Data Center Thermal Runaway

A review of cooling challenges in high density mission critical environments.

White Paper 105

2128 W. Braker Lane, BK12

Austin, Texas 78758-4028

www.activepower.com

Objective

This paper will review the impact on mechanical systems in the event of a power outage in medium to high density mission critical data centers. The paper will provide insight into the implications of loss of cooling to the critical IT equipment and how long it takes for the equipment to enter an automatic state of thermal shutdown.

Introduction

Research conducted in April 2007 by Aperture Research Institute suggests that as much as 22.3 percent of all outages in a data center are caused by servers overheating and shutting down on their own. This is neither fiction nor an incident reserved only for the ultra high density data centers currently under construction. Self-protective shut down due to high temperature in the server is very common from small to large data centers and computer rooms and typically occur during the event of a power outage where the UPS (uninterruptible power supply) continues to provide uninterrupted power, but the cooling systems are not.

DYNAMICS OF THE MECHANICAL SYSTEM

The UPS is tasked with feeding the server equipment conditioned and uninterrupted power. Unfortunately, it is not practical to run the data center cooling (mechanical) system off the UPS for a few different reasons:

- A typical cooling system in a data center consumes as much power and often more than the server equipment it is protecting.
- The motor load characteristics of the outdoor chiller or condensers caused by the constant on/off cycling leads to significant step-loads that can trip a UPS and drop the remaining load.
- As a result of the additional and possibly oversized UPS power required, cost usually prohibits the cooling system to be on UPS.

Larger data centers typically have a fuel-fired engine coupled with the UPS system to allow for continuous operation during events where main utility power is unavailable for several hours or days. Where engines are available, the cooling system will typically be backed-up to that. Engine or not, the result is similar. While the server load is kept up and running during a power outage by the UPS, the cooling system will fail until the engine restores critical power to all systems. Once power has been restored to the cooling system, chillers or condensers will cycle back on over a period of time that could be as long as 15 minutes. To avoid a rapid step load that would push the engine to its knees and shut down, systems are sequentially started – a little like starting from zero mph in the 4th gear, not a good idea. In fact, it can take as long as 10-15 minutes from the time power is restored by the engine until the cooling system receives the start signal. This is to avoid intermittent voltage drops and to ensure the engine is ready. Depending on the number of chillers or condensers in the facility, each will cycle on one by one until all are running at full capacity.

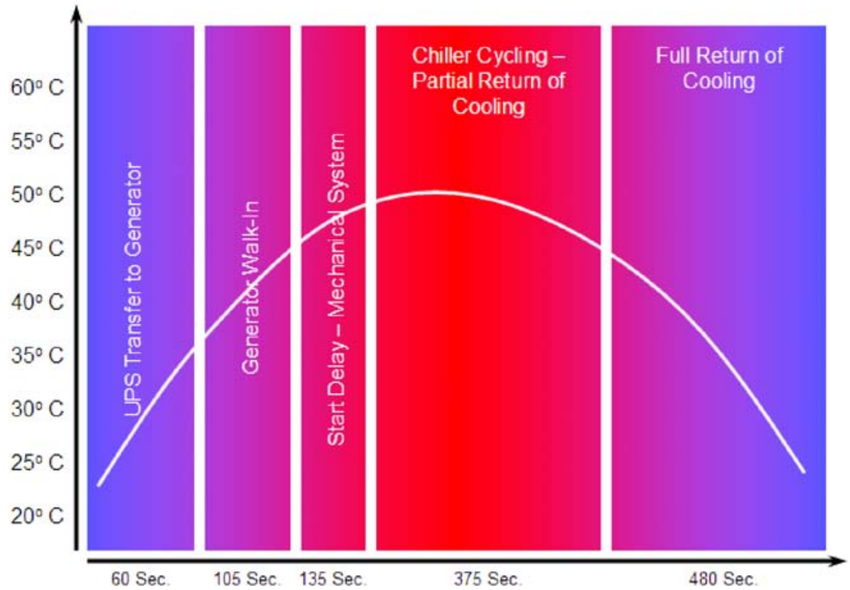

Figure 1: Simplified cause and effect diagram during a power outage with legacy engine start.

This process can take as long as 10-15 minutes depending on the type and size of the system all of which happens while the servers are powered and humming away without cooling. Unfortunately, the 15kW blade server cabinet would have called it quits at 60 seconds.

SELF-PROTECTIVE SERVER SHUTDOWN

The need for cooling has reached a point where we no longer count in hours or minutes of runtime without cooling, but in seconds. A typical rack-mount server features an embedded thermal management system. The system ensures the server does not sustain permanent damage in the event of rapid temperature increase.

ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) Technical Committee 9.9 recommends an inlet air temperature of 20-25 degrees Celsius (68-77 degrees Fahrenheit) which is endorsed and backed by the server vendors and their warranties.

ASHRAE
Advancing HVAC&R to serve humanity and promote a sustainable world

Datacom Equipment Environment Specifications

Class	Allowable Temp	Recommended Temp	Allowable Relative Humidity	Recommended Relative Humidity	Max Rate of Change
1	15-32° C 59-90° F	20-25° C 68-77° F	20-80%	40-55%	5° C 10° F

Figure 2: Equipment environment specifications in data centers (ASHRAE Technical Committee 9.9).

Temperature fluctuations should be no more than 5 degrees Celsius over 60 minutes. Fluctuations and constant exposure to temperatures outside of this band will degrade the useful life of the IT equipment and research conducted by the Uptime Institute has shown failure rates exceed normal field experience by more than four times. The series of temperature points at which actions are taken vary from one server manufacturer to the other. However, the default first level warning to the user is typically issued at 55 degrees Celsius. The second and critical threshold is at 65 degrees Celsius. The server will again notify the user through the operating system and if no action is taken and the temperature remains above 65 degrees Celsius for 20 seconds or more, an automatic shut down command of the operating system and power supply is issued possibly resulting in loss of data to the users. It is important to note the temperature measured inside the server equipment is substantially higher than the room ambient temperature.

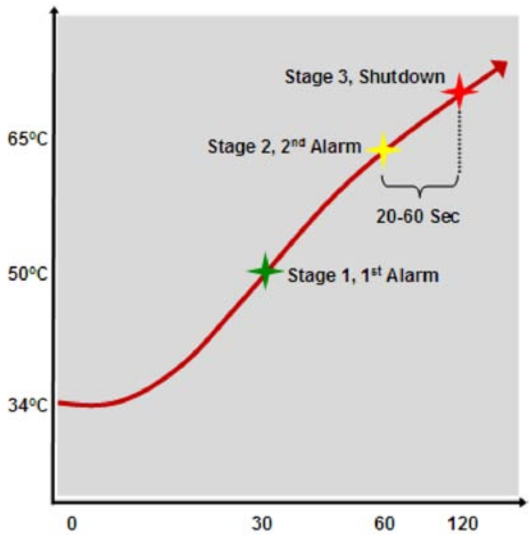

Figure 3: Simplified graphical representation of self-protective thermal shutdown.

Data polling from live servers has shown the delta between the inside server temperatures is typically 30-35 degrees Celsius higher than the ambient, although heavy dust and debris build up in and around the server can push this delta far higher. This means that 65 degrees Celsius measured inside the server equipment, as a thumb rule in an adequately designed data center, equates to 35-40 degrees Celsius in the room.

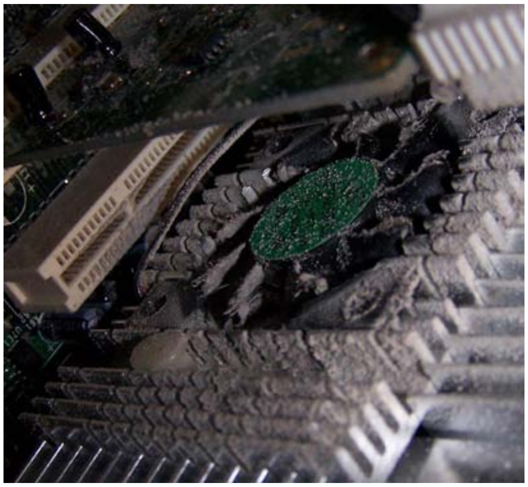

Figure 4: Example of significant dust build up on the heat sink.

Storage equipment typically has a lower temperature threshold due to its moving parts and especially spinning discs. In the case of storage equipment, research by Hitachi Data Systems has shown drive reliability degradation is a longer term challenge. The research showed temperatures fluctuating more than 5 degrees Celsius caused excessive out-gassing of the lubricants in the spindle motors driving failures up 15 percent in the following 30 days. By most standards, storage devices are more critical than the compute servers given their high data retention rate.

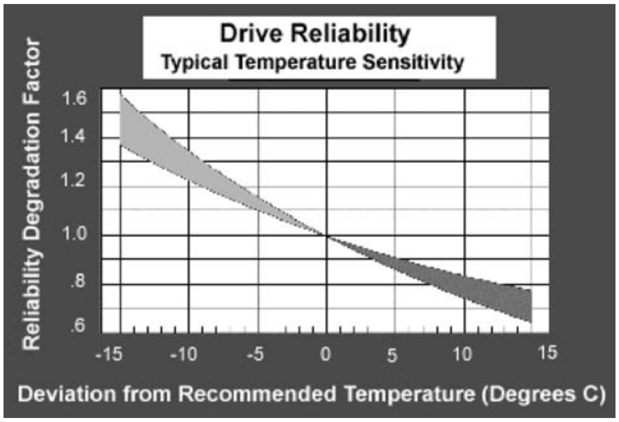

Figure 5: Drive reliability research (Hitachi Data Systems).

Figure 6: Example of a hard drive.

The threshold for network equipment, on the other hand, is somewhat higher than server equipment given their broad design use in remote unmanned sites, closets and conditioned data centers.

INCREASED COLD AISLE TEMPERATURES

It is worth noting that several studies on optimal data center environments suggests that the ASHRAE TC 9.9 recommended 20-25 degrees Celsius threshold can be increased to as much as 30-35 degrees Celsius. The main driver here is significant energy savings as a result of air conditioning systems using less energy to remove heat in the data center. Most server equipment is designed for an operating temperature between 10-35 degrees Celsius, although none of the server OEMs has endorsed a general data center wide operating temperature outside the current band. Data gathered in this report supports the recommended ASHRAE TC 9.9 temperature band of 20-25 degrees Celsius and any increase over and above that band will significantly reduce the reaction time during the event of a power loss and subsequent cooling outage.

> *"When the temperature increases, the rate of a chemical reaction increases also. Iron rusts or food spoils, rises more rapidly when conditions are warm than when they are cool. One result of this similarity is the useful generalization that for many reactions which occur near room temperature, a temperature increase of 10 degrees Celsius approximately doubles the rate of the reaction."*
>
> **Svante Arrhenius, 1903 Chemistry Nobel Prize winner and founder of the science of physical chemistry.**

THERMAL RUNAWAY

Several real life studies have been conducted on thermal shut down in data centers during cooling outages. Certain factors play into when the data center or parts thereof will shut down due to loss of cooling:

- Volumetric size of the room. The taller the ceiling, the more heat will rise and accumulate under the ceiling or ceiling tiles rather than in the aisles.
- Server workload. If the server workload and thereby the heat output is low, it will extend the time marginally.
- CRAC (computer room air conditioning) fans on UPS power. If the fans in the CRAC units continue to run, although the coil provides no heat removal, it will extend the time further.

The following chart represents nine data points collected through research by The Uptime Institute, Emerson Corp., EYP Mission Critical Facilities and Active Power, Inc. and plotted on a graph including the associated trend line for all nine data points.

		W/Rack	Seconds	Temp Rise - °C	Temp Rise / Sec - °C
The Uptime Institute, Menuet, Turner, May '06	■	600	180	3.5	0.0194
US Patent 6,170,561, O'Grady, Jan '01	▲	1,050	900	26.6	0.0296
The Uptime Institute, Menuet, Turner, May '06	■	1,200	180	9	0.0500
EYP, Kosik, Oct '06	●	2,500	90	5.6	0.0622
Active Power, Olsen, May '07	◆	3,500	40	3.9	0.0975
Emerson, Oct '05	■	4,000	360	66.7	0.1853
Active Power, Olsen, May '07	◆	5,000	50	10	0.2000
Emerson, Oct '05	■	8,000	180	66.7	0.3706
The Uptime Institute, Menuet, Turner, May '06	■	9,000	50	22	0.4400

Figure 7: The matrix represents the data points collected. It shows the recorded temperature rise in Celsius over a given period of time before the IT equipment went into a self-protective thermal shutdown. The column on the far right shows the normalized temperature rise in Celsius per second for each of the densities recorded.

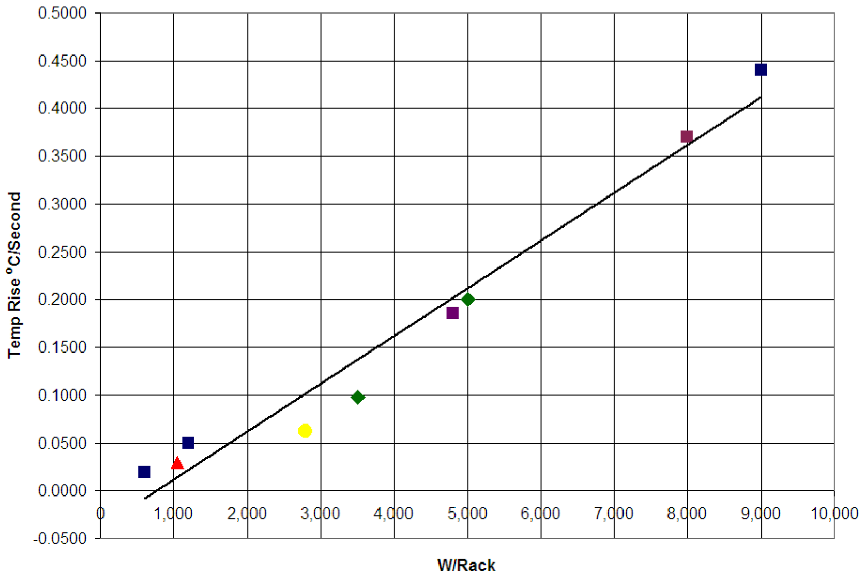

Figure 8: The graph is a visual representation of the matrix in figure 7. Temperature rise in Celsius per second for each of the densities recorded.

The trend line is expressed in this equation: $y = 0.00005x – 0.0381$ and can be applied as an approximate measure to any number of densities per rack to provide an indication of the time it takes for a given rack server load to overheat and shutdown as a result of loss of cooling.

Watts per Cabinet	Seconds								
	30	60	120	240	300	360	420	480	540
1500W	21.1	22.2	24.4	28.9	31.1	33.3	35.5	37.7	39.9
3000W	23.4	26.7	33.4	46.9	53.6	60.3	67.0	Shutdown Command	
5000W	26.4	32.7	45.5	70.9	Shutdown Command				
8000W	30.9	41.7	63.4	Shutdown Command					
10000W	33.9	47.7	75.4	Shutdown Command					
15000W	41.4	62.7	Shutdown Command						
20000W	48.9	Shutdown Command							
30000W	Shutdown Command								

Figure 9: The matrix represents the temperature increase over time (seconds) as it relates to the density per cabinet (Watts per Cabinet) in the event of cooling loss. It is assumed the base temperature is 20 degrees Celsius.

COMPUTATIONAL FLUID DYNAMICS (CFD) MODELING

CFD modeling is used extensively throughout the data center industry to accurately depict how a given fluid, in this case air in a data center, will behave given certain circumstances. Colors represent the temperature with blue being the coolest and red being the hottest. Figure 10 shows what a typical data center looks like when adequately designed and in normal operation.

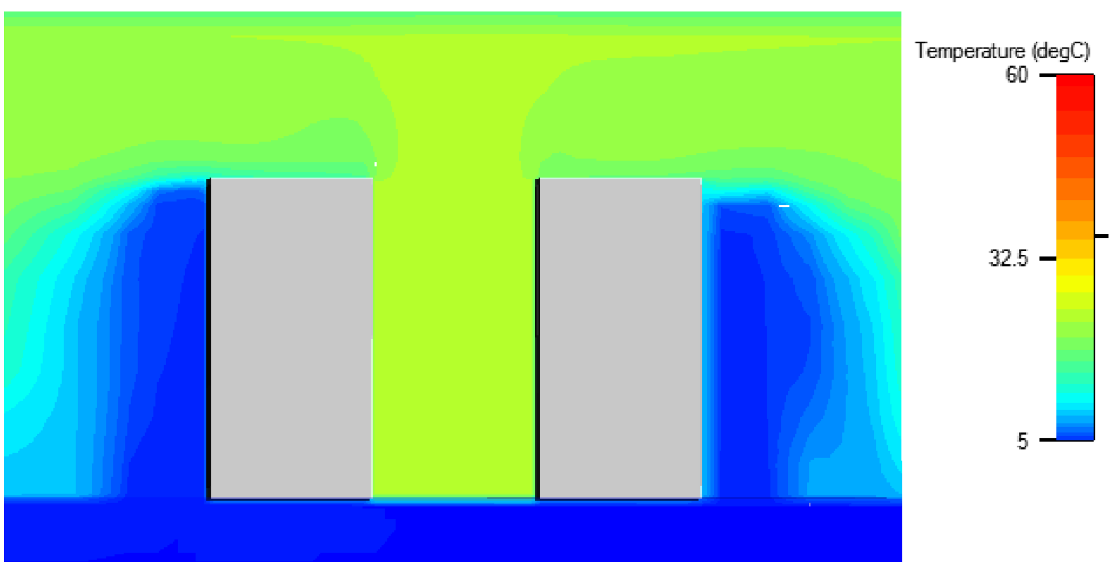

Figure 10: Side view looking down the middle of the hot aisle.

Figure 11 and 12 shows what happens in a room full of cabinets running at 5,000 watt load. The time stamp is 120 seconds after loss of cooling.

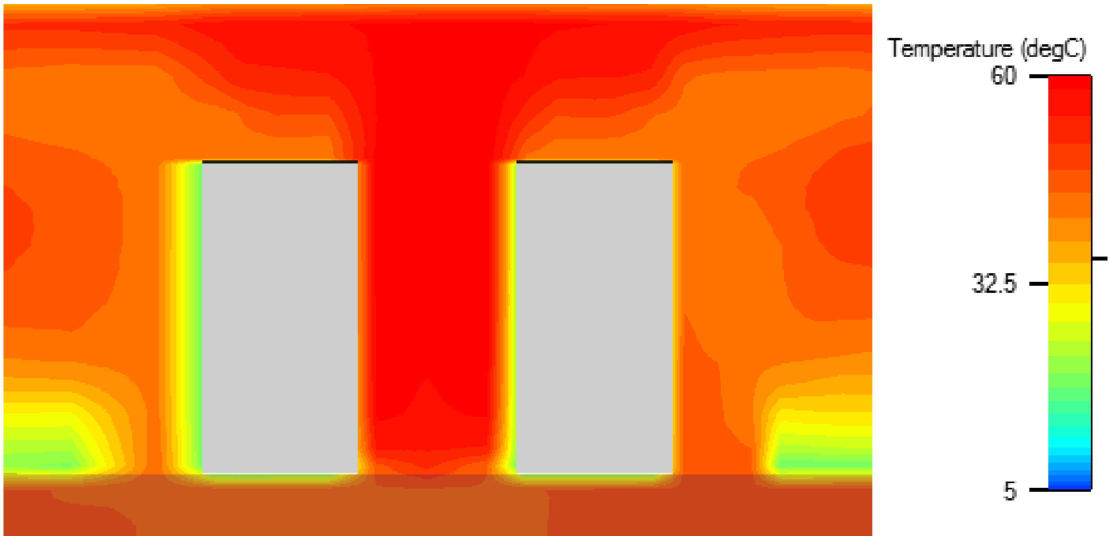

Figure 11: 5,000 watts per cabinet load after 120 seconds of cooling loss.

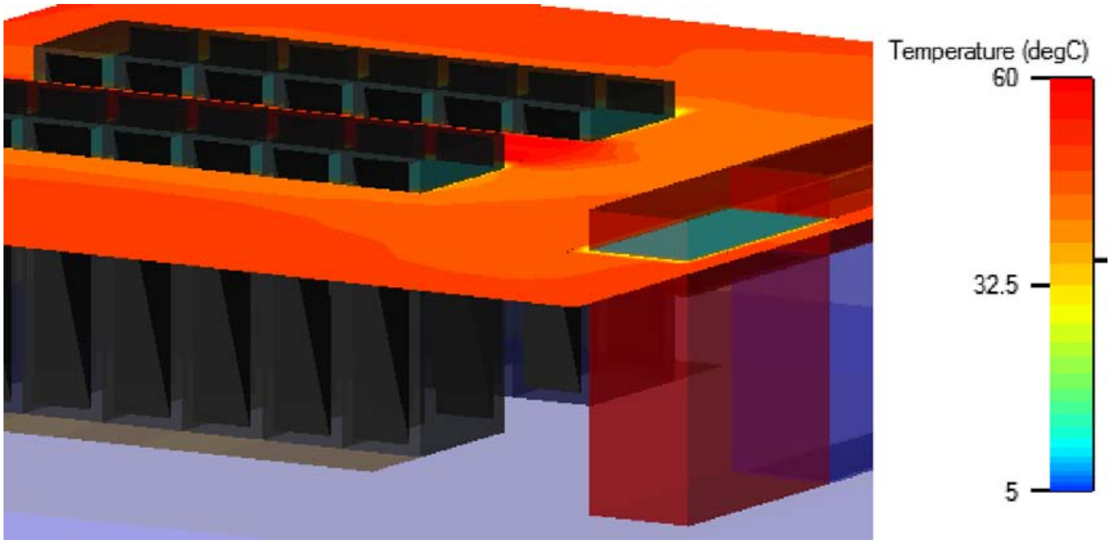

Figure 12: 5,000 watts per cabinet load after 120 seconds of cooling loss.

Figure 13 and 14 shows the same 5,000 watts/cabinet room with a time stamp of 240 seconds after loss of cooling resulting in a devastating effect on the ambient temperature in the room.

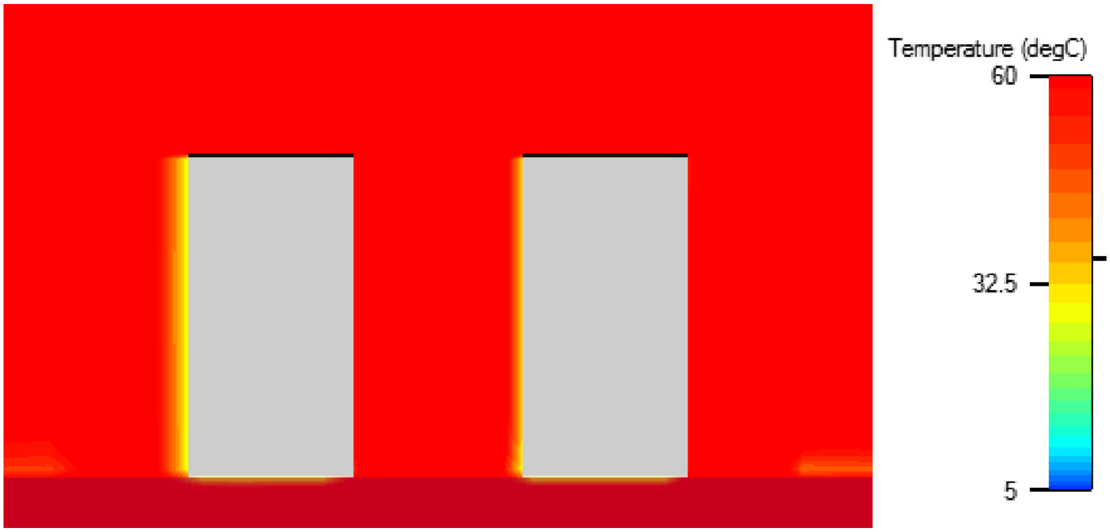

Figure 13: 5,000 watt per cabinet load after 240 seconds of cooling loss.

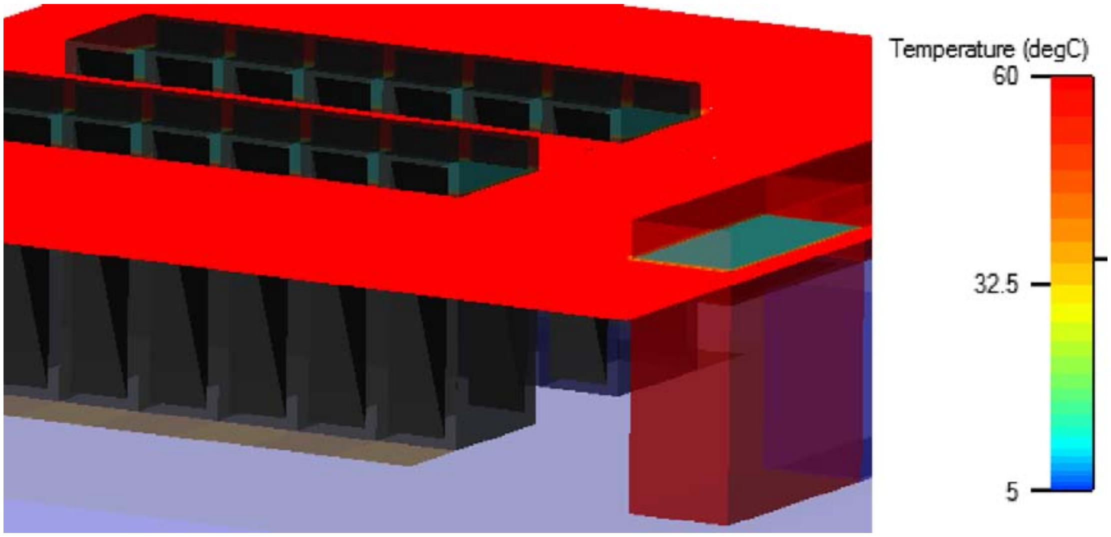

Figure 14: 5,000 watts per cabinet load after 240 seconds of cooling loss.

REDUCING EXPOSURE TO THERMAL RUNAWAY

The single most effective way of reducing the exposure to thermal runaway in a data center due to a power outage is to reduce the time it takes to reestablish power to the cooling (mechanical systems). One way of accomplishing this is to reduce the time it takes to signal the standby engine to crank up. To avoid the commonly understood ghost start phenomenon, where engines are signaled to start up even though power is restored within seconds of the original outage, the system has to be configured appropriately. According to EPRI (Electric Power Research Institute) nearly 99 percent of all outages last less than 10 seconds, which means the delay should be set at 5-10 seconds. In properly designed high availability critical power systems by definition, the diesel engine will start and assume (walk-in) load when commanded and typically within 5-6 seconds.

Synchronization time of multi-engine configurations for capacity or redundancy can further extend the walk-in time. However, the use of digital engine controllers can help reduce and control the walk-in time in a much tighter band close to that of single engines. A digital engine controller provides faster starting times by keeping track of the crankshaft angle as the machine spins down to rest. Knowing the crank angle on start up allows the best control of fuel injection during the next start, reducing the synchronization and walk-in time.

FIGURE 15: EXAMPLE OF A DIGITAL ENGINE CONTROLLER (COMAP CREATIVE ENGINEERING)

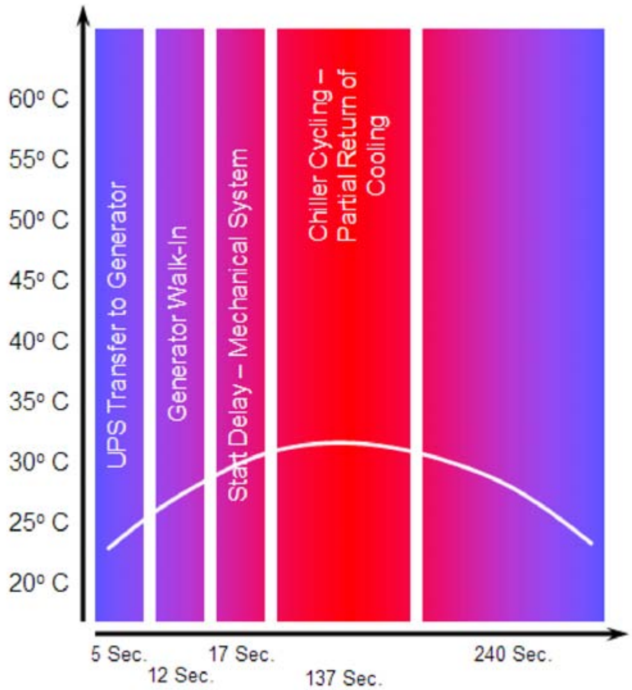

FIGURE 16: SIMPLIFIED CAUSE AND EFFECT DIAGRAM DURING A POWER OUTAGE WITH RAPID ENGINE START.

A combined 10-15 seconds from the time of the power outage until power is restored to the mechanical system stands in contrast to the commonly specified 100-120 seconds time laps in legacy power system designs today. A reduction of this magnitude will significantly reduce the ambient room temperature and thereby the risk of thermal runaway or self-protective shut down by the servers. As an example using the equation in figure 9, a 10,000 watt cabinet filled with blade servers will rise 0.4619 degrees Celsius per second during the event of a cooling outage. If the commonly specified 120 second time laps were employed in the power system architecture, the temperature rise would be 55 degrees Celsius or far higher than the servers own maximum temperature threshold. However, if a much shorter time laps of 15 seconds were employed in the same power system architecture, the temperature rise would be reduced to 6.9 degrees Celsius, an 87 percent improvement over that of commonly specified power system architectures.

Arguably, this would be a simple change even in existing facilities, but an even greater point is why then the need for the commonly recommended 15 minute of ride-through time? As densities in data centers continue to grow, it would appear the 15 minute battery bank is significantly over-rated from a number of standpoints:

1. A 15-minute allowance for "soft shutdown" of computer or other loads is entirely irrelevant since, by definition, shutdown after such times is intolerable to most businesses.

2. A 15-minute allowance for a "second crank" on the standby engine is also without merit because if, in the rarest of circumstances, the engine system does not start within the first five or six seconds, then like a car, it will likely not start within the next 15 minutes either. One should note the outage that occurred on July 24, 2007, at 365 Main, Inc., colocation facility in San Francisco. The engines failed to start up, taking a reported 37 minutes to manually crank the engines back up.

3. A 15-minute ride-through would have devastating effects on the temperature in the data center. Even a lightly loaded data center at 2,000 watts/cabinet would have increased 55 degrees Celsius according to figure 9. On November 12, 2007, it was reported data center developer and operator Rackspace, Inc., was impacted by a power outage in their Dallas/Fort Worth data center. During restart of the chillers, the data center recorded rapidly increasing temperatures and it was decided to turn off loads to mitigate a thermal runaway.

CONCLUSION

Based on real world data and laboratory modeling conducted in support of this paper, it can be concluded a rapid engine start in the range of 10-15 seconds rather than 100-120 seconds, common in many power system architectures today, is desired to help mitigate thermal runaway in the data center. Additionally, a tighter control of temperature fluctuations in the room can significantly mitigate long term reliability degradation of spinning storage drives and compute servers caused by excessive expansion and contraction of moving components. The commonly used practice of 15-minute ride-throughs is challenged by today's standard of completely irrelevant "soft-shutdown" of computers during a power outage and the time needed for a "second crank" on the standby engines.

REFERENCES

American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., Technical Committee 9.9 "Mission Critical Facilities, Technology Spaces and Electronic Equipment"
http://tc99.ashraetcs.org/

Aperture Research Institute, "Data Center Professionals Turn to High-Density Computing as Major Boom Continues," June 2007
http://www.aperture.com/about/aperture_research_institute.php

Opengate Data Systems
http://www.opengatedata.com/

Hitachi Data Systems
http://www.hitachidatasystems.com/

Electric Research Power Institute (EPRI)
http://www.epri.com/

The Uptime Institute
http://www.uptimeinstitute.org/

Institute of Electrical and Electronics Engineers, Inc. (IEEE)
http://www.ieee.org/

ComAp Creative Engineering
http://www.comap.cz/