OPEN.

FOR BUSINESS.

OCP SUMMIT

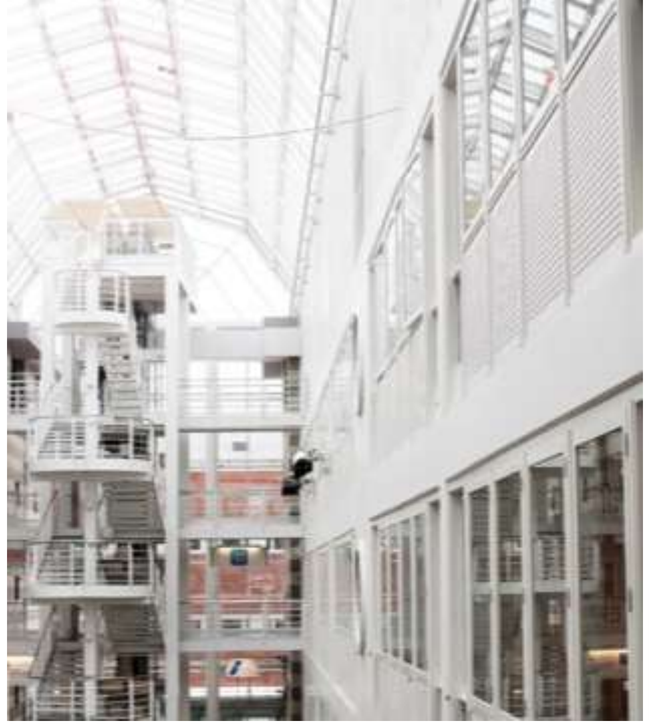
Europe Focus

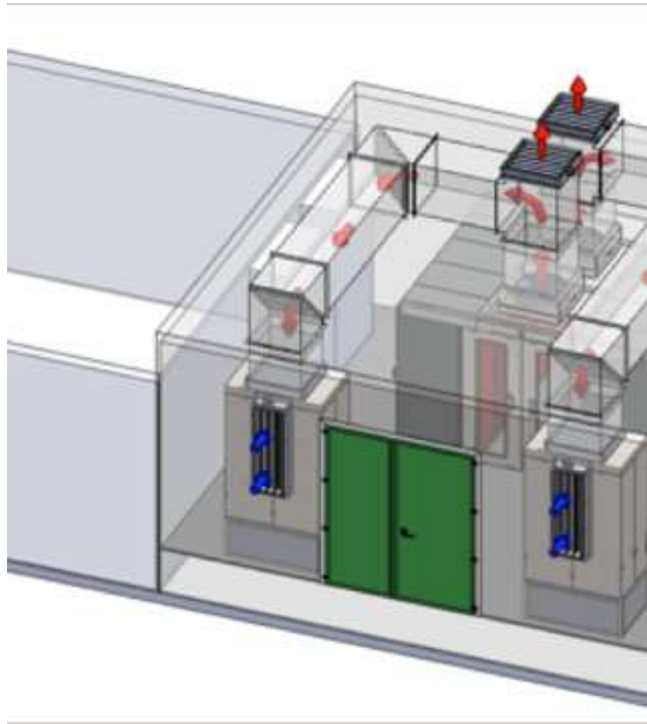
RISE and Open Source Monitoring of legacy OCP Systems funded by H2020

Tor Bjorn Minde, Chief Executive Officer
Dr Jon Summers, Research Leader in Data Centers
RISE SICS North, Lulea, Sweden

Agenda

- Introducing RISE ICE
- Introduce BTDC project
- Detailing the open source setup
- Showing results from the monitoring of the infrastructure
- Show results of the OCP systems operations
  - Heat sinks
  - Fans
  - Wind Tunnel and results
  - Workloads
- Summarise





SICS - ICE
A LARGE SCALE DATACENTER RESEARCH & TEST ENVIRONMENT

Mission and Vision

Mission: Turn great ideas into proven innovations in the data center space

- To build competence within the area of sustainable efficient data centers, as well as cloud applications and data analysis by enabling the use of testing in large-scale data centers
- To attract and thereby contribute to an increase of the number of scientists within the area of sustainable efficient data centers, as well as cloud applications and data analysis

Vision: To become the leading data center research institute in Europe

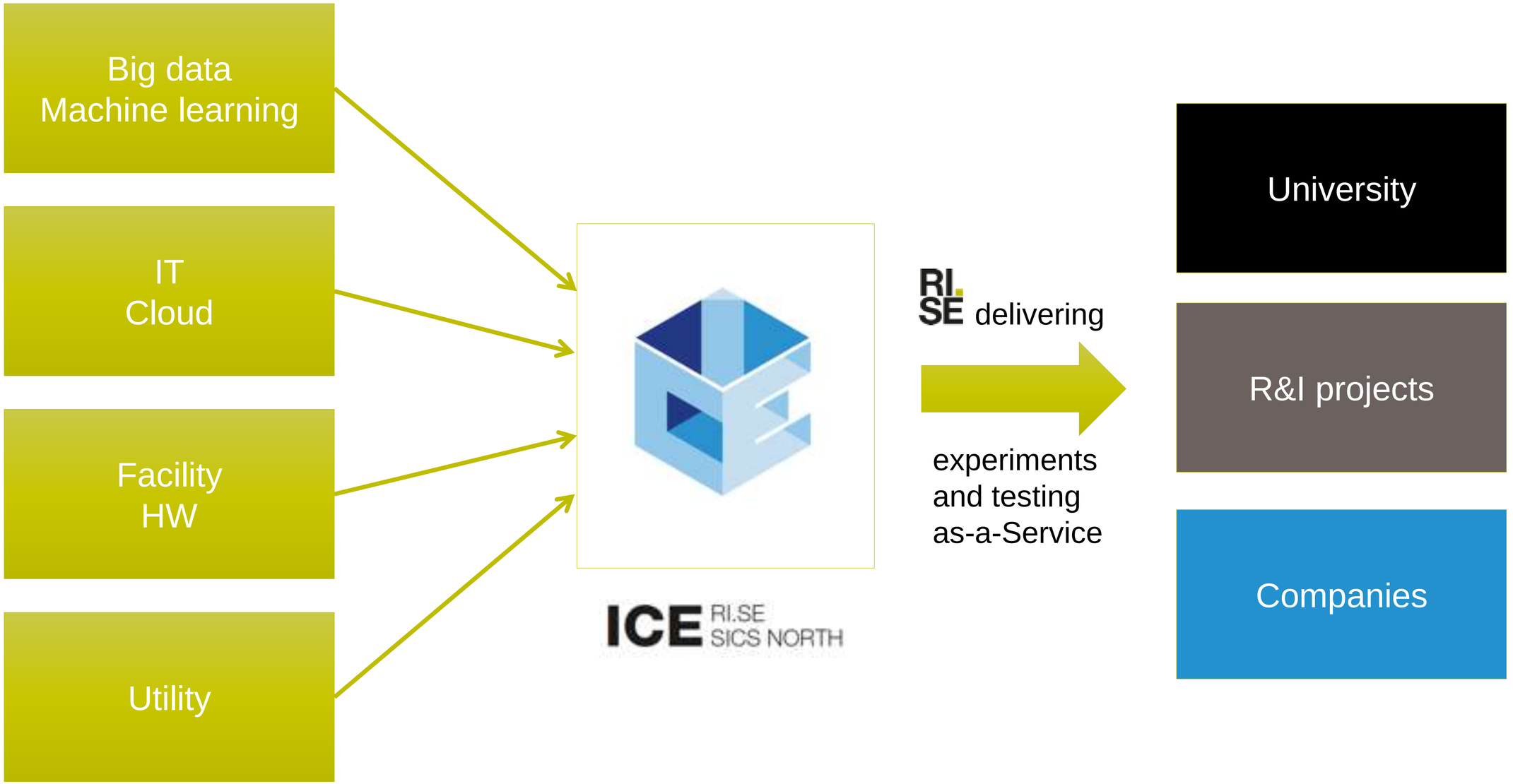
- The leading data center research institute in Europe within the area of sustainable efficient data centers, as well as cloud applications and data center centric data analysis

SICS ICE. A full-scale research datacenter and test environment with the purpose to increase knowledge, strengthen the DC ecosystem, and attract researchers.

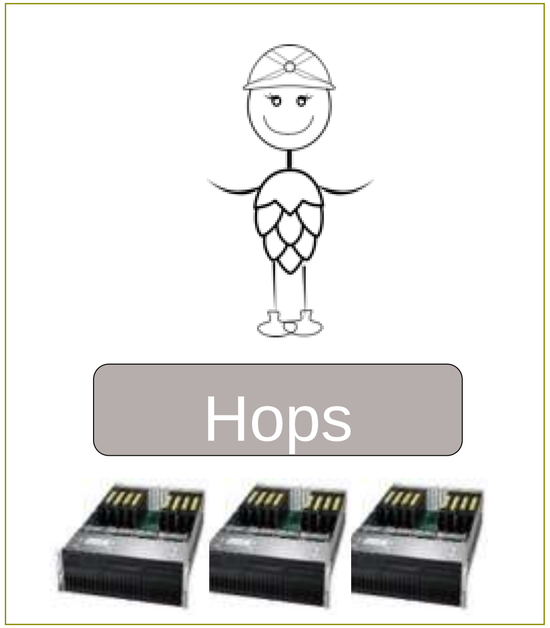
1000 physical servers
250 kW
200 TB RAM
Upto 10 petabyte storage
20 000 cores
80 GPUs
HDFS clusters
OpenStack ECC
OCP servers

RI.SE

Business model



SICS ICE web with all offerings

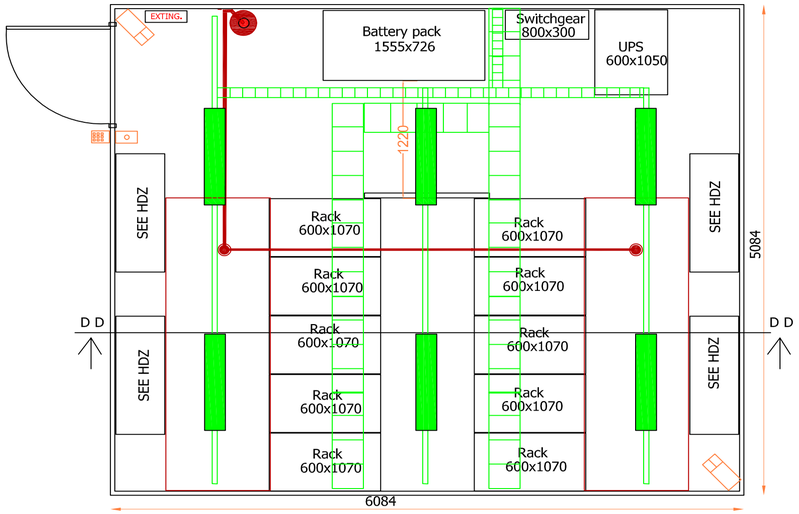


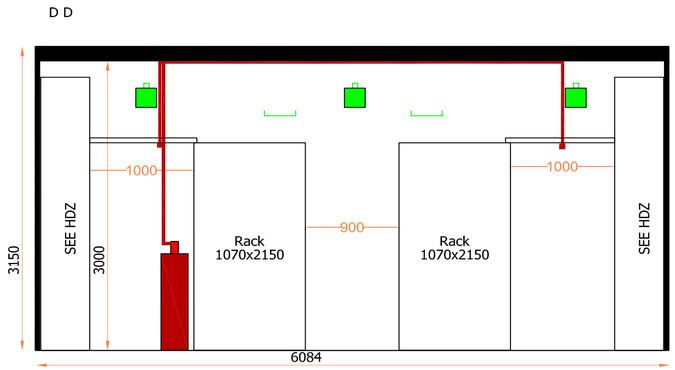
https://ice.sics.se

ICE - Module 1

Web-scale Datacenter Lab

HOT ISLE SEALED





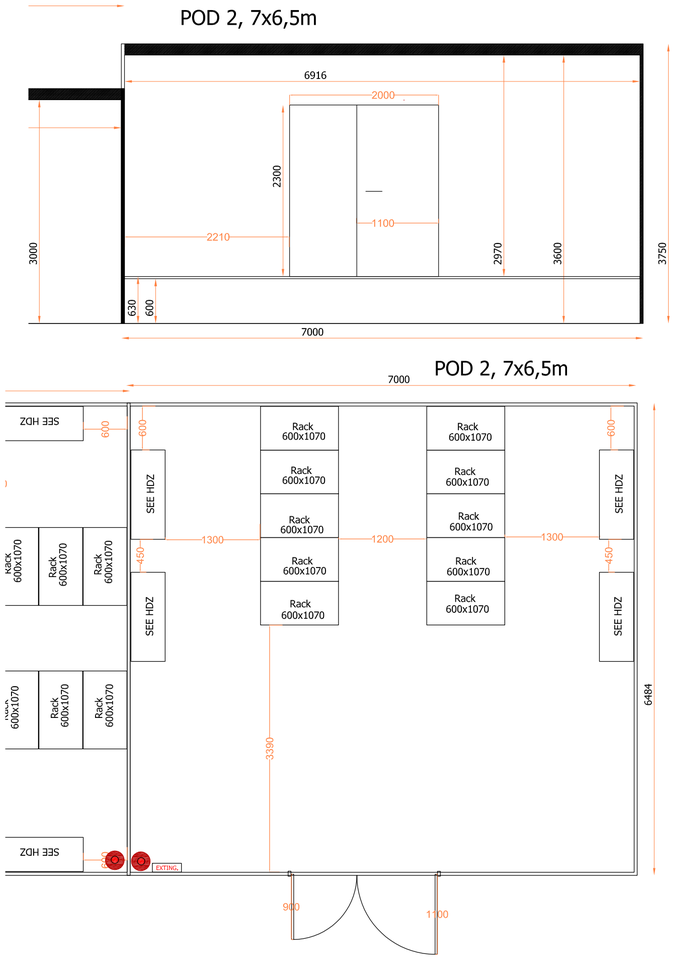
”Ordinary” web-scale data center with optimized and stable environmental conditions

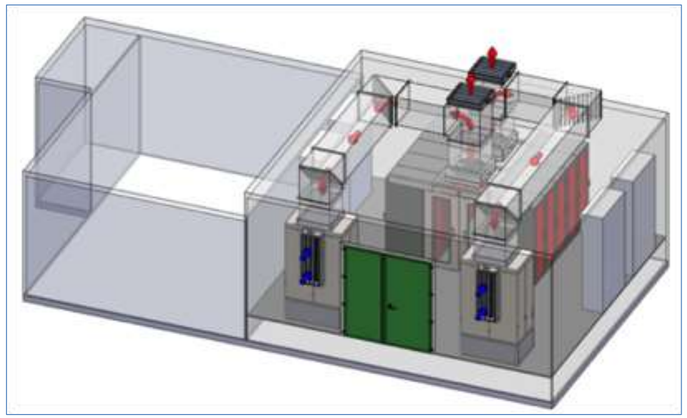
Mainly for datacenter monitoring and modeling and testing of data analytics and IT/cloud-related applications

- 5000 cores, Dell R730
- 40TB RAM
- Up to 7PB storage
- 80kW
- 10/40 Gb/s network
- Separate management network
- GPU accelerated Hadoop clusters
- 80 GPUs
- Openstack racks for ECC
- One OCP rack with 40 OCP windmill servers

ICE - Module 2

Datacenter Facility Lab





A flexible data center facility lab, with unique possibilities to vary environmental conditions

Mainly for testing of data center facility installations and equipment

- Dell servers R530/430
- SICS East HPE cluster, HOPS staging cluster, Nyx Hortonworks cluster
- Prepared for both liquid cooling and free air cooling solutions
- Quick couplings for water, power and fiber
- Flexible raised floor solution to be able to compare with slab floors – reconfigurable.

ICE – Module 3

Open Compute Project Lab

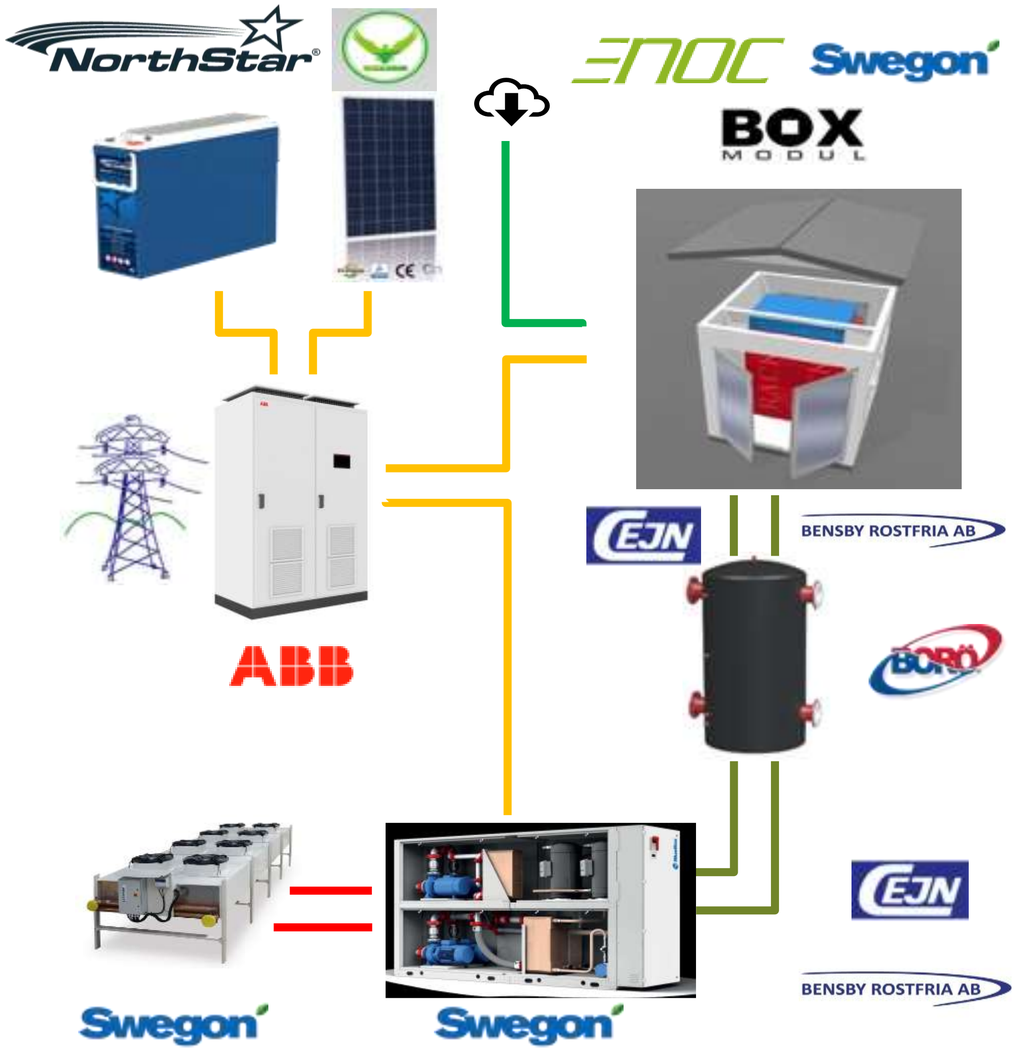
A flexible data center facility lab, focused on OCP equipment

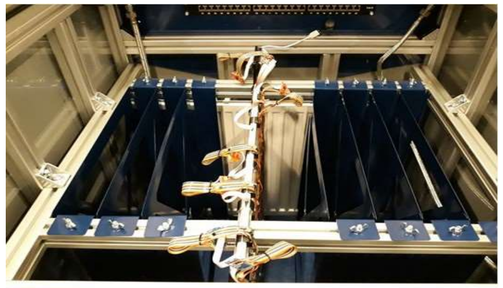
Mainly for testing of data center facility installations and OCP equipment

- OCP windmill servers
- 480 servers / 144 kW peak
- Prepared for free air cooling
- AC/DC power distribution using Trilogy HPDUs

ICE – Experiment area

Micro grid – Edge – Rack Wind tunnel





An experimental area for edge datacenter integration with a smart electrical micro grid and cooling network

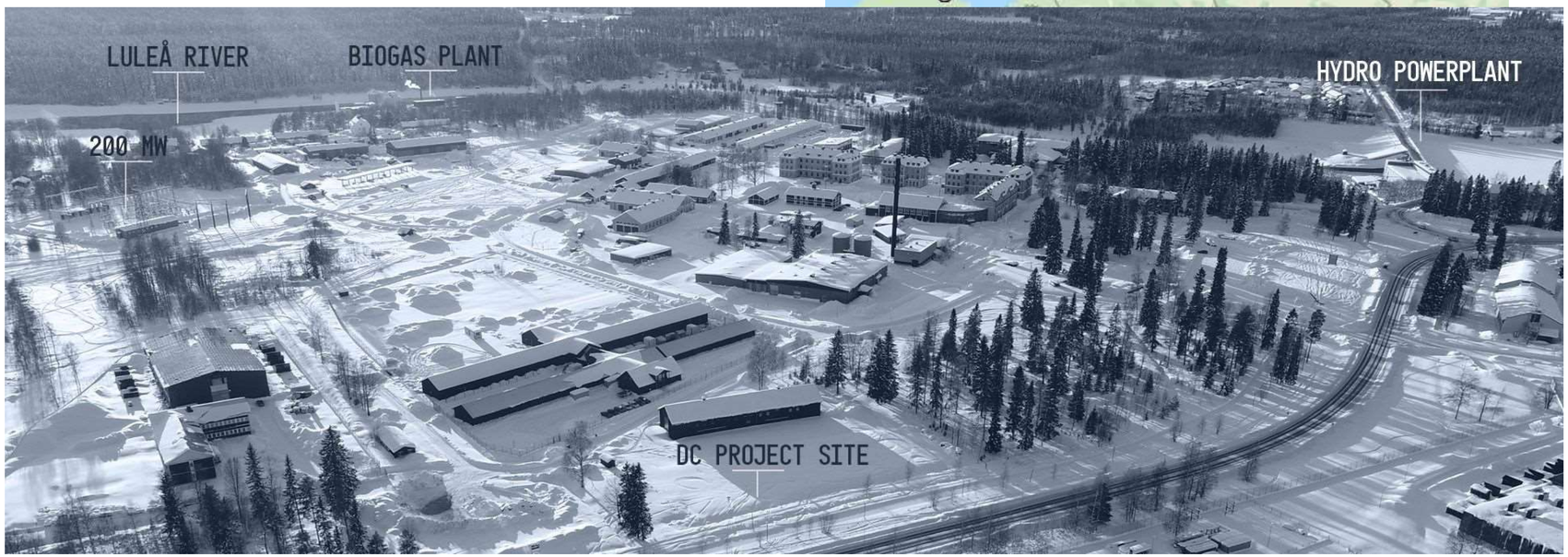
Mainly for testing of data center facility integration, peak shaving, operations

- 10 kW solar panels
- 30 kWh electricity storage
- 22 kW cooling power
- 2 m3 cold water storage
- 10 kW IT-load

Immersion based liquid cooling experimental set-up

Boden Type Data Center H2020 Project

- The Boden Type Data Center (BTDC) project is funded by the EU
- To build the most cost and energy efficient sub 1MW prototype data center in the world!



BODEN TYPE DC ONE

GA 768875

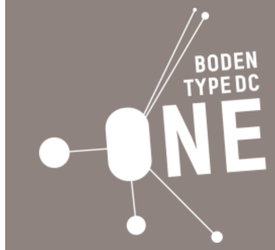
Visit https://bodentypeDC.eu

- Three Data Center PODs
- Designed for 500kW
- OCP Test POD will take contents of RISE North Module 3
- POD2 and 3 are designed for HPC/GPU donners for a target 350kW.

RI. SE

BodenType DC H2020 Project

- Rapid growth in data center energy consumption in Europe from 86TWh in 2013 to an anticipated 104TWh in 2020 (reference P.Bertoldi from the EU JRC).
- Aim of BodenType DC project to build a prototype data center that is energy efficient – needs a comprehensive monitoring and measurement tool.
- Efficiency is focussed on reducing power losses – no UPS, reducing cooling power consumption – using direct air (measured comparison with other methods) and better utilization of the IT systems – workload deployment and management.



GA 768875

Visit https://bodentypeDC.eu

- The Project partners are:

RI.
SE

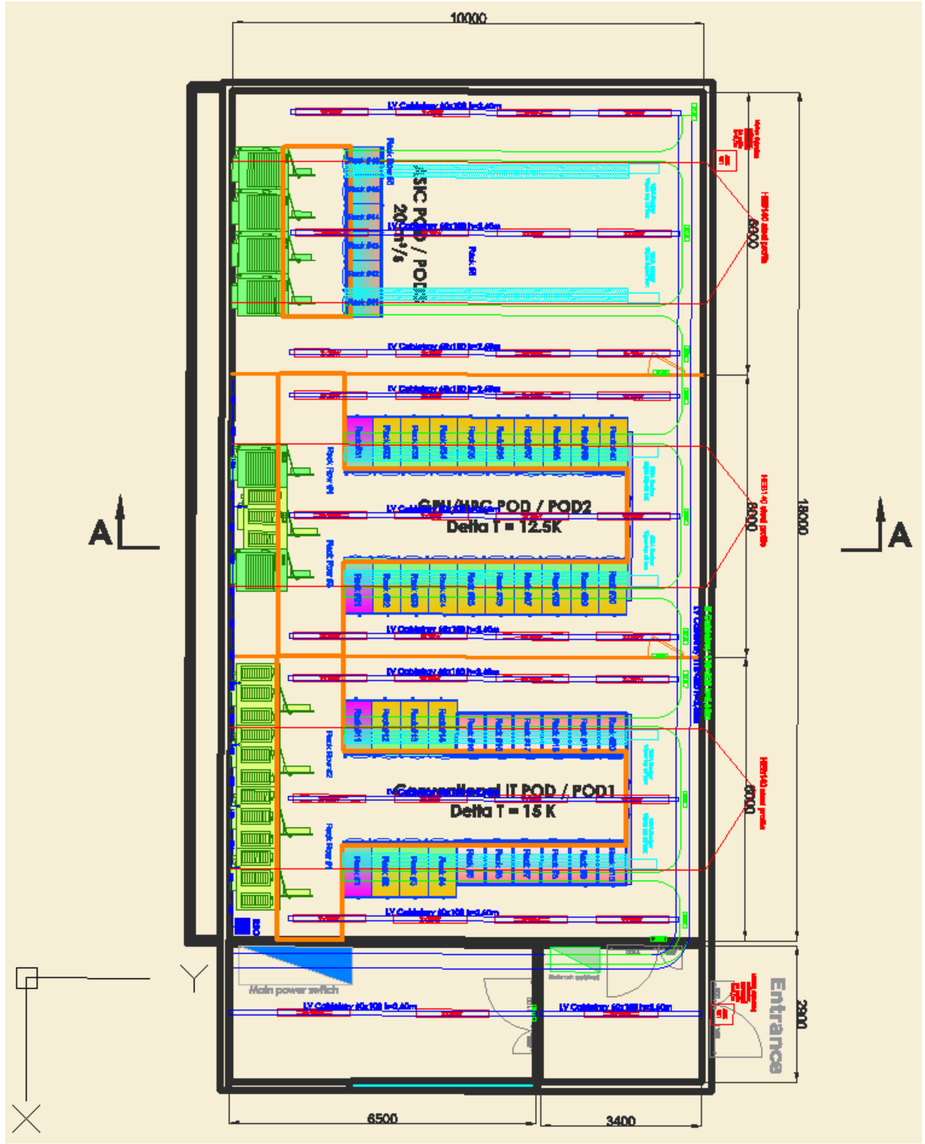
Construction of the Boden Type Data Center

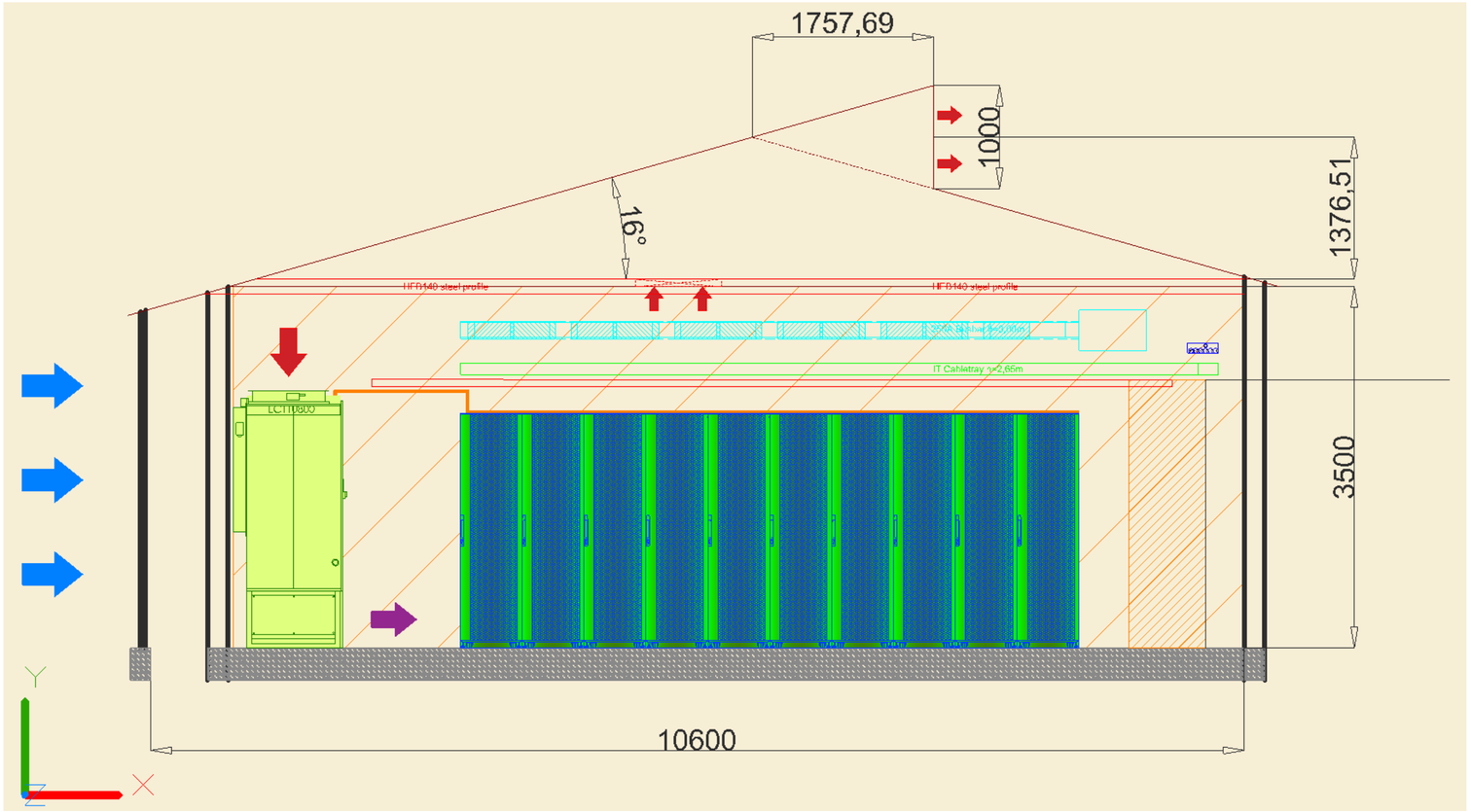
GA 768875

Visit https://bodentypeDC.eu

Conceptual design of the Boden Type Data Center

GA 768875



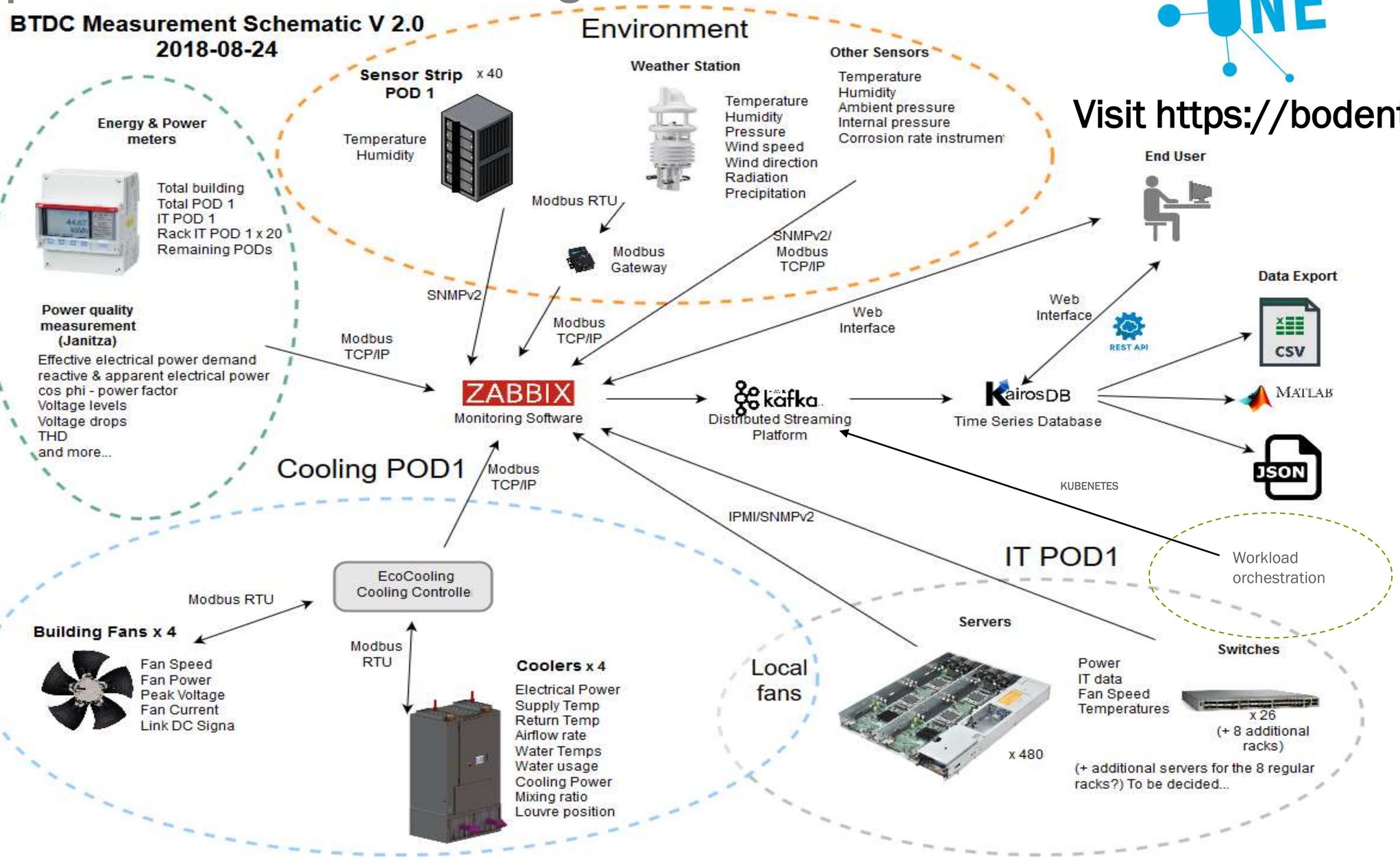


Visit https://bodentypeDC.eu

Open source monitoring of BTDC One

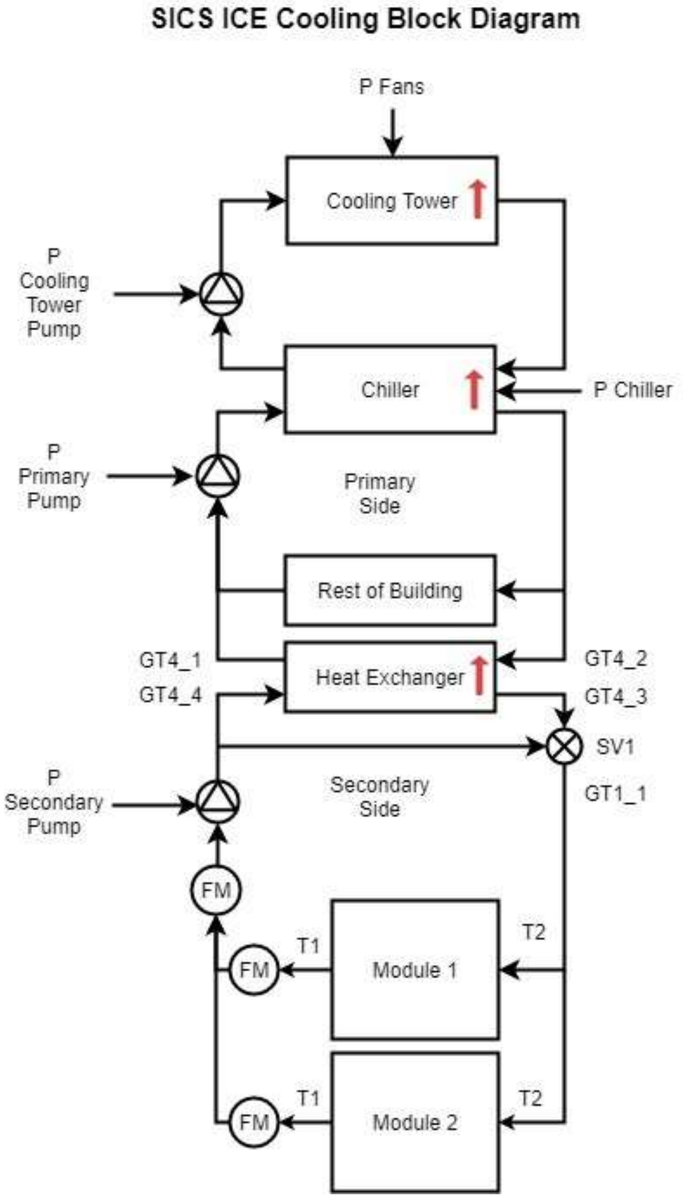
GA 768875

Visit https://bodentypeDC.eu



Montoring data center Modules 1 and 2 (RISE ICE)

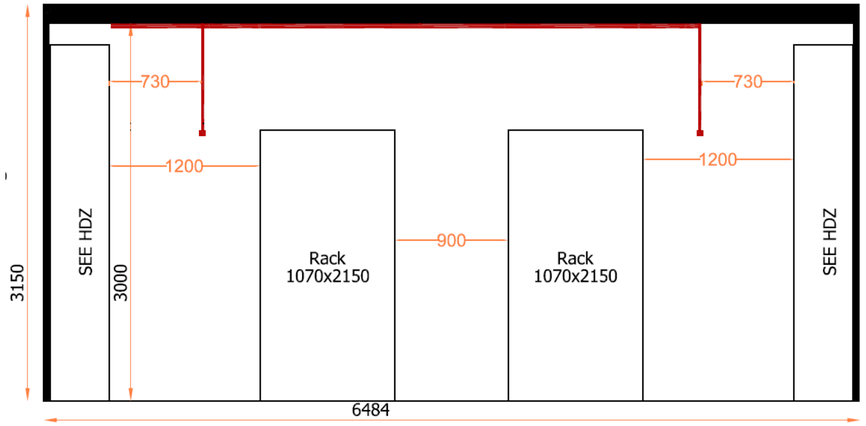
- IT
  - Server load
  - Server temps and fan metrics
  - Network load
  - Other IT information
- Facility
  - Temperature
  - Humidity
  - Power/per server/per rack
  - Cooling power
  - Cooling fan power
  - Etc. etc.



FM Flow Meter

Mixing Valve

Pump



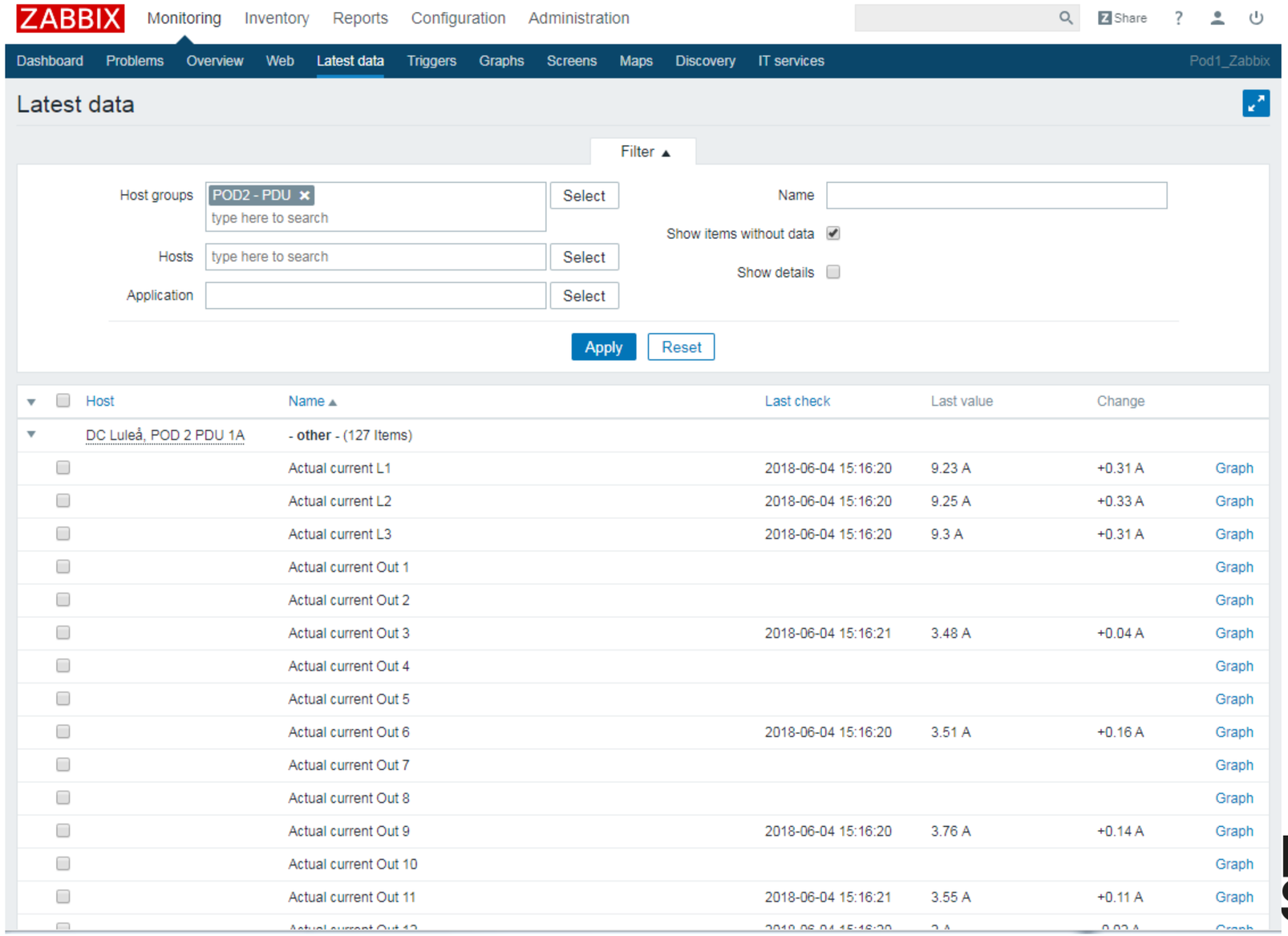
Analytics – What is the data be used for?

- The data has been used for
  - Energy statistics and operational metrics
    - PUE and other metrics. Continous or ”batch”.
    - Is the data center operating in the best way possible?
  - Condition monitoring
    - Server behaviour
    - Automatic fault identification and root cause analysis
  - Optimization and control (closing loop)
    - Load balancing
    - Operational schemes, temperatures
  - Experimental evidence and verification
    - Validating simulations results (e.g. CFD)
    - Fine grained energy savings from strategies and tools

RI.
SE

Data collection

- Open Source Tool
- Protocols to be read include
  - SNMP, IPMI, Modbus, M-bus
- Useful in data center operation
  - Track what is happening
- Hard to export data for post-analysis
- Difficult to do efficient searches.



Long term storage – Need a data structure

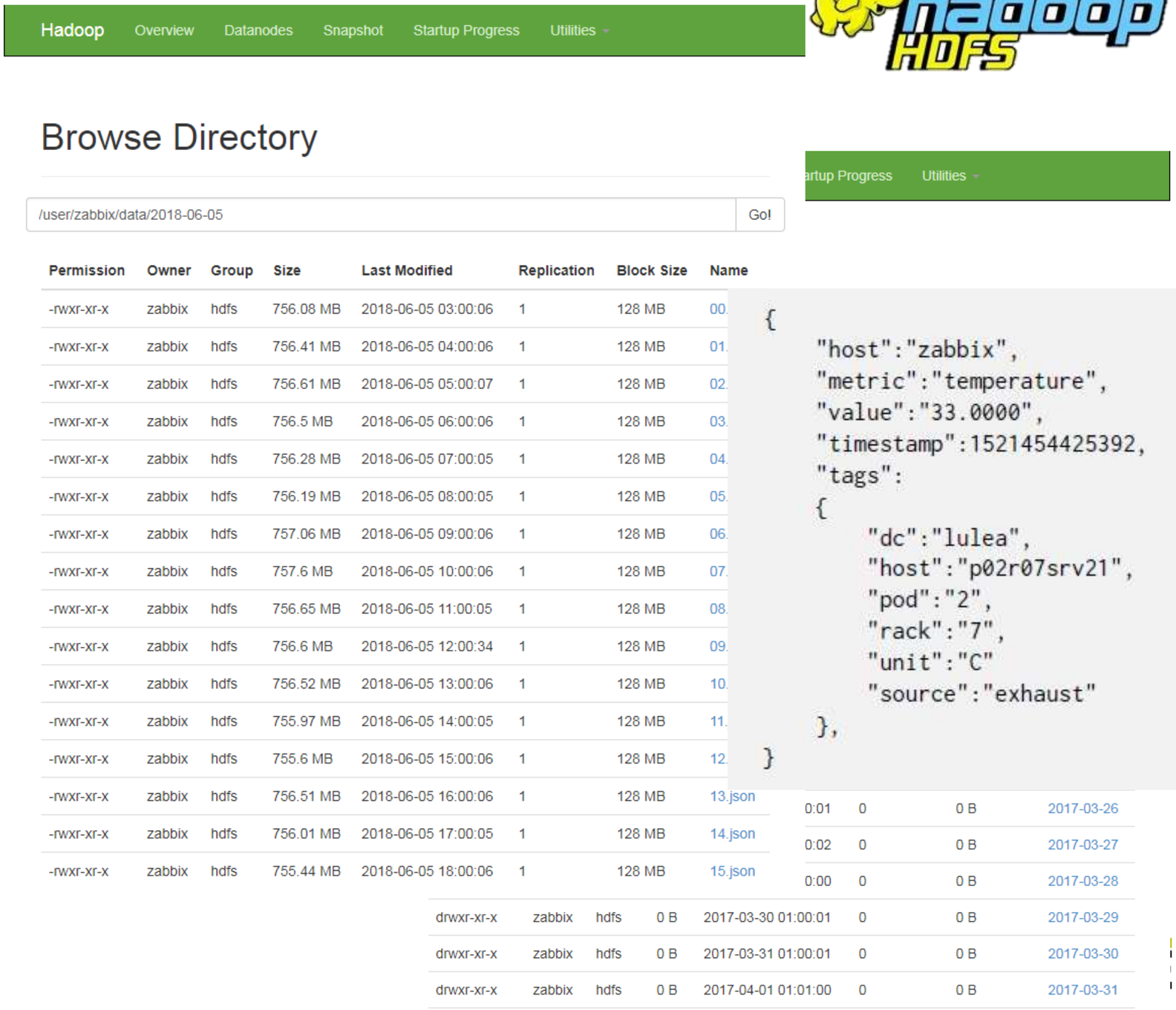

```
<metric name> <time stamp> <value> <tag> <tag>...
```

- **Metric.** The metric should represent a specific “thing” or measurement, like Ethernet packets or temperature. The definition of what the metric represents is defined with the <tag>s.
- **Tag(s).** The <tag> field applicable in the SICS ICE installation and can have the following keys:
  - **dc** Defines what specific data center in the case of several data centers being monitored.
  - **pod** Data center module (POD) number, 1, 2, 3 etc.
  - **rack** The rack number in the pod. (This is left empty for equipment not mounted in rack.)
  - **host** Defines which host the values are read from, servers, weather station, CRAC, etc. (This data will in some cases be redundant.)
  - **source** What is the source of the measurement on the current host. For temperature metrics this could e.g. be cpu, ram, exhaust, top-front, etc.
  - **unit** The unit the metric is represented in.
  - **id** If one host has multiple sources of the same type id is used to distinguish between them, e.g. there can be multiple cpus on the same host.
  - **opt** Optional key, very rarely used.

RI.SE

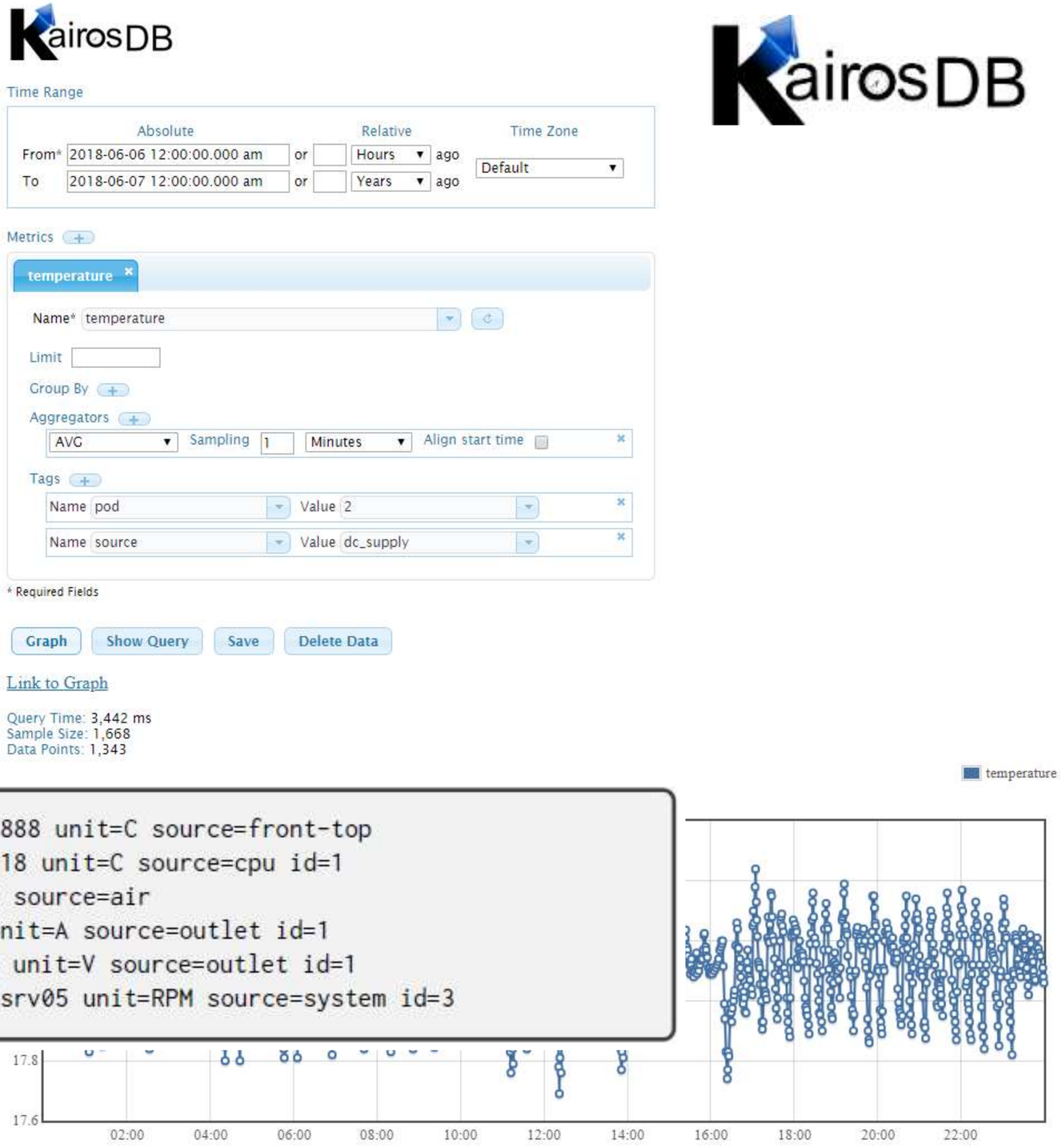
Data access - HDFS

- Divided in hourly files, in date folders.
- All data stored in plain text files.
- Human readable J-son format.



Data access
- KairosDB

- Same information stored as in the HDFS
- Web-interface
- Possible to search, plot and export.

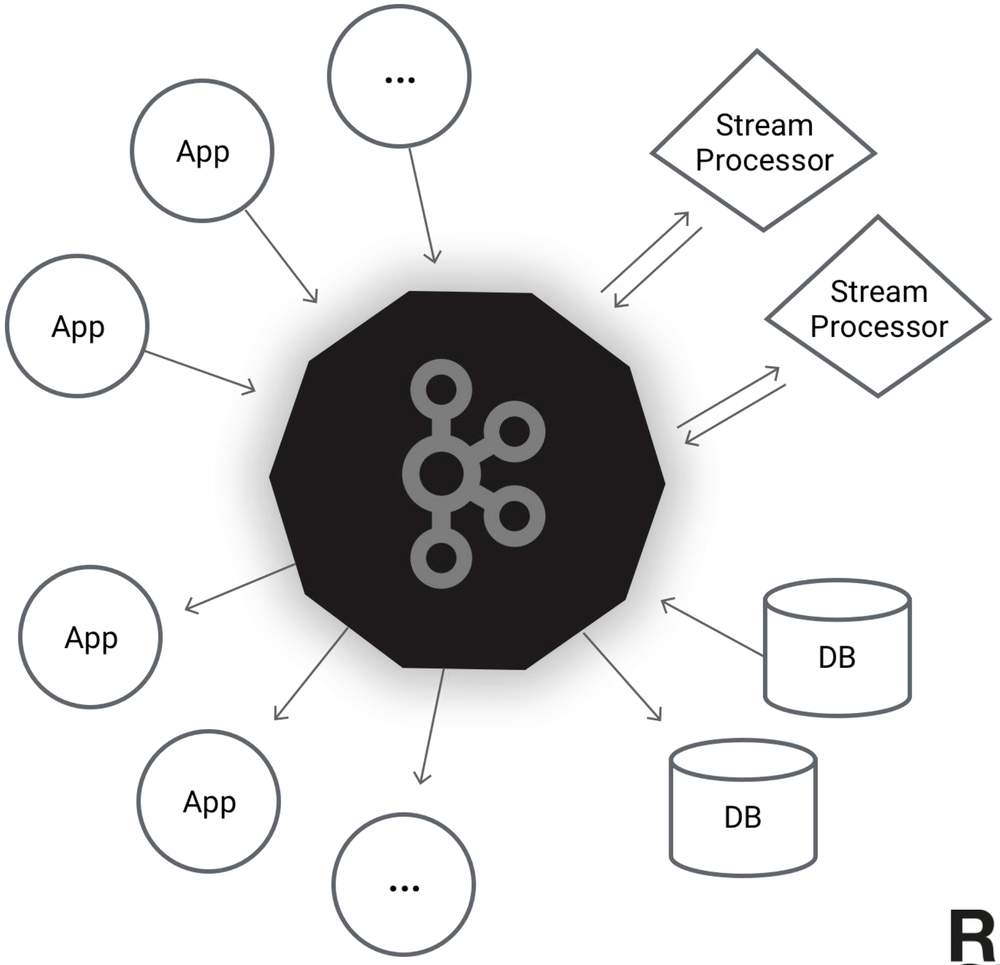



```
temperature <time> 23.1 dc=lulea pod=1 rack=3 host=p01r02emx888 unit=C source=front-top
temperature <time> 45.8 dc=lulea pod=1 rack=5 host=p01r05srv18 unit=C source=cpu id=1
temperature <time> -1.2 dc=lulea host=weather_station unit=C source=air
current <time> 1.1 dc=lulea pod=1 rack=5 host=p01r05hpdu18 unit=A source=outlet id=1
voltage <time> 231.1 dc=lulea pod=1 rack=5 host=p01r05hpdu18 unit=V source=outlet id=1
speed_rpm <time> 2280.0000 dc=lulea pod=2 rack=8 host=p02r08srv05 unit=RPM source=system id=3
```

Data ”transformation”

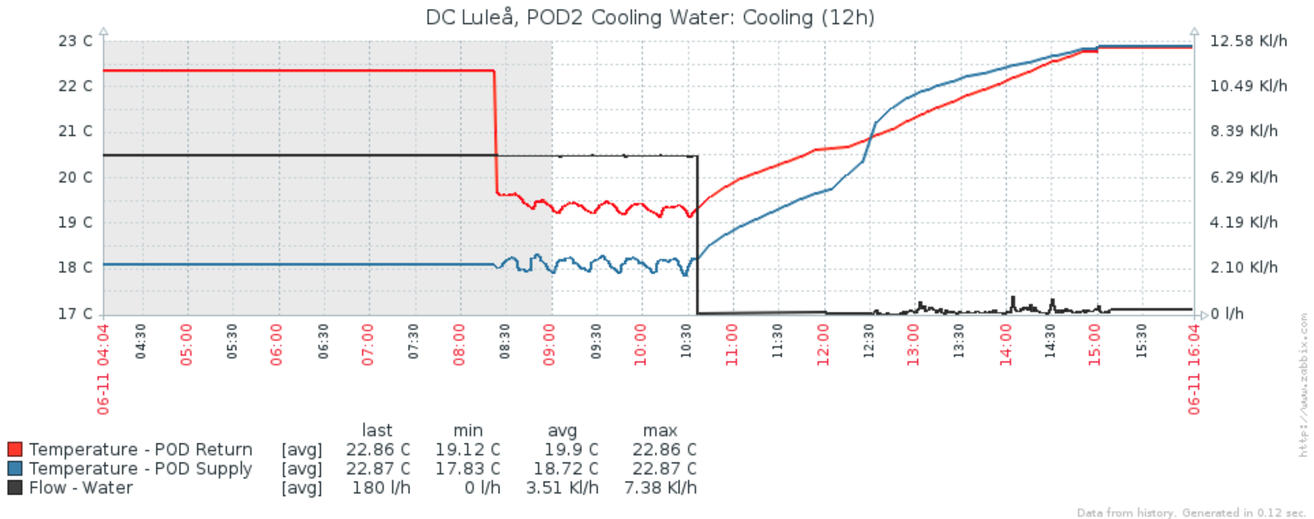
- How to get the data from Zabbix to the long term storage?

- Kafka
  - Distributed streaming platform
  - Reliably get data between systems or applications
  - Publish and subscribe (useful for smaller experiments)
  - Zabbix input module
    - Read data from Zabbix DB and get it into the Kafka ”stream”
  - KairosDB output module
    - Extract from Kafka and write to KairosDB
  - HDFS output module
    - Extract from Kafka and write to HDFS

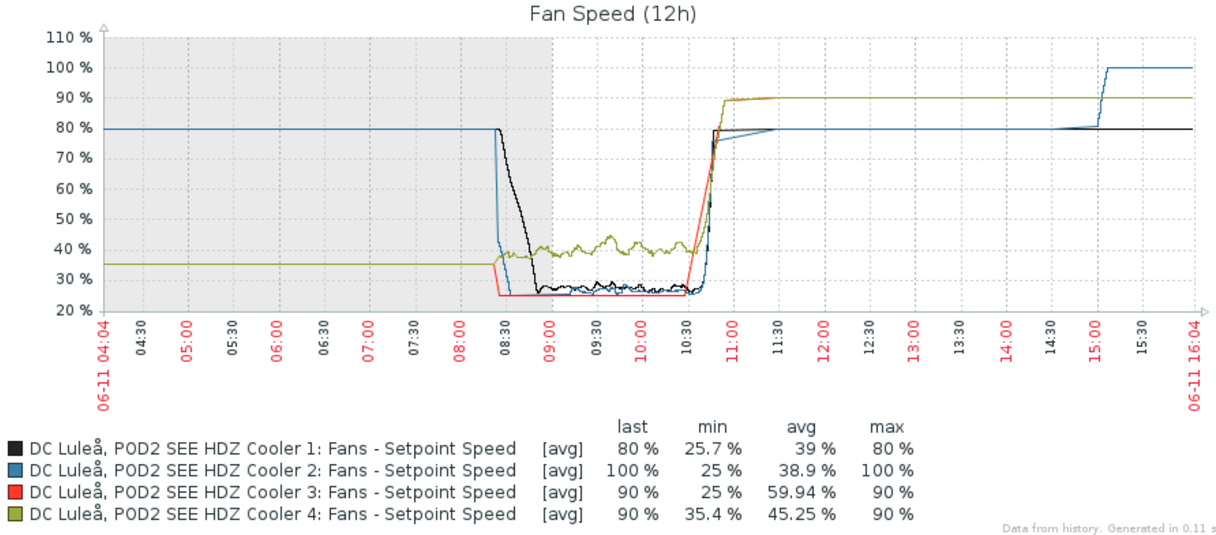


RI. SE

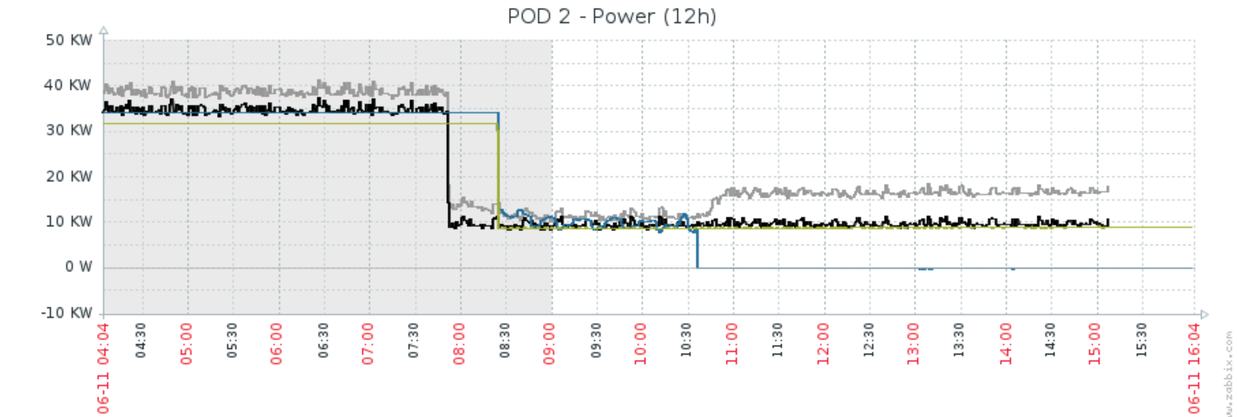
Example data, POD2 maintenance work



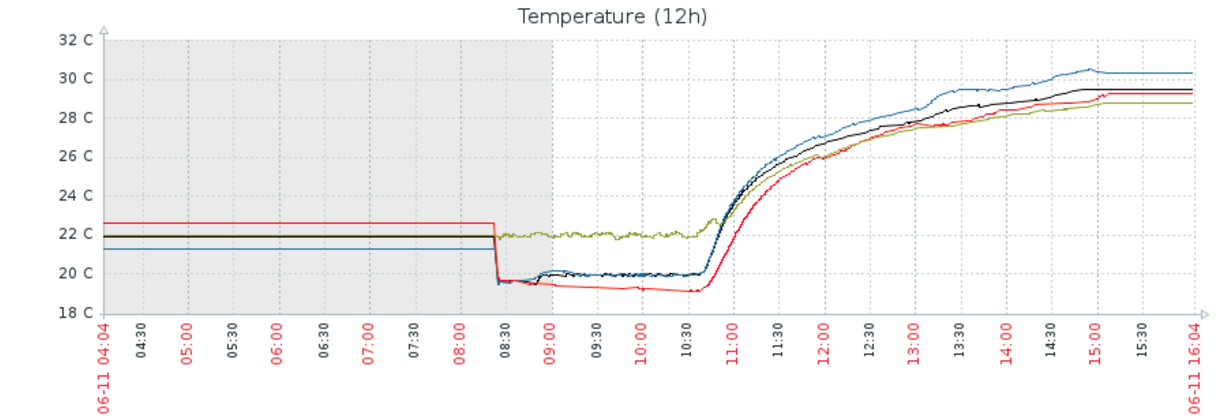
		last	min	avg	max
Temperature - POD Return	[avg]	22.86 C	19.12 C	19.9 C	22.86 C
Temperature - POD Supply	[avg]	22.87 C	17.83 C	18.72 C	22.87 C
Flow - Water	[avg]	180 l/h	0 l/h	3.51 Kl/h	7.38 Kl/h



		last	min	avg	max
DC Luleå, POD2 SEE HDZ Cooler 1: Fans - Setpoint Speed	[avg]	80 %	25.7 %	39 %	80 %
DC Luleå, POD2 SEE HDZ Cooler 2: Fans - Setpoint Speed	[avg]	100 %	25 %	38.9 %	100 %
DC Luleå, POD2 SEE HDZ Cooler 3: Fans - Setpoint Speed	[avg]	90 %	25 %	59.94 %	90 %
DC Luleå, POD2 SEE HDZ Cooler 4: Fans - Setpoint Speed	[avg]	90 %	35.4 %	45.25 %	90 %



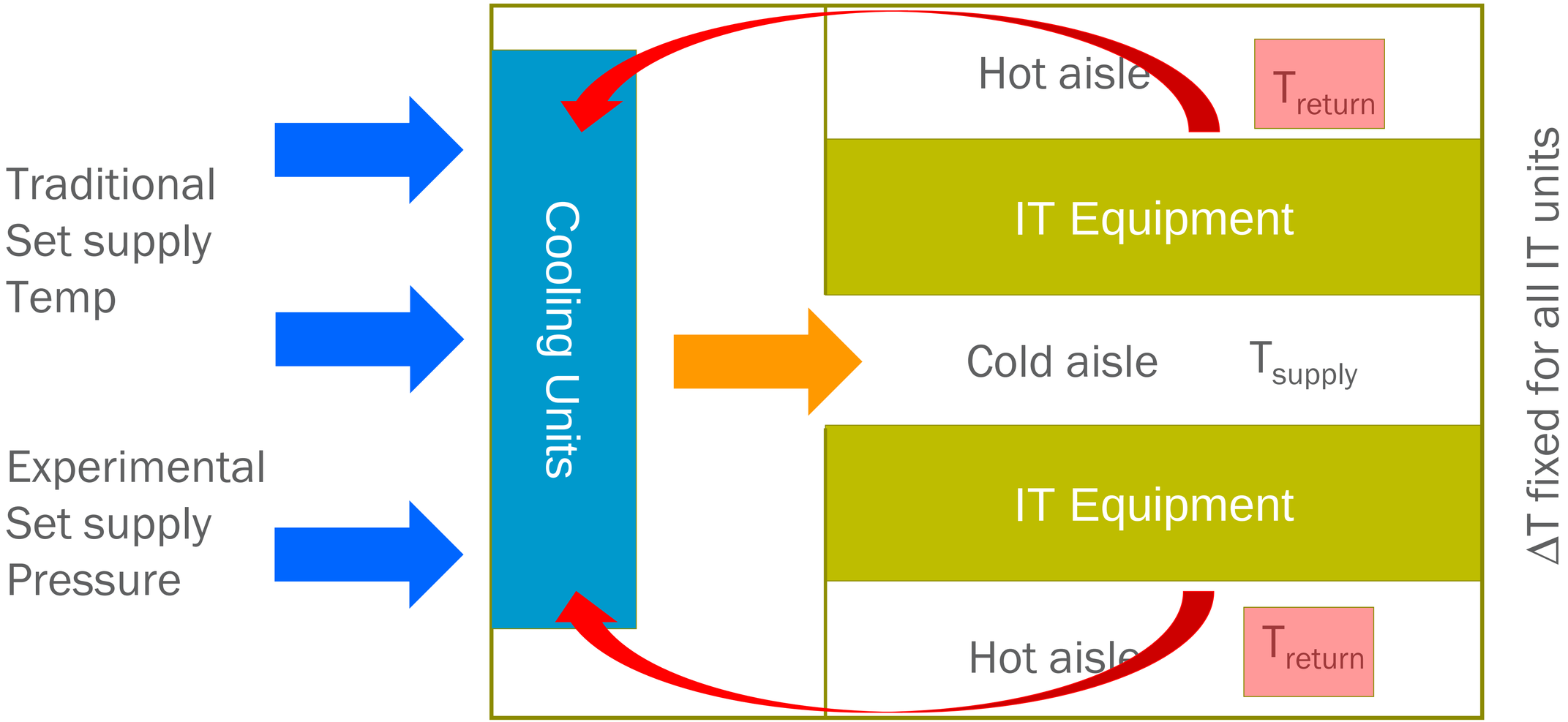
		last	min	avg	max
DC Luleå, POD2 Aggregator: P02 Electrical Power	[avg]	18.12 KW	10.25 KW	23 KW	41.43 KW
DC Luleå, POD 2 Energy Meter IT: Power - Total	[avg]	10.7 KW	8.4 KW	18.35 KW	37.52 KW
DC Luleå, POD2 Cooling Water: Cooling Power - Calculated_test2	[avg]	-1.98 W	-362.84 W	14.33 KW	34.29 KW
DC Luleå, POD2 Aggregator: P02 IT Power Janitza	[avg]	8.85 KW	8.59 KW	17.17 KW	31.71 KW



		last	min	avg	max
DC Luleå, POD2 SEE HDZ Cooler 1: Temperature - Room	[avg]	29.5 C	19.5 C	23.73 C	29.5 C
DC Luleå, POD2 SEE HDZ Cooler 2: Temperature - Room	[avg]	30.33 C	19.5 C	24.25 C	30.5 C
DC Luleå, POD2 SEE HDZ Cooler 3: Temperature - Room	[avg]	29.3 C	19.1 C	23.84 C	29.3 C
DC Luleå, POD2 SEE HDZ Cooler 4: Temperature - Room	[avg]	28.77 C	21.7 C	23.43 C	28.8 C

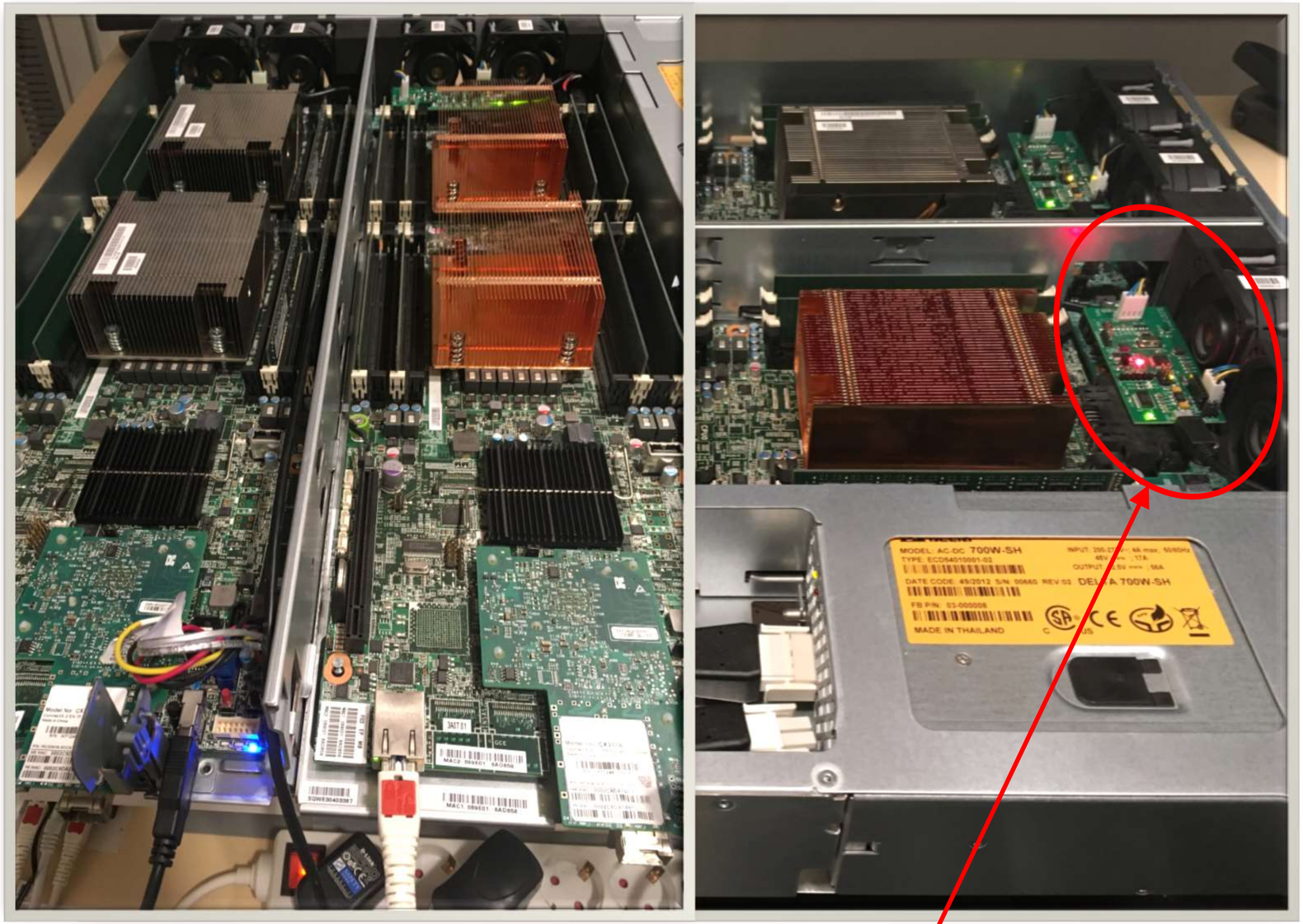
Cooling innovation with the fan controllers

POD 1 – **Cooling Units set T_{supply}**



Wind tunnel to determine ΔT – **IT equipment sets T_{return}**

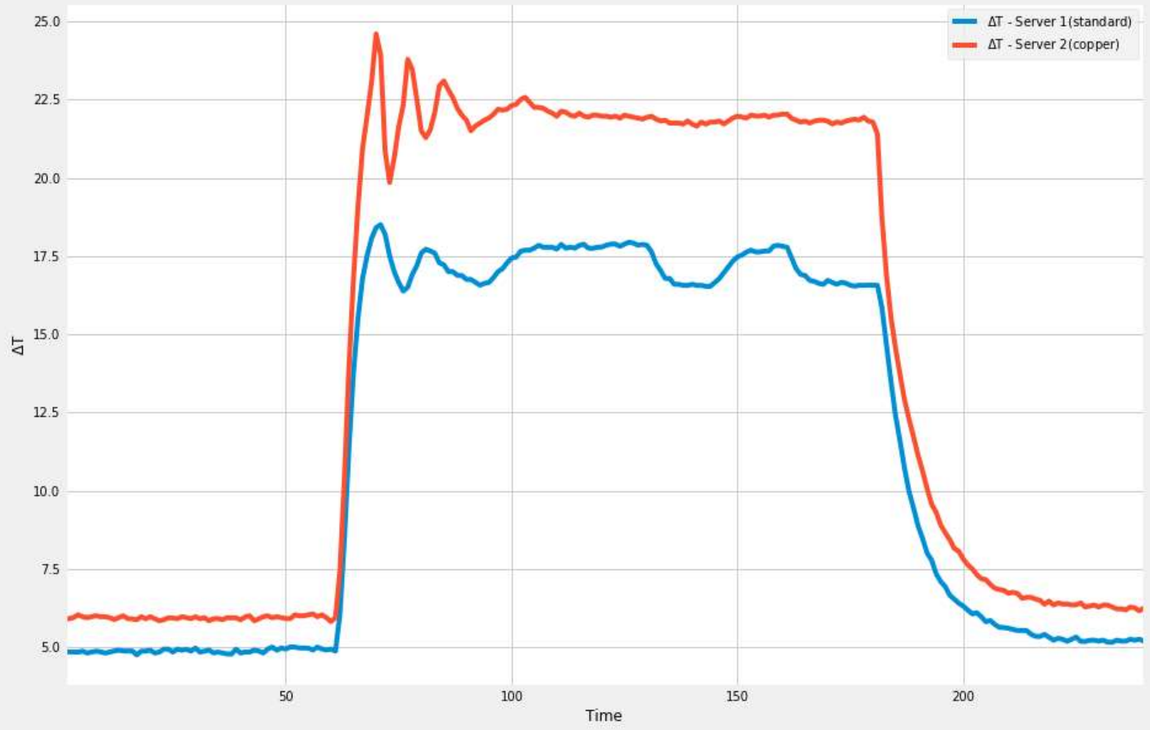
Heat sinks in OCP Servers



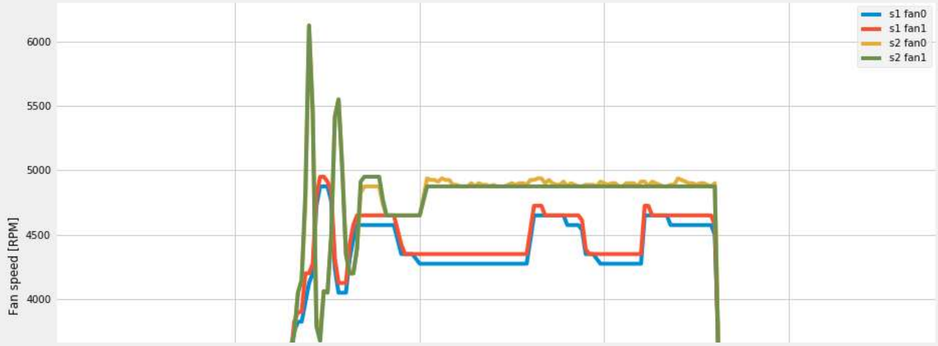
- Replaced aluminium heat sinks containing embedded heat pipes with copper heat sinks.
- Aim to achieve lower flow rates and higher delta temperature across the Windmill servers.
- Flowrate control is also achieved by the use of some developed fan controller cards.

Heat pipe aluminium versus copper heat sinks.

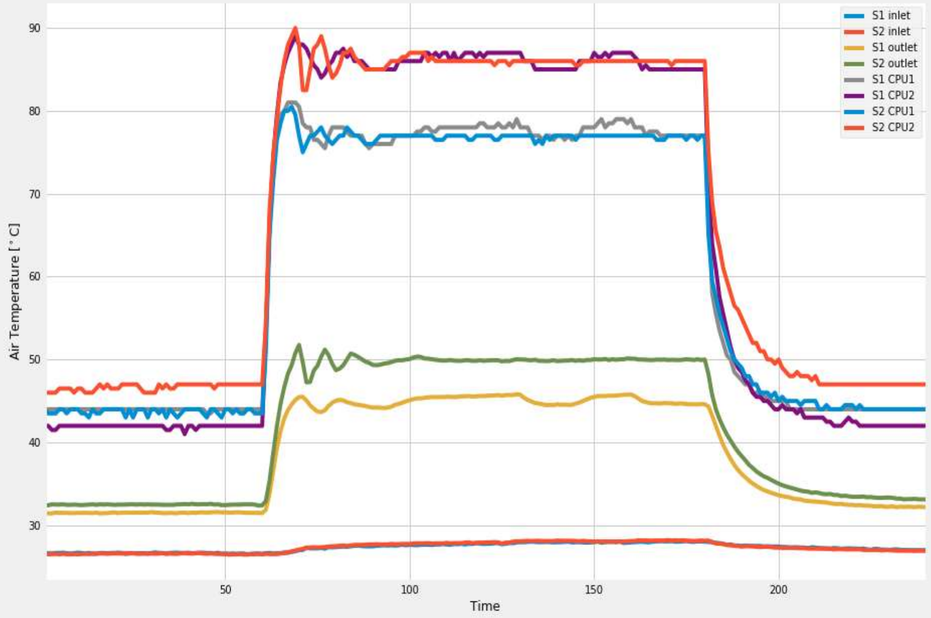
Delta T



Fan speeds.



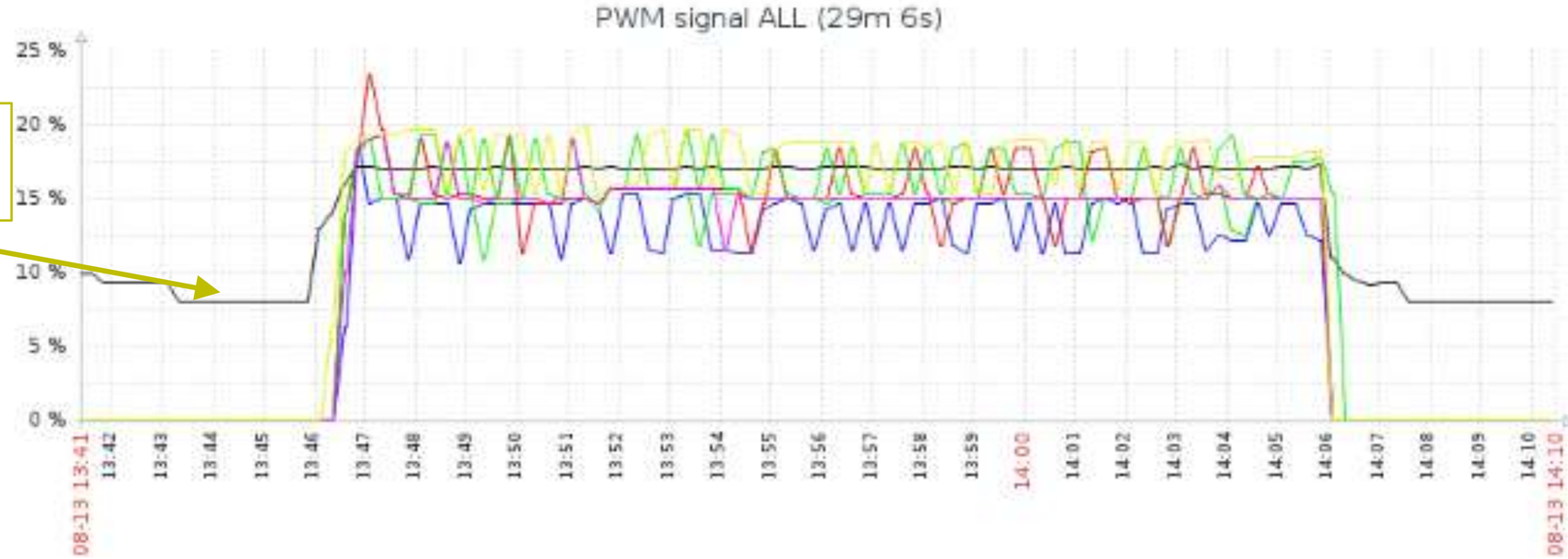
CPU temps.

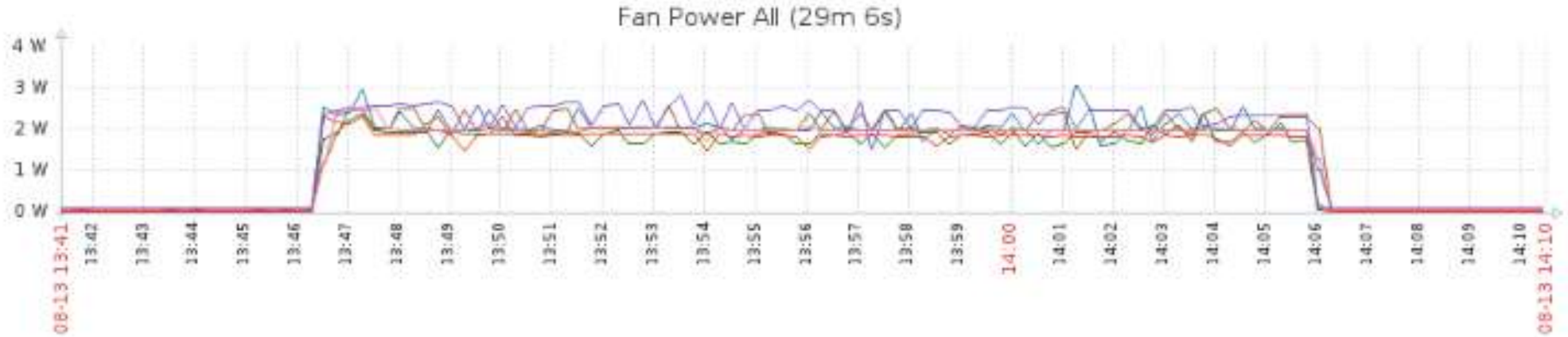


The above are initial results from a desktop experiment. Ongoing work in the wind tunnel for better control and hopefully more conclusive results.

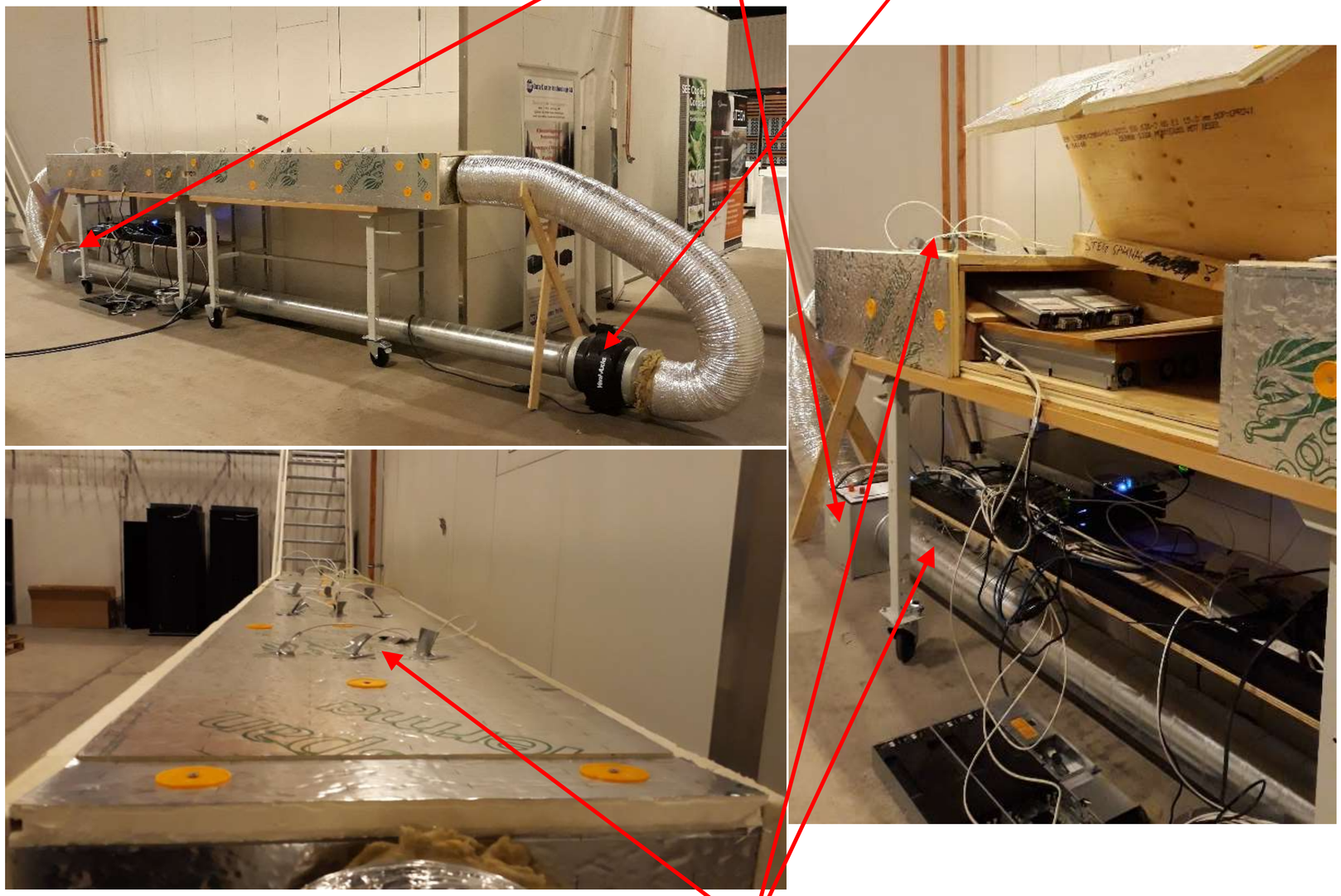
FAN Controller results – taking control over the BIOS

Fan signal from OCP motherboard





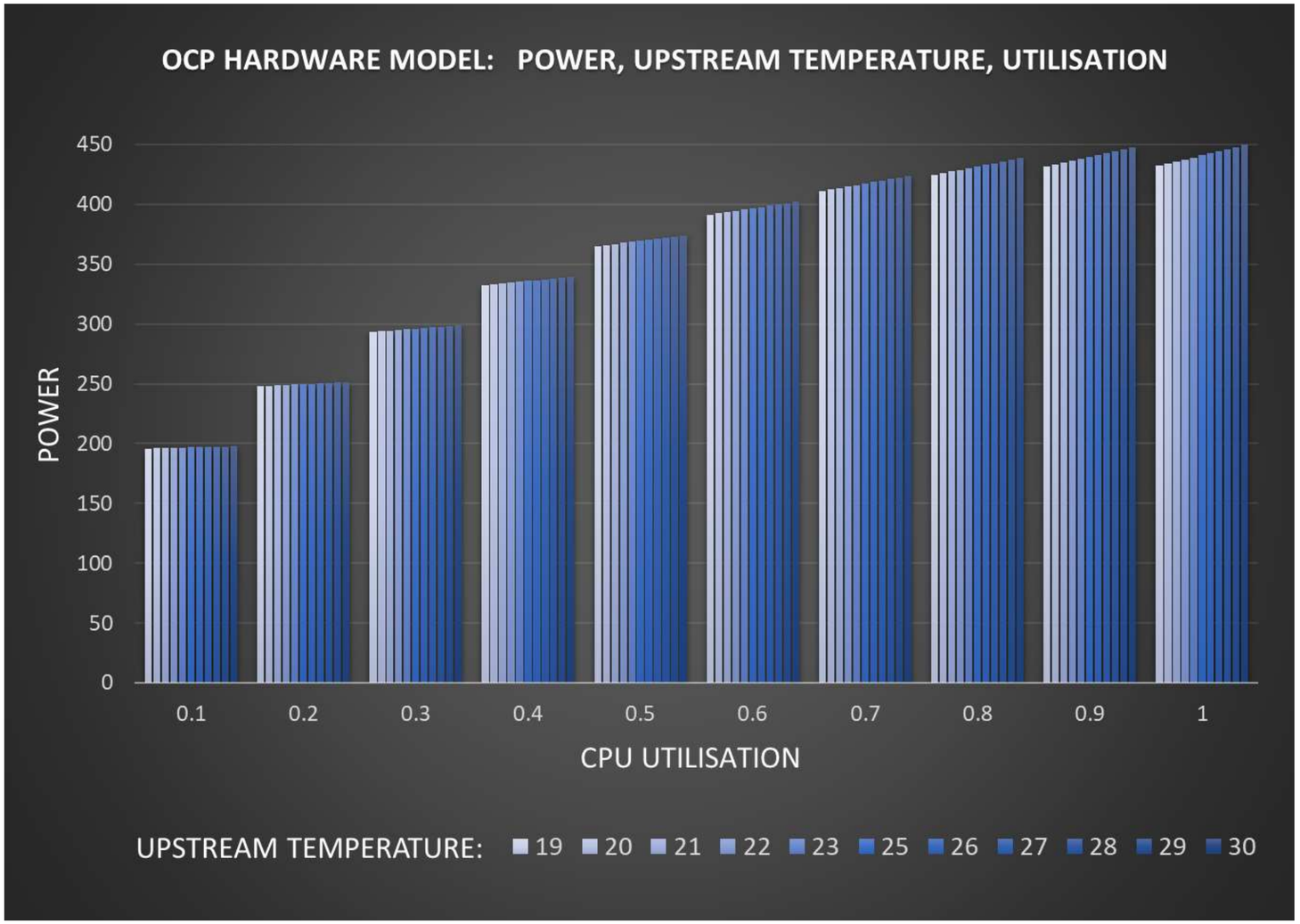
Server wind tunnel



- Set temperature with the HEX
- Flow rate with the FAN – can pressurize the cold aisle.
- Humidity?
- Upstream and downstream temperature measurements.
- Can support between 1U and 4U.

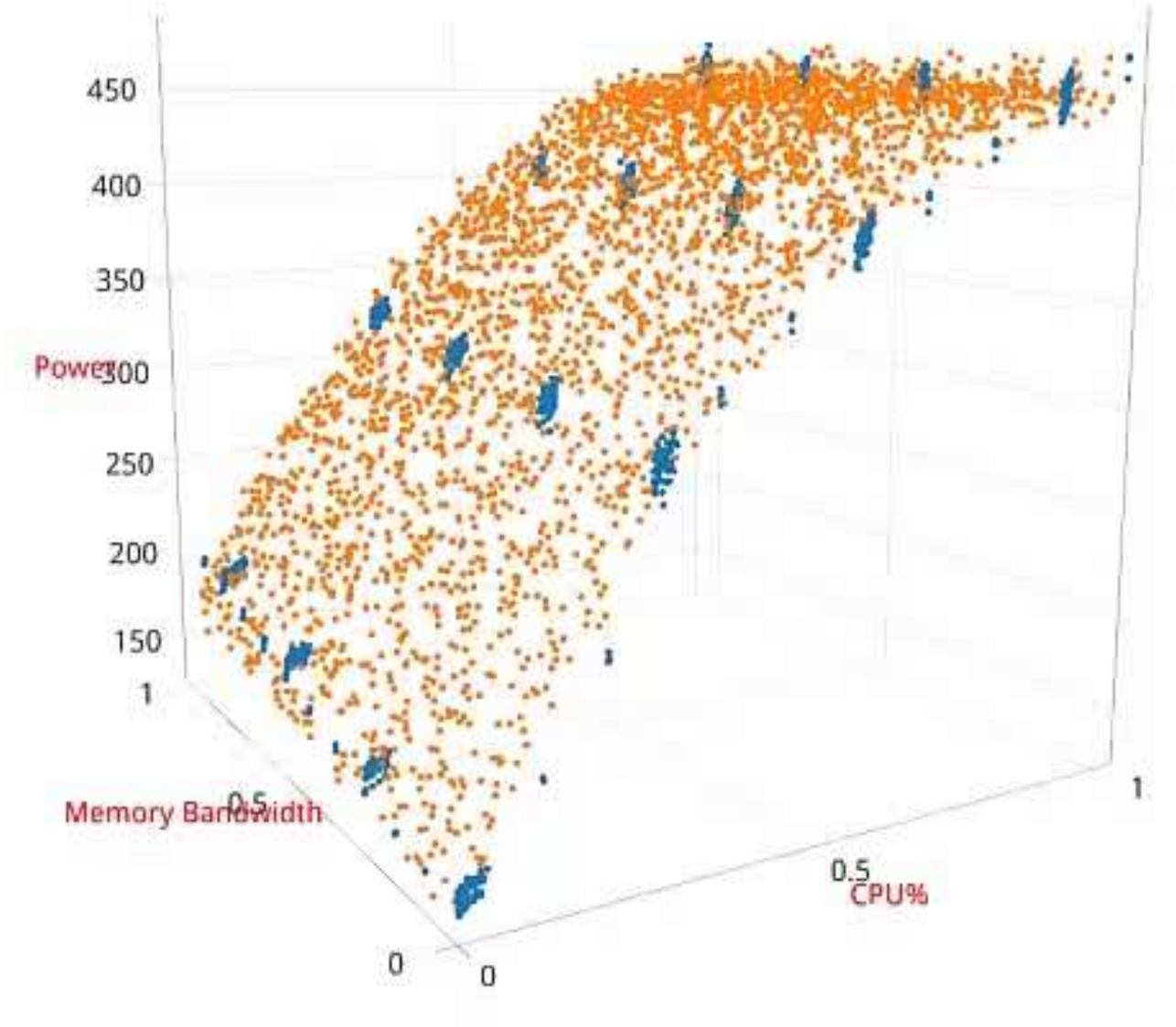
Results with the OCP Windmill Server

- Wind tunnel results are captured using the same data center monitoring tools.



- Uses the same monitoring system as the data center modules
- Enables remote access and control.
- Temperatures are in ^{o}C

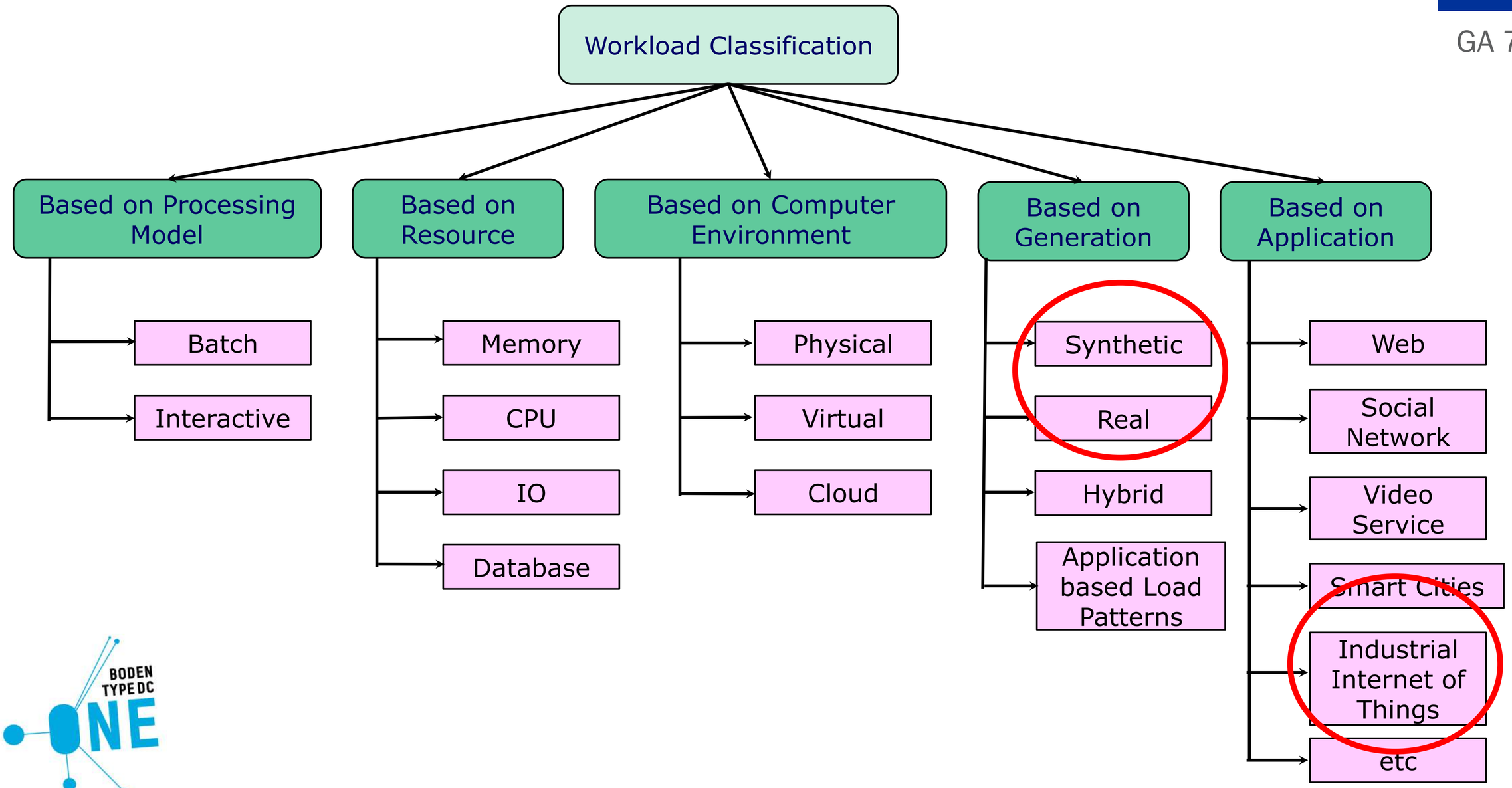
Results with the OCP Windmill Server



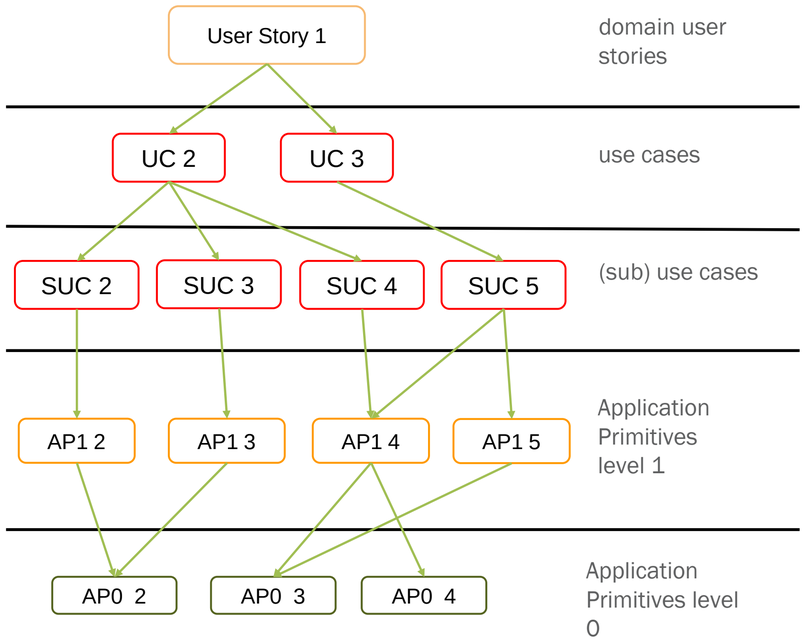
- Thanks to UK startup Edgetic for the results as a collaborator making use of the wind tunnel with its remote control capability and open source monitoring system.
- Blue points are measured and yellow are actually predicted from careful profiling of the OCP system behaviour.

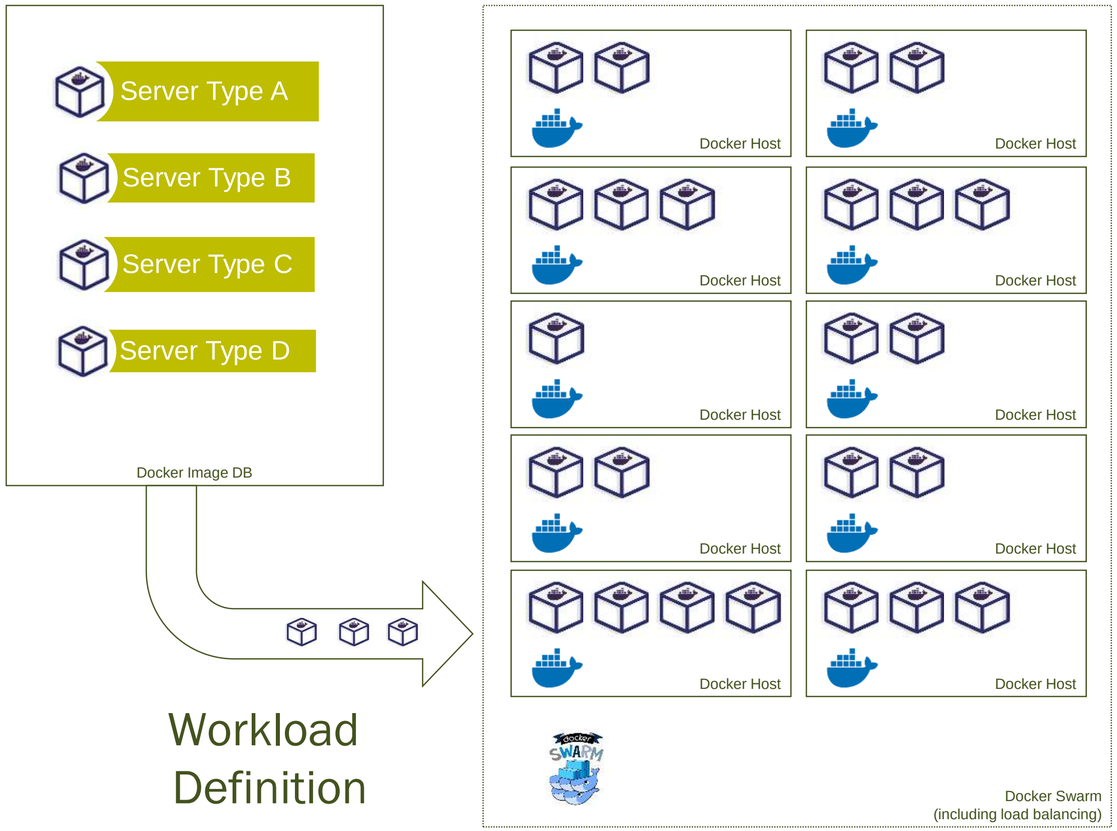
Types of workloads that will operate in the BTDC POD

GA 768875



Building realistic workloads for BTDC project





Building IOT equivalent workloads

GA 768875

SUMMARY

- Operate and monitor experimental test data centers.
- Platform based on open source software to monitor data center facilities, IT utilisation and lab based experiments.
- Created a 150kW IT footprint using legacy OCP Windmill servers, with power and network.
- IT footprint will go into a prototype data center being built in Boden and cooled by direct air with an ambitious linking of facility and server fans.
- Characterized the thermal and power of the OCP windmill server using a combination of heat sinks and locally programmed fan controllers, together with a uniquely developed server wind tunnel.
- The EU funded project has developed a strategy for deploying synthetic and near to real Industrial IOT workloads.

Visit https://bodentypeDC.eu

- Thanks to:
  Colleagues at RISE North
  Filip Blylod
  Sebastian Fredriksson
  Jonas Gustafsson
  Magnus Nilsson-Mäki
  Daniel Olsson
  Jeffrey Sarkinen
- Thanks to VINNOVA for funding SENDATE EXTEND
- Thanks to the EU H2020 funding and the BodenType DC Project partners
  H1 Systems
  Fraunhofer Institute of Optronics
  Ecocooling
  Boden Business Park

BODEN TYPE DC ONE

GA 768875

RI SE

OCP SUMMIT

OPEN. FOR BUSINESS.