Interpreting radiations from the Universe.

Site report 2017 IRFU

ARNAB SINHA
“Before I begin, I’d just like to make it known that I didn’t volunteer to do this presentation.”
SUMMARY

- Unix
- GRID
- Projects
- Infrastructure
- Security
- Management
- IT Infrastructure
  - Puppet 4, Foreman 1.14.2 (Dashboard)
  - CEPH
IT Infrastructure

- Puppet 4,
  Foreman 1.14.2 (Dashboard)
- CEPH
very high memory usage
• even with bluestore_cache_size=100MB

We WANT storage HA
• moved machines around to avoid downtimes when racks get power-cut (as usual)
• wanted to create service HA using storage HA, but needed snapshoting at the FS level for repository mirrors:
  • discovered that rbd cannot be mounted on 2 machines without ocfs2 or gfs2, but :
    • gfs2 has no snapshoting
    • ocfs2 has snapshoting, but modules not in any CentOS/SL, only in Oracle linux
  • cephfs snapshots are said to corrupt data
• ?? Any solutions for Fred…

got rid of EC pools for now
• EC pools cannot be highly available unless you have > (N+M) racks (coz’ racks get power-cut :’/ )
• because crush maps are all but flexible (for small sites ?).
• only choosleaf can really be used.
• other methods are useless, only the failure domain counts

only 1 VM in bluestore :
• started seeing daily PG scrub errors
• mark OSD out, mark it IN, fixed (usually). bug ?
Reinstalled with bluestore
HTCondor issues

We see intermittent cgroup issues

- condor fails at setting *some* cgroup settings
- causing jobs to be killed because of wrong memory limits
- so far, did not find any fix/reason

Machines management with Puppet 4... done

ARC CE started showing signs of scaling issues

134 condor WN, dynamic slots

- issues with all the ARC files
  - too many, cleanup not obvious (and was buggy)
  - too many files causing too many iops on a single server (no ssd)
- issues for updating the ARC BDII (see : files/iops)
2nd DELL Z9100 ordered

- we really want HA ;)
- our 1st Z9100 was a SPOF
  - not easy/impossible to temporarily bypass in case of issue

Ready and started fighting for 100Gbits, but our institute and NREN will require *years* to implement it.

10/10/2016: we deployed IPv6 on all our production servers – we’re now full dual stack

- it took us 3.5 years of fight to get this done (right).
Towards Cloud Computing

EOSC (European Open Science Cloud)
- create a trusted environment for hosting and processing research data to support EU science
- Participation in WP6

Indigo-datacloud
- OS data and computing platform targeted at scientific communities
- Participation in WP3
- Started testing Use Cases

Open for collaborations…
• Following a major air conditioning problem in 2016

• Adding a 220KW Carrier cooling system unit that supports high temperatures (for redundancy)

• Modification to the alarm system (which didn’t work in 2016)

• Water temperature/pressure
• Outdoor temperature
• Cooling systems (Carrier/Emerson)

• Targeting for 700KW from 500KW by 2017 end
INFRASTRUCTURE

- New cooling system
INFRASTRUCTURE

New real-time power usage (June 2017)
- 3 Cooling systems
- 3 server rooms

Advantages
- Calculate PUE precisely
- Work on simpler electricity savings

Planned
- Several stops this year
  - Starting week 22 (4 days)
PSSI: NEW SECURITY POLICY (CEA/IRFU)

• Reinforcement of security (France/CEA/IRFU norms)

• Inventory / encryption / updates / password
  • Secure VLANs

• Workstations
  • Windows (80% done)
  • Mac
  • Linux
  • Network equipments
SYSTEMS MANAGEMENT

GLPI

- The helpdesk tool launched recently
  - Apprentices examine as first level of support

Ongoing projects

- Present: Windows (Microsoft SCCM), Linux/Mac (FusionInventory and GLPI)
- Planned:
  - Automatic s/w inventory (all platforms)
  - H/w inventory
CONCLUSION

Major achievements

• Grid CEPH with bluestore
• GLPI

Roadblocks to overcome

• Network issues
• Shortage of personnel

Future planning

• EGI Federated Cloud
• Integrated Inventory management
THANK YOU!
QUESTIONS?
Interpreting radiations from the Universe.