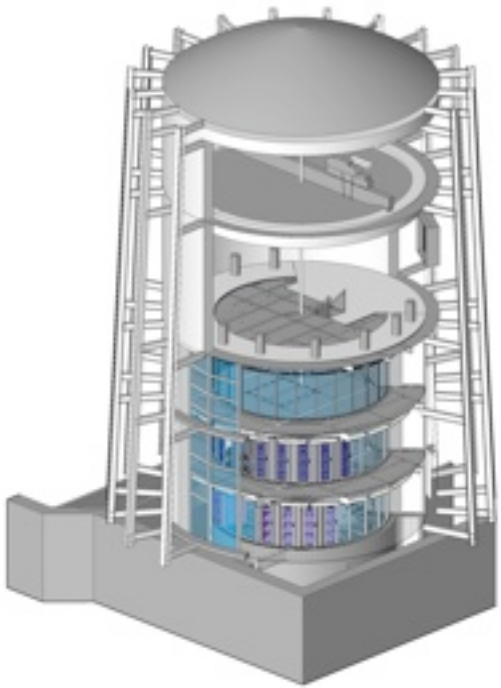
Lustre deployment and early experiences

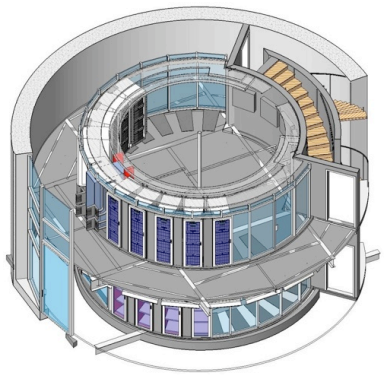
Florent Parent
Coordinator, Québec site

CLUMEQ



Outline

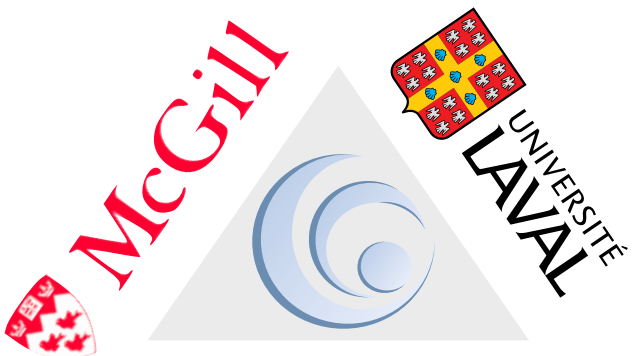
- What is CLUMEQ?
- CLUMEQ's new HPC cluster: *Colosse*
- Lustre experience



What is CLUMEQ?

- Consortium of 11 universities in the province of Québec, Canada
- Part of the *Compute Canada* national platform
- Two HPC sites:
  - ✓ Montréal
  - ✓ **Québec City**

Who is CLUMEQ?



UQÀM

UQAR

Université du Québec

Université du Québec en Outaouais

Université du Québec à Trois-Rivières

UQAC
Université du Québec à Chicoutimi

Université du Québec en Abitibi-Témiscamingue

École nationale d'administration publique

Université du Québec
Institut national de la recherche scientifique

Université du Québec
École de technologie supérieure

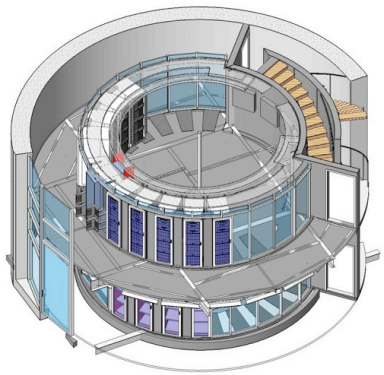
CLUMEQ's mission

- To serve the HPC needs of its member institutions in all fields of research
- To outreach non traditional and emerging HPC fields
- To train "highly qualified personel" (HQP)


Whitehorse
Yellowknife
Iqaluit
Enabling Canadian research excellence through high performance computing
Favoriser l'excellence en recherche au Canada avec le calcul de haute performance
Edmonton
Victoria
WestGrid
Calgary
Saskatoon
Regina
Winnipeg
SciNet
Toronto
RQCHP
CLUMEQ
Montréal
Ottawa
HPCVL
SHARCNET
St. John's
ACEnet
Charlottetown
Fredericton
Halifax
COMPUTE / CALCUL CANADA
PLANNED SYSTEMS
Capability -
Capacity -
Vector -
Major Storage -
Canadian Advanced Data Network
International Links

Outline

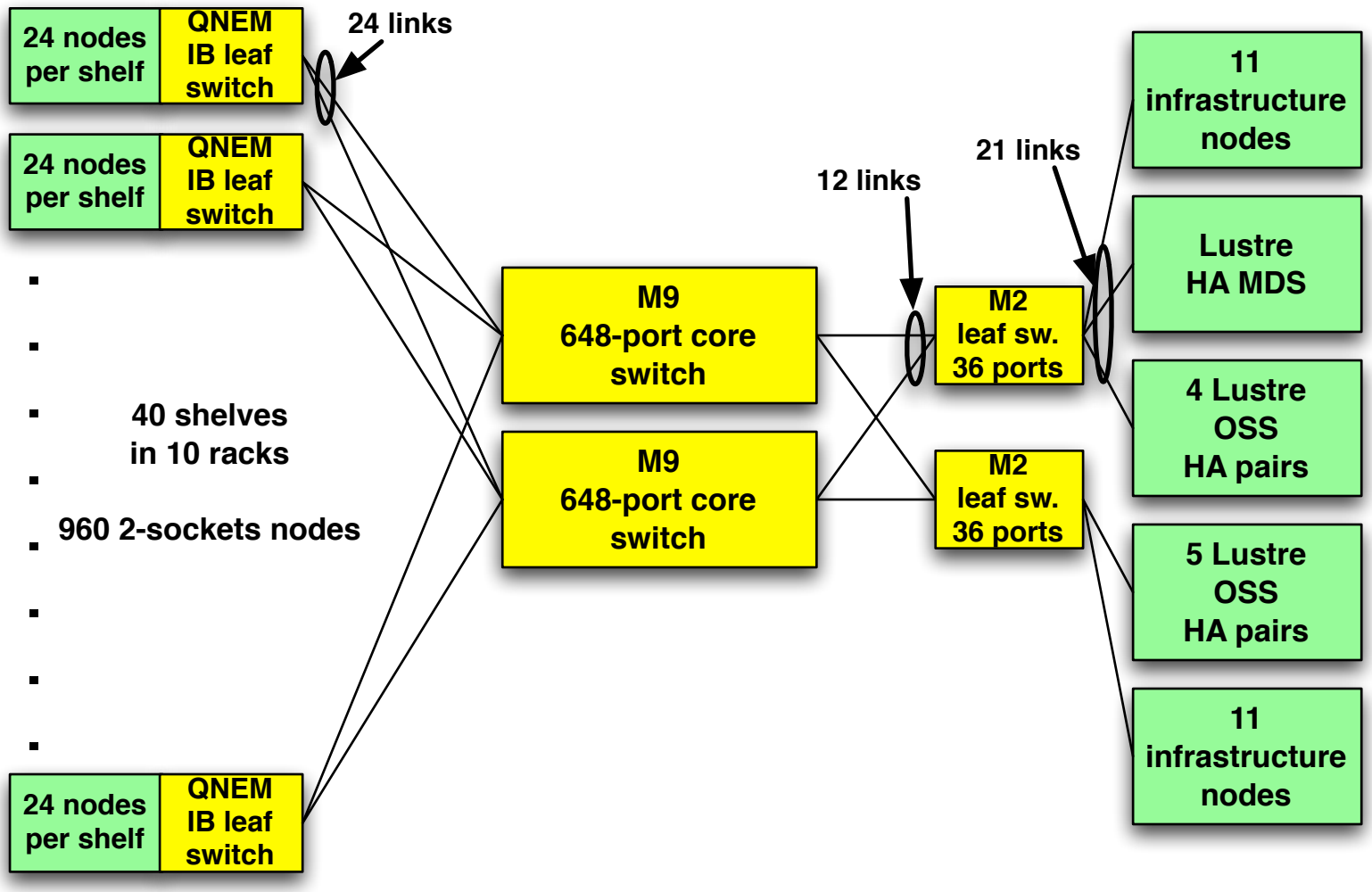
- *What is CLUMEQ?*
- CLUMEQ's new HPC cluster: *Colosse*
- Lustre experience

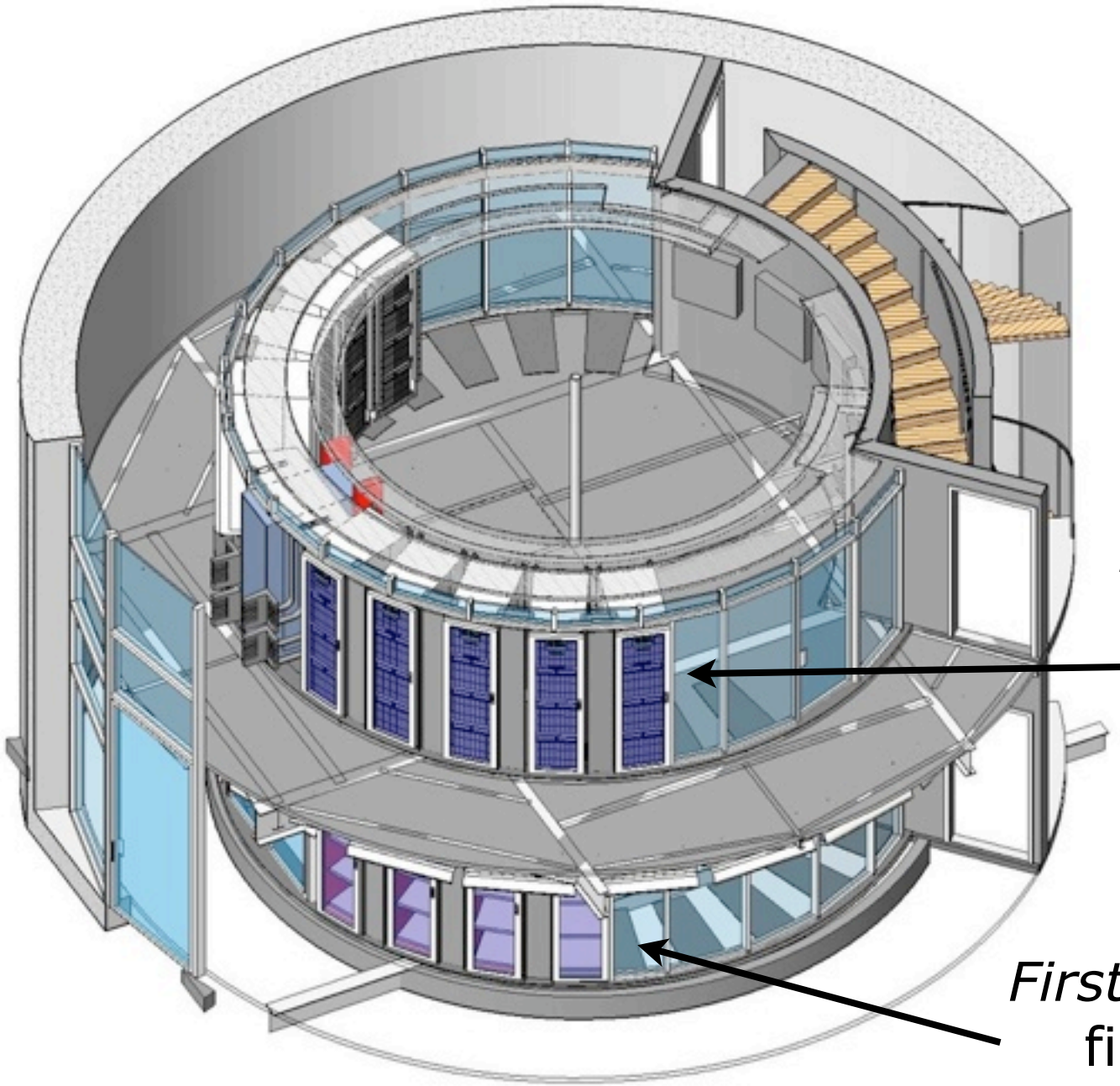


CLUMEQ *Colosse*

- Sun constellation system
  - ✓ 10 fully loaded Sun Blade 6048, with X6275 modules (double Nehalem EP blade, 2.8GHz, 24GB of RAM)
  - ✓ full-bisection IB-QDR interconnect (2xM9 switches)
  - ✓ 1 PB of Lustre storage in a high availability configuration, using 2 MDS and 9x2 OSS
  - ✓ Sun J4400 storage arrays
- 86 Tflops peak
  - ✓ 77 Tflops max (preliminary run)
  - ✓ ---> 80 Tflops ?

Infiniband Architecture





Racks aligned in a circle around a central hot core; outside ring is a cold aisle

Second floor contains all compute racks + core networking switches

First floor contains file system & infrastructure nodes

Street view...

Satellite view...

Daniel A. Denis ARCHITECTE

Free air cooling system

Main cooling system

Air blowers

cooling coils

et autres travaux dans l'entretoit.

Nouvelle échelle d'accès et plate-forme, voir document de l'ingénieur en strucutre

Pont roulant existant

Voir documents ingénieur en structure pour la réinstallation du palant. Sabler et peindre la poutre et les attaches au complet

Voir documents ingénieur en mécnaique pour les conduits de ventialtion et les nouvelles persiennes.

Garde-coprs amovibles du niveau 4, voir feuille A13

Conduit de ventilation au périmètre

Appareils d'éclairage suspendus, voir ing électrique

Support de fils électrique

Support de fils réseau

Plancher en lamelles d'acier pour tout le niveau

Conduit passant sous l'escalier mezz niv 9008 selon le niveau du plafond de la salle 00625A

Séparation coupe-feu degré de résistance au feu de 2 heures

Renforcir structure du mur avec des colombages fixés au mur

Limite de la cloison

Plafond suspendu

Nouveau plancher

Dessus pont roulant 26 764 346'-3"

Voir documents ingénieur en électricité pour modification du moteur et de l'alimentation électrique et ingénieur en structure pour travaux de modification sur le pont roulant.

Conduit de ventilation au périmètre

Dessus toit 31 413 361'-6"

Dessus dalle 30 498 358'-6"

Niveau 4 mécanique 20 539

Dessus ouv. mur rideau 20 287 325'-0"

Niveau 3 Serveurs 17 143

Niveau 2 Serveurs 14 143

Niveau1 Bas de la fenêtre 11 143 295'-0"

Plancher existant 10 000 291'-3"

Niveau mezzanine 9 008 288'-0"

Niveau 0 5 732 277'-3"

Niveau 00 5 046 275'-0"

18758 [61'-6 1/2"] Intérieur libre de la tour

9144 [30'-0"] mur rideau

3396 [11'-1 3/4"]

3000 [9'-10"]

3000 [9'-10"]

6000 [19'-8 1/4"]

3135 [10'-3"]

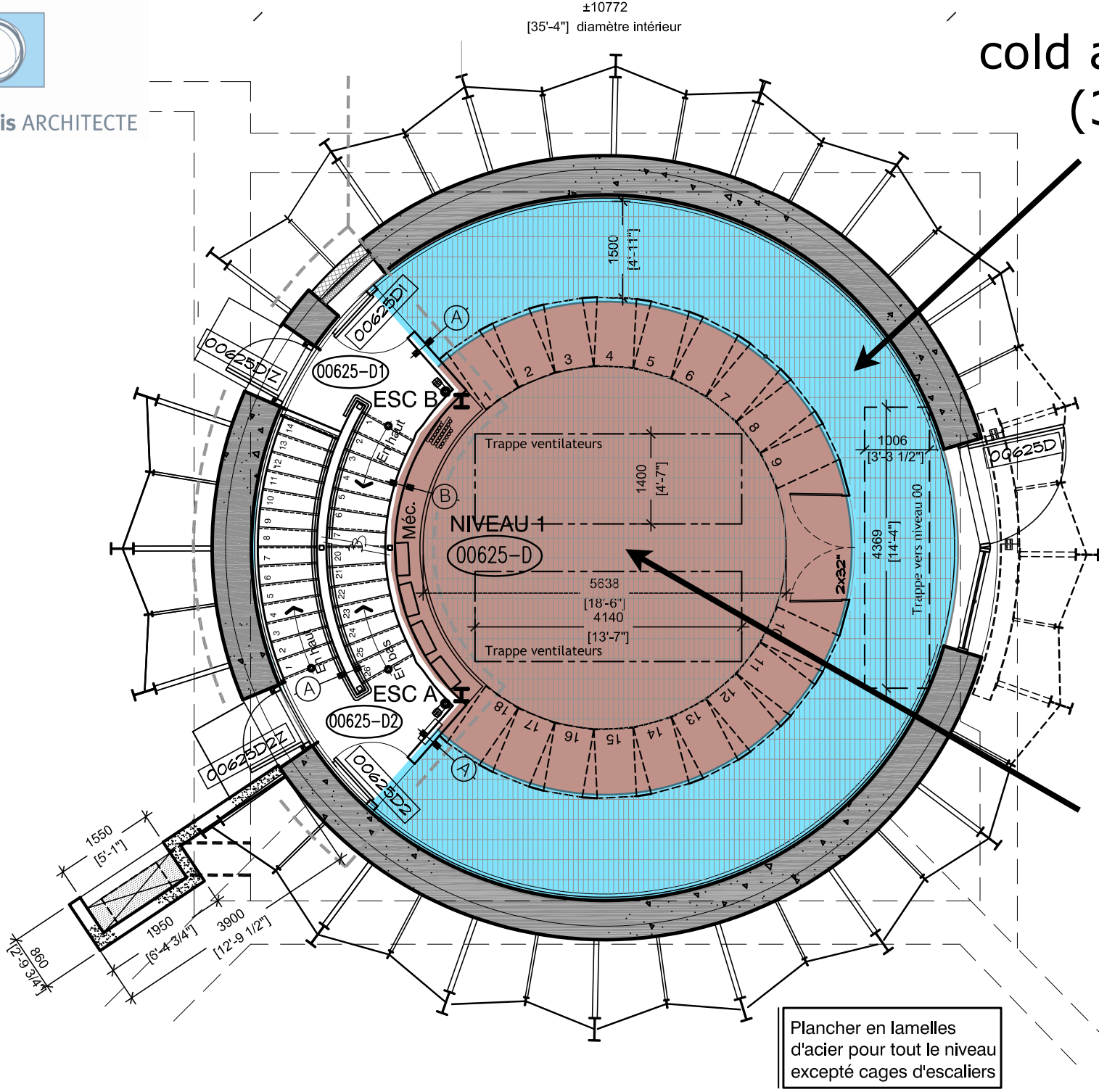
5411 [17'-9"]

3276 [10'-9"]

1143 [3'-9"]

3429 [11'-3"]

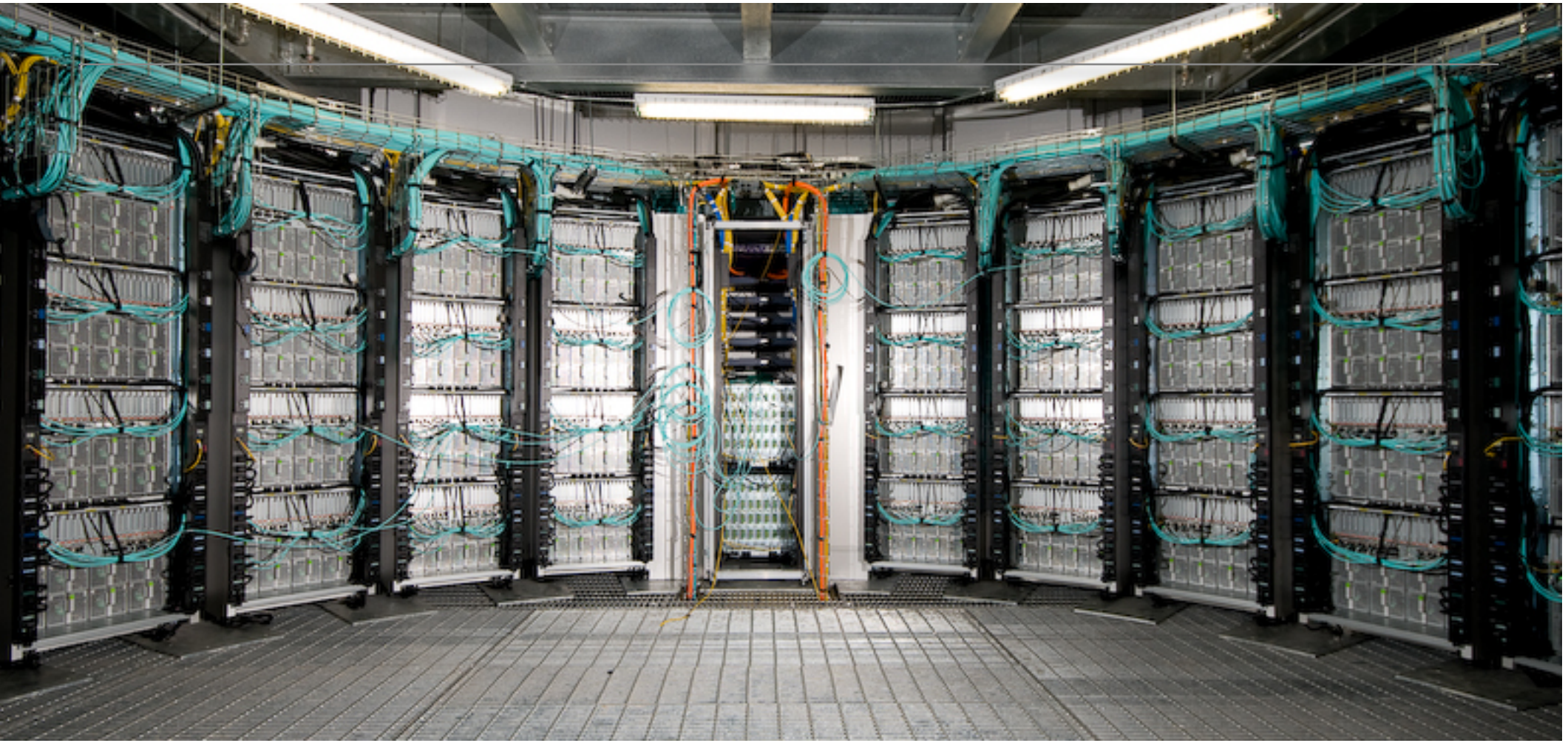
Daniel A. Denis ARCHITECTE



Plan niveau 1

Construction

View inside hot air core

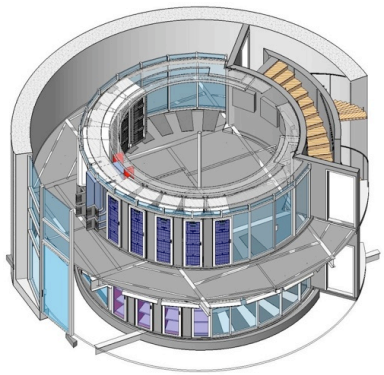


Main specifications

- Rack capacity: 56
- Cooling capacity: ~1.5 MW
- Electrical capacity: 1.1 MW (1.6 MW)
- Blowing capacity: 132,500 CFM
- Maximum air velocity: 2.4 m/s
- floor loading capacity: 940 lb/ft^2

Outline

- *What is CLUMEQ?*
- *CLUMEQ's new HPC cluster: Colossus*
- Lustre experience



Timeline

- New staff hired in Apr 2009 and July 2009
- Nov 12: System acceptance signed
- Dec 16: First “Beta users” on machine
- Nov - now: Learning, debugging, patching, tuning, helping users

Experience so far...

- Many technologies to get up to speed on installing, monitoring, debugging, tuning ...
  - ✓ Lustre
  - ✓ Infiniband
  - ✓ Grid Engine
- So far, pretty much everything is learned as we go
  - ✓ asking questions to SMEs
  - ✓ reading documentation, mailing list discussions

CLUMEQ Lustre deployment

- At acceptance:
  - ✓ CentOS 5.3 (2.6.18-128.2.1.el5)
  - ✓ OFED 1.4.1, Lustre 1.8.1.1
- Now
  - ✓ CentOS 5.3 (2.6.18-164.11.1.el5_lustre.1.8.2)
  - ✓ OFED 1.4.1(?) stock from RH
  - ✓ Lustre 1.8.2 + patch

Lustre fixes

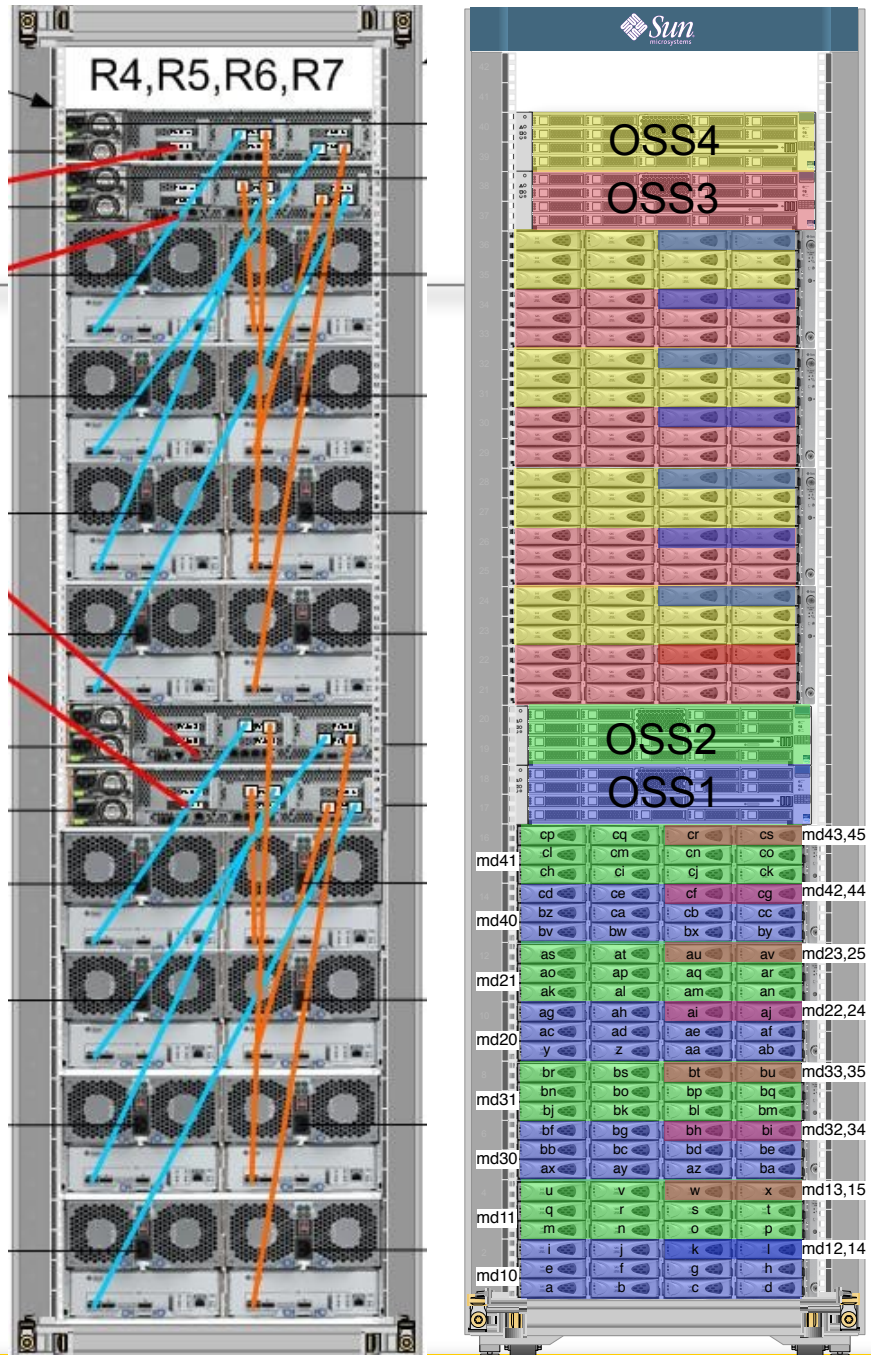
- OSS crash during heartbeat/failover
  - ✓ “CPU hog/soft lockups” bug (21612, 19557)
  - ✓ Patched in 1.8.2
- 1.8.2 installed when GA
  - ✓ Started to see high load on MDT, Lustre hanging on clients
  - ✓ inode link count fix (22177)
  - ✓ Installed patched version

Filesystem structure

- Lustre is the only FS on Colosse
  - ✓ /home for user accounts
  - ✓ /rap for group-shared space
  - ✓ /scratch for temporary data
- Lustre striping
  - ✓ /home and /rap use striping of 1 (typically small files)
  - ✓ /scratch uses striping of 72 (parallel IO performance)

OSS HA and RAID setup

- OSS HA pairs
  - ✓ 4 OST per OSS
  - ✓ Linux heartbeat used to signal failure
  - ✓ Have 8 OST on OSS in failure mode
- RAID 6
  - ✓ 10 1TB disk per OST
  - ✓ Software raid (md)

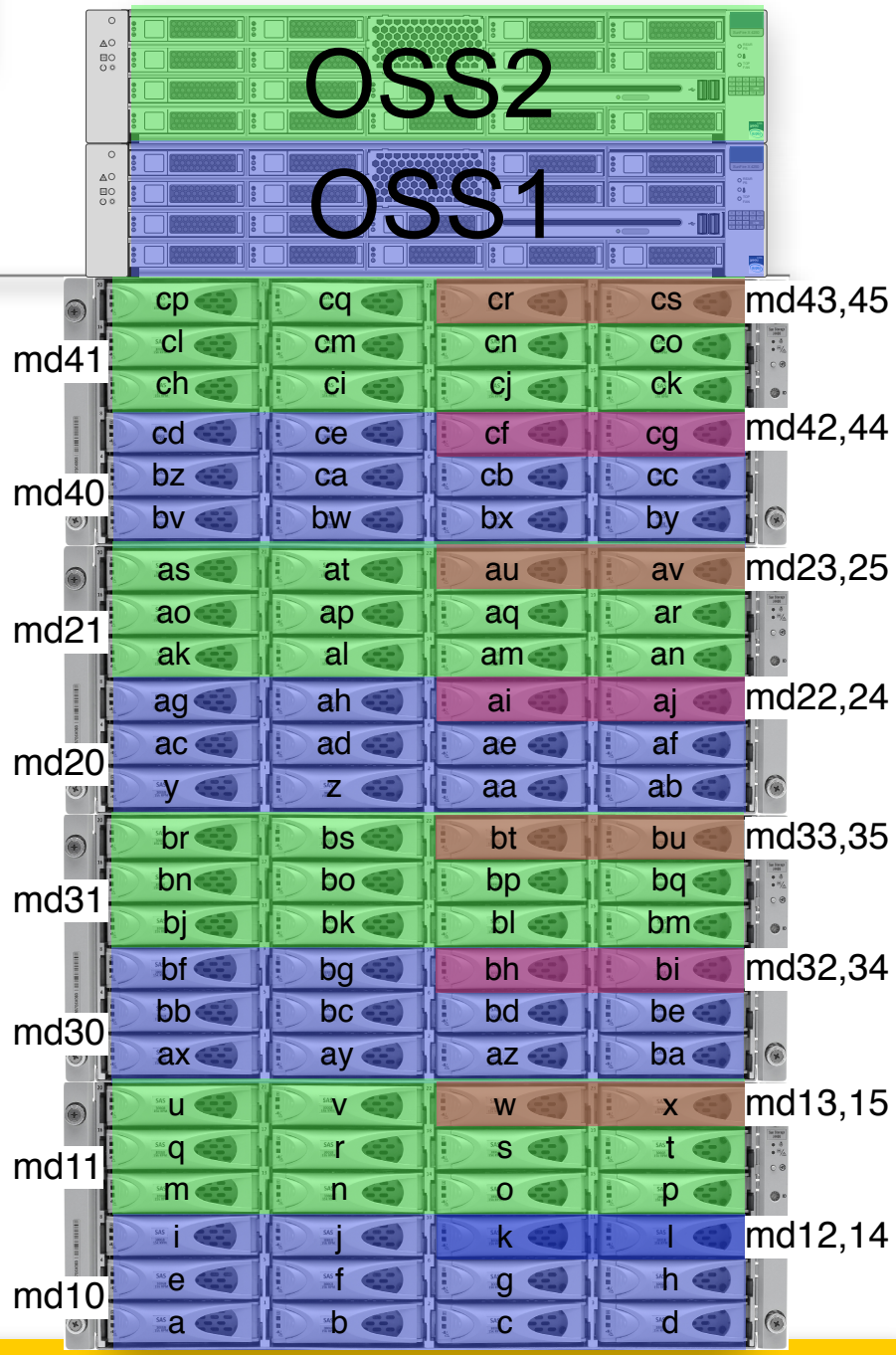


HA and heartbeat

- Linux heartbeat/HA
  - ✓ Does not always work. Seen issues w.r.t. node not able to “kill” its neighbor (IPMI issues, investigating)
  - ✓ Currently working in manual HA mode
- Observing that it takes a long time for clients to use the new OSS taking over OSTs.
  - ✓ Not sure yet why. Needs investigating. todo++

Finding a failed disk

- "Interesting" experience
- Device name to physical location is not reliable
  - ✓ need to double check with md* commands
- Then came "blinkenlights"
  - ✓ http://wikis.sun.com/display/HPCSoftware/JBOD+Troubleshooting+Utilities
  - ✓ REALLY useful!



Performance

- IOR used for parallel I/O measurements
  - ✓ IOR read performance = 33.6 GB/s
  - ✓ IOR write performance = 17.3 GB/s
  - ✓ all over IB
- Performance monitoring
  - ✓ Is there a BCP out there?

Conclusion

- Learned quite a lot in the past few months
- Lustre support team is key
- Looking to share experience