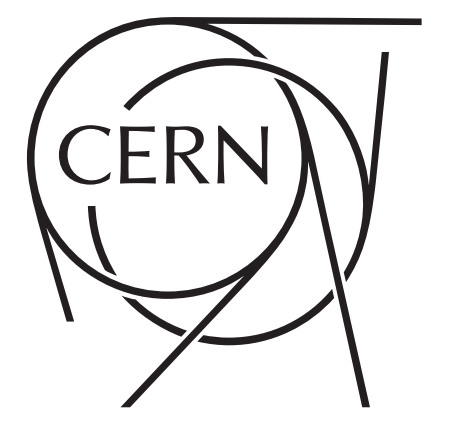


The ATLAS Wide-Range Database & Application Monitoring



Petya Vasileva (CERN), Andrea Formica (Université Paris-Saclay, IRFU (FR)), Gancho Dimitrov (CERN)
on behalf of the ATLAS Collaboration

Introduction

In HEP experiments at LHC the database applications often become complex by reflecting the ever demanding requirements of the researchers. The ATLAS experiment has several Oracle DB clusters with over 216 database schemas each with its own set of database objects. To effectively monitor them, we designed a modern and portable application with exceptionally good characteristics. Some of them include: concise view of the most important DB metrics; top SQL statements based on CPU, executions, block reads, etc.; volume growth plots per schema and DB object type; database jobs section with signaling for problematic ones; in-depth analysis in case of contention on data or processes.

A tool that allows developers to explore and tune application's database performance

See execution plans for each query



Get historical plots of database activity

```

PARSING_SCHEMA_NAME: ATLAS_PANDA_WRITER
Active execution plan: 2385149161 since 2018-06-18/03:27:34
SQL_ID: 63a5undrtabhf

SELECT status FROM ATLAS_PANDA.JEDI_Datasets WHERE
jediTaskID=:jediTaskID AND datasetID=:datasetID FOR UPDATE /*
DBProxy.getFuntRanges */

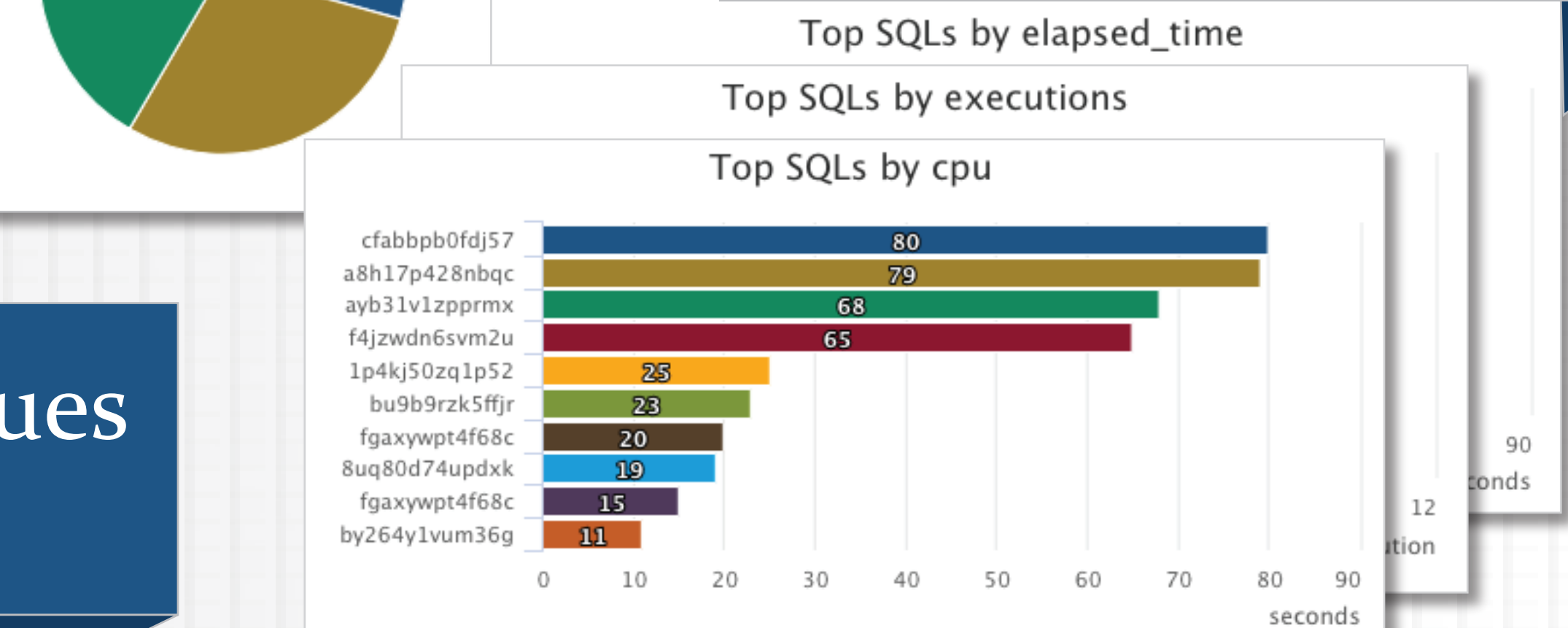
Plan hash value: 2385149161

+-----+-----+-----+-----+-----+-----+-----+-----+
| Id | Operation | Name | Rows | Bytes | Cost (%CPU) | Time | Pstart | Pstop |
+-----+-----+-----+-----+-----+-----+-----+-----+
| 0 | SELECT STATEMENT | | | | 3 (100) | | | |
| 1 | FOR UPDATE | | | | | | | |
| 2 | PARTITION RANGE SINGLE | | 1 | 19 | 3 (0) | 00:00:31 | KEY | KEY |
| 3 | TABLE ACCESS BY LOCAL INDEX ROWID | JEDI_DATASETS | 1 | 19 | 3 (0) | 00:00:31 | KEY | KEY |
| 4 | INDEX UNIQUE SCAN | JEDI_DATASETS_PK | 1 | | 2 (0) | 00:00:31 | KEY | KEY |
    
```

Top SQLs by rows_processed
Top SQLs by disk_reads



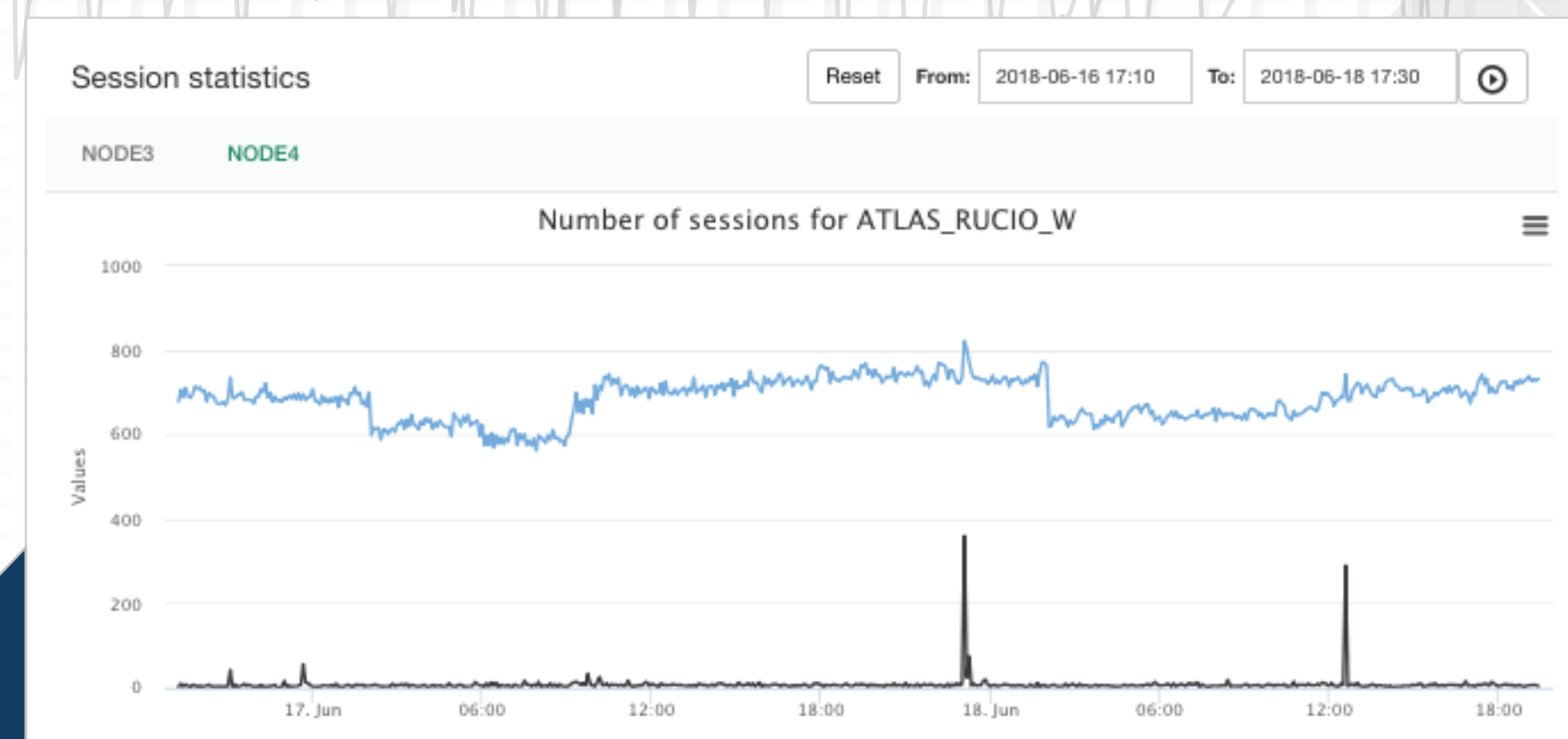
Monitor the top resource consumers



Notice increased values for basic DB metrics

Metric	Node1	Node2	Node3	Node4
User Transaction Per Sec	40	415	2487	70
SQL Service Response Time	1	0	0	0
Session Count	109	290	1309	138
Physical Reads Per Sec	18056	155	85	2
Logons Per Sec	3	3	32	3
Logical Reads Per Sec	34314	130510	130041	21307
Host CPU Utilization (%)	7	18	21	3
Current OS Load	3	5	5	1
Average Active Sessions	3	7	8	1

Track # active/inactive sessions per DB user



Check database jobs' state

Owner	Job Name	Last Start Date	Last Run Dur	Last Status	Current State	Next Run Time	Repeat Interval
ATLAS_PANDA	ADG_DAILYPART_PANDA	18-06-2018 10:00:00	0:00:00.03	SUCCEEDED	SCHEDULED	25-06-2018 10:00:00	FREQ=MONTHLY;INTERVAL=1;BYDAY=MON
ATLAS_DBA	ADG_SESSION_INFO_DROP_DATA_JOB	11-06-2018 09:00:00	0:00:00.00	SUCCEEDED	SCHEDULED	09-07-2018 09:00:00	FREQ=MONTHLY;INTERVAL=1;BYDAY=MON
ATLAS_DBA	ADG_SESSIONS_SNAPSHOT_INFO_JOB	18-06-2018 15:50:00	0:00:00.00	SUCCEEDED	SCHEDULED	18-06-2018 16:05:00	FREQ=MINUTELY;INTERVAL=15;
ATLAS_DBA	ARC-IVE_RUCIO_REQUESTS_HISTORY	18-06-2018 13:00:01	0:01:13.25	SUCCEEDED	SCHEDULED	18-06-2018 13:00:00	FREQ=DAILY;BYHOUR=13;BYMINUTE=0;BYSECOND=0
ATLAS_PANDA	BULKCOPY_PANCAPART_JOB	17-06-2018 18:00:00	0:00:46.25	SUCCEEDED	SCHEDULED	18-06-2018 18:00:00	FREQ=DAILY;INTERVAL=1
ATLAS_PANDAARCH	CACHE_JOBARCHIVED_DATA_BLOCKS	18-06-2018 16:00:00	0:00:00.28	SUCCEEDED	SCHEDULED	18-06-2018 16:07:30	FREQ=MINUTELY;INTERVAL=7;
ATLAS_PANDAARCH	CACHE_JOBARCHIVED_DATA_BLOCKS2	18-06-2018 16:00:00	0:00:00.34	SUCCEEDED	SCHEDULED	18-06-2018 16:10:30	FREQ=MINUTELY;INTERVAL=10;
ATLAS_PANDAARCH	CACHE_JOBARCHIVED_DATA_BLOCKS3	18-06-2018 16:00:00	0:00:00.42	SUCCEEDED	SCHEDULED	18-06-2018 16:10:30	FREQ=MINUTELY;INTERVAL=10;
ATLAS_PANDAARCH	CACHE_JOBARCHIVED_DATA_BLOCKS4	11-06-2018 10:31:00	-	DISABLED	DISABLED	11-06-2018 10:49:30	FREQ=MINUTELY;INTERVAL=18;
ATLAS_PANDAARCH	CACHE_JOBARCHIVED_INDEX	18-06-2018 15:58:00	0:00:02.89	SUCCEEDED	SCHEDULED	18-06-2018 16:07:30	FREQ=MINUTELY;INTERVAL=15;
ATLAS_DBA	CACHE_PANDAMON_JOBSPACE_ARCH	18-06-2018 16:04:00	0:00:00.41	SUCCEEDED	TUNING	18-06-2018 16:04:30	FREQ=MINUTELY;INTERVAL=4;

Detect blocking/blocked sessions

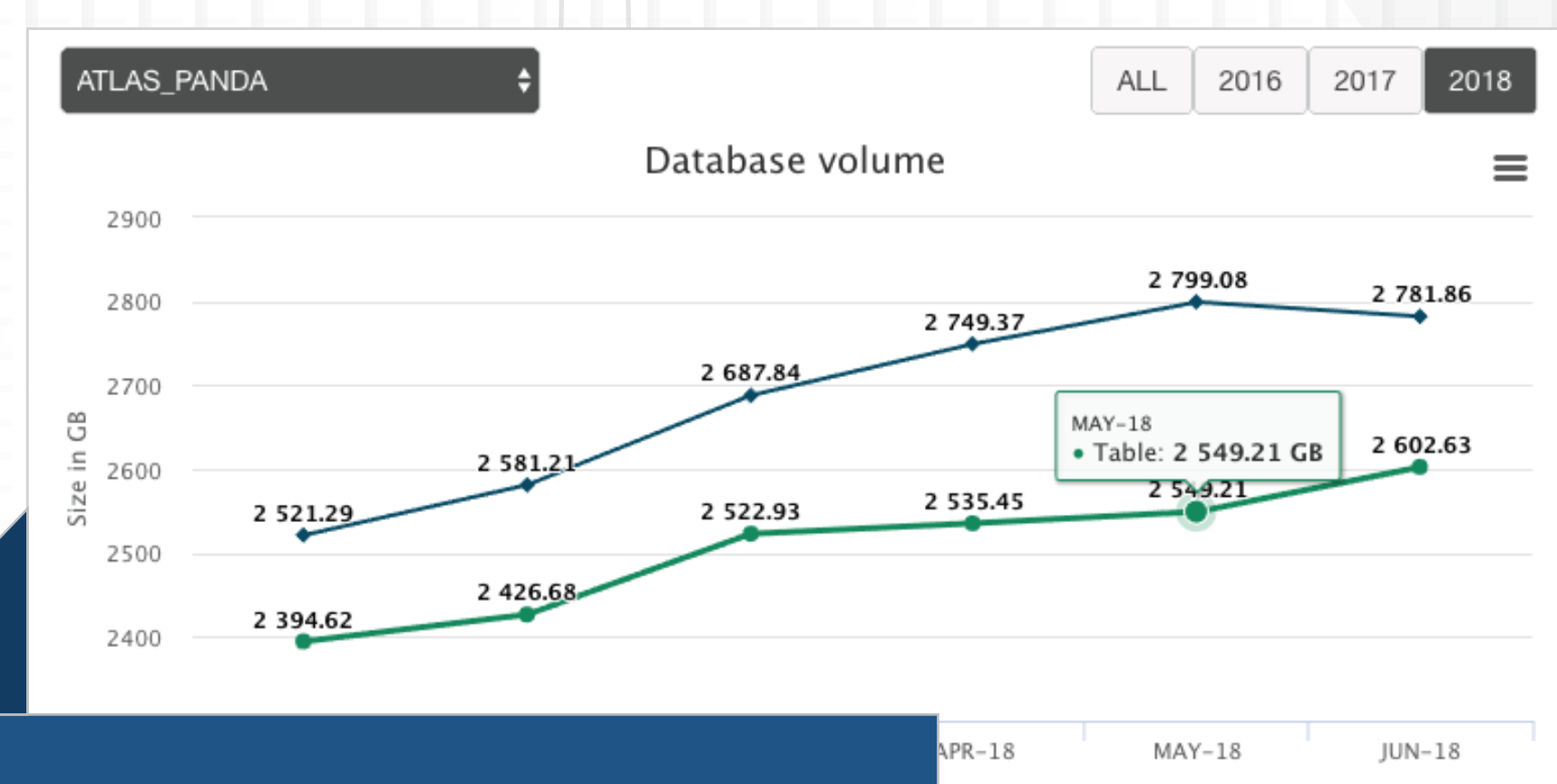
Session ID	Login Time	User Name	OS User	Program	Machine	Wait Class	SQL ID	ADRC	From	To	Run	Reset
56	18-06-2018 14:11:10	ATLAS_RUCIO_W	root	python@rucio-daemon-prod-03.cern.ch [TNS V1-V2]	rucio-daemon-prod-03.cern.ch	Idle	0					
2020	18-06-2018 14:19:45	ATLAS_RUCIO_W	root	python@rucio-daemon-prod-09.cern.ch [TNS V1-V2]	rucio-daemon-prod-09.cern.ch	Idle	2zibhwaDoe	7	ATLAS_RUCIO	REQUESTS	ALU18FAAAM8JUK	
121	18-06-2018 14:10:36	ATLAS_RUCIO_W	root	python@rucio-daemon-prod-03.cern.ch [TNS V1-V2]	rucio-daemon-prod-03.cern.ch	Idle	0					
987	18-06-2018 14:19:18	ATLAS_RUCIO_W	root	python@rucio-daemon-prod-09.cern.ch [TNS V1-V2]	rucio-daemon-prod-09.cern.ch	Idle	2zibhwaDoe	14	ATLAS_RUCIO	REQUESTS	ALU18FAAAM8JUK	
					apand033.cern.ch	Idle	hgrphub4p	250				

Spot significant variations in SQL execution parameters

Spot significant variations in SQL execution parameters

node	begin_time	plan_hash	module	parsing_schema	fetches	sorts	execs	prevsecs	loads	invalid	parallel_calls	disk_reads	buffer_gets	direct_writes	rows_proc	cpu_time	elapsed_time	etime_per_exec	lowwait	cluster_wait	app_wait	concurrency	plsql_time
1	18-06-18 10:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	31597	0	0	0	1	1	117791	740029	230913	0	60	69	1	7	2	0	0	0
1	12-06-18 23:17	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	50334	0	0	0	0	1	97281	942808	200589	0	58	65	1	5	1	0	0	0
1	10-06-18 17:44	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	32482	0	0	0	0	1	147678	1951636	228523	0	86	97	1	10	1	0	0	0
1	08-06-18 21:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	45970	0	0	0	0	1	206622	35507779	269163	0	129	141	1	11	1	0	0	0
1	07-06-18 10:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	15889	1	0	0	1	1	246470	40540346	268730	0	3448028	169	186	1	17	1	0	0
1	06-06-18 16:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	43848	0	0	0	0	1	229853	39040302	251662	0	158	179	1	19	3	0	0	0
1	06-06-18 15:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	55419	0	0	0	0	1	213406	34889964	246711	0	139	163	1	20	1	0	0	0
1	05-06-18 14:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	4267	0	0	0	0	1	201265	30853757	233857	0	127	147	1	16	1	0	0	0
1	04-06-18 12:41	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	55441	0	0	0	0	1	213216	33117208	233136	0	135	155	1	14	6	0	0	0
1	01-06-18 12:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	1980	0	0	0	0	1	170095	18022434	270312	0	93	103	1	8	2	0	0	0
1	29-05-18 09:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	57273	0	0	0	0	1	281694	35308119	301432	0	162	173	1	10	1	0	0	0
1	26-05-18 04:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	804	0	0	0	0	1	248400	36969128	306699	0	152	211	1	5	3	0	0	0
1	24-05-18 03:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	60571	0	0	0	0	1	79072	9096905	200476	0	53	63	1	5	5	0	0	0
1	23-05-18 21:04	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	41615	1	0	0	1	1	237399	41615516	260763	0	3842700	159	188	1	29	1	0	0
1	19-05-18 01:00	892639151	DEMS_SCHEDULER	ATLAS_PANDABIGMON	0	5806	0	0	0	0	1	186047	10074075	281571	0	86	99	1	10	3	0	0	0

Review DB volume growth per schema



And more...

Implementation

The project can be separated into three independent layers. The first layer consists in highly-optimized database objects hiding all complicated calculations in PL/SQL functions and procedures. The second layer represents a Java application providing REST access to the underlying database backend. The third layer is a JavaScript/AngularJS web interface visualizing the data in a descriptive manner. The monitoring uses gitlab-ci pipelines for basic testing, containerization and deployment on the CERN OpenShift infrastructure.