Ceph at the University of Maine System

Stephen Cousins
Forrest Flagg
Advanced Computing Group
University of Maine System
244 Neville Hall
Orono, Maine 04469
steve.cousins@maine.edu
raymond.flagg@maine.edu
Who am I?

- I have lived in Maine most of my life: born on the coast in Castine, grew up in Orono, work in Orono.
- Studied Math, Psychology and then Computer Science in Graduate School at UMaine
- 28 years as a computer professional
- 21 years working at University of Maine System
- 13 years working as HPC user with the Ocean Modeling Group at UMaine
- 12 years as HPC Administrator
Maine

- Became a state in 1820
- Famous for:
  - Lobster
  - Stephen King
  - Cold Snowy Winters
  - Beautiful Summers
- Infamous for:
  - Mosquitoes
What we’re missing in Maine right now:
University of Maine System (UMS)

- 7 Universities
- ~35,000 students
- Orono Campus (UMaine)
  - Flagship (main) campus
  - Started in 1865
  - about 12,000 students
  - Primary Research campus
  - One of only 11 Land, Sea and Space grant institutions in the country
- Hockey national champions in 1993 and 1999
ACG HPC System

- Infiniband QDR, FDR-10 and FDR HPC Infiniband network fabric: 40/56 Gbps
- 2012: Penguin 512 core AMD system (powered down)
- 2015: SGI 528 core Intel system
- 2016 – 2018: Supermicro 2752 core Intel Haswell/Broadwell/Skylake
- 2016: 3 nodes with 2 x nVidia K80 GPUs
- 2 Visualization systems with remote 3D capabilities
- 225 TB Lustre file system (not backed up)
- 1.1 PB NAS storage (ZFS served over NFS with IB)
- 1.3 PB NAS backup ZFS storage in Portland using 10Gbps Ethernet
UMS Advanced Computing Group

- 2000: UMaine: Army grant to build a 206 node “Beowulf cluster” (fairly new at the time) which made the Top500 list.
- 2004: received another Army grant to build a 256 node/512 CPU cluster of Mac Xserve nodes (successful prototype for 10,000 CPU cluster in Alabama).
- 2012: Transitioned to the UMS to meet the needs of Research Computing for the whole University of Maine System (UMS). Started using the name “Advanced Computing Group” (ACG)
- The ACG has grown to also include:
  - Cloud Computing: Openstack using Ceph storage
  - Storage for backup: Departments and groups, multiple storage solutions including Seafile and CIFS
  - Data Management Planning for grant requirements
  - Data Archiving for post-grant data preservation
  - Custom Data presentation: Portals
ACG Staff

Director
  Bruce Segee
Outreach Specialist
  Ami Gaspar
Data Architect
  Chris Wilson
Supercomputer Engineer/Administrator
  Steve Cousins
Cloud Administrator
  Forrest Flagg
Cyberinfrastructure Engineer
  Larry Whitsel
Current Primary Storage

- ZFS file servers composed of 81 or 84 drives
- 3 in Orono
- 3 in Portland
- Works well for majority of needs
- Performance lacking in some cases
- Scaling is a problem:
  - Add two systems every time we need to grow
Enter Ceph: 2017

- Wanted a Scalable HPC system (size and performance)
- Easy to manage, something multiple people can work with without much training (think vacations)
- Scalable performance
- Ideally:
  - Affordable
  - Something we can phase into with current hardware
- Didn’t think Ceph would work for us because at the time CephFS was not ready
- Found that SUSE SES was using CephFS and started dialog
Long process

- We’ve been working on this for over a year
- ACG is a small group and this has been a lower priority. I worked on it when I could
- It now is a high priority as capacity needs are increasing quickly

Process:

- Decide on CephFS
- Do it ourselves or go with Redhat, SUSE or other?
- Decided on SUSE to save time and money
- Create test system
- Install and test (ad nauseum)
- Deploy (happening now)
The Decision: SUSE SES 5

- SUSE has been very easy to work with
- Extremely helpful and patient
- Willing to try everything I’ve asked
- Cost model: per server vs capacity
- SES 5 has many features to ease management and monitoring
- Encourages input to know what features to add
Hardware design (testing phase) using old hardware

- 1 Master node
- 3 monitor nodes
- 2 Metadata nodes
- 4 OSD nodes each with
  - 42 4 TB SAS drives
  - 7 240 GB SSDs for Journals (6 OSD’s per SSD)
  - 2 480 GB SSDs for Metadata pool
  - Infiniband/10GbE VPI links (more on this soon)
Two sites

- Traditionally we have primary site in Orono plus backup site in Portland 100 miles away
- Would like to keep a backup site in case of disaster
- Dedicated fiber between sites without router
- No Infiniband in Portland
- Added IB-10GbE bridge to connect two sites
- Sub 4ms latency
- 2 OSD nodes in Portland, 2 in Orono
- Testing/benchmarking: dd, fio, bonnie++
Two sites: Motivation

- Backup: Disaster recovery: Attractive to not have to think about backups. Just set policy to have one copy in Portland.

- Alternative is to have a cluster in each Data Center and have periodic snapshots -> backup

- Licensing cost: more machines to license: Master, Monitors, Metadata

- On the other hand:
  - since this is for DR, performance isn’t a driver and can combine roles with OSD nodes, not use SSDs for journals
  - Performance cost might be too high
Two sites: Further Motivation: ERN-POC

- **ERN Project: Eastern Regional Network:**
  - **Vision:** *To simplify multi-campus collaborations and partnerships that advance the frontiers of research and innovation.*
  - **Proof of Concept (POC):** HPC clusters in New Jersey, Pennsylvania, Massachusetts, New Hampshire and Maine
  - **Federated logins between sites using Slurm**
  - **Shared storage between sites**
  - **Login to local HPC, run jobs at any of the sites**
  - **One Ceph cluster in New Jersey**
Two sites: Further Motivation: ERN, next step

- Preproposal for NSF Mid-scale Research Infrastructure Implementation grant submitted
- Hopefully will hear soon if accepted to submit a proposal
- $20,000,000 to expand to more sites and tackle many issues, including ...
- Multi-site Shared CephFS storage
  - I have data in Maine but want to submit a job in Pennsylvania that has Huge-Memory machines
- Maybe this is completely un-doable because of the distance/latency but worth trying to work on issues.
Some Results (Orono-Portland split test system)

- Deep Learning: Image analysis:
  - 2 million files in a single directory
  - First step: scan through all files: Metadata intensive:
    - NFS over ZFS: 10 hours
    - CephFS: Almost instantaneous
  - Second step: process each file: Bandwidth intensive
    - NFS over ZFS: 3 seconds per file
    - CephFS: 6 seconds per file.
- Overall: slower with CephFS using test system
- Hopefully can fix with newer equipment and tuning.
Some Results (Orono-Portland split test system)

Resiliency:

- Amazing resiliency. This is a test system but some people are actually using it. At least twice we have done something that has completely brought nodes offline while an HPC job was running on it. The job just paused until storage came back up and then continued.

- One outage was a network disconnection to Portland that lasted two hours.

- One outage was a problem I caused when moving two systems to Portland: two days and a complete rebalance of one OSD node.
Status and Future

- Actively testing Replications and Erasure Coding
- Decide on whether can live with performance hit with ~4 ms latency to Portland
- Would Cache Tiering help mitigate latency? Probably not?
- Lazy-IO: implemented in Kernel client? Help writes?
- Benchmark with IO500? Just heard about it yesterday (CERN)
- Soon to add 5 more OSD nodes and 3 new monitors:
  - 36 drive chassis: 30 Data drives, 5 SSDs for Journals, two SSDs for Metadata pool
- Tuning for Metadata server to take advantage of memory
- Once implemented, use as backing store for our next OpenStack system
Status and Future: 2

- Once production system is in place: migrate data off of ZFS systems
- Bring ZFS systems into Ceph cluster:
  - Add SSD’s to ZFS servers (journal and Metadata pool)
  - Purchase 30 new 10TB disks for each system
- End result: 14 OSD nodes, 4.2 PB raw storage
SUSE Experience

- What we are trying to do is kind of crazy
- Most vendors would probably say we are out of scope with multi-site live CephFS system
- SUSE has never pushed back on this
- They have always been receptive about trying things to try to come up with something that will meet our needs
SUSECON recommendation

- SUSE Enterprise Storage, from Requirements to Implementation – A Best Practice Guide
- TUT1130 Thursday 4:30 Belmont 3
Thanks very much to Alejandro Bonilla for all of his ongoing help!

Thanks also to Paul Augustyniak and David Byte for their advice and help in setting this system up.
Thank You

Questions?