

Architecture of a Noisy-Neighbor-Proof HPC Storage Solution

Ellis Wilson, Ph.D.
Software Architect

Neighbors Can (and Should) be Assets

In day-to-day life neighbors:

- Improve home safety
- Provide social opportunities for you and your family
- Allow for localized stores, activities, schools
- Can organize to provide a voice in local government

In HPC storage neighbors:

- Allow for enterprise-grade usability and features
- Provide far larger device pools, which can (with the right filesystem architecture):
 - Increase streaming bandwidth, and data and metadata IOPS
 - Reduce downtime in the face of failure
 - Improve capability for workloads to remain entirely in-core

Filesystem Neighbors: All Shapes and Sizes

Plan for diversity, or plan for failure

High BW Streaming Ingest

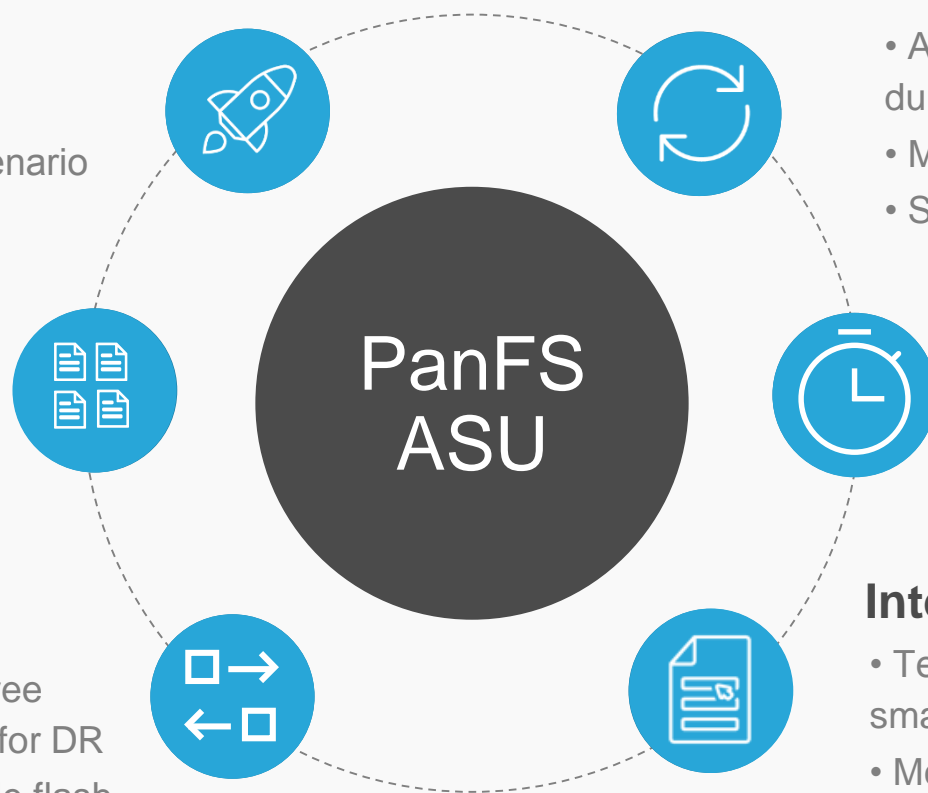
- All initial or periodic data ingest
- Medium or Large data
- Going directly to HDD still ideal scenario

Many-File

- Some apps treat FS like a database
- Small and not infrequently empty files
- Emerging memory + Flash SSD ideal

Disaster Recovery

- Administrator-driven snapshots + tree walks can be expensive/interruptive for DR
- Isolate to mostly metadata on NVMe flash
- Replicate large/small file changes from either HDD or SSD



High BW Streaming (Re-)Read

- AI has a high number of these workloads during training phases; back-testing similar
- Medium or Large data
- Serving directly from HDD still ideal scenario

Latency-Dominated

- Can be data or metadata latency-dominant
- Very sensitive to conflict with others
- Emerging memory + SSDs ideal

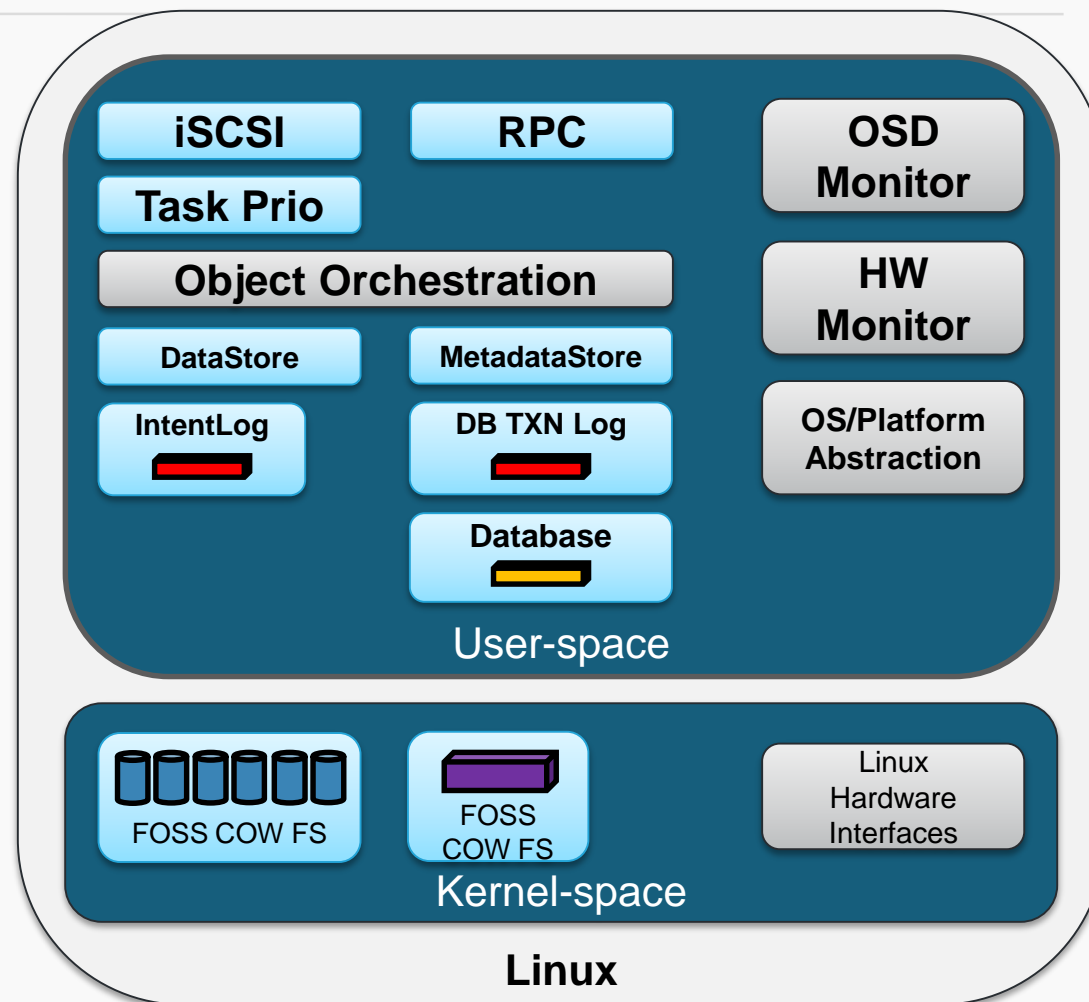
Interactive

- Tends to be metadata-dominated, with spikes of small file and latency-dominant workflow
- Most sensitive to conflict with others
- Only way to achieve: make good neighbors out of the other workloads

Your Friendly Neighborhood Ultra

Designed and built for real-world mixed workloads

- **Streaming Workloads:**
 - Served by array of HDDs; accelerate via NV-DIMM (W) + DRAM (R)
- **Many-File, Latency-Sensitive:**
 - Keep consistent via metadata in DB on NVMe SSD and small files on SATA SSD; accelerate via NV-DIMM
- **DR, Interactive**
 - Accelerate and normalize majority of workload (metadata) via DB/NVMe, and segment workloads to HDD/SSD based on file size.



PANASAS®

Questions?