SAND2019-9027D

Modeling Resilience Needs for Burst Buffers

Matthew L. Curry, 1423

Problem

- Burst Buffers (BBs) are necessary for checkpoint restart (CPR)
- BB data lifetimes are short, and data is easily regenerable – How much is CPR resilience worth?
- Previous work argues very low cost tolerance: < 2%
- This work addresses three concerns: *Generalizability, Applicability, Accuracy*

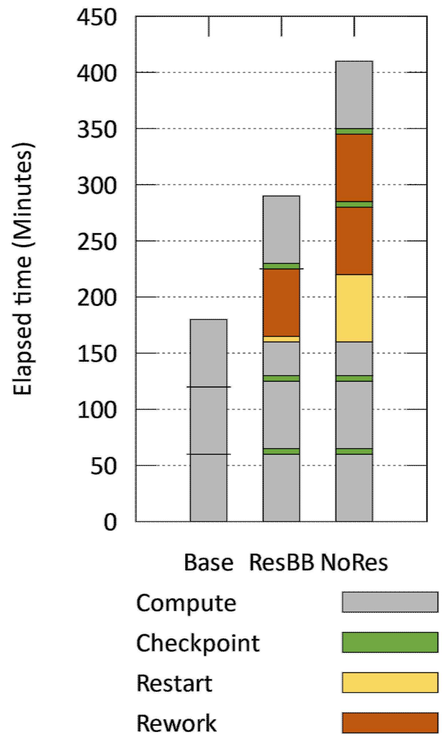

Figure 1: Illustration of resilience overhead in three architectures and situations

Applicability to Vanguard/Astra System Design

- 3x memory 240GB/s flash tier
- Large 30 GB/s disk tier
- Bleeding edge – Low spec'd MTTF for size

Approach

General Analytical Model

- Current models emphasize point design – BB as an exascale technology
- Recognize freedom in several dimensions: MTBF, size of flash tier, architecture decisions

Improved Fidelity

- Provide second-order estimates for system overheads: Rework, aggregate failure rates [Schroeder], and cost of resilient design

Optimized Use for Resilience Design

- Apply previous lessons in choosing checkpoint frequency [Daly] to optimize machine efficiency

Funding

- Exascale Computing Initiative – 0.5 FTE, FY19

Results

- Updated model assumptions show larger cost for nondurable storage. Major contributes include:
  - Component counts/failure modeling
  - Rework inaccuracies
  - Updated read cost assumptions
- Unreliable BBs have outsized impact on most efficient compute intervals
  - 25% longer compute interval with reliable burst buffer
  - Drain time is able to be increased
  - Increased reliability of burst buffer may decrease cost of capacity tier

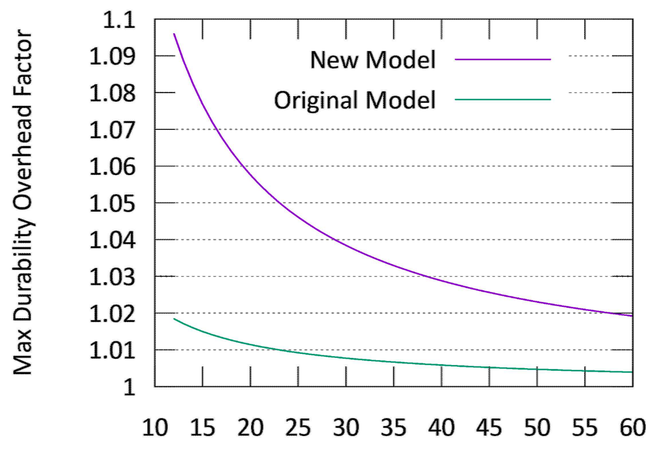

Figure 2: Improved model accuracy reveals more overhead associated with nondurable storage

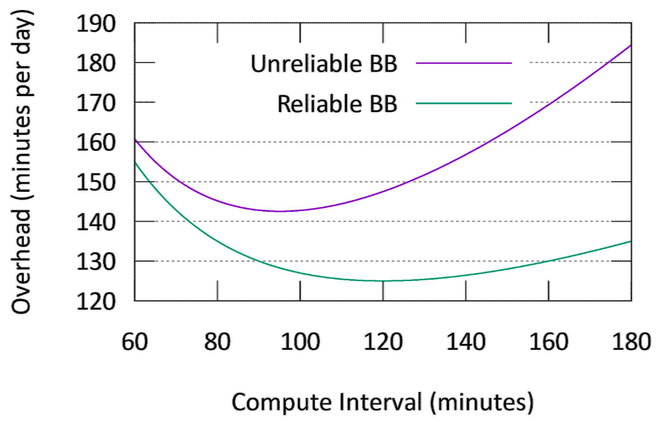

Figure 3: Extra efficiency with reliable burst buffers are available by tuning compute interval

Significance

- These results are useful for:
  - Codesign activities with storage vendors
  - System design activities in support of procurements
- Contributing an updated analytical model that is applicable to many system sizes
  - Analytical models are preferable because of understandability
- Even in the most unforgiving use case (i.e., 100% defensive I/O), durability of burst buffers is justifiable
  - This model is still very conservative; durability is likely even more important due to intra-device faults

References

- Bianca Schroeder, Raghav Lagisetty, and Arif Merchant. "Flash Reliability in Production: The Expected and the Unexpected." USENIX FAST 2016.
- John Bent, Brad Settlemyer, Nathan DeBardeleben, Sorin Faibish, Uday Gupta, Dennis Ting, Percy Tzelnic. "On the Non-Suitability of Non-Volatility." USENIX HotStorage 2015.
- John Daly. "A higher-order estimate of the optimum checkpoint interval for restart dumps." Future Generation Computer Systems, Vol 22. Issue 3, February 2006.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.
SAND No. _____

Don't forget the SAND No.