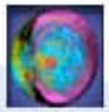
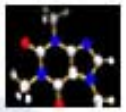
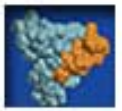
SciDAC
Scientific Discovery through Advanced Computing

Office of Science

Reliability/Resilience Panel

HEC File System & IO Workshop, Aug 3, 2010

Garth Gibson, Lin Xiao
Carnegie Mellon University

Carnegie Mellon
Parallel Data Laboratory

pdsi

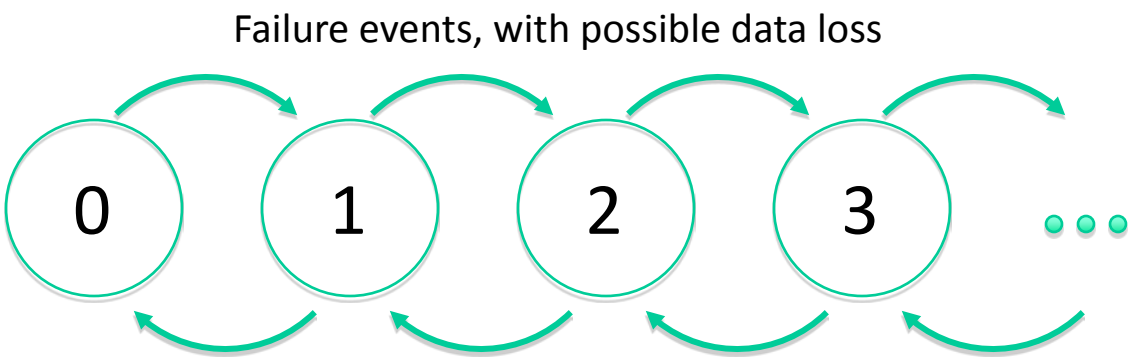
Metrics

- 1988: MTBF → MMTDL
- 2004 (Corbett): add RAID6 and a "correlated failures factor"
- Prob(DL < dt) α dt / MTTDL
  - Measurement equiv: Avg Data Loss Events / year
- But none of this is commonly used
- *Max simultaneous failures always tolerated* (N+i)
- Emphasize the illusion of perfection
- But probability of zero data loss approaches zero
  - Media errors, dependent failures, failed upgrades
  - A few lost files not the same as loss of millions or TBs
- Better, simple metrics: Avg Lost Bytes / year

$$MTTF_{RAID} = \frac{(MTTF_{Disk})^2}{(D+C*n_G)*(G+C-1)*MTTR}$$

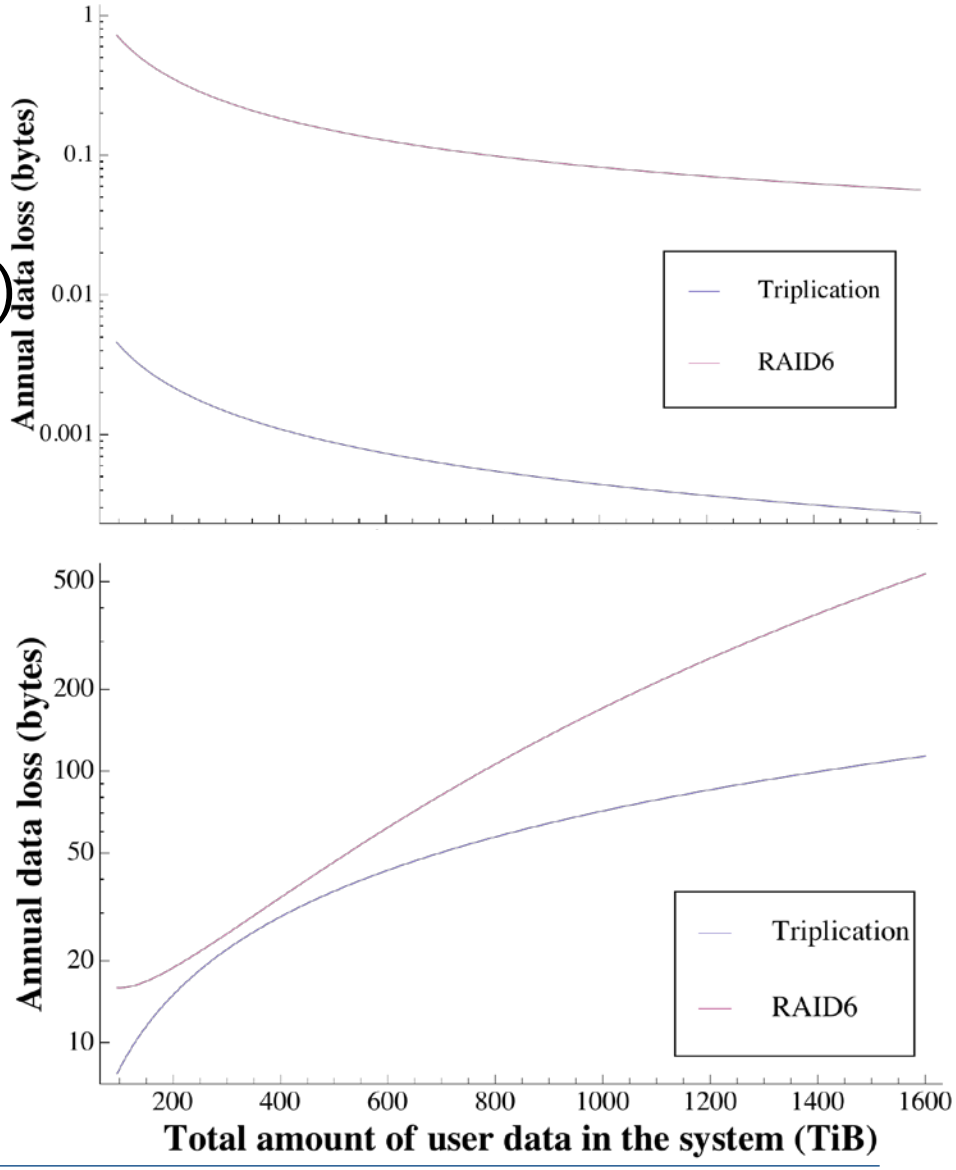
RAID 6 versus Triplication

- Model declustering / random replication
  - RAID 6 is cheaper, so triplication should be more reliable…
- Fundamental properties, not accurate distributions (ie., way too much Poisson), at least to start …
- Simple Renewal Reward Markov model
  - Each failure has an expected amount of lost data
  - Counting arguments for fraction of possible failures that lose data and how much



So How Big are Reliability Differences?

- Three copies more reliable
- Bigger systems less likely to have >2 blocks lost in any RAID set (8+2)/triple (3)
- Bigger systems repair in parallel faster (declustered)
- Triplication has ~3X disks for same user data, so ~3X faster failure & repair
- Delay in repairing is costly
- Assume .8TB/disk used, 64MB blocks, 1% AFR disk fail, exponential repair is either 12/N or 0.5+12/N (Markov model)



Closing

- The future is a world of failures everywhere
- We need tools to tame the raging beast
- Measure everything …. Know your rates
- Speak the unspeakable …. Loss happens
- Simple tools intuitive …. Educate always