Energy efficiency upgrades at PIC

HEPiX Fall 2015 at BNL, USA
12-16 October 2015

J. Flix*

** PIC Tier-1 project coordinator*

Institut de Física d'Altes Energies

Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Energy efficiency upgrades at PIC

HEPiX Fall 2015 at BNL, USA
12-16 October 2015

J. Flix*

On behalf of M. Delfino, V. Acín, A. Herbera, J. Hernández

PIC port d'informació científica

BROOKHAVEN NATIONAL LABORATORY

HEPiX

EXCELENCIA SEVERO OCHOA

Institut de Física d'Altes Energies

IFAE

Ciemat

Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

PIC computing centre

Port d'Informació Científica (PIC) is the largest Grid centre in Spain, supporting **research** involving **analysis of massive sets of distributed data**

It provides computing services for many **applications**

- host the **Spanish WLCG Tier-1 centre → ~85% of resources**
  - *** Offer 5.1% of Tier1 computing resources for ATLAS, CMS and 6.5% for LHCb**
- host resources of the Spanish federated **ATLAS Tier-2**
- provides an **ATLAS Tier-3 facility**

~5500 cores (~65 kHS06)
~6 PB disk
~12 PB magnetic tape

~25 Kms from Barcelona
Autonomous University of Barcelona

Energy savings: motivations

- We were (are) worried on the **energy consumption increases** at the site
  - We use the main IT room in the University Campus, where the University IT equipment is sitting (as well as our equipments)
  - The University covers most of these energy expenses
    - As of today, the electricity bill is basically dominated by <u>**our**</u> electrical costs...
- All the cooling elements, UPS, and the main IT room, <u>**needed a revision**</u>
  - **Old** elements (even difficult to find spare parts!)
  - **Inefficient** elements
  - **Not originally designed** (80s) for today's needs

PIC
port d'informació
científica

Historical actions taken...

- **2003:** PIC initial installation. IT from 40 to **200 KVA –** Minor modifications on the cold water cooling system
- **2008:** Major upgrade of power input (2 transformers, **1 MVA** each) and diesel backup (**4x500 KVA** generators)

→ **PUE: 2.3** (25 years-old chillers [+inefficient] in basement required 100 kW power to 'move' air)

- **2010:** <u>**two new chillers**</u> installed on the roof of the building

→ **PUE: 1.7** (high enough to think on doing some actions)

- **2010:** build of a <u>**compact IT room**</u>, handling 80 KVA of IT equipment. Cold/hot air separated, gas expansion air conditioning system

→ **PUE: ~1.6** (allowed to host part of the equipments during the 2014-2015 main room upgrades)

- **2013-2014:** <u>**adjust**</u> the PIC farm power to electricity cost
  - → Less CPU during high cost periods, and vice-versa, keeping annual pledges OK
  - → Reduction of electricity bill is **~10%**

(main) works in 2014 - 2015

- Modifying the already existing(s) IT room(s), as well as the introduction of **free cooling techniques**, towards establishing a sustainable site infrastructure

Simulations using historical weather data show that weather condition would allow some amount of free-cooling up to **6000 hours per year**

- The upgrades aimed to target **several distinct areas**:

- main IT room: segregation of hot/cold **air zones** - horizontal layout

- Introduction of **free-cooling** techniques

- Introduction of **new UPS systems** based on IGBT technology

- **Monitoring/control** tools for the new cooling and UPS systems

- (today) **Oil-immersion techniques** for future CPU purchases (new design for the compact IT module)

Free-Cooling at PIC

→ 2014: 15 weeks of work, w/o any major downtime, interruption and negative impact in Ops

Before:

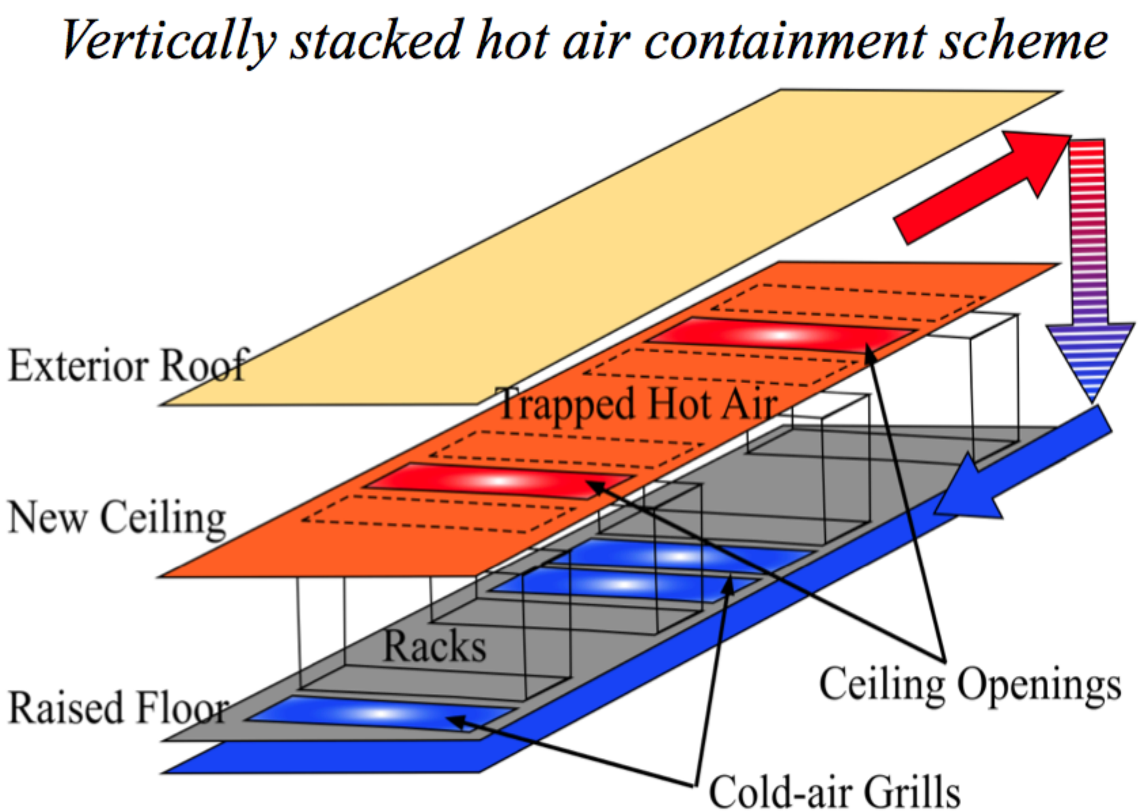
- No separation of cold/hot air in the main room (300m^2)
- Several CRAH's (Computer Room Air Handler) managing the air through a cold water battery, injecting air at 13°C to get a room temperature of 22-23°C (*inefficient*)
- PUE (Power Usage Effectiveness) was about **1.7**

After:

- IT room CRAH's replaced by **3 free-cooling units**: indirect heat exchangers with outside air and equipped with adiabatic cooling humidifiers
- Chilled water backup system integrated: regulated chilled water serpentine to bring air to the set point, as well as spray water drops to lower <5°C (if needed)
- Implemented separation of hot and cold flows in the room – no rack was moved
- Hot aisle containment and confinement + installation of ceiling to contain the hot air
- Increase of inlet temperature according to the ASHRAE recommendations
- Installation of dedicated monitors (ModBus gateways to Nagios/Ganglia) + control
- Aimed PUE in the range **1.45-1.3**

Trapping hot air in the ceiling

A **false ceiling** spanning the 400 m^2 of the entire computer room was built, creating a volume where hot air is trapped and guided towards the cooling units



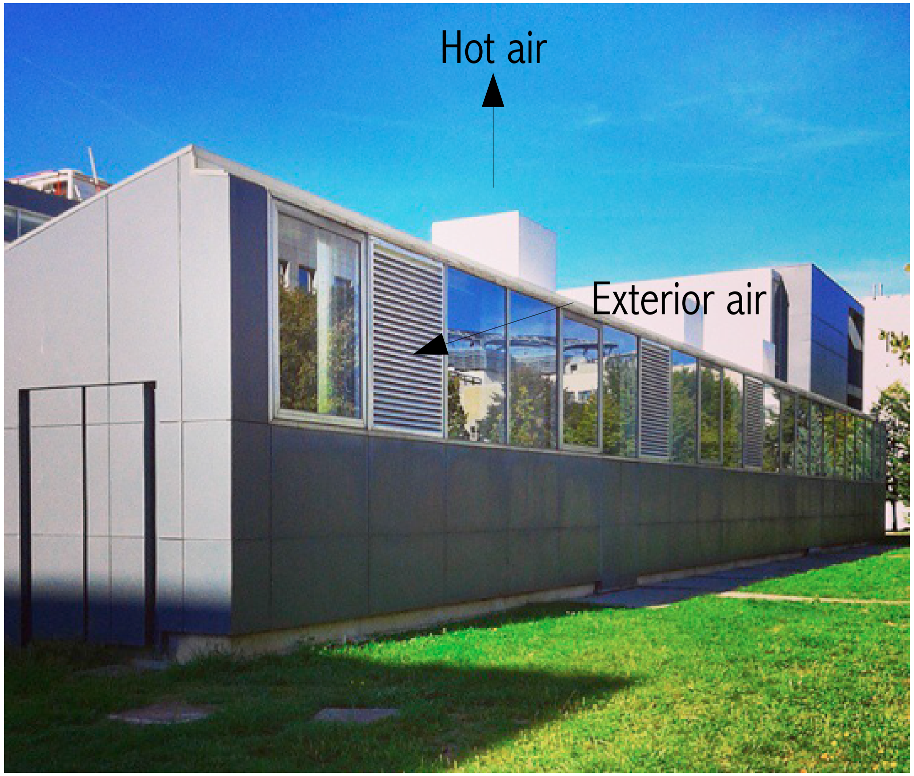
PIC port d'informació científica

Trapping hot air in the ceiling

The gap between the top of the racks and the false ceiling, as well as the vertical space at the edge of rack rows, is closed by plastic lame curtains hanging from the ceiling in order to delimit the hot air aisles

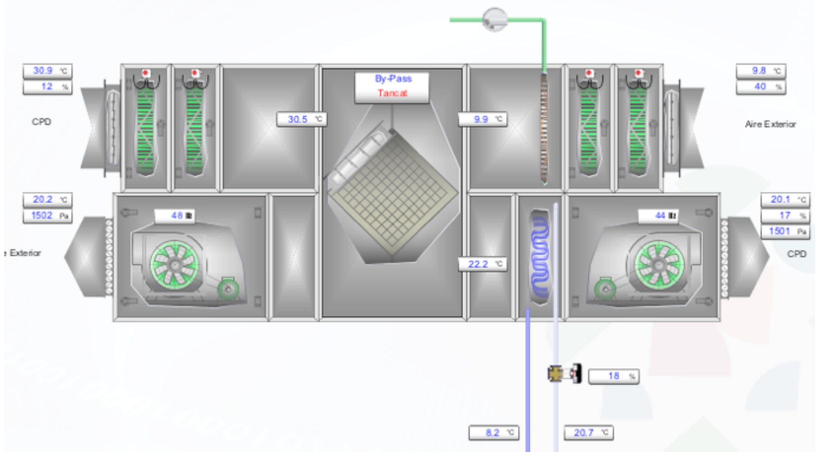
Free-Cooling at PIC

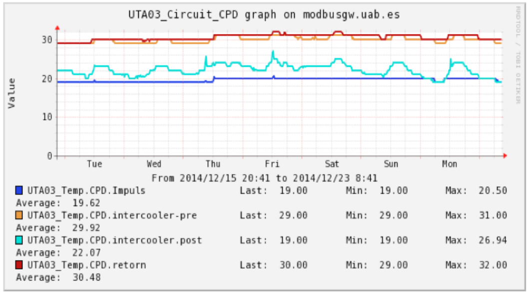
Installation of free-cooling units



New technical area

Free-cooling unit control/monitoring

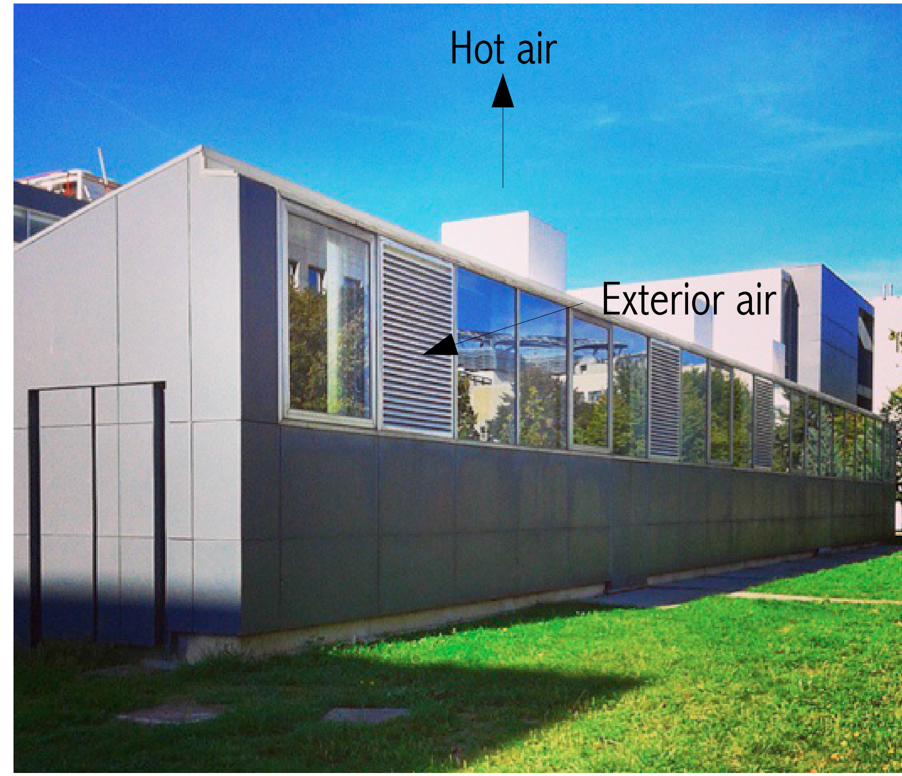




rdd graphs

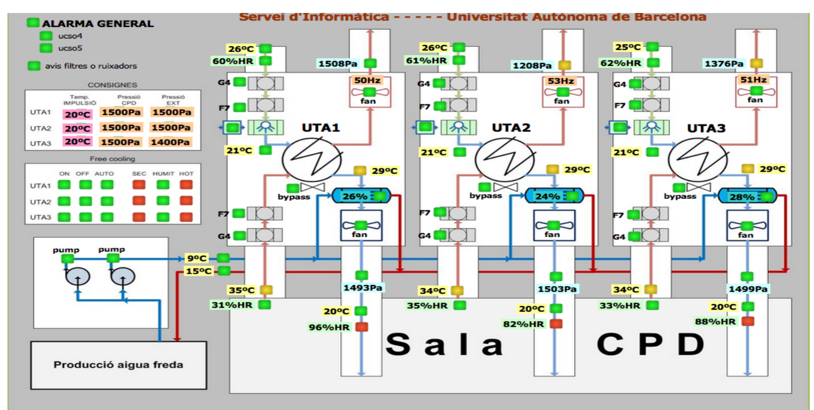
Free-Cooling at PIC

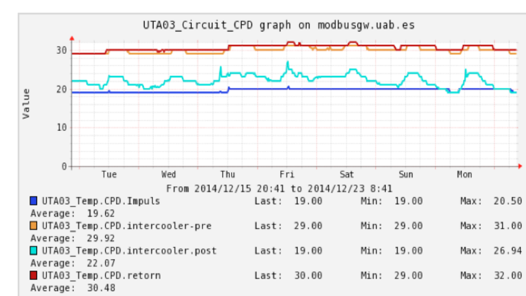
Installation of free-cooling units



New technical area

Free-cooling unit control/monitoring



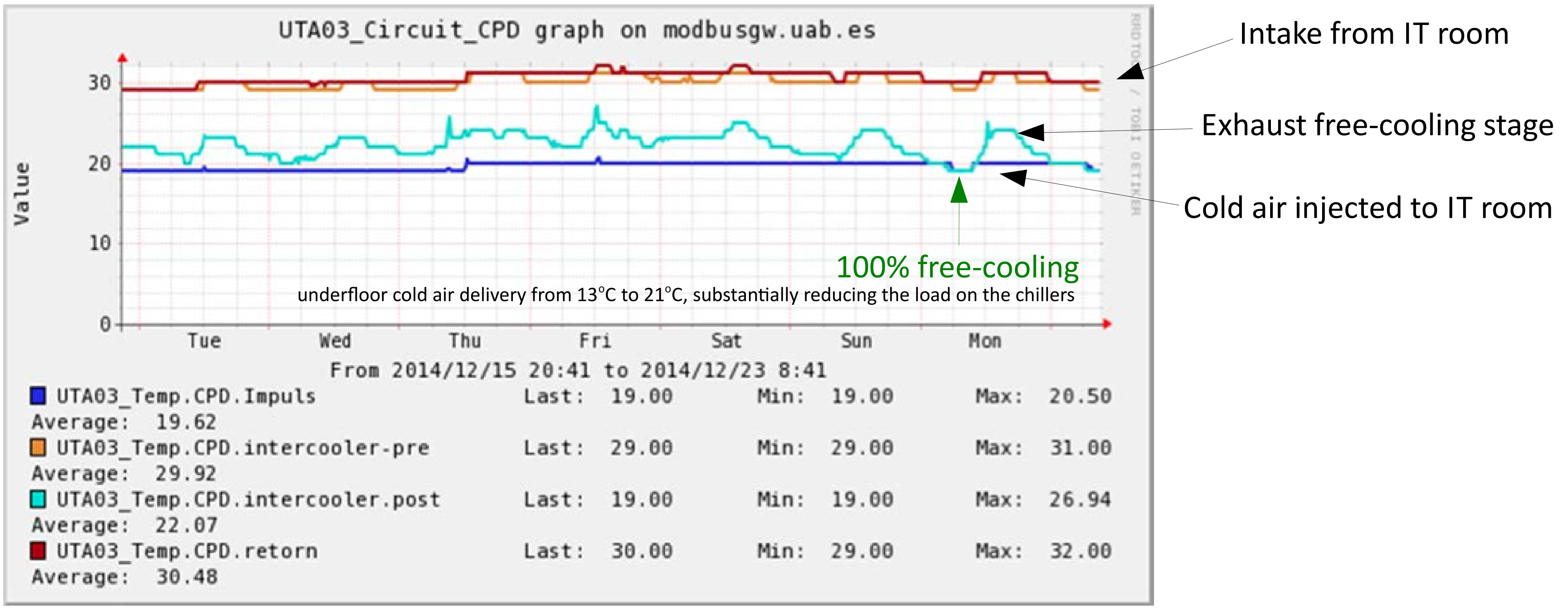


rdd graphs

Free-Cooling at PIC

The work was completed in **September 2014**

- Period ahead to study/adjust the system: reach maximum energy efficiency
- In Dec. 2014, we already reached PUE of 1.3. Averaged PUE of ~1.45 expected.



- Electricity costs savings in the next years estimated at **~100k€/year**

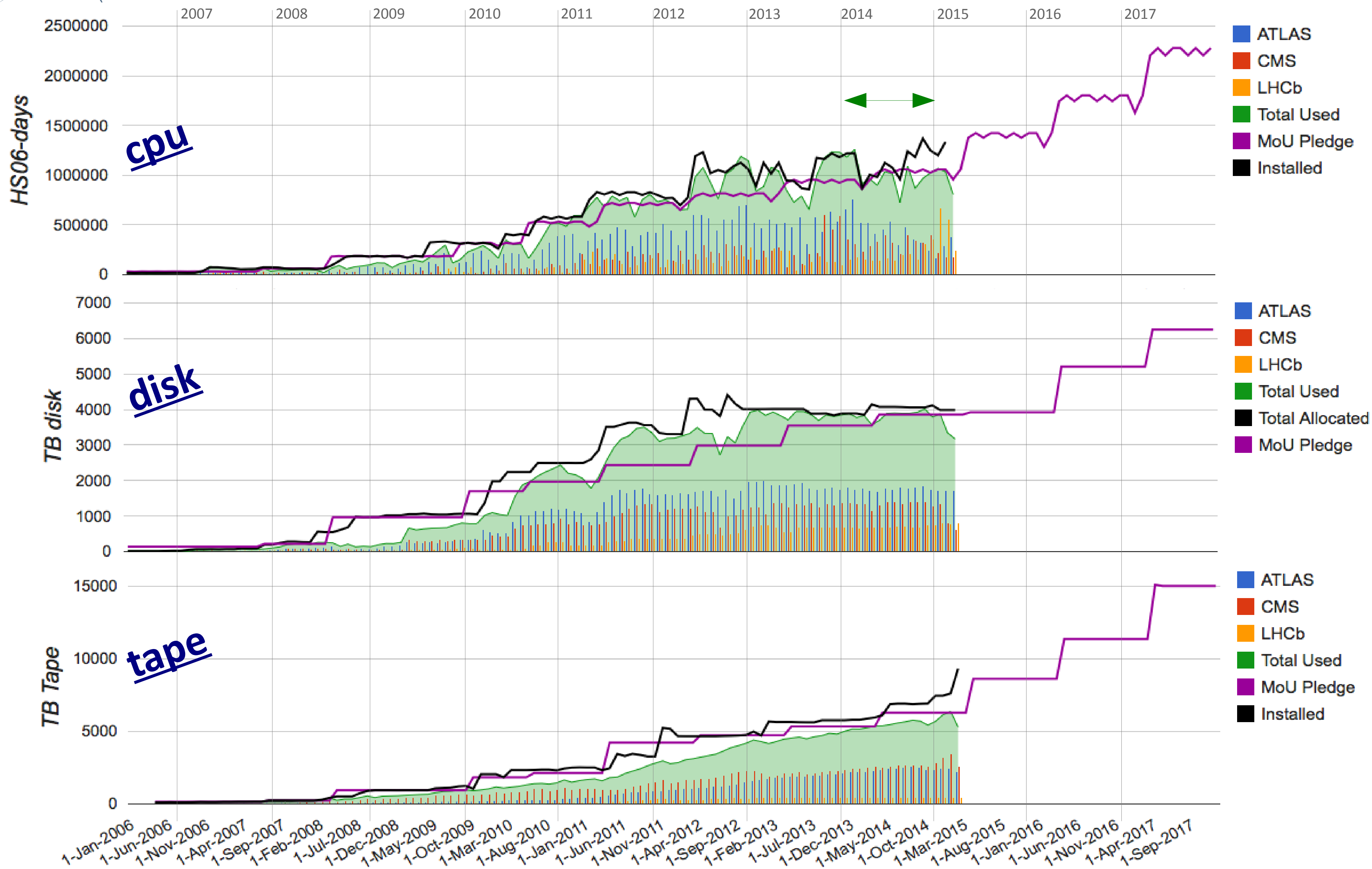
UPS with improved energy efficiency

- Old UPS system had **loses of ~15%**
  - **New UPS of 550 KVA** w/IGBT technology was installed in May 2014, w/efficiency in the range of 97%-99%

- **Eaton POWER XPERT 9395**: 2x275 KVA (2x250kW) modules
  - Placed in PIC basement
  - It can be configured to use only 1 module, and keep the other as reserve (when the consume is <250 kW)
  - Gives 10' to react in case of major outage
  - The batteries have a 10 years lifetime
  - It was the only model fitting in the designated area (**lucky!**)

- New electrical **distribution panels** (40x32A lines + multilple 16A lines)
  - In the main IT room (+ additional works in the basement)

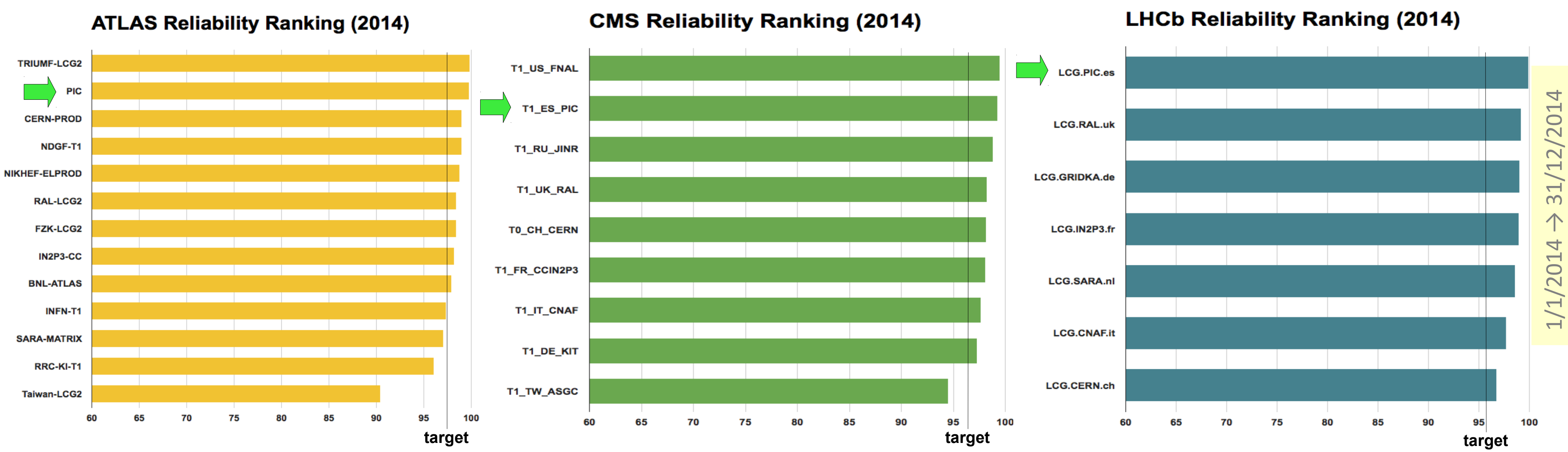
PIC Tier-1 CPU and storage capacity growth...



✔ PIC Tier1 delivering in terms of deploying pledged capacity

… with excellent reliability and efficiency

From Jan. 2014, WLCG measures the reliability using more detailed **experiment probes**



PIC Tier1 is at the **top of Reliability Rankings** (99.9% ATLAS, 99.4% CMS, 99.9% LHCb)

YES, being smaller makes it a bit easier to be reliable

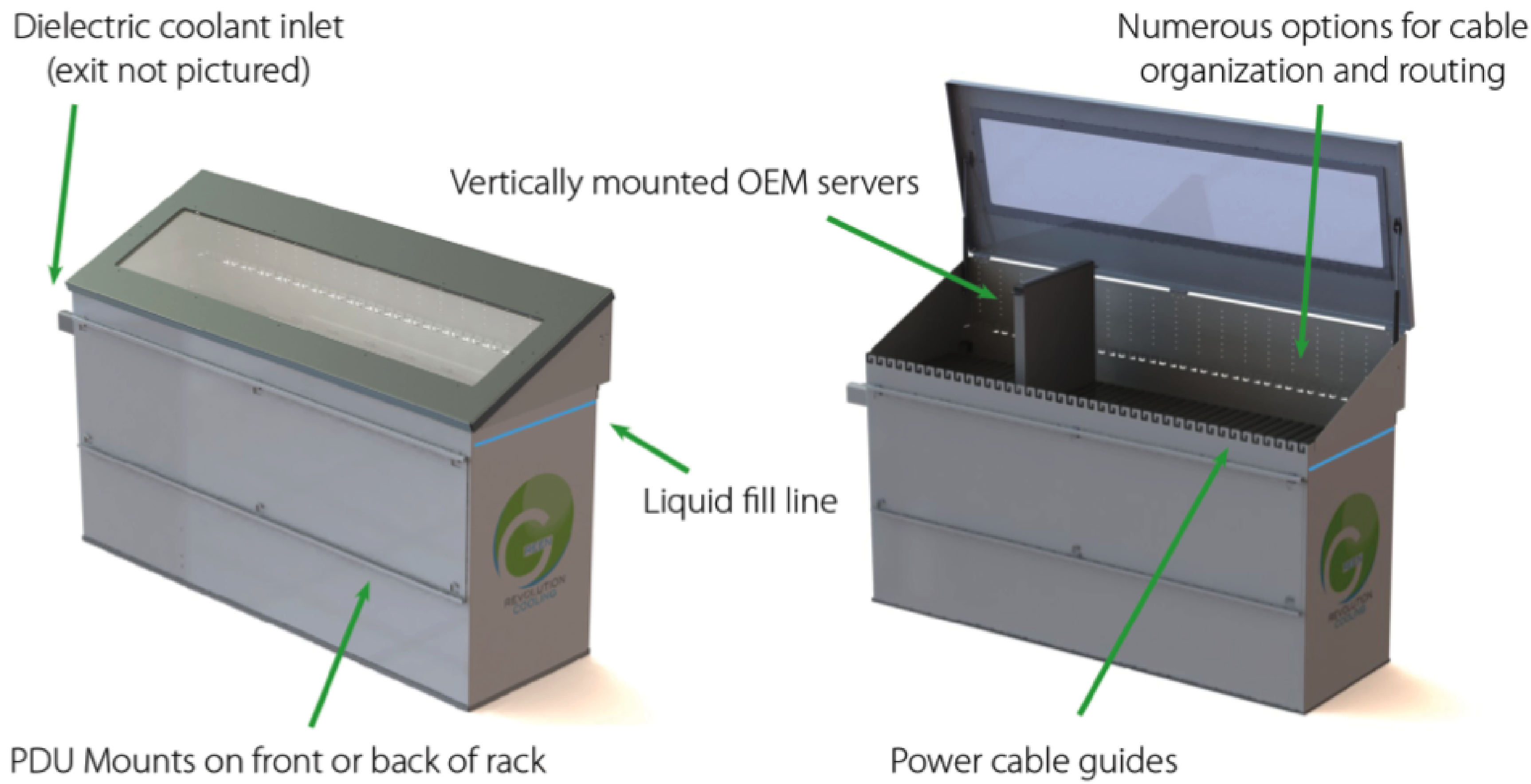
BUT, being a multi-experiment site makes it harder

✔ PIC Tier1 delivering in terms of service quality

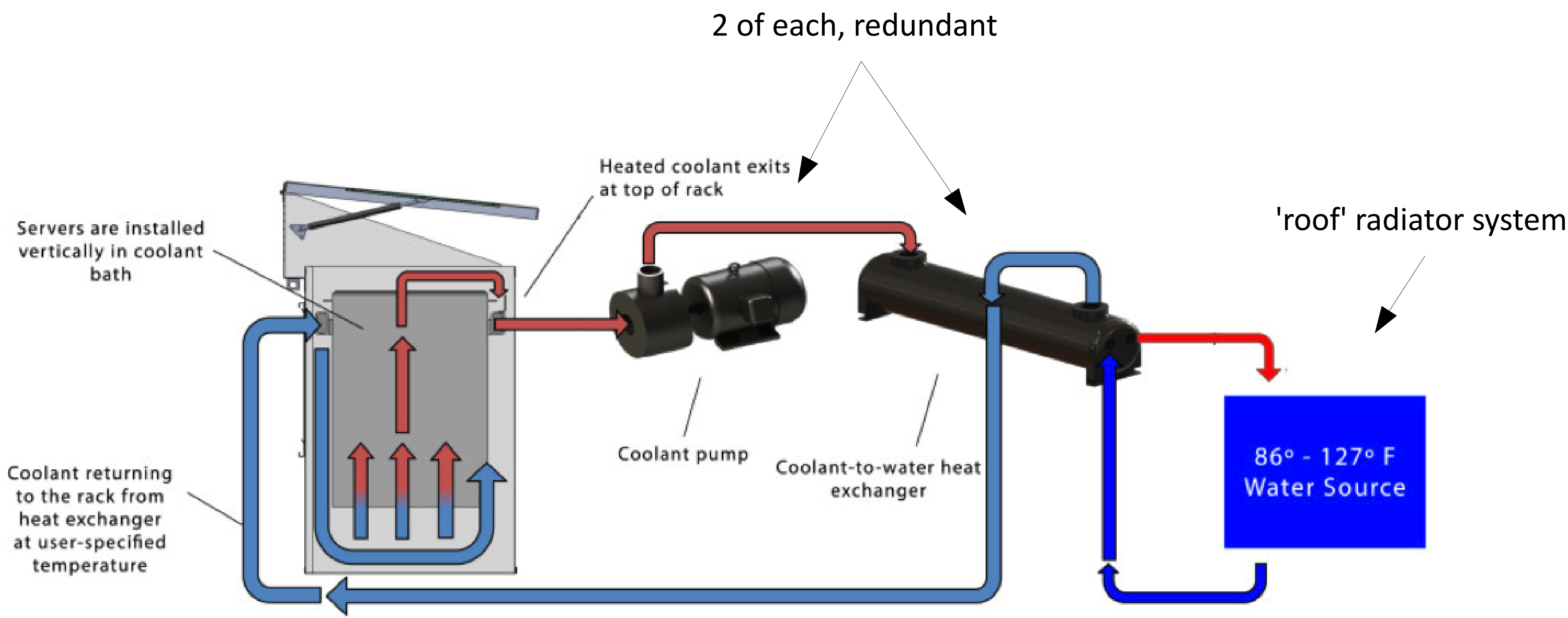
Oil immersion techniques

- Introduction of **<u>oil immersion techniques</u>** in PIC
  - Old compact module, containing 2/3 of CPU, was decommissioned
  - Ongoing installation of **GCR CarnoJet system**
    - *4 x 46U Tanks: each one capable to dissipate up to 45kW*
    - *2 redundant coolant pump modules + coolant-to-water heat exchanger*
    - *Circuit of water to the top of the building – 'radiators' to cold down water*
    - *Oil temperature 50ºC - Max. water input temperature 40ºC*

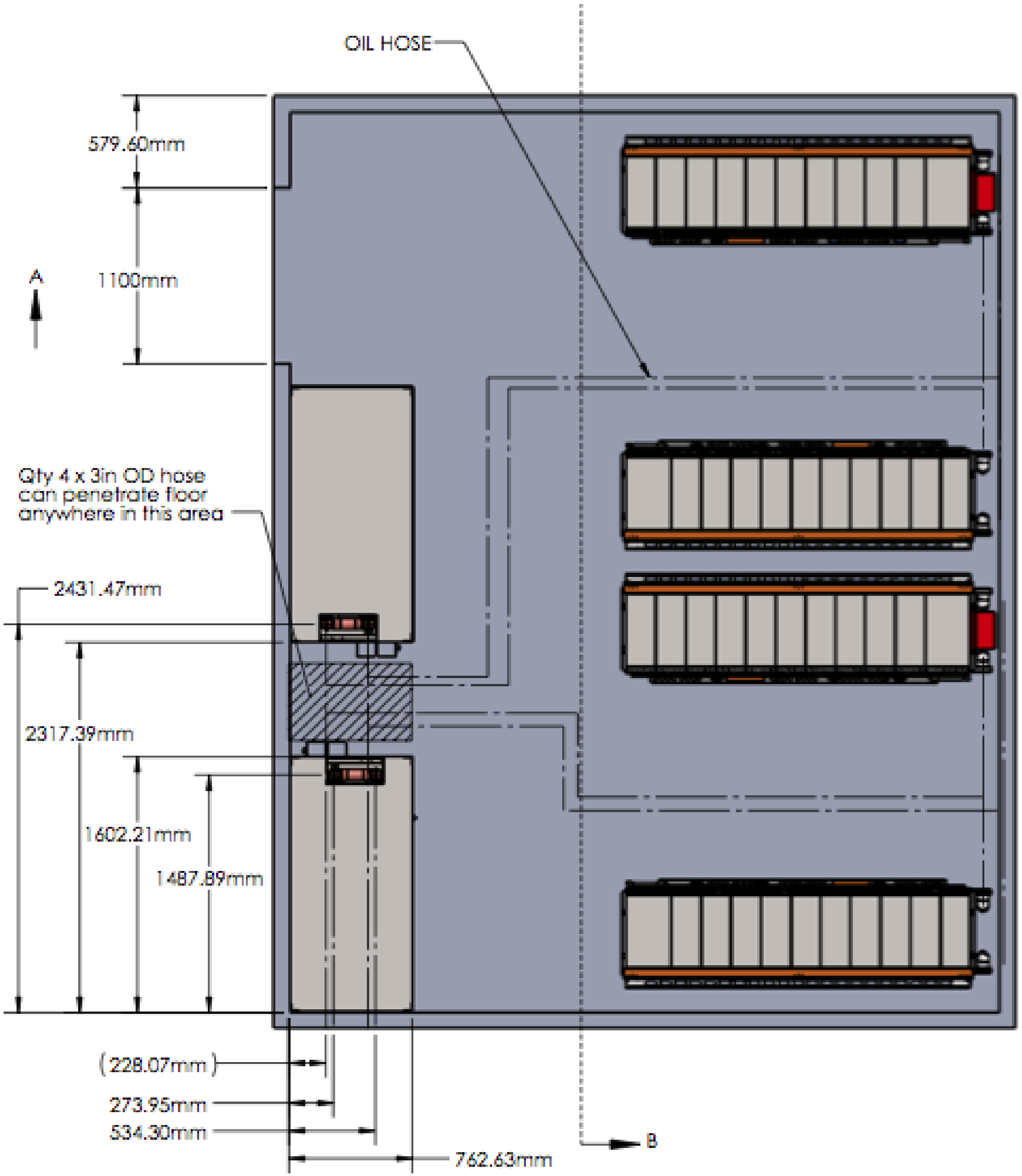
Oil immersion techniques

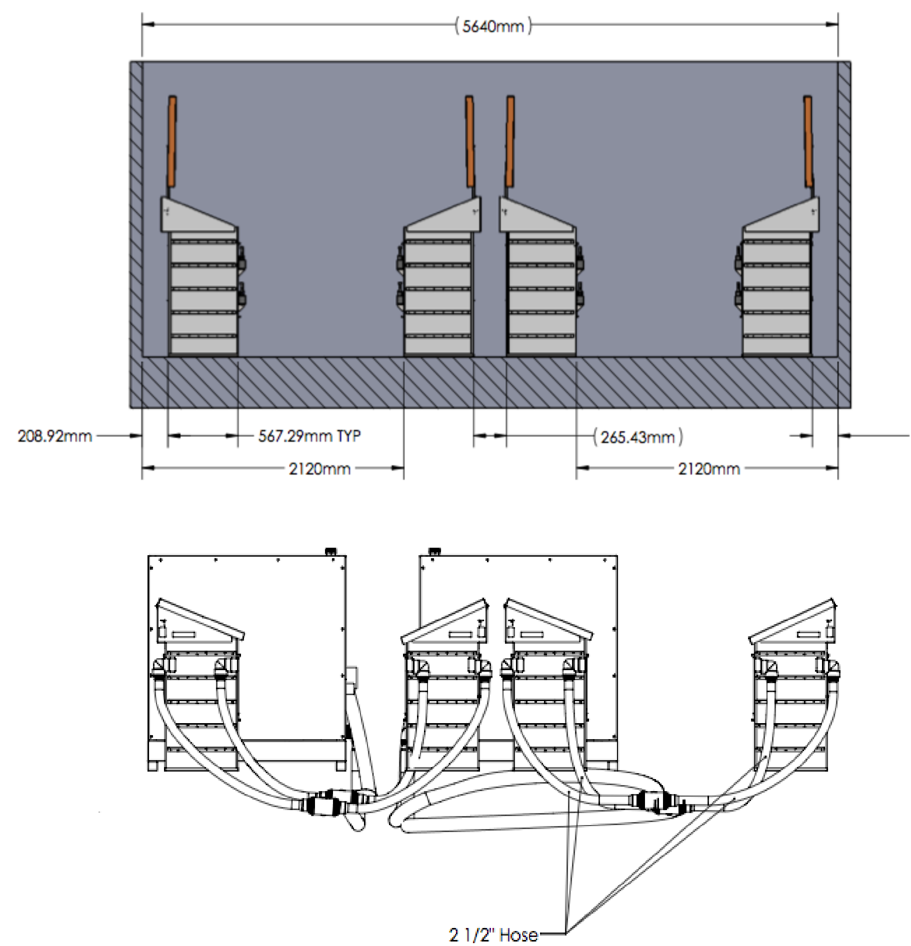


Oil immersion techniques



Oil immersion techniques





2 redundant coolant pump module
+ coolant-to-water heat exchanger

Oil immersion techniques

The tanks arrived yesterday to PIC

Oil immersion techniques

- Next weeks these systems will be **installed in PIC**
  - Then the water cooling system will come
- **CPU servers for immersion** are being purchased:
  - ~10 servers: 26xxv3, 500 GB SSD local disk, 4 GB/RAM/core DDR4
    - *~550 cores → 8,5 kW → ~7 kHS06*
    - The servers have **special configurations**: pizza-like, no fans (BIOS hack), no CPU coating, no fibers in oil please, …
    - models need to be certified by companies (otherwise, no warranty!)
    - Switches don't like oil immersion – should work at 50°C
- Aimed to work with PUE in the range 1.05-1.1
- Results will be presented at the **next HEPIX meeting**

PIC
port d'informació
científica

Questions?