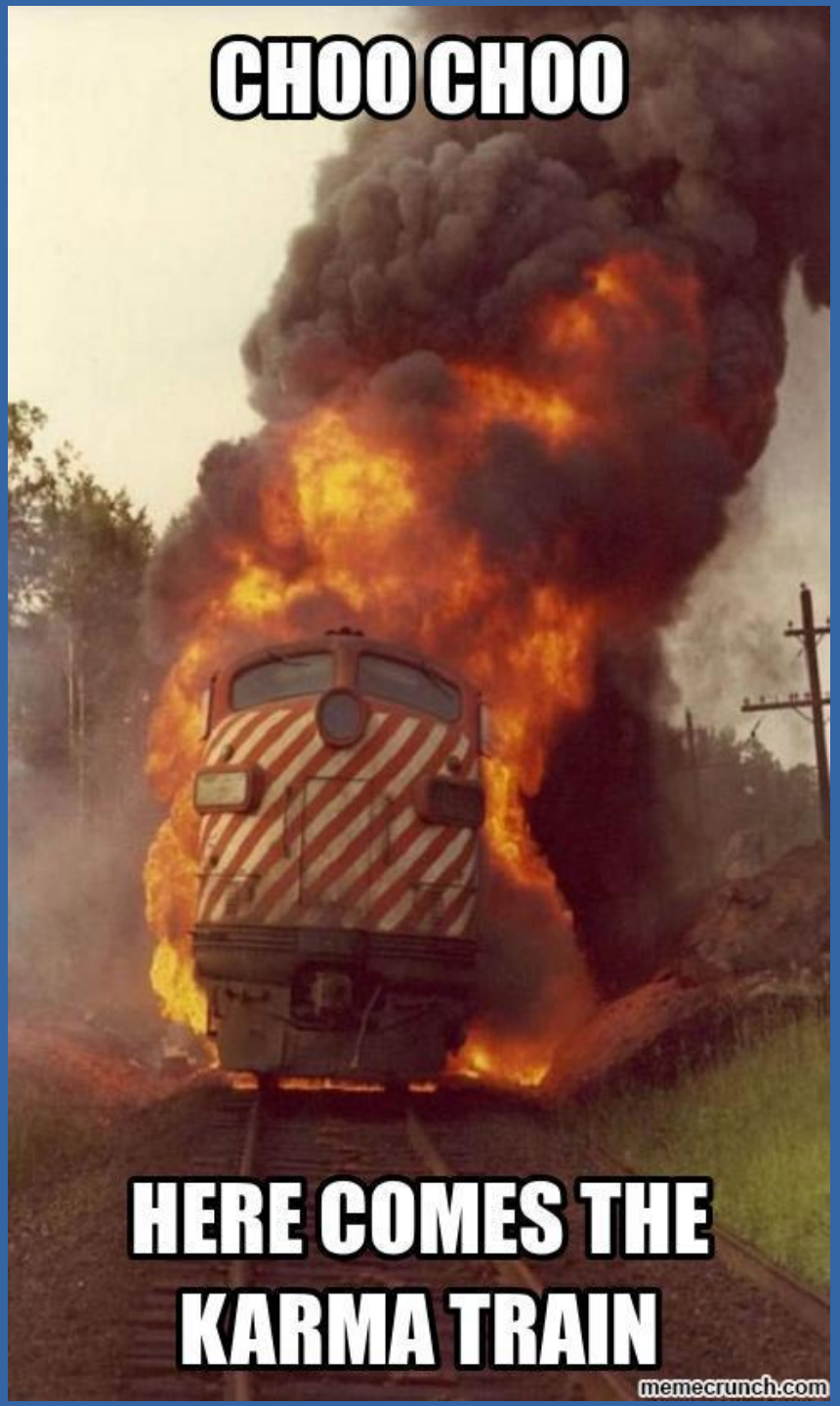
Some procurement tips

- Put a spreadsheet together, work out some builds
  - **Most** resellers give 40-45% discount easily
  - You need “Cisco Pricing Support” for more
- **Pick a good reseller**, explain what you want
- Be aware that bundles could be available
- Trade in those 7600s – it might be more worthwhile
- Don't rule out “refurbished” kit...

  You *might* be able to get what you need

Where are we now?

- Implemented a mixture of 9001 and 9006 chassis
  - Dependant on the power limits in PoPs
  - Front-to-back cooling a must for the 9006s
- Some Trident cards, but replacing with Typhoon
- All peering & transit is now run by ASR9k
- Targeting October to remove the last two 7600s
  - Bought some RSP880-RL!
- Running 5.3.x, but we have a 9001 on 6.0.2 and it's happy. Looking toward 6.1.x QNX later this year.


CHOO CHOO
HERE COMES THE KARMA TRAIN
memecrunch.com

Gotchas

- The 9006 baffle claims to be 3U in total. More like 3.3U.
- 3kW AC PSUs come with C22 sockets. Not C20. Whhhy?
- Kernel dumps on the 9001 don't get written; 4.3.2 included a workaround to have them written to eUSB
- On-board RSP ports are not in the regular forwarding path, which is frustrating for local Netflow Exporters
- Typhoon cards still need workarounds for BGP ASNs in IPv6 netflow exports; knob coming in 6.1.1

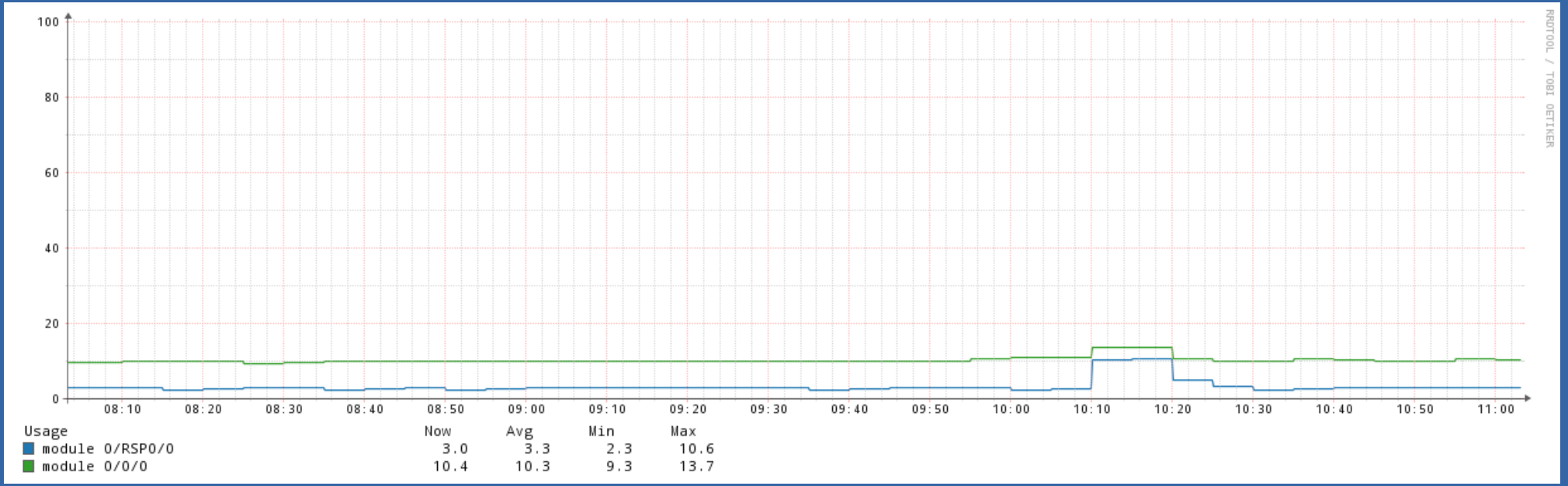
Gotchas

- MTUs are counted differently between 7600/6500 and ASR9k: remember to add/subtract 14b to MTUs
- Security caught us out: SSH ACLs are not good enough alone. 'Securing IOS-XR' guide is VERY useful.
- Don't delete the built in 'admin' (like UNIX root) account. It demands that you create a new one on the console(!)
- Upgrade, upgrades, upgrades. URGH. Looking forward to the Linux base, BUT it's Tomahawk-only.
    - The individual upgrade guides need careful study

Despite all of this, they do work

… And we're quite happy with them

Calm in an Emergency



This was during a period of heightened broadcast at an IXP

Key Takeaways

- Remember that NPUs aren't ASICs, and lose performance with features
  - Ingress ACLs lower packet throughput, for example.
  - The trade-off is less limitation on possible features
- Some of the feature benefits we get from this hardware
  - Segment Routing (death to 6PE!)
  - FAT-PW (et al) for L2 tunnels over port bundles
- Upgrading isn't just about bandwidth or port speeds
- Don't take my word for it; there's always something new *just* around the corner...

Stuff to Read

- **BRKARC-2003**, from Cisco Live (register, it's free)
  - Hardening IOS-XR devices
  - http://www.cisco.com/c/en/us/about/security-center/increase-security-ios-xr-devices.html
- IOS-XR upgrade guides
  - http://www.cisco.com/web/Cisco_IOS_XR_Software/index.html
- The Cisco GPL
  - http://www.ciscoprice.com

Thanks! Any questions?

Flammenwerfers willkommen...

Contact: @tomm3h on Twitter