Spectrum Scale Migrations and Challenges at St. Jude

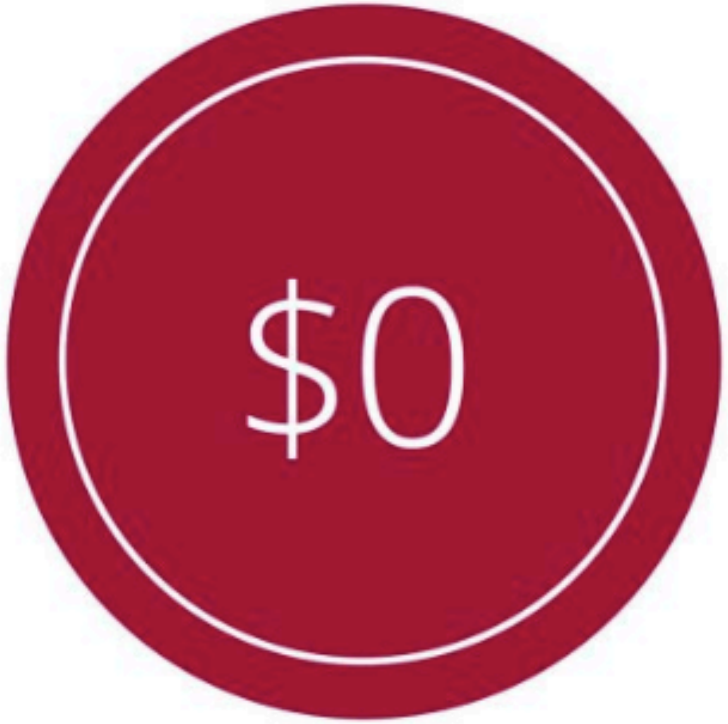
Will Schmied, Storage Architect

St. Jude Children's Research Hospital®

Opinions expressed are solely my own and do not express the views or opinions of my employer.

St. Jude 101

Cost to family: $0

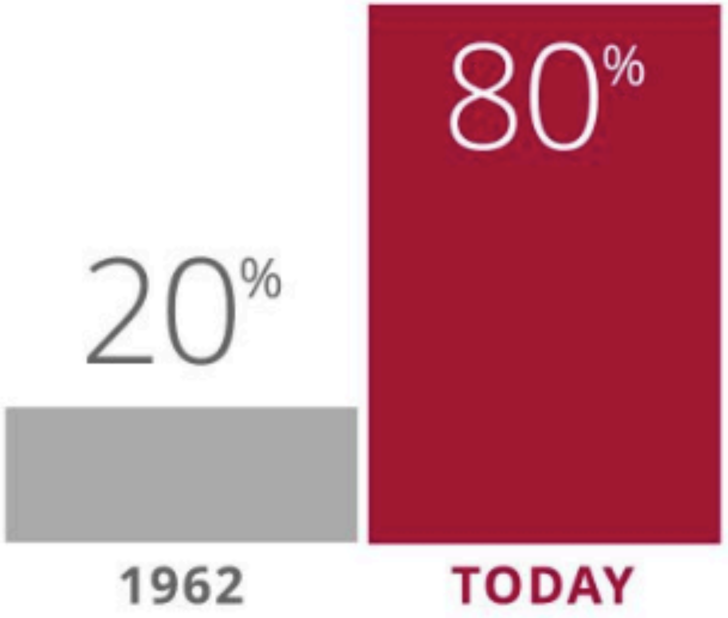


Families never receive a bill from St. Jude for treatment, travel, housing or food – because all a family should worry about is helping their child live.

Saving kids worldwide

St. Jude has treated children from all 50 states and **from around the world**.

Our goal: 100% survival



Treatments invented at St. Jude have helped push the overall **childhood cancer survival rate from 20% to more than 80%** since it opened more than 50 years ago.

A Brief History of Spectrum Scale at St. Jude

- St. Jude has used GPFS / Spectrum Scale since 2011.
  - Before that, various Linux “roll your own” based solutions.
- Currently on our third generation of SS clusters.
  - The fourth generation is in the procurement phase currently.
- Hardware currently single vendor (DDN).
  - Have used IBM and “roll your own” previously.
  - Supporting systems (for LSF, Bright) from Dell/EMC.
- I’ve been working with SS since May 2013.
  - Only person at St. Jude managing SS since February 2016.
  - *We have an open position. (Talk to me after!)*

1: In the beginning, there was SoNAS

- In the beginning, there was SoNAS.
  - Something more robust and ”expandable” was needed: St. Jude was doing real HPC compute work now!
  - The options in 2011 were few.
  - We joked that we had 4% of all installed SoNAS systems, since we had 4 installed on campus.
  - “S” was ~2 PB, “E” was ~3 PB. Also a replica system and an IBM development system on campus.
  - TSM HSM to LTO5 tape for survival due to lack of disk capacity.
  - SoNAS was OK (software). The S2A 6620 (storage) was not.

2: Native HPC was needed

- By early 2013, we knew we needed native HPC, not HPC over NFS.
  - We had experienced several significant downtimes due to S2A 6620 storage.
  - We were not sold on the ESS at that time, so we built our own clusters on top of DCS 3700 storage.
  - These would serve very specific computational workflows for the *Pediatric Cancer Genome Project.*
  - ”R” was ~1.9 PB, “C” was ~750 TB.
  - Over time, the native clients really showed how much work we could potentially get done with a properly designed cluster.
  - Combined data and metadata was a productivity killer.
  - AFM (IW) used for replication to home site (SoNAS “S”). This would result in data loss multiple times.

3: We need more power (and space)!

- By early 2015, the limitations of our DCS 3700 filesystems were very apparent.
  - Metadata on its own pool needed to be addressed.
  - We also wanted to get out of managing server/storage bare-metal provisioning.
- Enter the GRIDScaler.
  - These were to replace the “research” clusters, first and second generation.
  - Two of the SoNAS were still on site (“S” & “E”).
  - These GRIDScaler systems were to be an HPC specific cluster and a NAS specific cluster.

The GRIDScaler Ecosystem

- We currently have seven GRIDScaler clusters today.
  - SFA 7700, 12K, 14K, 7990, 18K.
  - All are running SS 4.2.3 (more on that later).
  - Largest cluster is ~22.5 PB (37 servers, 297 clients).
  - Total is ~48 PB, 90 servers, 383 native clients, ~4,500 protocols clients.
- Fourth generation cluster is coming early 2020, codenamed “Jude”.
  - Replace and consolidate two largest third generation clusters.
- Smaller SS 4 clusters will still have much ”capital lifetime” left as of 9/30/2020.
  - These will be upgraded in place to SS 5* for mainstream support.
  - More forklift upgrades in the future to get a clean SS 5 filesystem.

Why so many migrations?

- We find ourselves migrating often:
  - To meet research compute needs as data sets grow exponentially.
  - Due to undesirable conditions or restrictions within current system.
- Some reasons for forklift migrations:
  - SoNAS → anything else (restricted, closed environment)
  - V3.5 → v4.1 cleanly (no desire to update problematic source cluster)
  - V4.2 → v5.x cleanly (ensure full v5 functionality to prevent future constraints)
- Ironically, our goal since “generation 2” has been to upgrade the filesystem in place and rotate backing storage via NSD disk migration.
  - We have yet to accomplish this goal.

Migration ≠ Fun

- SoNAS → "Roll our own" (3.5 → 3.5): AFM. Not good.
  - We made it worse operating in IW mode after that!
- "Roll our own" → GRIDScaler (3.5 → 4.1): AFM for first system migration.
  - Bad: Required mmfsck later to fix corruption from migration.
  - Good: Cutover for compute nodes was a small window of time.
  - *By second system migration, parallel rsync was the only accepted method.*
- SoNAS → GRIDScaler (3.5 → 4.1): Many parallel rsync. Not great, not horrible.
  - Bad: Slow due to large, deep and complex folder structures. Customer communication and scheduling takes even longer in most cases though.
  - Good: Zero data consistency or availability issues.
  - Good: Overall, a nice experience (in other words: NO COMPLAINTS).

St. Jude Children's Research Hospital

What's next for migration?

- v4.2 → v5.x: ADA (Mira / Data Flow) with supplemental rsync in special cases.
  - Good: Easy to manage like rsync, but much faster to transfer data.
  - Good: Choose between keeping source system ACLs or writing new ACLs on destination via inheritance (we have use cases for both options).
  - Good: No additional extended attributes added (AFM).
  - Good: No worries about managing gateway node queues and memory (AFM).
  - Mixed: Compute node cutover still must occur at some point.
  - Mixed: Early migration data sets mounted on source cluster over NFS.
  - Bottom line: It is easy to configure, use and diagnose, with no undocumented features to configure.

St. Jude Children's Research Hospital

Questions?