



中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences



DQ2 Tracer Service

Donal Zang (IHEP)

ADC lab 

Outline

- Tracer service overview
- Improvements to the tracer service
 - Message broker: ActiveMQ
- Tracer monitoring

What's DQ2 Tracer Service

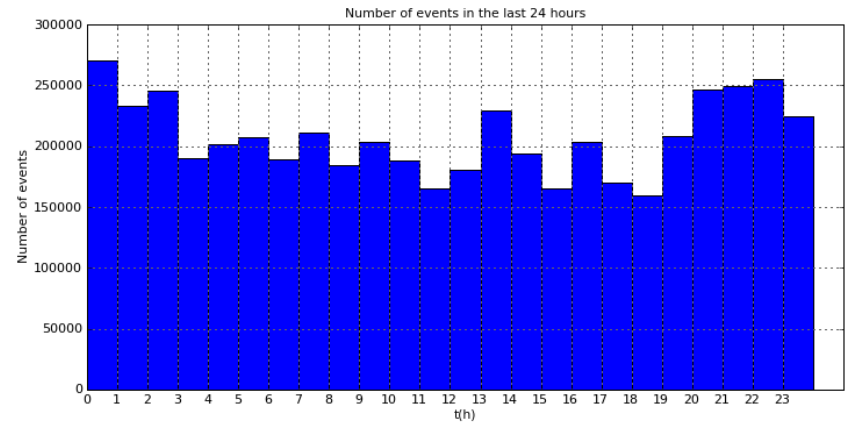
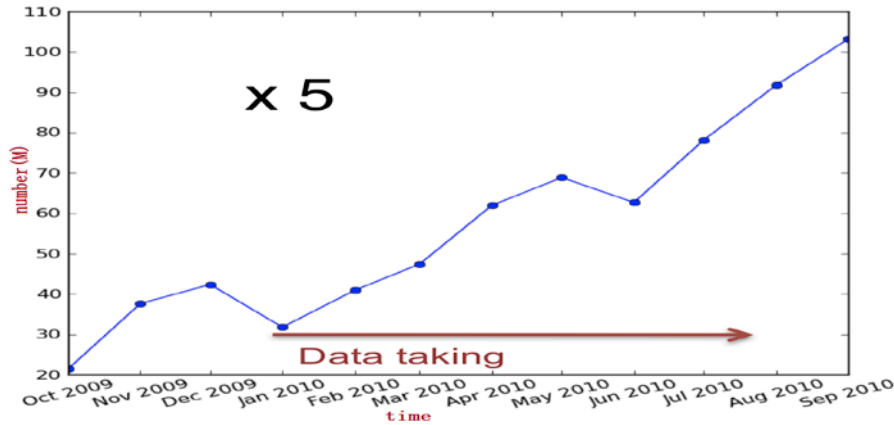
- DDM service, used to trace ATLAS file operations on the grid
 - type, status, local site, remote site, file size, time, etc
- Used by dq2 client tools (dq2-get,dq2-put) and other apps (PanDA)
- Traces can be analyzed for many purpose
 - dataset popularity (popularity.cern.ch by Angelos)
 - replay and simulations (see Martin's talk)
 - user behavior analysis
 - DDM system monitoring
 - ...

Huge Number of File Operations

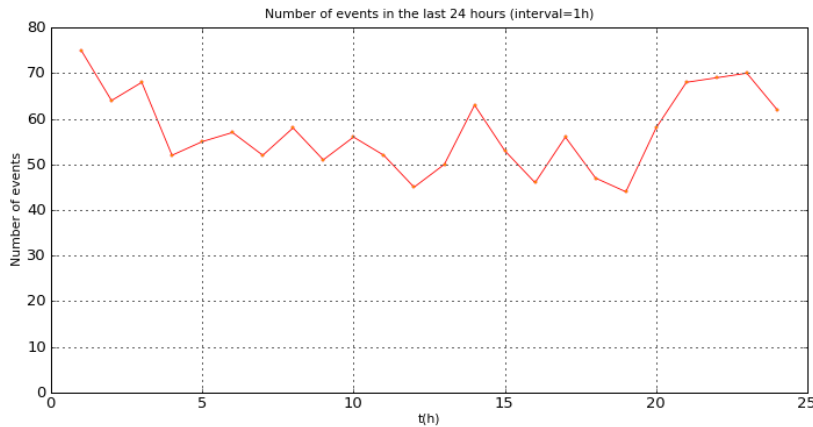
(based on online oracle db)

traces are growing fast

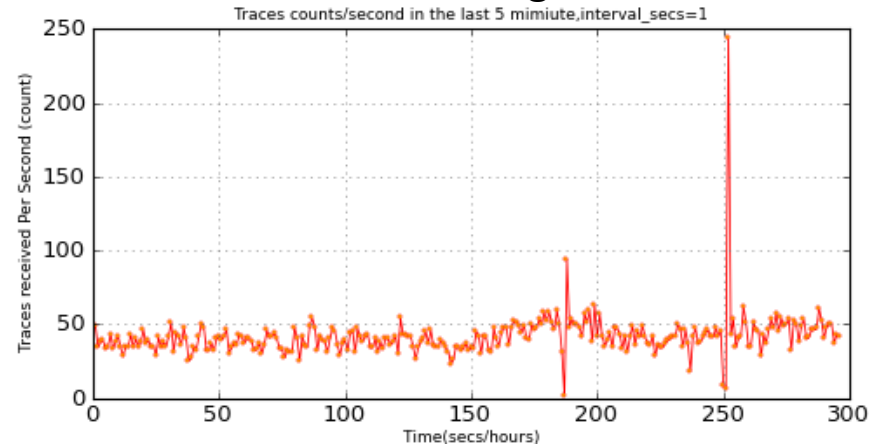
~5 million/day



average ~60/second

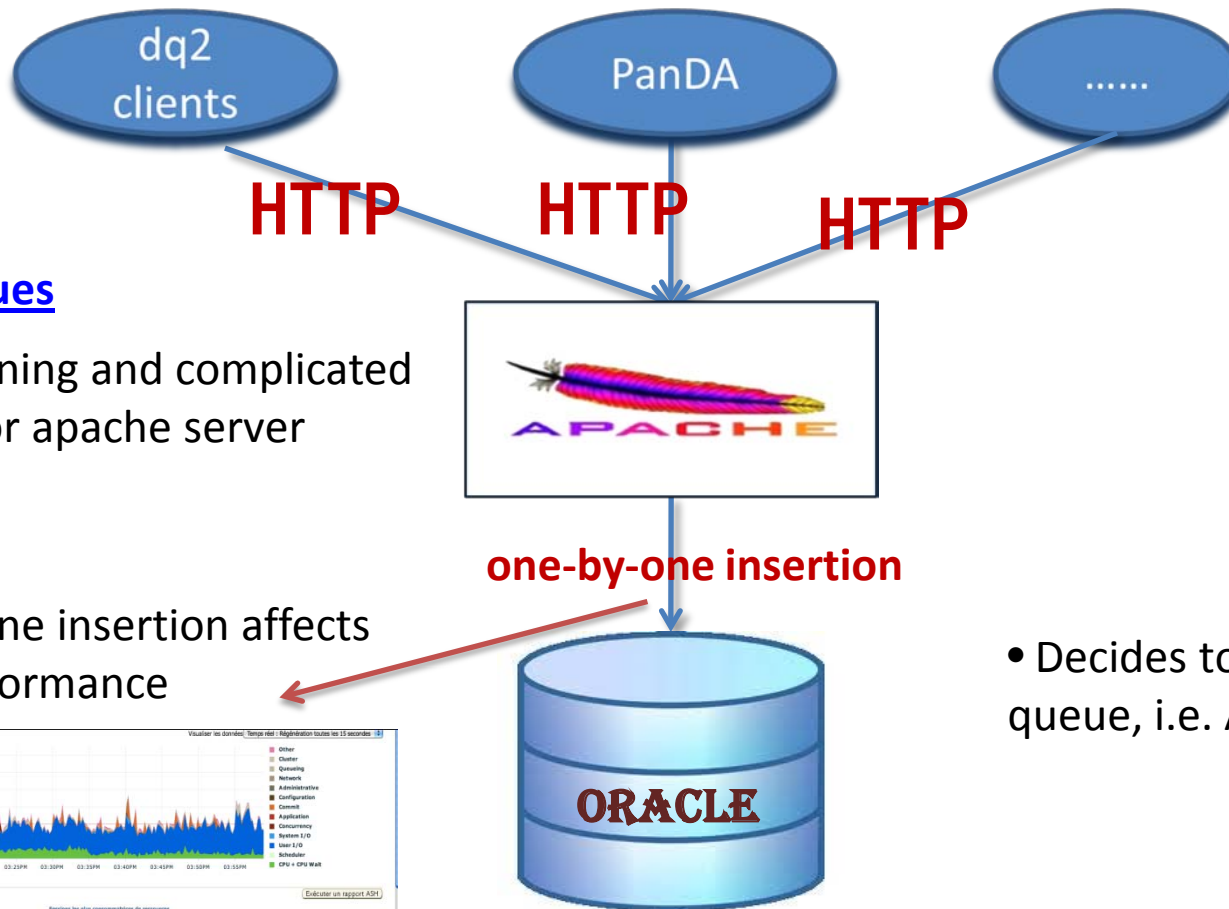


can be 250/second or higher sometimes



<http://bourricot.cern.ch/dq2/tracer/>
atlas-adc-ddm-lab@cern.ch

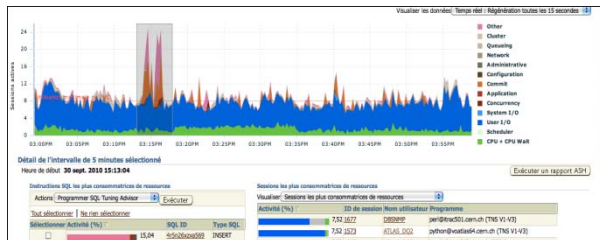
Problems with Current Structure



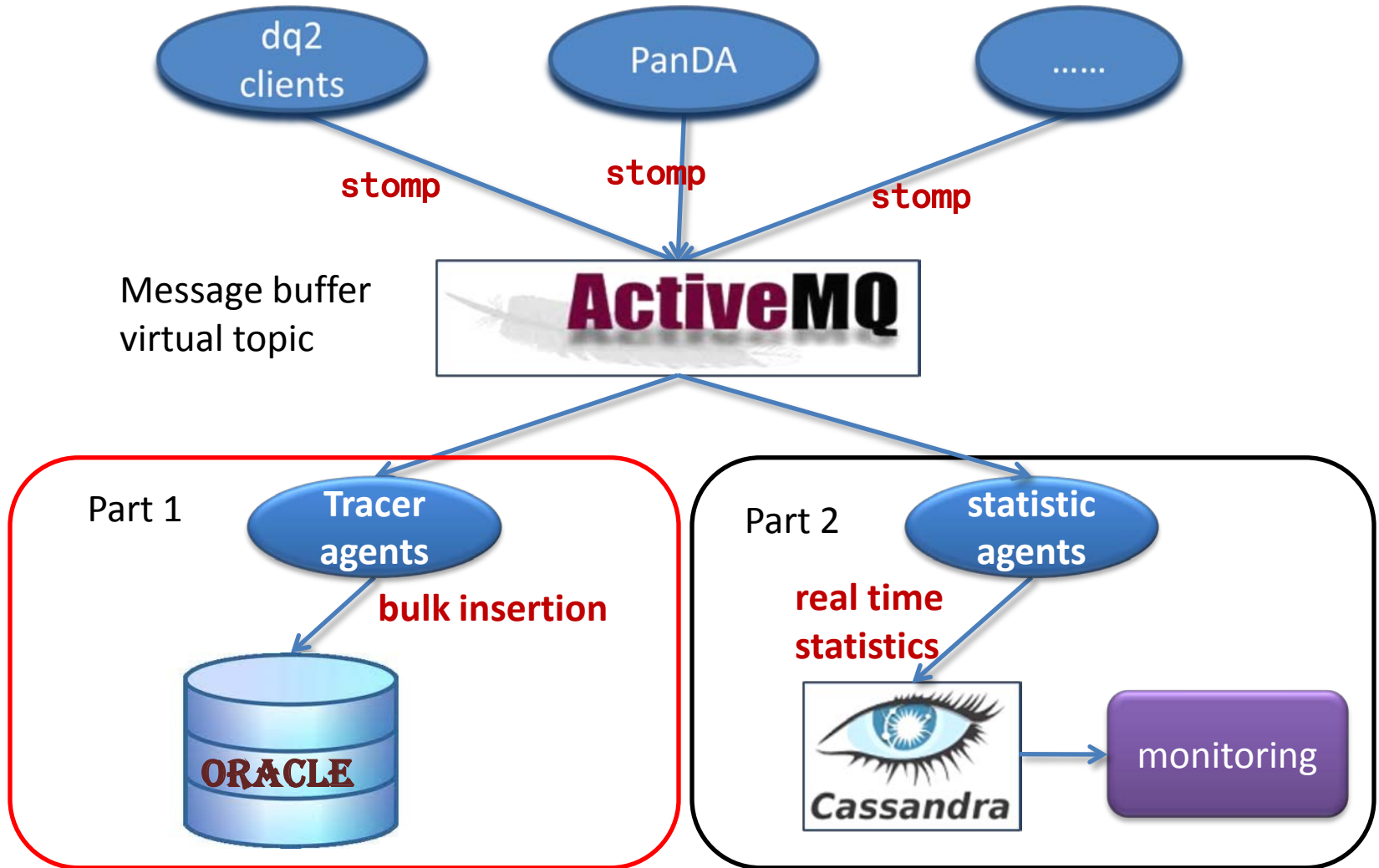
Scalability issues

- Lot of tuning and complicated settings for apache server
- One-by-one insertion affects oracle performance

- Decides to use a message queue, i.e. ActiveMQ



Using ActiveMQ



About ActiveMQ

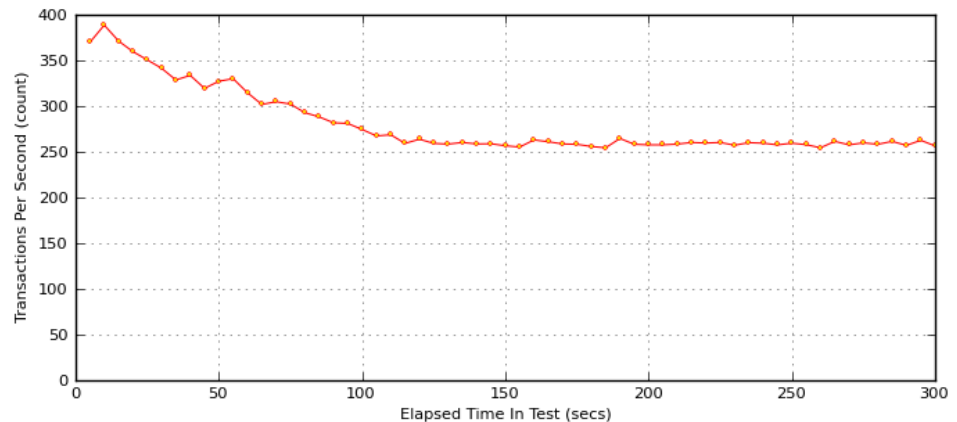
- Message Queue server
- Under Apache 2.0 License
- Fully supports JMS 1.1 and J2EE 1.4
- Supports stomp (Streaming Text Orientated Messaging Protocol)
 - Most popular languages: C, Ruby, Perl, Python, PHP, ActionScript/Flash, Smalltalk
- Supported by CERN IT as a general service

Stress Tests

- using multi-mechanize(<http://code.google.com/p/multi-mechanize/>)
- for all the test results:
 - http://atlddm25.cern.ch/multi-mechanize_1.010/projects/dq2/results/
- for the test case below:
 - 5 group,1000 threads for 300 seconds

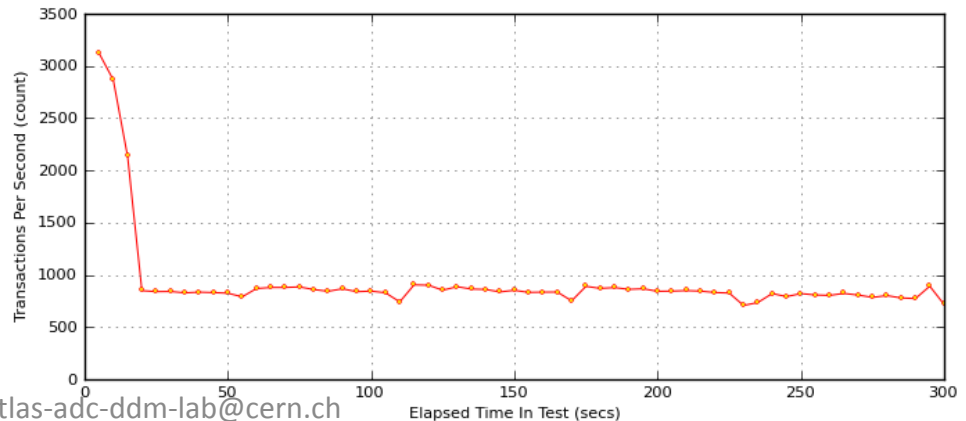
1.HTTP

- one server on the int. testbed
- same settings with central catalogs
- transactions: 85176
- ~300/second
- received: 83486
- lost:1690

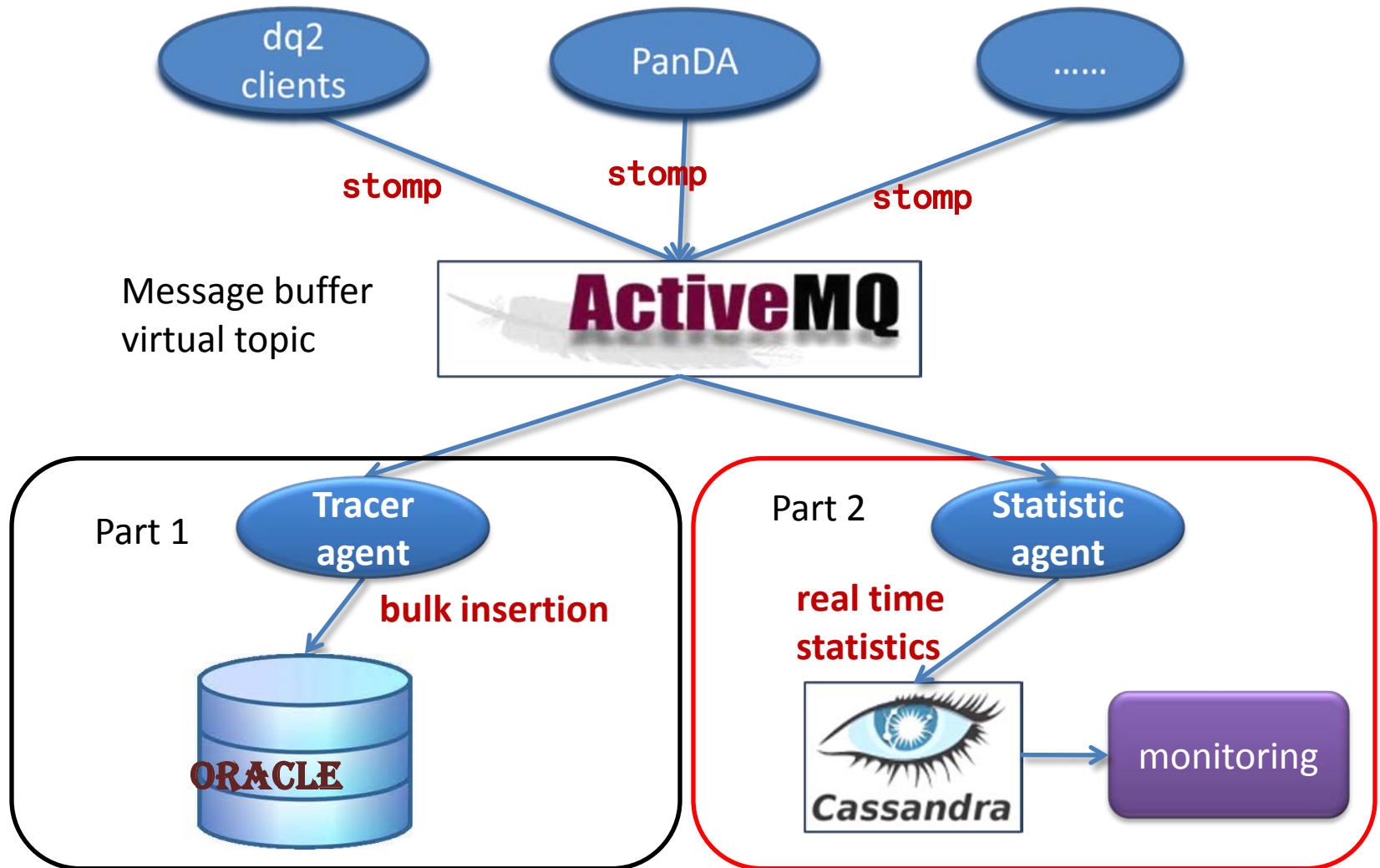


2.ActiveMQ

- service provided by CERN/IT
- events persistent
- transactions: 278827
- ~900/second
- no lost

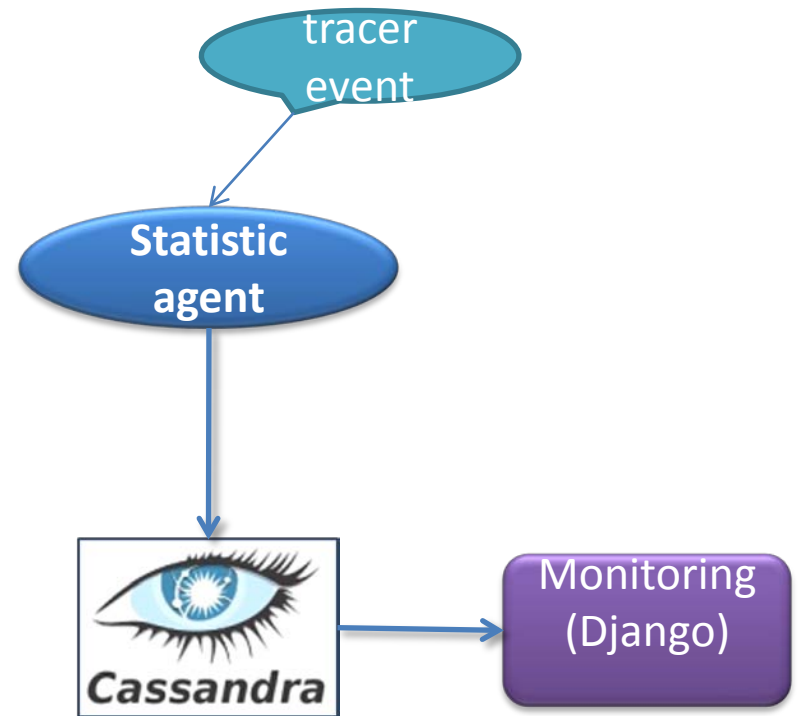


Using ActiveMQ



Monitoring Architecture

- Almost real time monitoring
- Do monitoring based on incremental statistics
- Easy to add new monitoring metrics
- Flush the results to storage at regular times
 - Oracle ?
 - key-value stores ?



<http://atlddm25/dq2/tracer/...>

Tracer Monitoring Metrics

- Metrics
 - Rate of requests
 - Rate of failures
 - Transfer rate
 - ...
- Metric attributes
 - count - the number of points accumulated in the interval
 - average - the average of all the file size over the interval
 - min - smallest observation
 - max - largest observation
 - sum - the sum of all the point values in the interval
 - median - the median of all the point values in the interval
- Intervals
 - 5 min
 - 60 min
 - 4 hr
 - 12 hr
 - 1 day
- Granularity
 - all sites
 - Remotesite
 - Localsite
 - combination: remotesite-localsite
 - event type
 - Users
 -

A lot of combinations and permutations !

⇒ Limitation with Oracle

⇒ Key-Value Store ? (Cassandra? Hbase? – Mario's talk)



- A structured key-value store
- Based on Dynamo(Amazon) and BigTable(Google)
- Used by Facebook, Twitter...
- Open source under apache license
- Fast, scalable

Cassandra Data Models

Column family **interval_5minutes**

Super Column Name
Column Name

201011241755	transferRate			local_read			...
	count	max	...	count	average	...	--
key	17703	5470629.0	--	11436	518120.07	--	

Column Value

Interval_5minutes

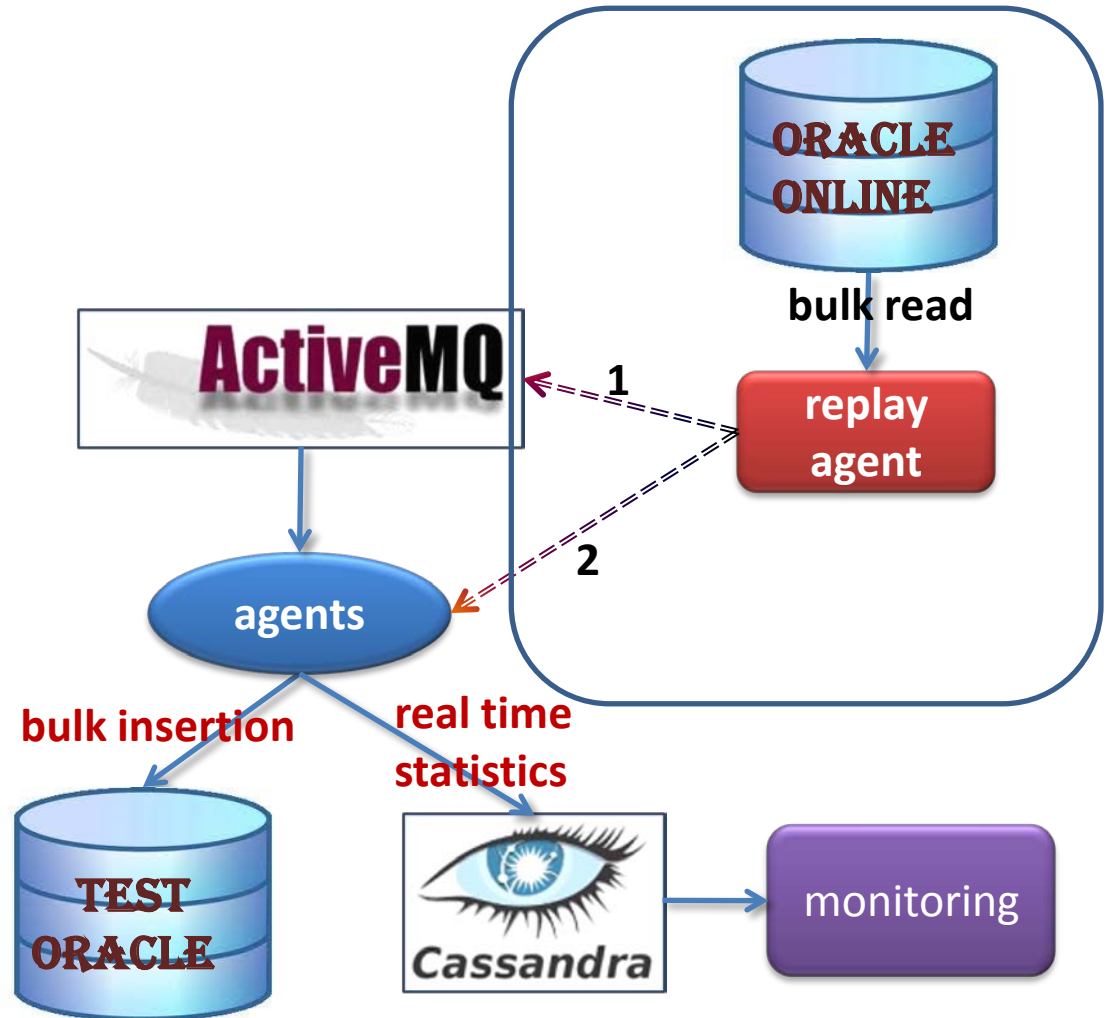
```

201011102105 : {
  local_read :
  {count:11436,min:0.0000,max:5507793.0,average:518120.07,median:467403173.0},
  transferRate :
  {count:17703,min:0.0,max:5470629.0,average:14121827.82,median:14518.892549},
  ....
}
    
```

Interval_1hour ...

Tracer Replay Agent

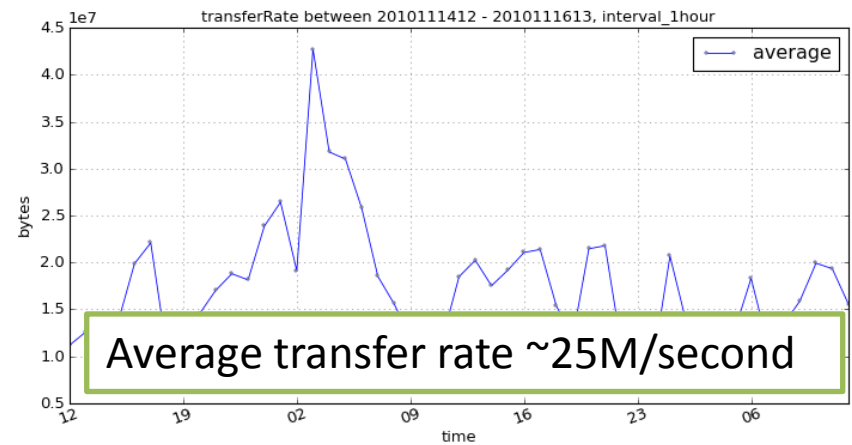
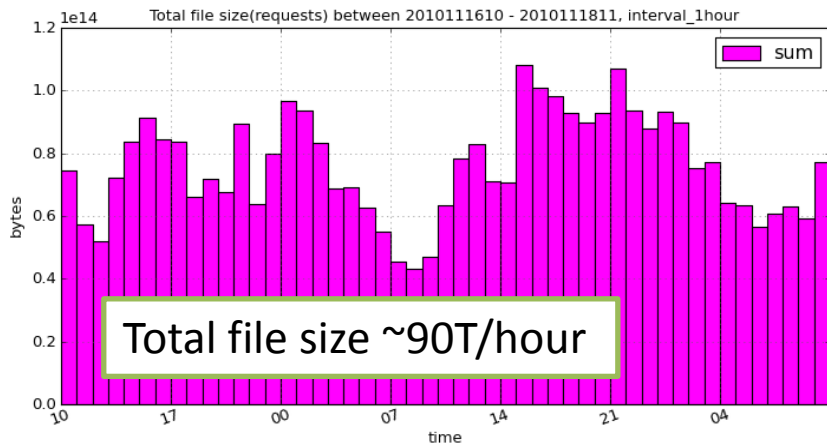
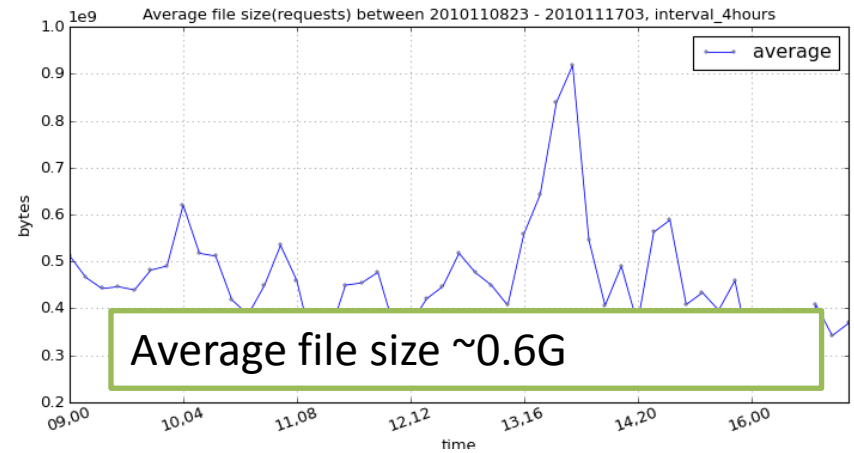
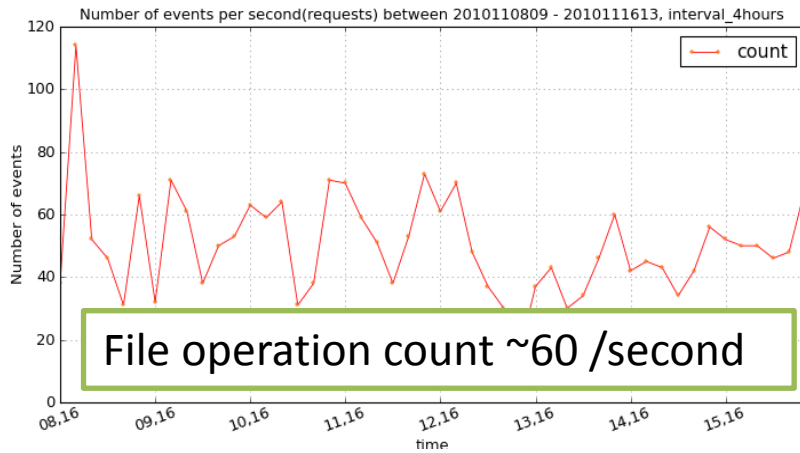
- Read and re-send the tracer events
- Used to test:
 - ActiveMQ
 - Tracer agents
 - New Oracle DB
 - Cassandra
- Have two ways:
 - Send to MQ
 - Send to Listeners



Monitoring Plots

(based on replay and cassandra)

Django + matplotlib <http://atlddm25/dq2/tracer/>



Conclusion

- The dq2 tracer monitoring is very useful
 - Immediate feedbacks of grid activities
- ActiveMQ plays a key role
 - Scalability
 - Used by other applications
- More plots will be provided...
- To be put in production before the end of this year, after more tests under the Tracer Replay Agents
 - New stable dq2clients and STOMP (Ref. Angelos'talk)

Thanks!
谢谢!