September 2020

Thermal and Load Profiling of a Data Center

Anonymous

Thermal and Load Profiling of a Data Center

ABSTRACT

Data center (DC) operators have an interest in maintaining an optimum temperature uniformly throughout their data halls and server racks. Too high a temperature can lead to equipment slowdowns or malfunctions; too low a temperature causes an unacceptable increase in energy consumption. Although ceiling temperature sensors enable the maintenance of an optimum average room temperature, it is often the case that temperatures closer to the computing module, e.g., CPU, GPU, QSFP optical transceivers, etc., are unacceptably high.

This disclosure describes techniques that leverage the existing high-density, distributed networks of thermal sensors in network equipment to determine local thermal conditions at various points in a data center. The techniques enable a fine-grained, wide-area, real-time visualization of thermal variations within a data center; optimization of data center airflow; an optimal placement (with reference to air supply) and interspace of devices, modules, servers, racks, and other equipment; the relocation temperature sensors if appropriate; etc.

KEYWORDS

- Data center
- Temperature monitoring
- Temperature distribution
- Thermal efficiency
- Thermal profile
- Load profile
- Air-handling unit
- Airflow

BACKGROUND

Data center (DC) operators have an interest in maintaining an optimum temperature at various points throughout their data halls and server racks. Too high a temperature can lead to equipment slowdowns or even shutdowns and malfunction; too low a temperature causes an unacceptable increase in energy consumption. To monitor temperature, data centers use temperature sensors, typically located on the ceilings. Readings from the sensors are used to monitor the temperature and adjust the air-speed at the outlet of the air-handling units.

Although ceiling temperature sensors enable maintenance of an optimum average room temperature, it is often the case that temperatures closer to a computing module, e.g., CPU, GPU, QSFP optical transceivers, etc., are unacceptably high. Despite an apparently optimum room temperature, a variety of thermal issues can arise localized at specific points, such as:

- Imbalanced air-pressure distribution between the inlet to a server rack (cold aisle) and its outlet (hot aisle).
- Recirculation of air from the outlet of a server rack (hot aisle) to its inlet (cold aisle).
- Non-uniform airflow distribution between racks and systems in the data center, in turn leading to non-uniform cooling (temperature distribution).
- Power loss due to the non-uniform airflow distribution.
- Insufficient airflow.

Various combinations of these effects can lead to hotspots (also known as dark locations), e.g., locations within the data center where the ambient temperature is chronically high; and wasted air and energy, since high ambient temperature, even if only local, can lead to faster spinning of the fans of the air-handling units. The resulting data center power profile can have serious consequences for data center capacity and for network resource allocations.

DESCRIPTION

This disclosure describes techniques that leverage the existing high-density, distributed networks of thermal sensors in network equipment to determine local thermal conditions at various points in a data center. Detected thermal conditions include the presence of recirculation, pressurization, hotspots, etc., and the distributions of traffic and consumed power. The production of local thermal data at a fine granularity and high sampling rate (e.g., on a minute-by-minute basis) enables mapping temperature variations, airflow conditions, recirculation, etc.

This data can enable data center operators to determine the locations of hotspots (if any) and critical in-unit operating conditions or airflow characteristics impeding operations, and to generally optimize airflow. Modeling thermal behavior at a unit or switch level and extending the model from rack to row to data hall enables creation of a detailed yet expansive view of temperature, airflow, hotspot conditions, etc. in the data center. The resulting fine-grained and wide-area map of thermal conditions informs the design of data centers and the placement of devices within data centers.

To determine the aforementioned map of thermal conditions, data is obtained from various sensors, e.g., the temperatures of the QSFP or other modules; the ASIC power, along with its components, core power and serdes power; the ambient temperature; the data hall air-supply temperature; speeds of fans (in revolutions per minute) within modules; motherboard temperatures; etc. The data is collected at a suitably fine sampling rate, e.g., on a minute-by-minute basis.

Detecting recirculation and determining its magnitude

As explained earlier, recirculation is said to take place when air from the outlet of a server rack (hot aisle) flows back to its inlet (cold aisle). Recirculation can increase server inlet temperatures, make cooling ineffective or inefficient, and lead to equipment slowdowns or shutdowns. To address this problem, a temperature sensor in the data hall reports global ambient temperature (T_{COLD}), also referred to as cold-aisle temperature, data hall temperature, supply temperature, or reference temperature. Each device typically has two sensors at its inlet, which report local temperature, e.g., the temperature close to the inlet of the server module. These two sensors typically read nearly the same temperature (to within 1°C), and their temperature is collectively referred to as server-inlet temperature (T_S). Each device also has a temperature sensor at its exit (near its fan), which reports a temperature referred to as motherboard exit temperature (T_E), also known as server-exit temperature or server-outlet temperature.

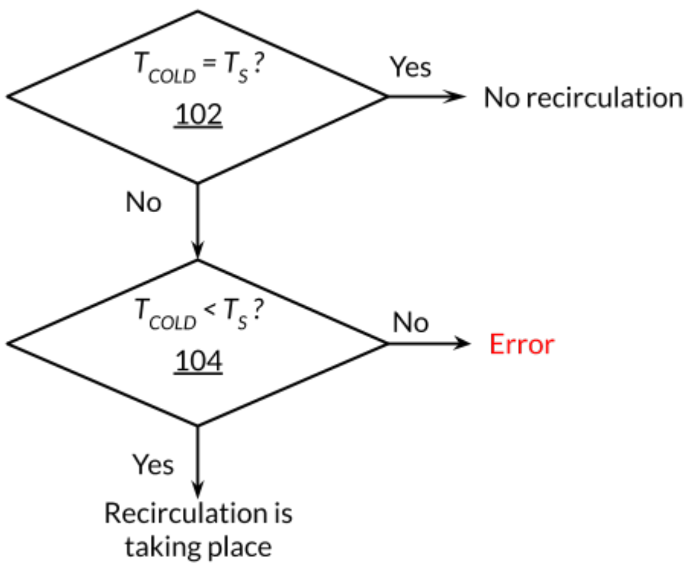


Fig. 1: Determining recirculation

Fig. 1 illustrates an example process to determine whether recirculation is taking place. If the cold-aisle temperature equals the server-inlet temperature (102), then no recirculation is taking place. If the cold-aisle temperature is less than the server-inlet temperature (104), then recirculation is determined to be taking place. The condition of server-inlet temperature being lower than the cold-aisle temperature is considered to be unphysical and possibly indicative of erroneous temperature sensing.

If recirculation does exist, its magnitude can be calculated by performing an energy-balance calculation as follows. Let V_{HOT} denote the volumetric flow rate in cubic feet per minute (cfm) of the recirculation. Let V_{COLD} denote the volumetric flow rate of cold air entering from the cold aisle into the server inlet. The total volumetric flow V_{TOTAL} coming from the server outlet can be correlated via experiment to the fan speed at the server outlet, and is hence a known quantity. The following equations apply:

$$V_{HOT} = \frac{T_S - T_{COLD}}{T_E - T_S} V_{COLD} \quad (1)$$

$$V_{HOT} + V_{COLD} = V_{TOTAL} \quad (2)$$

Equations (1) and (2) have two unknowns, V_{HOT} and V_{COLD}, in two equations, and hence, can be solved for the unknowns. In this manner, the magnitude V_{HOT} of the recirculation can be calculated.

Once recirculation is detected and its magnitude estimated, it can be rectified, e.g., brought down to zero, by checking for and clearing blockages, and by checking for and arresting outlet-to-inlet leakage pathways.

Detecting pressurization and determining its magnitude

Pressurization is said to be present when airflow takes place in a data center by virtue of its high pressure. The server-exit temperature of a module when measured outside a data center, e.g., in an unpressurized situation, is known as standalone temperature (T_{STAND}). From experiments, the unpressurized, server-exit temperature is known for each fan-speed at different ambient temperatures. The server-exit temperature of a module measured inside a data center, e.g., in a pressurized situation, is known as temperature under pressurization, denoted T_{DC}.

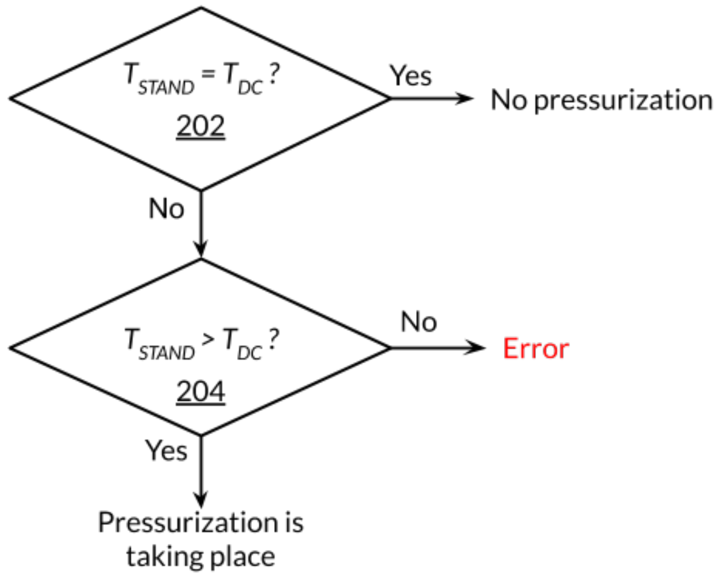


Fig. 2: Determining pressurization

Fig. 2 illustrates determining pressurization, per the techniques of this disclosure. If the standalone temperature of a module equals the temperature of the module when inside the data center (202), then pressurization (and the extra airflow induced by it) is determined to be absent. If the standalone temperature is greater than the temperature of the module when inside the data center (204), pressurization (and the extra airflow induced by it) is determined to be present. The condition of standalone temperature being lower than the temperature of the module when inside the data center is considered to be unphysical and possibly indicative of erroneous temperature sensing.

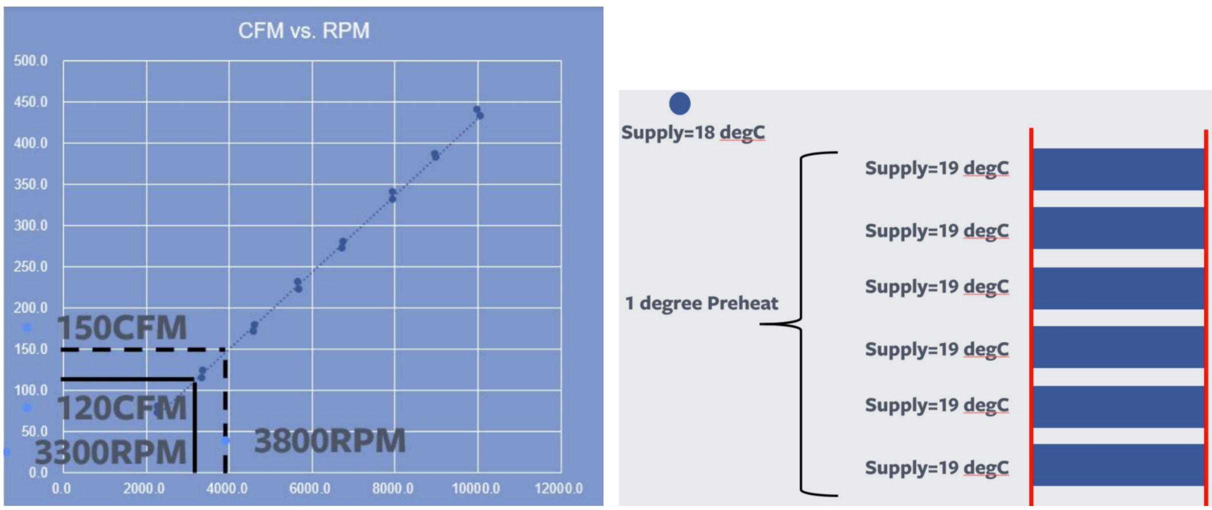
If pressurization is present, its magnitude can be calculated as follows. The thermal resistance is calculated using the formula:

$$\text{thermal resistance} = \frac{T_{STAND} - T_{DC}}{P},$$

where P is the power of the module. The thermal resistance can be keyed into a lookup table to find the extra airflow (in cubic feet per minute) through the fan.

The recirculation and pressurization calculations are repeated for every module in the server rack and data center to get an airflow (e.g., air velocity) and temperature profile for every unit in the data center, and construct a map of thermal conditions throughout the data center.

<u>*Example calculation*</u>



(a) **(b)**

Fig. 3: (a) Volumetric flow in cfm versus rotational speed in rpm for a fan (b) The cold-aisle (supply) temperature and temperatures at server inlets of a server rack.

Fig. 3(a) illustrates the volumetric flow rate in cubic feet per minute of a fan in a certain module versus its rotational speed in revolutions per minute (rpm) in an unpressurized (standalone) situation. At 3,300 rpm, the volumetric flow rate in an unpressurized situation is 120 cfm. At 3,800 rpm, the volumetric flow rate in an unpressurized situation is 150 cfm. The standalone temperature (T_{STAND}) for this module at 3,300 rpm is 50° C, whereas its temperature in a pressurized data center (T_{DC}) at the same rotational speed is 46° C, and its volumetric flow at this temperature is 150 cfm.

Fig. 3(b) shows the cold-aisle (supply) temperature (T_{COLD}) to be 18° C, while the temperature at the inlet of the module (T_S) is 19° C. The exit temperature is 30° C.

Questions:

- Determine if recirculation is taking place, and if so, its magnitude in cubic feet per minute.
- Determine if pressurization is present, and if so, its magnitude in cubic feet per minute.

Answer:

Since $T_{COLD} = 18 < 19 = T_S$, recirculation is taking place (see Fig. 1). Since $T_{STAND} = 50 > 46 = T_{DC}$, pressurization is present (see Fig. 2). From Equations 1 and 2, the magnitude of the recirculation is given by:

$$V_{HOT} = V_{TOTAL}\frac{T_S - T_{COLD}}{T_E - T_{COLD}} = 150\frac{19-18}{30-18} = 12.5 \text{ cfm.}$$

The extra airflow due to the pressurization in the data center is the difference between the airflows at 3,300 rpm within and outside the data center. Hence, the extra airflow equals 150 − 120 = 30 cfm. This extra airflow is attributable to recirculation, whose magnitude is calculated above as 12.5 cfm, and pressurization, whose magnitude is therefore 30 − 12.5 = 18.5 cfm.

In this manner, the techniques of this disclosure apply data analytics to thermal models to generate a visualization of thermal variations in a data center; to optimize data center airflow; to inform the placement (with reference to air supply) and interspace of devices, modules, servers, racks; to relocate temperature sensors if appropriate; etc. Thermal phenomena such as recirculation, pressurization, hotspots, etc. are mapped in a fine-grained, wide-area, and real-time manner. Airflows within the data center are determined on a per-unit basis, and the power of individual modules is profiled. The distribution of power loads across racks, servers, and services between servers within a rack can be determined and optimized. Substantial power savings, e.g., 5-20%, can be achieved by thermal and computational optimization.

CONCLUSION

This disclosure describes techniques that leverage the existing high-density, distributed networks of thermal sensors in network equipment to determine local thermal conditions at various points in a data center. The techniques enable a fine-grained, wide-area, real-time visualization of thermal variations within a data center; optimization of data center airflow; an optimal placement (with reference to air supply) and interspace of devices, modules, servers, racks, and other equipment; the relocation temperature sensors if appropriate; etc.