No, Bob, the “Cloud” is not the answer

Chris Malton

Who am I?

- Started as a Software Engineer at a CRM company
- Moved to an ISP in Southampton in 2014
- Left there in 2016 – Went to work for Swlines
  - IT Consultancy – primary client base involved with railways.
- Now – Technology consultant specialising in “difficult problems”
- What do I do?
  - Software Engineer
  - SysAdmin
  - Hardware Designer
  - Network Ops Team
- So a bit of everything then.....

The “Cloud” - What do I mean

- Lots of definitions
- In the case of the client this relates to “the Cloud” refers to “public cloud”
- That is your Amazon AWS, Microsoft Azure, Bytemark Cloud, etc. Their hardware, your virtual machine.

Public cloud is great...

- … for small projects.
- … for research and development purposes.
- … for large scale projects if you've got money to spend.

… but what if you haven't got money to spend?

- Running your own hardware is, in fact, sometimes cheaper.
- It costs money to set up.
- The annual costs are not bad for what you get.
- As long as you don’t spend hours dealing with unpredictable hardware failures.
- RIPE membership required for big blocks of IPs.

My client

- Information screens in the public transport sector
- Started out doing print media, and moved into digital.
- Started developing in AWS – because it was cheap.

The architecture

- Client machines connect back to servers over a VPN.
  - or over public internet.
- Servers provide data including service information, journey planning etc.
- Calls out to third party services for most of the data.

Cheap turns not cheap

- It's a problem of scale.
- The backend isn't resource-light.
- Three c4.large AWS instances load-balanced for 100 terminals.
  - What happens when we hit 200 terminals? What about 300?
- One availability zone
  - How do you reliably scale into two when you have VPNs?

The addressing headache

- Client uses a /16 in single subnet in Amazon – /28 exposed over VPNs
- Terminals have hard-coded IPs for load balancer in AWS.
- We have no control over the IP addresses used for terminals (and it's all IPv4!)
- Ends up getting messy – very quickly!

Future solutions to the addressing issue?

- NAT – but NAT is evil.
- VRFs – Gets complex to manage – but all customers can be kept separate
- Something else?

So... Why not run it ourselves?

- Management & admin headaches
- Public IP Addresses (and working failover)
- Our client has no IT support engineers (that's why they have us!)
- That huge (5-figure) setup cost.

The hardware went in two by two.

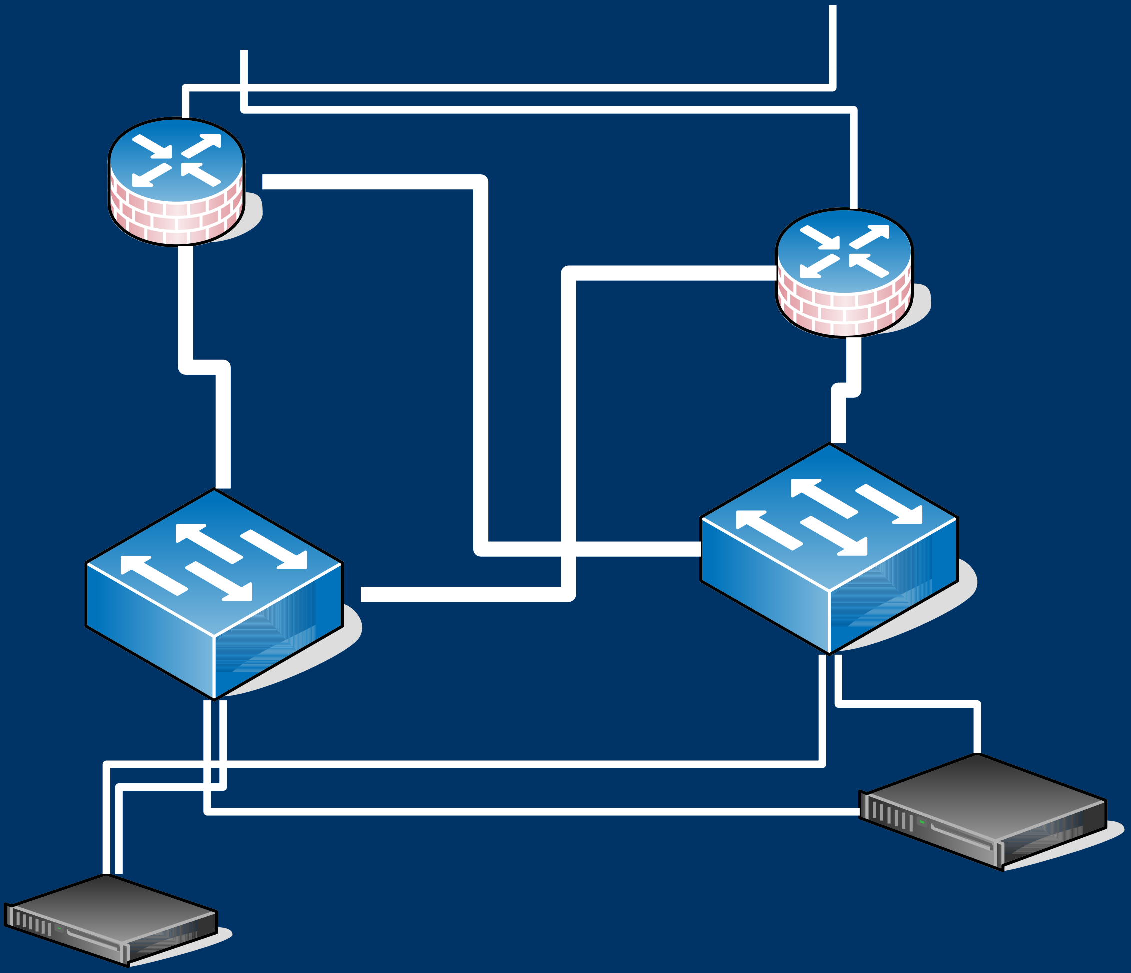
- 2 sites
- Each site with two firewalls
- Each site with two switches
- Each site with two servers

- And it's all wired for resilience....
- Servers have connections to each switch
- The switches are both connected to each firewall
- The firewalls are connected back to each switch
- And there's dual uplinks, one to each switch.

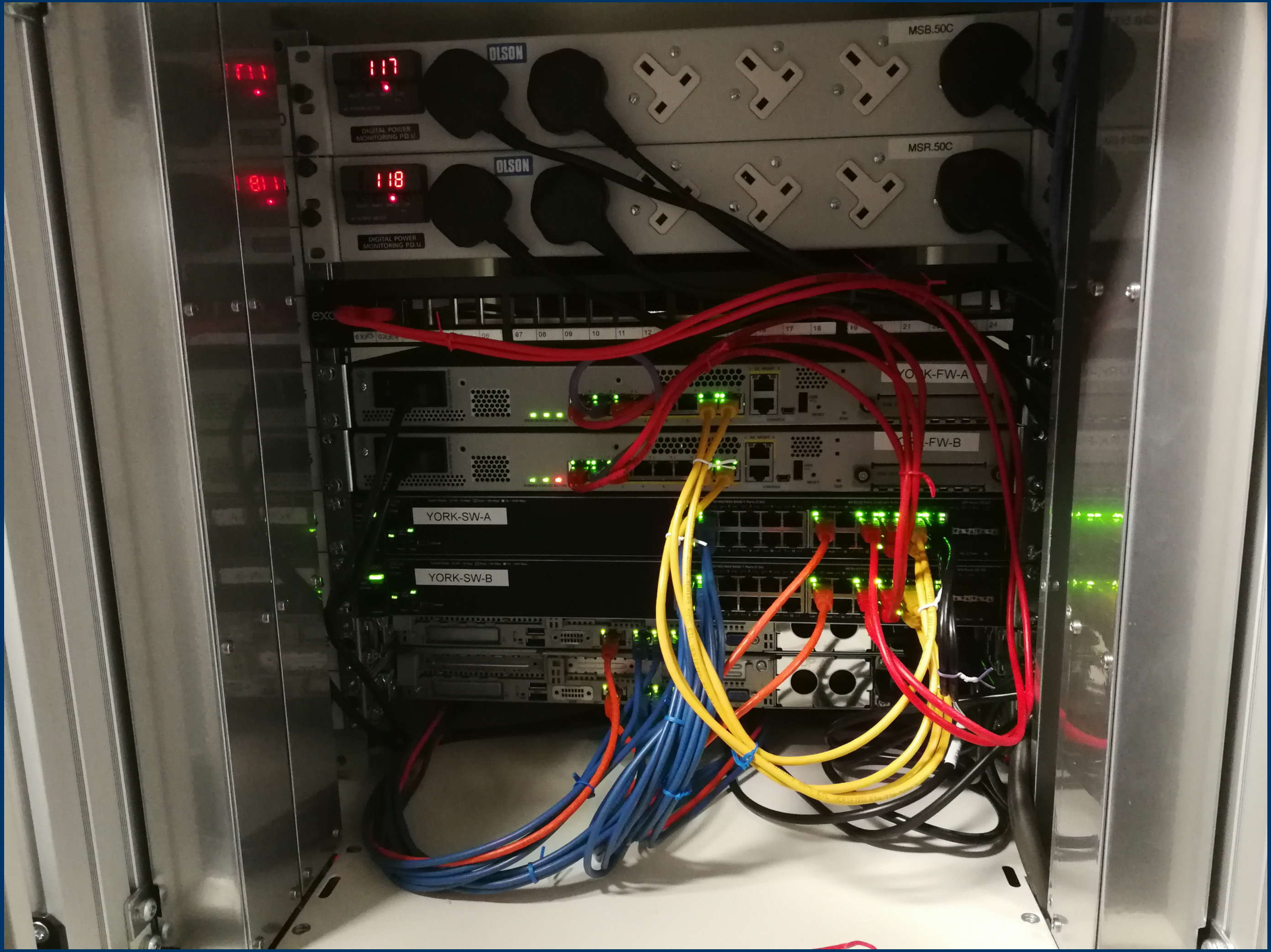
The hardware

- Cisco ASA 5508X with Firepower
- HP 3520 switches
- HP DL360 Gen 9 Servers
  - Dual processor, 8 core hyperthreaded
  - 128GB of RAM
  - 1.8TB of local storage per server

Here’s how we planned it



Here's what both racks look like


MSB.50C
OLSON
117
DIGITAL POWER MONITORING P.D.U.
MSR.50C
OLSON
118
DIGITAL POWER MONITORING P.D.U.
YORK-FW-A
-FW-B
YORK-SW-A
YORK-SW-B

The software

- Hypervisor: kvm + qemu
- High availability: keepalived
- Load balancer: haproxy
- Web server: nginx
- Management: puppet

- Theme here: It's all open source software

Two by two by two by two by…

- Two load balancers
  - keepalived managing the floating virtual IP
- Two database servers
  - keepalived managing the floating virtual IP
- Same for outbound proxy server
- Same for VPN servers
  - Works properly only if you use Dead Peer Detection.
- Same for DNS

Keepalived – The bit that makes it work!

- Allows you to have a whole group of machines.
- Uses IPv4 Multicast for v4 VRRP
- Give it an address, an interface, and a shared secret
- It just gets on with the job – no questions asked.

Getting stuff to the backends

- HAproxy is another awesome tool.
- Serves requests at backends that are up.
- Takes down backends out of the pool.
  - Upgrading a server is as simple as stopping the web server, upgrading it, and rebooting.
- Very high performance

How much does it actually cost to run?

- It costs less than 4 figures a month to run this.
  - In total across both sites!
- That’s under half the cost of running it in Amazon!
- Support costs have increased at the moment due to some serious teething issues with the servers which we’re working with HP on.

Moving forward

- HA plugin for the VPN servers – allows seamless failover.
- Auto-failover between sites (Fun and games with BGP).
- Clustered MySQL (difficult with 2 sites).

Who did we work with

- Stuart Hill & Andrew Doble (CDW)
- Alex Webb (4D Datacenters)
- Tom Hill and Nat Lasseter (Bytemark)

Questions?

- Ask me now
- Email me: chris@deltav-tech.co.uk