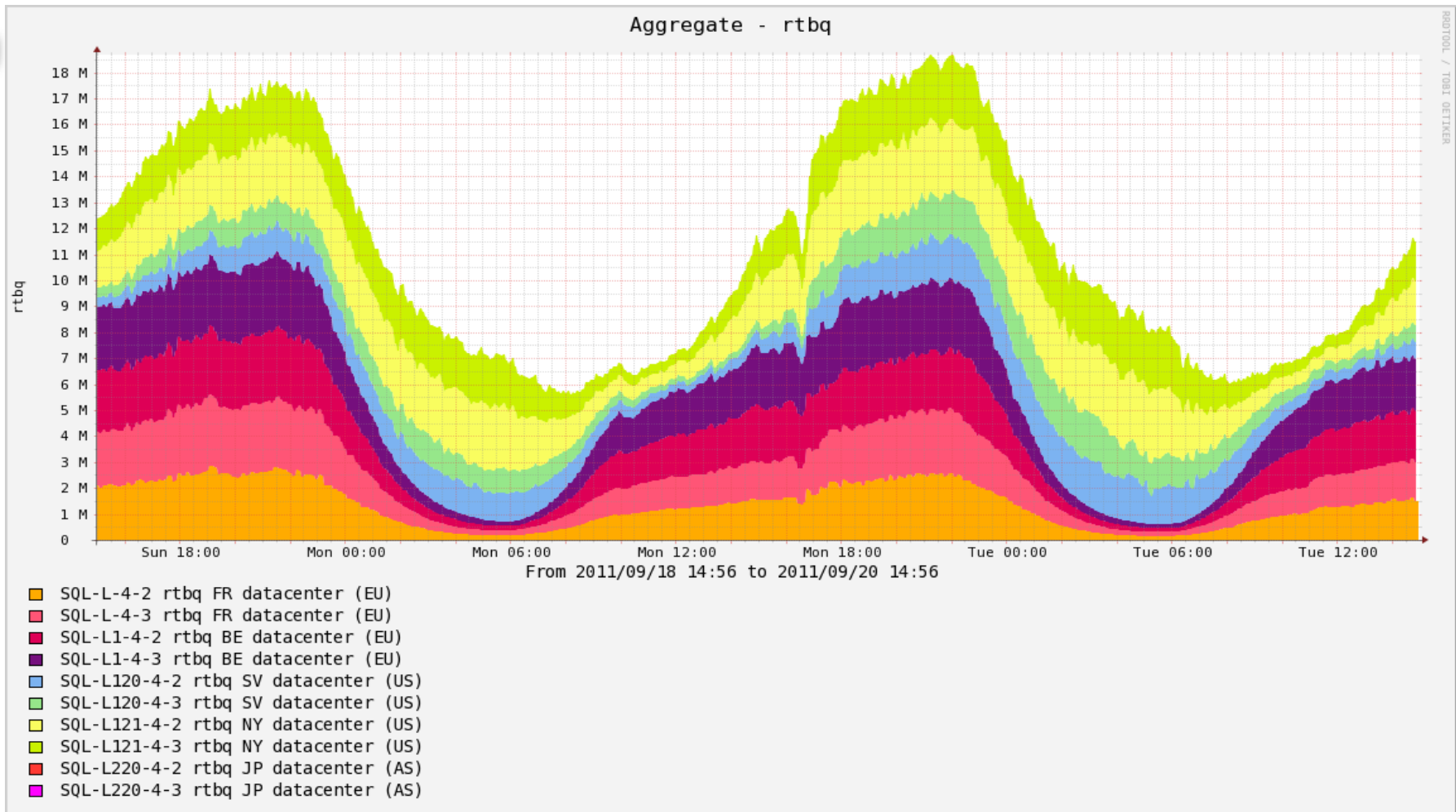# Publisher Team

Mission: find & buy 300+ million profitable ad placements every day

- Cookie Targeting: the work horse
  - 100 billion publisher arbitrations every day
  - 5 billion publisher pixels called every day
- Real-Time Bidding: the race horse
  - Started from scratch in May'10
  - Integrated since then with 8 RTB partners
  - 01/09/2011: 1.95 billion requests, 1.55 billion bids
- Criteo Publisher Marketplace : the mustang
  - (Re)started from scratch in Oct'10
  - Self-service for 2500+ publishers in 6 countries

1.9 million RTB requests per minute

Aggregate - rtbq

Legend:
- SQL-L-4-2 rtbq FR datacenter (EU)
- SQL-L-4-3 rtbq FR datacenter (EU)
- SQL-L1-4-2 rtbq BE datacenter (EU)
- SQL-L1-4-3 rtbq BE datacenter (EU)
- SQL-L120-4-2 rtbq SV datacenter (US)
- SQL-L120-4-3 rtbq SV datacenter (US)
- SQL-L121-4-2 rtbq NY datacenter (US)
- SQL-L121-4-3 rtbq NY datacenter (US)
- SQL-L220-4-2 rtbq JP datacenter (AS)
- SQL-L220-4-3 rtbq JP datacenter (AS)

From 2011/09/18 14:56 to 2011/09/20 14:56

criteo
Scalable Personalized Retargeting

Publisher Team: what's next

- Cookie Targeting
  - Quite mature, no huge new features on the radar
  - Major upgrades under the hood (Scalability team, more in a minute)

- Real-Time Audience: the next step in Cookie Targeting
  - Better arbitrage accuracy, pixel-free solution
  - Free retargeting alternative for publishers (Doubleclick)
  - Live on 16 publishers (inc. SevenOne, #3 in DE)

Publisher Team: what's next

- Criteo Publisher Marketplace
  - A zillion new features (no less)
  - Productivity features for Criteo teams

- Real-Time Bidding
  - More RTB partners (Yahoo coming)
  - Ramp-up in Europe, launch in Japan
  - Inventory quality : ad verification, blacklisting
  - Productivity tools for Criteo teams : publisher clustering, better stats, smarter logs, new Dashboard, etc.

Publisher Team: what's next

- Spin-off a new team in charge of hunting data
  - Use data already available in RTB requests (Google keywords, geotargeting, …)
  - Import socio-demographics (Quantcast)
  - Crawl publisher pages to extract valuable data (page content) and metadata (# of banners, above/below the fold, etc.)
  - Extract valuable keywords from URLs
  - Extract valuable keywords from Advertiser catalogs
- All this could be used to optimize buying
- Several experiments already in progress

criteo
Scalable Personalized Retargeting

Scalability Team

Mission: if it ain't broken, fix it before it breaks. If it IS broken, fix it already!

Main lines of work:

- Develop critical software modules used by all R&D teams
- Bleeding-edge technologies, from PoC to production
- Provide platform-wide coaching & expertise on performance topics
- Code refactoring
  - 50% drop in memory consumption on Cookie targeting servers (left)
  - 80% drop in memory consumption on Delivery servers (right)
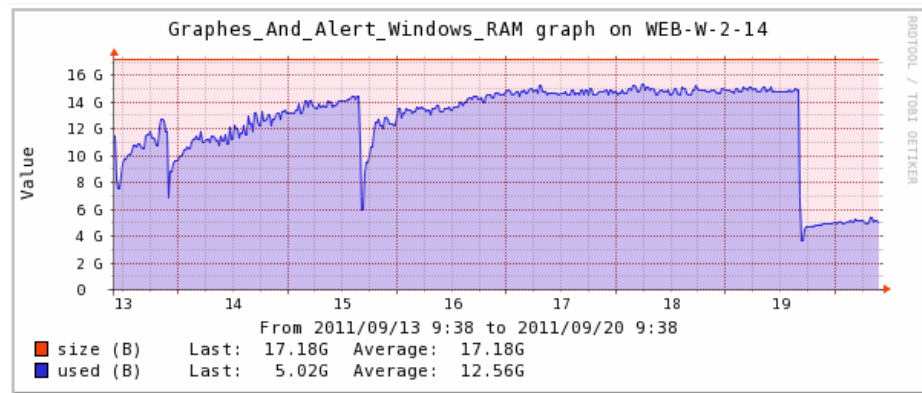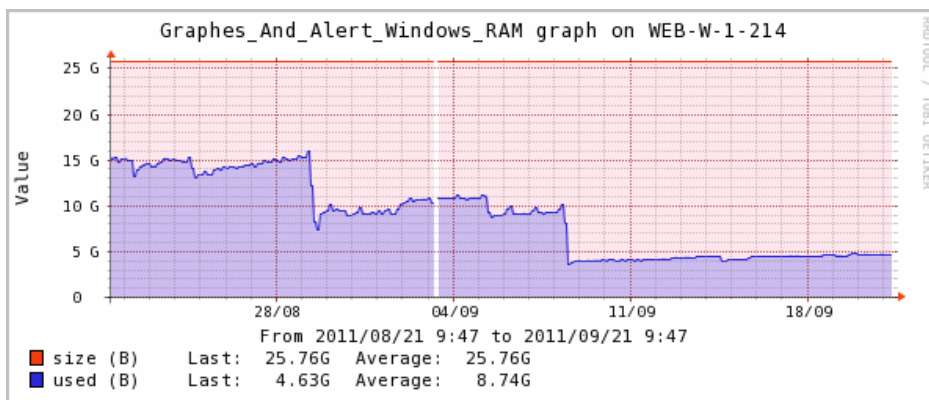
Scalability Team

- Catalog Import v2
  - Able to ingest & replicate monster US catalogs coast-to-coast
  - Completely redesigned
- Real-time publisher arbitration
  - Replaces pre-computed files loaded on web servers
  - Less latency in tagging/detagging
- Image delivery to banners: 690 million daily
- Log aggregation
  - over 1TB crunched daily
  - Hadoop cluster (700 cores, 700 TB)

![criteo logo] criteo
Scalable Personalized Retargeting

Production team

- Mission: 100% uptime, 24/7/365

- Design, procurement & setup:
  - Datacenters (5)
  - Servers + HPC (830)
  - Networking equipment & links
  - Databases (close to 4000)
- Monitoring and on-duty calls active on all of the above, as well as applications

criteo
Scalable Personalized Retargeting

Production team: hosting

New site outside Paris (40 racks)
This one's our blueprint

criteo
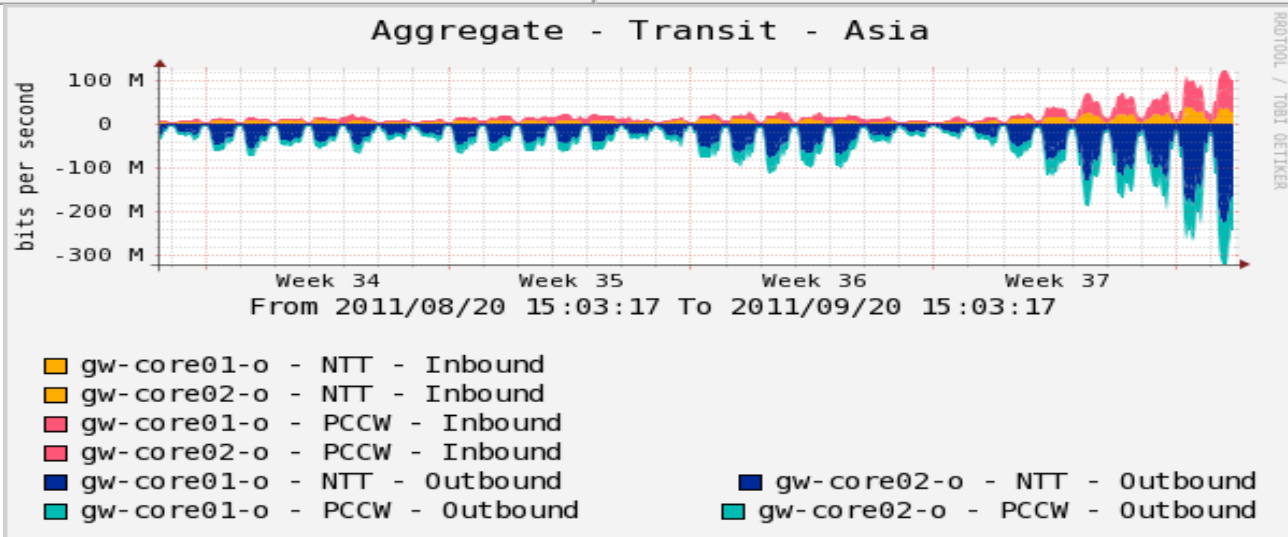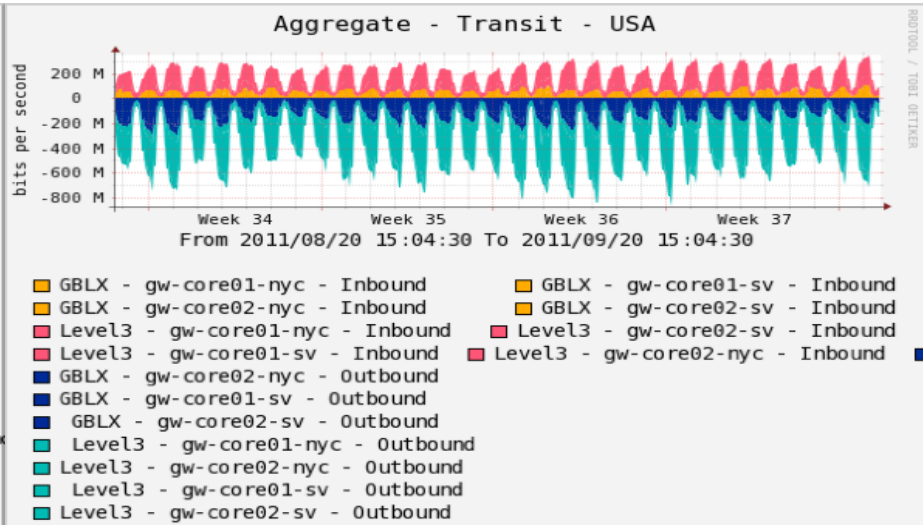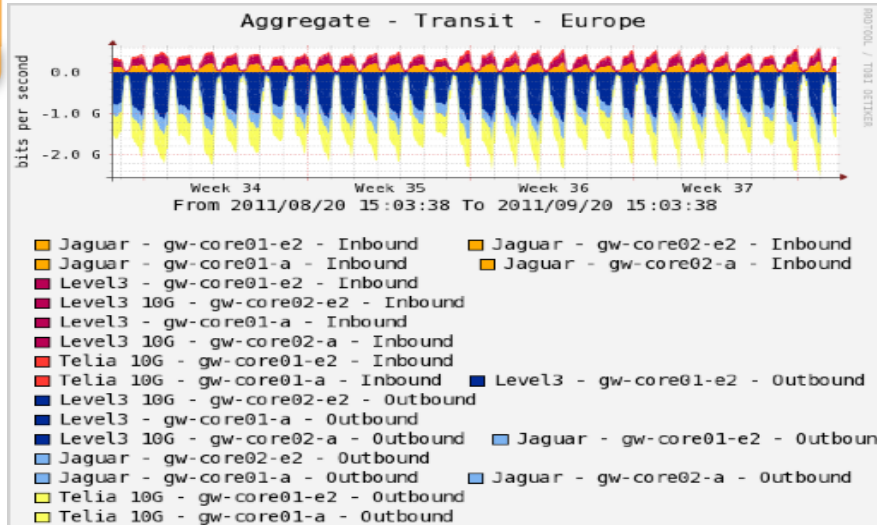Scalable Personalized Retargeting

Production team: network

- Up to **4Gb/s** of traffic in Europe
- Up to **1Gb/s** of traffic in the USA
- Up to **300Mb/s** of traffic in Asia
- 200 peering sessions
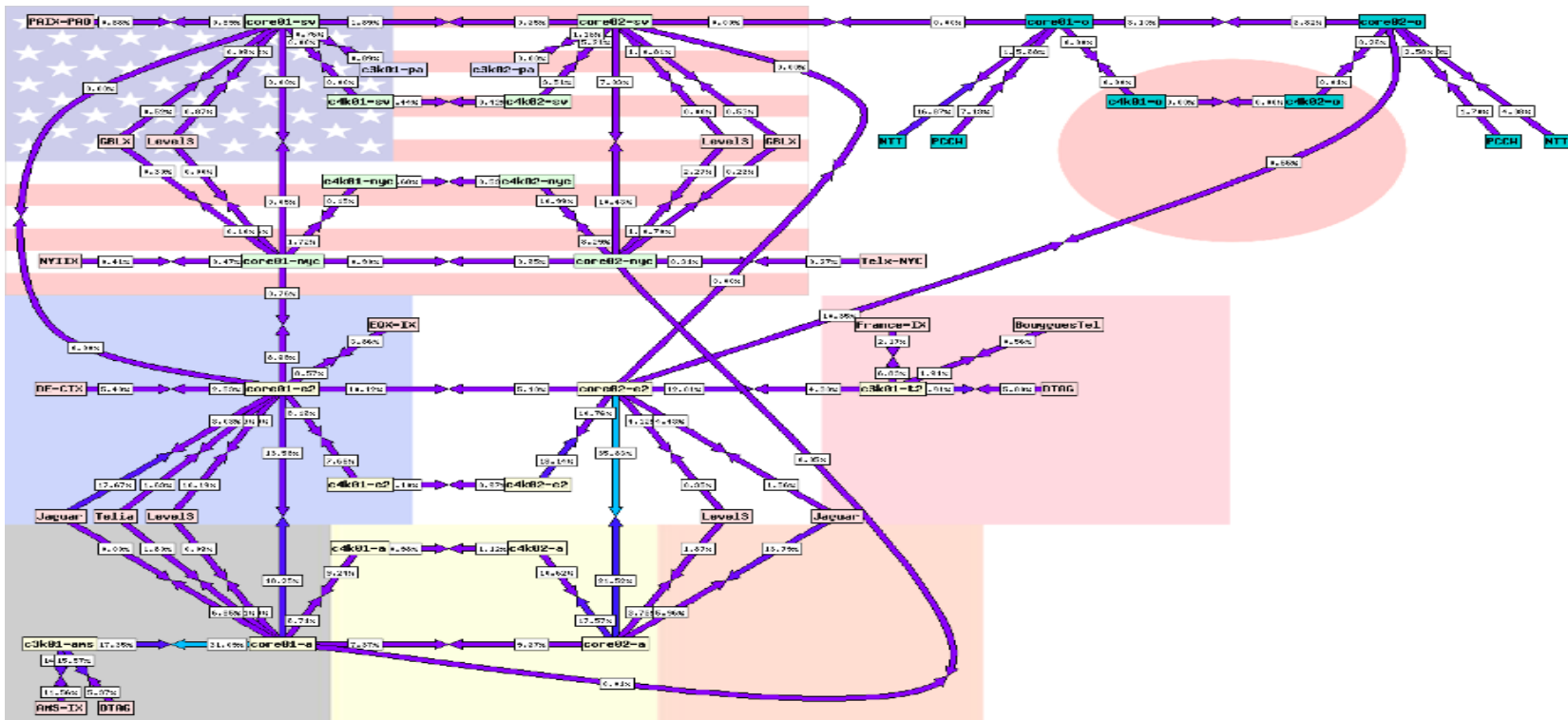
Production team: traffic

Peering

- Some references among hundreds of peers!

Production Team: what's next

- 2011
  - Finish migration & cleanup of Paris DC
  - Scale Tokyo DC (already?)
  - Double our hosting capacity in the US (NY & SV)
- 2012
  - Migrate Brussels DC to Amsterdam
  - Deploy HPC cluster in the US
  - 2nd datacenter in Asia
  - 3rd datacenter in the US (RTB-specific maybe?)
  - Eastern Europe, South America, … we'll see ☺