Pool Sizing

NDGF Manager
Mattias Wadenstein
<maswan@ndgf.org>

2019-04-09
NDGF All Hands
Ljubljana, Slovenia

NORDIC E-INFRASTRUCTURE COLLABORATION

Overview

- Current status
- Data
- Plans

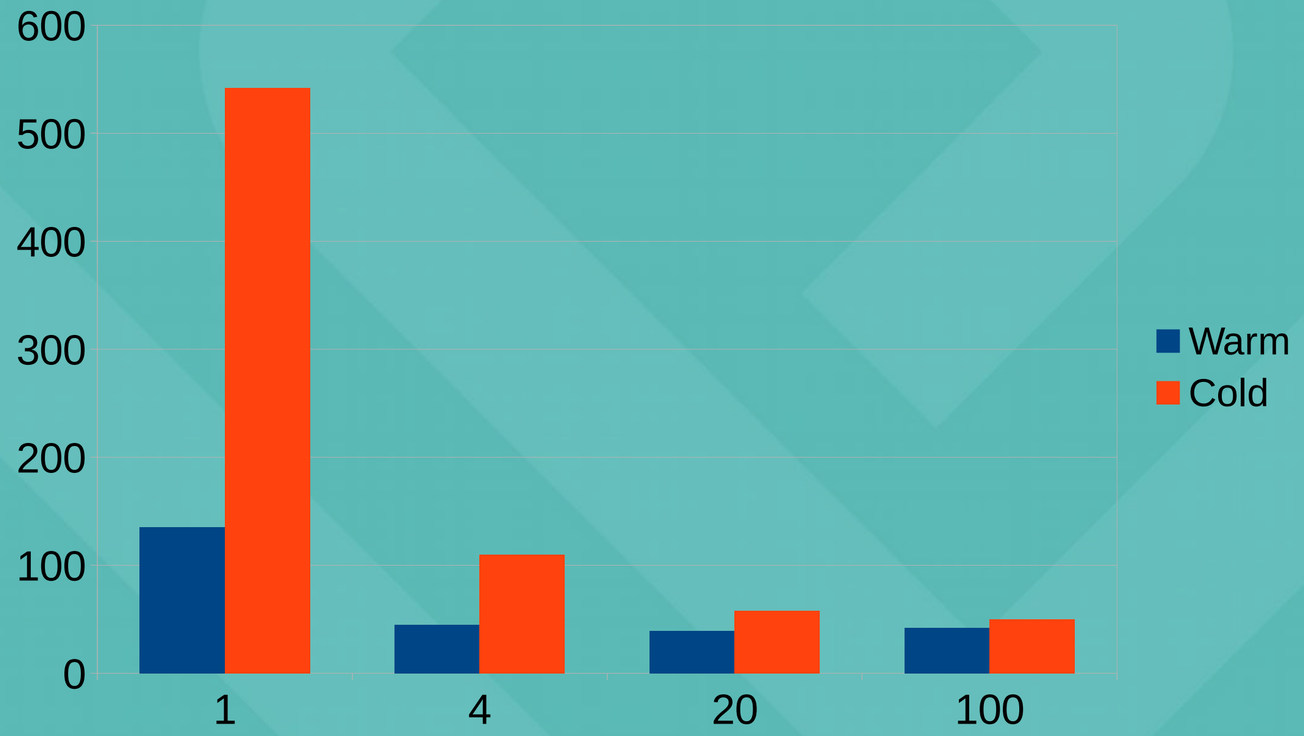
Current status

- One filesystem per raidset
- One pool per filesystem
- One or several pools/filesystems/raidsets per node

- Rationale
  - Pools should be independent, load on one should have minimal effect on others
  - Pool startup should depend on IOPS throughput in the filesystem for total number of stat(), not files per pool

Data

- The last assumption is not supported by data
  - Average pool startup time (in seconds) for ~850k files split over 1/4/20/100 pools



Future plans

- With tarpool scheme pool operators don’t need to care about how many pools NT1 ops create per fs
- Larger number of pools might increase RAM/TB requirements
  - If we do moderate number of pools, we should be able to keep within current recommendations
- Disks keep growing
  - 14TB today, 16/18TB probably this year
- Servers keep growing
  - Dell 740xd2, 24 disks on 2U

Questions?

NORDIC E-INFRASTRUCTURE COLLABORATION