



DKRZ

DEUTSCHES
KLIMARECHENZENTRUM

Lustre usages and experiences

at German Climate Computing Centre in Hamburg

Carsten Beyer

About DKRZ

- High Performance Computing Center
 - Exclusively for the German Climate Research
 - Limited Company, non-profit
- Staff: ~ 70
- Services for Climate Research:
 - Support for Scientific Computing and Simulation, Model Optimization, Parallelization
 - Data Management and Archiving
 - Data Visualization (3D Graphics and Video)
- University Research Group: HPC (Prof. Dr. Ludwig)

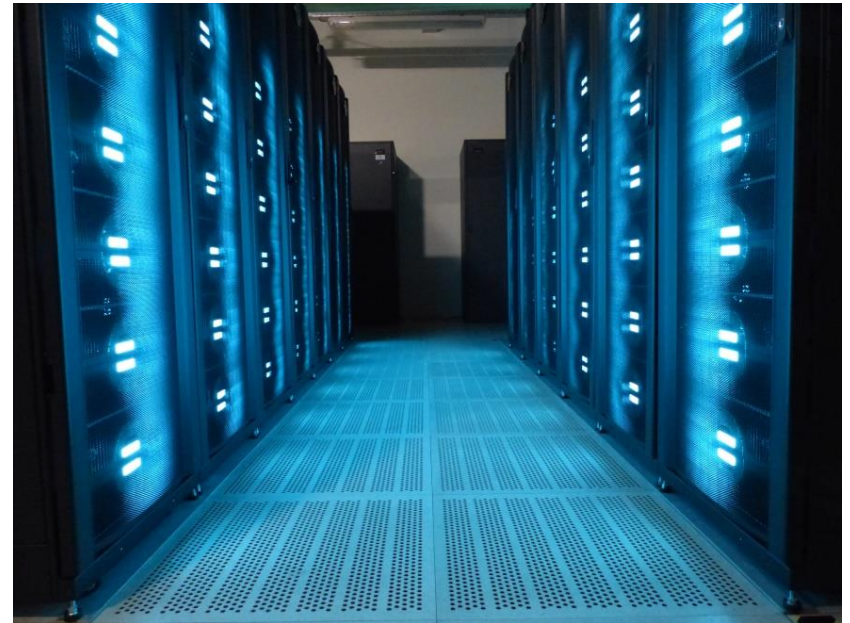
Mistral

- First phase 2015 (**second phase 2016**) , total cost: 41 Mio Euro
- Bull Supercomputer: 26 Mio Euro
 - Bullx B700 DLC-System
 - ~37.000 (**+ ~67.000**) cores (Intel Haswell/Intel Broadwell)
 - 1.550 nodes (2x 12 Cores) (**+1750 nodes 2x 18 Cores**)
 - 1,4 (**3,0**) PetaFLOPS
 - 115 TB (**266 TB**) main memory
 - Infiniband FDR
- Parallel file system:
 - Lustre, 21 (**+33**) PetaByte
 - Throughput > 0.5 TeraByte/s



Lustre - ClusterStor

- Seagate ClusterStor Setup (Phase 1 - CS9000/ **Phase 2 - L300**)
 - 62 OSS / 124 OST's / 6TB disks
 - 5 MDT
 - 21 PB / max. 6 Billion files
 - Lustre 2.5.1 / IB FDR
 - 455 million files / 90% filled
- **74 OSS / 148 OST's / 8TB disks**
- **7 MDT**
- **33 PB / max. 8 Billion files**
- **Lustre 2.5.1 / IB FDR**
- **196 million files / 40% filled**



Filesystem structure

➤ Lustre Phase 1

- HOME directories (/mnt/lustre01/pf => MDT0000)
- POOL directory (/mnt/lustre01/pool => MDT0000)
- Software tree (/mnt/lustre01/sw => MDT0000)
- SCRATCH dirs (/mnt/lustre01/scratch => MDT000[1-4])
- WORK-directories (/mnt/lustre01/work => MDT000[1-4])

➤ Lustre Phase 2

- WORK directories (/mnt/lustre02/work => MDT000[0-6])
 - extension of phase 1
 - Soft link /work -> /mnt/lustre01/work
 - Soft link /mnt/lustre01/work/<prj> -> /mnt/lustre02/work/<prj>

Migration GPFS to Lustre

- Copying 4.5 PB from GPFS (AIX) to Lustre (Linux)
 - Usage of GPFS policies to generate filelist (130 mio files)
 - Sort and split filelist (mix of big/medium/small files) as input for ,rsync`
 - Using SLURM on new system to schedule ~6000 Jobs for copying via IB gateway from NSD server

- Challenge
 - Not overloading NSD server on previous system (2x 10 GbE per node)
 - Changing UID/GID for some users from id's <1000 to >20000 during transfer (newer rsync version needed than in RHEL 6)
 - Time frame to achieve this
 - Because Benchmarks for approval of the system still ongoing

Migration GPFS to Lustre

- Where's my Quota and why takes 'mv' so long
 - All copied project data belonging to MDT0000 in a separate location
 - The ,new' project directories were distributed to MDT000[1-4]
 - How to copy the old data to the new directories ?
 - Could not be moved easily with ,mv' due to DNE phase 1
 - Trying tools like ,shift' for copying
- No quota could be used as before (user/fileset/user in fileset)
 - HOME/SCRATCH/WORK in one filesystem now
 - Using Robinhood for ,soft' quota (starting 2016)
 - User can see their ,quota' and amount of data on a DKRZ web frontend

Migration GPFS to Lustre

- Small files
 - HOME directories (~ 30 mio files / 6 TB)
 - Software tree
- Problem to run Backup with Calypso/Simpana
 - Full Backup of HOME takes days
- Large loading times for software (e.g. Matlab, Python)
 - Generating a 300GB image file stripped accross 16 OST's
 - Loop mount of this image file to interactive nodes (Login/Graphics)
 - Using caching on clients for higher performance

Tools used - Copytools

- Self-Healing Independent File Transfer (Shift) – Paul Kolano (NASA)
 - Dependent on 'mutil' (Paul Kolano / NASA)
 - Final rsync at the needed (Hardlinks)
- 'lfs find + rsync'
 - Generating filelist with 'lfs find' and split in equal parts
 - 'rsync' with split-lists (running as SLURM jobs in parallel)
 - Final rsync needed afterwards for hardlinks, directory ownership
- pftool – Los Alamos National Lab
 - Easy to use, scalable on several nodes with SLURM
 - Needs final 'rsync' afterwards for hardlinks, directory ownership

Tools used - Robinhood

- Report generation for
 - ‚User‘ reports -> HOME/SCRATCH
 - ‚Project‘ reports -> WORK
 - Total amount of data for each project (~ 220 active projects)
 - Per user in a project (up to 130 user per project)
 - Overview for project admin
 - ‚Soft‘ quota by mail to users / project admin
- Setup
 - Robinhood version 2.5.6-1 / reading Lustre changelogs
 - 2 Robinhood server for lustre01 (connected to 2/3 MDT)
 - 3 Robinhood server for lustre02 (connected to 2/2/3 MDT)

Tools used – HPSS pftp

- Tape Archive HPSS
 - Projects have to request storage on an annual basis (Quota)
 - Manual copy of data by user for their projects with ,pftp`
 - User can decide for single / double copy or long term archive (LTA)
 - Double copy is also accounted twice for project
 - LTA needs a description of data (to be identified later, stored up to 10 years)
- Currently no automatic migration from Lustre to HPSS
 - Possible tool could be Robinhood
 - Currently not tried yet
 - How to identify data to be migrated for projects

Monitoring – tools / sources

➤ Sources:

- Starting with own python daemon on CS (replaced by Seastream)
- Seastream API on ClusterStor (currently switched off on lustre02)
- Entries from ‚syslog‘ (Lustre clients)
- Lustre Illite (from clients)

➤ Tools:

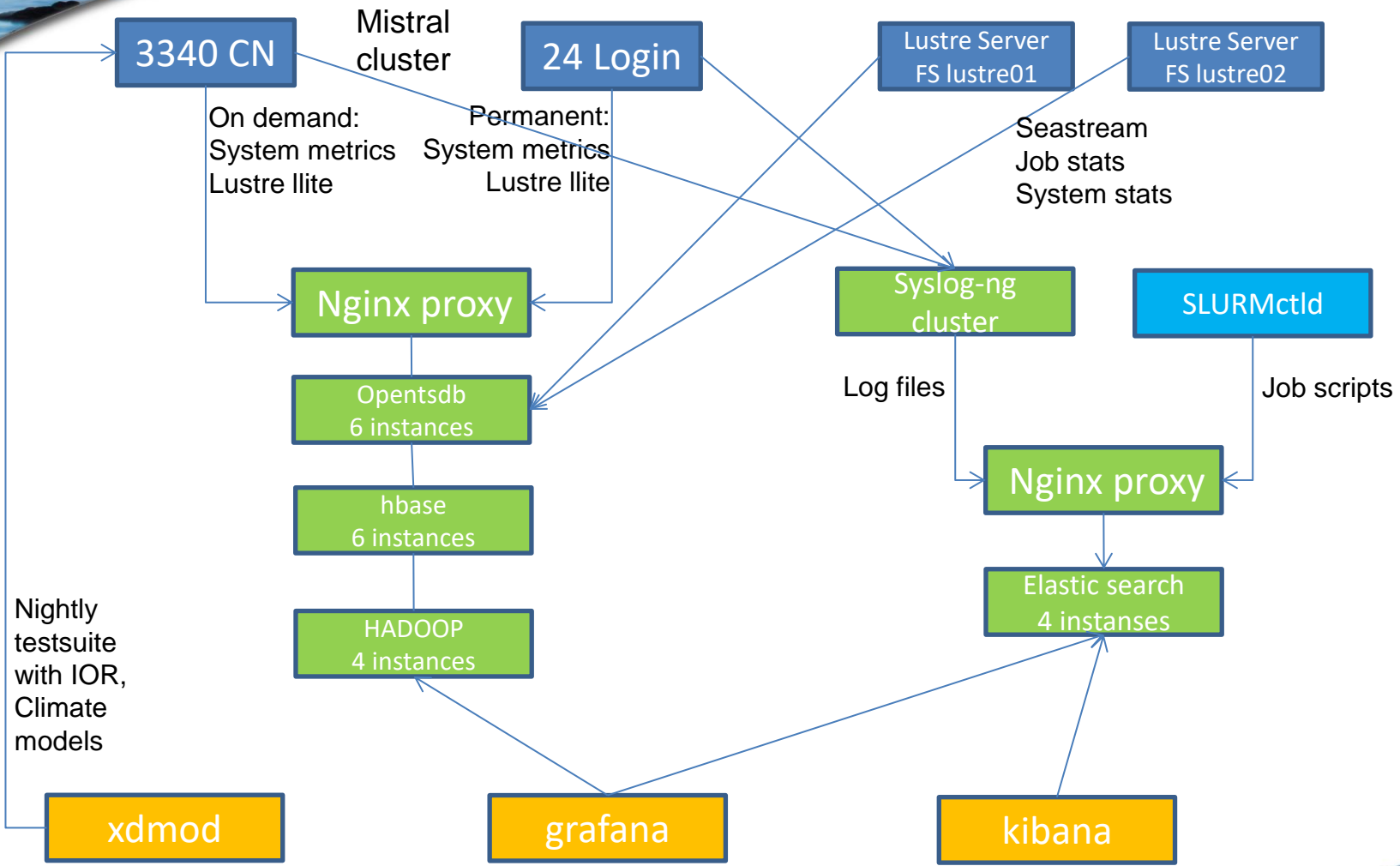
- Opentsdb / hbase / hadoop / elastic search

➤ Frontends:

- Xdmod / Grafana / Kibana
- Icinga on Clusterstor

Thanks to: Olaf Gellert, Hendrik Bockelmann, Josef Dvoracek (ATOS), Eugen Betke (Uni HH)

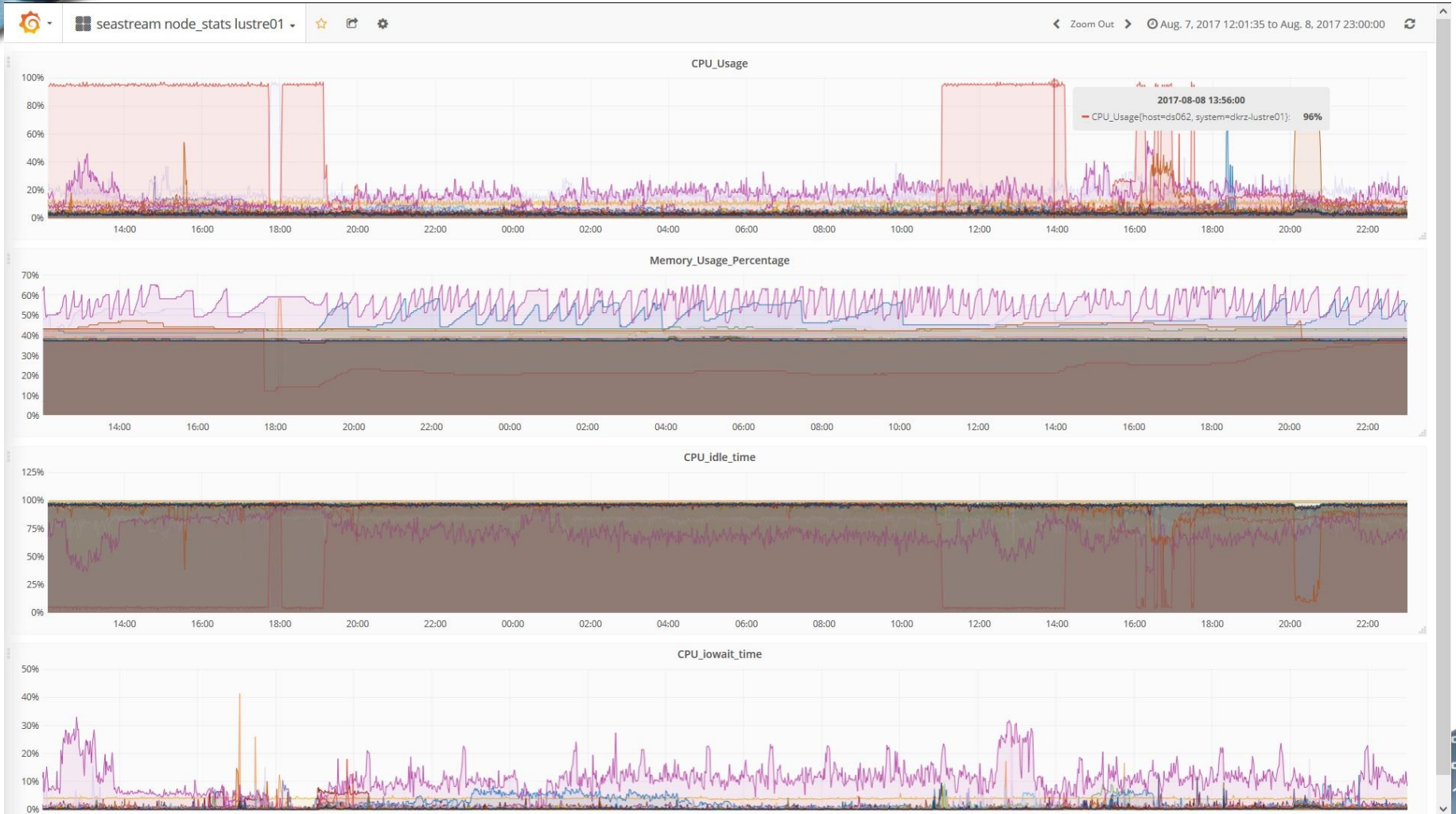
Monitoring - Overview



Monitoring - Grafana



Monitoring - Grafana



Pro's / Con's from daily business

- Most times fine ... but ...
- Degradation of performance when reaching >95% filling state and not symmetric usage of OST's (lustre01)
 - First solution was to deactivate OST'S / migrating data from OST's and shifting data to ,lustre02' filesystem
 - Newer solution is setting ,qos_threshold_rr' to 5% on all MDT
 - That was leading to about 1.2 PB of ,unusable' disk space
 - With the newer setting it is below 1 PB
- Degradation of performance by monthly RAID-check (lustre01)
 - Lowering the ,raid check priority'
 - RAID check runs now 2 1/2 weeks

Pro's / Con's from daily business

- Client dis-/reconnects from OST's / MDT's
 - Issue for SLURM jobs with IO (extended runtime)
 - Robinhood could not read changelog anymore
 - Workaround: Failover/reboot/failback affected MDT
- Firmware bug on hardware (lustre02)
 - Rare case of watchdog issue causes OSS HA pair to power down
 - After power up of OSS and fixing OST's reconnect from clients or amount of lustre02 not possible
 - Reboot of all Lustre clients needed (power reset)



Thank you for your Attention!

<http://www.dkrz.de>

Carsten Beyer

beyer@dkrz.de

Questions ?