



Elasticsearch in Dashboard Data Management Applications

David Tuckett
IT/SDC

30 August 2013
(Appendix 11 November 2013)



Contents

- Elasticsearch
- Data In
- Data Out
- Grouping
- Matrix Queries
- Plot Queries
- Conclusion
- Caveats

Elasticsearch

- **elasticsearch.**

*“flexible and powerful open source, distributed real-time search and analytics engine for the cloud
cool. bonsai cool”*

(<http://www.elasticsearch.org/>)

- Features: real time data, real time analytics, distributed, high availability, multi-tenancy, full text search, document oriented, conflict management, schema free, restful api, per-operation persistence, apache 2 open source license, build on top of apache lucene.

- Our set-up

- 6 VMs : 1 master node, 5 data nodes
- Latest release: 0.90.3



Data In : Example

- 1 month (July 2013) of statistics in 10 minute bins from WLCG Transfers Dashboard
- 12868127 rows
- Typical row

```
{  
  "src_site": "PIC",  
  "src_host": "srmcms.pic.es",  
  "dst_site": "GRIF",  
  "dst_host": "polgrid4.in2p3.fr",  
  "vo": "cms",  
  "server": "https://fts.pic.es:8443/glite-data-transfer-  
fts/services/FileTransfer",  
  "channel": "PIC-HTCMSDSTGROUP",  
  "technology": "fts",  
  "is_remote_access": 1,  
  "is_transfer": 1,  
  "period_end_time": "2013-07-01T00:10:00",  
  "files_xs": 1,  
  "errors_xs": 0,  
  "bytes_xs": 2692475011,  
  "streams": 10,  
  "transfer_time": 141,  
  "srm_preparation_time": 7,  
  "srm_finalization_time": 13  
}
```

Data In : Upload

- Define index

```
{
  "mappings" : {
    "tfr_metric_ng": {
      "_timestamp" : { "enabled" : "true", "store":"yes" },
      "_ttl" : { "enabled": "true"},
      "properties" : {
        "src_site" : {"type" : "string", "index" : "not_analyzed" },
        "dst_site" : {"type" : "string", "index" : "not_analyzed" },
        "src_host" : {"type" : "string", "index" : "not_analyzed" },
        "dst_host" : {"type" : "string", "index" : "not_analyzed" },
        "vo" : {"type" : "string", "index" : "not_analyzed" },
        "server" : {"type" : "string", "index" : "not_analyzed" },
        "channel" : {"type" : "string", "index" : "not_analyzed" },
        "technology" : {"type" : "string", "index" : "not_analyzed" },
        "is_remote_access" : {"type" : "boolean" },
        "is_transfer" : {"type" : "boolean" },
        "period_end_time" : {"type" : "date"},
        "files_xs" : {"type" : "integer"},
        "errors_xs" : {"type" : "integer"},
        "bytes_xs" : {"type" : "long"},
        "streams" : {"type" : "integer"},
        "transfer_time" : {"type" : "integer"},
        "srm_preparation_time" : {"type" : "integer"},
        "srm_finalization_time" : {"type" : "integer"}
      }
    }
  }
}
```

```
curl -XPOST dashb-ai-415:9200/tfr_metric_ng_201307 --data-binary @typedef_ng.txt
```

- Upload data

```
curl -s -XPOST dashb-ai-415:9200/_bulk --data-binary @table_export_ng_201307.txt
```

Data Out : Matrix

- Matrix statistics
 - bytes, successes, failures

Intervals

- 4h
- 24h
- 48h

Filtering:

- technology = fts
- vo = atlas

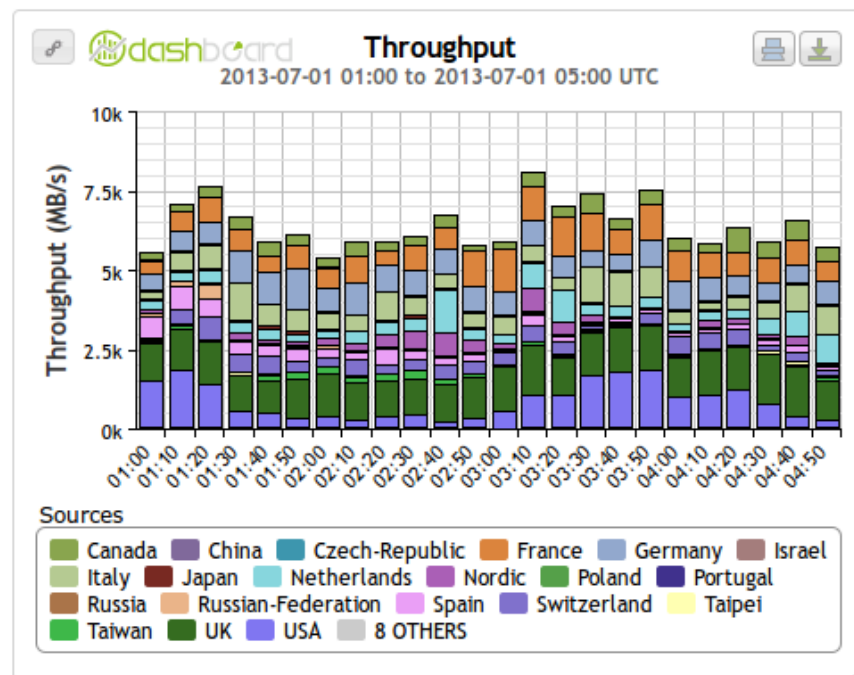
Grouping

- UI = src_country, dst_country
- DB = src_site, src_host, dst_site, dst_host, technology

	TOTAL	Australia+	Austria+	Canada+	China+	Czech-Republic+	France+	Germany+	Israel+	Italy+	Japan+	Netherlands+	Norfolk+	Poland+	Portugal+	Romania+	Russia+	Russian-Federation+	Slovak-Republic+	South-Africa+	
TOTAL	89 % 6 GB/s 148965 16890	14 % 655 kB/s 18	31 % 27 kB/s 45	83 % 387 MB/s 250	100 % 2 MB/s 163	17 % 3 MB/s 113	93 % 815 MB/s 16671	87 % 766 MB/s 22246	35 % 24 MB/s 1622	93 % 640 MB/s 13326	95 % 34 MB/s 883	92 % 462 MB/s 8925	93 % 236 MB/s 8167	78 % 2 MB/s 981	100 % 992 kB/s 83	90 % 237 kB/s 127	96 % 6 MB/s 839	99 % 63 MB/s 2326	100 % 398 kB/s 159	98 % 121 kB/s 80	
Armenia+	100 % 4 kB/s 10 0											100 % 4 kB/s 10 0									
Austria+	0 % 0 kB/s 0 1					0 % 0 kB/s 0 1															
Canada+	95 % 264 MB/s 7911 401			100 % 78 MB/s 1390 2			91 % 21 MB/s 234	98 % 12 MB/s 1502	0 % 0 kB/s 0	99 % 15 MB/s 1184		100 % 5 MB/s 60 0	100 % 0 kB/s 5 0								
China+	100 % 2 kB/s 5 0				100 % 2 kB/s 5 0																
Czech-Republic+	43 % 164 MB/s 981 1293			100 % 4 MB/s 0		0 % 0 kB/s 1	79 % 6 MB/s 22	1 % 849 kB/s 6	100 % 6 MB/s 25	100 % 35 MB/s 117	100 % 62 kB/s 2	100 % 25 MB/s 200	100 % 0 kB/s 5 0							100 % 2 MB/s 7 0	
France+	96 % 856 MB/s 14979 630			99 % 11 MB/s 205	100 % 2 MB/s 163		97 % 278 MB/s 4037	91 % 40 MB/s 133	11 % 3 kB/s 18	99 % 163 MB/s 2543	94 % 10 MB/s 137	100 % 3 MB/s 63	100 % 18 MB/s 445			98 % 98 kB/s 51	53 % 2 MB/s 39	0 % 0 kB/s 0			
Germany+	94 % 957 MB/s 21972 1361	21 % 4 kB/s 3	96 % 61 MB/s 180		0 % 0 kB/s 111	93 % 109 MB/s 698	94 % 213 MB/s 9162	14 % 60 kB/s 8	85 % 322 kB/s 495	100 % 54 MB/s 1	100 % 23 MB/s 2128	100 % 1701 5	78 % 2 MB/s 981				100 % 4 MB/s 800	93 % 4 MB/s 13	100 % 321 kB/s 149	100 % 4 MB/s 0	
Greece+	100 % 13 kB/s 12 0									100 % 13 kB/s 12 0											
Israel+	43 % 38 MB/s 555 744	14 % 655 kB/s 18	20 % 4 kB/s 12	13 % 76 kB/s 10	100 % 3 MB/s 22	100 % 153 kB/s 20	10 % 2 MB/s 23		63 % 9 MB/s 69	100 % 409 kB/s 4	87 % 14 MB/s 154	50 % 7 kB/s 24				83 % 76 kB/s 10		91 % 266 kB/s 2	100 % 76 kB/s 21	100 % 10 kB/s 10	
Italy+	87 % 595 MB/s 6437			100 % 2 MB/s 30		94 % 126 MB/s 1242	74 % 98 MB/s 720		75 % 86 MB/s 1977	100 % 675 kB/s 3	94 % 48 MB/s 394	100 % 36 MB/s 242									100 % 121 kB/s 80

Data Out : Plots

- Plot statistics
 - bytes, successes, failures
- Intervals
 - 4h (10 minute bins)
 - 24h (hourly bins)
 - 48h (hourly bins)
- Filtering:
 - technology = fts
 - vo = atlas
- Grouping
 - UI = src_country
 - DB = src_site, src_host, technology



Grouping : Elasticsearch Features

- *“The usual purpose of a full-text search engine is to return a small number of documents matching your query. Facets provide aggregated data based on a search query.”*
(<http://www.elasticsearch.org/guide/reference/api/search/facets/>)
- Facets
 - The `statistical` facet computes count, total, sum of squares, mean, min, max, var, and std dev on a numeric field.
 - The `terms_stats` facet groups statistical data by terms of a single field.
 - The `date_histogram` facet groups statistical data by time bins
- In the current release 0.90.3, facets cannot be combined, so grouping by terms of multiple fields for statistical aggregations is not supported.
- The future release 1.0, will have an Aggregation Module that will support combining Calc aggregators and Bucket aggregators.
<https://github.com/elasticsearch/elasticsearch/issues/3300>

Grouping : Oracle & Elasticsearch Methods

- OSG: Oracle Static Grouping
 - Query using “group by” on all possible grouping fields.
 - Further grouping for selected grouping fields in web action.
- ODG: Oracle Dynamic Grouping
 - Query using “group by” for selected grouping fields.
- ENG: Elasticsearch No Grouping
 - Query for all data.
 - Grouping in the web action.
- EIG: Elasticsearch Index Grouping
 - Add single field in index with all possible grouping fields concatenated.
 - Query using “terms_stats” facet to group by single field.
 - Further grouping for selected grouping fields in web action.
- EQG: Elasticsearch Query Grouping
 - Query using “terms” facet to list n distinct combinations of selected grouping field values.
 - Query using “date_histogram” facet n times filtering by distinct combinations.

Matrix Queries

- 4h, technology=fts, vo=atlas

Total rows	41782
Filtered rows	5780
DB grouped rows (src_site, src_host, dst_site, dst_host, technology)	834
UI grouped rows (src_country, dst_country)	218

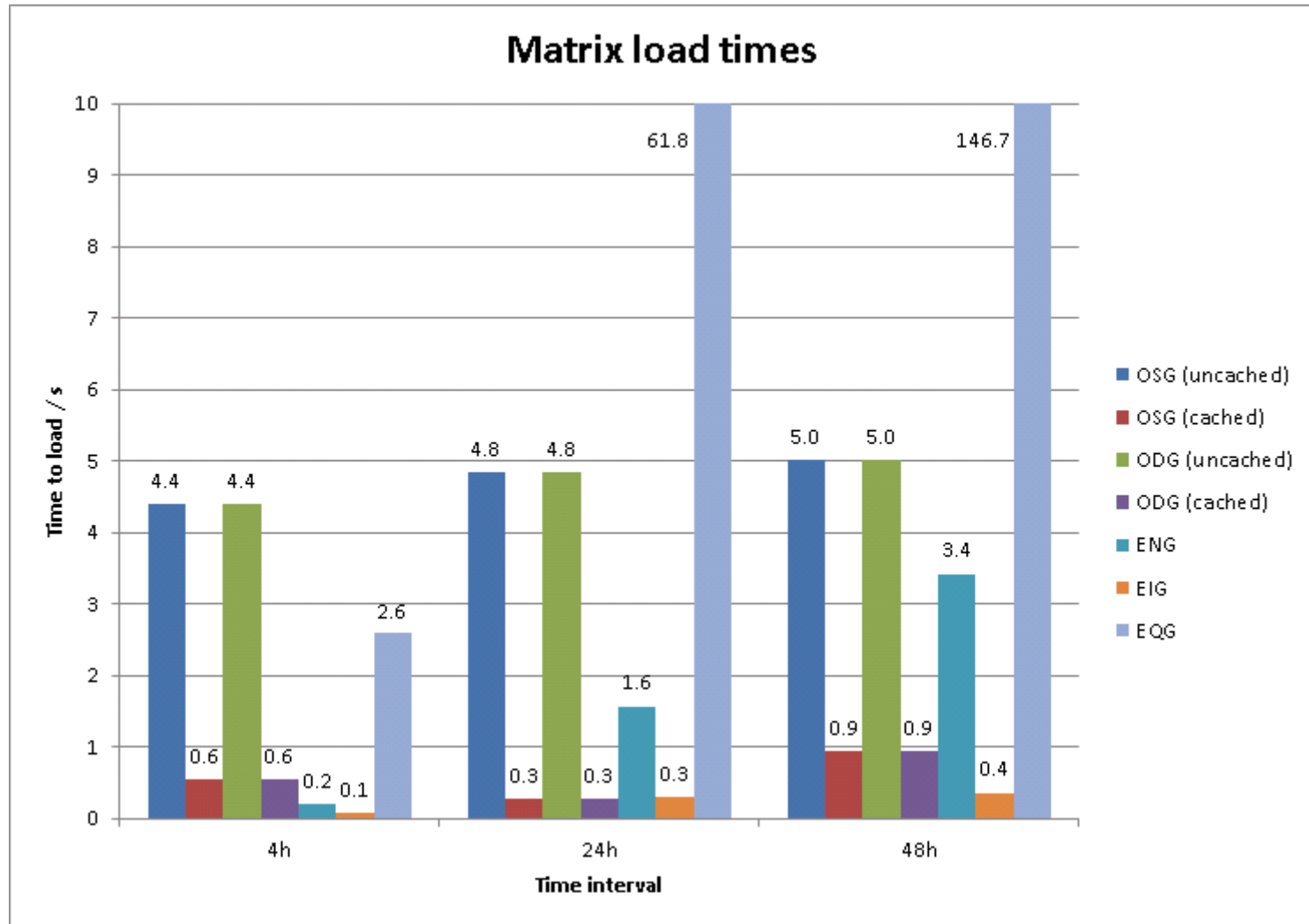
- 24h, technology=fts, vo=atlas

Total rows	261886
Filtered rows	37969
DB grouped rows (src_site, src_host, dst_site, dst_host, technology)	3027
UI grouped rows (src_country, dst_country)	419

- 48h, technology=fts, vo=atlas

Total rows	716218
Filtered rows	80713
DB grouped rows (src_site, src_host, dst_site, dst_host, technology)	4610
UI grouped rows (src_country, dst_country)	466

Matrix Queries



Plot Queries

- 4h, technology=fts, vo=atlas

Total rows	41782
Filtered rows	5780
DB grouped rows (all + bins, selected + bins, selected)	5780, 1611, 106
UI grouped rows (src_country, 10m bin)	473

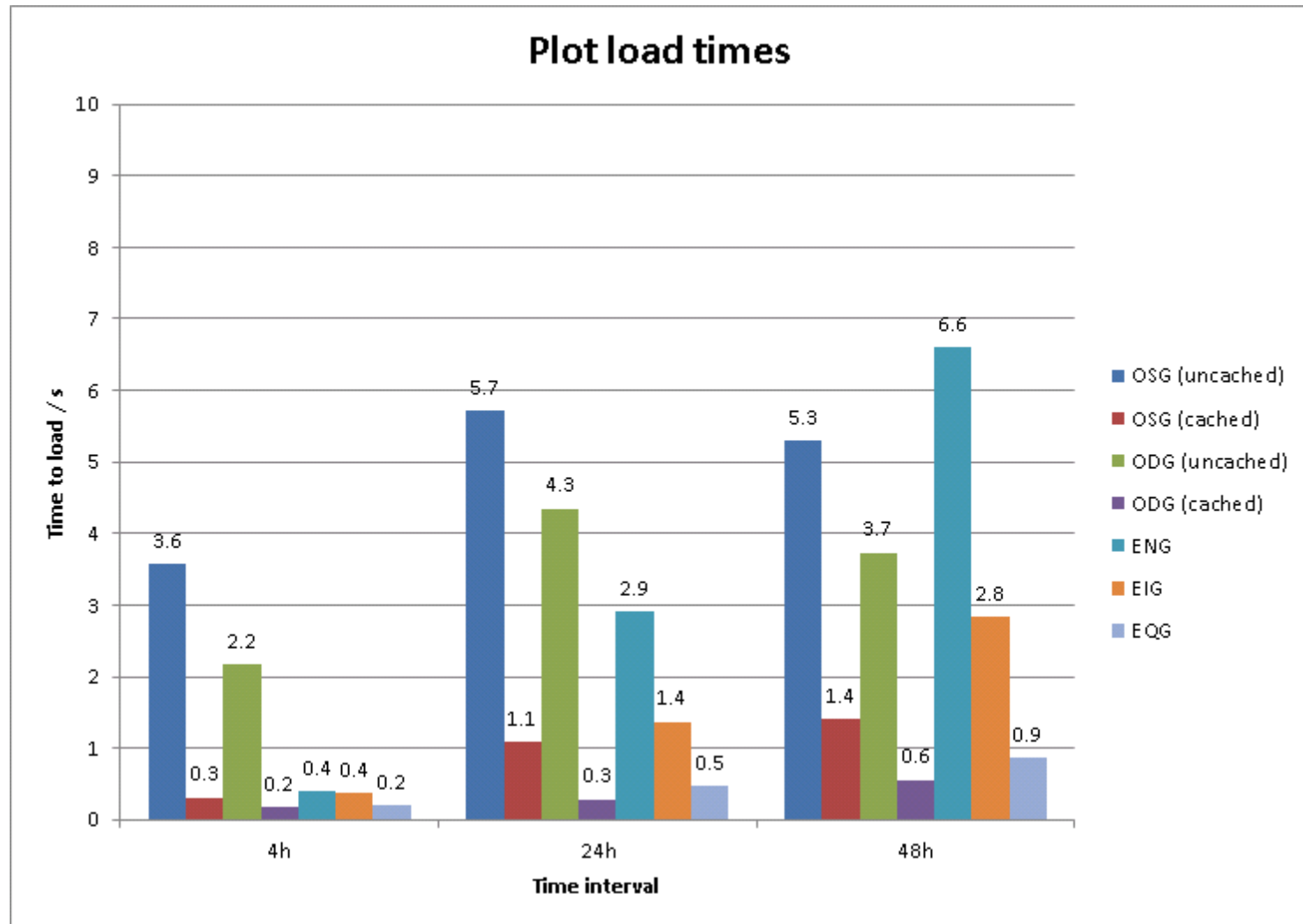
- 24h, technology=fts, vo=atlas

Total rows	261886
Filtered rows	37969
DB grouped rows (all + bins, selected + bins, selected)	14824, 2447, 118
UI grouped rows (src_country, 10m bin)	633

- 48h, technology=fts, vo=atlas

Total rows	716218
Filtered rows	80713
DB grouped rows (all + bins, selected + bins, selected)	31806, 5052, 124
UI grouped rows (src_country, 10m bin)	1308

Plot Queries



Conclusion

- Elasticsearch 0.90.3 does not support grouping by terms of multiple fields for statistical aggregations.
- Using Elasticsearch 0.90.3 for WLCG Transfer Monitoring we could achieve similar performance to 2nd hit, i.e. cached, Oracle performance but this means using diverse workarounds for multi-field grouping.
- Elasticsearch 1.0 includes a new Aggregation Module that will support grouping by terms of multiple fields for statistical aggregations.
- I would recommend re-evaluating ElasticSearch for WLCG Transfer Monitoring when 1.0 is available.

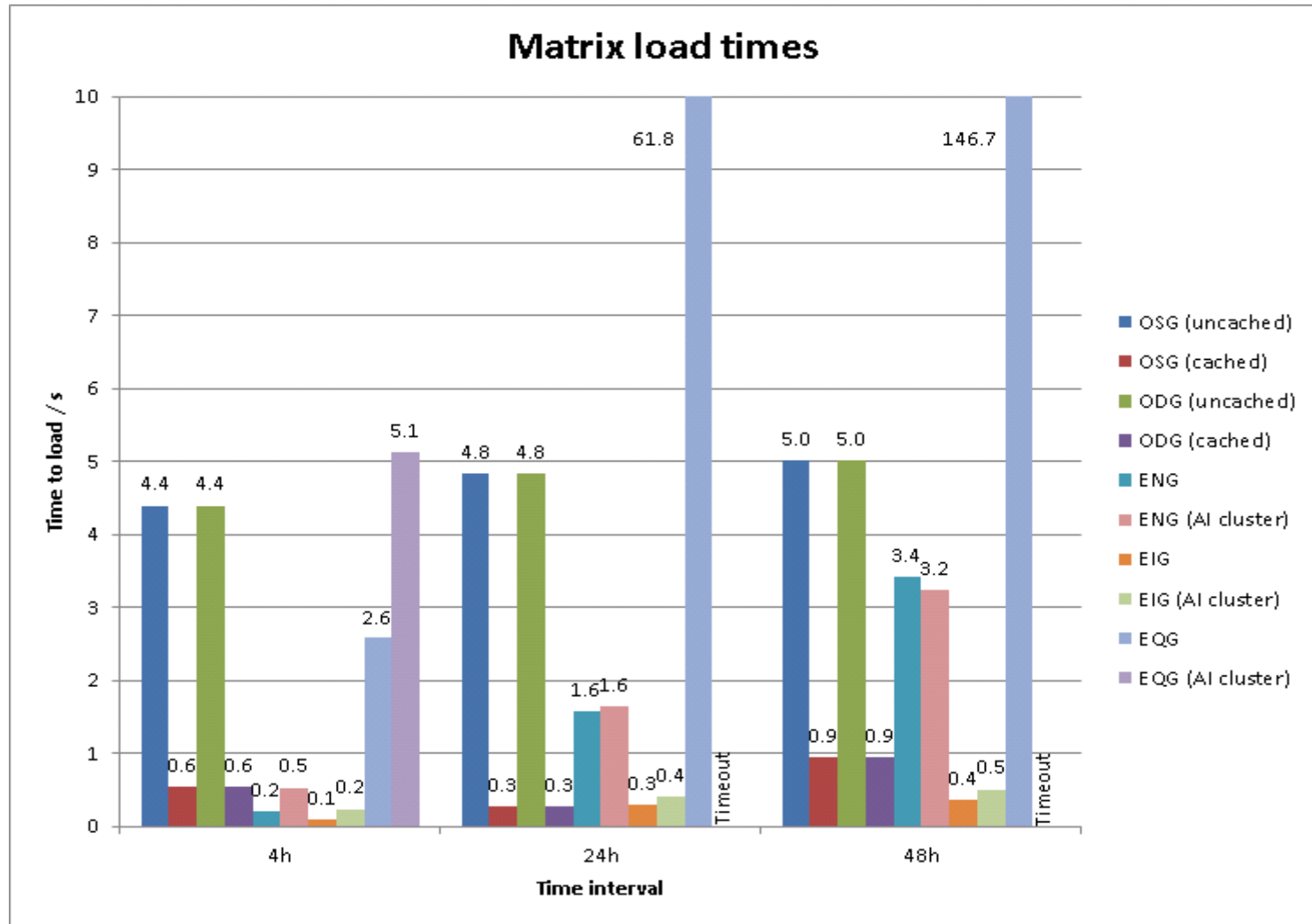
Caveats & Notes

- Oracle had many users when I ran the tests. Elasticsearch probably had 1 user.
- Oracle was running on top spec. hardware. Elasticsearch was running on VMs.
- I didn't use the "parallel" hint that Oracle supports. I probably missed some optimisation options in Elasticsearch.
- ...
- Pablo and Ivan have looked at other aspects of Elasticsearch such as importing data, using Python clients, ...
- I can create a twiki page with the data and queries used in these tests

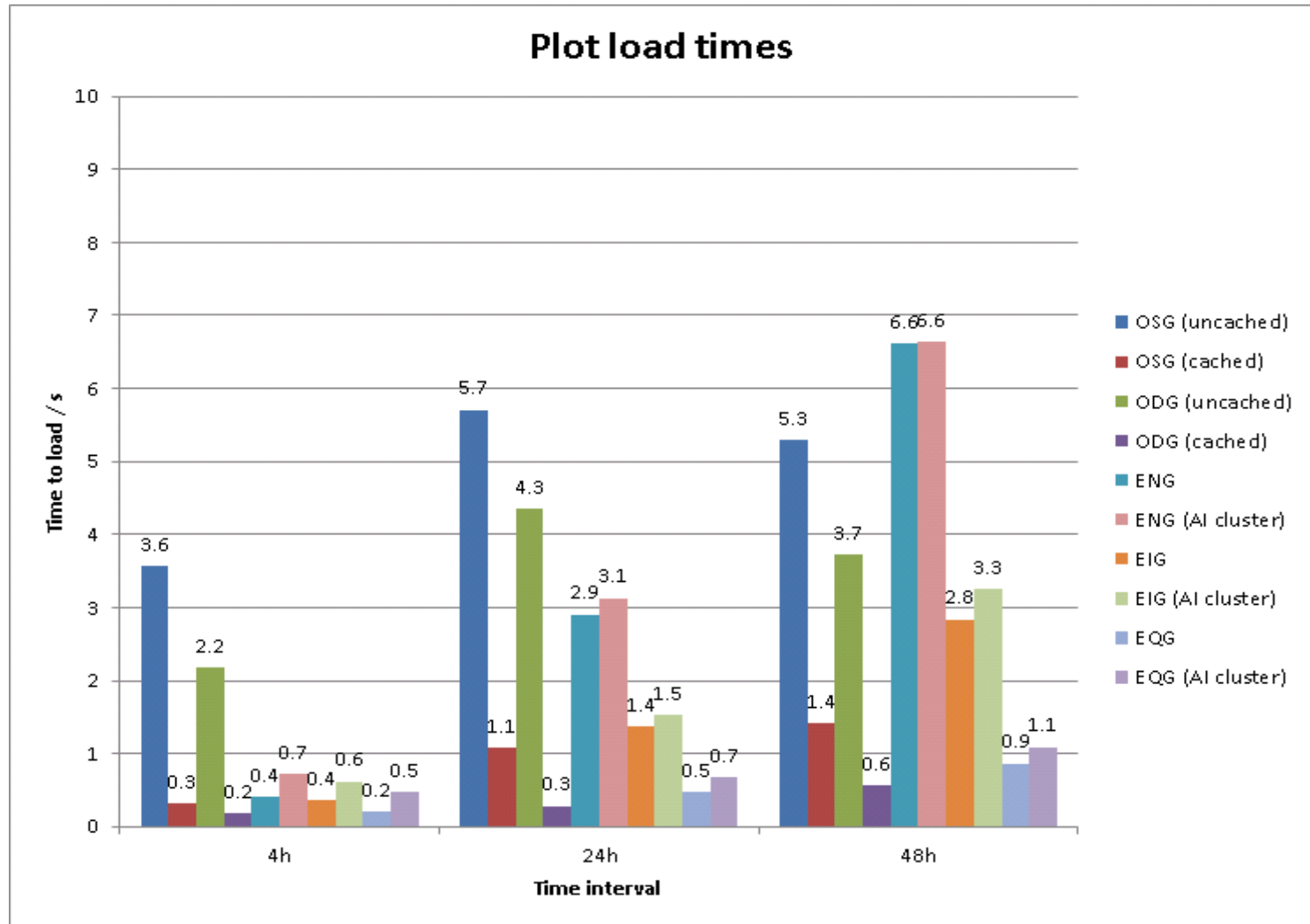
Appendix – 11 November 2013

- This appendix includes charts showing load times for matrix and plot queries using the Agile Infrastructure Monitoring (AI-Mon) Elasticsearch cluster based on physical machines
- The AI-Mon cluster is shared and under continuous load indexing lemon metrics and syslog records
- The comparison is made with those results shown in “Matrix Queries” and “Plot Queries” charts (pages 11 & 13) using the IT-SDC-MI Elasticsearch cluster based on virtual machines
- AI-Mon cluster
 - 11 Physical machines
 - 2 master nodes: 8 cores, 16 GB RAM
 - 1 search node: 8 cores, 16 GB RAM
 - 8 data nodes: 24 cores, 48 GB RAM
 - Elasticsearch release: 0.90.3
 - SSL enabled without authentication for read access

Appendix - Matrix Queries



Appendix - Plot Queries



Appendix – Conclusion

- In most cases, load times for matrix and plot queries using the AI-Mon cluster are slower than using the IT-SDC-MI cluster but still faster than using Oracle without caching
- The increase in load time is no more than 0.5 seconds* and does not generally increase with longer time period queries
(* excluding the EQG method for matrix queries which does not perform well on either cluster)
- It seems that any performance gains expected from using higher specification machines on the shared physical cluster are compensated by performance losses due to load from other applications