SHARKFEST 2015

WIRESHARK DEVELOPER AND USER CONFERENCE

Root cause in complex networks

Tips and lessons from fast-paced and enterprise financials

Chris Bidwell, chris@cbidwell.com

Chris Bidwell

- Computer scientist & packethead
- Lapsed software developer
- 9 years in networking for financial companies
- All views expressed are my own

Disclosure

- I like when things work properly
- My experience is based on demanding clients to whom **any** interruption is an outage
- Stability is paramount in my world
- ... expect a lot of TLAs!

Contents

- Background
- Recommendations
  - 'Best practices'
  - Tools
  - Strategy
- Examples/anecdotes

Complex Networks

- As networks grow, usually their complexity grows
- 'Creative'/point solutions can be unwieldy
- Documentation is rarely accurate
- Shadow IT solutions
- Software evolves
- Automation can go horribly wrong

Root Cause

- Hunt it to improve efficiency and productivity... sometimes!
- Important in large-scale operations
- Essential to be multidisciplinary
- Know your environment
- Accept that others may not care or understand

Recommendations

Keep Good Time

- Be consistent
- NTP (4+ srcs), PTP
- Use correct timezones + DST
- Validate and monitor your clocks, sync status, grandmaster etc.
- Offset can indicate asymmetric latencies

Log Everything

- emerg,alert,crit,err,warn,notice,[info]
- Local + remote, timestamp both
- 30day+ retention
- Correlate, group, aggregate
- Validate and monitor your logging
- Syslog is (usually) UDP

Poll/Trap Everything

- Interfaces: in/out bits, pkts, err/drops, speed/duplex
  - 1min 'aggressive' but necessary
- Inventories: h/w models/SNs, s/w state: MAC, ARP
  - Routing/CEF?
- Env: PSU + Fans
- Storage, RAM

Backup Everything

- Test your configs are recoverable (!!!)
- RTCD + nightly downloads
- Know how to search your archives
- Track changes over time

e.g. Cisco ASA:
Spot The Difference

```
ciscoasa# show run | inc community
snmp-server host inside 192.168.0.161 community ***** version 2c
snmp-server community *****
ciscoasa#
```

```
ciscoasa# more system:running-config | inc community
snmp-server host inside 192.168.0.161 community private version 2c
snmp-server community private
ciscoasa#
```

(Also affects PSKs for VPNs)

Capture Everything

- OK, maybe not everything
- Tag, slice, filter where it makes sense
- 'Packet brokers'
- Acknowledge inaccuracy, be grateful for the insight
  - SPANs, TAPs each introduce sources of error
- NetFlow/sFlow/IPFIX etc.
- Hosts: sysdig

Test For The Unexpected

- Beware of writing tests that only prove what you expect or only test correct configuration & state
- Baseline your setup
- Prove your tools work before you need them
- Check for regression

Underpinnings

Media + L1 Foundations

- Copper
  - Faulty plugs, shorts/cuts: CRC/100Mb
- Fibre/optical
  - Low RX => errors, flapping, err-dis
  - Dirt, kinks, pinching, droop
- RF
  - Line of sight(ish) obstructions, weather

L2+3 Foundations

- MAC, ARP, DHCP, NAT + aging/expiry
- RPVST, LACP/LAG (+LB)
- FHRPs + SSO/NSF
- Controlled roots, HA priorities
- Static/dynamic routing, redistribution
- ECMP (+LB), costs/metrics
- RPF

Hardware Foundations

- Everything's finite
- RAM, CAM, TCAM
- Dedicated + shared buffers
- Timers + clocking rates
- Encapsulation, fragmentation
- **Virtualisation = resource sharing**

WAN Connectivity

- Different failure modes, SLAs
  - Internet, MPLS, Leased line
  - Wavelength, Dark fibre
  - Microwave, Satellite
- Loss, jitter due to queuing and QoS

Higher/Application Level

(Dynamic) Services

- DNS (bleugh WINS.. NIS?!)
- DHCP
- LDAP, AD/KRB, TACACS, RADIUS
- Load balancers
- Application proxies
- Firewalls, IPS
- ... Should all support logging + audit!

Application Latency

- RTT (BDP), loss, OOOPs
- DNS/Auth timeouts
- Host resource contention
- Crappy code

Tools

Your Tools

- Make them work for you
- Spend some time to learn and customise
- Configure shortcuts and keep cheatsheets

The Armoury

Packet capture tools	***Wireshark, tshark, dumpcap, editcap, mergecap***
Text Editor	***emacs/vi/vim, notepad++, TextMate***
GNU tools	***awk, sed, grep, cut, sort, uniq***
Scanners/probes/active discovery	***nmap, zmap, Angry IP, nessus, nc, portqry***
Terminal emulator	***PuTTY (+CM)***
Quick references	***cheatography.com, this new thing called google.com***
Traffic/packet generators	***iperf, hrping, scapy***

*examples, not necessarily endorsements

$#!@ Happens

The Perfect Storm

- Lurking inefficiencies
- Conflicting dynamic protocols
- Traffic growth, contention
- Lack/loss of resiliency
- Upgrades (any)
- Config changes
- The Boss is in town

Situational Awareness

- See it for yourself if possible
- Keep good notes
- Look out for patients lying (give them the opportunity to tell the truth)
- Many problems, not all are yours
  Not all problems are problems...
- Organisational factors

Tracing The Path

- L2 path: blockedports, root
- LACP load-share hashing
- L3 path: candidate routes in RIB
- FIB
- Load, loss, QoS
- Filtering, NAT, tunnelling

Got Root?

Test The Hypothesis

- Sanity-check the logic
- Be wary of extremely complex ideas
- Be wary of extremely simple ideas
- Explain it to yourself, question assumptions
- Double-check your measurements
- Bounce it off colleagues/peers

Prepare the ammunition

- Write up the case, be verbose
- Timeline often invaluable
- Include evidence
- Ensure all symptoms are addressed
- Proof read, re-read, peer review
- Be diplomatic rather than critical
- Line up fixes

Examples

1. 2K8Std + RAM + OOOPs

- Users begin complaining that file transfers are slow

- 1Gbps bottleneck, 3ms RTT, <5Mbps?!
- Affected seemingly random hosts

- Multi-point packet capture
- Flow analysis

- In-line Crypto + HW Cfg/OS

2. Baseline test: ARP?

Q: How many ARPs per minute do you expect on a LAN?

- 10?
- 100?
- 1,000?

2. Baseline test: ARP?

Q: How many ARPs per minute do you expect on a LAN?

- Count request/response pair as 1
- 4x hosts (3 appliances, 1 Win2012)
- 4x L3 switches
- /24 subnet

2. Baseline test: ARP?

Q: How many ARPs per minute do you expect on a LAN?

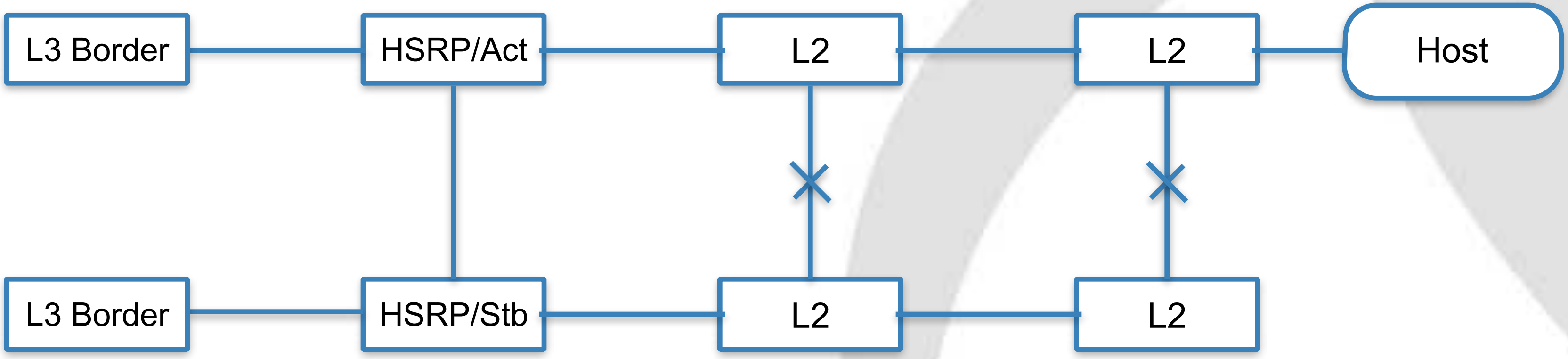
- Count request/response pair as 1
- 4x hosts (3 appliances, 1 Win2012)
- 4x L3 switches
- /24 subnet

A: ... ~120,000 ARPs/min (WTF?)

2. Baseline test: ARP?

- Identified with Wireshark on Win2012
  - 2x 10Gbps NICs across 2 switches (paired)
  - Security software, including pre-boot network authorisation for full-disk encryption... suspicious!

3. 'Big' L2 domain flooding



3. 'Big' L2 domain flooding

- High utilisation reported on nearly every uplink (1Gbps)
- No observable negative impact