OpenSFS Test Cluster Donation

Dr. Nihat Altiparmak, Assistant Professor
Computer Engineering & Computer Science Department
University of Louisville
Donation Details

- Jan 5, 2018
  - OpenSFS announced call for proposals for donation of their Lustre test cluster

- Feb 9, 2018
  - University of Louisville submitted a proposal

- March 8, 2018
  - University of Louisville has been selected as the recipient of the OpenSFS Test Cluster equipment

- Now we are here to describe our plan with it

- THANK YOU OpenSFS!
Cluster Description

● 32-node cluster
  ○ 4 MDS nodes, each with
    ■ 2 x 2.4 GHz 4-core CPU
    ■ 64 GB memory
    ■ 4 x 100 GB SSD and 8 x 2 TB HDD storage
  ○ 4 OSS nodes, each with
    ■ 2 x 2.4 GHz 4-core CPU
    ■ 64 GB memory
    ■ 2 x 100 GB SSD storage and 22 x 2 TB HDD storage
  ○ 24 compute nodes, each with
    ■ 1 x 2.4 GHz 4-core CPU
    ■ 32 GB memory

● Including necessary networking, management, and infrastructure equipment
Our Research Group

● Computer Systems Laboratory (CSL) @ UofL
  ○ http://cecs.louisville.edu/csl/

● One faculty, two PhD students, two Masters students, and four undergraduate students

● Contemporary topics being investigated:
  ○ High Performance Parallel I/O
    ■ [TC '16], [MASCOTS '14], [TOS '13], [CLUSTER '12], [ICPP '12], [TPDS '12]
  ○ Self-Optimizing Storage Systems
    ■ [HiPC '16]
  ○ Energy-efficient Data Retrieval and Task Placement
    ■ [CLUSTER '17], [BDCAT '17], [MASCOTS '16]
Our Plan with the Cluster

Main Goals:

- Promote innovation and adoption of open-source scalable storage technologies
- Generate practitioners and researchers with expertise in storage systems
Our Plan with the Cluster

● Research Goals:
  ○ Developing novel, high-performance, and energy-aware distributed:
    ■ data placement,
    ■ data reorganization, and
    ■ data retrieval strategies
  considering the internal characteristics of
    ■ SATA HDD,
    ■ SATA SSD, and
    ■ NVMe SSD devices.
  ○ Collaborating with other groups and running real world applications on the cluster:
    ■ RNA-Seq pipeline, tumor growth simulation, etc.
Our Plan with the Cluster

- **Research Goals:**
  - NVMe-aware caching, tiering, buffering, and I/O scheduling techniques, and their effect on system’s overall energy consumption
    - MDS and OSS nodes will be upgraded with NVMe SSD devices to enable NVMe-related research
    - Power meters will be used to measure the energy consumption
  - Host managed (aka. Software-defined or “user programmable”) open channel and/or multi-stream SSDs will be investigated for better performance through increased internal parallel I/O and reduced garbage collection
Our Plan with the Cluster

- Teaching/Training Goals:
  - The infrastructure will be integrated into Operating Systems (undergraduate), Distributed Systems (graduate), and Storage Systems (independent study) courses
  - Course projects will be assigned and students will be provided access to the cluster to implement their projects
Research and Data Dissemination

- All developed software and tools will be released under free and open source GNU GPLv3 license
- Teaching and outreach material will be released under Creative Commons (CC) license allowing free adoption, modification, and redistribution
- Produced I/O traces will be released under SNIA IOTTA license allowing free redistribution
- Publications will be released under the copyrights of the publishers
- All materials will be shared publicly through the webpage of the Computer Systems Lab at:
  - http://cecs.louisville.edu/csL
Thank You!

Questions?
Related CSL Publications


