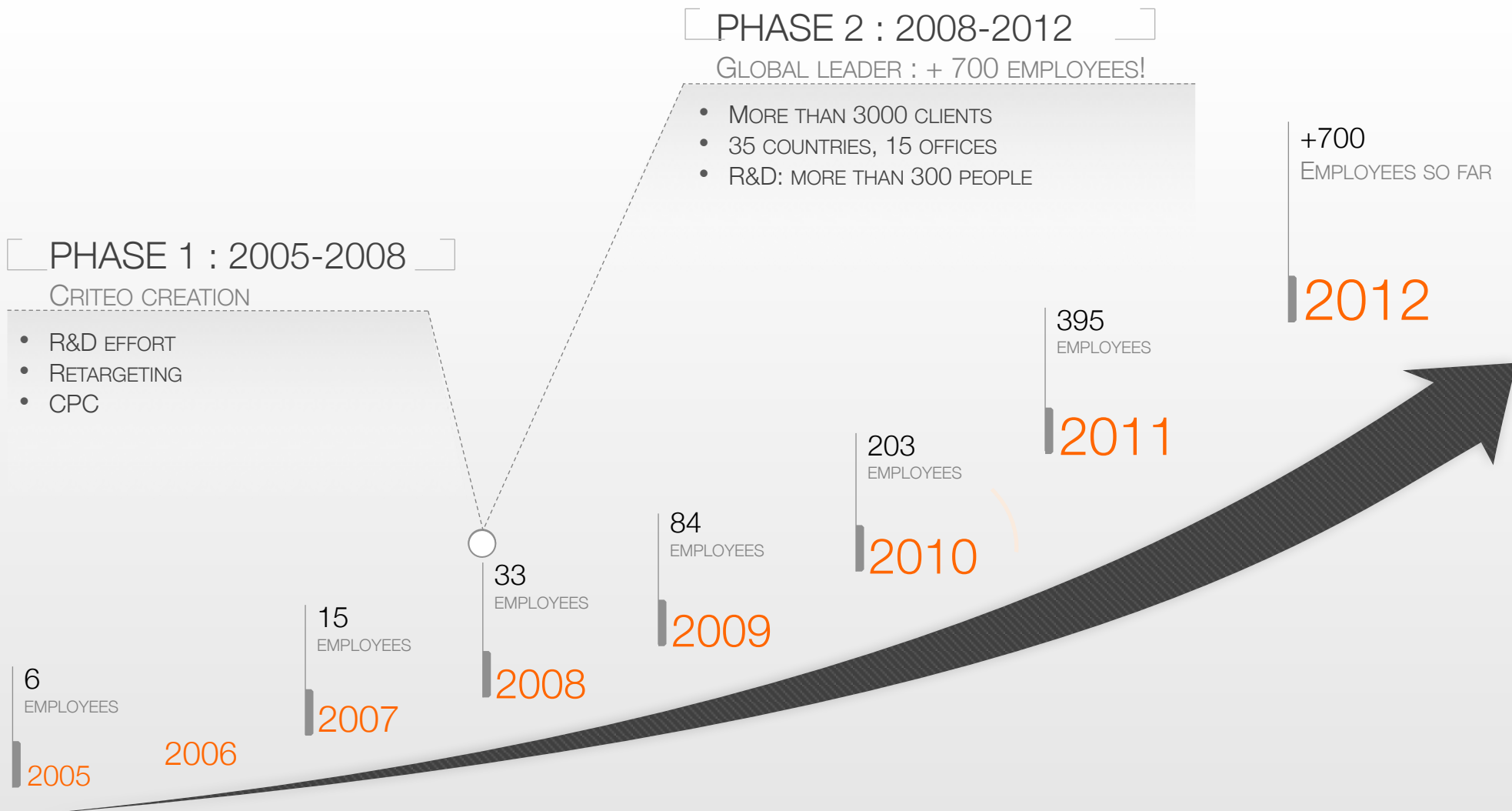


How Criteo Scaled and Supported Massive Growth with MongoDB

MongoDB Conference
New York City, June 2013



Julien SIMON
Vice President, Engineering
j.simon@criteo.com @julsimon

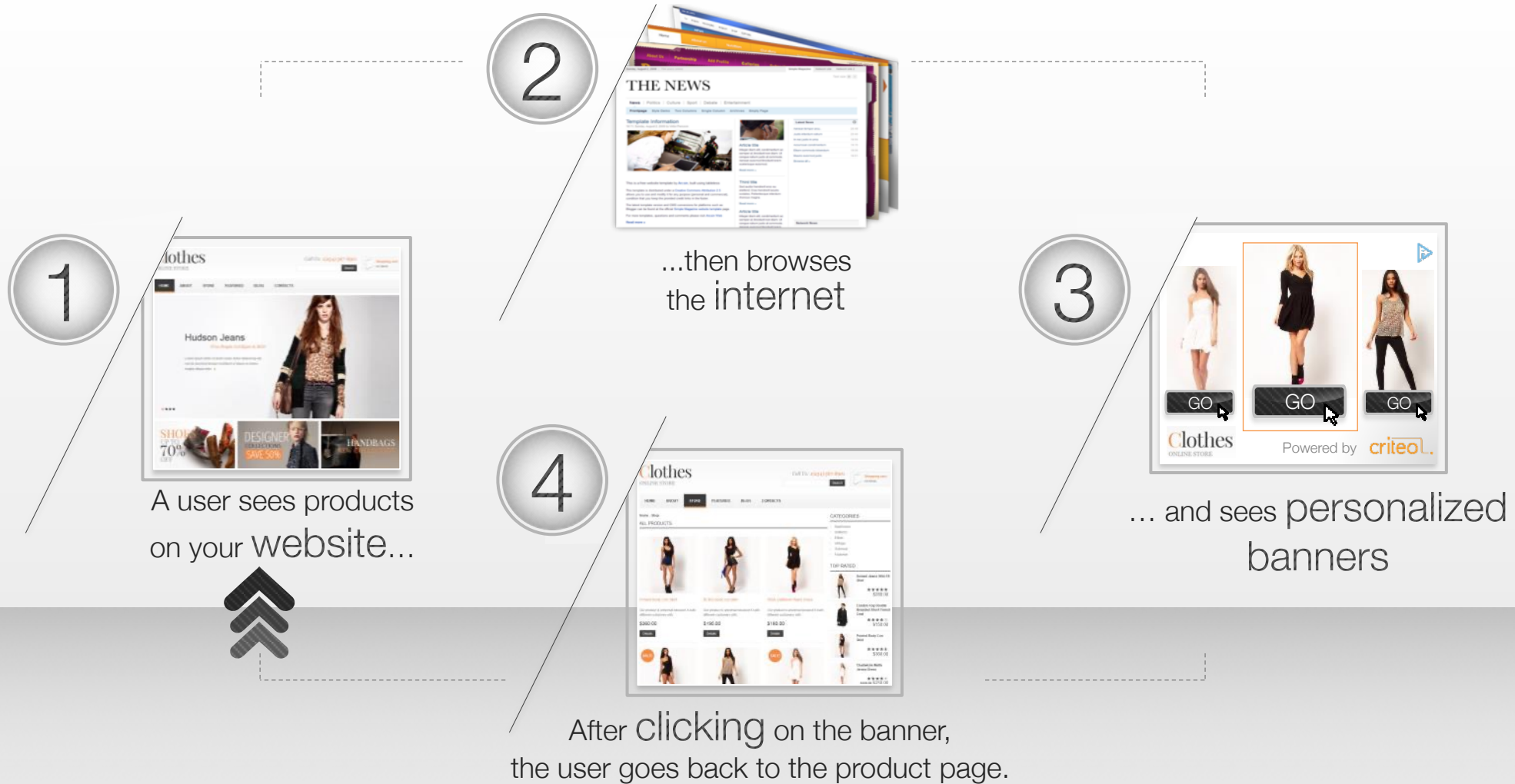


GLOBAL PRESENCE

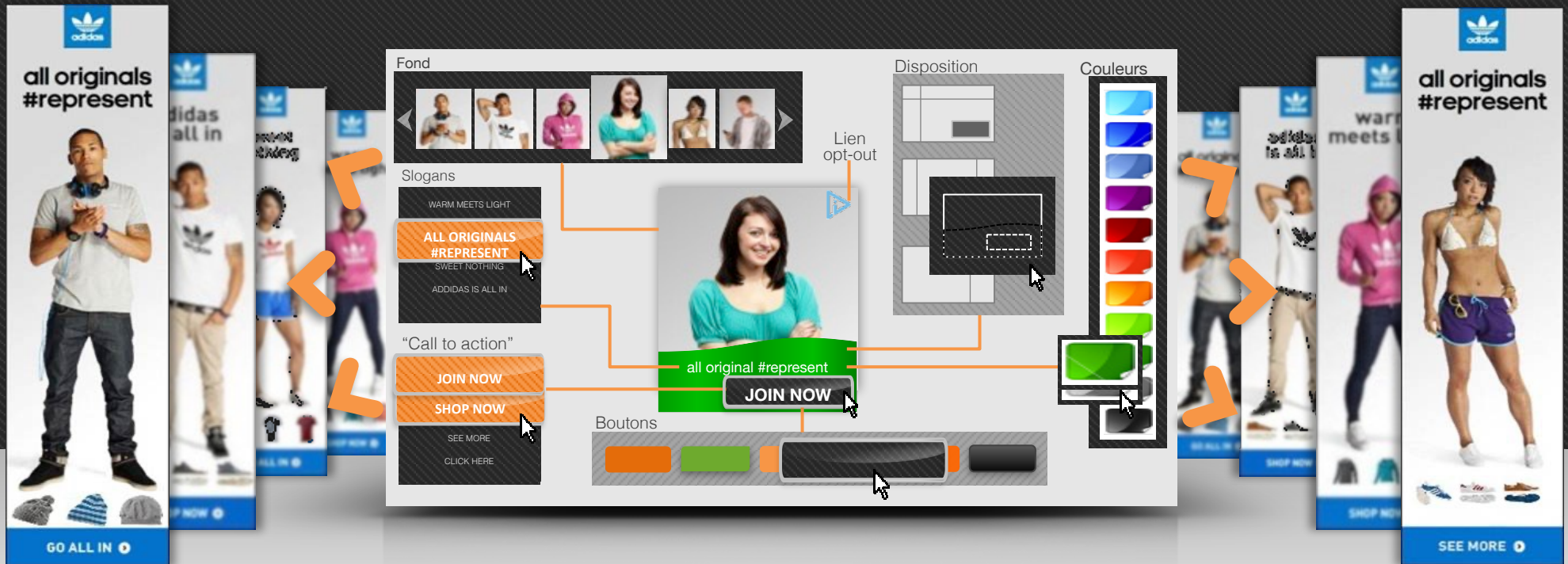
15 OFFICES, 30+ COUNTRIES



PERFORMANCE DISPLAY

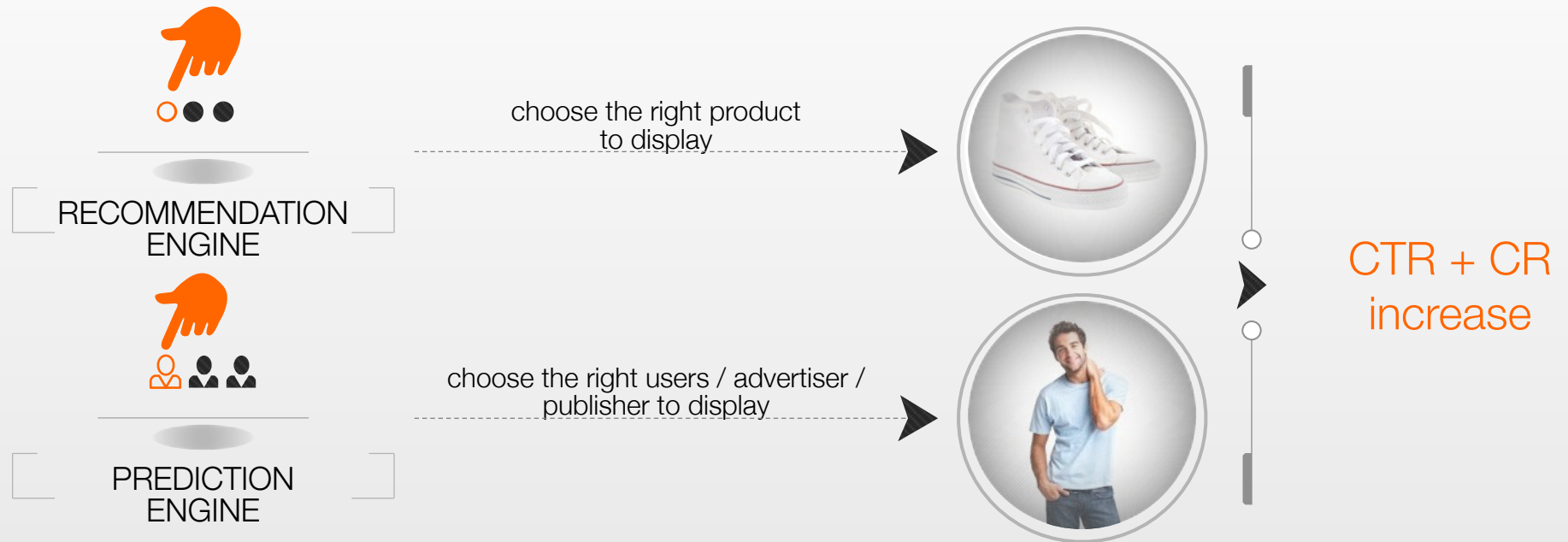


REAL-TIME PERSONALIZATION



PREDICTION & RECOMMENDATION

2 CORE TECHNOLOGIES



INFRASTRUCTURE



»» 7 DATA CENTERS

»» SET UP AND MANAGED
IN-HOUSE

»» AVAILABILITY > 99.95%

»» DAILY TRAFFIC

- HTTP REQUESTS: 30+ BILLION
- BANNERS SERVED: 1+ BILLION

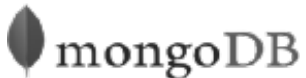
»» PEAK TRAFFIC (PER SECOND)

- HTTP REQUESTS: 500,000+
- BANNERS: 25,000+

HIGH PERFORMANCE COMPUTING

FETCH, STORE, CRUNCH, QUERY **20 additional TB** EVERY DAY ?

...SUBTITLED « HOW I LEARNED TO STOP WORRYING AND LOVE HPC »



Storm Kafka

PRODUCT CATALOGUES

- Catalogue = product feed provided by advertisers (product id, description, category, price, URL, etc)
- 3000+ catalogues, ranging from a few MB to several tens of GB
- About 50% of products change every day
- Imported at least once a day by an in-house application
- Data replicated within a geographical zone
- Accessed through a cache layer by web servers
- Microsoft SQL Server used from day 1
- Running fine in Europe, but...
 - Number of databases (1 per advertiser)... and servers
 - Size of databases
 - SQL Server issues hard to debug and understand
- Running kind of fine in the US, until dead end in Q1 2011
 - transactional replication over high latency links



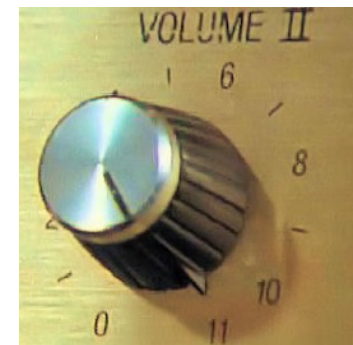
REQUIREMENTS FOR A NEW DB

- Scale-out architecture running on commodity hardware (aka « Intel CPUs in metal boxes »)
- No transactions needed, eventual consistency OK
- High availability
- Distributed clusters, with replication over high latency links
- Requestable (key-value not enough)
- Open source
 - ... with active user community
 - ... backed by a stable organization with long-term commitment (not one guy in a garage)
 - ... no licence fees for production use
 - ... commercial support available at reasonable cost
- Easy to learn, (re)deploy, monitor and upgrade
- « Low maintenance » (don't need a 10-people team just to run it)
- Multi-language support



FROM SQL SERVER TO MONGODB

- Ah, database migrations... everyone loves them ☺
- 1st step: solve replication issue
 - Import and replicate catalogues in MongoDB
 - Push content to SQL Server, still queried by web servers
- 2nd step: prove that MongoDB can survive our web traffic
 - Modify web applications to query MongoDB
 - C-a-r-e-f-u-l-l-y switch web queries to MongoDB for a small set of catalogues
 - Observe, measure, A/B test... and generally make sure that the system still works
- 3rd step: scale !
 - Migrate thousands of catalogues away from SQL Server
 - Monitor and tweak the MongoDB clusters
 - Add more MongoDB servers... and more shards
 - Update ops processes (monitoring, backups, etc)



OUR MONGODB DEPLOYMENT

- Europe
 - 18 3-server shards (1+1+1)
 - 800M products, 1TB
 - 1B requests/day (peak at 40K/s)
 - 350M updates/day (peak at 11K/s)
- US
 - 14 4-server shards (2+2)
 - 400M products, 650GB
- APAC
 - 12 3-server shards (2+1)
 - 300M products, 500GB
- 146 servers total :
2.0 (+ Criteo patches) → 2.2 → 2.4.3



MONGODB, 2+ YEARS LATER

- Stable (2.4.3 much better)
- Easy to (re)install and administer
- Great for small *datasets* (i.e. smaller than server RAM)
- Good performance if read/write ratio is high
- *Failover* and inter-DC replication work (but shard early!)
- Performance suffers when :
 - *dataset* much larger than RAM
 - read/write ratio is low
 - Multiple applications coexist on the same cluster
- Some scalability issues remain (master-slave, connections)
- Criteo is very interested in the 10gen *roadmap* ☺



THANKS A LOT FOR YOUR ATTENTION!



www.criteo.com
engineering.criteo.com



