DataCenter 2020: first results for energy-optimization at existing data centers

July 2010

Powered by

T-Systems

intel

DataCenter 2020: first results for energy-optimization at existing data centers

Introduction

Environmentally friendly and sustainable IT is a hot topic given the current climate change and ecological challenges. According to Gartner's market researchers, IT equipment is responsible for two percent of all CO_2 emissions worldwide. This is almost equal to the amount of carbon dioxide generated by air traffic. With rising energy costs added to the mix, more and more companies are paying greater heed to their consumption of IT resources and ensuring that the design of their IT infrastructure is as environmentally friendly ("green") as possible. The objective is to use energy resources more efficiently while also reducing CO_2 emissions. Much focus lies on achieving maximum energy efficiency – that is, to optimize the amount of energy used per unit of output.

Energy consumption at data centers can be reduced using methods that are easy to effect.

Those responsible for IT in companies and data centers face major challenges. On the one hand, the number of IT applications and the data volumes to be stored are increasing rapidly. This requires greater processing power, and therefore more server hardware. In addition to operational electricity costs, there is also the added expense of cooling systems. According to analysts at IDC, for every dollar that the CIO spends on a new server, another 50 cents is spent on electricity and cooling.

The most important tool for combating power consumption and improving the climate footprint of the data centers is higher energy efficiency. Needless to say, IT's commitment to the environment is not solely limited to lower power consumption. However, energy efficiency plays an important role and it is a fundamental prerequisite for an environmentally friendly data center. This poses the question: how can the existing infrastructure (cooling, power, space etc.) use the latest technology with maximum efficiency – that is, how can the best possible ratio of power consumption to required performance be achieved?

This is the subject of the technology partnership between Intel and T-Systems in the context of the DataCenter 2020 project. The two companies are working together on solutions for the industrialization and automation of ICT services with the aim of making these marketready with maximum efficiency and cost effectiveness. The nucleus of this collaboration for energy efficiency is a test laboratory at the Euroindustriepark in Munich. The aim of the Intel and T-Systems laboratory is to calculate measured values that can be used as a starting point for optimizing existing data centers and making them more energy efficient. These findings will also be used to create a model for the data center of the future.

Partnership between Intel and T-Systems

The technology partnership between Intel and T-Systems centers on essential questions regarding the use of technology today and in the near future. Together, the two companies want to develop future-proof ICT solutions that use energy efficiently, save costs and pave the way for new ICT services. To increase the pace of innovation throughout the IT industry, Intel and TSystems are making the results of their collaboration and the recommended courses of action publicly available.

The partnership focuses on DataCenter 2020 for the optimization of existing data centers and for joint research on the data center of the future. The collaboration between Intel and T-Systems facilitates new knowledge and methods as part of a holistic approach (End-to-End). Both partners place enormous value on the practicability and actual application of these, since existing data centers already harbor huge energy-saving potential. Procedures such as the arrangement of racks in the warm aisle/cold aisle or housing concepts are already standard nowadays, but there is still clearly room for improvement. At DataCenter 2020, Intel and T-Systems are developing a realistic roadmap that details how the energy efficiency of data centers can be greatly increased.

This white paper describes the experimental environment of DataCenter 2020 and presents the first research results from the test lab for energy-optimization at existing data centers. Essentially, to date the experts from Intel and T-Systems have focused on two measures: through the strict separation of cold and warm air, they were able to reduce the fan speed of the forced air cooling devices, and with the increase in room temperature, they were able to prolong the time dedicated to free cooling, which occurs in conjunction with the outside temperature.

DataCenter 2020: the test lab at Munich's Euroindustriepark

The test lab is equipped with around 180 rack servers, as well as the latest in energy, air-conditioning, measuring and control technology. Approximately 1800 data points record parameters such as humidity, room temperature, temperature difference between the incoming and outgoing air, processor load, and fan speeds. The most important instrument is the electricity meter. In addition, the raised floor houses a smoke generator that monitors airflow. The smoke that is generated makes the direction and speed of the airflow visible. At the same time, the engineers can identify short-circuits in the power supply, i.e. detect permeable areas (leakage air) and places into which air is absolutely not allowed to flow.

At the data center, airflow plays a vital part in climate control. The engineers carry out various tests, for instance on the cold aisle containment. Here, the forced air cooling device channels cold air through the raised floor to cool the servers, and the cold air is kept strictly separate from the warm air. In accordance with the set room temperature, humidity and CPU load, the testers can then optimize parameters such as cold water supply or ventilator speed.

In addition to the cold aisle containment, the partners examine systems with warm aisle containment (where warm air is sucked out and then cooled). They test the efficiency of liquid-cooled racks and the properties of various forced air cooling devices that use different technology. To simulate data centers with different ceiling heights, the test lab employs a lift slab with a variable height of between 2.5 and 3.7 meters. The tests conducted up to now were based on a ceiling height of 2.7 meters. The primary objective is to determine which solution offers the data center the best in energy efficiency and under which conditions. To achieve this, the engineers examine the entire process chain, from energy supply through to consumption. Besides improved air-conditioning technology, much focus is placed on the use of energy-saving IT, for instance in server and processor technology.

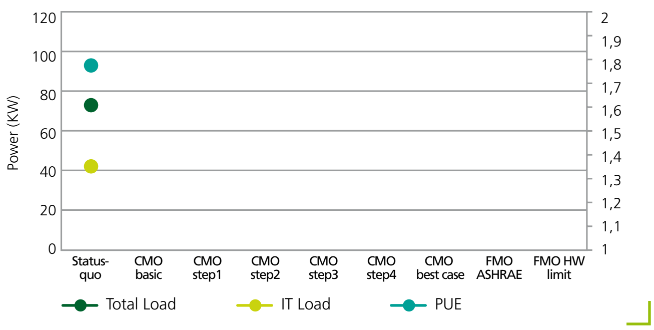
PUE value as a measurement for energy efficiency

To measure energy efficiency at DataCenter 2020, T-Systems and Intel use the Power Usage Effectiveness (PUE) industry standard defined by the Green Grid organization – that is, the efficiency of energy consumption. This value measures how much consumed energy is actually converted into processing power. PUE is the ratio of total facility power consumption to IT equipment power consumption. At present, the average PUE value at existing data centers is 1.9.

IT equipment power consumption is the amount of power used by all IT devices in the data center, including servers and other computers, storage devices and network systems, switches, monitors, and other peripheral and telecommunications devices. Total facility power includes, in addition to the power consumption for IT, the electricity costs for the infrastructure that supports the IT operations. This comprises systems such as UPS, switching systems, batteries, cooling systems, pumps and lighting. Total facility power is quite easy to read on the electricity meter.

Starting point: today's standard data center

In conventionally designed data centers, the server cooling system eats up around half of all the energy needed. This high level of consumption is also based on different inefficiencies. At DataCenter 2020, the researchers from Intel and T-Systems produced their first measurements by simulating the conditions at current up-and-running data centers with all their faults. The PUE value resulting from this environment was approximately 1.8. The following conditions were defined for the measurements.



Starting conditions: 174 Servers @ 100% load = 42 kW
8 Racks → 5kW/Rack: Chiller mode yearly average at Munich DC site

1. Leakage in the raised floor, racks and cable feedthroughs causes thermal short circuits.

Background: at the data center, airflow plays a vital part in climate control. It is essential to clearly separate cold air from warm air and to avoid unnecessary airstreams within the data center. Take the example of a raised floor: in many data centers, a raised floor is used to separate the cold air for cooling under high pressure. It is also generally used for routing power cables. The raised floor must be expertly sealed to avoid so-called leakage air, which causes cold and warm air to mix, leading to thermal short circuits. Leakage must be prevented in the racks and cable feedthroughs too.

2. The speed of the forced air cooling devices is set to the maximum (100 percent) to ensure that, despite leakages in the raised floor, enough air reaches the servers in the cold aisle.

3. The IT load is limited to approximately 5 kW/rack (8 racks with a total of 174 servers are in use) or 2 kW/m².

4. The inlet temperature (air) in the raised floor is set to 18°C. This results in a server intake temperature of around 22°C.

Background: the forced air cooling device draws warm air in, cools it down and blows it back into the raised floor at the appropriate temperature. Cooling is achieved through an internal heat exchanger that is supplied with water from chillers located outside of the building. Inside the chiller, a compression refrigerator raises the temperature of the coolant to the extent that it can be chilled with ambient air. This consumes a massive amount of energy. At lower ambient temperatures, the coolant can instead be chilled directly using external air without the need for a chiller. This process is referred to as indirect free cooling and it is far more efficient than cooling using a compression chiller.

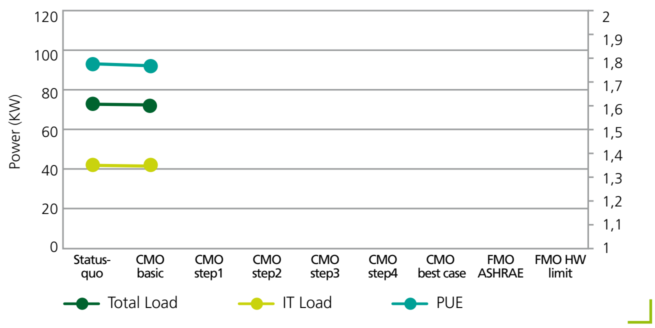
5. Since PUE represents an annual mean value, the experts at DataCenter 2020 developed a mathematical yearly model for the water chiller so as to eliminate non-representative snapshot readings. They also incorporated the average temperatures for the Munich area into the model. As a result, the temperature gradation can be used to forecast the potential for using indirect free cooling in the annual mean.

Under these conditions – including inefficiencies – the measurements gave a PUE value of approximately 1.8.

Optimization phase I: separating cold and warm air

Having simulated the conditions in a conventionally designed data center, the experts at DataCenter 2020 began the first optimization phase. This phase ended with a PUE value of 1.48. It focused on the forced separation of cold and warm air in three steps:

1. Sealing of leakages

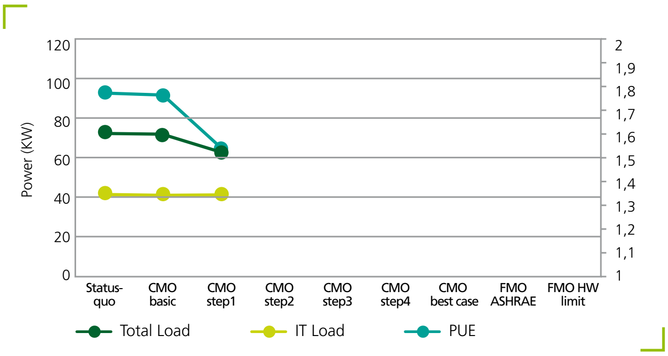


Blind blanks
Floor leakage reduced to 20%

The first step taken by the experts was to make the compartments airtight, thereby preventing air from escaping unnecessarily. They eliminated leakage air by sealing the raised floor (for example, where power cables were fed through) and by inserting dummy panels in the racks (between the rack units that house the servers).

Sealing the leakages had no impact on the PUE, however, since all devices were still running in the same way. Indeed, the fact that air could no longer unnecessarily escape led to a considerable rise in air pressure within the raised floor. This forms the precondition for the next step.

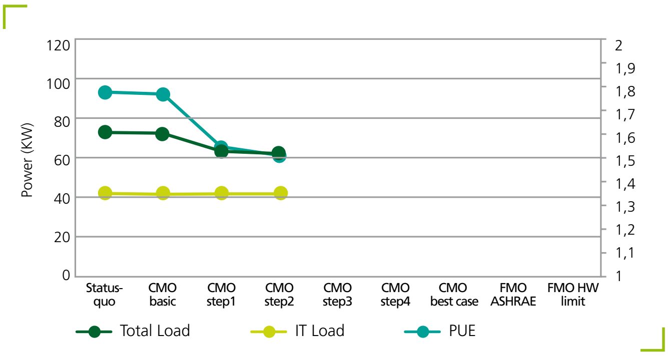
2. Adjustment of fan speed in the forced air cooling device



CRAC fan speed optimization
(to be repeated in any of the following steps)

The experts at DataCenter 2020 were now able to reduce pressure in the raised floor to the minimum required, thereby ensuring that the intake temperature remained sufficient throughout the rack height. They achieved this by lowering the fan speed in the forced air cooling device. Since the fan was now rotating much more slowly and therefore using less energy, the PUE value decreased from 1.8 to 1.55.

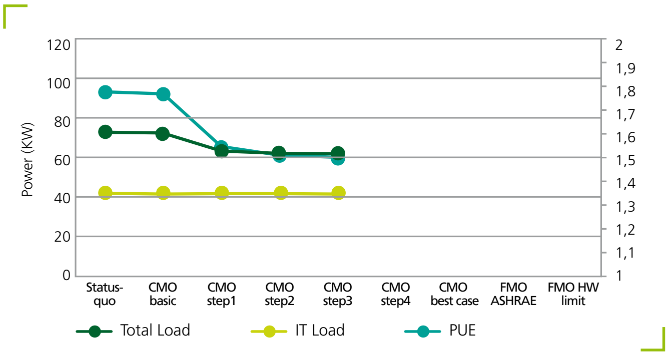
Given that energy saving effects could only be achieved in this phase (separating the airflows) by adjusting the fan speed, this optimization step is repeated in all subsequent steps, even without specific reference.



Doors installed at the end of the aisle
(without ceiling)

3. Gradual improvement of the cold and warm air isolation

The strict separation of cold and warm air was applied in this step to future-proof against thermal short circuits. Doors were fitted at the ends of the aisles, preventing cold and warm air from mixing through the rack sides at the beginning and end of an aisle.



Replacement of perforated tiles (38% opening) with
grating tiles (98% opening) in cold aisle

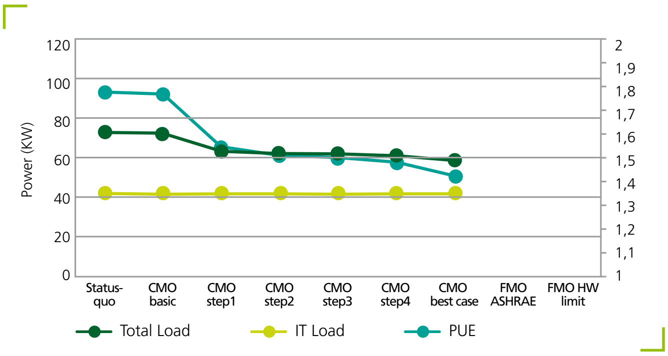
In addition, replacing the usual perforated tiles in the cold aisle with grating tiles achieved a reduction in aerodynamic resistance, since air flows through the larger openings with more ease (Slide 8). This allowed for a greater decrease in the fan speed of the forced air cooling device. Lastly, the cold air ceiling was closed to prevent the cold and warm air from mixing across the upper side of the rack. Following the third step of the first optimization phase, the PUE value sank once more, from 1.55 to 1.48.

Optimization phase II: Increasing the inlet temperature

The second optimization phase focused on increasing the inlet temperature. In principle, the ideal temperature range for humans is around 22°C, but the same does not apply for servers. However, many data centers fail to use energy efficiently because they are actually too cold. Today, many servers can tolerate ambient temperatures of 30 to 35 degrees Celsius, while the air intake temperature in most data centers is between 20 and 25 degrees Celsius. If the intake air does not need to be cooled as much, the air-conditioning system will use less energy.

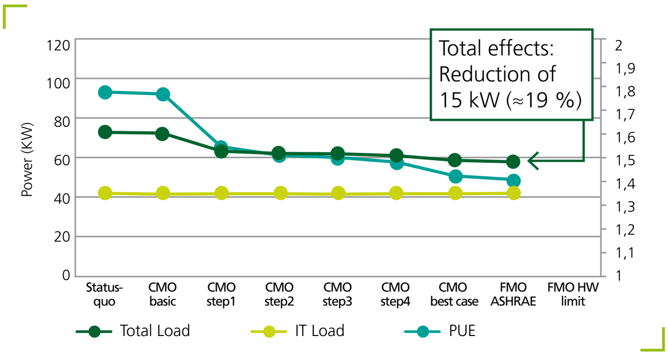
Therefore: if the inlet temperature is increased by raising the coolant temperature (that is, the water supply temperature), there will be less need for forced cooling, generated for example using compressors. Forced air cooling devices feature a heat exchanger that is supplied with water from chillers located outside of the building. Inside the chiller, a compression refrigerator raises the temperature of the coolant to the extent that it can be chilled with ambient air. At lower ambient temperatures, the coolant can be chilled directly using external air without the need for a chiller. This process is referred to as indirect free cooling. The higher the temperature of the cooling agent is and the lower the outside temperature, the less energy is required. The efficiency of the chiller also increases with every degree in room temperature. The experts at DataCenter 2020 must therefore clarify the following question: what is the ideal water supply temperature?

4. Reaching the limits of the existing cooling infrastructure
With standard chillers, the water supply temperature is limited to 14°C. However, an increase in this limit would see a drop in the need for forced cooling and indirect free cooling would feature in the annual mean for longer. In a best case test, the experts at DataCenter 2020 introduced an air cooling device with an EC motor featuring a direct-current ventilator and raised the flow temperature to 14°C. EC motors require around 30 percent less power than conventional AC motors. This setup allowed the experts at DataCenter 2020 to reduce the PUE value even further, this time from 1.48 to 1.43.



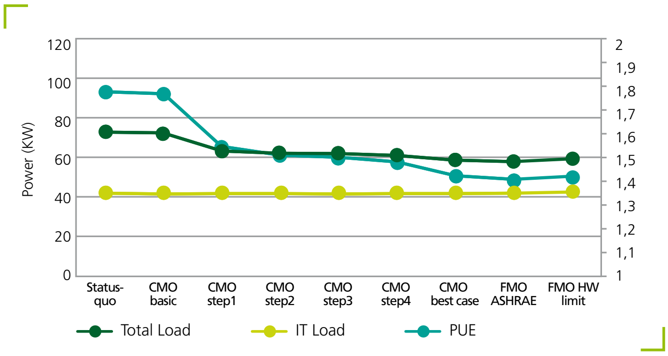
Raise water supply temperature from 8° C to 14° C
Fix coolant flow (valve to 75% manual)

5. Increasing the inlet temperature in line with current ASHRAE recommendations
The next step taken by the DataCenter 2020 team was to raise the water supply temperature to 24°C as per the recommendations of ASHRAE (American Society of Heating, Refrigerating and Air Conditioning Engineers), corresponding to a server intake temperature of 27°C. The PUE value thus dropped to an optimal 1.4.



Water supply temp increased from 14° C to 24° C

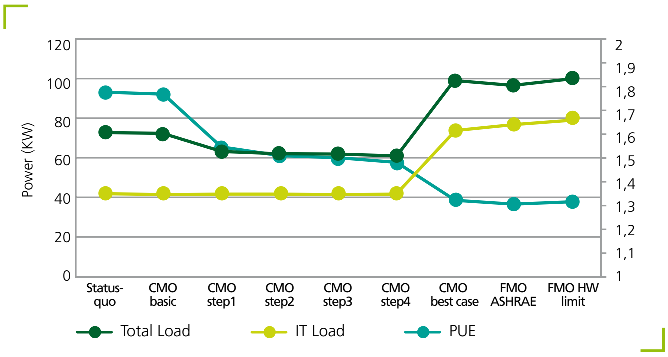
ASHRAE (www.ashrae.org) is the American Society of Heating, Refrigerating and Air-Conditioning Engineers. The four-volume ASHRAE manual is a reference guide to air conditioning technology. A new volume is published every year. ASHRAE also publishes standards and guidelines in the area of air conditioning technology, which are referred to in building regulations.



Maximum raise of water supply temp to 34°C

6. Reaching the server's thermal limits
Since the specifications of most servers today allow intake temperatures of up to 35°C, the researchers' next step was to reach this limit and raise the inlet temperature to 35°C. However, this made the server fans run faster, using more energy, and thus in this experiment the PUE value increased again slightly to 1.43.

7. Increasing the IT load to 10 kW/rack or 4 kW/m²
In the last test, the experts at DataCenter 2020 doubled the IT load from around 5 kW/rack or 2 kW/m2 (= initial load) to around 10 kW/rack or 4 kW/m². Since this doubles the IT equipment power from 40 kW to 80 kW, the total facility power increases too. This experimental procedure achieved another improvement in efficiency and a PUE value of 1.3.



Doubling

Summary

Energy consumption at data centers can be reduced using methods that are easy to effect. To lower the PUE value, the researchers from Intel and T-Systems experimented with different scenarios at DataCenter 2020. The improved energy efficiency at existing data centers is due mainly to the following two measures:

1. The strict separation of cold and warm air and the resulting optimization of airflow, which allows for a reduction in the fan speed of the forced air cooling device. This result forms the basis for all further steps and must be implemented with comparatively cost-efficient measures.

2. Increasing the room temperature or inlet temperature. This measure shortens the time for forced cooling and lengthens the time for indirect free cooling. The experts achieved the best result at a server intake temperature of 27°C in accordance with ASHRAE's recommendations. This necessitates a detailed inspection of the existing infrastructure and buildings, so that they can be used as effectively as possible in line with design options. Location, power supply and customer focus are important criteria too for completing a holistic assessment of existing data centers.

By implementing straightforward changes, energy consumption at an existing data center can quickly be reduced to moderate costs. To achieve this, the individual measures must be expertly coordinated and stringently implemented. The integrated approach described here reduces operating costs where the load remains constant, or frees up space to allow for an increase in IT capacity where total energy consumption remains constant. To achieve maximum efficiency, the experts at DataCenter 2020 recommend always using the available cooling capacities and power supply performance at their maximum levels.

Outlook

T-Systems and Intel will implement the initial findings from research conducted at DataCenter 2020 into their own projects for the optimization of data centers. In the next step of the project, the experts will examine to what extent the data center infrastructure can be controlled by the IT load so that energy consumption is also optimized in the partial-load operational range. The long-term objective is "infrastructure on demand", whereby the infrastructure in the data center supplies the servers with the exact amount of cooling output that they need at this moment in time.

For further information www.datacenter2020.com

Powered by

T-Systems · intel