



Los Alamos

NATIONAL LABORATORY

————— EST. 1943 —————



STILTS

Or why LANL doesn't use HSM



Brad Settlemyer

May 22, 2019



Why won't LANL use HSM

- The problem is NOT that we don't believe HSM claims
 - We believe HSMs can move the data
 - We believe HSMs can tolerate failures
 - Maybe HSMs could even be batch scheduler aware
- The problem is what policies could we reasonably set?!
 - Input decks – keep it forever, but also modify it over the course of a campaign
 - Checkpoints – save some
 - Analysis Data – save all – for a while
 - Processed Data – save all – for a while
 - Movie files – save forever

What HSM policy would we implement?

Workflows (campaigns) come in approximately 3 types:

1. Wildly successful!
 - User runs a series of test calculations that show the large scale calculation will succeed, runs the large calculation and succeeds
2. Successful with modifications
 - User runs test calculations that show large calculation will succeed, but large calculation surprises user. User modifies small calculation with this feedback, modifies large calculation and after a few false starts finishes successfully
3. Failure
 - Same as above, but the user eventually decides the large scale calculation can't succeed on the available compute platform (e.g. not enough RAM)

The spectrum of HSM

- We believe there are scientific workflows that match HSM well
 - Streaming/experimental/observational data processing pipelines
 - Shorter scientific campaigns
- But there are workflows that don't match HSM well
 - PI driven science (chasing a hypothesis every which way they can)
- Risk! Consequences!
 - Run out of tapes
 - Waste 6.5 months of calculation
 - Surprising users is a problem

Goal of STILTS

- **Make our slower, less agile storage tiers easier to use**
 - Designed to protect data
 - Designed to grow capacity and performance over time
- **Buy smaller, faster agile/bursty tiers**
 - Designed for tens of thousands of mounts
 - Absorb ugly workloads efficiently
- **Minimize data movement, accelerate scientific workflows**

Use the batch scheduler to drive data movement

- STILTS-CS (Short-Term, Intermediate, and Long-Term Scaffolding for Campaign Storage)
 - Implemented as a SLURM plugin
 - Enable stage-in/stage-out from arbitrary file systems
 - Leverage LANL's parallel file movement tools
 - Do not create backups!