An Energy Efficient Datacenter in Orsay

Michel Jouvin
LAL, Orsay
jouvin@lal.in2p3.fr

HEPiX, Bologna, April 2013

Outline

- Project background
- Early phase: convincing lab directors
- Feasibility and design studies
- Datacenter status
- What next?
- Lessons learnt and conclusions

Background: P2IO Initiative

- P2IO: a common initiative in 2009 by all HEP, NP and Astrophysics labs in Orsay-Saclay
  - Physique des 2 Infinis et des Origines
  - Includes 8 labs, including CNRS/LAL and CEA/Irfu
    - Many pre-existing collaborations between these labs
  - Goal 1 : foster synergies between each others
  - Goal 2 : make fundamental physics visible in the new emerging Paris Saclay University
    - May be a funding opportunity…
- P2IO built around several technological platforms
  - Electronics, cryogenics, accelerators, space technologies…
  - VirtualData = computing
    - Simulation, storage and analysis of large data volumes

P2IO VirtualData WG

- Only P2IO WG with all the 8 labs involved in
  - VirtualData : virtualization + real data!
- Goal since 2009: build a computing expertise network around a shared computing platform
  - 130 people involved in computing
    - Development (~75) and Operations (~55)
    - Covering all main areas of expertise in computing
    - Strong links with computer sciences at Univ. Paris Sud
  - Shared computing platform hosted in 2 shared facilities to enable redundancy when needed
    - Use cloud expertise gained with StratusLab
- Several failed attempts to get funding for building these shared facilities (2010, 2011)

P2IO Computing Resources

- GRIF: a 8-year old successful experience of a multi-lab experience to build and manage a unique grid site
  - 4 (of 6) GRIF partners involved in VirtualData
    - ¾ of GRIF resources located at VirtualData partners
    - GRIF consolidated resources: 8000 cores, 4 PB of disk
  - Resources are distributed in each lab
  - Asset: confidence between people used to work together
- StratusLab: a small production IaaS cloud
  - Currently 500 cores, 50 TB
  - Growing number of users at Univ. Paris Sud and outside
  - GRIF may bring its resources if we succeed to implement grid services as one of the cloud service
- Computing resources for astrophysics, in particular IDOC
  - 2000 cores, 2 PB of disk, high availability (data) services

2011/12 : Context Changes

- Paris Saclay University recognized that there was not enough money to move all Orsay labs to Saclay
  - Until then, impossible to make any significant project in Orsay part of the future university (Univ. Paris Sud)
- Computing facility crisis in the 3 main Orsay labs
  - LAL: major cooling incident end of 2010, need of major work to restore normal cooling conditions for the long term
    - ~150 k€ of infrastructure work needed in 2013
  - IAS: hosting of new computing resources required in 2013-14 but no space to do it
    - ~100 k€ of infrastructure work required in 2013-14
  - IPNO: cooling problems similar to LAL became critical late 2012
    - ~100 k€ of infrastructure work required
- LABEX P2IO ready to participate to funding: ~250 k€

P2IO Map


CEA/Irfu
Plan
Trafic
LLR
IMNC
LAL
IAS
IPNO +
CSNSM
LPT
500 m
2000 pieds

Early Project Phase…

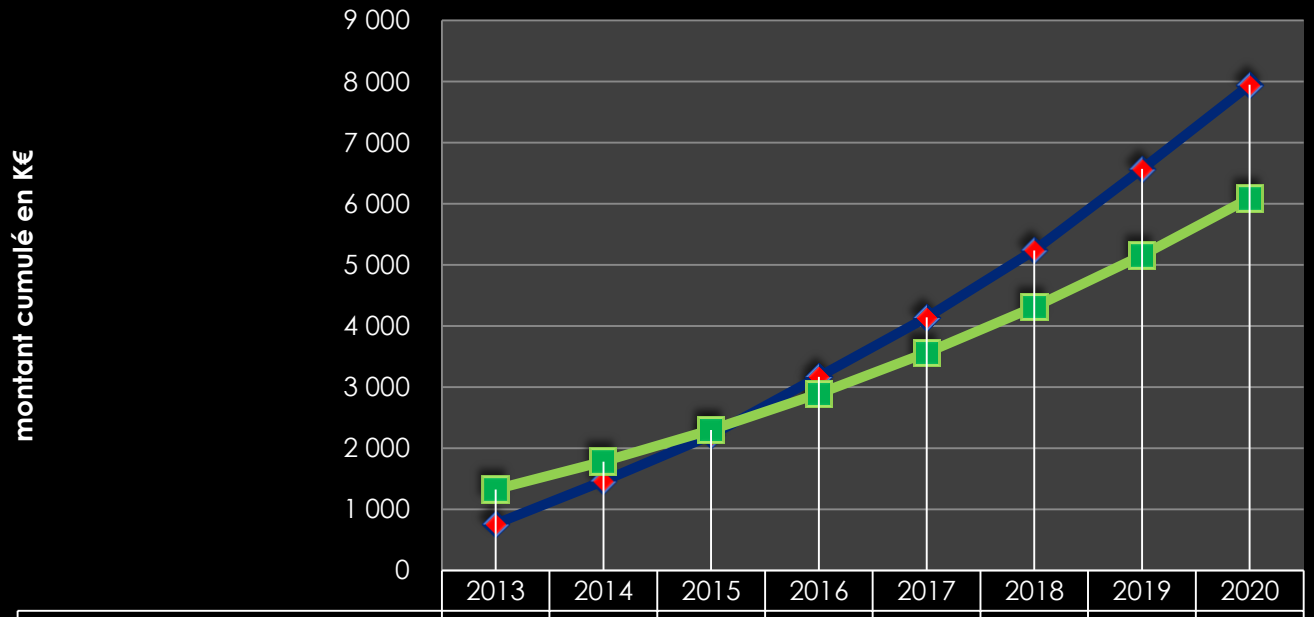
- P2IO directors convinced end of 2011 that it was the right time to bootstrap a new computing hosting facility in Orsay
  - Ready to spend the money planned for local needs in a common project
  - Ready to look for additional funding to make a common project feasible
- 2 early design requirements
  - Build a new facility designed around high energy efficiency
    - Save on operation costs to enable some return on investment
  - Design a modular facility that can start small and grow according to the needs without significant increase of consolidated cost

… Early Project Phase

- 2 existing buildings identified as potentially suitable in Orsay
  - 1 "technical building": 2 storey building, used to host electrical installations (transformers) at the ground floor, racks at 1st floor…
    - 220 m2 per floor
  - Ground floor area of an office building
    - 400 m2 on one level
- Budget naively estimated to 800 k€ (without VAT)…
  - ~100 m2 appropriate for high density for the first phase
    - Initial goal : replace main existing problematic computing rooms
  - ~8 k€/m2 to refurbish an existing building into an efficient computing facility
    - Electricity outlets and cooling
- Target date: 1st October 2013! 18 months left…
  - Latest possible date for avoiding local investments in labs

Expected RoI

Consolidated operational costs comparison with and without the new room (old rooms PUE = 2, new room PUE = 1,3)

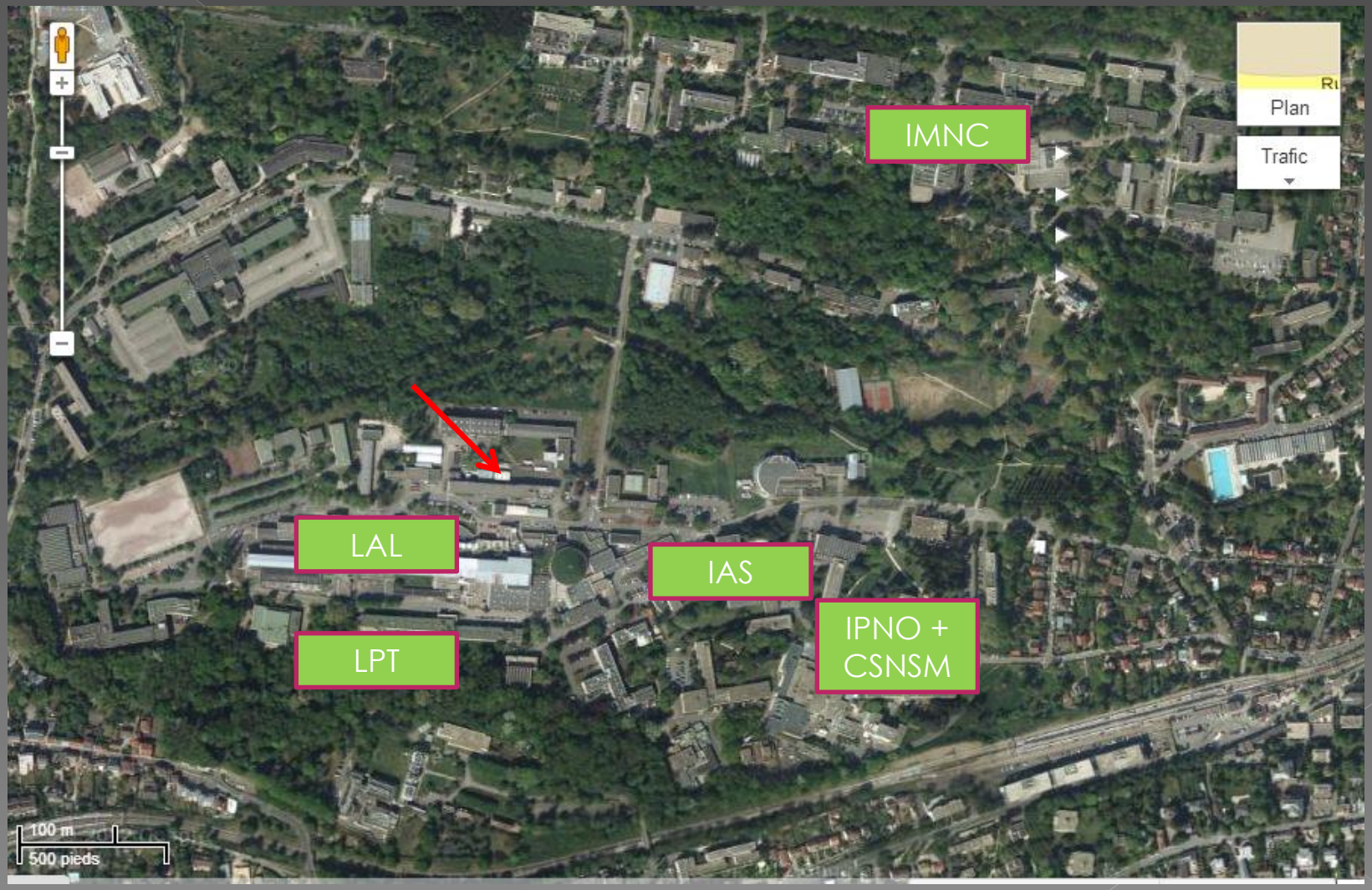


	2013	2014	2015	2016	2017	2018	2019	2020
coût cumulé en l'état avec réinvestissement pour poursuite activité	760	1 461	2 211	3 167	4 140	5 237	6 570	7 949
Coût cumulé avec mutualisation	1 325	1 780	2 300	2 892	3 560	4 312	5 153	6 092

Feasibility (May-June 12)

- April 2012: datacenter consultants (APIS) hired to assess feasibility of our project in the foreseen budget/timeline
  - Design proposal (APS) for reaching a PUE=1,3
    - Based on existing projects, seemed ambitious but feasible
    - GSI-like PUE (~1) seemed too difficult to reach in our context: lack of expertise, too constrained budget, too much R&D
  - Study based on the 2 proposed locations with recommendations/criteria for choice early June
- Design proposal: water-cooled racks only with cold water produced by chillers + free-chilling units
  - PUE simulation based on last 5 years weather data confirmed that PUE=1,3 was realistic
  - No UPS: highly available HV feed (redundant feeds)
- Minimum budget : ~850 k€

Chosen Building



Detailed Design (Summer 12)

- New datacenter consultants hired beginning of July to act as "project manager" until the end of the project
  - Critical Buildings
  - 2 storey technical building chosen
- Capacity final target defined to 1,5 MW in 220 m2
  - Designed for the "next 10 years"
  - 84 "high density" racks, 50% computing, 50% storage
    - Computing rack: up to 30 kW and 1T
    - Storage rack: up to 15 kW and 1,5 T
- Initial phase : 100 m2, 400 kW
  - 28 racks up to 15 kW
    - Existing resources are less dense than target
  - 2 water chillers **required**
    - Water chiller problems are the main source of unavailability

Detailed Design (Summer 12)

- New datacenter consultants hired beginning of July to act as "M
  -
  -
- Ca
  -
  -
- Initi
  -
    - Existing resources are less dense than target
  - 2 water chillers **required**
    - Water chiller problems are the main source of unavailability

Vue global du modèle de la Salle 206

Detailed Design (Summer 12)

- Initial design refined with modularity in mind
  - Based on multi-feed concept: up to 6 "technical poles" providing 300 kW IT of electricity and cooling each
    - Each pole has a HV switch gear, a HV transformer and LV distribution panel
    - 1 water chiller attached to each electrical feed
    - 6 chosen because of the building layout: 6 transformer rooms (30 m2)
  - N+1 configuration providing a 2N configuration for machines
    - Both for electricity and cooling
  - Phase 1 : 1 pole with an oversized electrical capacity (400 kW) + 2 water chillers
    - 1 chiller will be moved to the second pole when built
  - Multiple path from initial to target capacity
- Coasting also refined: 900 k€ minimum!
  - Including 100 k€ of studies and consulting
  - Final target estimated ~2,8 M€
  - Green light given by directors to do the call for tender mid-oct.

Call for Tender

- Prepared in October 2012, published mid-November
  - Offers to be received by January 7
  - Criteria: 60% technical value, 40% price
  - 4 different lots: structural/finishing work, electricity, cooling, raised floor
- Many offers received for each lot
  - Record: 11 offers for the electricity lot…
  - Several offers with a cost below the estimates
    - But wide range of prices for almost each lot
- 1 month negotiation with the 2 best offers for each lot
  - End result: total cost inline with estimates
  - Medium-size companies: more flexibility for this medium-size but time-constrained project
- Enterprise chosen on Feb. 22, work started March 15…

Construction Phase

- In fact started by a significant demolition phase!
  - Completed in 3 weeks
  - Main source of uncertainties when refurbishing an existing building…
- No major issue found except (unexpected) asbestos presence in a small room
  - Despite worries, should not impact the overall planning by more than 2 weeks
  - Last uncertainty to be cleared next week after digging out a trench (outdoor) for water pipes
    - Risk of unknown presence of an "obstacle"
- End of work expected end of July, 1st rack in October!
  - August is off in France, September to fix last issues


HILTI

Moving in

- Water-cooled racks needed to move in existing resources
  - Room designed for racks with rear door heat exchangers
    - Due to its form factor
  - Not yet decided between active and passive rear door
    - Overall consumption should be the same even if PUE is not…
    - Probably a mix: active for racks > 15 kW
    - Emerson/Knurr is the only rack maker proposing passive racks -> 30 kW
  - Passive or active: must be regulated with monitoring/alarming capabilities
- PDU: 2 32A (3 phases) in each racks by default
  - ~45 C13 outlets, monitored, manageable by blocks (PDU subsets)
- Planned over 6 months
  - Partly for budget reasons: racks funding over 2 years

The Other Room…

- Short term: second room planned at Ecole Polytechnique
  - LLR is already part of GRIF
  - Much smaller: ~70 m2 for P2IO in a 200 m2 room
    - Not extensible
  - Critical for implementation of highly available services
    - Will allow to start playing with 2 different rooms…
- Driven very differently than the Orsay datacenter project…
  - Driven by Ecole Polytechnique DSI
  - Still not clear after 2 years what they want to do
- Still hope to have something ready by the end of the year

… The Other Room

- Long term solution not yet clear
  - Depends on the Campus Paris Saclay
    - 23 partners: universities, engineering schools, research centers…
  - A WG about to start to discuss scientific computing needs for the new university
  - Hope to convince partners of doing something ambitious along the P2IO project ideas
    - 2 or 3 well interconnected shared computing facilities spread around the campus
    - High availability by distributing services around rooms

Lessons Learnt (so far..)

- Importance of open-minded, creative datacenter experts
- Importance of infrastructure specialists at the lab
  - Technical details about the building are critical for the design
- Importance of collective work: project managed by a group of ~10 people with computing expertise
  - ~1 meeting every 2 weeks since 1 year!
- Tight links between everybody: short iteration cycles
  - Learn from each others, be prepared to tackle new issues!
- Keep the project at the right size!
  - Avoid unnecessary complexity: always a temptation
- Reuse experience!
  - HEPiX was a source of inspiration…

Conclusions

- P2IO VirtualData: a new horizon for LAL computing
  - LAL very much involved in the VirtualData WG
  - LAL will move its machine room to the new shared facility
- The new energy efficient datacenter is (should be!) an asset for the computing of our labs
  - Lower cost of computing operations
  - Better resiliency than a single lab can afford
- Main challenge is still ahead of us: build a unique platform managed by people from different labs
  - Importance of links with users: don't want to create computing division independent from the labs
  - A first governance draft has been proposed and is being discussed