REGIONALES HOCHSCHUL-
RECHENZENTRUM KAISERSLAUTERN

TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN

RHRK Information

High Performance Computing with the Cluster „Elwetritsch“

Fokus: largefiles on /scratch

Course instructor: Dr. Josef Schüle

/scratch

- Every User has his own /scratch/<userid>
- Environment Variable $SCRATCH

File system without quota (limits)
Temporary files

Though large - it lives from your cleaning

/scratch

- **Optimal bandwidth requires system policy:**
  - Most files are ~ 1 MB in size
- **Best usage:**
  - Medium sizes files
  - 100 KB - 50 GB
  - max. 1000 files per directory

/work

Help me - I have many more small files

- **10,000 files sized 100 kB or smaller**

-> email hotline@rhrk.uni-kl.de

will provide you "smallfile"region on filesystem /work.

largefiles

Help me - my files are much larger.
Files > 500 GB are causing real problems

But:
/scratch/userid/large_files

is ready to take those files (only those)

largefiles

Files in large_files/ contain opaque information
If your files are in large_files/ but they should not:

- **Don't move them - they keep the opaque information**
- **Copy them**
  - command line cp
  - tar cf - large_files/misplaced directory | tar xf - . to copy the directory "misplaced"

Same for files in /scratch which should be in large_files/:

- **Don't move - copy them**

IT-Gurus - How /scratch is working

file content and file properties (directory, etc, called META information) is separated

file content is split into chunks - normally 512K sized

- **files smaller then 512K occupy 512K**
- **files larger then 512K are split and put on max. 4 Servers**

IT-Gurus - How /scratch is working

If a file has 1,024 TB, each Server has to write $2x10^8/4$ chunks, that is in total 256 GB

Writing $5x10^7$ times to a server takes time,

- **2GB/s Bandwidth -> 128 sec**
- **10 us per write -> 500 sec**

These 4 Servers are unbalanced loaded (amount of data and no. of requests)

IT-Gurus - How /scratch is working

If a file is located in large_files

- **chunk size is 1M**
- **16 servers are used to write the files**

If used for 100K sized file - 1M is used, nothing won, just a lot space spilled.

IT-Gurus - How /scratch is working

If a file in large_file has 1,024 TB, each Server has to write $2x10^8/16$ chunks, that is in total 64 GB

Writing $1.2x10^7$ times to a server takes time,

- **2GB/s Bandwidth -> 32 sec**
- **10 us per write -> 120 sec**

Instead of 628 sec now 152 sec

Space used for file is equally spread - no unbalance

Meta data contains chunk information

Moving files changes only directory entry, but not the chunks

- High Performance Computing on Elwetritsch
- Largefiles

Vielen Dank
Thank You