Opportunities for Exploiting Social Awareness in Overlay Networks

Bruce Maggs
Duke University
Akamai Technologies

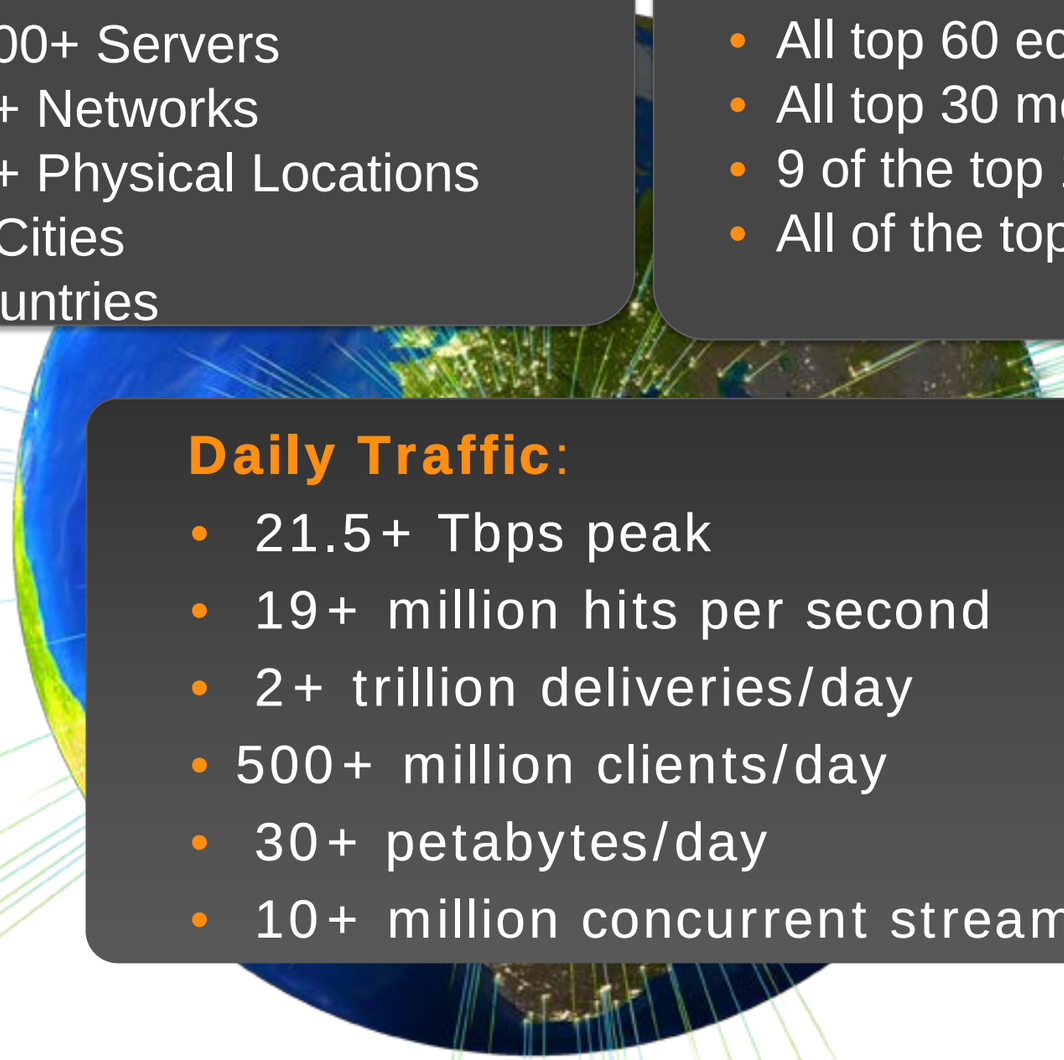
The Akamai Intelligent Platform

A Global Platform:

- 127,000+ Servers
- 1,100+ Networks
- 2,500+ Physical Locations
- 650+ Cities
- 81 Countries

Delivering 130,000+ Domains

- All top 60 ecommerce sites
- All top 30 media & entertainment companies
- 9 of the top 10 banks
- All of the top Internet portals



Daily Traffic:

- 21.5+ Tbps peak
- 19+ million hits per second
- 2+ trillion deliveries/day
- 500+ million clients/day
- 30+ petabytes/day
- 10+ million concurrent streams

Systems and Opportunities

1. Akamai KONA – protection against denial of service attacks

2. Akamai NetSession – peer-to-peer content delivery

Attacks on Akamai Customers

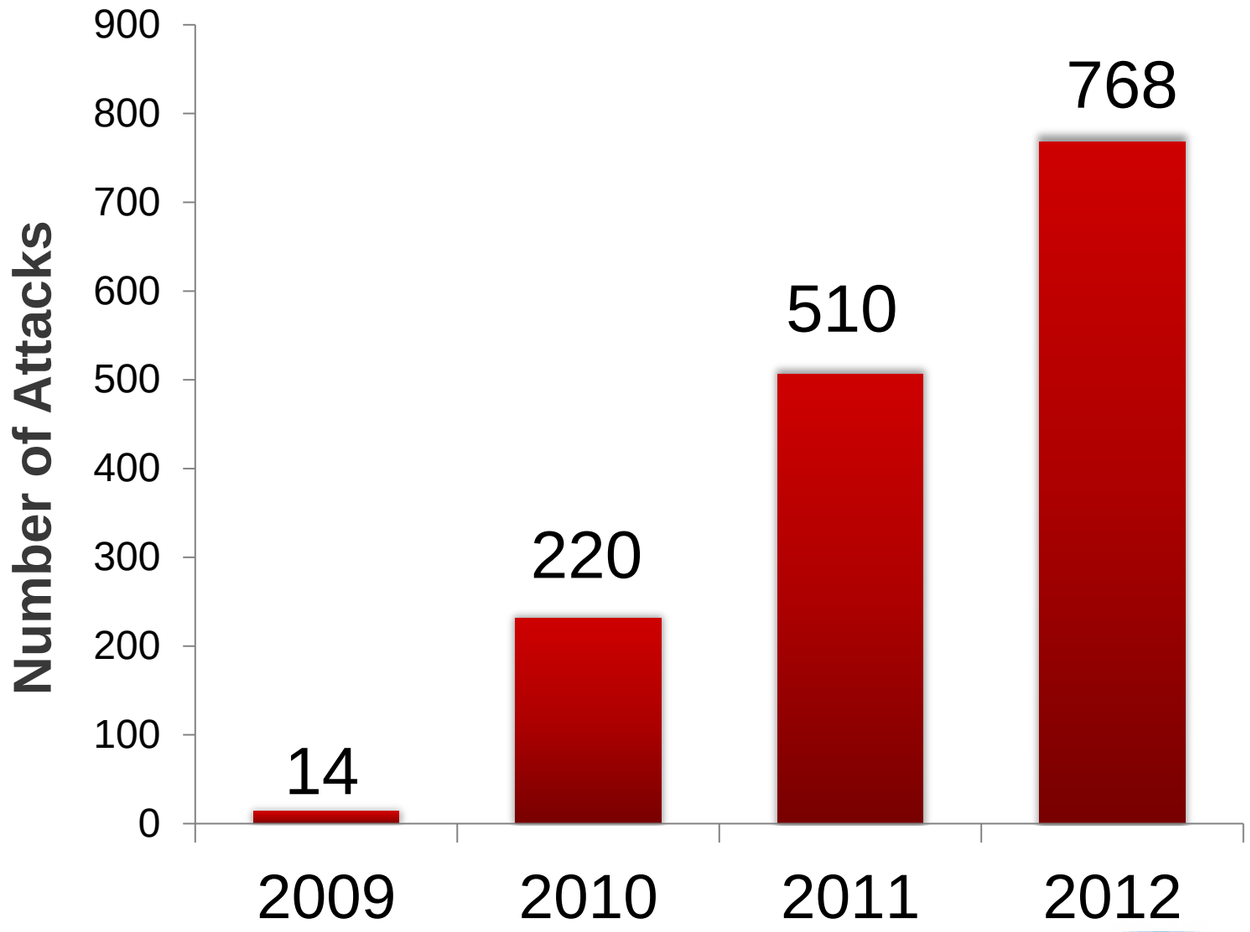
Typical Attack Size

10 Gbps

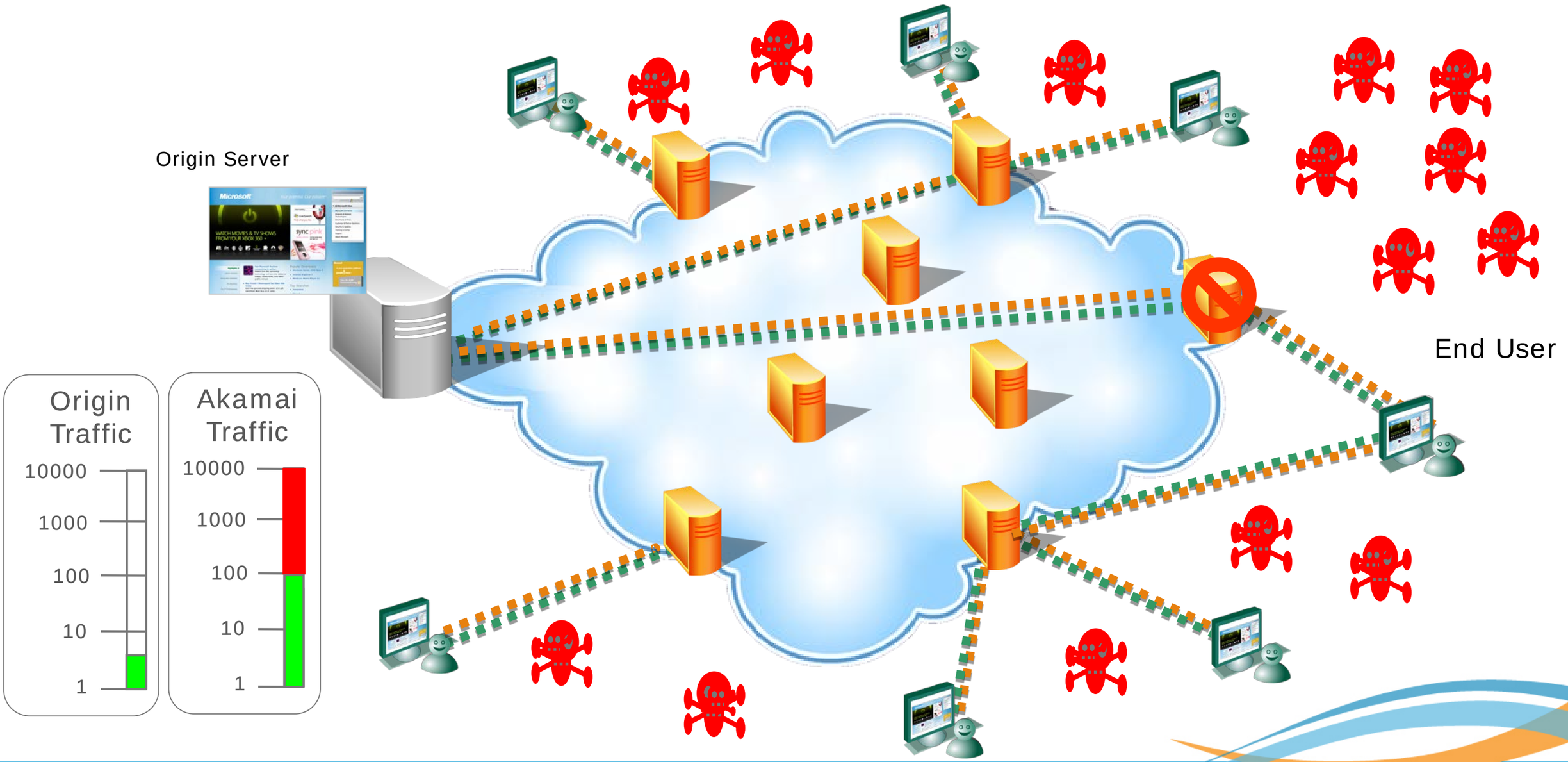
Large Attack Size

100+ Gbps

Attacks are originating from all geographies and are moving between geographies during the attack

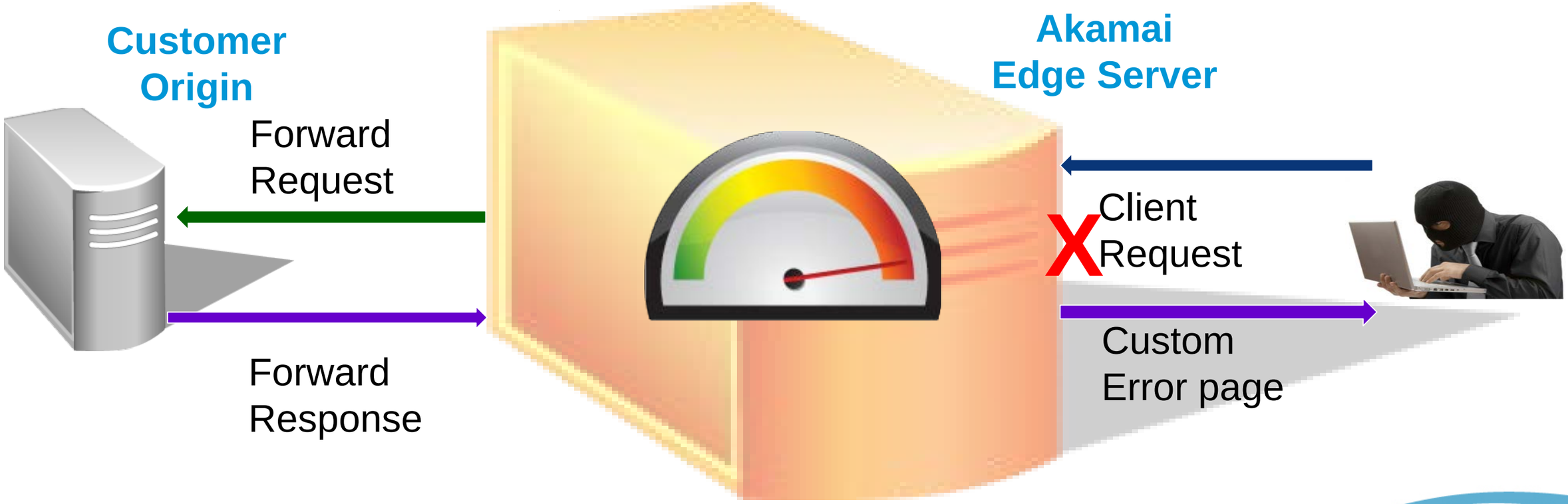


The Akamai Platform Provides a Perimeter Defense



Defeating HTTP flooding attacks – Rate Controls

1. Count the number of Forward Requests
2. Block any IP address with excessive forward requests



Operation Ababil

“none of the U.S banks will be safe from our attacks”

Phase 1	Phase 2	Phase 3
Sep 12 – Early Nov 2012	Dec 12, 2012 – Jan 29	Late Feb 2013 – Now

- DNS packets with “AAAAA” payload
- Limited Layer 7 attacks
- Early-mid Oct 2012 announced names of banks where attacks succeeded
- (Did not announce bank names if attacks were unsuccessful)
- Began use of HTTP dynamic content to circumvent static caching defenses

- Incorporate random query strings and values
- Addition of random query strings against PDFs
- Additions to bot army
- Burst probes to bypass rate-limiting controls
- Addition of valid argument names, random values

- Multiple probes
- Multiple targets
- Increased focus on Layer 7 attacks
- Target banks where attacks work
- Fraudsters take advantage

Phase 1 Attack – Sept 2012

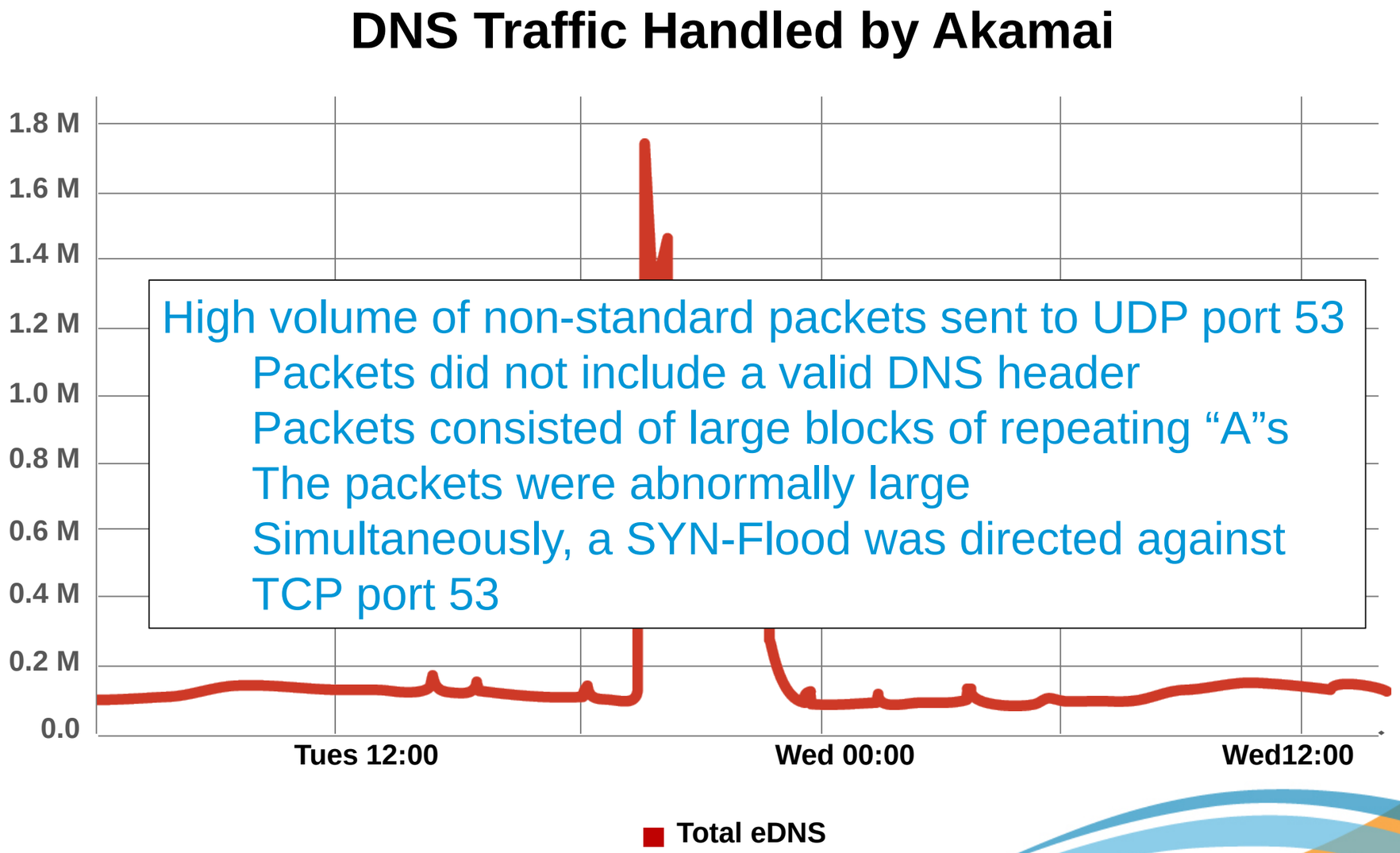
Attack Traffic:

23 Gbps

(10,000X normal)

Duration:

4.5 Hours



Phase 2 Attacks - January 2nd, 2013

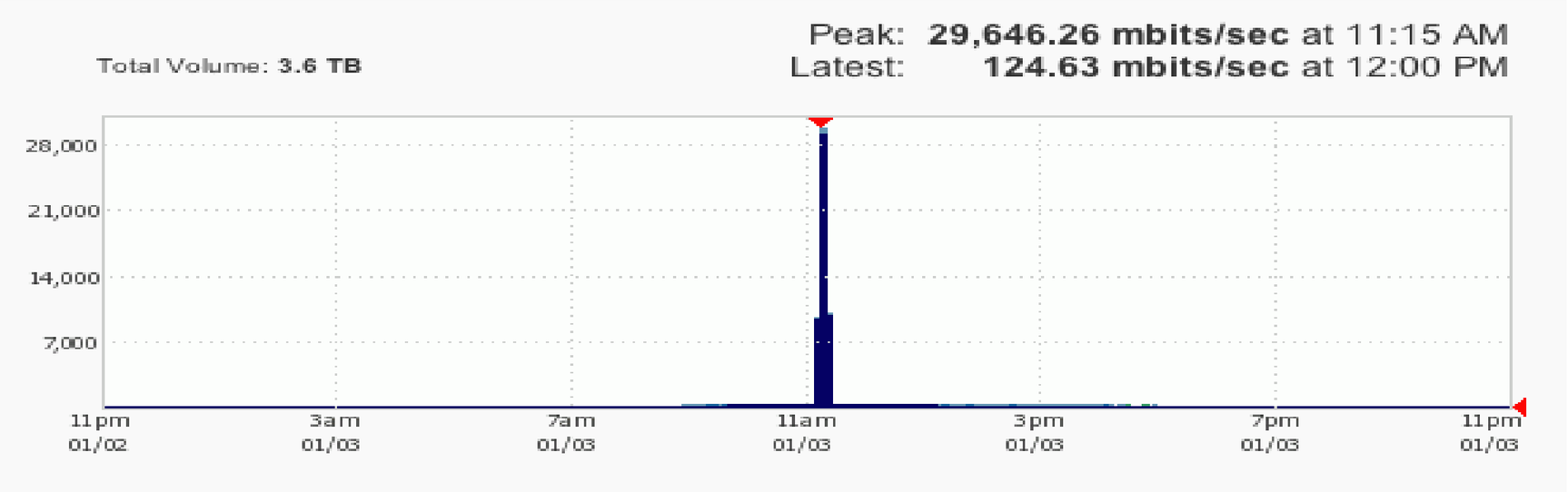
Bank #1

Bank #2

Bank #3

Bank #4

Bank #5



QCF targeted PDF files

Akamai Dynamic Caching Rules offloaded 100% of the traffic

No Origin Impact

	TOTAL VOLUME	% VOLUME
Edge Responses	1.9 TB	97.3 %
Midgress Responses	3.5 GB	0.2 %
Requests	48 GB	2.5 %
Origin Responses	348.9 MB	0 %

Phase 2 Attacks - January 2nd, 2013

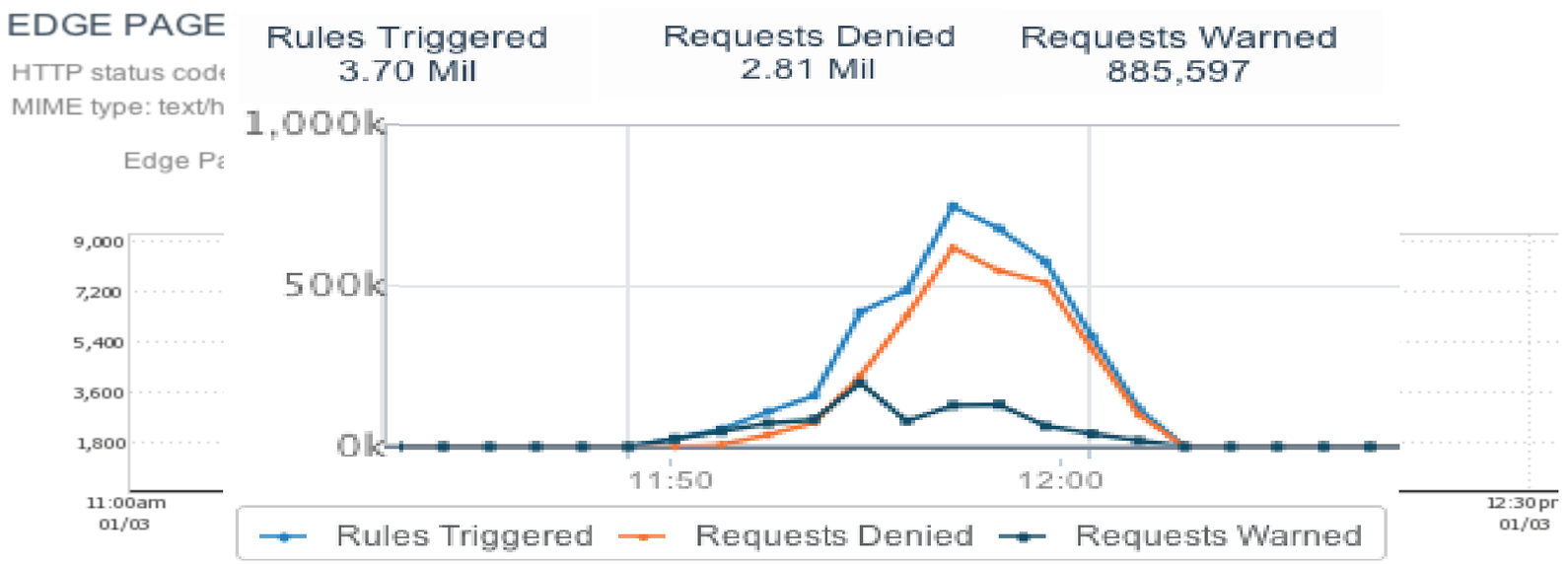
Bank #1

Bank #2

Bank #3

Bank #4

Bank #5



QCF targeted marketing web pages

Rate controls automatically activated
Attack was deflected, far from bank's datacenter

No Origin Impact

Phase 2 Attacks - January 2nd, 2013

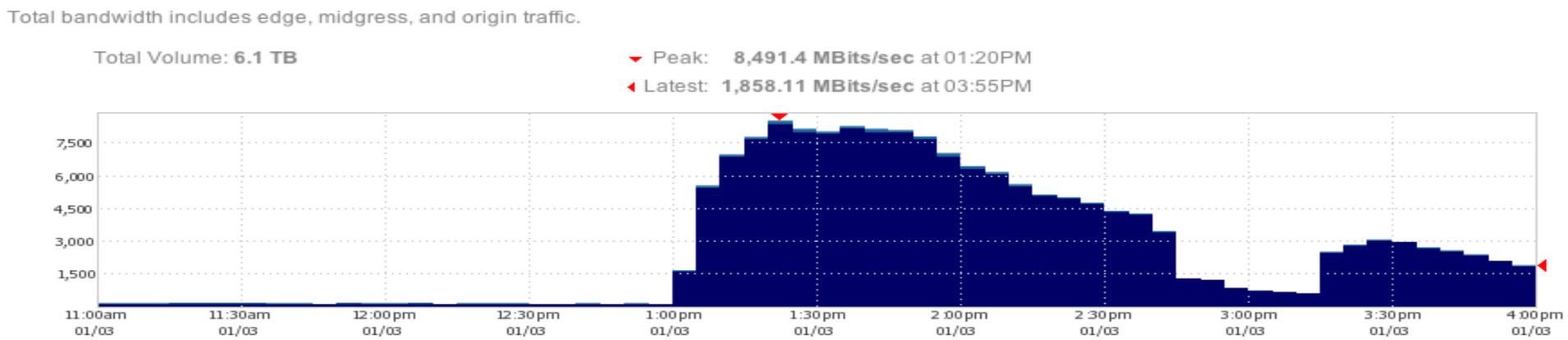
Bank #1

Bank #2

Bank #3

Bank #4

Bank #5



QCF targeted SSL

Akamai offloaded 99% of the traffic

No Origin Impact

	TOTAL VOLUME	% VOLUME
Edge Traffic	6 TB	98.1%
Midgress Traffic	68.5 GB	1.1%
Origin Traffic	46.3 GB	0.8%

Phase 2 Attacks - January 2nd, 2013

Bank #1

Bank #2

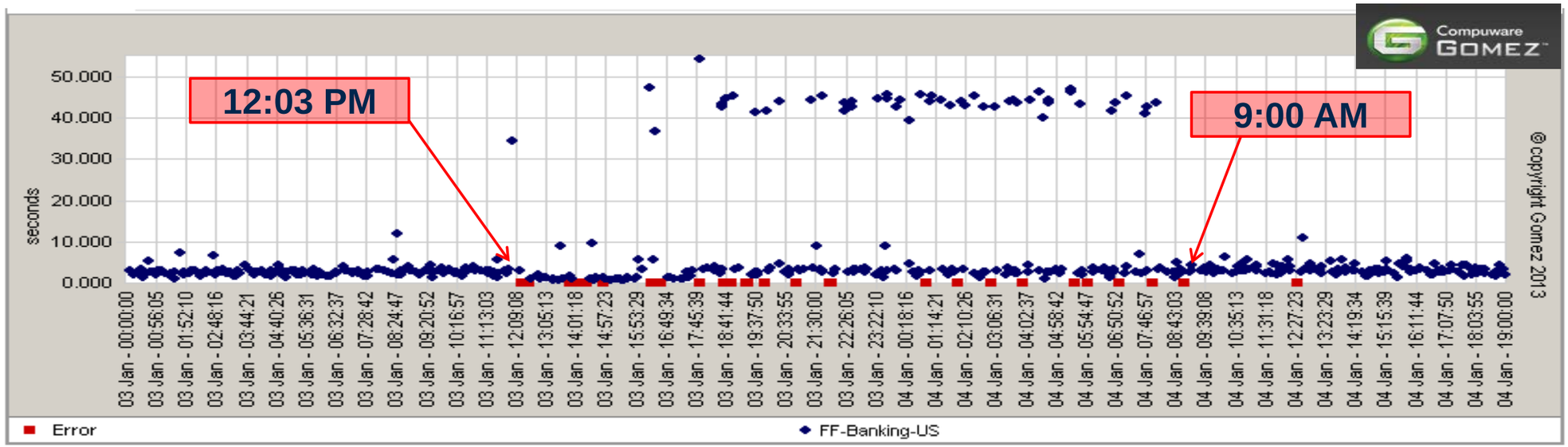
Bank #3

Bank #4

Bank #5

NOT on Akamai

Gomez agents in 12 cities measuring hourly



- Error/Outage—site not responding

Phase 2 Attacks - January 2nd, 2013

Bank #1

Bank #2

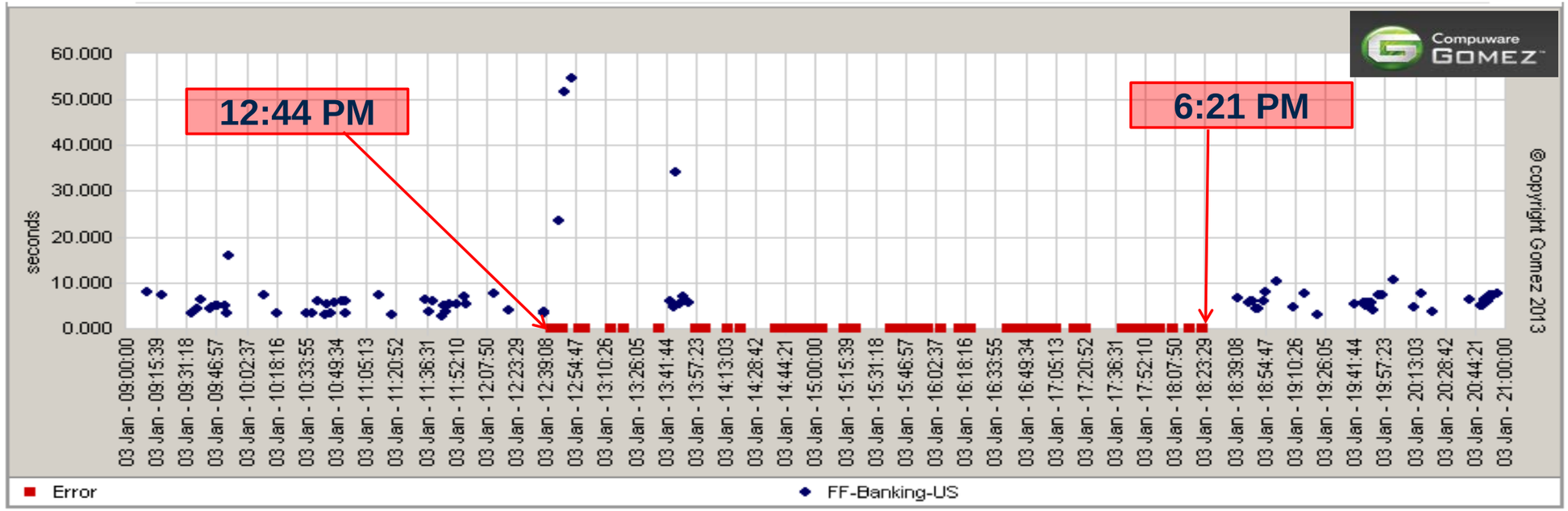
Bank #3

Bank #4

Bank #5

NOT on Akamai

Gomez agents in 12 cities measuring hourly



- Error/Outage—site not responding

Phase 3 Attack Example

- Attack started at March 5, 2013 morning
- Peak Attack Traffic > 126 thousand requests per second
- 70x normal Edge Bandwidth (29Gbps)
  - Origin Traffic stayed at normal levels
- ~2000 Agents participated in the 20 minute assault
  - 80% of the agents were new IP addresses that had not participated in earlier campaigns

Attack Tactics - Pre-attack Reconnaissance

Attackers test the site with short burst high speed probes

- Short bursts of attack requests on non-cacheable content every 10 minutes
- Peak of 18 million requests per second

Example of Why it is Critical to Provide An Always-On Defense

If the site falters, they announce that they will attack that bank and return later with a full scale attack

If the site is resilient they move on

Key Defense Tactics

Query String white-listing

- Ignore query strings intended to “break” caching - don’t include it in the cache key

Rate controls to prevent layer 7 flooding attacks

- Blacklisting known IP attackers was not sufficient

Login page protections

- Combine rate controls with request validation strategies and edge validation offload

Opportunity for Exploiting Social Awareness

Idea: estimate authenticity of client by evaluating client's social network connections

- Requires access to social network data, but not back-end bank data
- Compromised machines may present legitimate social connections
- Start-up ``Socure'' pursuing this approach

Download Manager Product/Service

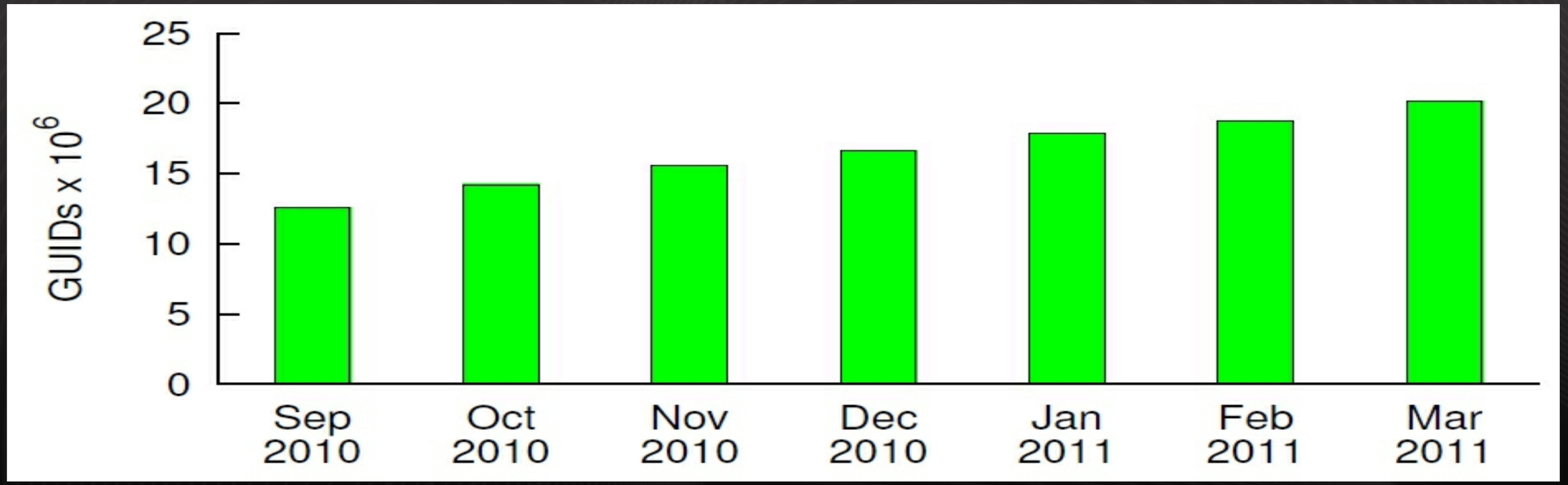
- Hybrid between a fixed-infrastructure CDN and a pure peer-to-peer delivery system
- Based on technology developed by

  Red Swoosh (acquired by Akamai in 2007)
- Goal is to deliver large files at lower cost
- In beta operation today

NetSession Interface

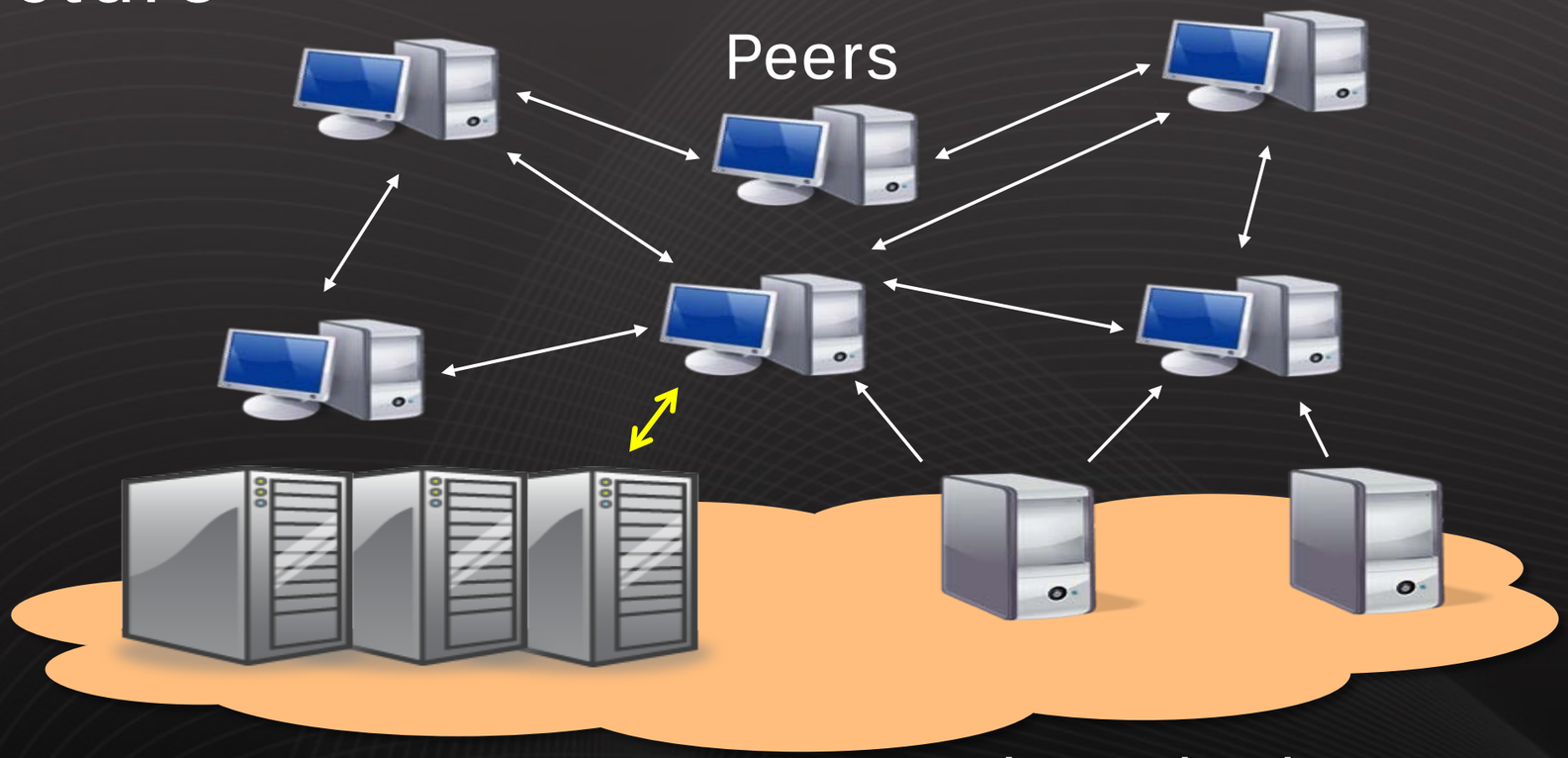
- client-side software
- runs as a service on Windows or MacOS
- provides an API to installation software or browser
- stays connected to an Akamai Control Node
- each installation is supposed to have a unique GUID

Growth in Number of Installations



Full product roll-out expected summer 2011.

Architecture



- BitTorrent-like protocol with control nodes serving as “trackers” and ***assigning*** peers
- CDN acts as a backstop

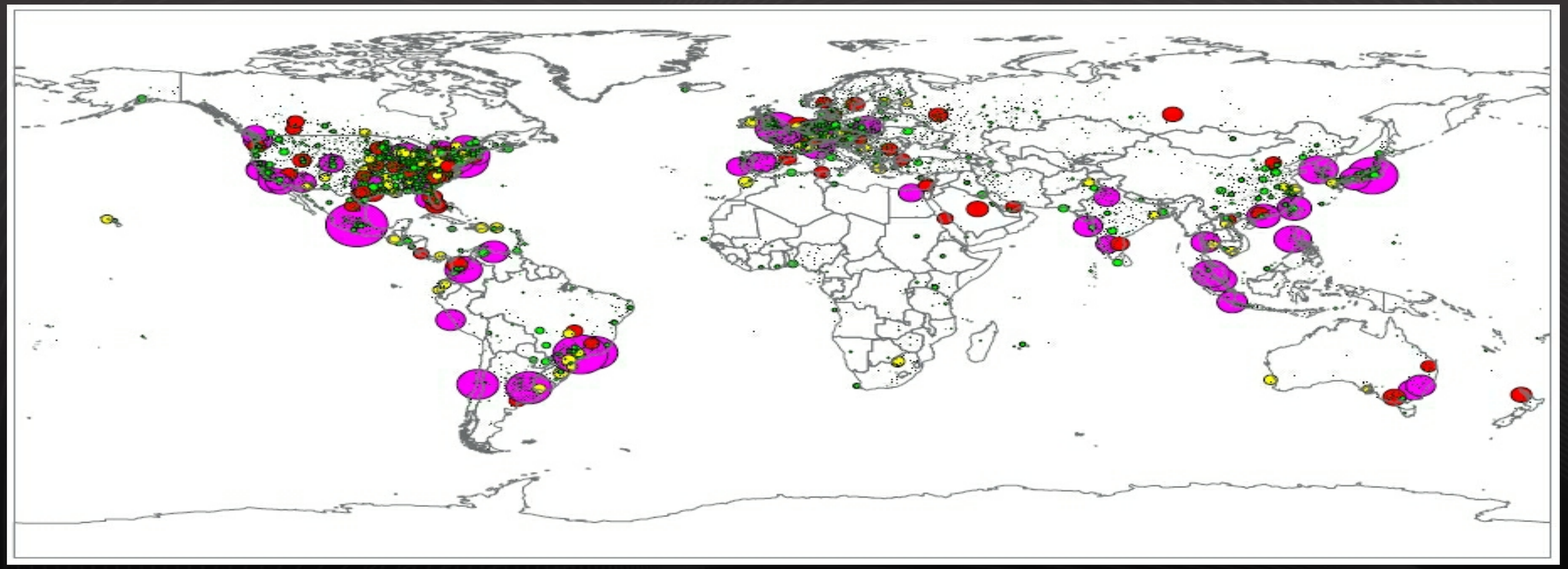
Logging

- Content providers want all downloads to be logged
- Peers ***must*** report P2P downloads to control nodes as they occur
- Akamai Edge Servers also log downloads from peers

Statistics from December 2010 Logs

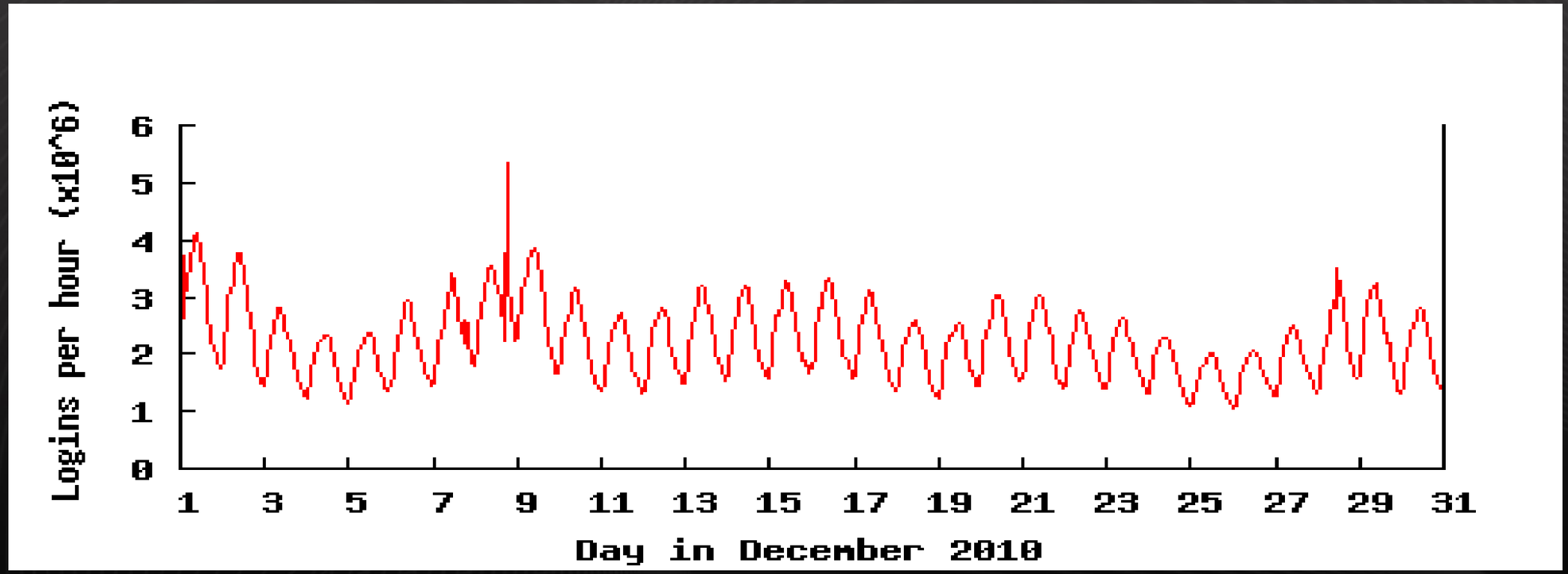
Control plane logs:	
Time period covered	12/1 – 12/31, 2010
Log entries	4,266,622,391
Number of GUIDs	16,608,613
Control plane servers	100
Distinct URLs	299,889
Distinct IPs	115,790,140
Downloads initiated	18,037,706
Geolocation data:	
Distinct IPs	115,790,140
Distinct locations	32,864
Distinct domains	213,899
Distinct countries	234

Locations of Clients per EdgeScape



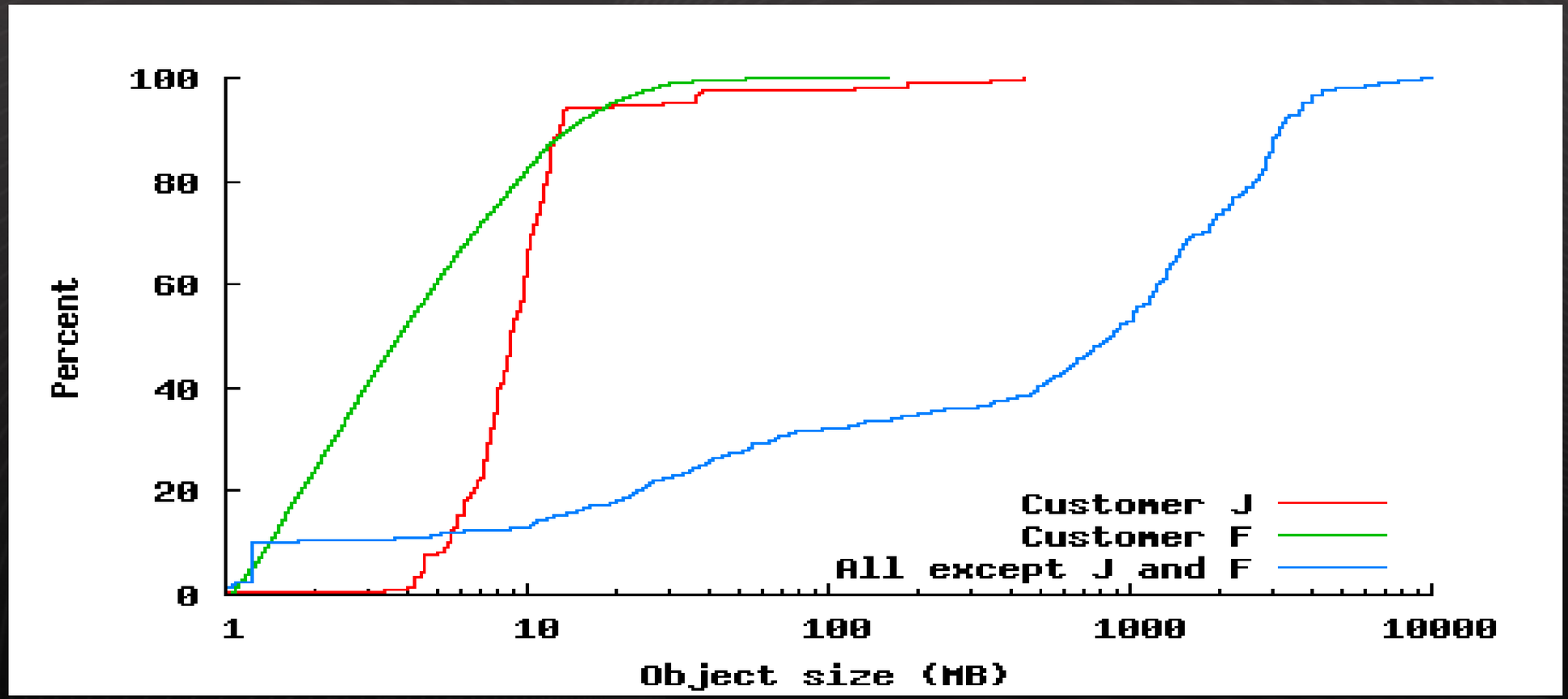
Granularity of predictions varies by continent.

December Client Logins



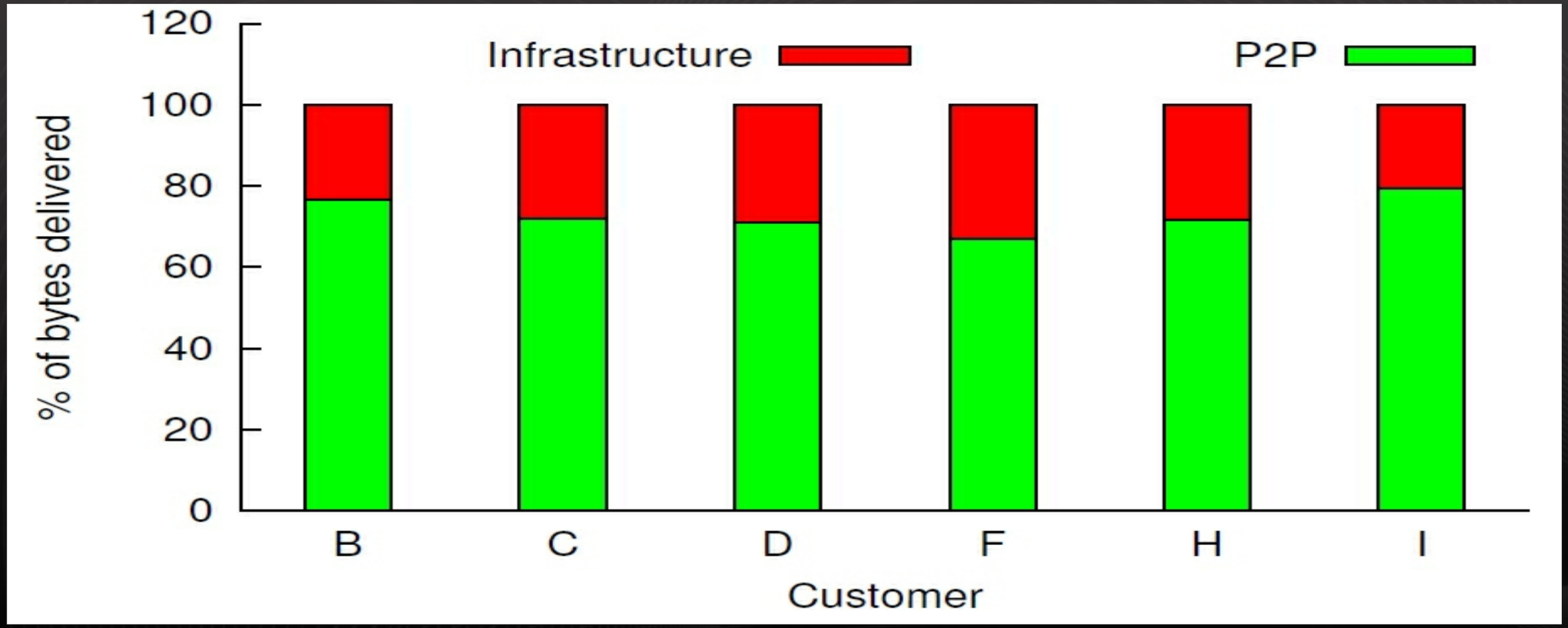
What happened December 8?

File Size Distribution



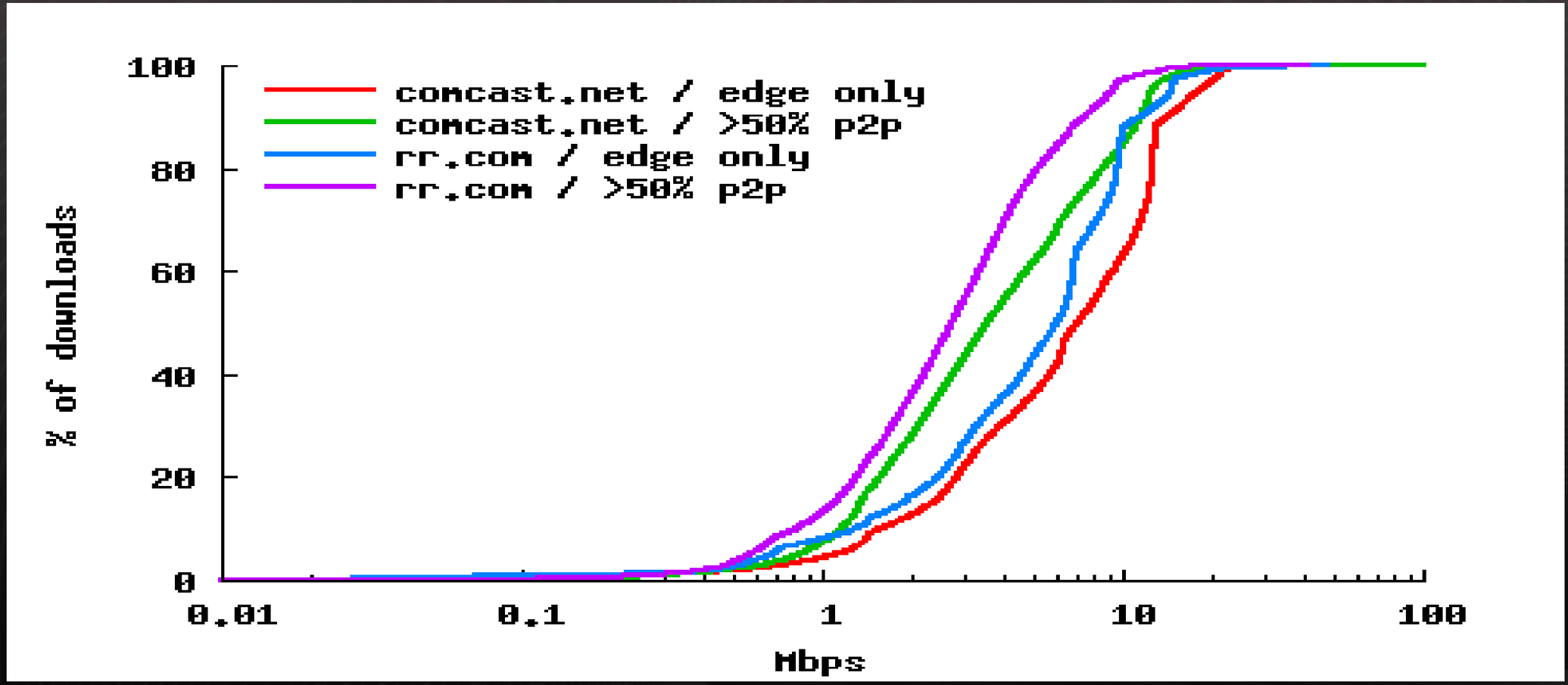
For customers other than J and F, median file size was 872MB.

P2P Efficiency for Largest Enabled Customers



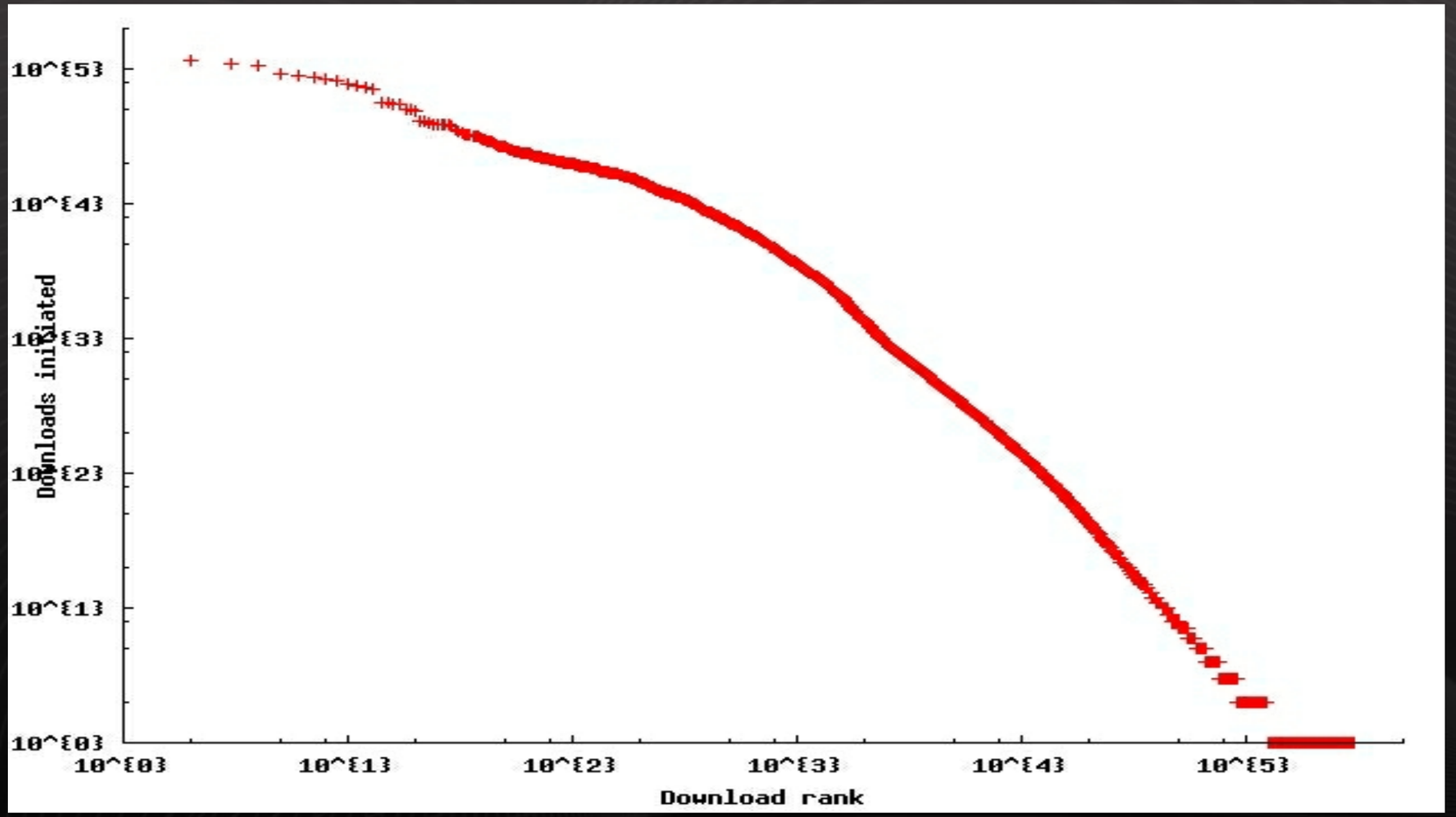
Matches or exceeds earlier predictions.

Performance



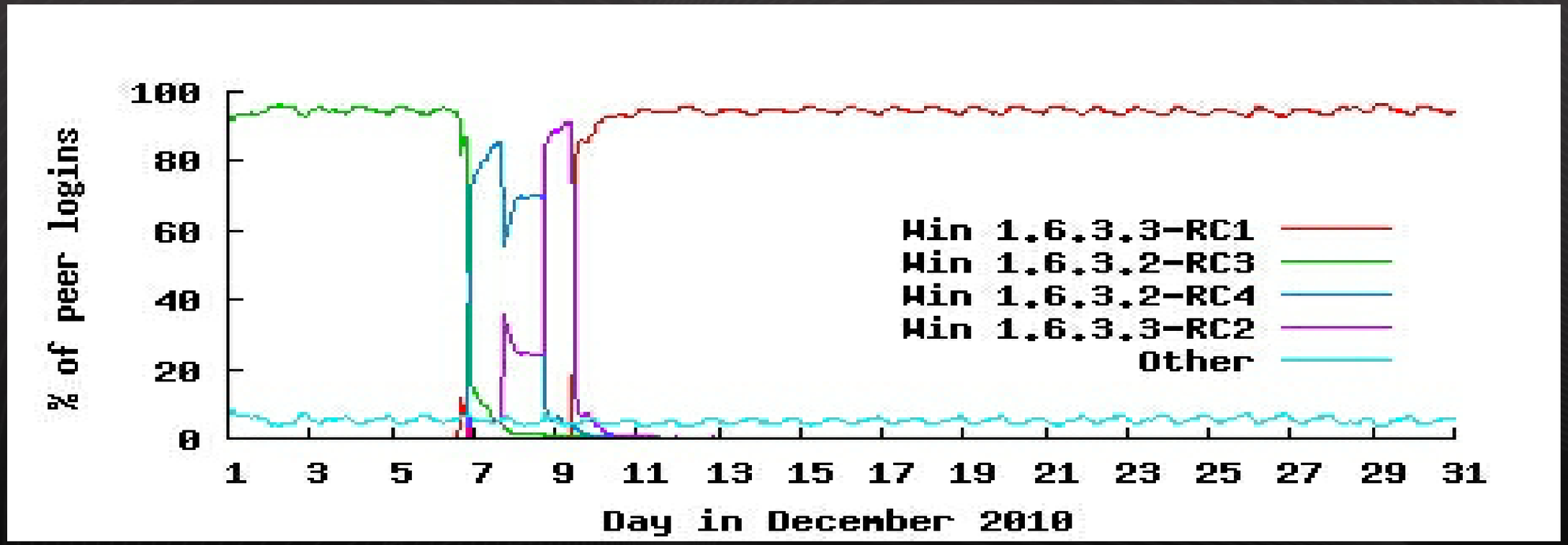
Majority P2P downloads average several Mbps.

Content Popularity



Typical?

NetSession Versions Observed in Field



NetSession should excel at downloading itself.

Opportunity for Exploiting Social Awareness

- Idea: use peer-to-peer system to deliver content for social network
- Motivated by privacy concerns, peer-to-peer social networks have been suggested before
- CDN often suffers cache misses when serving this content
- First search social network contacts for content