Evolution of the Building Management System in the INFN CNAF Tier-1 datacenter facility

Pier Paolo Ricci

on behalf of INFN CNAF Tier1 Infrastructure

pierpaolo.ricci@cnaf.infn.it

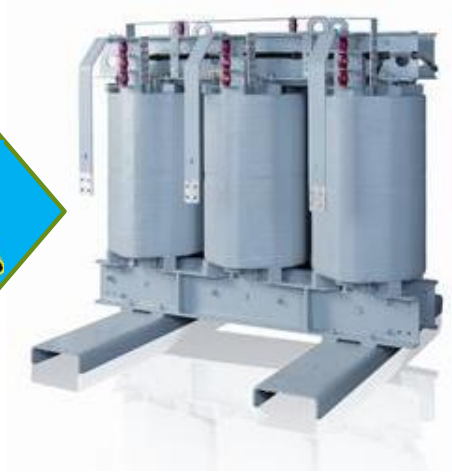
CHEP 2016 San Francisco

Summary

- The INFN CNAF Tier-1 resources: Infrastructure and IT
- Development of a new Building Management System (BMS): requirements and primary objectives
- The StruxureWare™ Building Operation software (SBO) architecture
- The Power Usage Effectiveness (PUE) and other relevant metrics
- Integration with Open Standard (Web Services) and the Tier-1 Monitoring tool
- Future development and conclusion

INFN CNAF Tier-1 resources

- **1 Dedicated electrical line** from energy distribution network
- **15,000 V** Voltage
- Up to **4.0 MW** Power



- **3 Tranformers** (2 in production, 1 reserve)
- **2.5 MVA each**
- **15,000 V => 400 V**
- **2 separated 4000 A** electrical distribution **lines** (RED & GREEN)

- **6 EMERSON Chiller**
- **300 KW each** of **cooling capacity**
- Periodic auto-rotation of active vs standby (N+2 rendundancy)
- Integrated freecooling

- **2 EuroDiesel Rotary diesel UPS**
- **1700 KVA** each (1340 KW Real Power with a ~0.8 cosϕ)
- 3-5 days autonomy

+

- **1 Standard diesel generator** (no UPS)
- 1250 KVA

can be used for emergency power or not-IT equipment (e.g. chillers)

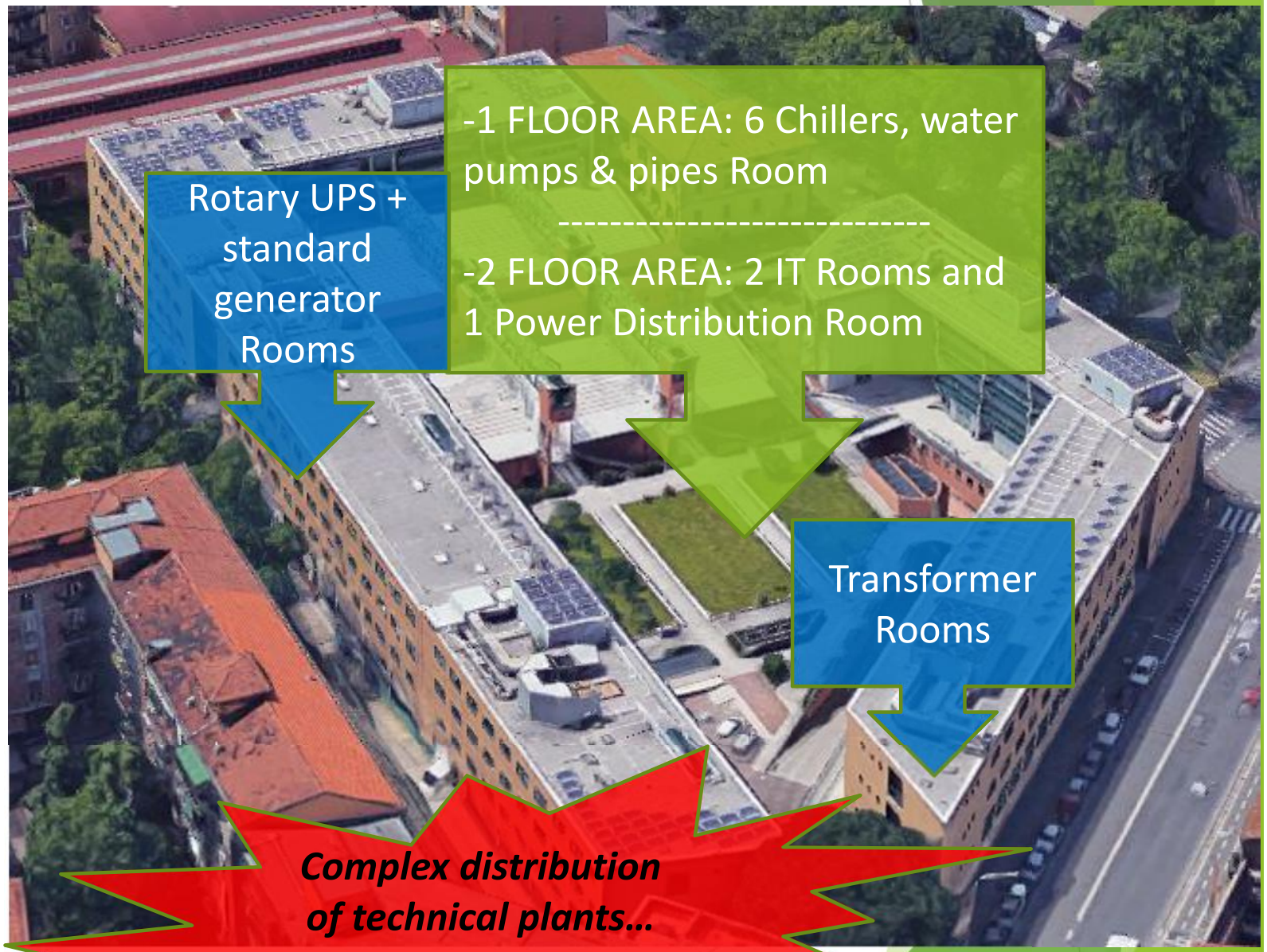
UPS

IT RESOURCES

INFN CNAF Tier-1 resources rooms

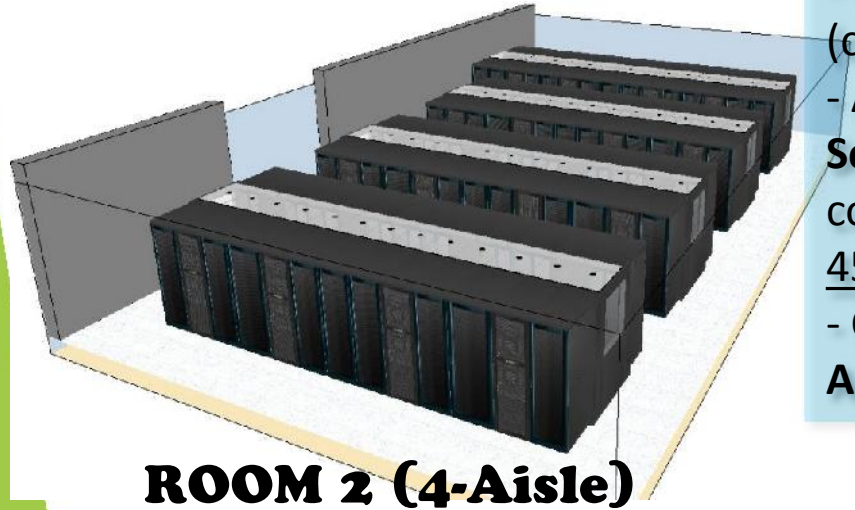
- The INFN CNAF Tier1 is hosted in a complex university building ☹
- Installation used all avaliable space in 2009
- A total of 1950 m^2 is used for:

- 2 IT Resources Rooms (250 m^2 + 350 m^2)
- 4 Additional locations:
. Tranformers Room
. UPS Room
. Chiller Room
. Power Room



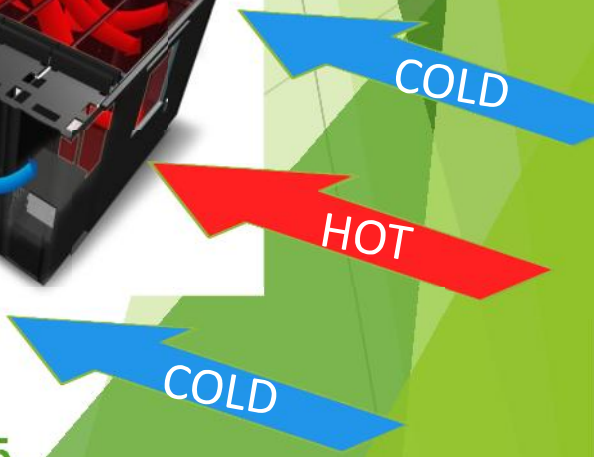
INFN CNAF Tier-1 IT rooms

- **44 APC InRow RP (IRP)** Precision Cooling with 2-ways valves, 3 Fans and humidity control. (Cold water provided by chillers)
- 50 KW of cooling capacity each IRP => **1600 KW** with N+2 redundancy in 6 "Aisle" over 2 different rooms
- **48** Racks for IT equipment with a 10KW cooling capacity (Room 1=> 2-Aisle)
  76 Racks for IT equipment with a 20KW cooling capacity (Room 2=> 4-Aisle)
- Hot Aisle containment (without floating floor) with environmental sensors (T&H) in HOT and COLD corridor



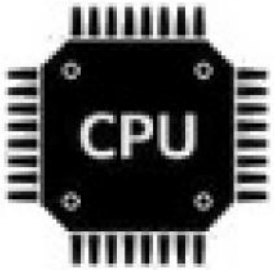
ROOM 2 (4-Aisle)

- Water Setpoint: (chillers): 15°C /20°C
- **Air T & Humidity Setpoint:** (IRP group control): 24°C & 45%/60%
- **Cold/Hot detected Air T:** 24°C/31°C



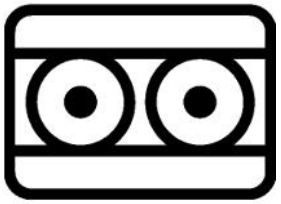
INFN CNAF Tier-1 IT resources

Tier-1 for all LHC experiments (Alice, Atlas, CMS, LHCb)
and ~20 others non-LHC (including Astroparticle Physics)

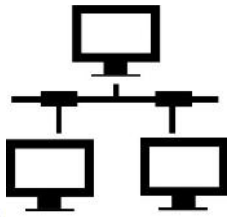


- Tier-1 cluster computing power:
  205,000 HEP-SPEC06
  provided by 22,000 CPU cores

- Tier-1 disk space capacity:
  ~20 PByte net used space
  80 disk servers (8 GPFS clusters)



- Tier-1 tape space capacity:
  1 Tape library with 10,000 slot capacity (Oracle SL8500)
  ~ 34 PByte tape used space (8.5 TByte cardridge)



- Tier-1 network facility:
  4 Core Switches
  350 x 10Gb/s Ports
  468x 1Gb/s Ports
  6+2 (general purpose) x 10Gb/s WAN Connections

A new BMS => Objectives:

- The old BMS system was the TAC VISTA software:
  - currently phasing out… ☹
  - Many "cons" (difficult to edit, no compatibility with open protocols, GUI based on Java etc…)

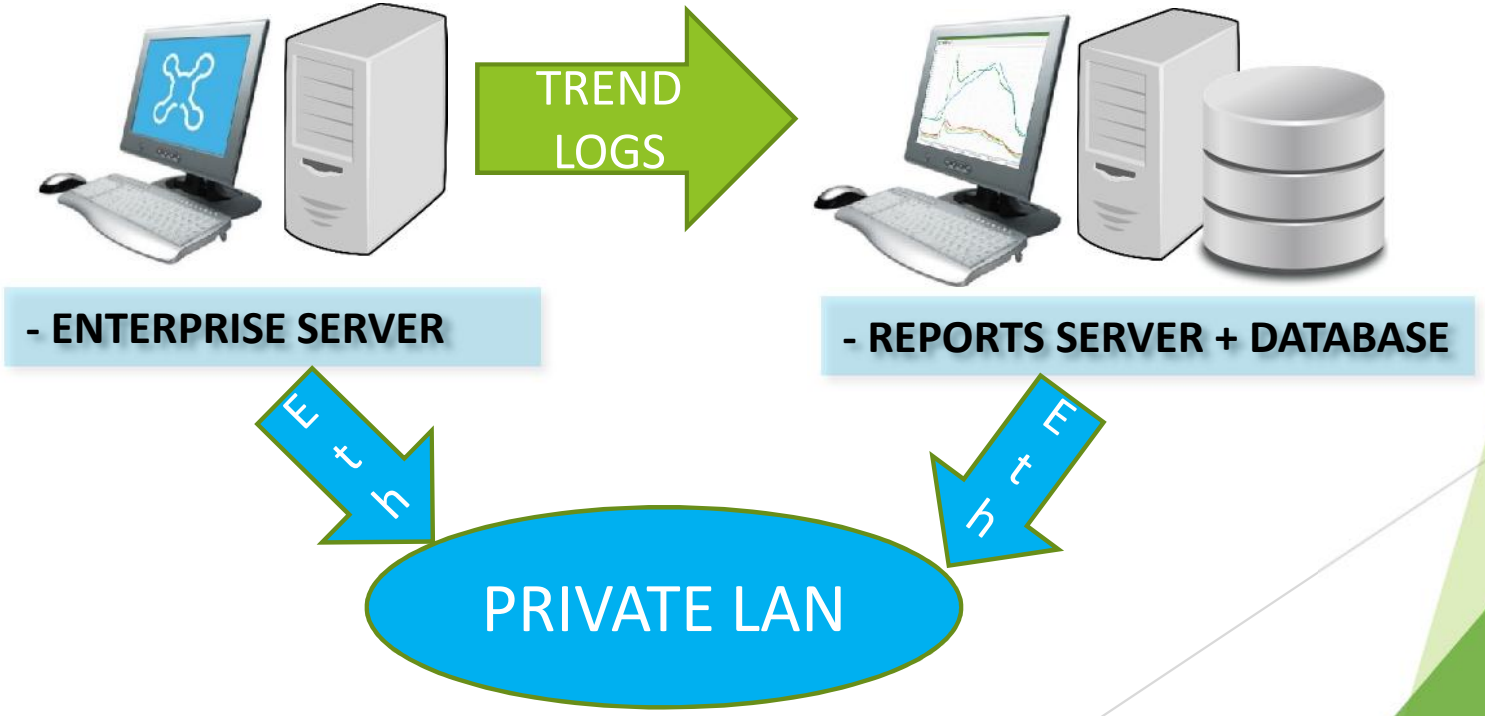
- We need a more flexible system that could re-use the "sensors" and "collectors" hardware (to limit the hardware cost of the migration) and improve user-friendly management

- The "natural" choice is the **Schneider StruxureWare™ Building Operation software (SBO)** architecture:

  - Full compatible with Modbus protocol (TCP/IP and serial)
  - Full compatible with the TAC VISTA Lonworks network (re-use cabling&HW)
  - The webstation GUI user interface just need a standard browser (no Java or other plugin needed); works directly on mobile devices! ☺
  - Open to standard protocols (e.g. webservices serve&consume)
  - Migration from TAC VISTA was easily split into 3 phases over 8 weeks period:
    1) TCP Modbus "aisles" & "PLCs" => can run both on VISTA and SBO (6 weeks test phase)
    2) Serial Modbus => OUT on VISTA & IN only on SBO (2 weeks "critical")
    3) Lonworks Network & ALL alarms/trends => ONLY SBO ACTIVE! (2 weeks for finalize)

The Schneider SBO architecture

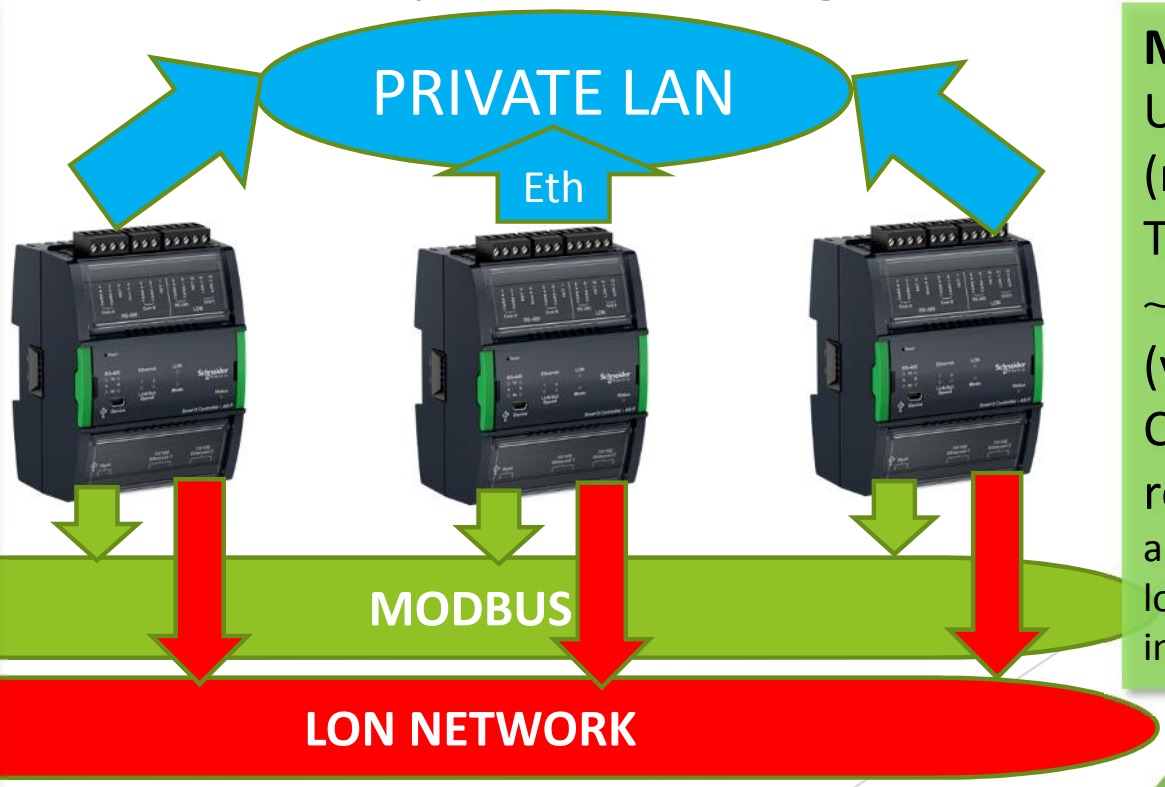
- The core of the software management, web user interface and archiving is based on 2 servers:
  - **Enterprise software server:** runs the core software for management & backup of the configuration.
  - **Report Server**: used for archiving the long-term trends (Microsoft SQL Server) and adds advanced reporting options.



The Schneider SBO architecture

- **3 Automation Servers** provide the "engines" for the **BMS system**, in 3 "strategic" physical locations (-1 Floor Area, Power Distribution Room & Tranformer Room):
  - Runs stands-alone
  - Collect data directly from the Lonwork Network, Modbus Serial and Modbus TCP (PLCs) and provides control logic

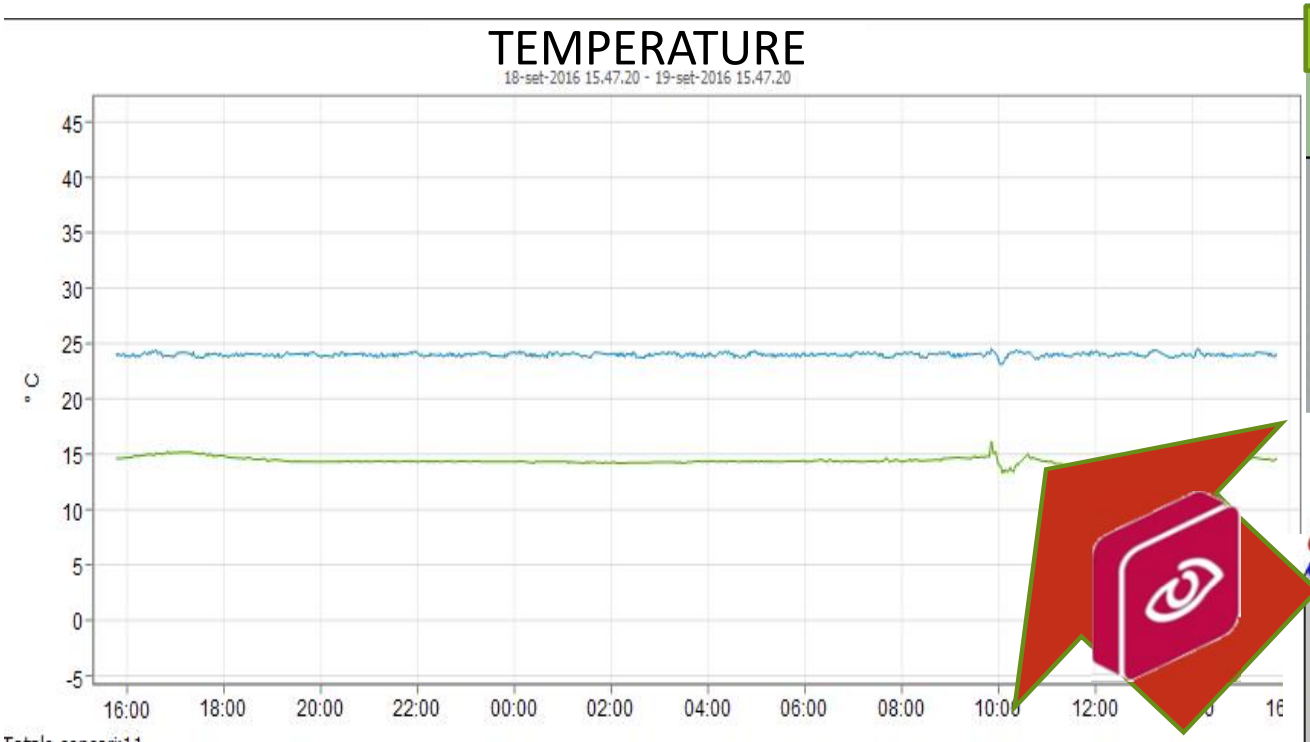
LON NETWORK
~30 "Xenta Family" programmable modules (reused from TAC)
~300 "Points" (variables or states)
2 separated bus
Implements control logic (e.g. chiller rotation)



MODBUS NETWORK
Use both Serial (rs484/rs232) and TCP/IP
~4500 "Points" (variables or states)
Connection to 6 reduntand PLCs (PLCs are used for electrical switch logic & electric measuring instrument)

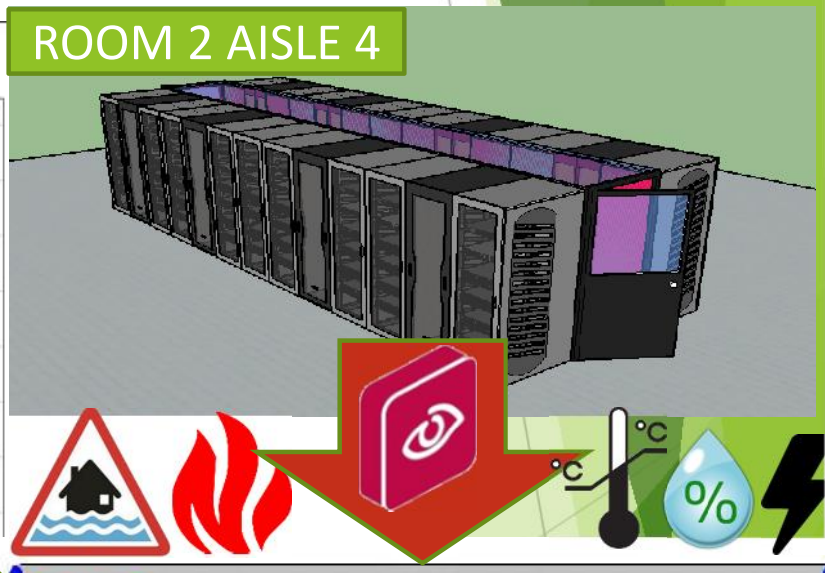
The Data Center Expert

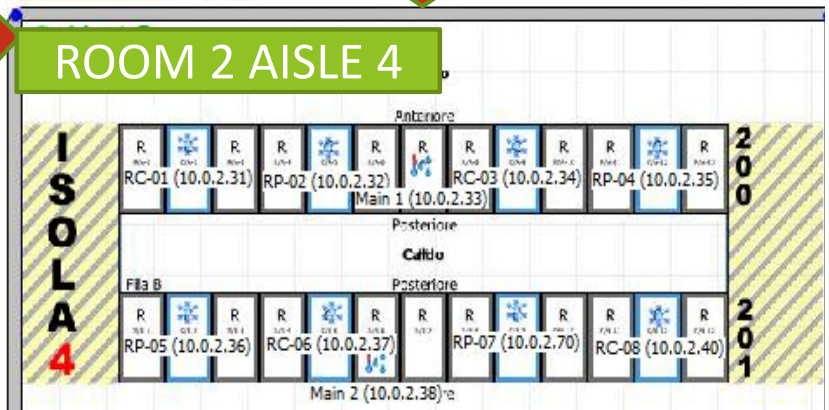
- The Tier-1 BMS is integrated by another Schneider (formerly APC) software package: **The StruXureware Data Center Expert (DCE)**
- Fine monitoring, tuning and notification over the Datacenter "Aisles" components and Metered Power Distribution Units (PDUs)
- With the MODBUS TCP Output Module => **DIRECT INTEGRATION IN SBO**



Totale sensori:11

Colore	Dispositivo monitorato	Sensore	Unità	Ultimo val...	Valore mi...	Valore ma...
	RP-02 (10.0.2.32)	Temperature Fluid IN		14,4	13,3	16,1
	RP-02 (10.0.2.32)	Temperature Air OUT		23,9	23,1	24,5

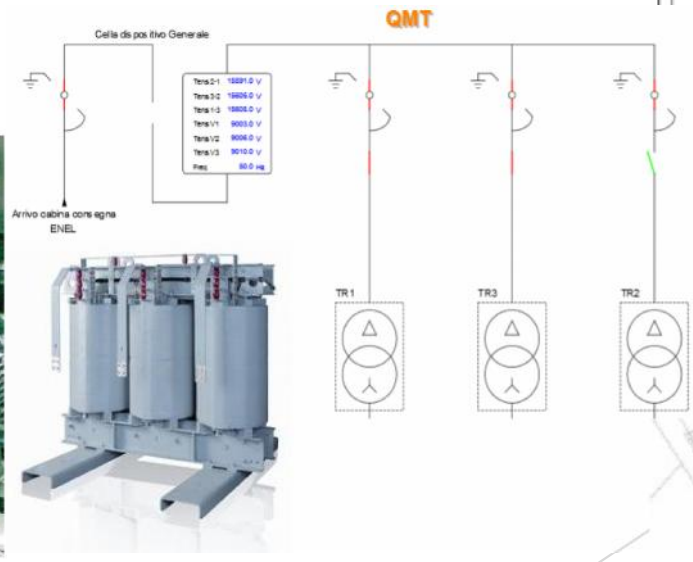


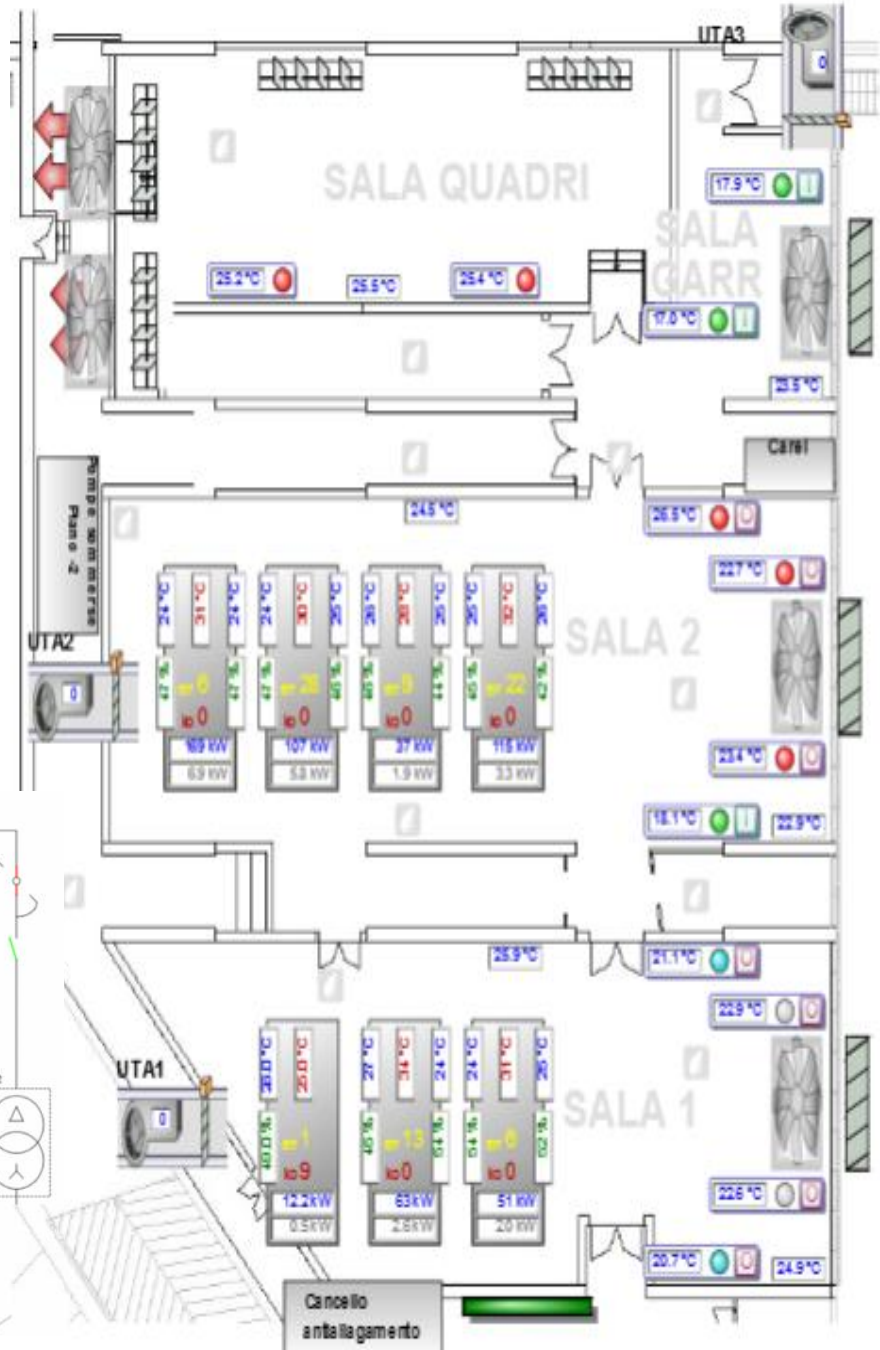


The INFN CNAF Tier-1 BMS

- Power Switch state, Electrical values, Temperature & Humidity values, Fan states etc…
- Mail alarms notification fully customizable
- User-friendly Grafic Pages for HVAC *(Heat, Ventilation and Air Conditioning)*, Electrical circuit diagrams, Mechanical Pages, Fire prevention system, Flooding&Water Leaking (-2 Floor) etc…

WATER PUMPS

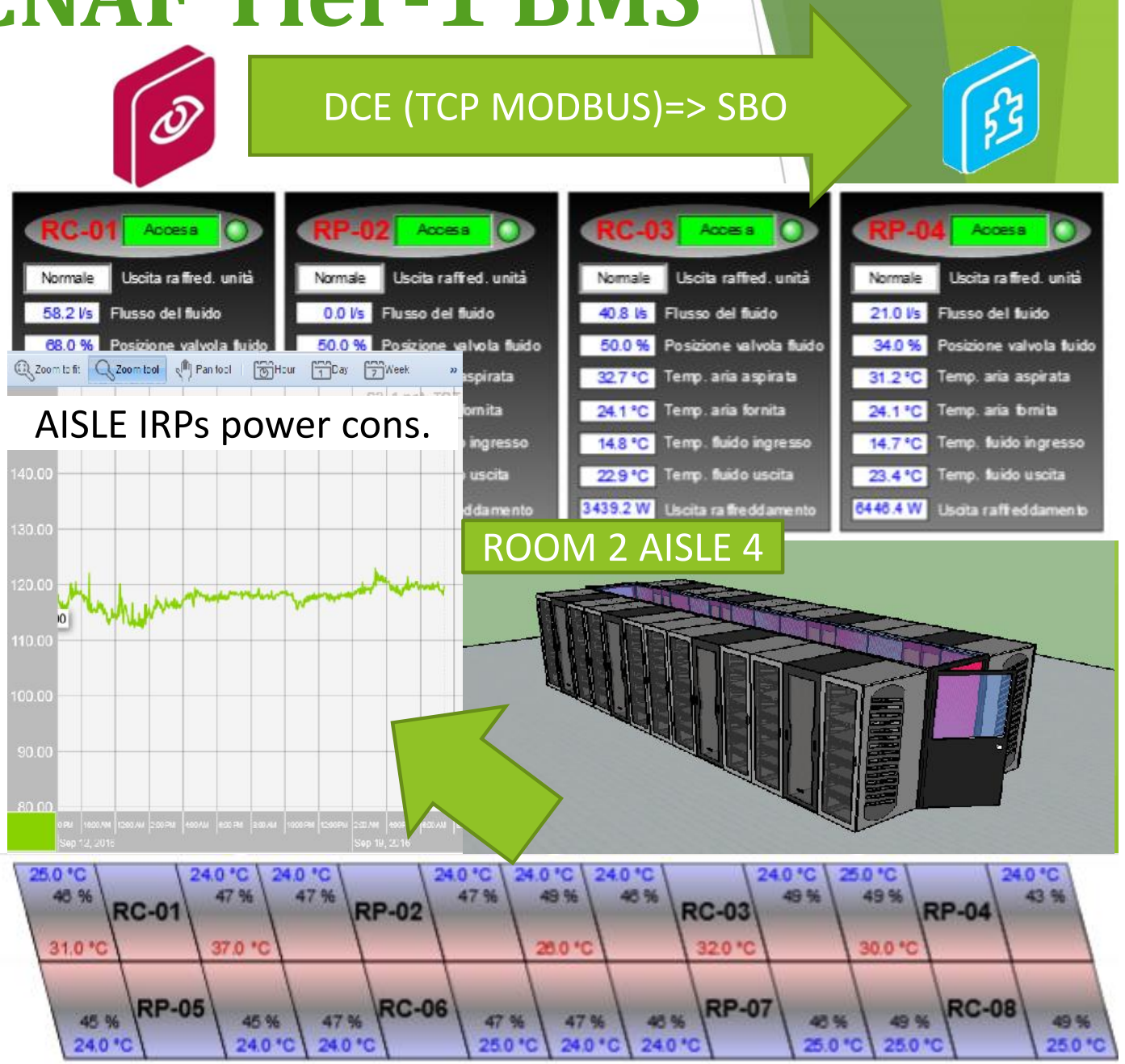




The INFN CNAF Tier-1 BMS

- Information from the DCE software are integrated in SBO

SBO is the central "entry point" for the **INFN CNAF Tier1 BMS!**



BMS METRICS (LOGS & TRENDS)

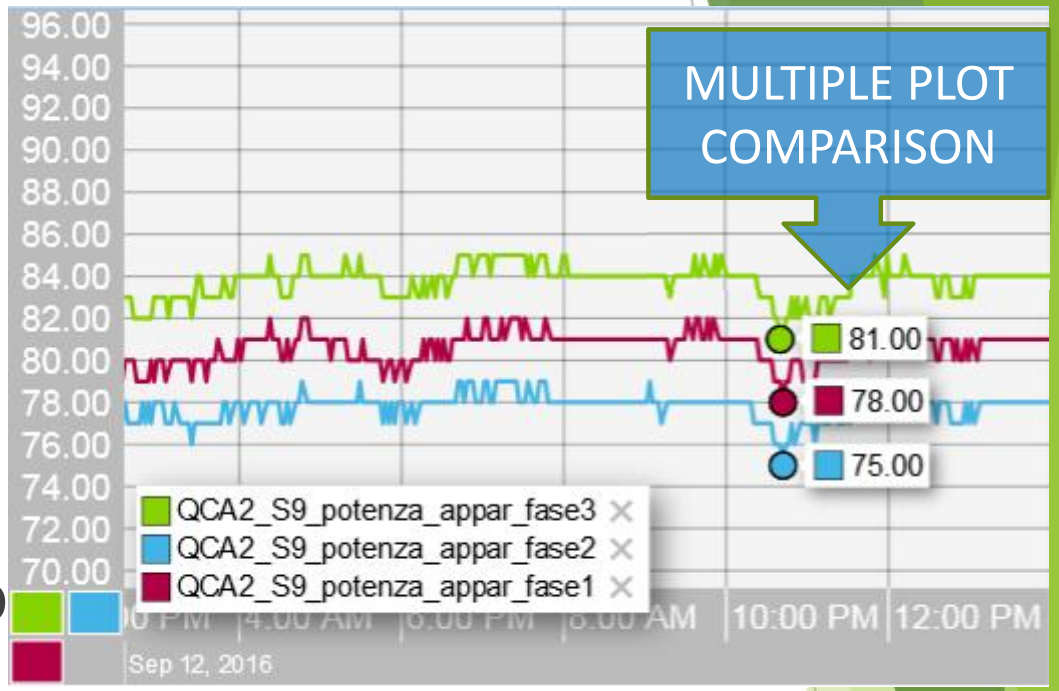
- The **metrics** (*state* and *variable* values) are one of the most important features of an optimal **BMS.**

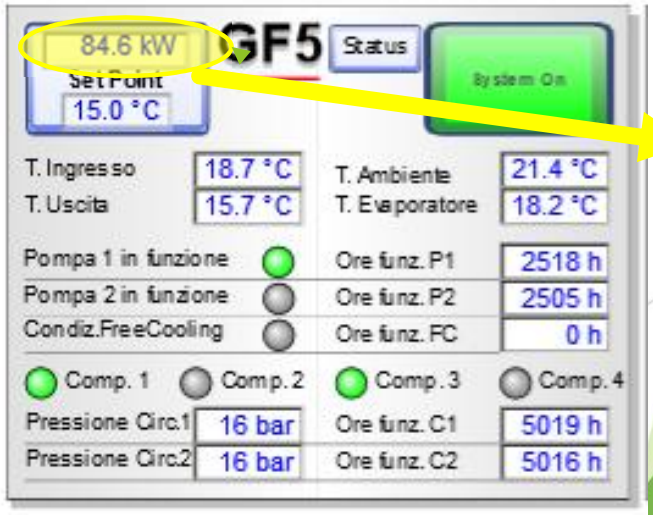
- LOGS & TRENDS represents the "history" of the physical quantities in the technical plants & datacenter rooms.

- Essential for:
  - Identification of periodic "hot spots" or critical area
  - Datacenter HVAC/Power optimization
  - Reverse engineering of specific conditions
  - Data center Power/HVAC outage or major failure (fire) analysis

BMS METRICS (LOGS & TRENDS)

- The SBO architecture has increased the number of metrics we are collectiong and the archiving duration (=> Report Server & database)
  - Over 2500 LOGS & TRENDS collected! => HIGH DETAILS
  - Intuitive system GUI => TREND
  - 15 minutes granularity
  - Optimized for variable state (e.g. a power switch condition is logged only when a change occurs)
  - Over 10 years of history possible (but it is just started ☺!)





BMS METRICS (PUE and pPUE)

From Wikipedia => "*Power usage effectiveness (PUE) is a measure of how efficiently a computer data center uses energy; specifically, how much energy is used by the computing equipment (in contrast to cooling and other overhead).*"

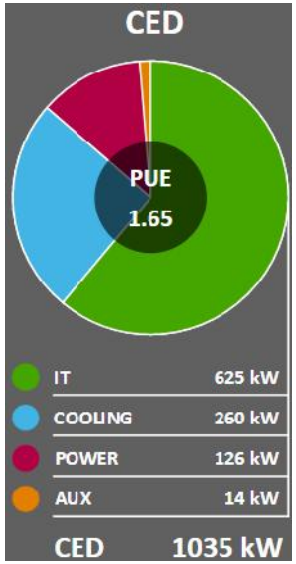
- Increasing cost of Electrical Power => PUE reduction is fundamental!!!

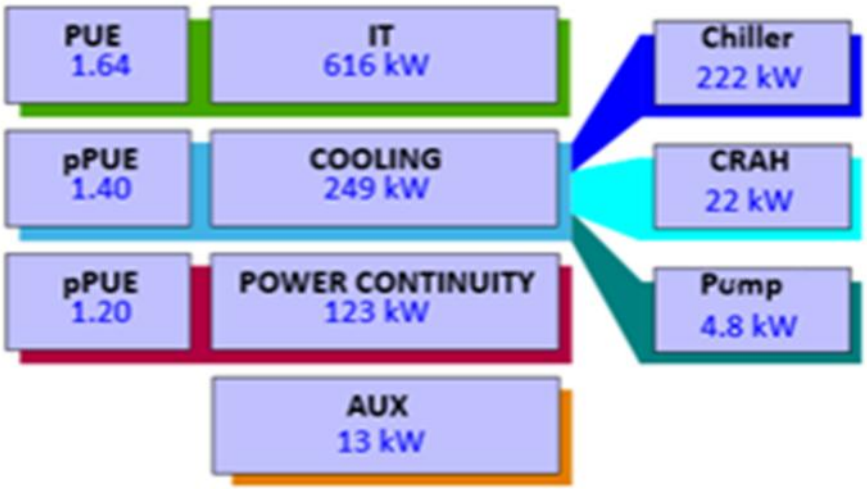
$$PUE = \frac{Total\ Energy}{IT\ Energy}$$

- We also introduce the partialPUE (pPUE) in order to monitor the power demands specific area of the infrastructure and optimize them

$$e.g.\ POWER\ CONTINUITY\ pPUE = \frac{UPS\ Energy + IT\ Energy}{IT\ Energy}$$

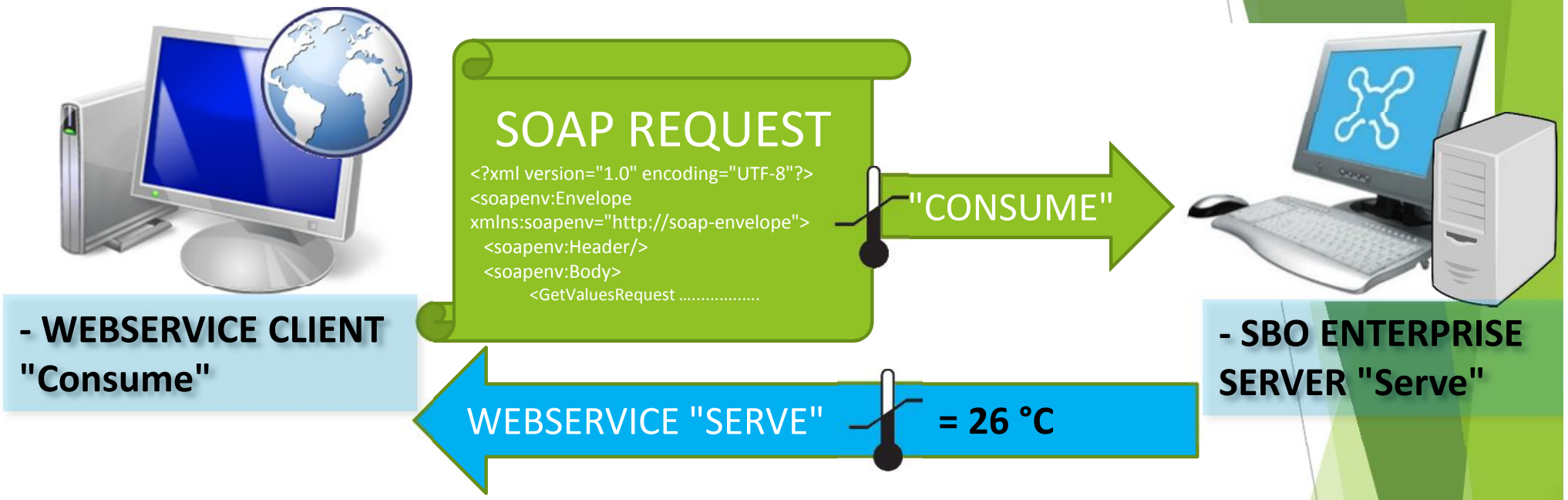
Indicator of the Rotary Diesel Power Continuity Energy Loss





TR1 On 451 kW TR2 Off 0 kW TR3 On 555 kW

BMS open standards (Web Services)



- **Serve & Consume** Web Services capability of the SBO software
- **SOAP** request using standard code
- NDA Agreement signed with Schneider for the Web Services **Serve** of "strategic" variables
- Access to the "current value" of BMS variables from the INFN CNAF Tier-1 Monitoring Infrastructure *(see next slide)*
- Future implementation fot the Web Services SBO **Consume** e.g. connecting to external weather forecast => Automatic adjust of HVAC scheduling

Tier-1 Monitoring Infrastructure

- Monitoring made with Opensource components:

Backend

- Metrics: Sensu for standard metric measurements + custom python scripts (implementation of SOAP requests)

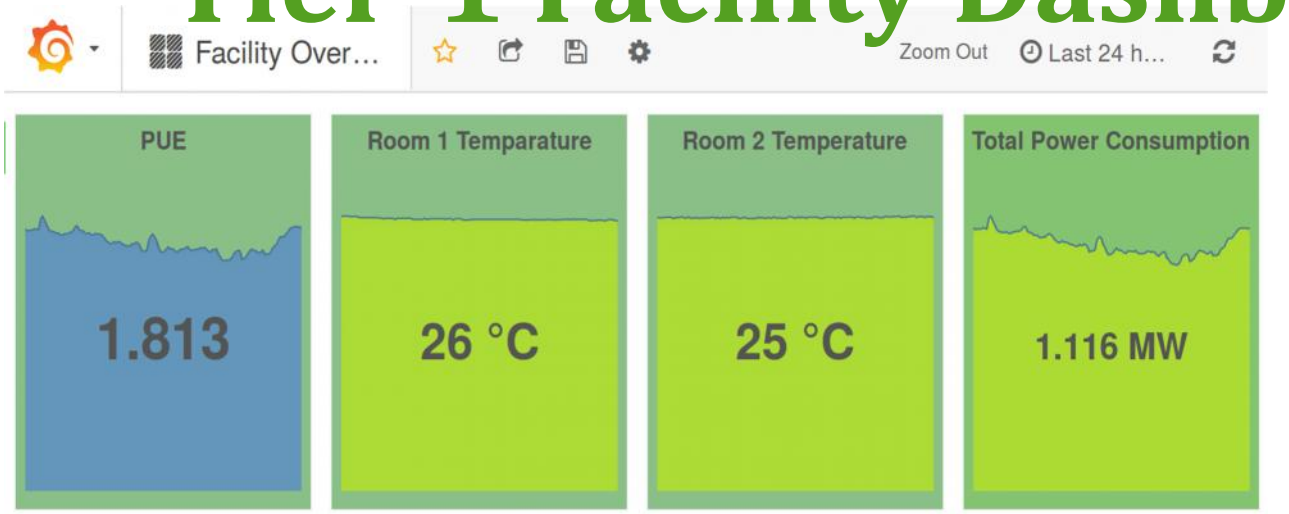
- Data Storage: InfluxDB, Whisper

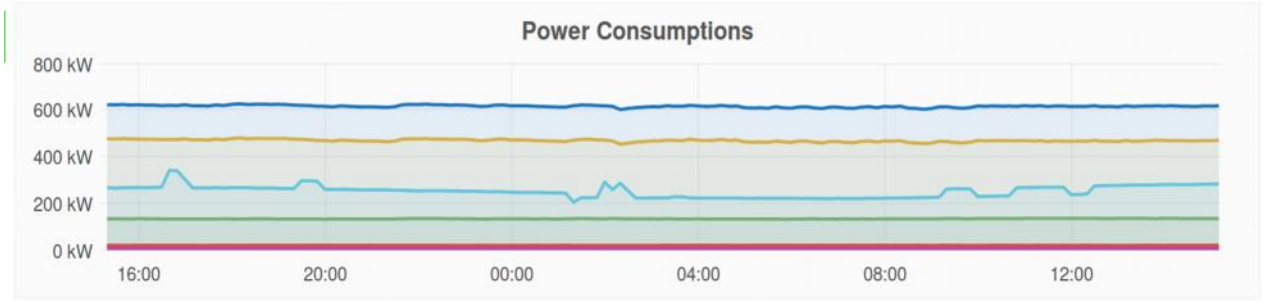
Frontend

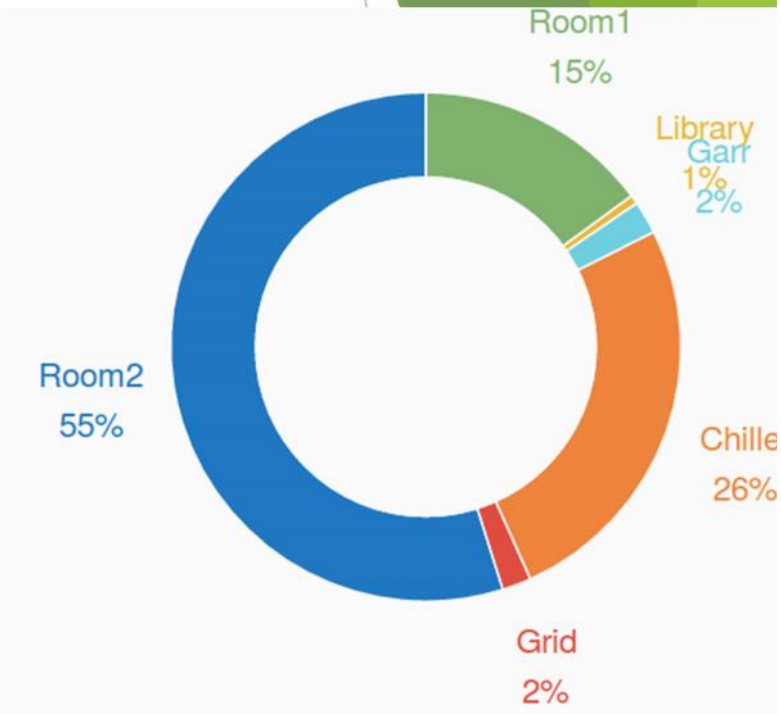
- Uchiwa for notifications and status

- Grafana for charting

Tier-1 Facility Dashboard

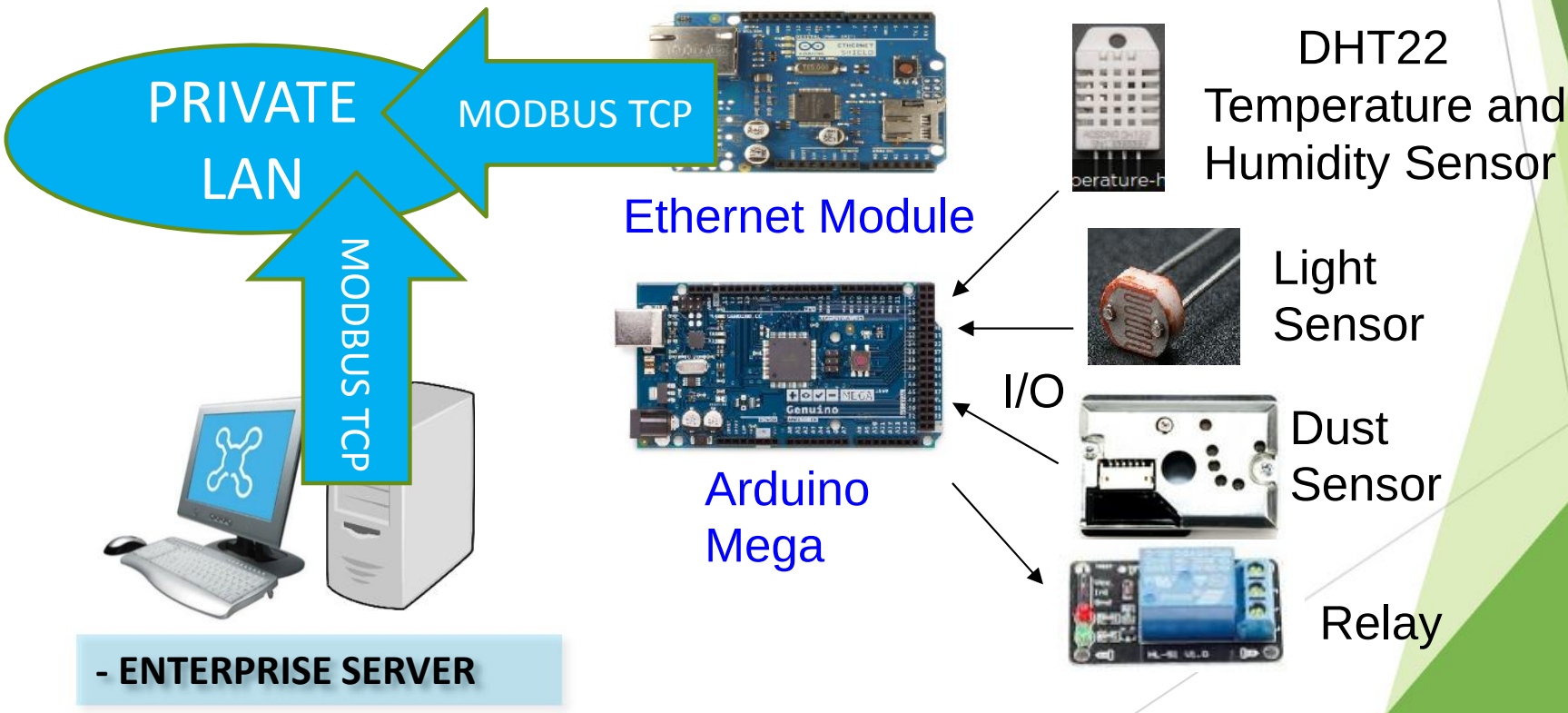






TCP MODBUS sensors expansions

- A study for using low-cost platforms (Arduino with specific modules) as Modbus compatible sensors collector is under way…
- Could be used for additional/redundant monitoring with a minimum economical effort.
- Could help providing very "custom" sensors (e.g. "home made" dust sensor)



Conclusion

- The BMS choice for the DCE/SBO software for migration was successful
  => Good reliability and Great compatibility!

- Our PUE and pPUE analysis show we need big improvement!
  - Increase chiller efficiency => project for a chiller tech refresh
  - Fine granularity of rack power consumption => increase the number of metered PDUs
  - Rotary UPS power loss => big issue ☹

- The SBO compatibility with open standard (WEB SERVICES) improve our BMS integration and "open-mindedness"
- Standardization of comunication (MODBUS) => Possibility to integrate different platforms (e.g. the Arduino tests)

THANKS FOR YOUR ATTENTION!