Measurement data made available in a public data repository ORDP version

Project acronym: BodenTypeDC  
GA Number 768875  
Call identifier H2020-EE-2017-RIA-IA

Deliverable no. D5.4 version ORDP  
Dissemination Level: [PU]  
Deliverable type: [Other]

Due Date of deliverable: 31 03 2020  
Completion date of deliverable: 08 07 2020  
Revision date of deliverable: 25 09 2020

Lead beneficiary responsible for deliverable:

Related work package: **WP5 Prototype testing, measurement**

Author: Filip Blylod (RISE)  
Sebastian Fredricksson (RISE)  
Jonas Gustafsson (RISE)  
Daniel Olsson (RISE)  
Jeffrey Sarkinen (RISE)  
Jon Summers (RISE)  
Reinhard Herzog (Fraunhofer IOSB)  
Thomas Batz (Fraunhofer IOSB)  
Alan Beresford (EcoCooling)

**Document history:**

|  |  |  |
| --- | --- | --- |
| **Revision** | **Date** | **Status** |
| 0.1 | 30/03/2020 | Initial Draft |
| 0.2 | 16/06/2020 | Decision on ORD Pilot route |
| 0.9 | 08/07/2020 | Final Draft |
| 2.0 | 25/09/2020 | Updated for ORDP release |

**Disclaimer:** The European Commission support for the production of this publication does not constitute endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

**Notice on using EU funds:** This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 768875

**Copyright Notice:** BodenTypeDC Consortium 2017. All rights reserved.

# Executive Summary

The operational period of the BodenType DC project from the perspective of detailed data collection ran during the period of 22nd January 2019 when the data collection system was commissioned to 20th May 2020 when the PODs were put into a state of stasis.

This report supports the data made available as an Open Research Data (ORD) Pilot. Details of what data has been collected has been addressed in earlier project deliverables some of which are closed and confidential documents, but elements of the documents are included for completeness in this version of the ORDP deliverable, D5.4.

**List of abbreviations**

|  |  |
| --- | --- |
| ASIC | Application-Specific Integrated Circuits |
| BTDC | Boden Type Data Center |
| CPU | Central Processing Unit |
| DCIM | Data Center Infrastructure Management |
| DOI | Document Object Identifier |
| IOT | Internet Of Things |
| ORD | Open Research Data |
| PUE | Power Usage Effectiveness |
| UPS | Uninterruptible Power Supply |

Table of Contents

[1. Executive Summary 3](#_Toc52656303)

[2. Introduction 5](#_Toc52656304)

[3. The available datasets 6](#_Toc52656305)

[3.1. Hardware providing data 7](#_Toc52656306)

[3.2. The data center operations behind the data 9](#_Toc52656307)

[3.2.1. Workload mixtures 9](#_Toc52656308)

[3.2.2. Workload distribution strategies 9](#_Toc52656309)

[3.3. The excel spreadsheet 12](#_Toc52656310)

[4. Conclusions 14](#_Toc52656311)

[5. References 15](#_Toc52656312)

[6. List of figures 16](#_Toc52656313)

# Introduction

The BTDC One data center has been operational since January 2019. The data center has 3 PODS which are configured with slight variations in the fresh air-based cooling systems. The datasets that have been collected since operation commenced in PODs 1 and 3 with the operational period from the perspective of data collection using the database system and DCIM system installed and described in an earlier deliverable being from 5th March 2019 to 2nd March 2020 inclusive, so 52 weeks, 364 days and 8,736 hours.

The datasets for POD 1 comprehensively include energy, temperature, relative humidity, fan speeds, average server chip temperatures, average server fan speeds, percentage server utilization and dew point. The POD 3 datasets include average server power consumption and percentage fan speeds, but the period of operation is shorter than POD 1, being only from 26th April 2019 to 24th February 2020 inclusive. The data set also includes weather data from an in-situ weather station and power quality reports from proprietary power analysis software over the full 364-day period.

The data is made available as an excel spreadsheet with the finest time resolution being for hourly averaged data, but the data is for all 364 days, so 8,736 hours.

The datasets are available long term from the Research Data Repository of Fraunhofer-Gesellschaft, Fordatis, which is indexed by the Registry of Research Data Repositories at re3data.org and the data has a Document Object Identifier (DOI) of <http://dx.doi.org/10.24406/fordatis/87> for permanent reference and which should be used as a unique reference for this research data being associated with any scientific publications that makes direct use of the datasets.

# The available datasets

The collection and management of the datasets are depicted in Figure 1, which highlights the different software components that have been configured to enable the implementation of data collection.

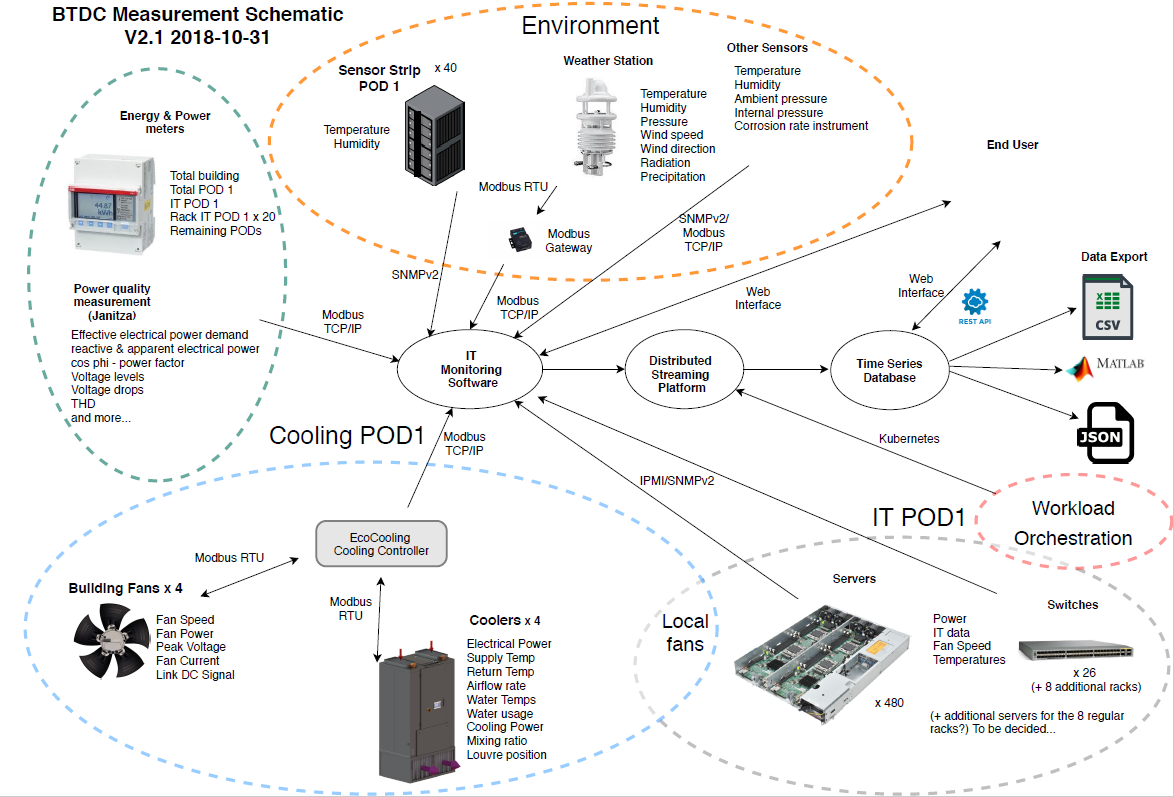
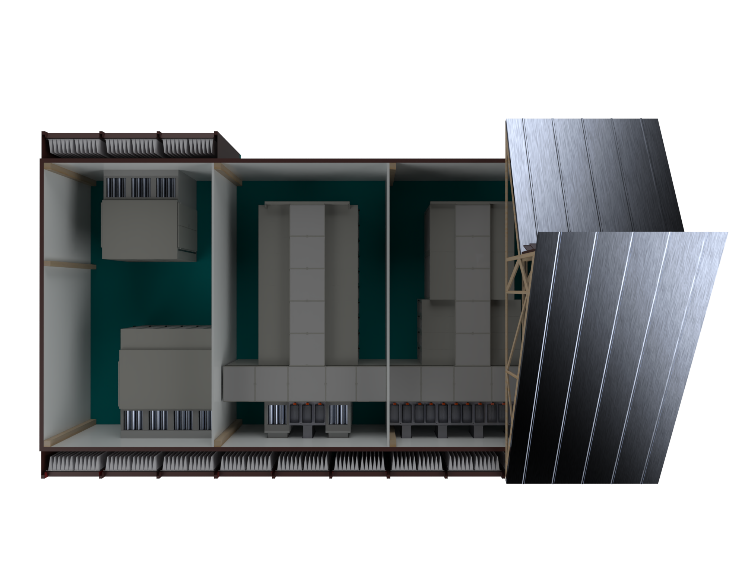


Figure : The data center data collection arrangement.

The physical layout and shape of the data center from both the inside and outside is shown in Figure 2.

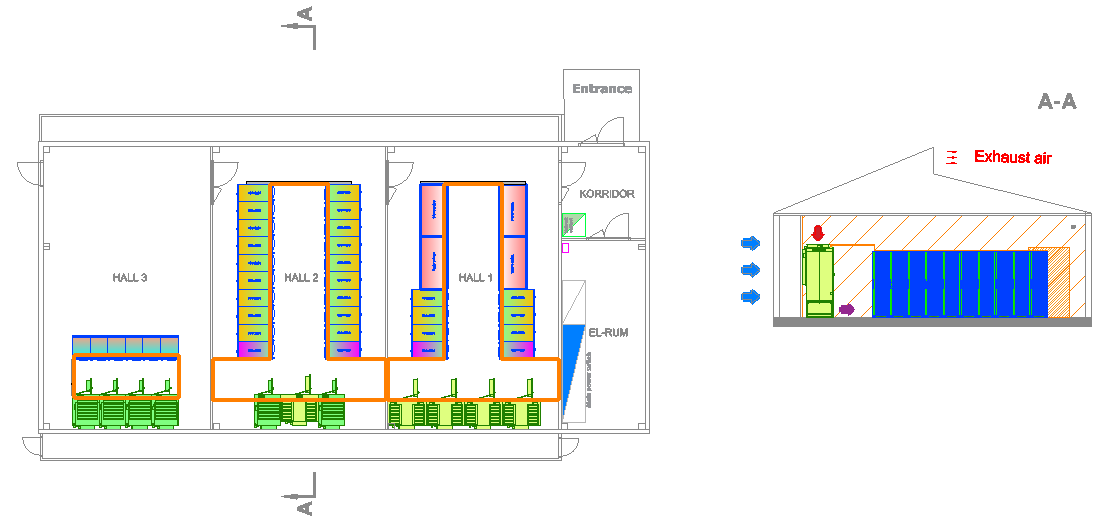
 

X

X

Figure : Left: 3D schematic of the inside of the data center and the sections with an X have not been operation for data collection. Right: Photo of the data center from the outside.

The plan view showing the two PODs mentioned so far, PODs 1 and 3, is presented in Figure 3.



POD 1

POD 3

Figure : Plan view of the internal data center layout showing PODs 1 and 3.

## Hardware providing data

The data that has been collected from equipment is broken down into measured *power* from the devices listed in Table 1 and the movement and consumption of fluids in Table 2. Note that the cooling components ECT10800 provide evaporative function for cooling and humidifying the incoming air as well as ventilation in POD 1. In cooling or humidification mode the coolers consume water and therefore the water meter is monitored throughout the operational period of the data center.

Table Power Measurement Equipment

|  |  |  |
| --- | --- | --- |
| **Measurement Point** | **Device** | **Quantity** |
| Facility Level Power | Janitza UMG604-PRO | 11 |
| Rack Level Power | Janitza UMG604-PRO | 20 |
| Total Building Power | Janitza UMG96RM | 1 |

Table Monitored Facility Level Air Movers and Water Provision to the coolers

|  |  |  |
| --- | --- | --- |
| **Equipment** | **Type** | **Quantity** |
| Cooler | EcoCooling ECT10800 | 4 |
| Fan | Ziehl Abegg EC Blue | 8 |
| Inlet Water Meter | Siemens WFK30 | 1 |
| Drain Water Meter | Siemens MAG5100W | 1 |

The BodenType DC has three PODs and POD 1 is populated with legacy Open Compute Project equipment networked with the IT equipment listed in Table 3.

Table Information Technology Equipment in POD 1

|  |  |  |
| --- | --- | --- |
| **Equipment** | **Type** | **Quantity** |
| Server type 1 | OCP Windmill v2[[1]](#footnote-1) | 61 |
| Server type 2 | OCP Windmill v2[[2]](#footnote-2) | 419 |
| Switch type 1 | Cisco 3064-x | 13 |
| Switch type 2 | Cisco 3048TP | 13 |

The monitoring is also obtained in the backdrop of both internal and external environmental conditions that vary with the seasons and local weather, which has been monitored by a range of equipment listed in Table 4.

Table Environmental Monitoring Hardware Components and Devices

|  |  |  |
| --- | --- | --- |
| **Equipment** | **Type** | **Quantity** |
| Weather Station | Lufft WS10 | 1 |
| Sensor Strip | SICS Prototype v2 | 28 |
| Pressure/Humidity/Temp Sensor | BME280 | 2 |
| Temperature Sensor | MCP9808 | 7 |
| Corrosion Measurement | Cosasco ECM | 1 |

The data has been stored in a time series database (TSDB) based on the widely used opensource software, KairosDB[[3]](#footnote-3), which has a Web User Interface to enable access to the time series data in the form of a plot or via download of the data in either CSV or JSON format.

POD 3 contains 4 (ventilation only) free coolers, namely the ECV18000 units from EcoCooling and therefore do not consume water. POD 3 contains a maximum of 180kW of ASIC units with some variation in workload, but far more limited than the general-purpose compute systems based on x86 instruction set CPUs.

Keyword search is based on aspects such as metric, devicetype, source, tag, etc. and these include terms such as humidity, temperature, etc. The keywords are specific to the applications and are not provided with this dataset but have been used to create the large Excel spreadsheet file.

## The data center operations behind the data

The testing phase of the data center took place between project months 17 - 30 which corresponds to the calendar dates February 2019 - March 2020 inclusive. During the test period, the IT equipment and data collection system of POD 1 was operated continuously. The power and thermal environment of POD 3 was monitored over a shorter period with IT equipment from an external donor. POD 3 cooling equipment and control was operated independently of POD 1.

### Workload mixtures

* IoT workloads proposed and constructed by project partner Fraunhofer
  + Predictive maintenance
  + Smart cities
    - Specifying the number of parallel running workloads
* Synthetic workloads packaged by project partner RISE
  + Specifying level of CPU, RAM, I/O and Network workloads

### Workload distribution strategies

Strategies for the IT workload distribution were configured within the workload generation tool. Workloads identified as IOT are application-orientated and can be distributed in the three ways bulleted as follows:

* A - Uniformly Distributed Workload
  + This will spread the workload evenly throughout the data center. For the synthetic workloads this involves running the exact same workload on every server. In the IoT workload case this involves running the docker deployment engine so that the docker containers are spread evenly as possible throughout the data center. Each IoT workload consists of 14 docker containers.
* B - Semi-Uniformly Distributed Workload
  + This will involve spreading a workload thought the data center with groups of servers, racks or sub racks receiving different workloads. This involves making groups that contain servers from every other rack or every other server within a rack. In the synthetic workload case this operates by running 100% on group 1 and 25% on group 2. In the IoT case this involves the running of docker type 1 on group 1 and docker type 2 on group 2. Equally it is possible to involve only running a workload on group 1 and leaving group 2 servers idling. It is possible to create more than two groups, two was only used as an example.
* C - Unevenly Distributed Workload
  + This case will be the most extreme case, where the workload will be applied as a physical section of the data center. This could involve placing the load on half or quarter of the servers which are located together physically.

In each test campaign several workloads were executed. Table 6 below shows the type of the workloads. ***S*** stands for a synthetic, ***IOT*** for an Internet of Things oriented application. The different synthetic workload patterns are given by S1, S2, S3 and S4 and application oriented by IOT1 and IOT2.

Table : Timeframe for the measurement campaigns.

|  |  |
| --- | --- |
| **Project Month** | **Load** |
| **17** (February 2019) | Calibration test campaign |
| **18-24** (March – August 2019) | Campaigns with V1 workloads |
| **25-30** (September 2019 – March 2020) | Campaigns with V2 workloads |

The test phase was divided into smaller time periods identified as campaigns where a campaign is defined as one 5-day (Monday to Friday) or 6-day (Monday to Saturday) test period with the version V1 or V2 workload profiles respectively – see Table 5 above and Table 7 below. These campaigns were run every third week of the data center operation with two different test slots per day (0:00 – 12:00 and 12:00 – 24:00) – see Table 6 below. Synthetic (S) and application-oriented (IOT) workloads are deployed as detailed in public deliverables D3.1 (BodenTypeDC 2018) and D3.2 (BodenTypeDC 2020). The IOT workloads are running for only 11 hours, since the twelfth hour is used to shut down all sub applications so that the next workload can start without any running rogue/background applications.

Table 6: Workloads used during the test campaign periods

|  |  |  |
| --- | --- | --- |
| **Day** | **Slot 1  (0:00 —12:00)** | **Slot 2  (12:00 —24:00)** |
| D1: Monday | **S1.A** | **IoT1.A** |
| D2: Thursday | **S2.A** | **S1.C** |
| D3: Wednesday | **IOT2.B** | **S1.B** |
| D4: Thursday | **S1.C** | **IoT2.C** |
| D5: Friday | **S3.A** | **S4.B** |
| D6: Saturday | **IOT2.B1** | **IOT2.C1** |

Table 7: Timetable over the year for the test campaign

|  |  |  |
| --- | --- | --- |
| **Test-campaign** | **Calendar-week** | **Duration** |
| TC\_01 | 11 | 18.03. — 22.03.2019 |
| TC\_02 | 14 | 01.04. — 05.04.2019 |
| TC\_03 | 17 | 22.04. — 26.04.2019 |
| TC\_04 | 20 | 13.05. — 17.05.2019 |
| TC\_05 | 23 | 03.06. — 07.06.2019 |
| TC\_06 | 26 | 24.06. — 28.06.2019 |
| TC\_07 | 29 | 15.07. — 19.07.2019 |
| TC\_08 | 32 | 05.08. — 09.08.2019 |
| **Test campaign Phase II  (with additional tests)** | | |
| TC\_09 | 35 | 26.08. — 30.08.2019 |
| TC\_10 | 38 | 16.09. — 21.09.2019 |
| TC\_11 | 41 | 07.10. — 12.10.2019 |
| TC\_12 | 44 | 28.10. — 02.11.2019 |
| TC\_13 | 47 | 18.11. — 23.11.2019 |
| TC\_14 | 50 | 09.12. —14.12.2019 |
| TC\_15 | 01 | 30.12. — 04.01.2020 |
| TC\_16 | 04 | 20.01. — 25.01.2020 |
| TC\_17 | 07 | 10.02. — 15.02.2020 |
| TC\_18 | 10 | 02.03. — 07.03.2020 |

The data center operated continuously during the test phase for over one year so that the facility could be exposed to the wide range of climatic conditions and temperatures that are found at the location. For the geographic location of Boden, the monthly average temperature is -13 in January and the corresponding temperature for July is 15 (data n.d.). The weather data indicates a minimum of   
-24.5 and a maximum of 32.6 with an average temperature of 2.0 over the test phase.

## The excel spreadsheet

The long-term storage database was interrogated to create several datasheets. The sheet labelled *Parameters* list the power meters in the switch room, the tap off box meters in POD 1 and shows the layout of POD 1. Note that the power meters were not configured to use NTP and therefore the offset “time difference” was taken on the 27th March 2020 for data that was collected internally by the meters. The data collection system used an NTP timestamp.

The second sheet labelled *Humidity POD1* contains the hour by hour data for the outside ambient temperature, specific humidity and dew point as well as the temperature, relative humidity and dew point of the supply air, that being the air that is supplied to the front of the servers in POD 1.

The data that is presented in the sheet labelled *PUE* is the hourly averaged instantaneous Power Usage Effectiveness of the whole data center and the individuals PODs, namely 1 and 2 that were in use during the ORDP range of dates. It should be noted that the instantaneous PUE does not conform to the IEC/ISO standard on reporting PUE, which should be annualized to take into account seasonal variations in the metric (ISO/IEC 30134-2 2016/Amd 1:2018).

The next three sheets labelled *H*, *D*, and *W* contain the operational data for the PODs on an hourly, daily and weekly average respectively. The hourly average also contains additional information such us sub-metering of the power in POD 1 as well as three differential pressures between the cold and hot aisles at three different locations. The columns identified as pressure sensor 1 and 3 are loaned EBTRON HTN104-1 devices to the project by Ebtron[[4]](#footnote-4). The pressure sensors have been calibrated in the RISE windtunnel. These three main sheets also report server percentage utilization, which is based on CPU threads and not the cores since the servers were operated with hyperthreading[[5]](#footnote-5) enabled. This distorts the relationship between workload and power consumption of a server since each server has 2 CPUs each with 8 cores and therefore 32 threads and therefore a workload that employs more than 16 threads will in fact draw similar power so for workloads greater than 50% utilization will little variation in power consumption.

The remaining sheets are self-explanatory and since the data center has not been fitted with a centralized UPS the LET report will be of interest as it highlights the quality of the power provided to the data center by the local power company throughout the operational period.

Snapshots of the spreadsheet can be seen below for daily, Figure 4, and hourly, Figure 5, data.

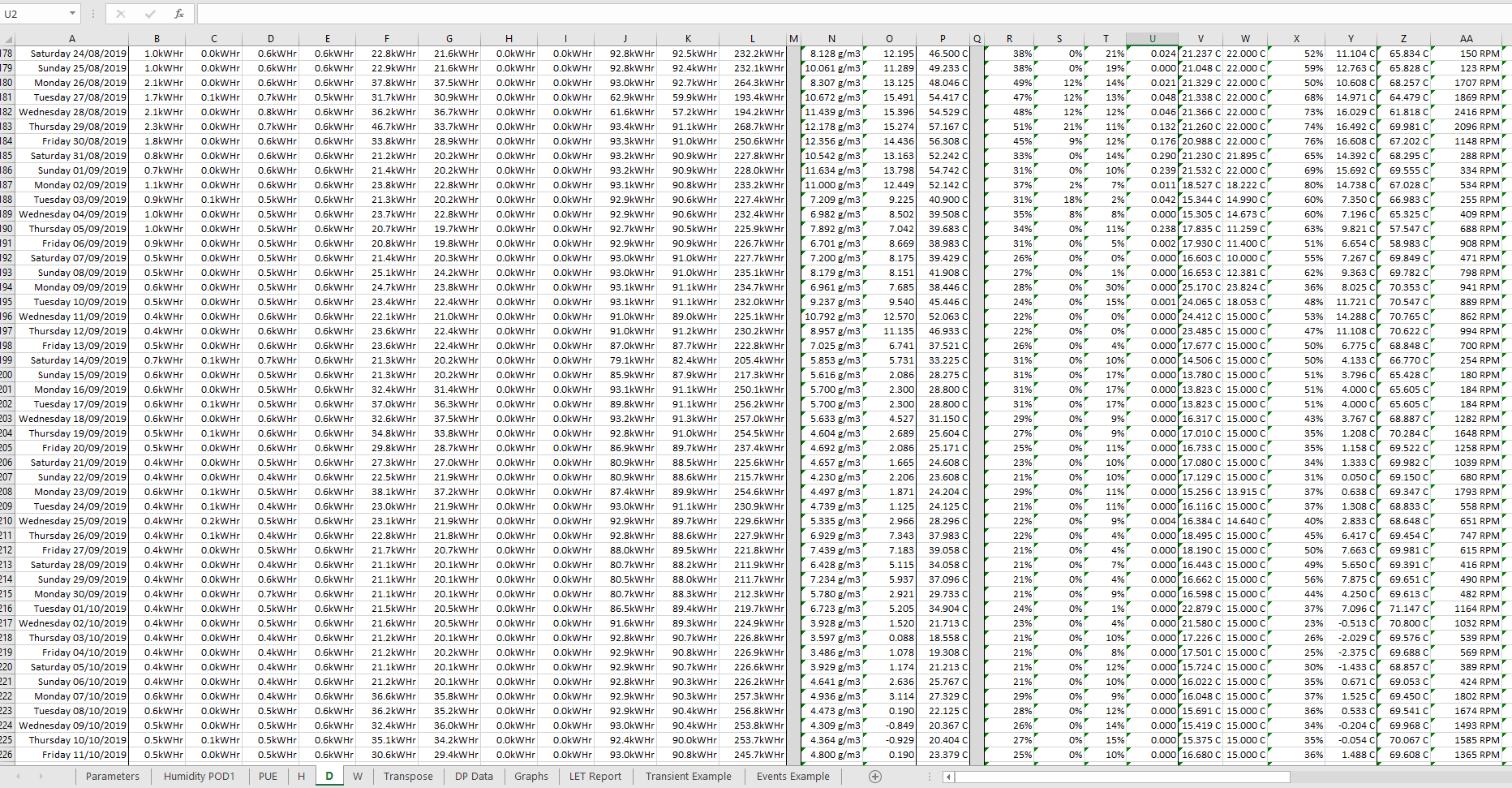


Figure : Snapshot of part of the Excel spreadsheet with hourly averaged data.

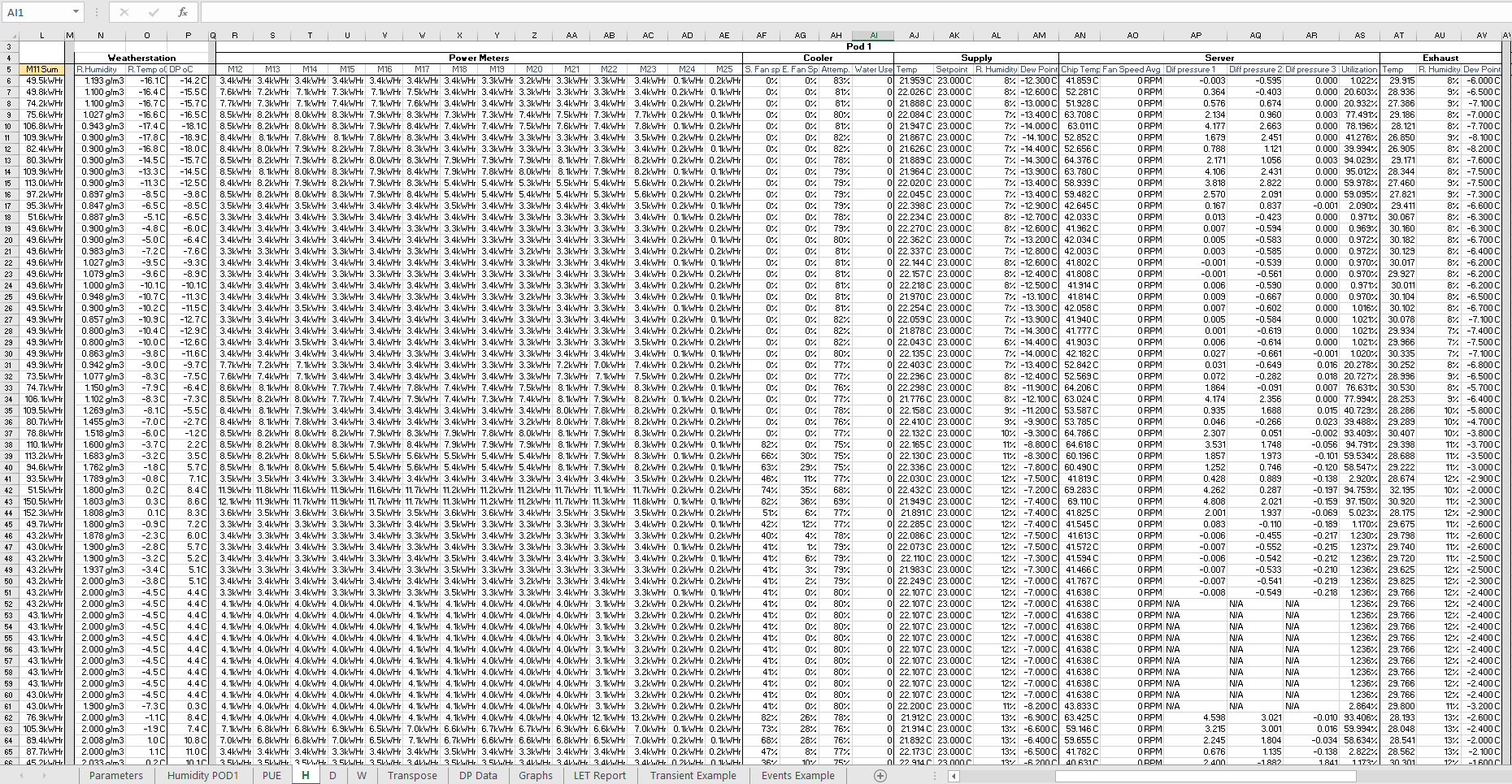


Figure : Snapshot of part of the Excel spreadsheet with daily averaged data.

***Raw data by  
the hour of  
each test  
campaign.***

# Conclusions

The BodenType DC project has successfully created a very rich dataset of one year of operation of comprehensively monitored data center providing the data in a number of excel worksheets for the community to access and make use of in the form of an Open Research Data repository with relevant hour, day and week averaged monitored values for both PODs 1 and 3 of the BTDC One.

# References

BodenTypeDC. 2018. "Deliverable 3.1: Benchmark Modeling Methodology." https://bodentypedc.eu/upload/images/publications/bodentypedc\_d3\_1\_benchmark-modelling-methodology.pdf.

BodenTypeDC. 2020. "Deliverable 3.2: Benchmark Models based on real application scenarios." https://bodentypedc.eu/upload/images/publications/3\_2\_Benchmark\_models\_based\_on\_real\_appliation\_scenarios.pdf.

data, SMHI Öppna. n.d. https://www.smhi.se/klimatdata/meteorologi/temperatur.

ISO/IEC 30134-2. 2016/Amd 1:2018. "Information technology – Data centres – Key Performance Indicators – Part 2: Power Usage Effectiveness (PUE)."

# List of figures

[Figure 1: The data center data collection arrangement. 6](#_Toc52656322)

[Figure 2: Left: 3D schematic of the inside of the data center and the sections with an X have not been operation for data collection. Right: Photo of the data center from the outside. 6](#_Toc52656323)

[Figure 3: Plan view of the internal data center layout showing PODs 1 and 3. 7](#_Toc52656324)

[Figure 4: Snapshot of part of the Excel spreadsheet with hourly averaged data. 13](file:///C:\Users\Jon%20Summers\ownCloud\WP5-H2020\Deliverables\D5.4\BodenTypeDC_D5.4_RISE_2020-09-25-version2-ORDP.docx#_Toc52656325)

[Figure 5: Snapshot of part of the Excel spreadsheet with daily averaged data. 13](#_Toc52656326)

1. Contains two Intel Xeon E5-2770 v1 CPU @ 2.6GHz each with 8 cores and a total of 16GB of RAM [↑](#footnote-ref-1)
2. Contains two Intel Xeon E5-2660 v1 CPU @ 2.2GHz each with 8 cores and a total of 128GB of RAM [↑](#footnote-ref-2)
3. https://kairosdb.github.io/ [↑](#footnote-ref-3)
4. https://ebtron.com/ [↑](#footnote-ref-4)
5. https://www.intel.com/content/www/us/en/gaming/resources/hyper-threading.html [↑](#footnote-ref-5)