Comprehensive Burst Buffer Evaluation

Eugen Betke, Julian Kunkel

Research Group
German Climate Computing Center
2017-11-12

Objectives

- Understanding how burst buffers can be used in an alternative way
  - Burst buffers are mainly used for catching I/O peaks
- Improving runtime of I/O intensive application by better workflows
- Reducing procurement costs by intelligent usage of burst buffers

Test systems and evaluation tools

Test systems

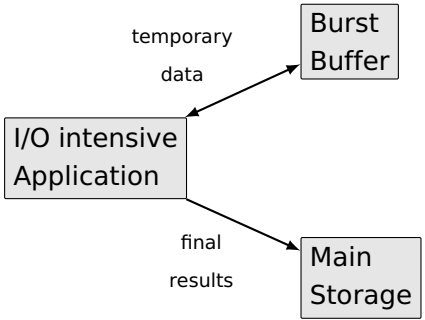
- Kove XPD [3]
  - In-memory storage
- DDN IME [5]
  - SSD-based
- Cray DataWarp [2]
  - SSD-based

Parallel I/O benchmark tools

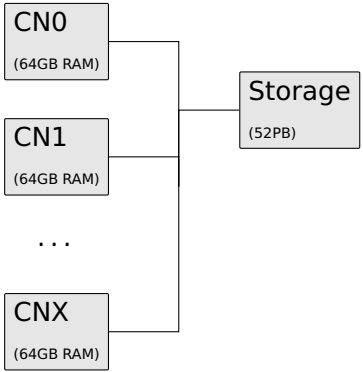
- NetCDF-Bench [4]
  - is a parallel NetCDF benchmark
  - generates I/O load to a shared NetCDF file
  - mimics scientific data
    - Many climate scientist favor NetCDF to other formats, because it offers powerful features and has a simple interface.
- IOR
  - uses MPI-IO interface in our tests
  - generates I/O load to individual files in order to get best I/O performance

Short-term campaign storage space

- Purpose
  - Reduction of I/O load on main storage
- Basic idea
  - Storing temporary data on main storage may be inefficient when
    - Temporary data is stored on burst buffer
    - Results are stored on main storage
- Expectation
  - Speed up of I/O intensive applications
- Evaluation methodology
  - Gathering of burst buffer characteristics
- Goal
  - Intelligent and efficient workflows



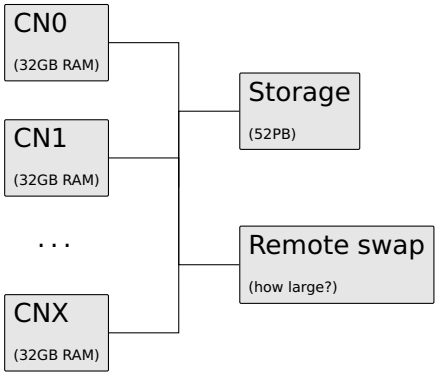
Reducing procurement costs of HPCs [1]



Observations made on Mistral [1] (HPC of DKRZ)

- Most applications are using only a fraction of available memory
- A few memory intensive applications have high memory requirements

Reducing procurement costs of HPCs [2]



- Purpose
  - Reducing total HPC costs
- Basic idea
  - Equip compute nodes with less memory
  - For memory intensive application use remote swap file system
- Expectation
  - Most programs are not affected
  - Memory intensive application are affected by swap overhead
- Evaluation methodology
  - Tracing of swap in/out with kprobes
- Goal
  - Cost model

References

- **HLRE-3 "Mistral"**. https://www.dkrz.de/Klimarechner/hpc. Accessed on 2017-03-22.
- Cray Inc. **Cray XC40 DataWarp's applications I/O accelerator**. Cray Inc. Cray Inc. 901 Fifth Avenue, Suite 1000 Seattle, WA 98164, Oct. 2015.
- Kove. **Kove XPD**. http://kove.net/downloads/Kove-XPD-L3-datasheet.pdf. Accessed on 2017-08-24. 2017.
- **NetCDF-Bench**. https://github.com/joobog/netcdf-bench. Accessed on 2017-08-25.
- DDN Storage. **Burst buffer & beyond; I/O & Application Acceleration Technology**. DDN Storage. Sept. 2015.