

# TCP Anycast - Don't believe the FUD

Operational experience with TCP and Anycast.

Matt Levine (CacheNetworks), Barrett Lyon  
(BitGravity), Todd Underwood (Renesys)

# What's (IPv4) Anycast?

- From a network perspective: nothing special.
- Just another route with multiple next-hops.
- Service(s) exist on each next-hop and respond from the anycast ip address.

# It's the packets, stupid.

- The Perceived Problem: Anycast is not a stable environment for stateful protocols (notably TCP), especially “long lived” sessions.
- eg: High-Def Porn Downloads.
- Other presentations have made reference to existing deployments..that's us!
- TCP Anycast not only works, it has been used in production for years.

# Anycast at CacheFly

- Deployed in 2002.
- Prefix announced on 3 continents.
- 3 POP's in the US
- 5 “common carriers (transit)” + peering
- Effective BGP Communities from upstreams is key.

# Proxy Anycast

- Proxy traffic is easy to anycast!
- Customers are isolated on a VIP/virtual address.
- The virtual address lives over common carriers allowing even distribution of traffic.
- State is accomplished with custom hardware.

# Node Geography

- Anycast nodes that do not keep state must be geographical separated.
- Coasts and countries appear to provide enough padding for route instability.
- Nodes that are near by could possibly require state between each node if routes are unstable.

# IP Utilization

- 'Anycast is wasteful'
- Really? How much IP Space do you need to advertise from 4 sites via unicast?

# Carriers and Peering

- For content players, having even peering and carriers is key.
- Having a European centric transit provider in the US without having the same routes in Europe could cause European traffic to home in the United States.
- Use quality global providers to keep traffic balanced.



# Carriers and Peering

- When peering...
  - Keep in mind that a peer may isolate traffic to a specific anycast node.
  - Try to peer with networks where it makes sense.
  - Try to make sure your peers know what you're doing, and/or have a good community set.

# Benefits of Anycast

- For content players, anycast can help with:
  - Moving traffic without major impact or DNS lag.
  - Provides buffers for major failures.
  - Allows for simplistic traffic management, with a major (potential) performance upside.
  - The ability to interface tools to traffic management.

# Data - May 9, 2006

- Renesys: Monitored changes in atomic-aggregator for a CacheFly anycast prefix.
- Keynote: Monitored availability and performance of 30k file.
- Revision3: Monitored behaviour of 'long lived' downloads of DiggNation videocast - Total of 7TB transferred.

# Renesisys Data

- 130 BGP updates for May 9.
- Observed 34 distinct 'POP Changes' (monitoring atomic aggregator property on routes).
- 130 updates considered 'quite stable prefix'.

# SJC “Event”

- Between 07:00 and 07:35 UTC - observed 98 updates, 20 aggregator changes.
- Unable to correlate these shifts with any traffic changes - most likely we don't have a big enough sample size.

Okay..so BGP seems  
ok..what about actual  
application stability, is  
TCP really stable??

NO :(

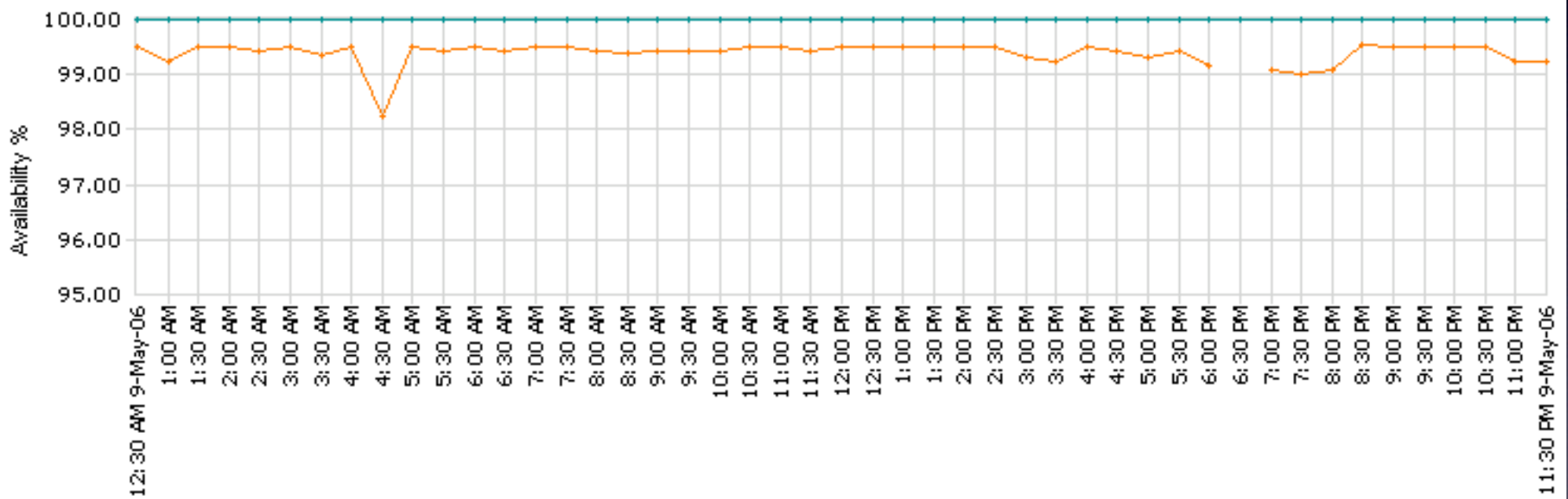
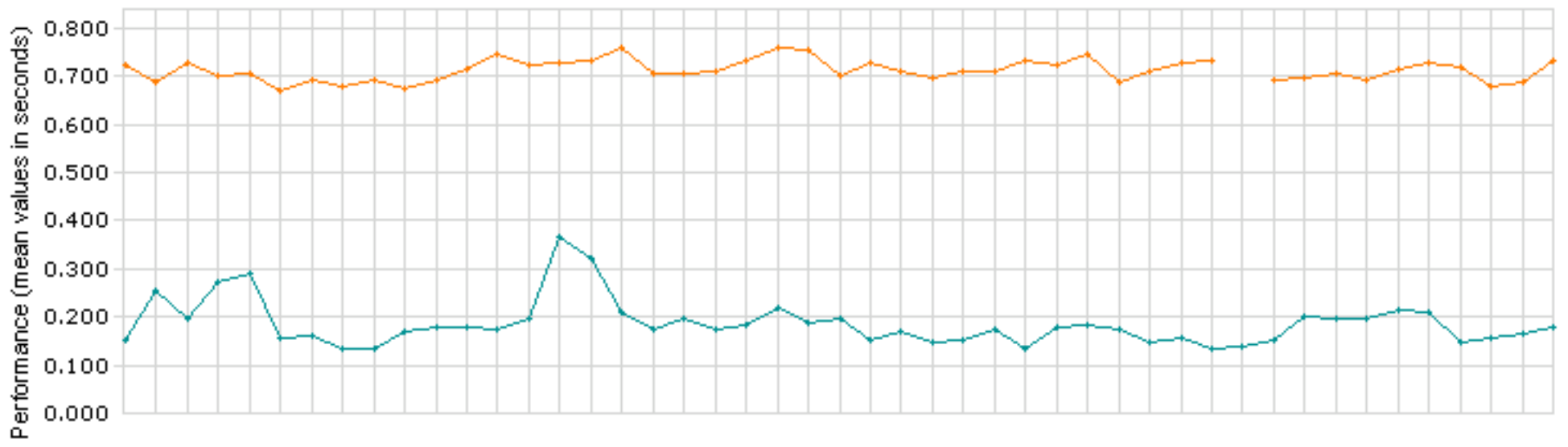
Let's go shopping

Just Kidding

# (Short-Lived) Keynote Data

- 30k download sampled from 31 locations every 5 minutes. (or an average of 1 poll every 9.6 seconds)
- Compared against the 'keynote business 40'
- Data Collected from May 9, 2006





© 1998-2005 Keynote Systems, Inc.

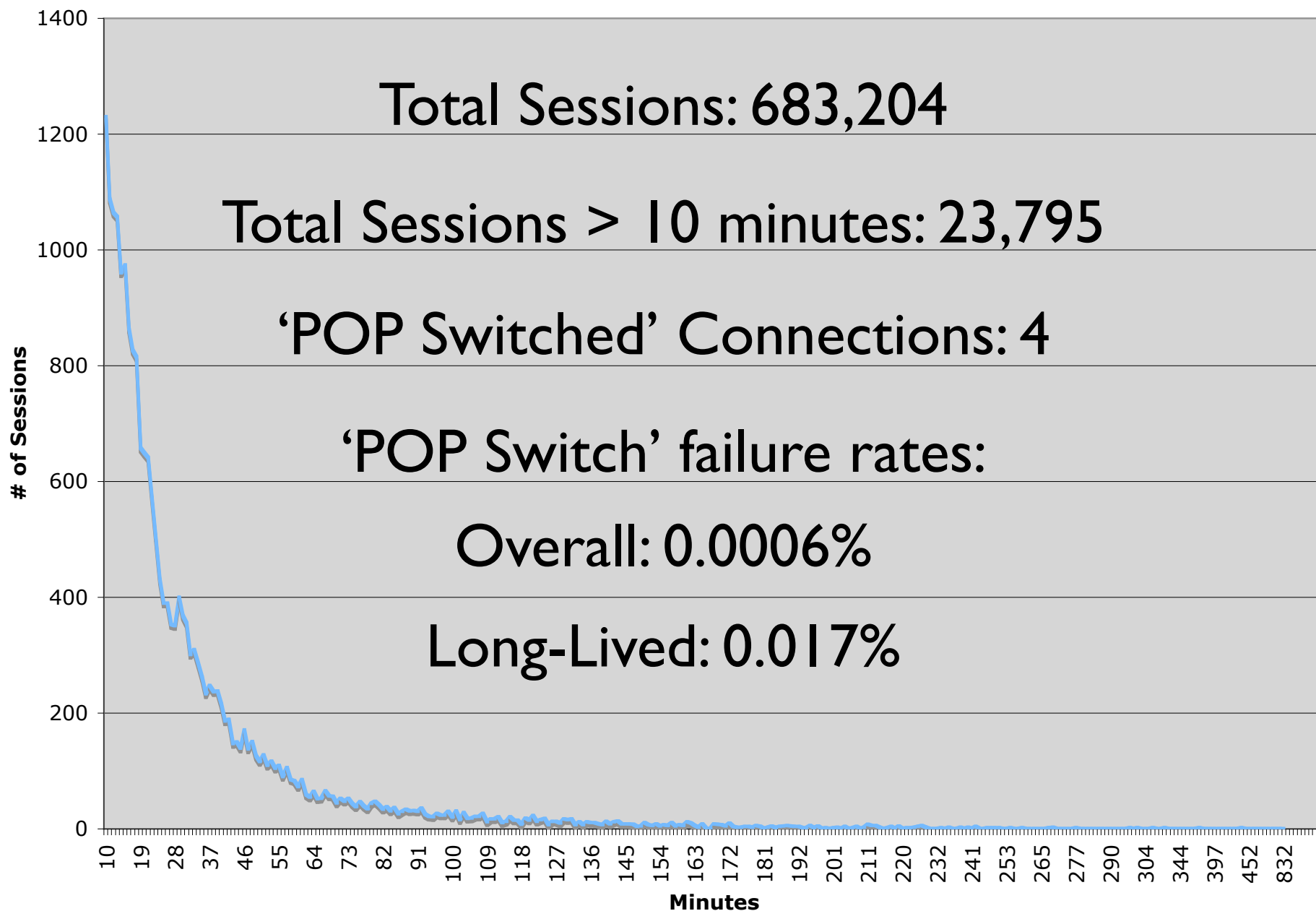
Green line: Anycast  
 Orange: Keynote Business 40

# Analyzing Revision3 Data

- Monitored IPTV downloads for Revision3 Downloads for 24 hours (thanks, Jay).
- Methodology: Analyze packet captures - look for new TCP sessions not beginning with SYN.
- Compare that against global active connection table.

# Long-Lived Data

TCP Session Length



# Anycast Gotchas

- Large-Scale Changes in provider policies can impact your traffic, and it's up to you to figure out what changed.
- “Things that are bad” become worse, notably per-packet load balancing across provider or topological boundaries.

# Conclusions

- In our experience, stateful anycast is not inherently unstable, and failure/disconnect rates are inline with offering unicast services.
- This is counter-intuitive to some published data from previously published data.
- “Trust us, it works.” (tm)
- Widespread failures cause havok; however the internet doesn't go crazy \*that\* often.

# Transitioning to IPv6

- We have a plan!
- The plan consists of being dead by the time customers demand v6.

# What you can do

- Stop telling people anycast doesn't work for TCP if you haven't tested it, it just makes us mad.
- If your application cannot handle TCP/IP failures gracefully, do not run anycast - in fact, don't run it on the internet.
- Experiment
- Share your experience - we want to know if we're crazy or not.

Questions?