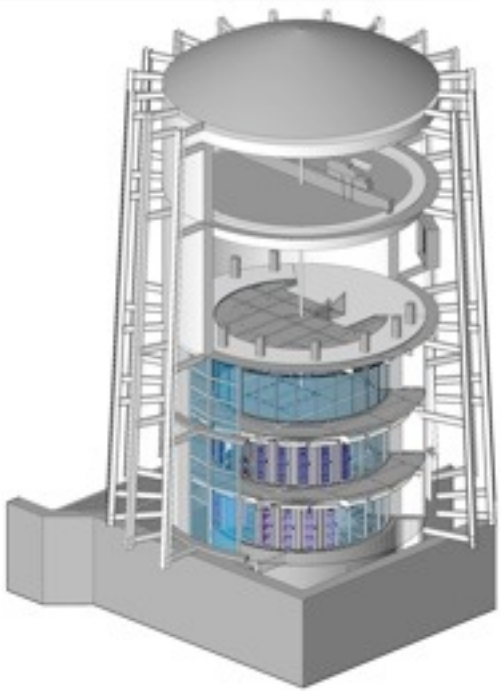
Colossus : A cool HPC tower!

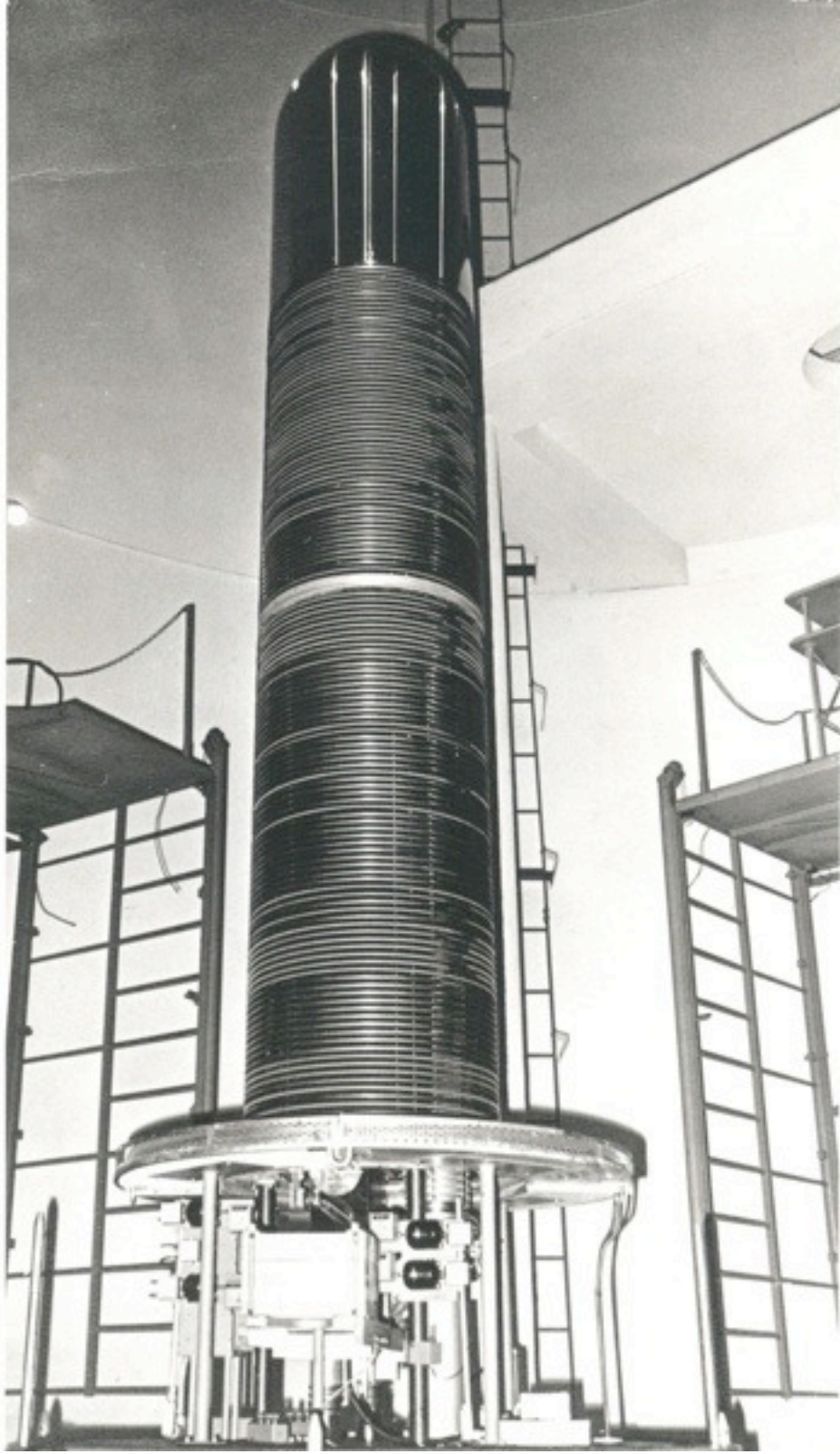
Marc Parizeau, Ph.D.
Deputy Director of CLUMEQ



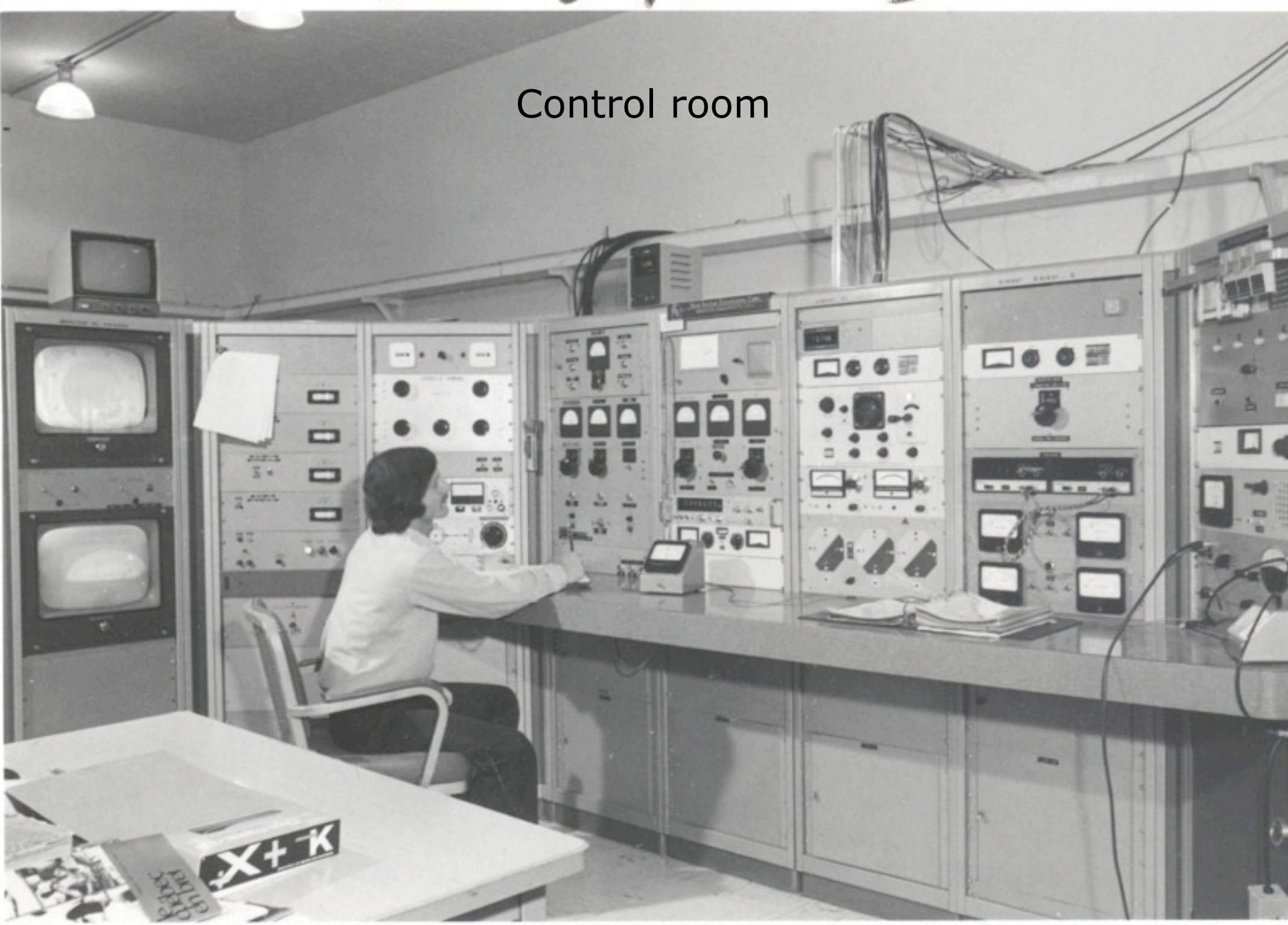
What to do with a 36 feet wide by 65 feet high **nuclear** grade silo with 2 feet thick concrete walls ?

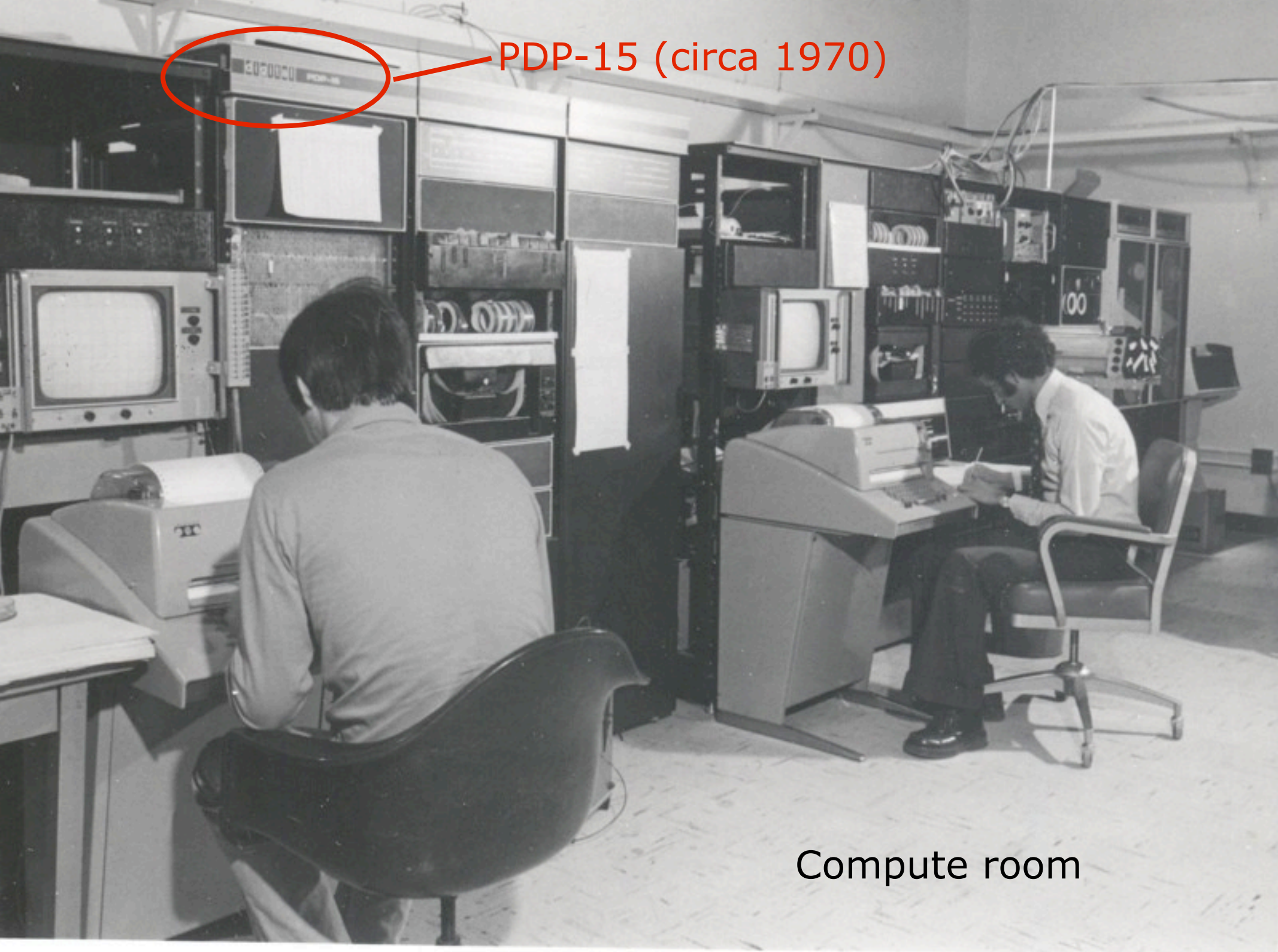
Saturday, November 14, 2009

- Van Der Graaff particle accelerator (circa 1965)
- At the time a 20M$ infrastructure




circa 1965


Control room
X+ K-

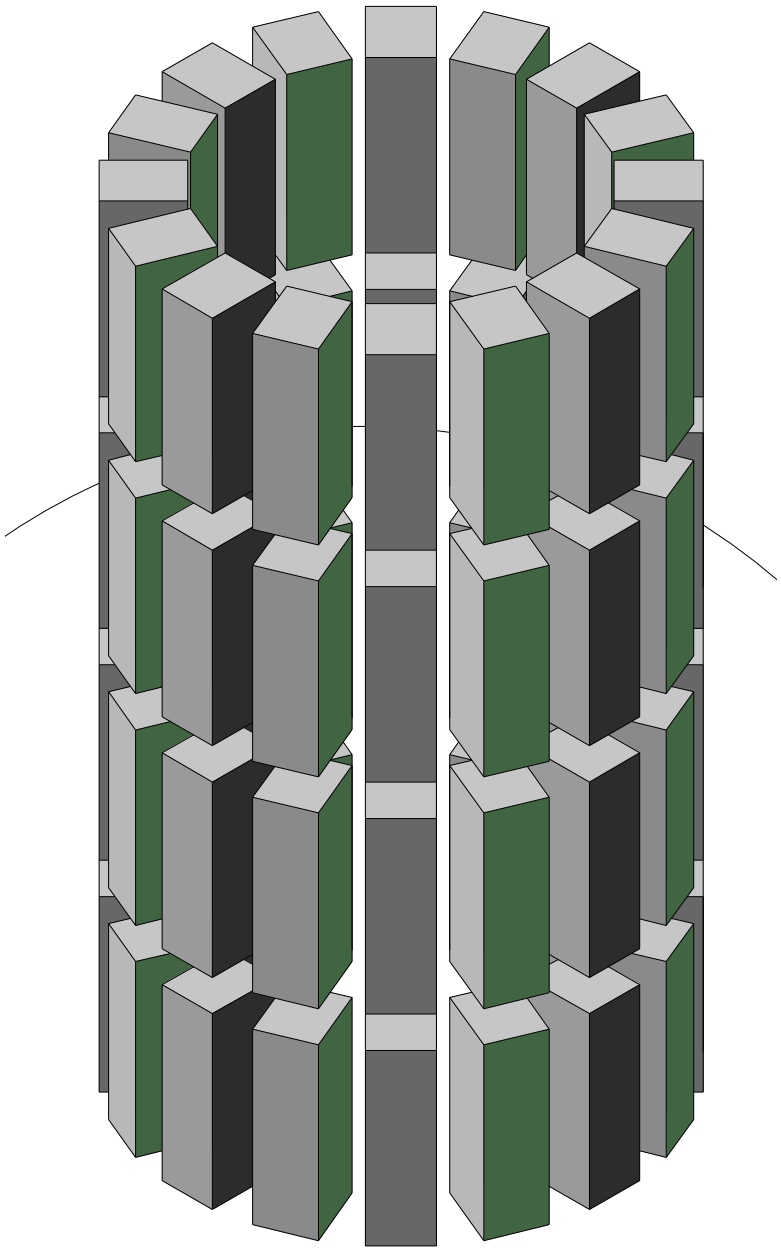

PDP-15 (circa 1970)
Compute room


Target room

So what to do?

The basic idea

- Cylindrical hot core
- Ring shape cold plenum
  - ✓ large cross-section
  - ✓ low air velocity
  - ✓ no corners to create turbulence
  - ✓ shorter cable lengths
- The building is the CRAC!



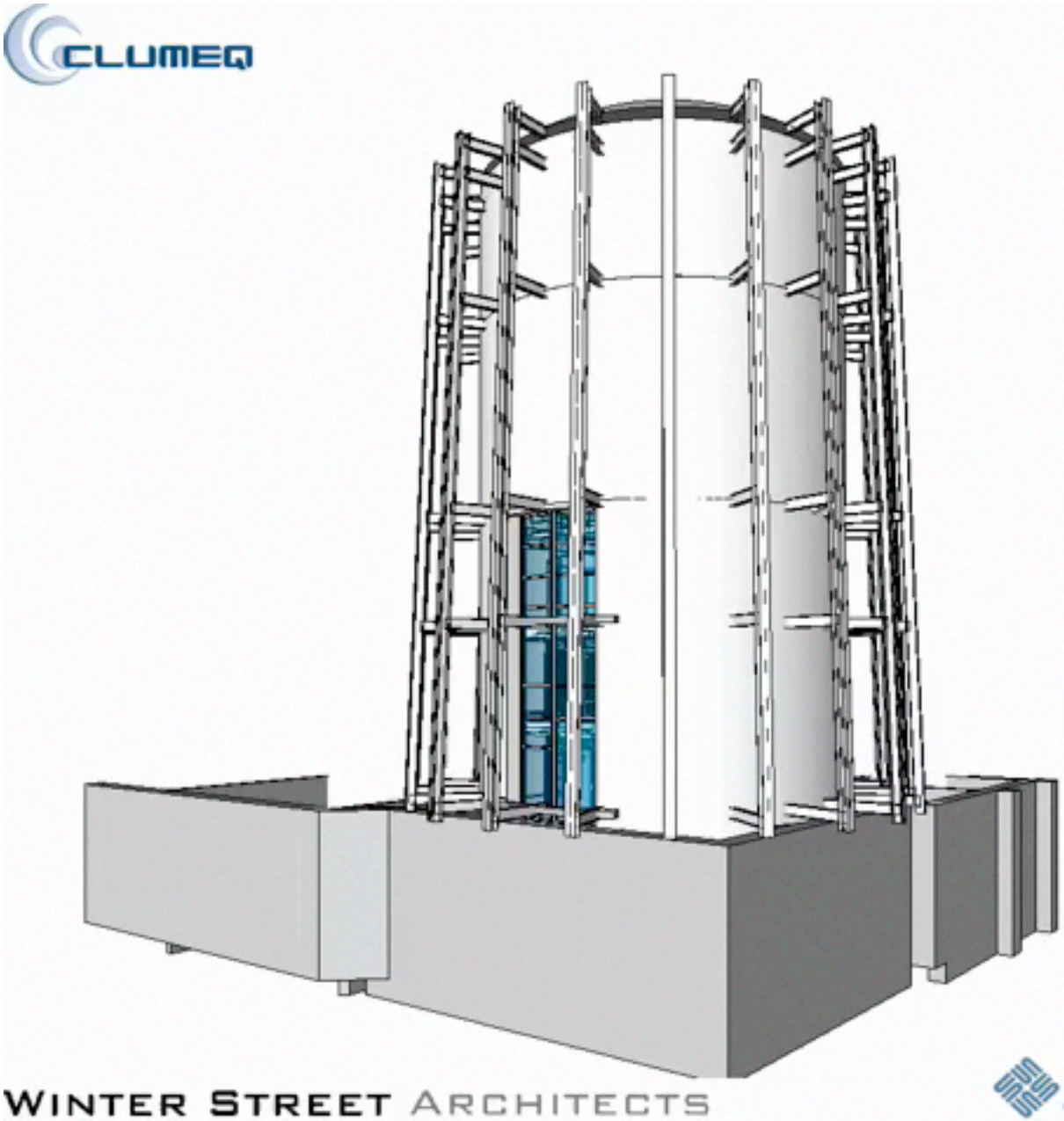
The steps

- Hire creative architect and engineers
- Take the time to explore several concepts (12)
- Win a few arguments with bureaucrats who don't have a clue about what your trying to achieve
- Make sure you get precise and complete plans and specifications before going to RFP
- Make sure you have a good project manager to keep control over the budget!

- Also, if you can, find a very bright and creative student to help you sort out all of the *nuts and bits* !

The result

- Meet the CLUMEQ Colossus...
  - ✓ high efficiency,
  - ✓ high capacity,
  - ✓ cool design



The Rhodes Colossus, one of the Seven Wonders of the ancient world...

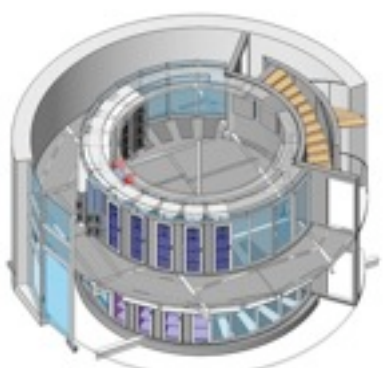
constructed to honor the Rhodian patron ***Helios****, the god of Sun*

circa 1970

The Forbin project!

Outline

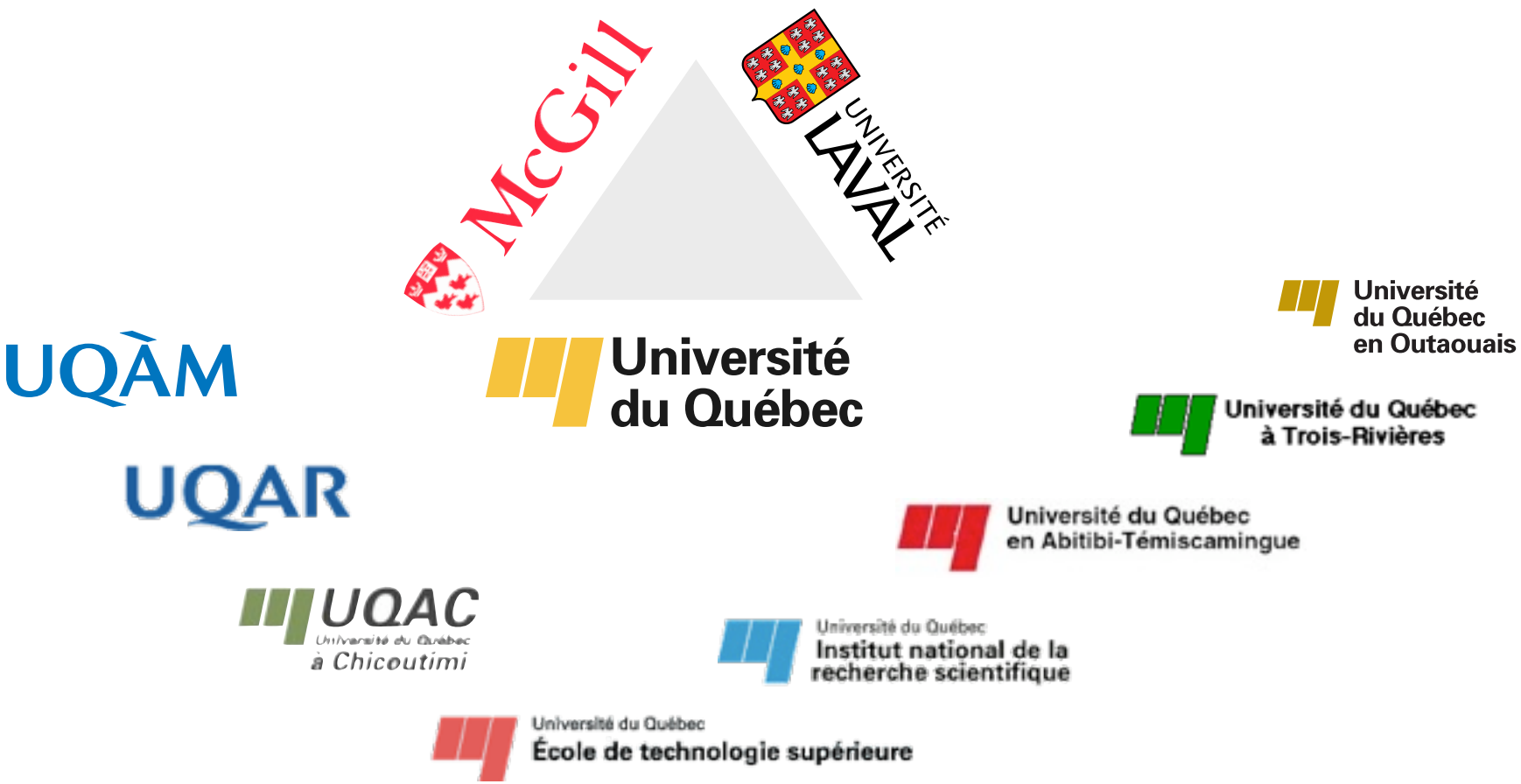
- CLUMEQ consortium & Compute Canada
- CLUMEQ's new HPC cluster: *Colossus*
- Québec City site
  - ✓ old Van Der Graaff particle accelerator recycled as a cooling enclosure
  - ✓ cylindrical layered design
  - ✓ high cooling capacity
  - ✓ energy efficiency



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?



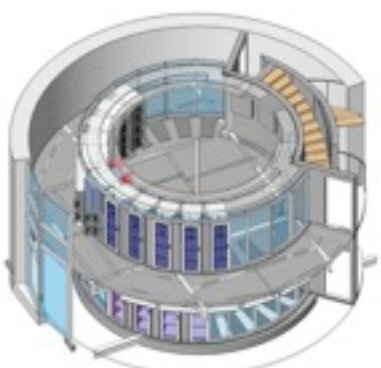
CLUMEQ's mission

- 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
Ottawa
Montréal
CLUMEQ
Fredericton
Charlottetown
Halifax
St. John's
ACENET
COMPUTE / CALCUL CANADA
PLANNED SYSTEMS
Capability -
Capacity -
Vector -
Major Storage -
Canadian Advanced Data Network
International Links

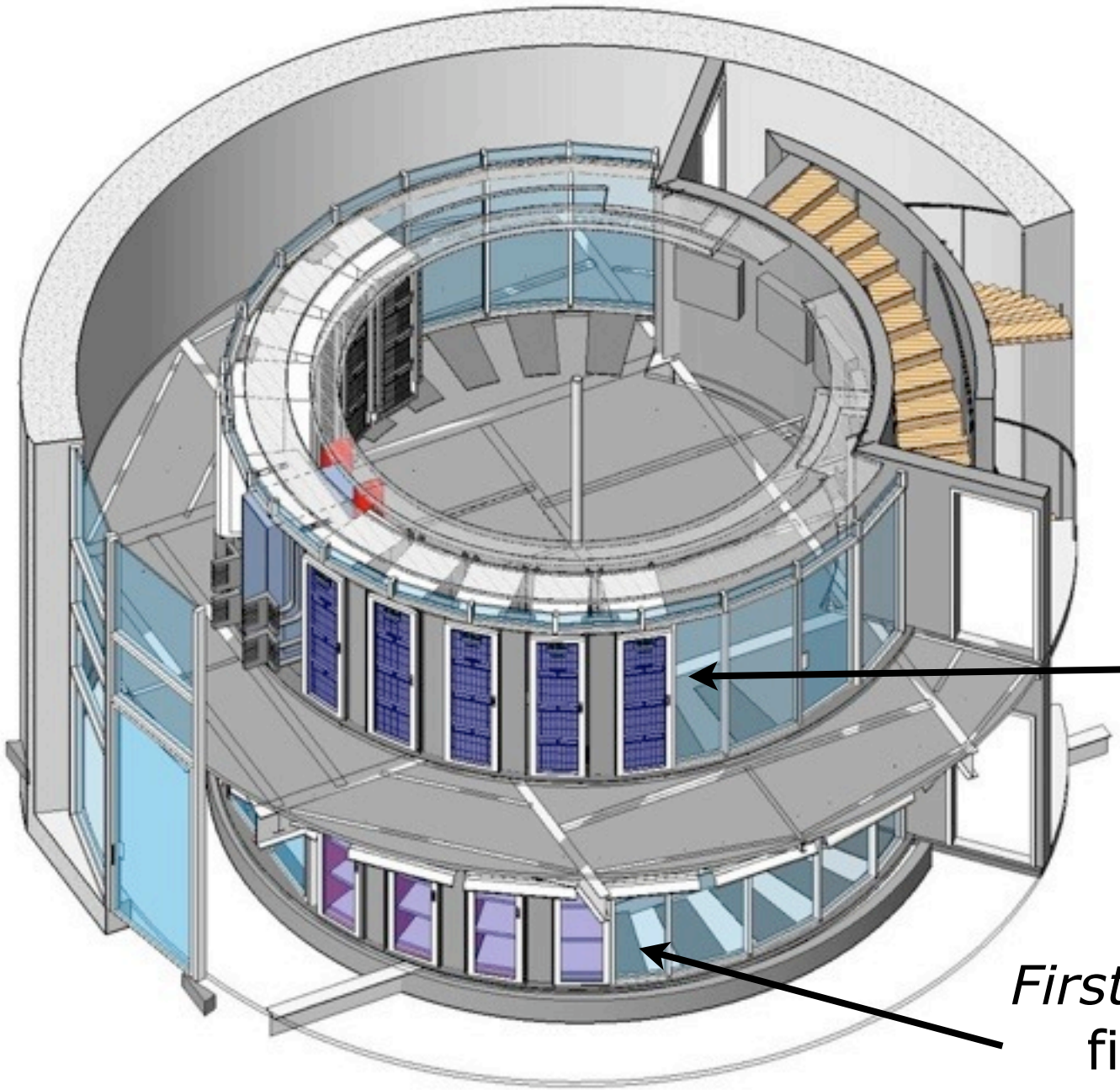
Outline

- *CLUMEQ consortium & Compute Canada*
- CLUMEQ's new HPC cluster: *Colossus*
- CLUMEQ's Québec City site
  - ✓ old Van Der Graaff particle accelerator recycled as a cooling enclosure
  - ✓ cylindrical layered design
  - ✓ high cooling capacity
  - ✓ energy efficiency



CLUMEQ *Colossus*

- 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 ?



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

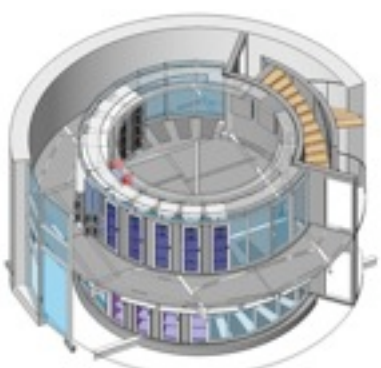
Performance

- Compute nodes:
  - ✓ 36.6 < STREAM < 37.9 GB/s
  - ✓ SPECint = 233
  - ✓ 189 < SPECfp < 190
- Interconnect:
  - ✓ MPI ping-pong latency < 2 usec
  - ✓ MPI ping-pong bandwidth > 3.1 GB/s
  - ✓ MPI all-to-all bandwidth > 1.1 GB/s
  - ✓ iPerf > 9.2 Gb/s

- Lustre file system (18 OSS):
  - ✓ IOR read performance = 33.6 GB/s
  - ✓ IOR write performance = 17.3 GB/s
  - ✓ all over IB
- Boot time: 4 minutes 58 seconds
  - ✓ over IB
- Max power HPL: 332 kW

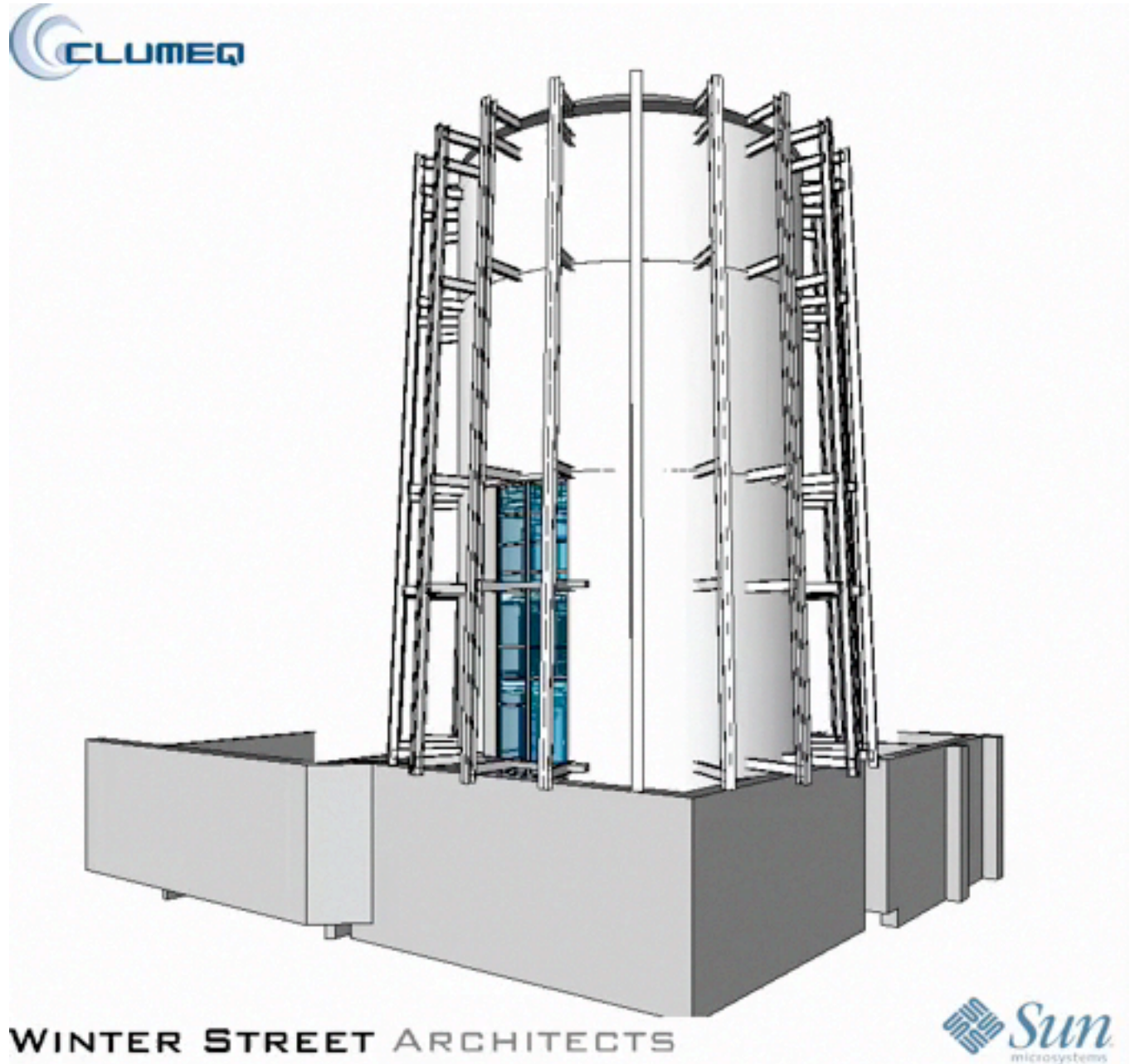
Outline

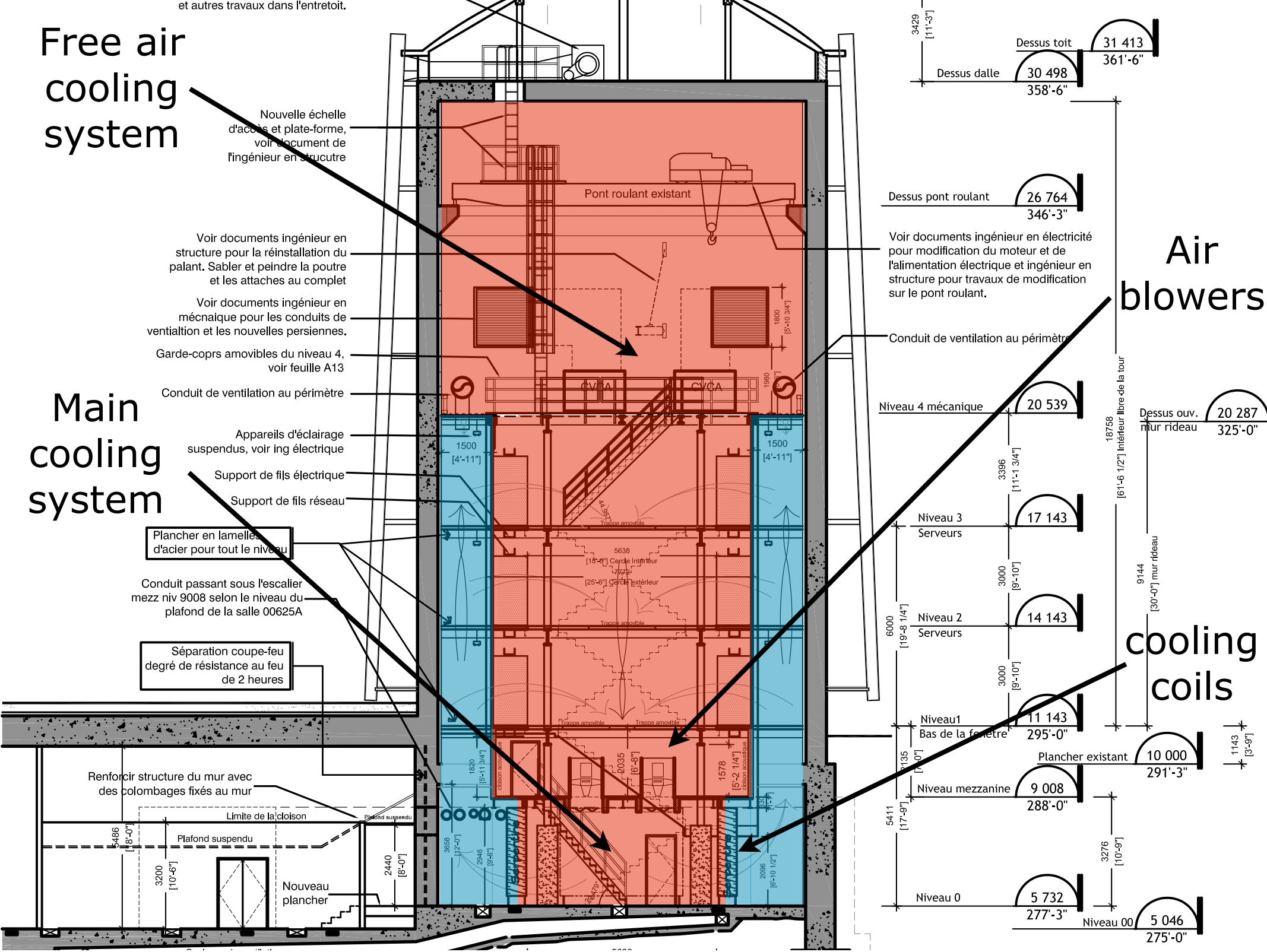
- *CLUMEQ consortium & Compute Canada*
- *CLUMEQ's new HPC cluster: Colossus*
- CLUMEQ's Québec City site
  - ✓ old Van Der Graaff particle accelerator recycled as a cooling enclosure
  - ✓ cylindrical layered design
  - ✓ high cooling capacity
  - ✓ energy efficiency

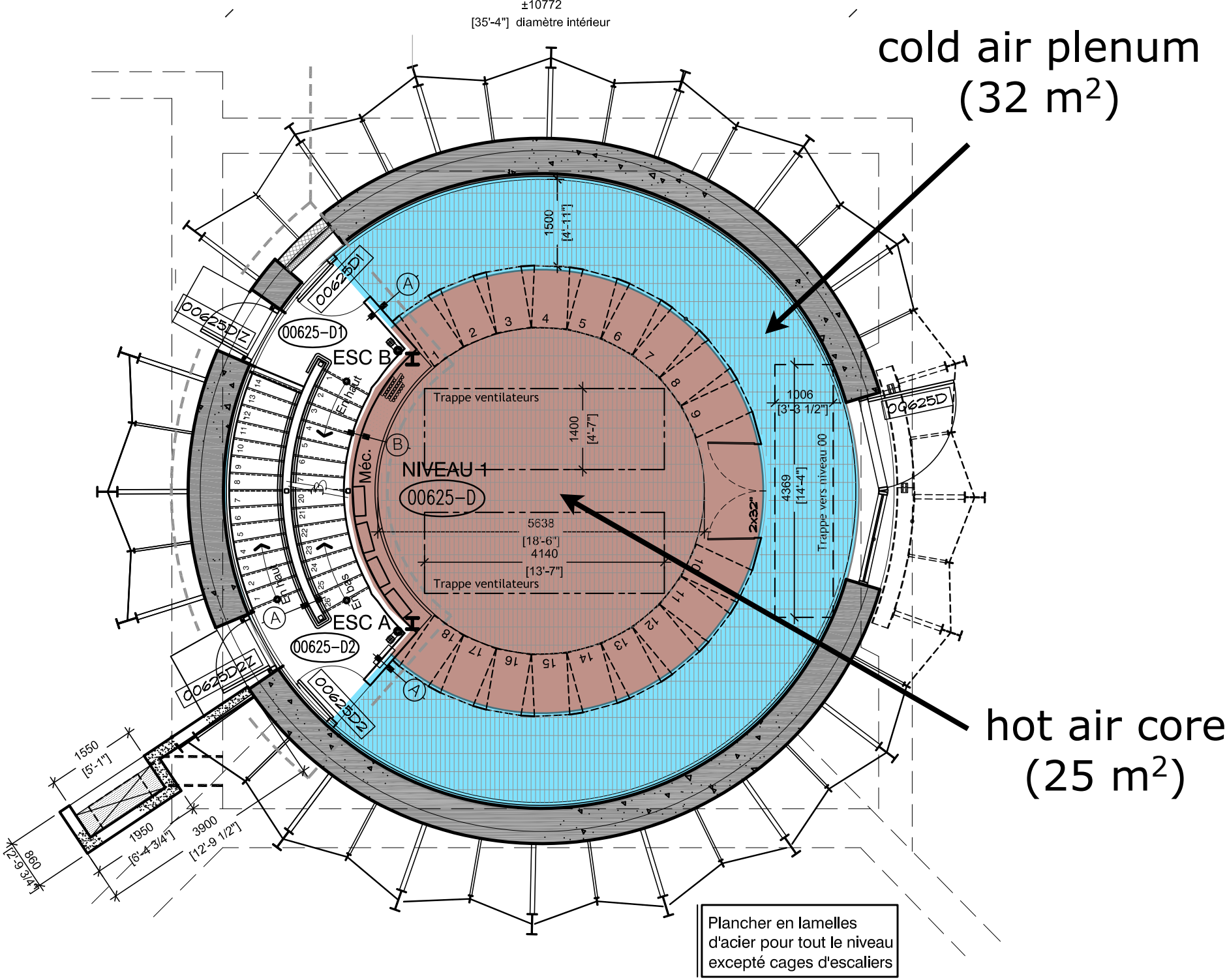


Street view...

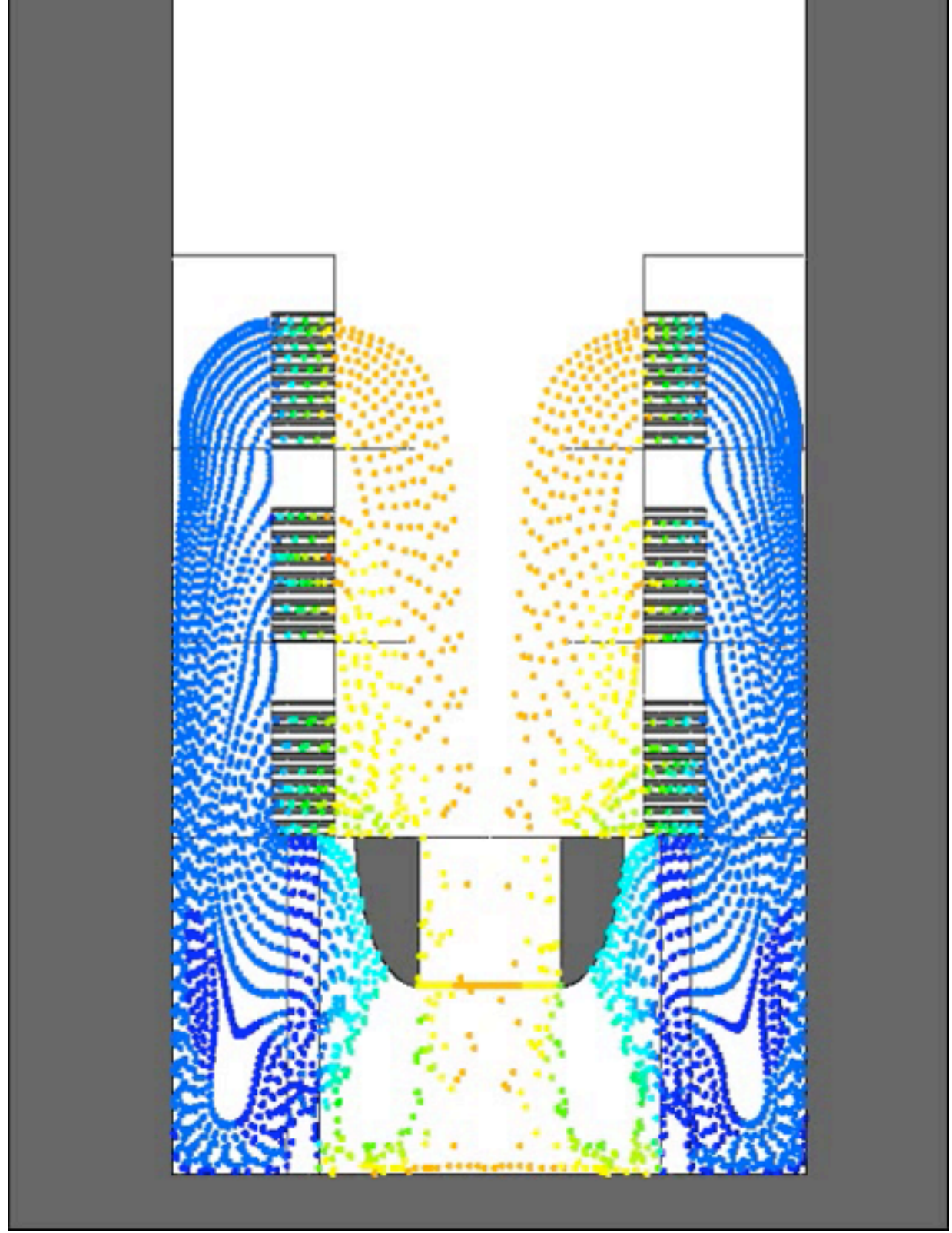
Satellite view...


CLUMEQ
WINTER STREET ARCHITECTS
Sun
microsystems


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 structure
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
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
CVCA
Trappe amovible
Cloison acoustique
Dessus toit 31 413 361'-6"
Dessus dalle 30 498 358'-6"
Dessus pont roulant 26 764 346'-3"
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"]
6000 [19'-8 1/4"]
5411 [17'-9"]
3276 [10'-9"]
1143 [3'-9"]
Saturday, November 14, 2009



Saturday, November 14, 2009



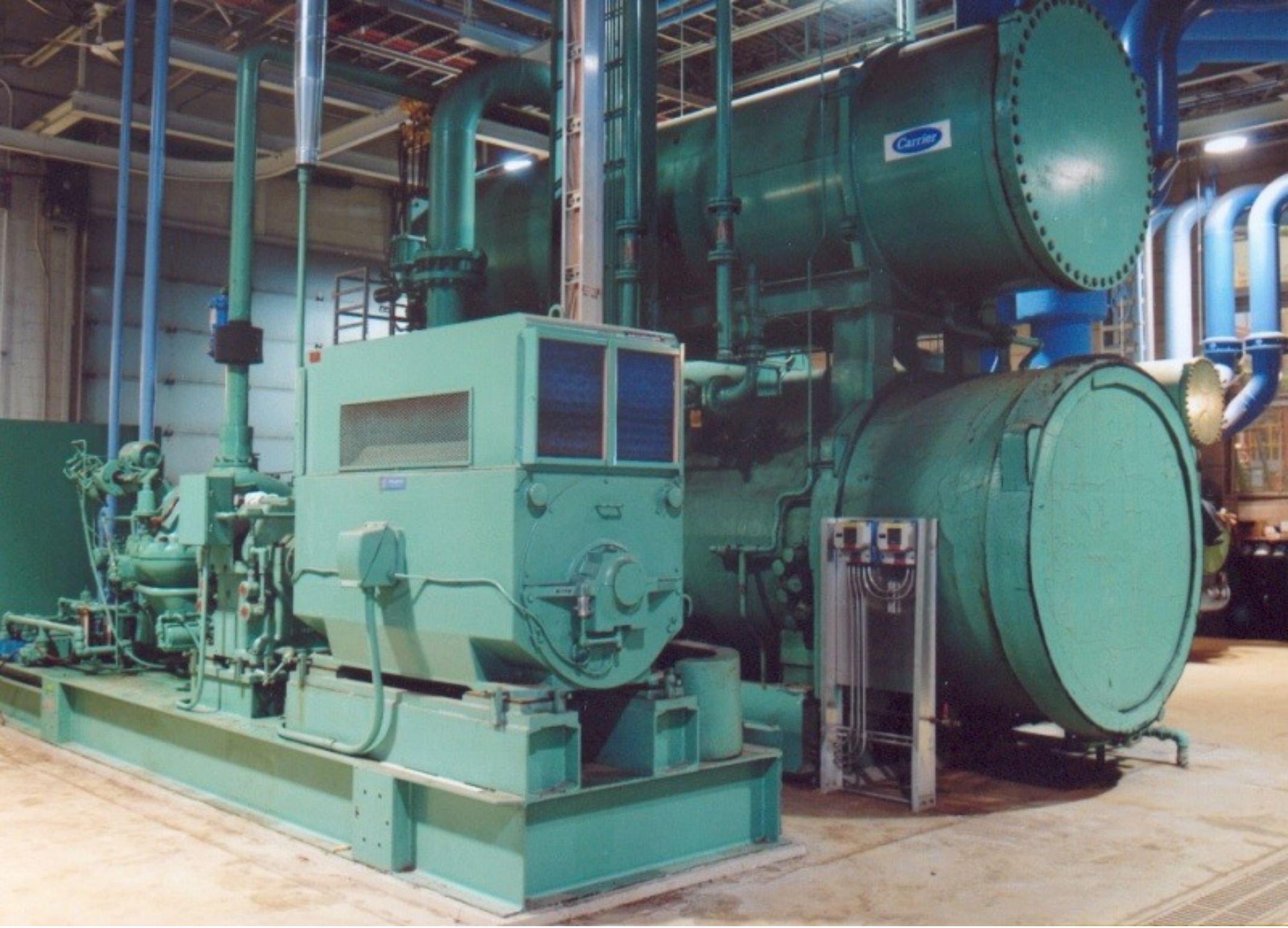
Main specifications

- 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

Energy efficiency

- Campus wide hot-water loop for heating
- Campus wide chilled-water loop for cooling
- Partiel free air cooling (up to 300 kW)
  - ✓ Preheating of outside air using computer heat
- Residual heat transfered to hot-water loop
- Controlled differential pressure in the cold air plenum

High efficiency gaz furnaces

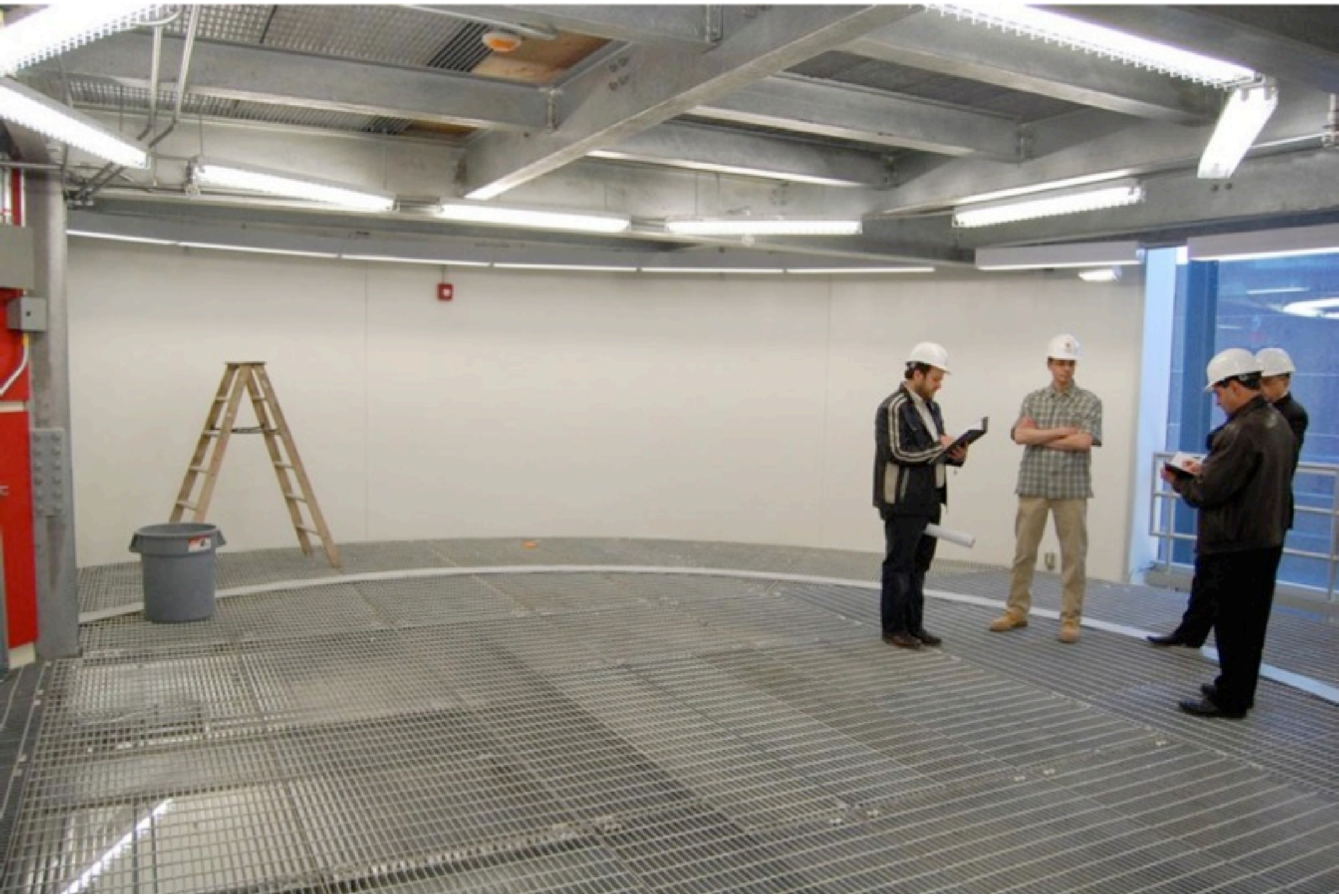


Saturday, November 14, 2009

Deployment process

First level - early stage

Second level - early stage



first level with floor plates to hold racks
also ~400kW of 208V-3P power

Second level with floor plates

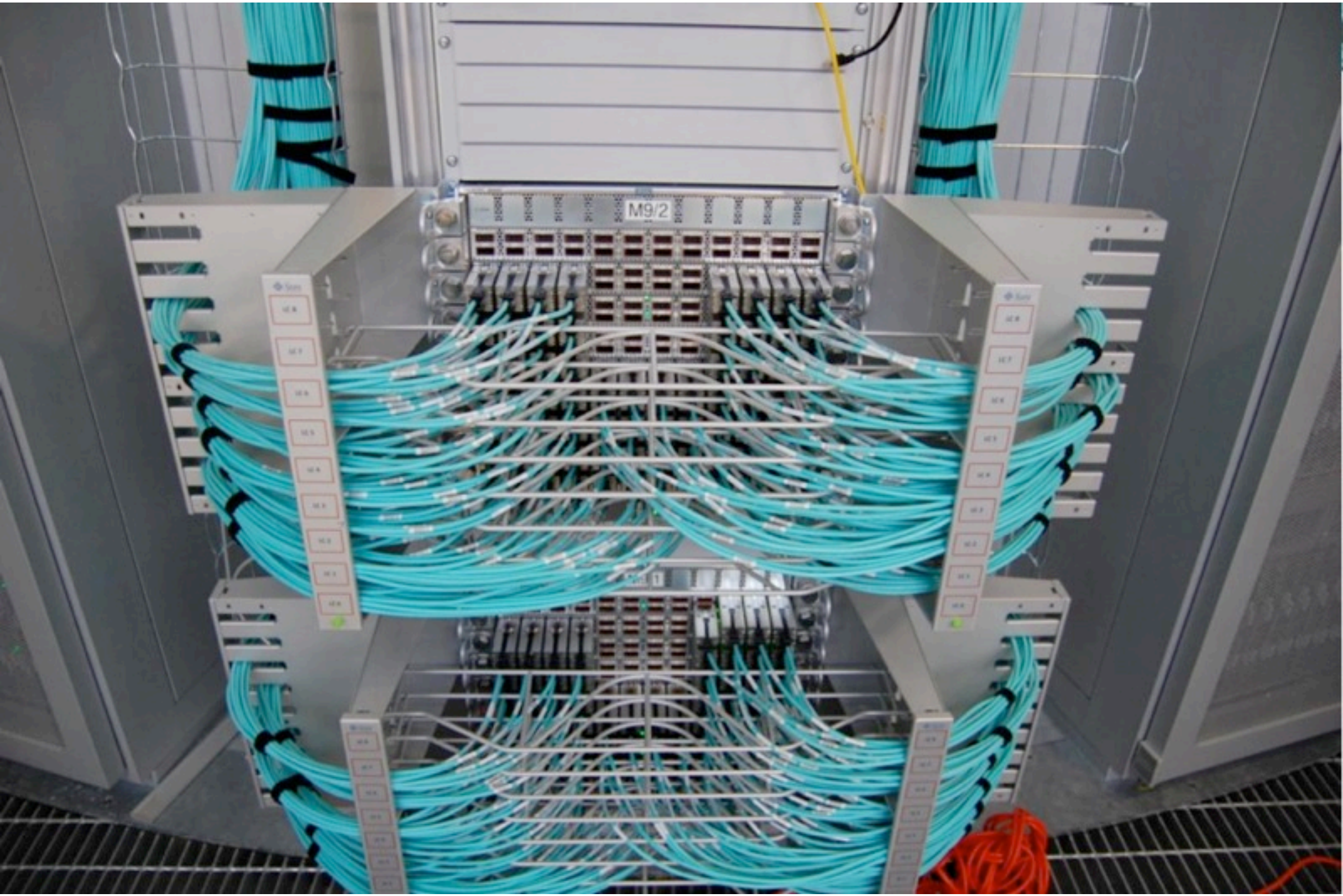
Empty third level

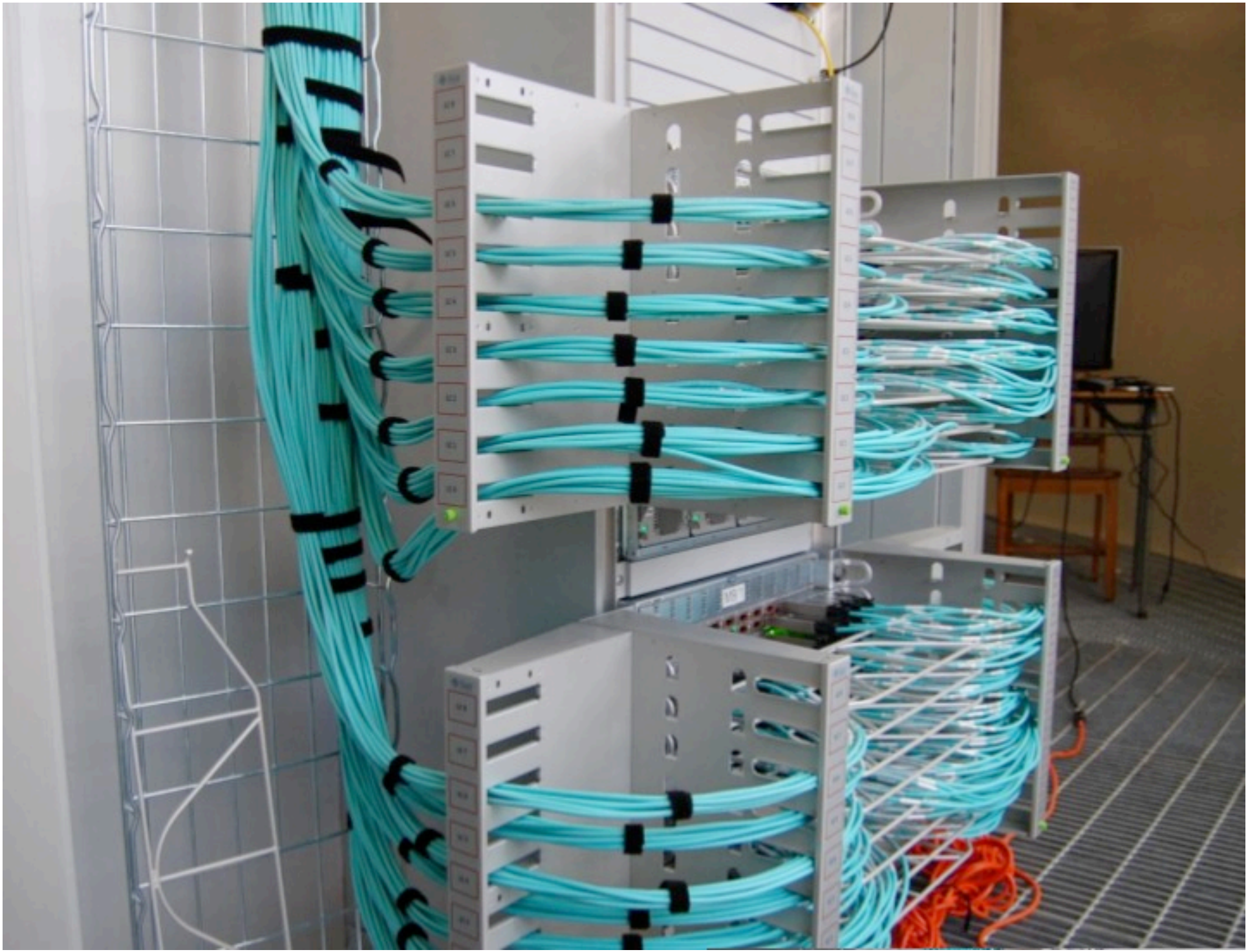
Infrastructure racks arrive on first level

Compute racks ready for IB cabling

First IB cable!

IB cabling complete





Saturday, November 14, 2009

Compute rack IB cabling

Compute racks complete

Demolition process

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009


Architect
project
manager

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Construction process

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Saturday, November 14, 2009

Conclusion

- A cool super-efficient cooling enclosure for a supercool supercomputer!
- Second largest computer in Canada
  - ✓ but highest performance computer in Canada, for real world applications that don't scale as well as HPL

Acknowledgments

- Silo design team
  - ✓ Daniel Denis, Alexandre Frenière (architects)
  - ✓ Denis Lemieux, Eng. (mechanical/electrical)
  - ✓ Lucien Viel, Eng. (structural)
  - ✓ Pierre Lévesque, Eng. (energy efficiency)
- Silo project Management
  - ✓ Jean l'Heureux (UL)

- Sun People:

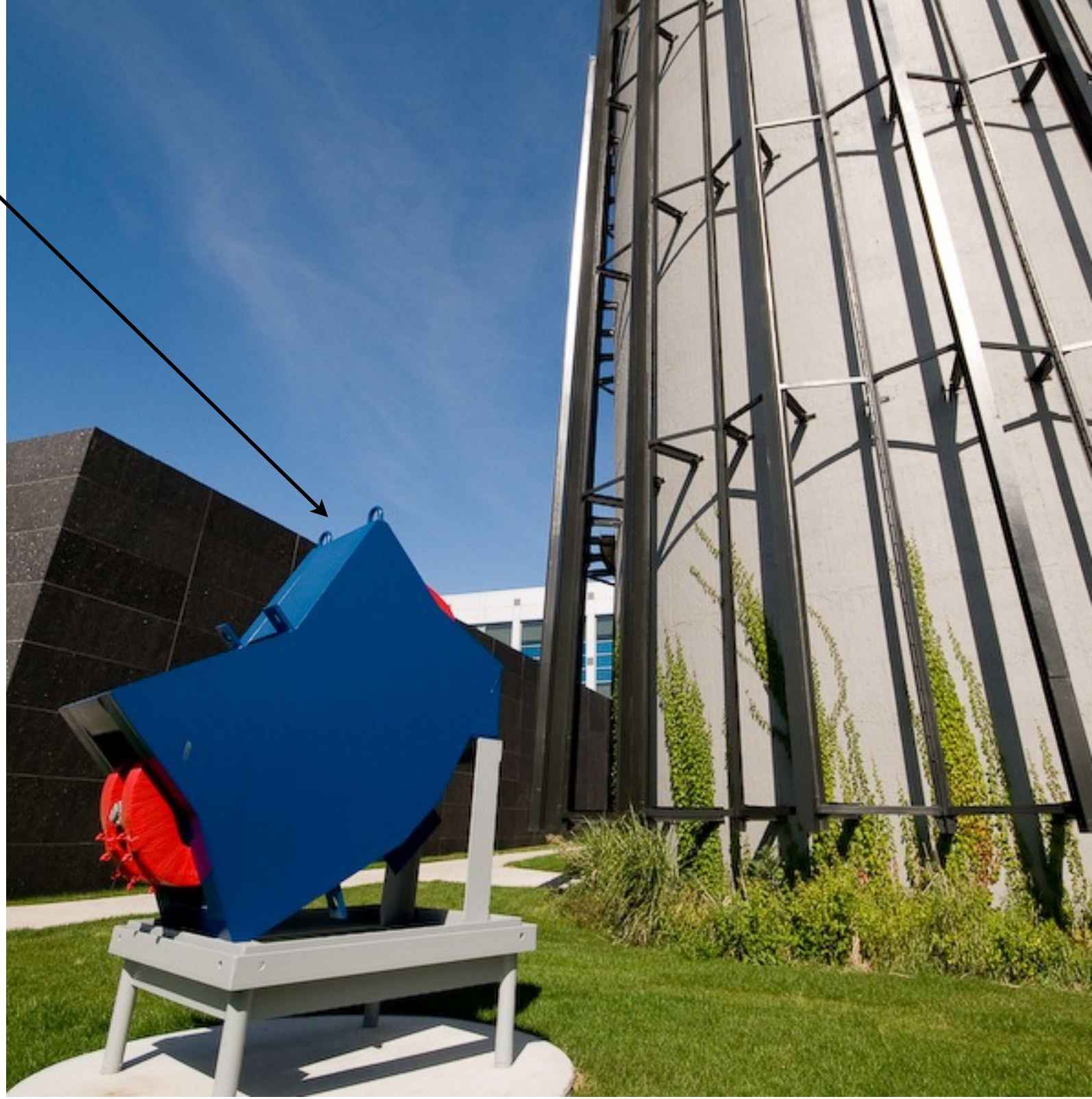
✓ Mario Cantin (CA)
✓ Nicolas Dubé (USA)
✓ John Fragalla (USA)
✓ Joey Jablonski (USA)
✓ Robert Lafosse (CA)
✓ Frank Leers (USA)
✓ Paul Lodrige (USA)
✓ Madhu Mahadevan (CA)
✓ Louis Rhéaume (CA)
✓ Michael Sims (CA)
✓ Matt Stanley (CA)
✓ Réal Trudel (CA)

- CLUMEQ staff:
  - ✓ Michèle L'Allier-Davies
  - ✓ Louis Demers
  - ✓ Jean Sébastien Landry
  - ✓ Frédéric Lefebvre
  - ✓ Jeffrey McDonald
  - ✓ Florent Parent
  - ✓ Vladimir Timochevski
  - ✓ Lee Wei Huynh
- CLUMEQ Director
  - ✓ Jorge Vinals (McGill professor of Physics)

October 17th,
2006

Nicolas Dubé,
at the time Ph.D.
candidate
at Université Laval

electromagnet
of the accelerator



Questions?