Advanced User Group

Meeting

2018-06-25

Issue: random failing jobs

Cluster has two networks:

- Fast IB network only for jobs
- 10G network for admin, control

Sango and Slurm controller, and the newer nodes,
are on different 10G segments, connected via a gateway

Problem: This connection intermittently fails

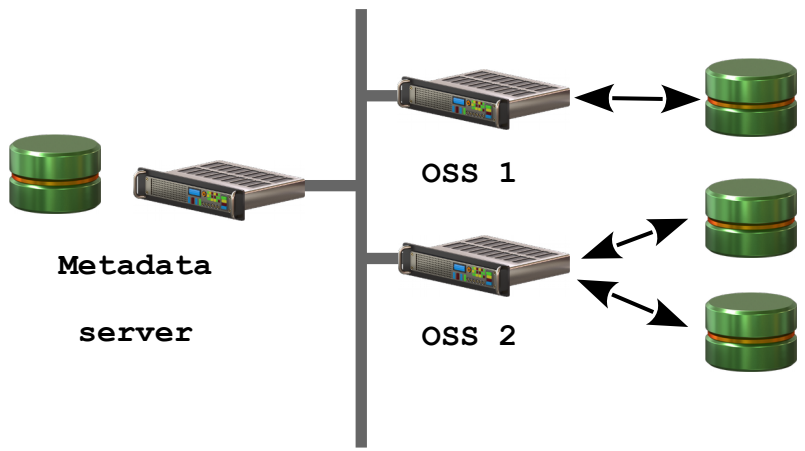
Solution: Move all nodes into same segment

Timeframe: August:ish

Issue: slow file operations

- Create and write or delete many files: heavy load on MDS
- Read or write a file from many clients: heavy load on OSS
- MDS and OSS interact
- When either is overloaded, requests queue up and file ops get delayed

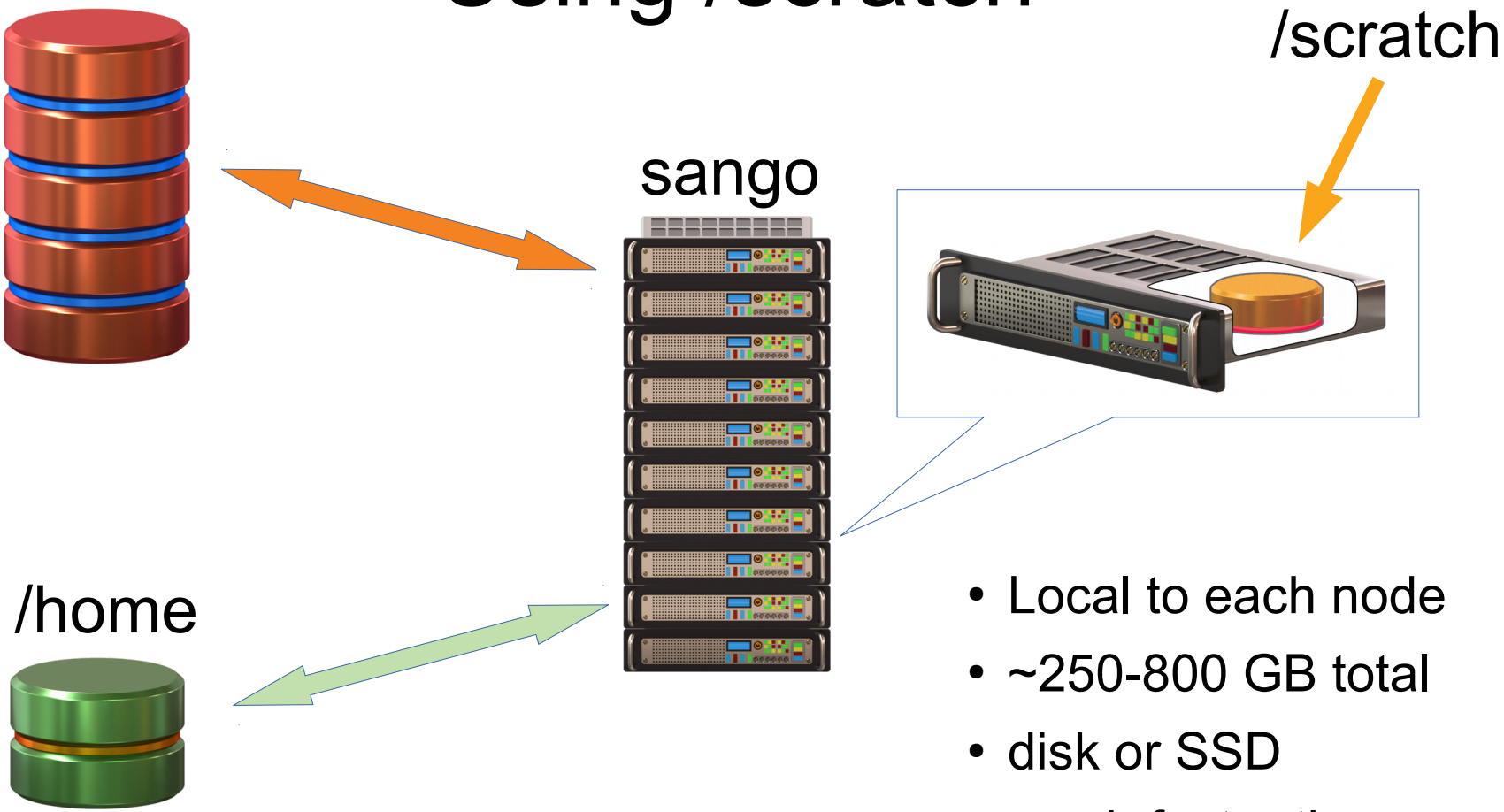
Solution: Avoid the above two situations.



OSS: **O**bject **S**torage **S**erver
read and write data

MDS: **M**eta-**D**ata **S**erver
map file name, attributes to data

Using /scratch



- Local to each node
- ~250-800 GB total
- disk or SSD
- much faster than any network storage

Use case:

Many processes reading a large input file:

- At start, copy file to each nodes /scratch
- Run computation on the /scratch copy
- Effective even if you only read a few times

Each process creates small intermediate files

- create, write on /scratch on each node
- At end of job: collate results, copy back to /work
- **Note:** only if your job does not need global access to intermediate data

Copy files

```
#!/bin/bash
#SBATCH -N 4
#SBATCH -c 16
#SBATCH -n 32
# directory name “/scratch/username.job_id”
d=/scratch/$SLURM_JOB_USER.$SLURM_JOB_ID

# Create temporary directory on each node.
# Processes run one per node if the numbers match
srun -n $SLURM_NNODES mkdir $d

# Copy data
srun -n $SLURM_NNODES cp -r /work/Unit/datafolder $d
# or even
sbcast /work/Unit/myfile $d

# delete everything at the end
srun -n $SLURM_NNODES rm -rf $d
```