First things first...

- IXLeeds director since March 2013

- Looks like this:

- Runs Portfast Ltd (http://www.portfast.co.uk/)

IXLeeds

- UK's first regional IXP that still exists
- ~3Gbps (as of March 2014)
- 11 members, mix of CDN, hosting, academic & access

Hardware History 1

- Switch 1 – HP Procurve.

- 48 x copper GE

- Few other features

- Pros – Free!

- Cons – Everything else about it
  - Including broken ports!

Hardware History 2

- Switch 2 – Brocade (Foundry) RX
- Modular chassis with mix of 1 & 10G
- Generously lent to us by LINX (Thanks!)

- Pros –10G, plenty of ports, Free!
- Cons – power consumption in the region of 2kW

New hardware!

- Edgecore ES5652BT1 – 48x10G, 4x40G
- Quanta T1048-LB9 – 48x1G, 4x10G

- Pros
  - Cheap! Same broadcom forwarding ASIC as switches many times the price
  - Running Debian!
- Cons – Unknown territory

Anatomy of a small switch

PSU

PowerPC SoC

Broadcom forwarding ASIC

Ports
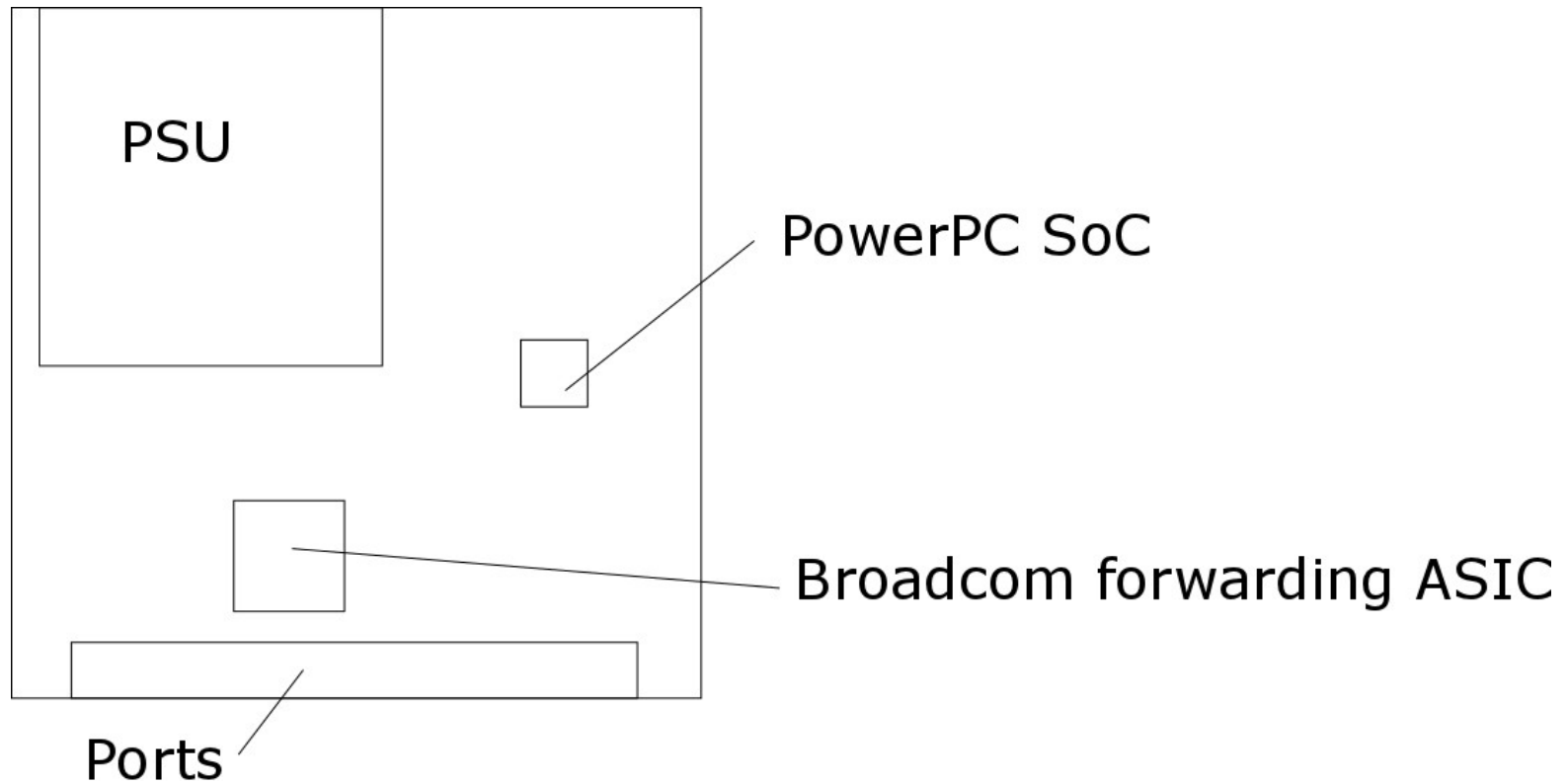
Why Cumulus?

- All other IXPs in the UK apart from LINX Juniper are Extreme based

- Some vendors completely out of our price range

- Interesting and open new development

- We are small and want to innovate & stand out

Design

- Two switch design
  - Separate 1G and 10G connections due to "head of line" buffering issues inherent in the Broadcom design (affects all vendors that use these chips)
  - Only 9MB buffer space for the whole switch
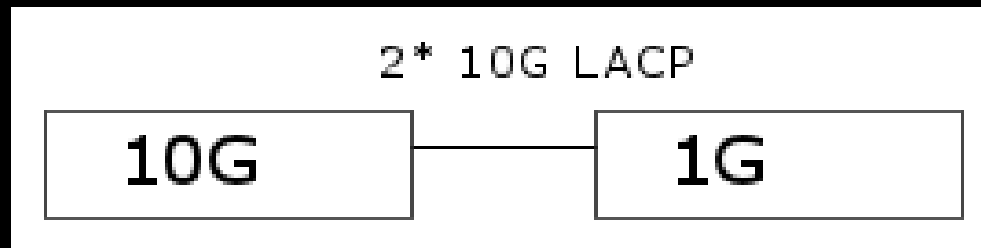
Deployment

- Switch configuration was prepared, switches were tested in the lab

Disaster!

- 1G Switch CPU was crashing

- Forwarding plane continued to work (somewhat)

- Software was changed, features disabled, still the switch crashes

Problem solved

- Changed the hardware, everything now happy.
- Cumulus were very helpful throughout
- They sent free Cu SFPs – just in case!

Internals

```
tom@ixlsw1:~$ uname -a
Linux ixlsw1 3.2.46-1+deb7u1+cl1 #3.2.46-1+deb7u1+cl1 SMP Fri Feb 7 13:15:34 PST 2014 ppc powerpc GNU/Linux

tom@ixlsw1:~$ free
             total       used       free     shared    buffers     cached
Mem:       2008316     263848    1744468          0      24532     127408
-/+ buffers/cache:      111908    1896408
Swap:            0          0          0

tom@ixlsw1:~$ cat /proc/cpuinfo
processor       : 0
cpu             : e500v2
clock           : 999.999996MHz
revision        : 5.1 (pvr 8021 1051)
physical id     : 0
core id         : 0
bogomips        : 125.00

processor       : 1
cpu             : e500v2
clock           : 999.999996MHz
revision        : 5.1 (pvr 8021 1051)
physical id     : 0
core id         : 1
bogomips        : 125.00

total bogomips  : 250.00
timebase        : 62500000
platform        : Accton Technology Corporation ES5652BT1
model           : accton,es5652bt1
Vendor          : Freescale Semiconductor
PVR             : 0x80211051
SVR             : 0x80ea0021
PLL setting     : 0x4
Memory          : 2048 MB
Memory          : 2048 MB
```

Configuration

- It's just a Linux box (this is great!)
  - /etc/network/interfaces
  - ifup / ifdown
  - ethtool
  - brctl
  - bwm-ng
  - apt-get

Port security & L2 hygiene

- Does anyone make a "port security" feature that actually works?

- Layer 2 providers leak

- Filter it all!
  - Ethertype
  - MAC address
  - Storm control

Port security & L2 hygiene

```
# Akamai
auto swp1
iface swp1 inet manual
    pre-up ethtool -s $IFACE speed 10000 autoneg off
    up ip link set dev $IFACE alias "PEER: Akamai"
    up ip link set dev $IFACE mtu 9000
    up ip link set $IFACE up
    down ip link set $IFACE down


# swp1 Akamai
-A INPUT -s c8:4c:75:e9:11:3f -p 0x0800 --in-interface swp1 -j ACCEPT
-A INPUT -s c8:4c:75:e9:11:3f -p 0x086dd --in-interface swp1 -j ACCEPT
-A INPUT -s c8:4c:75:e9:11:3f -p 0x0806 --in-interface swp1 -j ACCEPT
-A INPUT --in-interface swp1 -j DROP
-A FORWARD -s c8:4c:75:e9:11:3f -p 0x0800 --in-interface swp1 -j ACCEPT
-A FORWARD -s c8:4c:75:e9:11:3f -p 0x086dd --in-interface swp1 -j ACCEPT
-A FORWARD -s c8:4c:75:e9:11:3f -p 0x0806 --in-interface swp1 -j ACCEPT
-A FORWARD --in-interface swp1 -j DROP
```