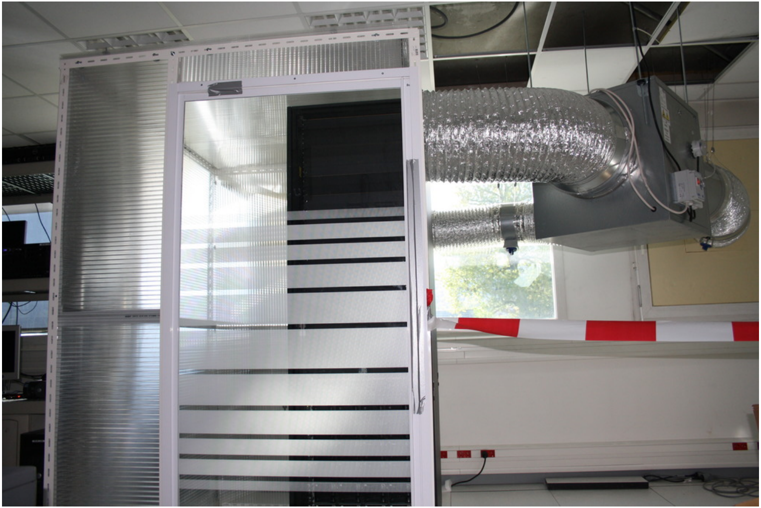
FrigID’R, extreme freecooling

Bruno Bzeznik, Olivier Richard, Pierre Neyron, Françoise Roch, Christian Seguy, Romain Cavagna

CIMENT, LIG

November 2012

FRIGID'R : free air-conditioning for supercomputer !



Outline

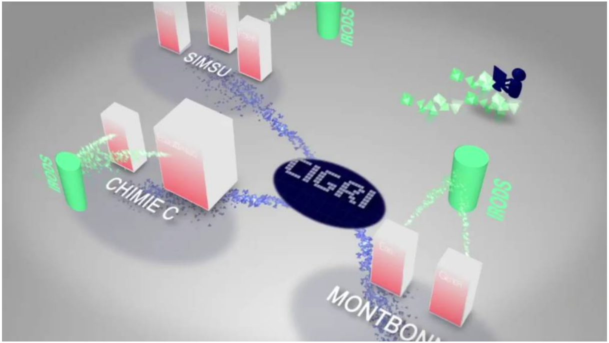
1. Context
2. Genesis
3. The project
4. Results

CIMENT

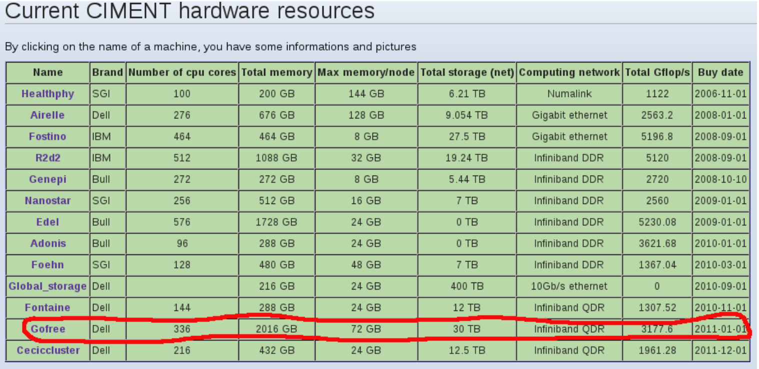
- CIMENT is the High Performance Computing (HPC) Centre of Grenoble University
- It provides researchers and engineers with an easy access to local HPC resources to develop and test their codes
- It is composed of about 3500 cpu cores (2012, 5500 expected in 2013) in a dozen of supercomputers

CiGri

- CiGri is the grid middleware aggregating the computing power of the supercomputers
- Its goal is to optimize the usage of the (free) resources with regard to multi-parametric applications



CIMENT Resources



Current CIMENT hardware resources

By clicking on the name of a machine, you have some informations and pictures

Name	Brand	Number of cpu cores	Total memory	Max memory/node	Total storage (net)	Computing network	Total Gflop/s	Buy date
Healthphy	SGI	100	200 GB	144 GB	6.21 TB	Numalink	1122	2006-11-01
Airelle	Dell	276	676 GB	128 GB	9.054 TB	Gigabit ethernet	2563.2	2008-01-01
Fostino	IBM	464	464 GB	8 GB	27.5 TB	Gigabit ethernet	5196.8	2008-09-01
R2d2	IBM	512	1088 GB	32 GB	19.24 TB	Infiniband DDR	5120	2008-09-01
Genepi	Bull	272	272 GB	8 GB	5.44 TB	Infiniband DDR	2720	2008-10-10
Nanostar	SGI	256	512 GB	16 GB	7 TB	Infiniband DDR	2560	2009-01-01
Edel	Bull	576	1728 GB	24 GB	0 TB	Infiniband DDR	5230.08	2009-01-01
Adonis	Bull	96	288 GB	24 GB	0 TB	Infiniband DDR	3621.68	2010-01-01
Foehn	SGI	128	480 GB	48 GB	7 TB	Infiniband DDR	1367.04	2010-03-01
Global_storage	Dell		216 GB	24 GB	400 TB	10Gb/s ethernet	0	2010-09-01
Fontaine	Dell	144	288 GB	24 GB	12 TB	Infiniband QDR	1307.52	2010-11-01
Gofree	Dell	336	2016 GB	72 GB	30 TB	Infiniband QDR	3177.6	2011-01-01
Ceciccluster	Dell	216	432 GB	24 GB	12.5 TB	Infiniband QDR	1961.28	2011-12-01

This presentation tells the story of "Gofree"...

Outline

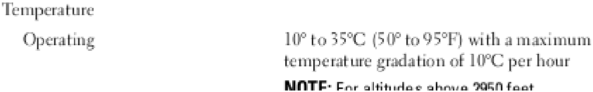
1. Context
2. Genesis
3. The project
4. Results

Some facts

- 2008 : Intel's free-cooling proof of concept : put 450 blade servers into a dusty free-cooling (pulsed air from the outside of the building) environement and compares the failure rates with 450 blades into conditionned and filtered air.
- 2008 : Ecoclim LPSC (IN2P3 Lab, Grenoble, France) : builded a datacenter using direct freecooling 85% of the year.

Some facts

- 2010 : Computers are more permissive with regard to operating temperature, for instance :



Temperature	
Operating	10° to 35°C (50° to 95°F) with a maximum temperature gradation of 10°C per hour **NOTE:** For altitudes above 2950 feet

- 2011 : New ASHRAE classes
- 2012 : "5°C to 10°C and 35°C to 40°C during 10% of the year"

Some facts

- In Grenoble, temperature is below 25 °C 85% of the year
- In Grenoble, temperature is below 32 °C 99% of the year
- We own a best-effort computing grid (CiGri)
- We turn off computing nodes when there's no job (OAR energy saving)
- Fact : A lot of energy is just wasted for cooling our supercomputers

Yet another HPC project

- Grenoble's observatory project for buying a supercomputer of 3TFlop/s
- But all of our datacenters have reached their thermal limits !
- One of our building has a small datacenter with a big electrical line, but no air conditionning

An idea

Extreme Freecooling

- Make an extreme freecooling solution :
    - no chilling system
    - if temperature is too hot, stop the computing nodes
- Handle resources which are only unavailable from time to time :
    - work on suspend/resume solutions to avoid killing the jobs
    - try and predict shutdowns thanks to weather forecast information.

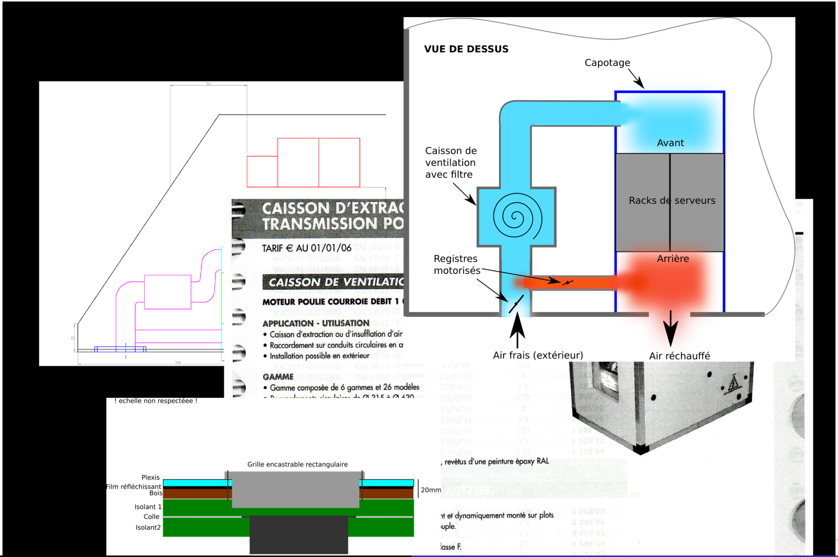
Open-minded researchers

- Q : Are you OK with the idea ? Do you accept computing capability cuts some days during summer ?
- A : Yes !

Outline

1. Context
2. Genesis
3. The project
4. Results

Study



CAISSON D'EXTRAC
TRANSMISSION PO

TARIF € AU 01/01/06

CAISSON DE VENTILATIC

MOTEUR POULIE COURROIE DEBIT 1 (

APPLICATION - UTILISATION

- Caisson d'extraction ou d'insufflation d'air
- Raccordement sur conduits circulaires en a
- Installation possible en extérieur

GAMME

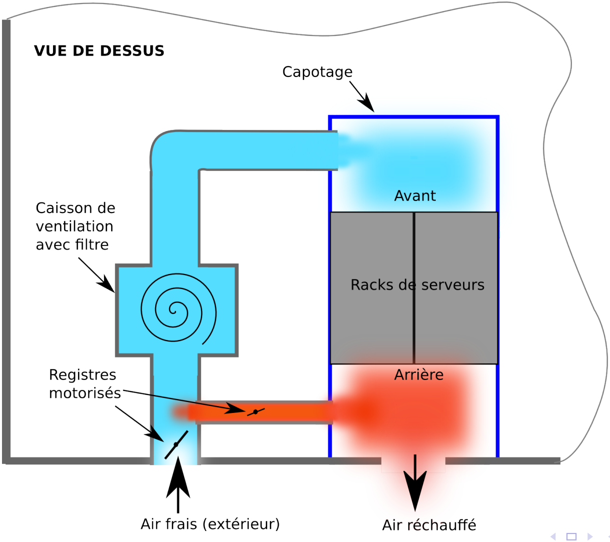
- Gamme composée de 6 gammes et 26 modèles

, revêtus d'une peinture époxy RAL

ıt et dynamiquement monté sur plots
ıuple.

lasse F.

Technical principle



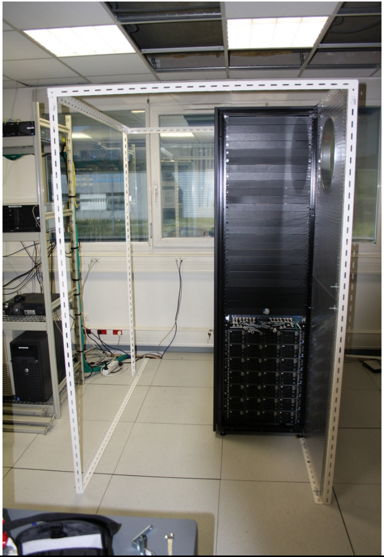
A simple DIY design

- Funding : less than **4000 euros TTC**
- FAN : 6000m3/h max, 800W max
- Engine variator
- Simple air filter that can be cleaned with water
- Monitored PDUS
- 2 electrical air-flow valves
- 1 arduino micro-controller to handle the valves
- Structure : perforated angles, polycarbonate panels
- 3 days of "meccano"
- Some electronics and scripting

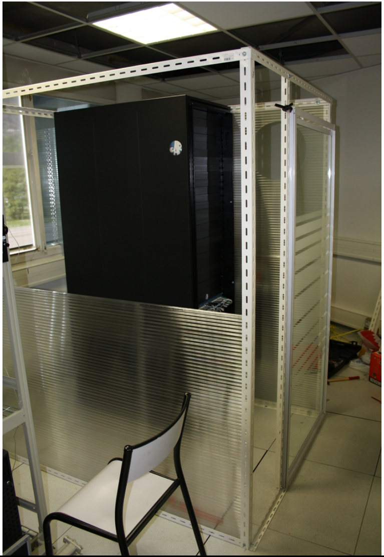
Automation

- Current version :
  - 4 temperature sensors + ipmi sensors of the chassis
  - Arduino to control the the valves
  - Scripts on the cluster's frontend to control the Arduino
- Work in progress :
  - a dozen of 1-wire temperature sensors
  - Autonomous arduino to control the valves

Construction

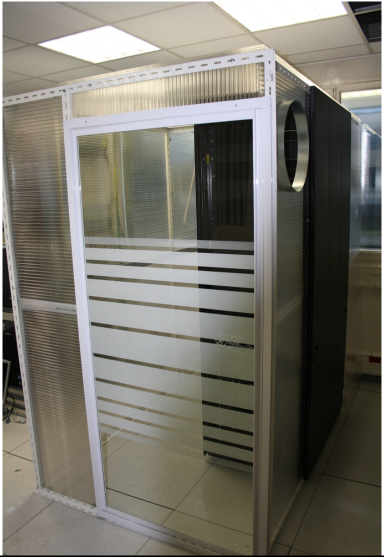


Construction



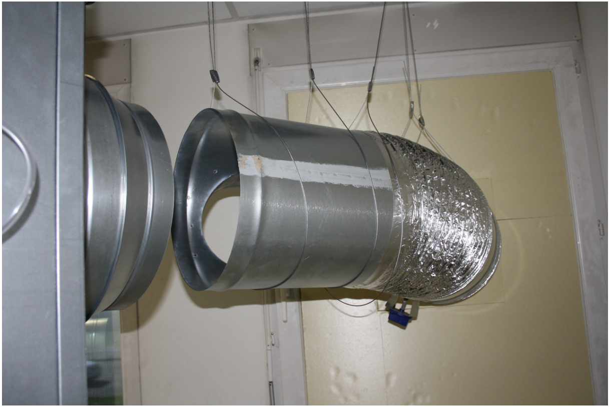
Construction

Construction



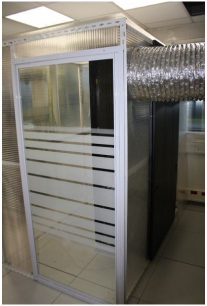
Construction

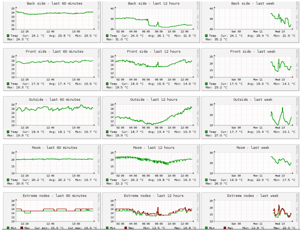
Construction



Construction

Et voila !




Back side - last 60 minutes
Temp Cur: 24.1 °C Avg: 23.8 °C Min: 23.5 °C Max: 24.3 °C
Back side - last 12 hours
Temp Cur: 24.0 °C Avg: 26.1 °C Min: 21.6 °C Max: 31.0 °C
Back side - last week
Temp Cur: 24.1 °C Avg: 29.4 °C Min: 21.9 °C Max: 35.2 °C
Front side - last 60 minutes
Temp Cur: 17.9 °C Avg: 17.4 °C Min: 16.6 °C Max: 18.3 °C
Front side - last 12 hours
Temp Cur: 18.0 °C Avg: 16.6 °C Min: 14.9 °C Max: 18.5 °C
Front side - last week
Temp Cur: 17.5 °C Avg: 18.0 °C Min: 14.1 °C Max: 23.2 °C
Outside - last 60 minutes
Temp Cur: 18.4 °C Avg: 18.1 °C Min: 16.7 °C Max: 19.9 °C
Outside - last 12 hours
Temp Cur: 18.7 °C Avg: 13.4 °C Min: 10.3 °C Max: 18.9 °C
Outside - last week
Temp Cur: 17.7 °C Avg: 15.4 °C Min: 10.1 °C Max: 27.0 °C
Room - last 60 minutes
Temp Cur: 20.2 °C Avg: 20.2 °C Min: 19.7 °C Max: 20.6 °C
Room - last 12 hours
Temp Cur: 20.2 °C Avg: 19.8 °C Min: 16.3 °C Max: 22.2 °C
Room - last week
Temp Cur: 19.9 °C Avg: 22.0 °C Min: 17.5 °C Max: 26.0 °C
Extreme nodes - last 60 minutes
Min Max Cur min: 15.0 °C Cur max: 16.0 °C
Extreme nodes - last 12 hours
Min Max Min: 12.5 °C Max: 16.8 °C
Extreme nodes - last week
Min Max Min: 12.8 °C Max: 22.0 °C
12:20 12:40 13:00
02:00 04:00 06:00 08:00 10:00 12:00
Sat 09 Mon 11 Wed 13

Shutdown of the computing nodes

- Shutdown if :
  - Motherboard temperature of the hotest node is above **46**°C
  - OR host room temperature is above **35**°C
- Restore power if :
  - Motherboard temperature of the hotest node is below **28**°C
  - AND host room is below **33**°C
- Manual actions to minimize interruptions : slow down processors and prevent besteffort jobs when we are close to the limits
- But not really effective as the temperature of the computers depends more on the outside temperature than the load

Outline

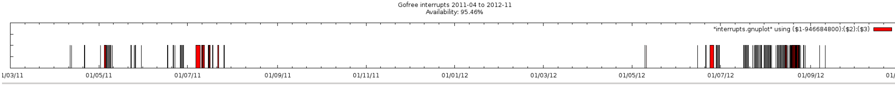
1. Context
2. Genesis
3. The project
4. Results

Availability

- System up and running for 19 months now (since April 2011)
- **95.46**% availability, while taking into account :
  - the tests during the first 2 months
  - the shutdowns for the maintenances (2 days work for to improve the isolation (sillicon) during 2011 summer !)
  - 2 summers and only 1 winter periods
- Estimated availability for 2 years of operation (April 2013) : **96.4**% ! !

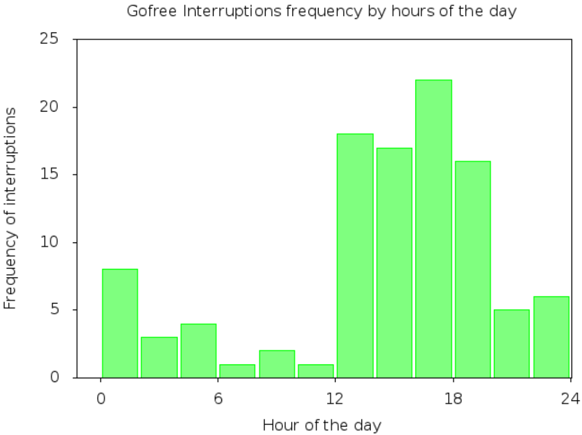
Interruptions

- **103** interruptions in 19 months
- **75** days with at least one interruption
- BUT : most of the interruptions are due to the host room temperature (remember 35˚C)
- Average downtime duration : **6** hours
- Event distribution during the year :



Interruptions

- Interruptions are predictible : mostly on afternoons



PUE

- Average electrical power of the FAN on a year : 524W
- Maximum Mesured IT power : 7098W
- Average mesure IT power : 3633W
- PUE between **1.08** and **1.14**
- (A better PUE may be obtained with a better FAN variator)

Troubles and solutions

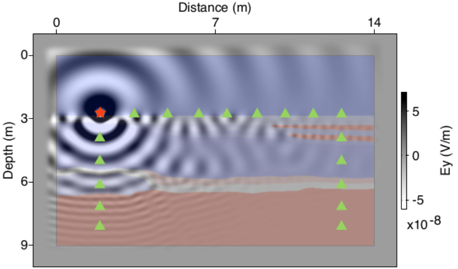
- Air-tightness : tap is not good, silicon is ok
- Neutrality for the hosting room : avoid installation inside a cooled datacenter !
- Suspend/resume of infiniband network cards : IB comms are lost on resume... no solution for now except checkpoint.
- Pollen in May : have to clean the filter once a week at some times ; easy with a mosquito net
- Size of the holes of the windows : too much pressure inside, limits the air flow ; have to enlarge the holes.

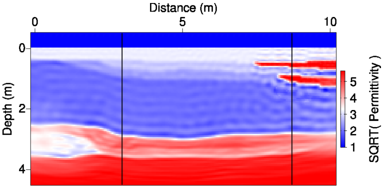
Users feedback

- "During this summer, I didn't compute ; I explored other research fields waiting for Gofree to wake up”
- "The operational mode was ok for me. I don't have checkpointing into my code, so I anticipated the availability by using weather forecasts and CIMENT graphs to know when to start my jobs.”
- "During the hot days of this summer, I adapted myself, using another CIMENT supercomputer. I used Gofree only in the morning when it was available for smaller jobs”
- "It's not hard to deal with periods when the computer is not available as this is well focalized in time. During most of the year, the difference with a cooled supercomputer is indiscernible”

Thanks !

Scientific results, computed on Gofree (Electrical field emitted by a GPR antenna, simulated on the ground of the campus of Grenoble)


Distance (m)
Depth (m)
x10^{-8}
Ey (V/m)


Distance (m)
Depth (m)
SQRT(Permittivity)