



October 1<sup>st</sup> 2014, Oaktable World 2014

Eric Grancher, CERN-IT

# Using ASH to understand IO issues

CERN DB blog: <http://cern.ch/db-blog/>



# Slow IO is different than IO outlier

- IO tests and planning: SLOB, fio, etc. help to size
- Latency variation is the user experience (C. Millsap)
- 1, 19, 0, 20, 10: average 10
- 9, 11, 10, 9.5, 10.5 : average 10
- 10.01, 9.99, 10, 10.1, 9.9: average 10
- 10.01, 9.99, 10.01, 9.99, **500**, 10.01, 9.99: average ~10
- Ex: web page: 5 SQL statements, 10 IOs per request
- 50 IOs at 0.2 ms = 5 ms
- 50 los at 300 ms = 1500 ms = 1.5s

# Pathological cases, latency

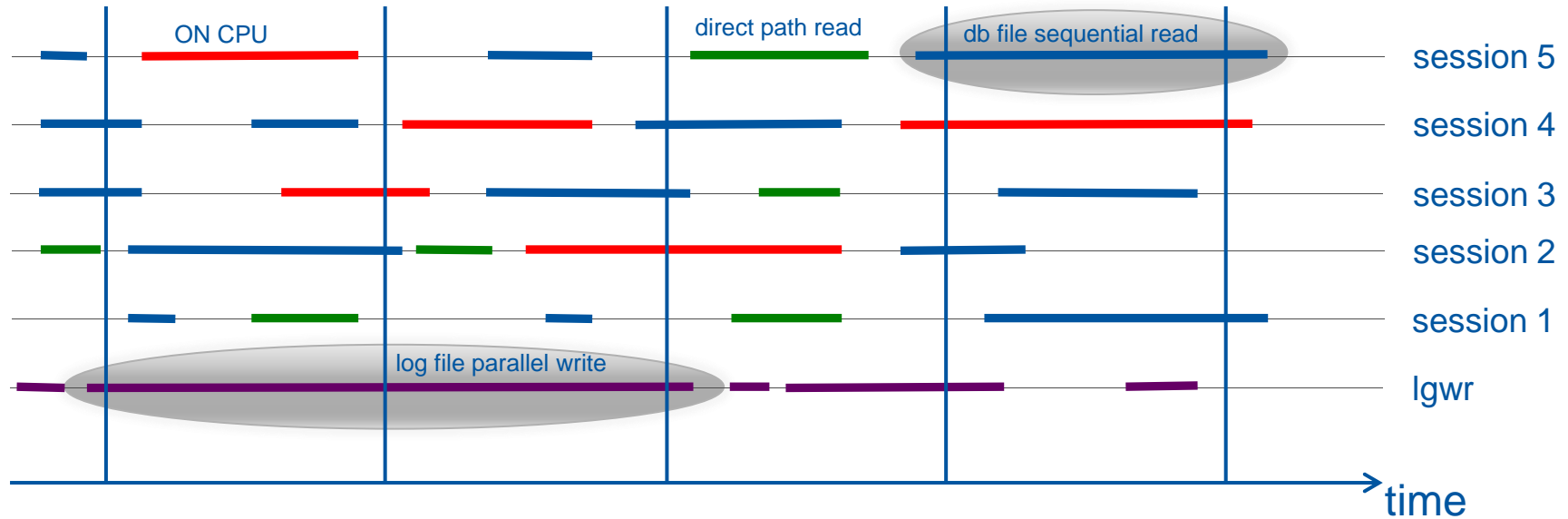
- Spinning disk latency  $O(10\text{ms})$ , Flash  $O(0.1\text{ms})$
- Is it always in this order of magnitude? If not, you should know, with detailed and timing(\*) information.
- Reasons include bugs, disk failures, temporary or global overload, etc.

(\*) correlation with other sources (OS logs, storage sub-system logs, ASM logs, etc.)

# Complementing AWR for IO

- AWR captures histogram information, not single IO
- AWR does not capture information on ADG
- In addition, desirable to extract information
  - Longer term (across migrations, upgrades, capacity planning)
  - With timing, slow IO operations from ASH

# Capture Active Session History long IO



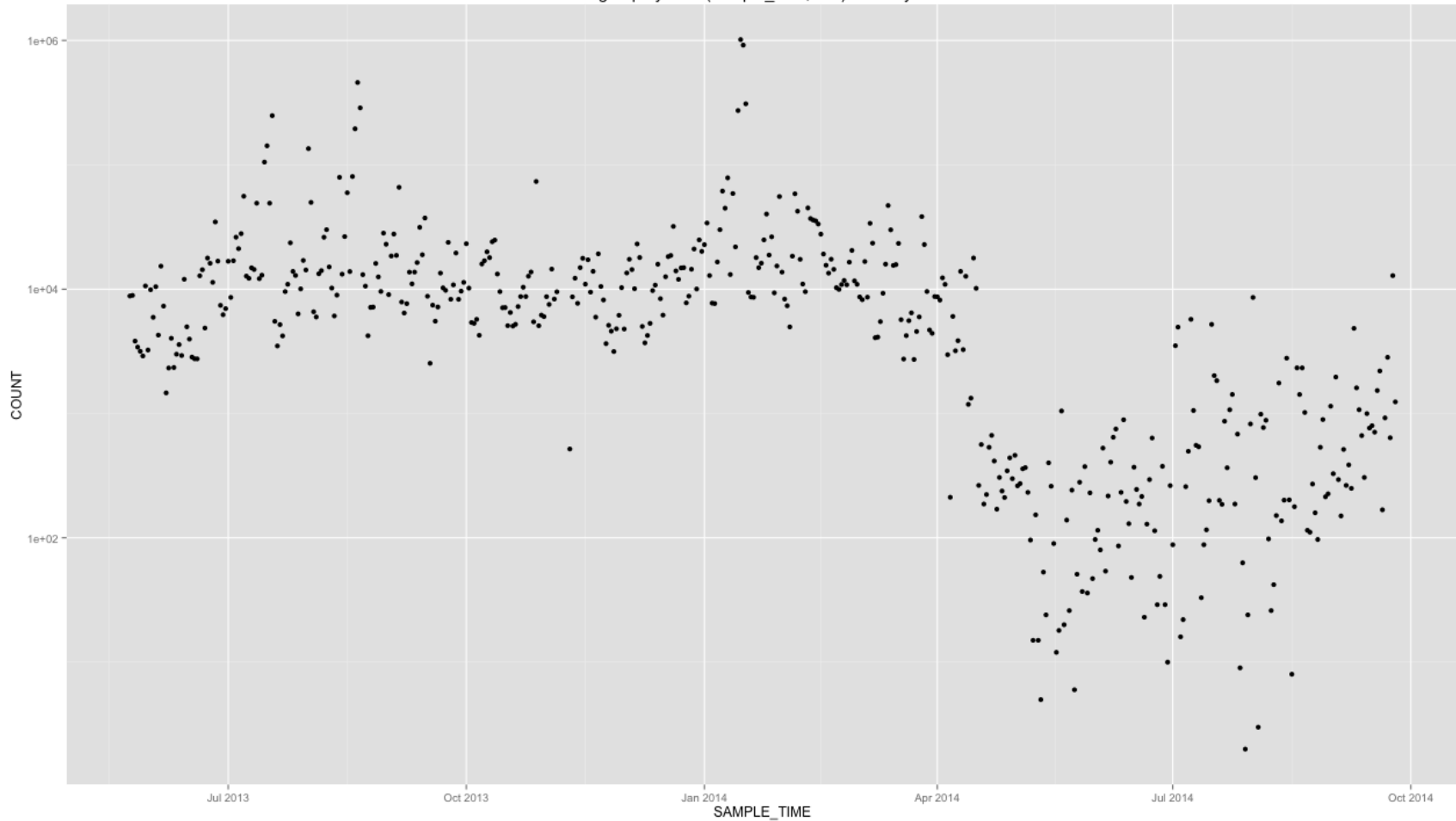
Relative wait times indicative...

# ASH long IO repository

- Stores information about long(\*) IO operations
- Query it to identify major issues
- Correlate with histograms and total IO operations

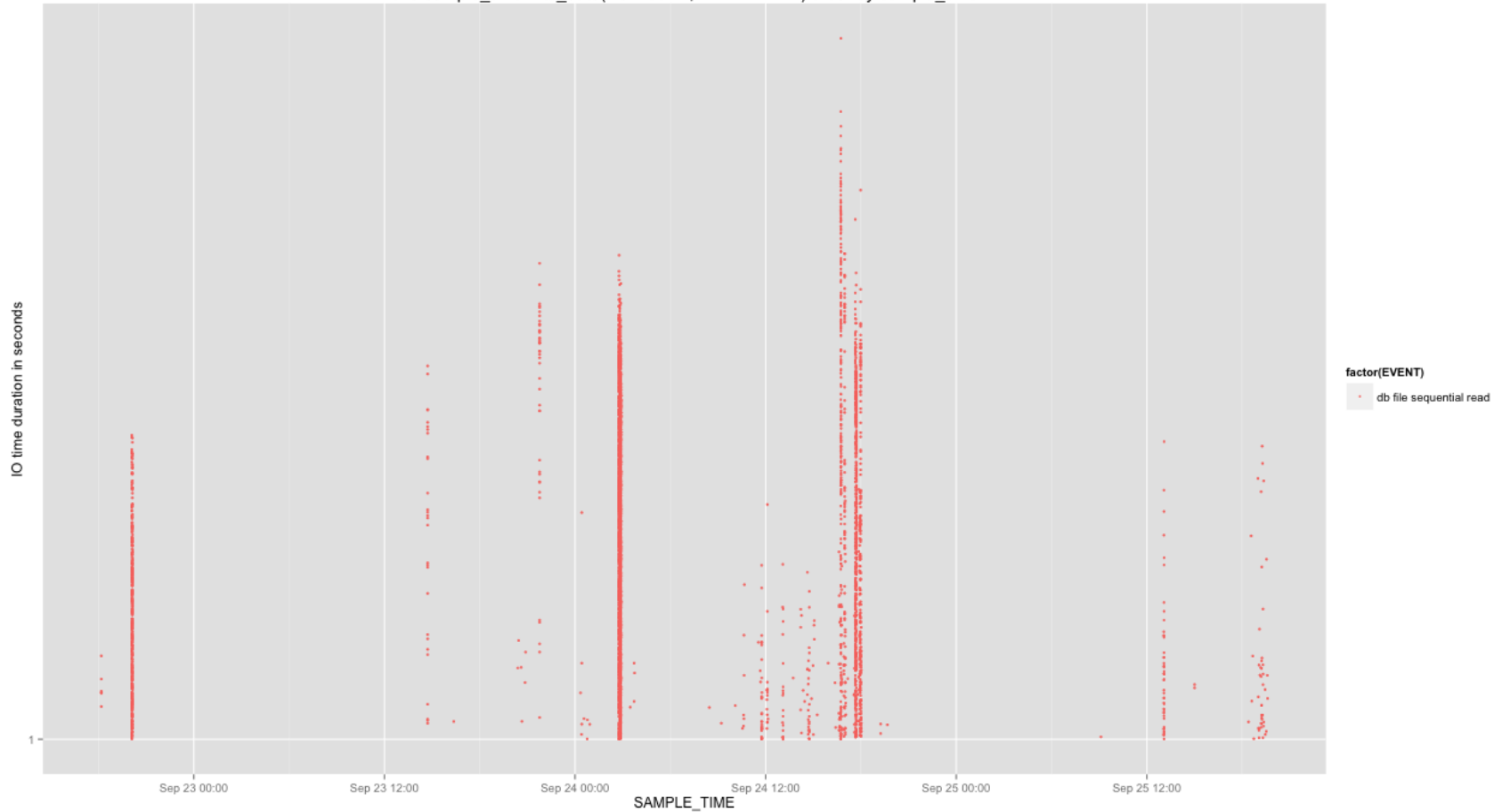
(\*) longer than expected, >1s, >100ms, >10ms

```
A1, Number of IO operations taking more than 1 second (aggregated per day):
select trunc(SAMPLE_time,'DD') sample_time, count(*) count from cerndb_ash_longwait where db_unique_name in ('ADCR_RAC50','adcr')
and sample_time>sysdate-3000 and time_waited>1000000 and open_mode='READ WRITE'
group by trunc(sample_time,'DD') order by 1
```

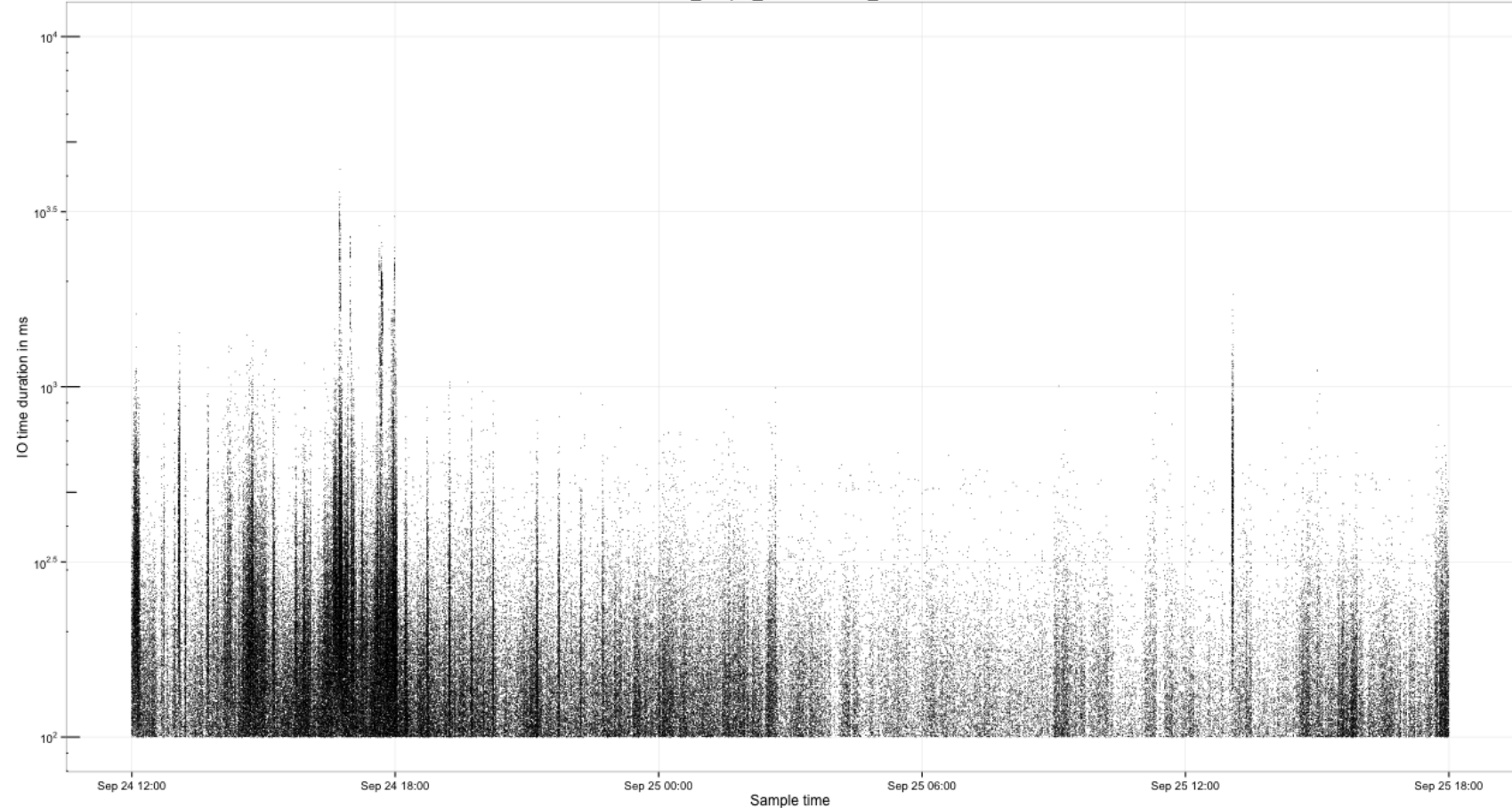




```
A2, IO latency seen from the database, 8890 IO operations:  
select SAMPLE_time, event, time_waited/1000000 time_waited  
from cerndb_ash_longwait where db_unique_name='ADCR_RAC50'  
and event='db file sequential read'  
and time_waited>1000000  
and sample_time > to_date('20140922','YYYYMMDD') order by sample_time
```



A3, IO latency seen from the database,  
IO operations taking more than 100ms ( 463010 IO operations )  
select sample\_time,event, time\_waited/1000 time\_waited from cerndb\_ash\_longwait  
where event ='db file sequential read'  
and sample\_time between to\_date('20140924-1200','YYYYMMDD-HH24MI') and to\_date('20140925-1800','YYYYMMDD-HH24MI')  
and db\_unique\_name='ADCR\_RAC50'





[www.cern.ch](http://www.cern.ch)