



Keeping Track of 70,000+ Servers

The Akamai Query System

Jeff Cohen, Thomas Repantis, Sean McDermott, Scott Smith, Joel Wein

The Akamai Platform

- Various web infrastructure services
- Over 70,000 machines
- Over 1 million distributed components
- Over 1000 autonomous systems
- 24/7/365 operation
- Failures, usage changes
- Massive, real-time monitoring

Query

- Distributed data collection
- Aggregation at several hundred points
- SQL-style interface

A Sample Query



```
SELECT
  c.continent_name,
  SUM(l.hits) hits
FROM
  load_info l,
  region_data r,
  continent_data c
WHERE
  l.georegion=r.id AND
  r.continent=c.continent
GROUP BY
  c.continent_name
ORDER BY
  hits DESC;
```

c.continent_name	hits
-----	-----
North America	4,620,551
Europe	3,392,102
South America	655,175
Asia	552,258
Africa	106,781
Oceania	39,905
Antarctica	135

Outline



- Design goals
- Architecture
- Uses
- Scale

Design Goals

- Reliability
- Completeness
- Scalability
- Latency
 - Data
 - Query
- Consistency
- Fault tolerance
- Synchronization

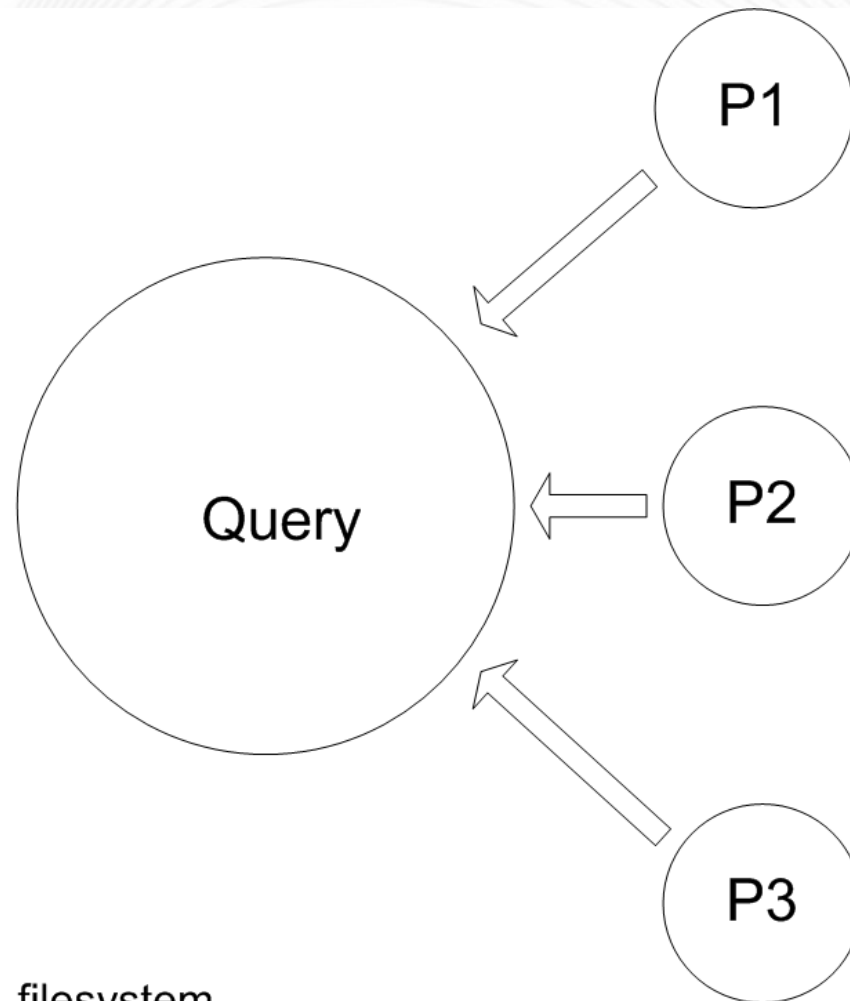
Architecture



- Collection at edge machines
- Aggregation at cluster level
- Aggregation at global level
- Answering queries

Query at the Edge

- Each machine collects its own data
- Many processes may publish
- Snapshots every two minutes



filesystem

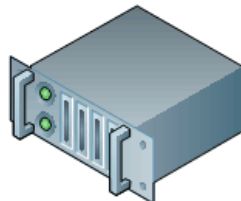
machineip	mountp	blocks	bavail	bsize
10.123.123.1	/var	1500000	36665	4096

Cluster Proxies

- Collect data for the whole cluster
- Include themselves

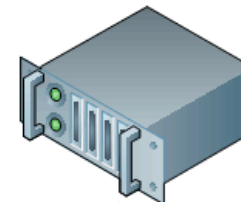
Edge machine

filesystem				
machineip	mountp	blocks	bavail	bsize
10.123.123.9	/var	1500000	51575	4096

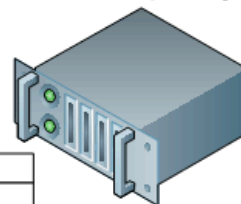


Edge machine

filesystem				
machineip	mountp	blocks	bavail	bsize
10.123.123.7	/var	1500000	17631	4096



Cluster proxy



filesystem				
machineip	mountp	blocks	bavail	bsize
10.123.123.1	/var	1500000	36665	4096

filesystem				
machineip	mountp	blocks	bavail	bsize
10.123.123.1	/var	1500000	36665	4096
10.123.123.7	/var	1500000	17631	4096
10.123.123.9	/var	1500000	51575	4096

Aggregated table

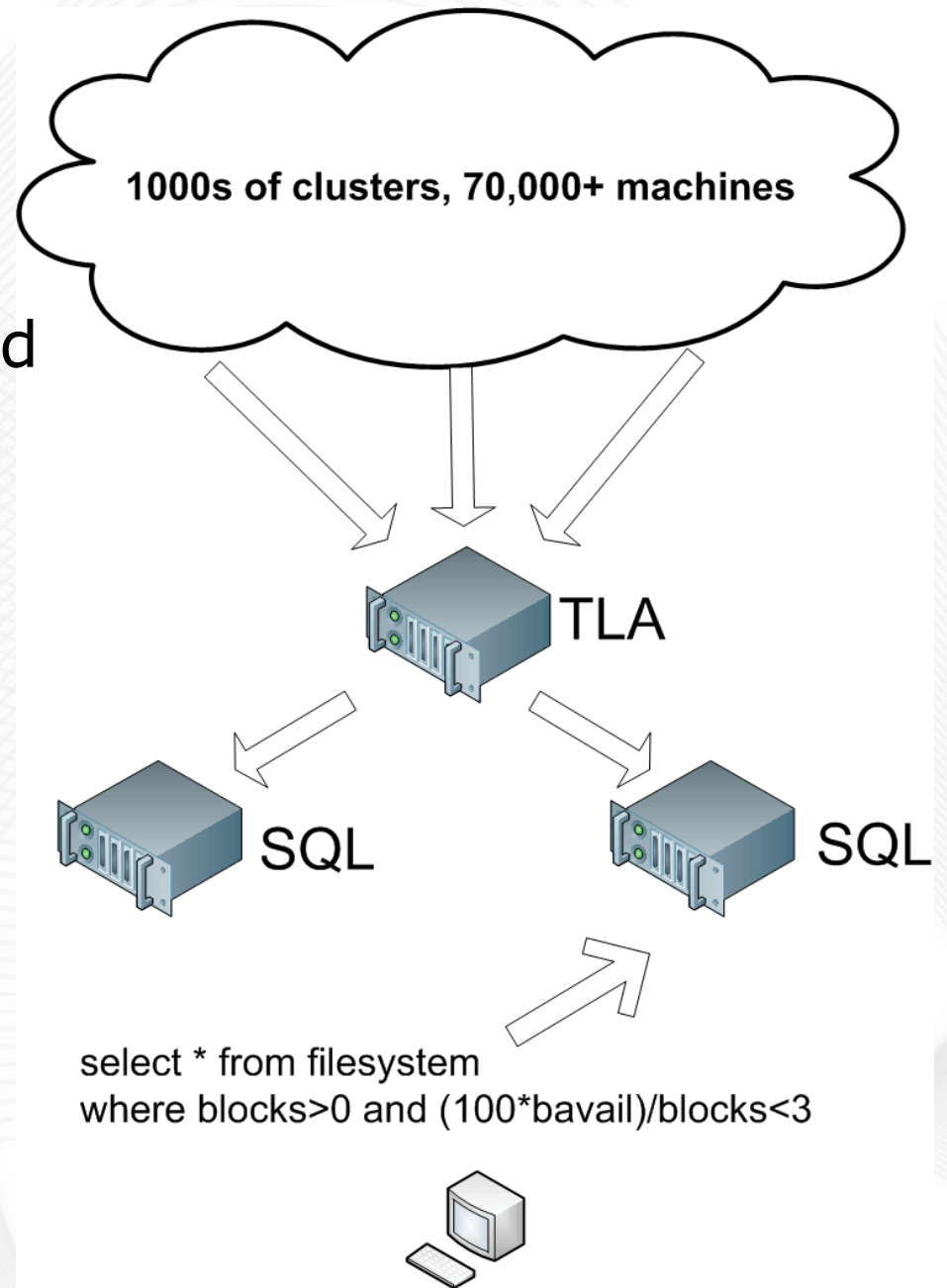
Top-Level Aggregators



- Collect data for the whole network
- Snapshots every two minutes
- Static tables for data that doesn't change much

SQL parsers

- Get tables from 1 TLA
- Only get the ones we need
- Answer queries based on them



- Span different parts of the network
- Designated for different purposes
- Several replicated TLAs & SQLs
- Combined TLA/SQLs
- Shared hostnames
- Help meet reliability guarantees
- Help tolerate faults & keep localized

Uses



- Alerts
- Graphical monitoring
- Incident response
- Capacity management
- Many others

Alerts

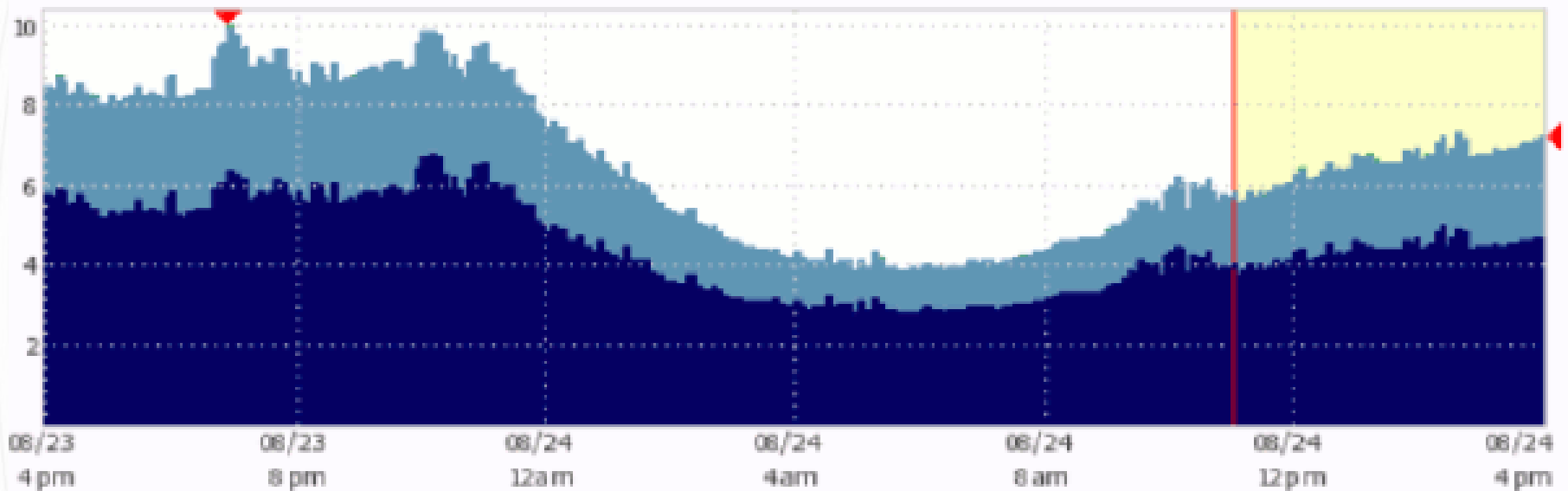


- Several thousand alerts
- Important way of detecting problems
- Goal is to find & fix before customer impact
- SQL and procedure for each alert
- Highly customizable
 - Priorities
 - Times to start, clear
 - Notification methods
- Akamai alerts, customer alerts

Graphical monitoring



- Several automated users graph Query data
 - Akamai web site
 - Portal graphs
 - Historical graphs



Eastern Daylight Time

Estimated data begins at 11:00AM

Scale



- Several hundred TLAs, SQLs, TLA/SQLs
- Thousands of queries per minute
- Tens of GB in the system
- Up to 6 GB per TLA (and growing fast)
 - Internet usage
 - Network growth
 - Customer growth
 - Data/customer
 - More queries
- Age of data typically a few minutes



Thank you!



Jeff Cohen
Akamai Technologies
jecohen@akamai.com