

Sebastian Oeste

Monthly Storage Talks - Metrics for procurement

Dresden, April 2024

Motivation - What I want to discuss with you today

- Procurement of storage systems is difficult!
- Storage system is critical component! Can stop the whole cluster.
 - A lot of side-effects with other cluster components.
 - What are meaningful metrics for a storage procurement?
 - What are useful feature requirements for storage systems in HPC?
 - How to design meaningful acceptance tests?
- Sharing experiences in the procurement process.
- Discuss what was good what could be better in the future.

Procurement Process - In three steps

A tedious and bureaucratic process.

1. **Leistungsverzeichnis** - Definition of requirements
2. Vendor offers - maybe additional refinement phase
3. Supplement - Installation - **Acceptance phase**

Performance Metrics

Define the performance of a file system. Typical metrics are:

- Bandwidth - throughput
- IOP/s - rate of create, stat, read, write, ... operations
- Easy to define on paper but not really meaningful.
- Hard to get a real estimation of systems performance in production.
- Acceptance happen on empty cluster with vendor tuning.
- Wrong estimates can prevent offers.

Benchmarks

- IO-500¹ - Benchmarks suite with defined I/O pattern
- elbencho²
- fio³ - best single node benchmark Tool but not helpful for HPC file systems.
- Requires deep understanding of the benchmarks to prevent hacks for best results.
- E.g. `mdtest_easy_write` - zero byte writes could be cached with features like DoM.

¹<https://io500.org/>

²<https://github.com/breuner/elbencho>

³<https://fio.readthedocs.io/en/latest/>

Applications

- Identify real world applications from production.
- Define expected values of I/O metrics on these applications.
- Recommended method for systems with homogenous workloads.
- Hard to identify for systems with a lot of small- or mid-sized jobs.

Features

- Interfaces - POSIX, GDS, S3
- Redundancy / Data safety: RAID, Backup, Erasure Coding, Snapshots
- Availability - HA features failover of components
- Security - Encryption
- Administration - Monitoring, QoS, Quotas

Features

- More tied to an actual product than performance metrics.
- Hard to test during acceptance phase.
- Need to define what happens if the vendor changes the roadmap (feature get kicked / is discontinued)

What if?

Other things to test for ...

- Outages of network components - a hiccup in the network should not stop the file system.
- Outages of single storage components e.g. failure of a MDS or OSS.
- Writing very large files or filling up the file system to a certain level by one job.

Lessons Learned Acceptance Test

- Lower demands on the peak performance of the system.
- Try to define a mix of workloads that need to run in a certain time (time-to-solution as metric!)
- Outage test are fun and provide interesting insides.
- Define a clear feature set.