

Mean PB to Failure – Initial Results from a long-term study of disk storage patterns at the RACF

CHEP 2015 – Okinawa

Primary Author: Tony Wong (tony@bnl.gov)

Co-Authors: Costin Caramarcu, Chris Hollowell, Tejas Rao, William
Strecker-Kellogg, Alexandr Zaytsev

Brookhaven National Laboratory

Background

- RHIC-ATLAS Computing Facility (RACF) is a dedicated facility for the high-energy and nuclear physics community at Brookhaven National Lab (BNL).
- Tape + central storage usage at the RACF is currently ~53 PB.
- Additional, over 12,000 disk drives in Linux Farm cluster provide ~24 PB of raw storage capacity.

RACF Distributed Storage

- Nearly 9,000 drives with ~21 PB raw capacity used for RHIC distributed storage.
- Drives have varying capacities (0.5-4 TB), age (0-5 years) and brand (Seagate, Hitachi, Western Digital, etc).
- DDM via dCache (PHENIX) and xROOTd (STAR).
- Initial attempt to track long-term drive usage patterns at the RACF.
- This study inspired by Western Digital's presentation at **HEPIX Fall 2013** (see www.hepix.org) on MPTF as an alternative to MTF.

Methodology

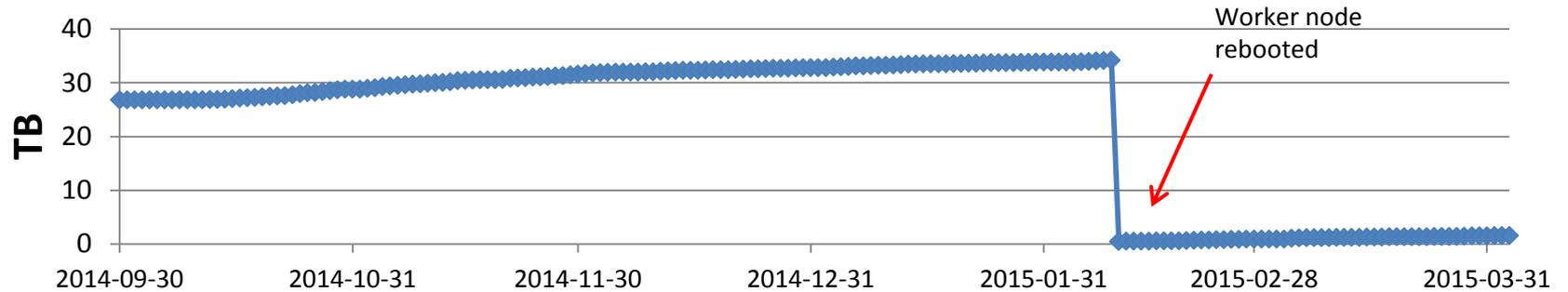
- Use SMART data to record activity (Reads and Writes) daily.
- Store historical data in MySQL (collected ~186 days of data so far).
- For manageability, record data from selected drives (10% of total) only – not all 12,000 drives.
- Different experiments (ATLAS, PHENIX and STAR) represented among selected drives for heterogeneity.
- Extrapolate usage over full year.
- Compare with manufacturer published specs (where available) and correlate to drive reliability.

US ATLAS

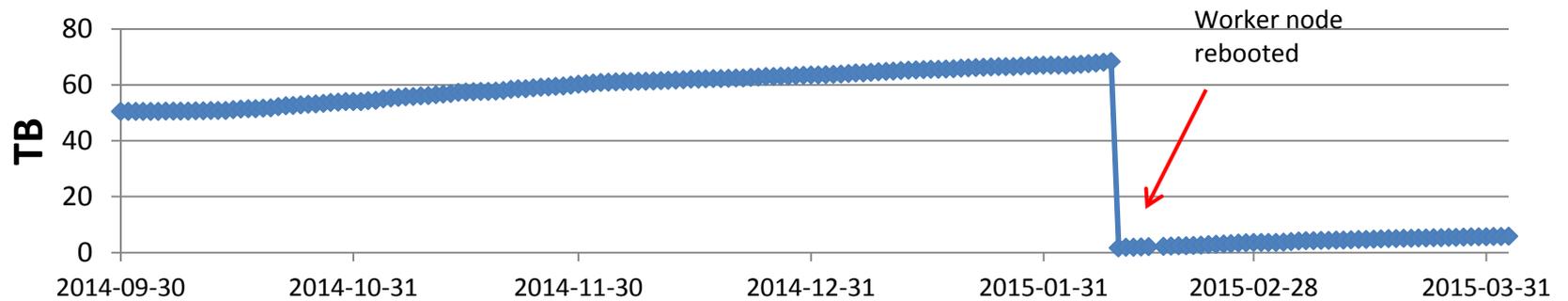
- Storage-light worker nodes.
- Storage used only as scratch space for jobs – no distributed storage on worker nodes.
- Little I/O required at the worker-node level.
- Average read is ≈ 0.05 TB/day (per drive).
- Average write is ≈ 0.13 TB/day (per drive).

ATLAS (acas1093)

Read



Write

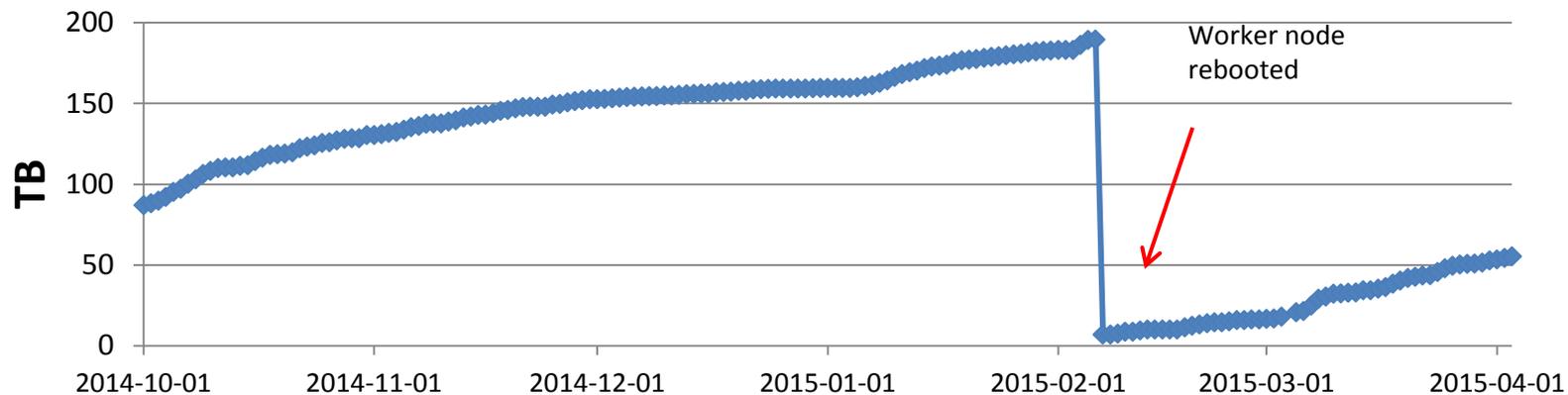


PHENIX

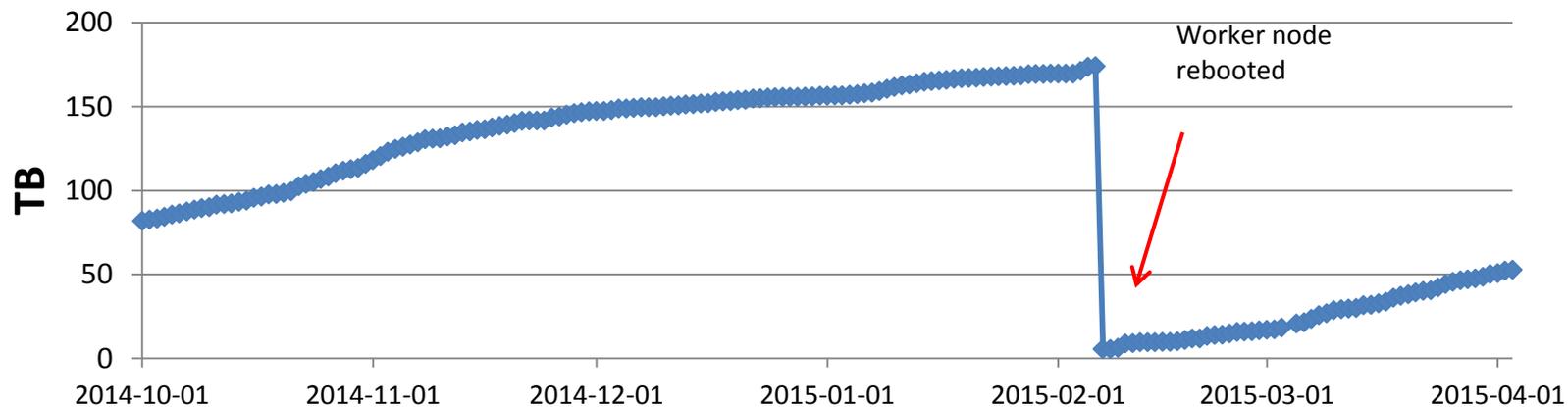
- Storage-heavy worker nodes.
- Distributed storage on worker nodes.
- Equally high Read-Write rates at the worker-node level.
- Average read is ≈ 0.82 TB/day (per drive).
- Average write is ≈ 0.75 TB/day (per drive).

PHENIX (rcas2714)

Read



Write

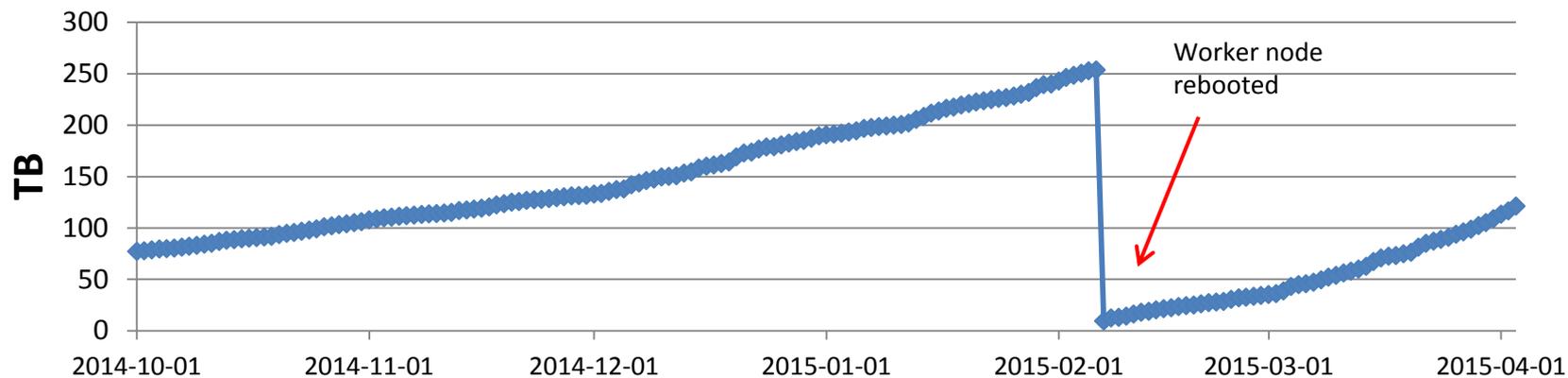


STAR

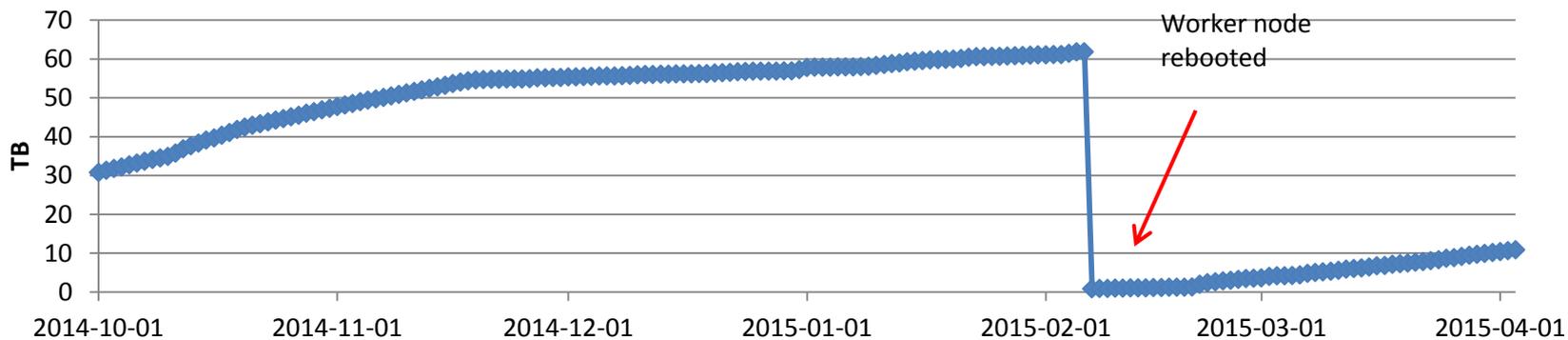
- Storage heavy worker nodes.
- Distributed storage on worker nodes.
- High Read and medium Write rates at the worker node level.
- Average read is ≈ 1.56 TB/day (per drive).
- Average write is ≈ 0.22 TB/day (per drive).

STAR (rcrs6229)

Read



Write



MPTF

- Western Digital published a paper on MTTF and MPTF in June 2013 (<http://www.wdc.com/wdproducts/library/other/2579-772003.pdf>).
 - Correlation between MPTF and HDD failure rate.
 - WD publishes MPTF as maximum “workload” (reads + writes) to achieve stated MTTF.
- Seagate also lists maximum workload to achieve stated MTTF (for example, see <http://www.seagate.com/www-content/product-content/enterprise-hdd-fam/enterprise-capacity-3-5-hdd/constellation-es-4/en-us/docs/100740544d.pdf>).

Workload at the RACF

- Yearly (linear) extrapolated usage (per drive):
 - ATLAS Read/Write (~ 17.2 TB/ ~ 45.8 TB)
 - PHENIX Read/Write (~ 300.0 TB/ ~ 275.6 TB)
 - STAR Read/Write (~ 567.8 TB/ ~ 81.4 TB)
- Define workload = Reads + Writes.
 - ATLAS ≈ 63.0 TB/yr
 - PHENIX ≈ 575.6 TB/yr
 - STAR ≈ 649.2 TB/yr

In Perspective

- RACF has Seagate and Western Digital Re drives:
 - Seagate published workload – 550 TB/yr
 - Re published workload – 550 TB/yr
 - Se published workload – 180 TB/yr
- Within these workloads, expected failure rates should be between 0.44% to 0.73%, depending on manufacturer.
- For example, see <http://www.wdc.com/en/products/internal/enterprise/> for more data on Western Digital drives.

Disk Failure Rates (Preliminary)

- Only includes hardware replaced due to unrecoverable failure.
- 2013-2014 data only.
- Annual failure rate (AFR) as a fraction of total disk pool in each cluster:
 - ATLAS $\approx 0.35\%$
 - PHENIX $\approx 0.94\%$ → highest Write workload
 - STAR $\approx 0.41\%$ → highest Read workload
- Contribution from other factors (thermal conditions, power quality, etc) folded in – not measured separately yet.

Summary

- Some RACF usage cases exceed specified maximum workload by a significant margin.
 - Should be a factor in disk failure rate.
 - Early data suggests higher disk mortality rates in heavy write usage (ie, PHENIX) environment when compared to other usage cases.
- Even then, AFR is acceptable – validation of worker node-based distributed storage model as **a cost-effective solution**.
- Long-term study will yield more precise results.
 - Need more data to cover entire cycle (data collection, reconstruction, analysis, etc) multiple times
 - Assuming Reads/Writes have the same impact on AFR—this may be incorrect

Acknowledgement

The authors would like to gratefully acknowledge useful discussions with Western Digital on the merits of the MPTF metric and its correlation to disk reliability.