Using Sierra Burst Buffers

Sierra Roadshow, SNL & LANL

December 4-5, 2018

Elsa Gonsiorowski
HPC I/O Specialist

LLNL-PRES-763458

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Hardware Specifications

Each compute node has 1 burst buffer device

- 1.6 TB NVMeoF
- 5840 MB/s Sequential Read
- 2100 MB/s Sequential Write
- 1000 KIOPS Random Read
- 140 KIOPS Random Write

LLNL-PRES-763458

Burst Buffer Capability

NVMeoF: Non-Volitial Memory over Fabrics

- Asynchronous data transfer to/from GPFS servers
- *Decoupled from Compute Node usage*
  - pre-stage: move data to compute nodes
  - during-job transfers to/from GPFS
  - post-stage: move data to GPFS

LLNL-PRES-763458

Burst Buffer Access

- LSF: integration to allocate & de-allocate devices
- bbcmd: command line interface
- BBAPI: C-API for transfers from node-local file system
- BSCFS: C-API for interaction through a FUSE file system supports shared files across nodes

LLNL-PRES-763458

Burst Buffer Access

- LSF: integration to allocate & de-allocate devices
- bbcmd: command line interface
- BBAPI: C-API for transfers from node-local file system
- BSCFS: C-API for interaction through a FUSE file system supports shared files across nodes

- SCR: Scheduler integration and C library for codes with existing checkpointing capabilities.

LLNL-PRES-763458

Burst Buffer Access

- LSF: integration to allocate & de-allocate devices
- bbcmd: command line interface
- BBAPI: C-API for transfers from node-local file system
- BSCFS: C-API for interaction through a FUSE file system supports shared files across nodes

- SCR: Scheduler integration and C library for codes with existing checkpointing capabilities.

More tools coming soon

LLNL-PRES-763458

All Documentation

lc.llnl.gov/confluence/display/SIERRA/Burst+Buffers

LLNL-PRES-763458

Basic Use Case

1. Allocate with LFS
   `bsub -stage "storage=2"`
2. XFS file system available for use on compute nodes
   `$BBPATH`
3. Initiate transfers via `bbcmd`, `bbAPI`, and/or `bscfs`.

LLNL-PRES-763458

Basic Use Case

1. Allocate with LFS
   `bsub -stage "storage=2"`
2. XFS file system available for use on compute nodes
   `$BBPATH`
3. Initiate transfers via `bbcmd`, `bbAPI`, and/or `bscfs`.

Details on Sierra confluence

LLNL-PRES-763458

LSF Options

Sub-options to `bsub -stage`

`storage=`	allocated space on each SSD, in GB
`in=`	full path to stage-in script
`out=,`	full paths to 2 stage-out scripts

LLNL-PRES-763458

LSF Options

```
bsub -stage "storage=2:in=/in.sh:out=/out1.sh,/out2.sh"
```

```
#!/bin/bash
#BSUB -stage "storage=2:in=/in.sh:out=/out1.sh,/out2.sh"
```


LLNL-PRES-763458

Basic Example

```
cp /opt/ibm/bb/test/test_basic_xfer.c ./
mpicc test_basic_xfer.c -o test_basic_xfer    \
    -I/opt/ibm -L/opt/ibm/bb/lib              \
    -Wl,--rpath=/opt/ibm/bb/lib -lbbAPI
```


LLNL-PRES-763458

Documentation and Help

Documentation available on Sierra Systems Confluence

`lc.llnl.gov/confluence/display/SIERRA/Burst+Buffers`

Elsa Gonsiorowski <gonsiorowski1@llnl.gov>

Disclaimer
This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.