Thermal management in data centers

For data centers, maintaining a continuous flow of information requires near 100% uptime. As 24/7/365 operations become the norm for business, ensuring that data center infrastructure stays up and running is mission critical. Hard failures can cost a large business millions of dollars in lost productivity and opportunity costs. If this were not a big enough challenge, data centers consume a lot of energy, and in times of rising worldwide demand, high operating expenses can put a serious dent in the information technology (IT) organization's budget. Because of the need for high uptime rates, the ever-increasing cost of electricity, higher server densities and limits on electrical grid capacity, data center operators are now looking for ways to optimize performance and increase their kW per square foot rating—all while reducing costs. Reducing the consumption of energy in a data center while maintaining high availability is no small task, but the rewards are high.

To ensure reliability and economical operating costs, the power distribution and cooling infrastructure must be actively managed. Outages stemming from electrical or mechanical failure can be prevented by physical redundancy practices and predictive/preventive maintenance (P/PM) and are currently being used by most data centers. To reduce electrical consumption and address green operating mandates, data center operators are now exploring ways to raise temperature set points.

With the increase in electronic traffic, heat generated by electronic equipment, and the concomitant costs of powering cooling systems in electronic data centers are increasing continually. Various research groups working in academic institutions, research laboratories and industries have been applying a variety of tools to study and improve performance of data centers.

Periodic inspections of data centers are very important to ensuring the reliability, continuity and sustainability of the systems they house. In fact, such inspections are often mandated by user-founded organizations such as the Uptime Institute and/or by insurance carriers, who do not want to pay damages for lost data due to failed equipment.

One important tool for performing data center inspections is the thermal imager, also known as an infrared (IR) camera. The following step-by-step account describes how to use a thermal imager to inspect data center systems from the electrical source – a transformer or substation – to the server racks and everything in between, including the critical heating, ventilation and air-conditioning (HVAC) system.

Figure 1. – Thermal mapping of data center

A thermal imager displays and can store two-dimensional images of an object's surface temperatures. Using an imager, can be easily detected anomalies in the temperatures of electrical or mechanical components – items that are hotter or colder than similar objects in the same environment. Overheating components usually indicate a potential problem that requires maintenance before failure occurs. In data centers, where cooling is important to keep servers from overheating, uncharacteristically cool surfaces might also indicate a problem, perhaps an imbalance in the HVAC system that requires correcting.

In addition to easily detecting comparative temperatures of equipment surfaces, thermal cameras can also record actual surface temperatures. This helps detect situations such as an overheating transformer or motor, allowing for repair or replacement before failure.

When thermal images reveal potential problems, capture them on the imager and upload them to a computer that runs software for reporting and analysis. By regularly monitoring equipment and keeping a thermal "track record" on a computer for long term comparison, one can better detect abnormal readings and changes in the trend. To ensure the consistency required for side-by-side comparison, follow a pre-established sampling route and scan the same objects or areas each time from the same vantage points. Along with repair records, thermal trending information provides a documented data trail for insurance carriers, management, and any others who require confirmation of a reliable operation.

In a data center, the components are like a series of dominoes. If one fails, it takes everything downstream with it. It makes sense to "begin at the beginning". For a meaningful inspection session, the system must be operating and should be pulling as large an electrical load as possible. More current running through the wires produces more heat energy, and that's what an infrared camera "sees."

- *Transformers* are usually owned by the electric utility, although sometimes they are the property of the data center's owner. On transformers, check the secondary windings and coils. Look at terminations and lugs (bolted connections) "inside of the box." Look for thermal anomalies, i.e., differences in temperature – ΔTs – of similar components. Also, look for physical damage and debris that might interfere with the operation of the transformer, and scan for load imbalance. The latter is signaled by a ΔT between circuit phases.

Figure 2. – Electrical transformer thermal imaging

- Many data centers have an alternate source of power for redundancy. This second source could be another utility transformer on a different grid or a standby generator. Alternate power sources must be scanned and inspected, too, and while they are in use and under load.
- *Standby generators* should be inspected while they are powered up with everything downstream running off them. Here, too, check lugs and terminations and look for damage and debris. To detect problems with cooling or exhaust systems, the operator will need to record actual temperatures rather than observing ΔTs.

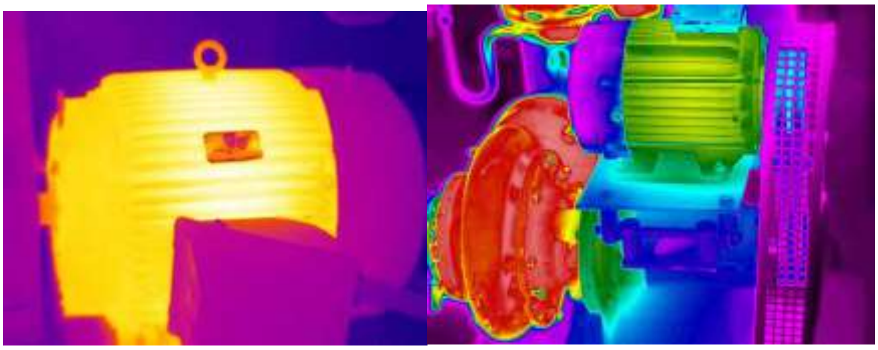

Figure 3 - Thermographic representation for an electric generator

- *When a transfer switch* is functioning correctly, it senses where the power is coming from (main or standby) and switches to that source. Don't overlook that switch during inspection, because if it fails, it won't matter how good maintenance procedures are downstream. With current running through the transfer switch, scan it and look for heating that might signal loose connections (e.g., insufficient torque or compression on a lug or termination).

- The *main switchboard* is a large enclosure with many switches. The cabinet houses various components including busbars, bolted connections and fuse clips. Look for thermal anomalies in connections (including bus connections), terminations, fuses and fuse clips. Also look for imbalance, damage, and debris.

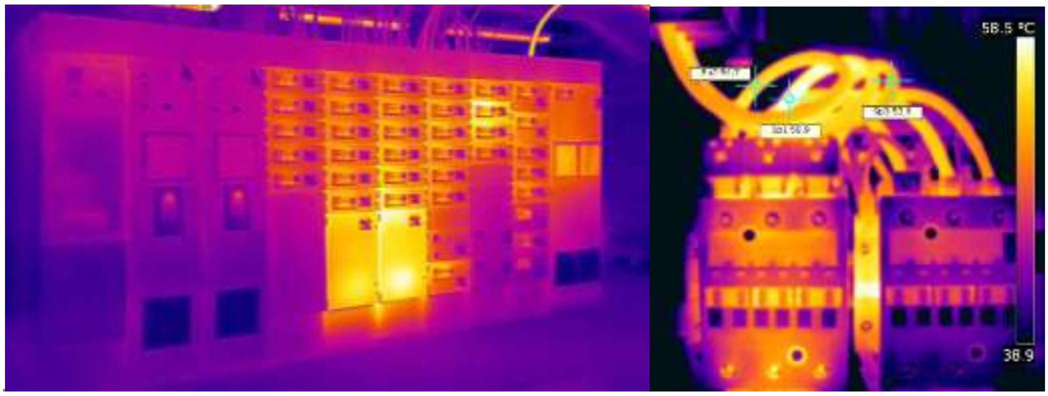

Figure 4. – Thermal imaging of main swithcboard

- *A UPS (uninterruptible power supply)* is usually immediately downstream of the switchboard. When inspecting a UPS, scan the input connections, the terminals, and the inverter section, where there are small fuses and capacitors. Under load, must be used thermal imager to check the battery section. Look at terminal posts, casings and feeders. A bad cell heats up very quickly under load. After the load scan, immediately scan the batteries not loaded. Bad cells cool very quickly when the load is removed. Finally, check the on-board transformer (if present).

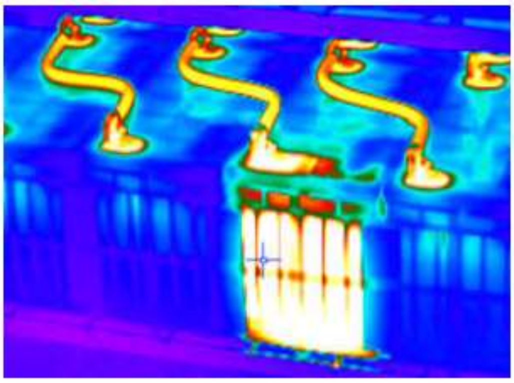

Figure 5. – IR scan of an UPS battery during a load test

- *Power distribution units (PDUs)* are downstream of the UPS and are typically located close to the servers, to which they distribute power. Normally, a PDU will have a circuit breaker panel and sometimes a transformer. In scanning PDUs, look at lugs and terminals, including circuit breaker terminals. Visually check for damage and debris, and if a PDU is not a straight-through-voltage model, scan the on-board transformer.

- *Server racks* are becoming increasingly more compact, opening up space for more servers in existing data centers, but they are also increasing demand on the centers' power and cooling capabilities. In fact, the heat generated by the today's blade servers has some experienced thermographers reporting that they no longer spend much time scanning server racks. The high heat makes comparative temperatures difficult. Still, the thermal imager is useful for monitoring power strips and power supplies built into the racks as well as wiring connections, plugs and plug strips. Look for overheating due to loose connections and loose or bent plugs. A thermal scan can also detect broken cords and broken conductors in wires. To detect the latter condition, look for what is called "the barber pole effect," in which can be observed the thermal differences of the twisted strands.

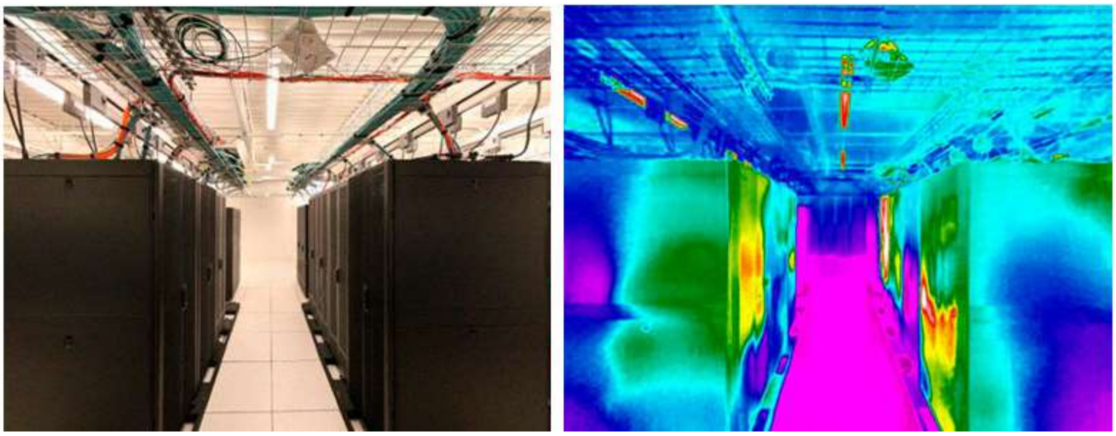

Figure 6. – Server racks thermal IR mapping

Traditionally, data centers have been air-cooled. Still today, the typical data center is air-cooled, utilizing the hot aisle/cold aisle layout. Cooled air is fed from the computer room air conditioning (CRAC) units to the cool aisles under a raised floor through perforated tiles (diffusers) up into the cool aisle, into the equipment and out the hot aisle. The heated air is then returned to the CRAC units.

Data center cooling systems have changed little over the past 25 years, but owing to the issues discussed above, new designs are being developed and tested, two are notable; cold aisle containment and liquid cooling. Cold aisle containment uses a raised floor, but contains the cold air between the cold aisle racks, sending the cold air directly to the server inlets, greatly reducing air mixing and short-cycling. Liquid cooling is used within most CRAC units, but liquid-cooled racks take advantage of the enhanced heat transfer characteristics of liquids. Since the CRAC units can be installed outside the main floor area, this design eliminates short-cycling. These systems are significantly more complex and expensive now, but may become more and more important as server densities increase beyond air cooling capabilities.

1. Tools for analyzing the data center thermal environment

Data center thermal/power control strategies to date have never had visibility into the thermal dynamics and energy dynamics inside the IT server assets. As such, data center ambient temperatures are presently controlled without insight into substantial energy wastage mechanisms that are present inside all enterprise server assets and storage assets that contain spinning hard disk drives (HDDs). These parasitic energy-wastage mechanisms include:

1. nonlinear power draw from the server fan motors: for many enterprise servers these days, the fan motors use more power than the CPUs, and much more than the memory; moreover, the aggregate fan motor power goes up with the cubic power of the fan RPMs.
2. nonlinear “leakage power” inside the CPU chips: leakage power is completely wasted (doesn't help with the computing), and goes up exponentially with chip temperature.
3. substantial energy wastage due to ambient vibrations in the servers (from internal fans and blowers) and the metal racks housing the servers (from nearby AC equipment, PDU internal fans, and flow-induced vibrations from fluid and fluid/air chillers).

CFD Thermal Modeling

The data center‘s cooling system must be designed and engineered to provide cooling to computer components. The objective of the design of the cooling system is to provide a clear path from the source of the cooled air to the intakes of the servers and to return the heated exhaust air to the CRAC efficiently. Data centers are usually designed and drawn with computer-aided drafting and design (CADD) software and modeled using *computational fluid dynamics* (CFD) modeling. These tools are available to predict performance for the design of a new data center. CFD is a valuable means of predicting data center thermal performance. Acceptable performance depends on accurate modeling of the energy-consuming components and the heat that they produce. CFD is, however, limited by the granularity of input data and as a result, requires many questionable assumptions.

No matter how complex and well-prepared, CFD modeling is not reality. Simple things like under floor cable or ducting installations have significant impact on theoretical flows of cooled air. Deviations from ideal performance will only show up after physical testing. Also, during and after construction, changes happen. Unforeseen issues like adding servers or increasing server densities are rarely re-modeled after construction.

Contractors move equipment, change cabling and conduit routes and HVAC ductwork, inadvertently creating voids and obstructions, reducing or increasing air pressure and diverting the flow of cooled and heated air. Obviously, these types of unforeseen changes to the thermal dynamics of the cooling system are seldom modeled. So, IRt is used to validate the CFD model (in a normal operating condition) and direct HVAC technicians and IT managers to heating problems (hot and cold spots). After repairs have been accomplished, IRt is used to check the repairs.

Thermal Mapping of a Data Center

Thermal mapping is a new approach to capturing the full ―in-situ‖ thermal condition of a data center and all of its equipment. The key advantage is that it is possible to get an overall view of the thermal condition of the entire room for a given point of time while still having the capability to zoom in on specific problems. This is very different from more traditional methods because it allows overall context and viewpoint selection, much like one gets with CFD modeling, but this is actual thermal imagery. For example, reports can demonstrate how a local thermal pattern visible in one aisle is actually the sign of a cooling air blockage across several aisles. When the overall layout of the servers, floor, walls and ceiling is available, what appears to be good thermal performance in one image may actually be wasteful excess cooling when the entire thermal map is analyzed. These problem areas are easy to see only with the overall image.

Temperature sensors have the advantage of monitoring temperatures continuously. Placed in various strategic locations, sensors are a good idea to monitor overall changes, but are certainly no replacement for thermal imaging. Typically, a single thermal image has over 75,000 thousand temperature points and, for instance, a five thousand square foot data center thermal image will have many millions. Another positive aspect of thermal mapping is that it allows for trending (comparing data gathered at different times). Information that did not appear to be important at the time of the survey can be used to discover a change in the cooling system. For instance, on one survey an area with no visible problems might be covered by thermal mapping where more conventional surveys might not capture that spot because it showed no apparent problems. When a subsequent survey shows a problem, it is possible to see the change or show that nothing has changed. It is possible to see changes like new cable runs or equipment installed which impacted free air flow. Every data center has equipment that is important to different groups of people. By methodically capturing all the thermal data and carefully post-processing it into user-friendly displays, these different people can see what is important to them, without having to be on scene during the survey. Also, experts in different disciplines can review the imagery and prepare reports at remote locations. To create a thermal map, one must collect the thermal and visual imagery in an ordered manner, carefully post-process it into mosaics, and create the construct to display it in 2-D and/or 3-D. To create meaningful reports, the thermal imagery must then be analyzed.

3-D Thermal Mapping View of Data Center Floors

Three-dimensional thermal mapping is a new approach to capturing the thermal condition of a data center and all of its equipment and is the most powerful of all tools for presentation to operators, consultants, contractors and HVAC professionals wanting to accomplish adjustments, repairs and redesigns. A 3-D model can be rotated and viewed from any angle. The under floor blockage can also be better appreciated or if there is an issue of conflict between the cooling and the humidity control systems. Overcooling below the dew point can trigger [electrical] heating to raise the temperature. This can have significant energy ramifications as the two systems fight one another. Expert HVAC personnel with access to the total picture can perform this kind of analysis without even having to be on site.

2. Multiscale thermal management of data centers

Proper operation of computing equipment imposes unique thermal management requirements. The typical approach currently used for thermal management of data centers consists of computer room air conditioning (CRAC or AC) units that deliver cold air to the racks arranged in alternate cold/hot aisles through perforated tiles placed over an under-floor plenum, see figure below.

A multiscale thermal solution is centered on the multiscale (multilevel) nature of the data centers. Understanding the different levels (scales) of thermal problems in electronic enclosures, we can select technologies to address the problems at component levels and to achieve the best overall effectiveness at the system level".

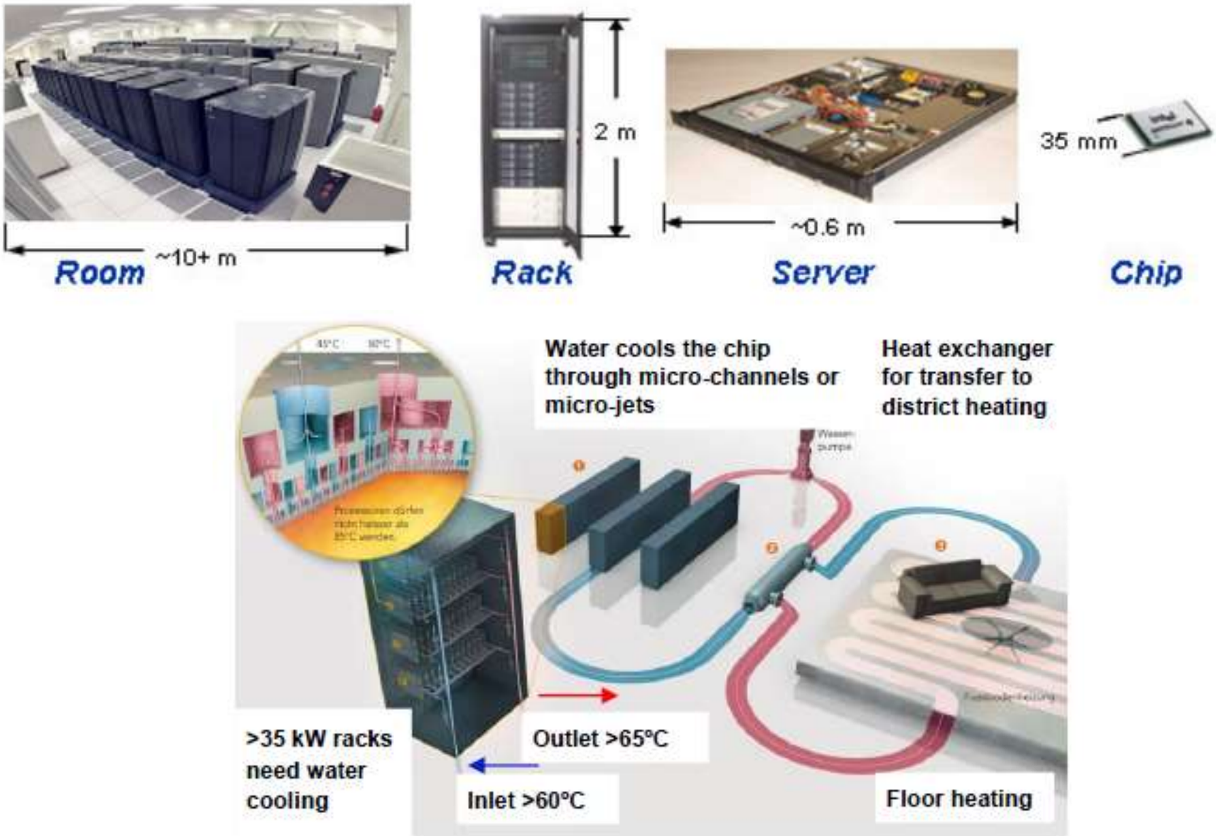

Figure 7 - High-performance chip-level cooling with improved energy efficiency and reduced carbon emission

In fact, designing and connecting cooling systems at different scales (levels) of a datacenter increases design freedom and results in a greater flexibility in configuring the system to achieve desired behavior and so enables designers to achieve a high power dissipating rack that was not possible before. All of the current air-cooling solutions and the innovative future approaches can be addressed within a multi-scale framework. Also, a multi-scale solution leads designers to achieve several innovative methods to achieve the design requirements of the next generation data centers. The obtained manageable heat load through an ideal multiscale thermal solution is believed to be the maximum heat load at chip and rack levels which could be effectively cooled in the next generation data centers.

1. Scale 1- Chip Level: This scale solution includes different methods to enhance heat dissipation from the chip itself. For example, it includes the design of effective heat sinks and micro heat exchangers attached to the chip for single or two-phase heat transfer. Also, design of new high conductive thermal adhesives or innovative methods for the attachment the die to the heat spreader, including the use of solder may be considered in this scale.

The chip level determines the rate of the heat generation in the data center, while the CRAC units at the room level are responsible to provide the cooling solution to keep the chip temperatures in a safe range. Satisfying the design requirements of an energy efficient open cooling system in today's and future air-cooled data centers is challenged by thermal modeling, inherent variability management, and having multiple objectives. These challenges are solved through a Proper Orthogonal Decomposition (POD) based reduced order thermal modeling, robust design principles, and the compromise Decision Support Problem (cDSP) construct.

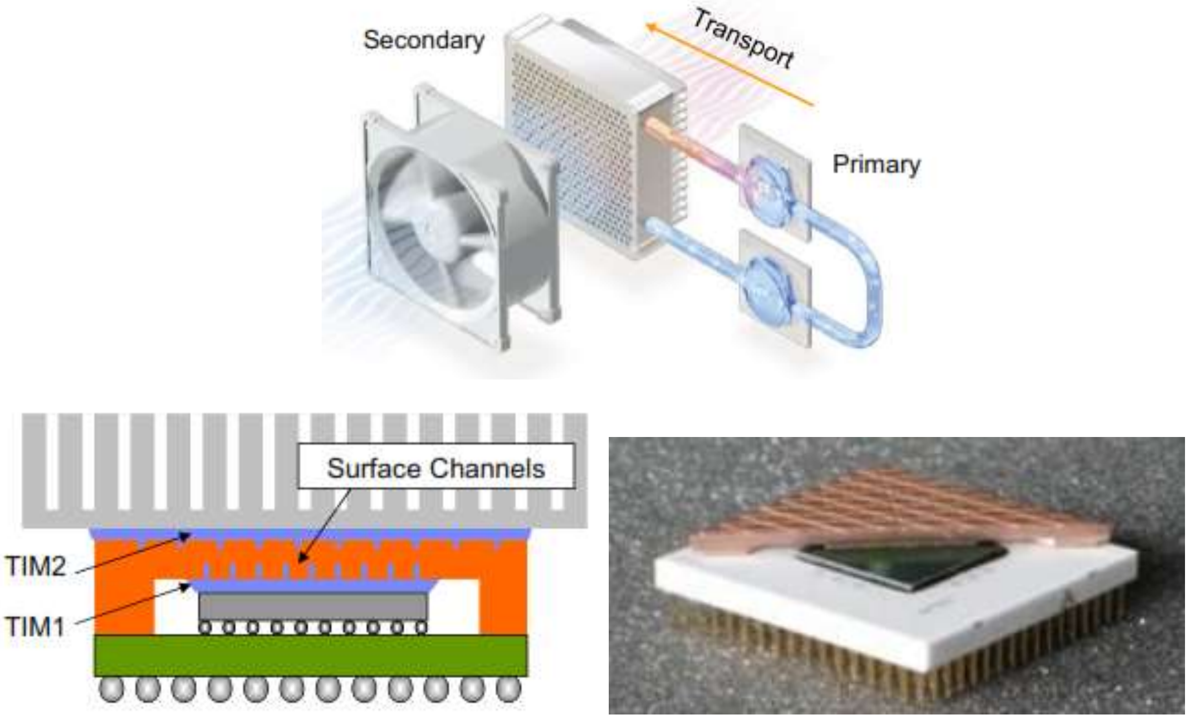

Figure 8. – IBM chip level thermal optimization with nested channels

2. Scale 2 - Server Level: These scale solutions happen within or on the printed circuit board of the server. These solutions mostly are related to designing of different cold plates in combination with the chip scale solutions. A Liquid Cooling System is a good example to use in this scale. Also, the innovative solutions in this scale should focus on effective directing of heat through a path from the chip to the board and finally to the ultimate ambient. The heat rejection through path of chip/substrate/board can be done by, for example, using additional solder balls as the thermal interconnects and heat spreaders or board-integrated liquid cooling as the board heat removal means.

3. Scale 3 - Chassis Level: In current air cooling systems, this scale is used to install the designed fans. This scale encompassing the space in front of each set of the blade servers could play a great role in applying various solutions to the rack, especially in combination with two previous scales. For example, this large space can be used by one or more macro heat exchangers. These heat exchangers can transfer the heat from the chips of the servers into the

cold air flowed by CRAC units in the rack scale (Scale 4). Also, installing a plate in this scale can provide a support for some components of a compact refrigeration system installed within these three scales to maintain the chip operating temperature as low as -70°C, if required.

4. Scale 4 - Rack Level: This scale is suitable for using macro heat exchanger for heat removal from the hot air exiting through the rack before entering the room (Scale 5). IBM rear door water-cooled heat exchanger is an innovative solution in this scale offered by IBM to cool the hot air before entering the hot aisle of a data center. Also, an air-water or air-refrigerant heat exchanger can be designed to install on the top or sides of the rack while they can directly use the chilled water flowing through the tubes in the plenum of the data center.

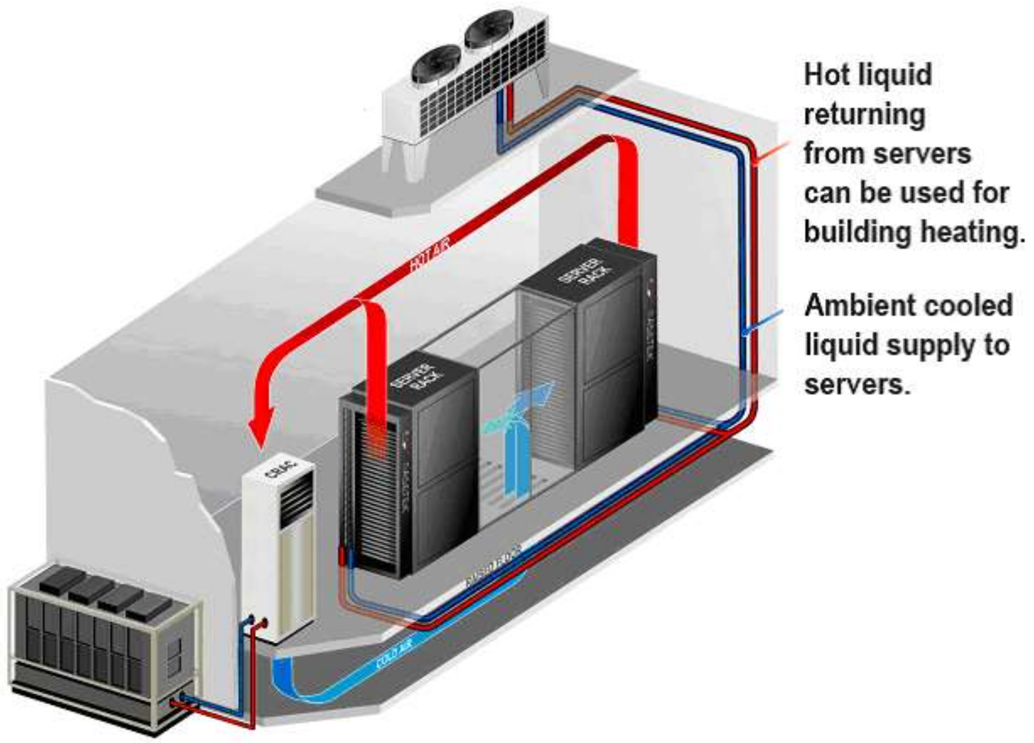

Figure 9. - Liquid and dry cooling by recycling waste heat

5. Scale 5 - Room Level: This scale has been of interest of various researchers in the recent years working to enhance the effectiveness of the typical air-cooling systems and prevent the recirculation. The efforts to optimize the configurations of the racks and CRAC units and the dimensions of the racks and room to prevent the recirculation are some of these works. Also, the heat transfer in this scale can be enhanced by using the different heat exchangers in possible connection with the rack (Scale 4) and their tubing through the room, if needed, such as the solutions offered by APC and Liebert.

6. Scale 6 - Plenum Level: Some done research in this scale includes optimization of the air flow through the plenum and perforated tile considering the plenum depth, under-floor

partitions, and tile specifications. Moreover, because the chilled water pipes pass through the plenum and below the racks, the plenum can have more roles in cooling the future data centers. The effective use of this scale combined with the previous scales can bring a lot of options in configuring the liquid-cooling systems for racks of data centers.

Finally, by properly designing and effectively combining the various solutions at each scale, an effective and efficient cooling solution for the next generation data center would be designed and deployed.

Software for data center design

Data center infrastructure management (DCIM) is a category of solutions which were created to extend the traditional data center management function to include all of the physical assets and resources found in the Facilities and IT domains. DCIM deployments over time will integrate information technology (IT) and facility management disciplines to centralize monitoring, management and intelligent capacity planning of a data center's critical systems. Since DCIM is a broadly used term which covers a wide range of data center management values, each deployment will include a subset of the full DCIM value needed and expected over time.

Full DCIM deployments will involve specialized software, hardware and sensors. With more than 75 vendors in 2014 self-identifying their offerings to be part of the DCIM market segment, the rapid evolution of the DCIM category is leading to the creation of many associated data center performance management and measurement metrics, including industry standard metrics like PUE, CUE and DCeP – Data Center Energy Productivity as well as vendor-driven metrics such as PAR4 - Server Power Usage and DCPM – Data Center Predictive Modeling with the intention of providing increasingly cost-effective planning and operations support for certain aspects of the data center and its contained devices.

Since its identification as a missing component for optimized data center management, the broad DCIM category has been flooded with a wide range of point-solutions and hardware-vendor offerings intended to address this void. The analyst firm Gartner Research has started using a term to try and focus on DCIM vendors with a more comprehensive set of capabilities. DCIM Suite vendors number less than two dozen in 2014, and consist of software offering which are wide-ranging and integrated in nature. The existing suites touch upon both IT and Facilities and depending upon the vendor's heritage may have a bias towards either IT asset lifecycle management or facilities monitoring and access. It is likely that for an extended period of time, the DCIM Suites that exist will continue to have their core strength in one discipline or the other, but not equally addressing both. Important to note is that there are dozens of other vendors whose technologies directly support or enhance the DCIM suites. In general, these specialists' offerings can also be used as viable stand-alone solution to a specific set of data center management needs. In the fourth quarter of 2014, Gartner released their Magic Quadrant and Critical Capabilities reports which are the first tangible approach to a quantitative comparison of the values each vendor has to offer. The Magic Quadrant focused on 17 vendors, while the Critical Capabilities report examined just 7 that they considered broad enough to compare.

The large framework providers are re-tooling their own wares and creating DCIM alliances and partnerships with various other DCIM vendors to complete their own management picture. The inefficiencies seen previously by having limited visibility and control at the physical layer of the data center is simply too costly for end-users and vendors alike in the energy-conscious world we live in. These multi-billion dollar large framework providers include Hewlett-Packard, BMC, CA and IBM/Tivoli and have promised DCIM will be part of their overall management structure. Today, each is defining their approach in doing so through organic and partnership efforts.

While the physical layer of the data center has historically been viewed as a hardware exercise, there are a number of DCIM Suite and DCIM Specialist software vendors who offer varied DCIM capabilities including one or more of the following; Capacity Planning, high-fidelity visualization, Real-Time Monitoring, Cable/Connectivity management, Environmental/Energy sensors, business analytics (including financial modeling), Process/Change Management and integration well with various types of external management systems and data sources.

DCIM software is used to benchmark current power consumption and equipment temperature through real-time feeds and equipment ratings, then model the effects of "green" initiatives on the data center's power usage effectiveness (PUE) and data center infrastructure efficiency before committing resources to an implementation.

nlyte software, the only provider of performance-based solutions for Data Center Infrastructure Management (DCIM). Accelerated customer acquisition, global partnerships, product innovation and market expansion are the cornerstones of nlyte Software’s recent success in the DCIM market. *nlyte s*oftware provides intelligent capacity planning (figure 12). Its performance-based solution enables the world’s largest companies to optimally place data center assets to make the most efficient use of power, cooling and space. Clients often experience a reduction in operating expenses by as much as 20% annually. The software can generate reports for monitoring the data center’s operation (figure 13) by real-time monitoring and alarming of power & environmental parameters from a powerstrip, PDU, server level CPU or a sensor, using the most popular protocols.

Another design software, also agreed by specialists, is *6SigmaRoomLite*, which provides a simple approach to the build of a scientifically accurate Virtual Facility. The software offers a clear, simple and accurate view of conceptual design and environmental issues within data centers, by modeling, testing and analyzing the infrastructure, cooling, cabinets, equipment and much more. The *6SigmaRoomLite* modules allow the build of virtual data center at different levels of complexity and allow closer examinations of facility workings at rack level, chip level and management level. The complex CFD calculations are hidden from the user, but provide simulations (figure 14) including airflows, power calculations, network mapping and efficiency calculations (PUE). Its library contains thousands of ACUs, PDUs, UPS, floor grilles, servers to ensure the model matches the reality.

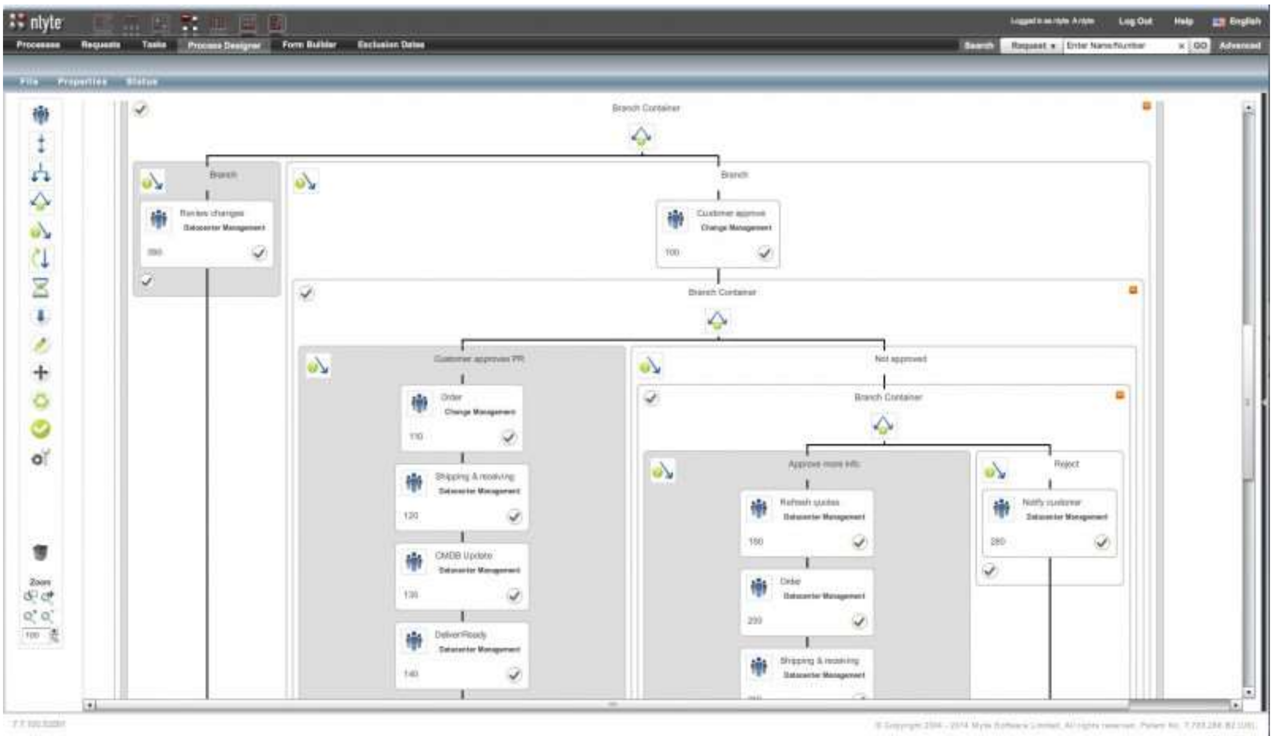

Figure 10. - Workflow management od fata center with *nlyte*

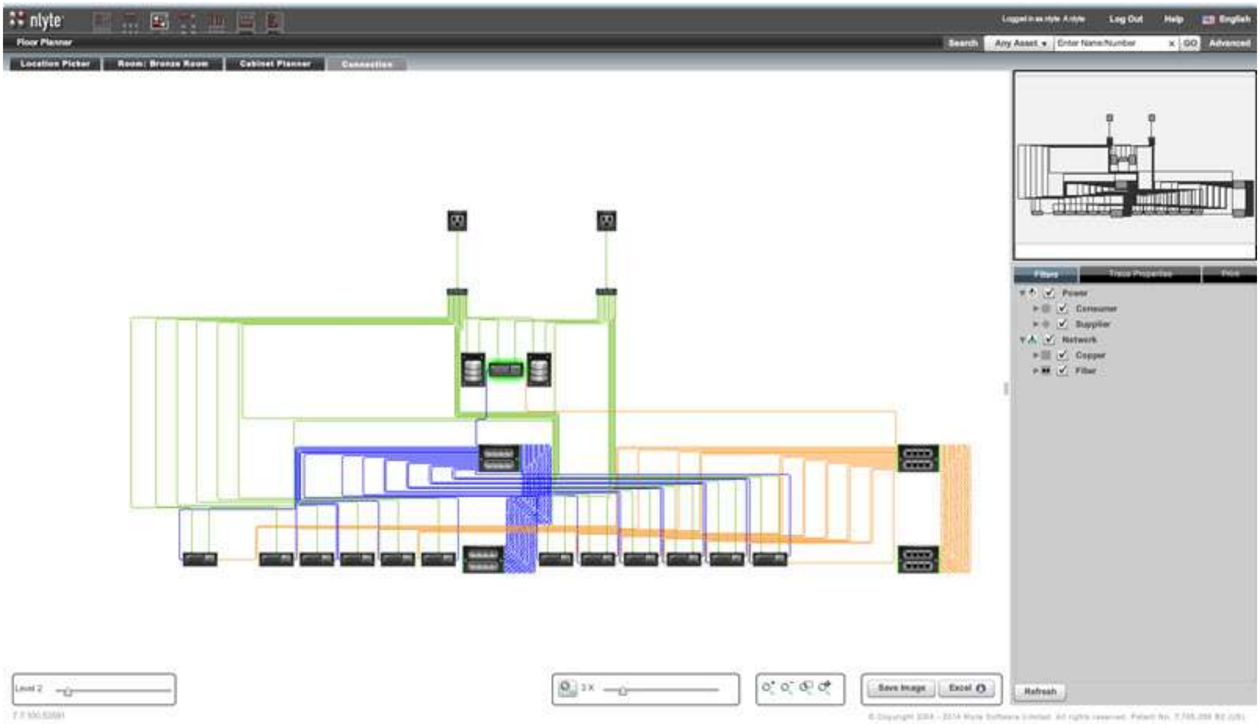

Figure 11. - Graphical approach to tracing both power and data connections a data center

Figure 12 - Capacity planning – rack rear view

Figure 13 -Nlyte reports on assets's capacity and on cooling consumption by asset type

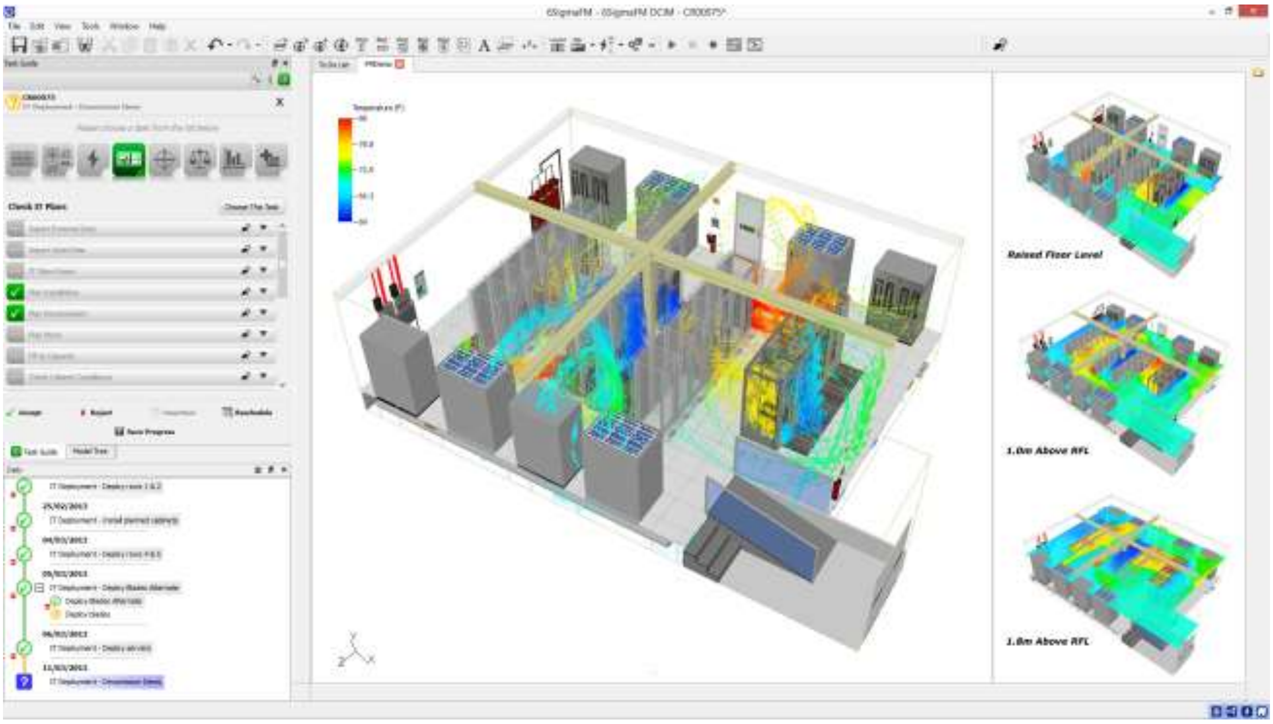

Figure 14. – Data center design with CFD simulations

Using DCIM with simulations is the ultimate way to improve data center practices. The overall benefits of using simulation in an integral way to design, optimise and manage data centres are great: a user can test its facility throughly in any way with the virtual facility acting as a buffer zone, protecting the real one from harm. A bespoke virtaul model can contain and record every necessary detail of design, operation and management. This approach removes the chance for inaccurate assumptions to be made in design – all possible configurations cand be tested for best results before any equipment is deployed.

References:

1. Schmidt, Roger R., E. E. Cruz, and Madhusudan K. Iyengar, *Challenges of data center thermal management*, IBM Journal of Research and Development 49.4.5 (2005): 709-72
2. Buyya, Rajkumar, Anton Beloglazov, and Jemal Abawajy, *Energy-efficient management of data center resources for cloud computing: a vision, architectural elements, and open challenges*, arXiv preprint arXiv:1006.0308 (2010).
3. Shah, Amip J., et al., *Exergy analysis of data center thermal management systems*, Journal of Heat Transfer 130.2 (2008): 021401.
4. Joshi, Yogendra, and Pramod Kumar, *Energy efficient thermal management of data centers*, Springer Science & Business Media, 2012.
5. http://www.infraredimagingservices.com/electrical-infrared
6. http://www.fluke.com/fluke/sgen/community/fluke-news-plus/articlecategories/ti/datacenter-ti
7. Gregory Stockton, *Using Thermal Mapping at the Data* Center, Stockton Infrared Thermographic Services, InfraMation 2010 Proceedings.
8. Patterson M, Tschudi B, Vangeet O, Cooley J, Azevedo D., *A metric for measuring the benefit of reuse energy from a data center*.,The Green Grid; 2010.
9. Kenny Gross, Kalyan Vaidyanathan, Oracle White Paper, *MegaWatt-Hours Av*oided, Oracle Quality Division Oracle San Diego Physical Sciences Research Center, 2015.