



Dissipazione termica nei data center

*Un esame delle sfide poste dal
raffreddamento in ambienti
mission critical ad alta densità.*

Libro bianco 105

2128 W. Braker Lane, BK12

Austin, Texas 78758-4028

www.activepower.com

OBIETTIVO

Il presente documento esamina l'impatto sui sistemi meccanici in caso di un'interruzione di rete in data center mission critical di media o alta densità. Il documento contiene informazioni sulle implicazioni scaturite dalla perdita di raffreddamento dell'apparecchiatura IT di prima necessità e il tempo necessario affinché la stessa passi automaticamente allo stato di arresto termico.

INTRODUZIONE

La ricerca condotta nell'aprile 2007 da Aperture Research Institute suggerisce che, tra tutte le interruzioni che avvengono nei data center, il 22,3 per cento sono imputabili al surriscaldamento dei server e al loro conseguente autospegnimento. Questo risultato non è né un'esagerazione, né riguarda i soli data center ad altissima densità come quelli attualmente in fase di realizzazione. L'arresto autoprotettivo che segue un surriscaldamento è molto comune nei server, sia nei piccoli e grandi data center, sia nei locali computer, e di solito si verifica durante un'interruzione di rete quando l'UPS (gruppo di continuità) continua a fornire alimentazione mentre gli impianti di raffreddamento non fanno altrettanto.

DINAMICA DEL SISTEMA MECCANICO

L'UPS ha il compito di alimentare i server con energia condizionata e continua. Per i motivi elencati di seguito, non è pratico alimentare l'impianto di raffreddamento (meccanico) di un data center con un UPS:

- Un tipico impianto di raffreddamento per data center consuma più energia delle apparecchiature server che protegge.
- Le caratteristiche di carico del motore di un raffreddatore o dei condensatori esterni generate dal costante ciclo di accensione e spegnimento comportano un considerevole carico incrementale che può innescare un UPS e provocare il conseguente abbandono del carico residuo.
- In molti casi, ne consegue che i costi da sostenere per garantire l'enorme potenza aggiuntiva necessaria per alimentare un impianto di raffreddamento con un UPS sono del tutto proibitivi.

Data center sovradimensionati di solito impiegano un motore a combustibile abbinato a un sistema UPS. Questo binomio consente il funzionamento continuo anche nel caso in cui l'alimentazione di rete non sia disponibile per ore o addirittura giorni. Dove presente, l'impianto di raffreddamento di solito viene commutato sui motori. Indipendentemente dalla presenza o meno di un motore, il risultato è simile. Mentre l'UPS sostiene e mantiene in funzione il carico del server durante un'interruzione di rete, l'impianto di raffreddamento continuerà a non funzionare fino a quando il motore non interviene per ripristinare la potenza critica di tutti gli impianti. Una volta ripristinata l'alimentazione all'impianto di raffreddamento, i raffreddatori o i condensatori continueranno ad accendersi e spegnersi per un periodo che può durare anche 15 minuti. Per evitare un repentino carico incrementale, debilitante per il motore al punto di provocarne l'arresto, gli impianti vengono avviati in maniera sequenziale, proprio come avviene quando si evita di avviare un veicolo con motore freddo inserendo la quarta marcia. Di fatto, tra il momento in cui viene ripristinata l'alimentazione fornita dal motore e quando l'impianto di raffreddamento riceve il segnale di avvio potrebbero trascorrere anche 10-15 minuti. Questo intervallo serve a evitare cadute di tensione intermittenti e far sì che il motore sia pronto. Ogni singolo raffreddatore o condensatore presente nella struttura continuerà ad accendersi e spegnersi sequenzialmente fino a raggiungere la piena capacità.

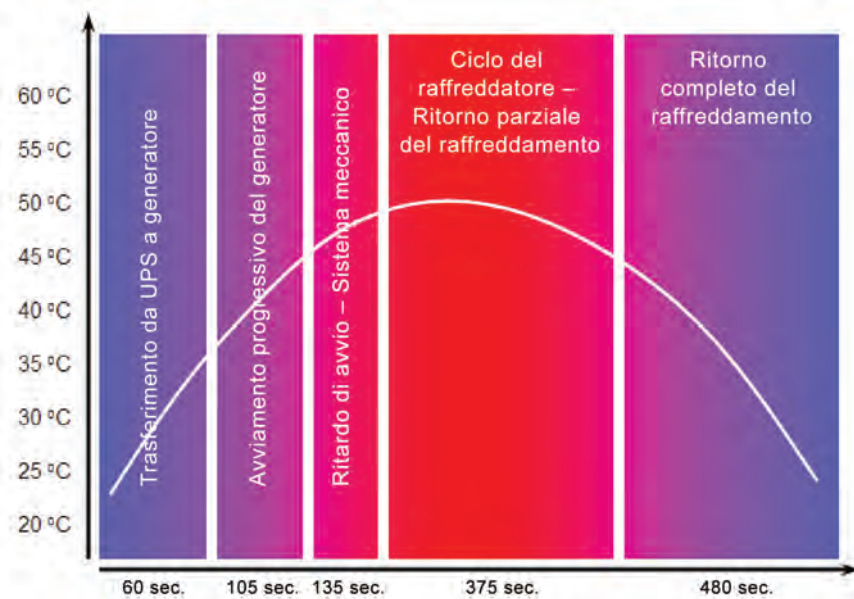


FIGURA 1: DIAGRAMMA SEMPLIFICATO CAUSA-EFFETTO DURANTE UN'INTERUZIONE DI RETE CON L'AVVIAMENTO PROGRAMMATO DEL MOTORE.

Questo processo può impiegare anche 10-15 minuti in funzione del tipo e delle dimensioni dell'impianto. Tutto ciò mentre i server continuano a essere alimentati e a funzionare senza l'ausilio di raffreddamento. Nelle circostanze descritte, l'armadio di un server blade da 15 kW avrebbe innescato l'autospegnimento trascorsi appena 60 secondi.

ARRESTO AUTOPROTETTIVO DEL SERVER

La necessità di raffreddamento è tale che ormai non si parla più di ore o minuti di funzionamento in sua assenza, ma di secondi. Un server tipico montato in rack è caratterizzato da un impianto di dissipazione di calore integrato. L'impianto garantisce che il server non subisca danni permanenti in caso di un repentino aumento della temperatura.

La temperatura della presa d'aria consigliata dal Comitato tecnico ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers, società americana dei tecnici di riscaldamento, refrigerazione e condizionamento dell'aria) 9.9 è compresa tra 20 e 25 gradi Celsius (da 68 a 77 gradi Fahrenheit), valore approvato e sostenuto anche dai produttori di server e dalle relative garanzie.



Specifiche ambientali per le apparecchiature Datacom

Classe	Temperatura consentita	Temperatura consigliata	Umidità relativa consentita	Umidità relativa consigliata	Tasso di variazione max
1	da 15 a 32 °C da 59 a 90 °F	da 20 a 25 °C da 68 a 77 °F	20-80%	40-55%	5 °C 10 °F

FIGURA 2: SPECIFICHE PER GLI AMBIENTI CON APPARECCHIATURE NEI DATA CENTER (ASHRAE TECHNICAL COMMITTEE 9.9).

Nell'arco di 60 minuti non devono verificarsi escursioni termiche superiori ai 5 gradi Celsius. Le escursioni termiche e la costante esposizione a temperature esterne a questa gamma hanno effetto sulla vita utile dell'apparecchiatura IT. La ricerca condotta da Uptime Institute ha dimostrato infatti che i tassi di guasto superano di quattro volte la normale esperienza sul campo. La serie di soglie di temperatura secondo cui è previsto un intervento variano da un produttore di server all'altro. Tuttavia, il primo livello di avviso predefinito viene emesso all'utente quando vengono raggiunti i 55 gradi Celsius. La seconda soglia, quella critica, è posta a 65 gradi Celsius. Il server segnalerà nuovamente il problema all'utente attraverso il sistema operativo; se non si interviene e il livello della temperatura supera i 65 gradi Celsius per 20 secondi, il sistema operativo e l'alimentatore emettono il comando di arresto automatico, che può comportare la perdita di dati per gli utenti. È importante sottolineare che la temperatura misurata all'interno delle apparecchiature server è in genere sensibilmente superiore alla temperatura ambiente.

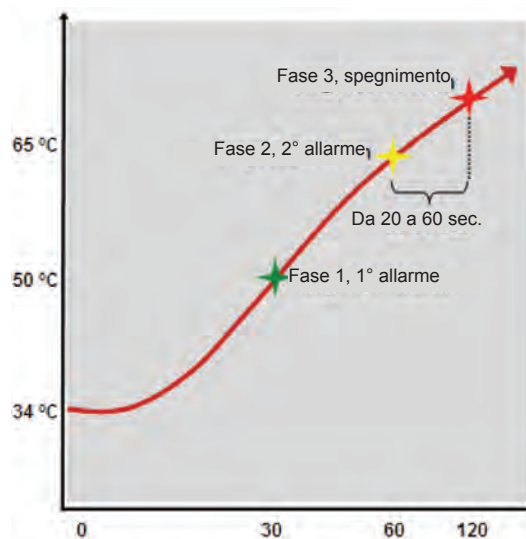


FIGURA 3: RAPPRESENTAZIONE GRAFICA SEMPLIFICATA DELL'ARRESTO TERMICO AUTOPROTETTIVO.

La raccolta di dati da server attivi ha dimostrato che il divario tra la temperatura interna del server e quella ambiente è di 30-35 gradi Celsius, anche se l'accumulo di polveri pesanti e residui dentro e intorno al server può incrementare ulteriormente questo divario. Ciò equivale a dire che, in un data center progettato adeguatamente, una temperatura di 65 gradi Celsius misurata all'interno di un server di regola corrisponde a 35-40 gradi Celsius nell'ambiente esterno.

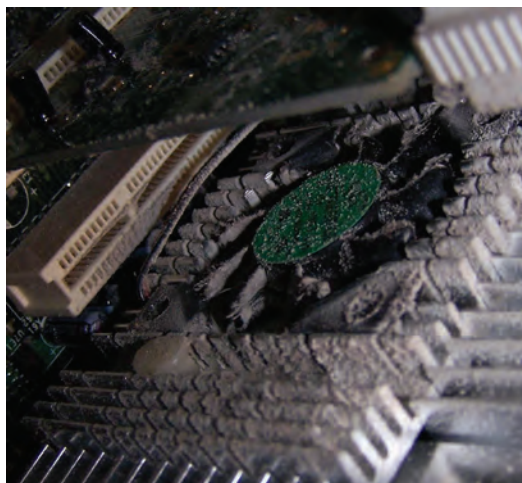


FIGURA 4: ESEMPIO DI ACCUMULO SIGNIFICATIVO DI POLVERE NEL DISSIPATORE DI CALORE.

In genere, la soglia della temperatura nelle apparecchiature di archiviazione è inferiore per la presenza di parti mobili e dischi rotanti. Nel caso di dispositivi di archiviazione, una ricerca condotta da Hitachi Data Systems ha dimostrato che il degrado prestazionale dell'unità rappresenta una sfida a più lungo termine. La ricerca ha infatti dimostrato che un'escursione termica di oltre 5 gradi Celsius provoca l'eccessivo degassamento dei lubrificanti nel motore dell'albero, con un aumento del 15 per cento dei guasti nell'arco dei 30 giorni successivi. Secondo numerosi standard, i dispositivi di archiviazione sono più importanti dei server, dato il loro alto tasso di memorizzazione dei dati.

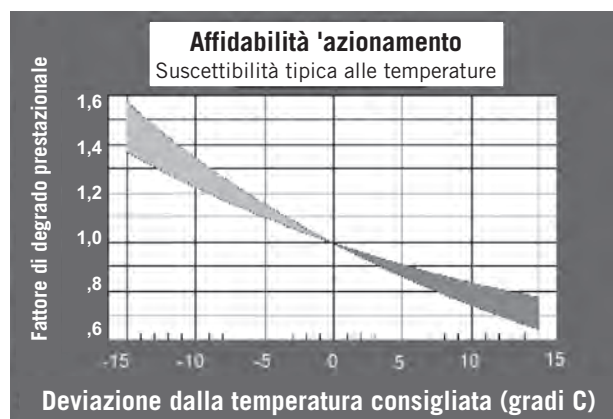


FIGURA 5: RICERCA SULL'AFFIDABILITÀ DELLE UNITÀ (HITACHI DATA SYSTEMS).



FIGURA 6: ESEMPIO DI DISCO RIGIDO.

D'altro canto, la temperatura di soglia per le apparecchiature di rete è molto più elevata delle apparecchiature server, considerata l'eterogeneità delle installazioni in sedi remote non presidiate, armadi e data center condizionati.

AUMENTO DELLE TEMPERATURE NEL CORRIDOIO FREDDO

Vale la pena di sottolineare che vari studi svolti su ambienti ottimali per data center suggeriscono la possibilità di un aumento a 30-35 gradi Celsius della soglia di 20-25 gradi Celsius consigliata dal the ASHRAE TC 9.9. A motivare una tale scelta ci sarebbe il risparmio energetico generato dall'impianto di aria condizionata, il cui dispendio di energia per rimuovere il calore dal data center subirebbe una riduzione. La maggior parte delle apparecchiature server è studiata per funzionare a temperature comprese tra 10 e 35 gradi Celsius, tuttavia nessun OEM di server ha omologato una temperatura di funzionamento generale per tutti i data center esterna all'attuale gamma. I dati raccolti nella presente relazione supportano la gamma di temperature comprese tra 20 e 25 gradi Celsius consigliata dall'ASHRAE TC 9.9 e rivelano che qualsiasi incremento oltre detta gamma ridurrebbe significativamente il tempo di reazione durante un evento di mancanza di alimentazione e successiva perdita di raffreddamento.

“Quando la temperatura aumenta, con essa aumenta anche la velocità di reazione di un elemento chimico. Il ferro si arrugginisce e il cibo si guasta molto più rapidamente al caldo che non al freddo. Un risultato di tale similitudine è l'utile generalizzazione che per molte reazioni che si verificano in prossimità della temperatura ambiente, un aumento di 10 gradi Celsius all'incirca raddoppia la velocità di reazione.”

Svante Arrhenius, vincitore del Premio Nobel del 1903 per la chimica e fondatore della chimica fisica.

FUGA TERMICA

Sono stati condotti numerosi studi dal vivo sull'arresto termico nei data center durante le perdite di raffreddamento. In questo ambito, alcuni fattori sembrano contribuire più di altri durante l'arresto intero o parziale di un data center; tra questi è possibile citare:

- Dimensione volumetrica dell'ambiente. Più alto è il soffitto, maggiore è il calore che sale e si accumula sotto il soffitto o i pannelli, senza disperdersi nei corridoi.
- Carico di lavoro del server. Se il carico di lavoro del server e la relativa produzione di calore sono bassi, il tempo disponibile subisce una marginale estensione.
- Unità ventilanti per il raffreddamento del locale computer alimentate da UPS. Se le unità ventilanti continuano a funzionare, anche se la bobina non riesce a rimuovere il calore, il tempo disponibile subisce un'ulteriore estensione.

Il seguente diagramma rappresenta i nove punti dato raccolti attraverso la ricerca condotta da The Uptime Institute, Emerson Corp., EYP Mission Critical Facilities e Active Power, Inc. con cui si è tracciato un grafico che ha generato la relativa linea di tendenza.

	Con Rack	Secondi	Aumento temp. - °C	Aumento temp./sec. - °C
The Uptime Institute, Menuet, Turner, maggio 2006	600	180	3,5	0,0194
Brevetto US 6.170.561, O'Grady, gennaio 2001	1.050	900	26,6	0,0296
The Uptime Institute, Menuet, Turner, maggio 2006	1.200	180	9	0,0500
EYP, Kosik, ottobre 2006	2.500	90	5,6	0,0622
Active Power, Olsen, maggio 2007	3.500	40	3,9	0,0975
Emerson, ottobre 2005	4.000	360	66,7	0,1853
Active Power, Olsen, maggio 2007	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

FIGURA 7: LA MATRICE RAPPRESENTA I PUNTI DATO RACCOLTI. MOSTRA L'AUMENTO DELLA TEMPERATURA IN GRADI CELSIUS REGISTRATO IN UN DATO PERIODO DI TEMPO PRIMA CHE L'APPARECCHIATURA IT AVVIA L'ARRESTO TERMICO AUTOPROTETTIVO. LA COLONNA ALL'ESTREMA DESTRA MOSTRA L'AUMENTO DELLA TEMPERATURA NORMALIZZATA IN GRADI CELSIUS AL SECONDO PER OGNUNA DELLE DENSITÀ REGISTRATE.

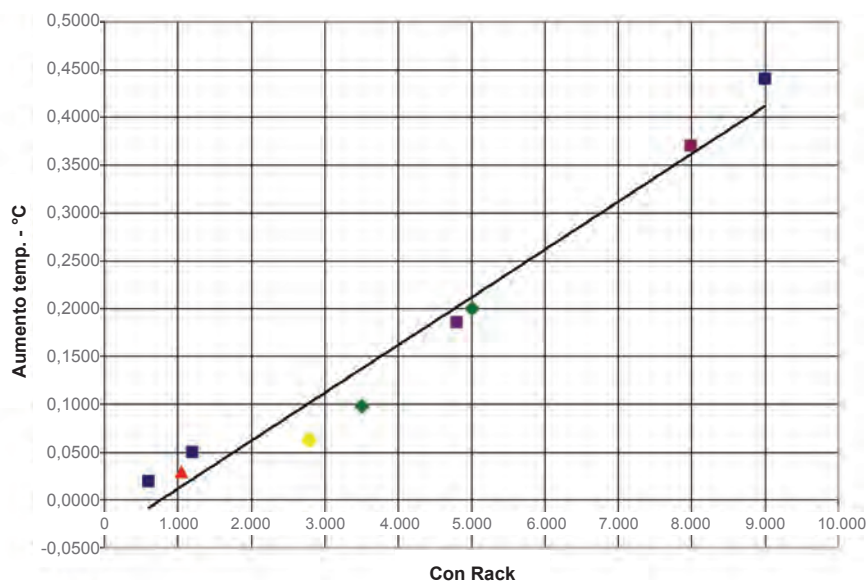


FIGURA 8: IL GRAFICO È LA RAPPRESENTAZIONE VISIVA DELLA MATRICE NELLA FIGURA 7. L'AUMENTO DELLA TEMPERATURA È INDICATO IN GRADI CELSIUS AL SECONDO PER OGNUNA DELLE DENSITÀ REGISTRATE.

La linea di tendenza viene espressa con la seguente equazione: $y = 0,00005x - 0,0381$ e può essere applicata come misura approssimativa a un qualsiasi numero di densità su rack al fine di fornire indicazioni sui tempi necessari al surriscaldamento di un dato carico su server rack e al conseguente arresto per perdita di raffreddamento.

		Secondi								
		30	60	120	240	300	360	420	480	540
Watt per armadio	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	Comando di arresto	
	5000W	26,4	32,7	45,5	70,9	Comando di arresto				
	8000W	30,9	41,7	63,4	Comando di arresto					
	10000W	33,9	47,7	75,4	Comando di arresto					
	15000W	41,4	62,7	Comando di arresto						
	20000W	48,9	Comando di arresto							
	30000W	Comando di arresto								

FIGURA 9: LA MATRICE RAPPRESENTA L'INCREMENTO DELLA TEMPERATURA CON IL PASSARE DEL TEMPO (SECONDI) RELATIVAMENTE ALLA DENSITÀ PER ARMADIO (WATT PER ARMADIO) NEL CASO DI UNA PERDITA DI RAFFREDDAMENTO. LA TEMPERATURA BASE PRESUNTA È 20 GRADI CELSIUS.

MODELLO DI FLUIDODINAMICA COMPUTAZIONALE (CFD)

L'uso del modello di fluidodinamica computazionale è molto diffuso nel settore dei data center per offrire un quadro preciso di come un dato fluido, in questo caso l'aria in un data center, si comporta in determinate circostanze. I colori rappresentano la temperatura, con il blu e il rosso usati per indicare rispettivamente il freddo estremo e il caldo estremo. La figura 10 mostra l'aspetto di un tipico data center progettato adeguatamente in condizioni normali di funzionamento.



FIGURA 10: PROSPETTIVA LATERALE OSSERVANDO L'AREA CENTRALE DEL CORRIDOIO CALDO.

Le figure 11 e 12 mostrano cosa accade in un locale pieno di armadi con un carico di 5.000 watt. L'orario registrato è 120 secondi dopo la perdita di raffreddamento.



FIGURA 11: CARICO DI 5.000 WATT PER ARMADIO TRASCORSI 120 SECONDI DALLA PERDITA DI RAFFREDDAMENTO.

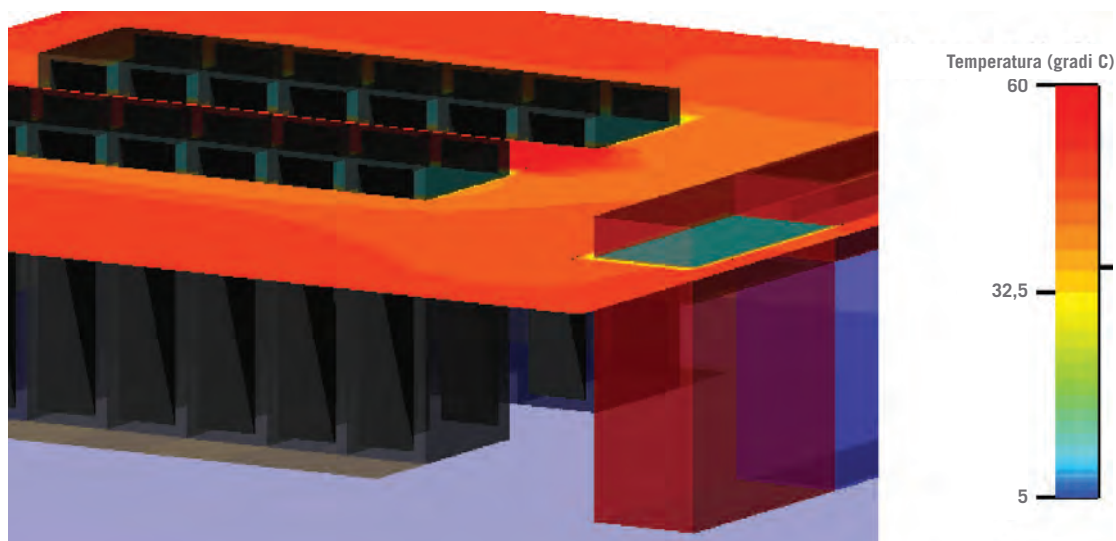


FIGURA 12: CARICO DI 5.000 WATT PER ARMADIO TRASCORSI 120 SECONDI DALLA PERDITA DI RAFFREDDAMENTO.

Le figure 13 e 14 mostrano lo stesso locale con armadi a 5.000 watt trascorsi 240 secondi dalla perdita di raffreddamento e gli effetti devastanti sulla sua temperatura ambiente.



FIGURA 13: CARICO DI 5.000 WATT PER ARMADIO TRASCORSI 240 SECONDI DALLA PERDITA DI RAFFREDDAMENTO.

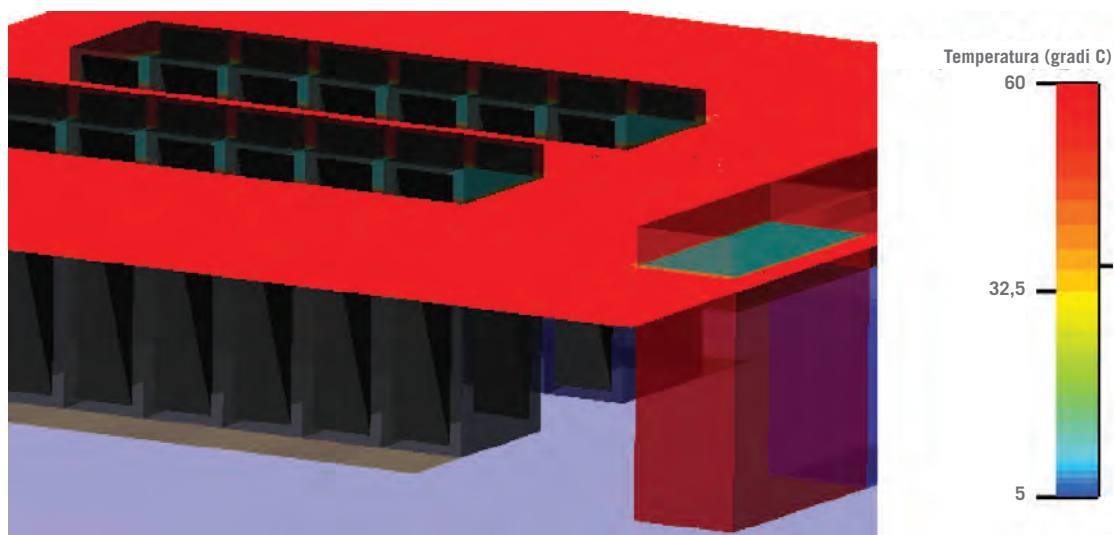


FIGURA 14: CARICO DI 5.000 WATT PER ARMADIO TRASCORSI 240 SECONDI DALLA PERDITA DI RAFFREDDAMENTO.

RIDUZIONE DELL'ESPOSIZIONE ALLE DERIVE TERMICHE

Il modo più efficace per ridurre l'esposizione alle derive termiche in un data center causate dalla mancanza di alimentazione consiste nel ridurre il tempo necessario per ripristinare l'alimentazione dell'impianto di raffreddamento (meccanico). Un metodo è ridurre il tempo necessario per richiedere l'avviamento del motore in standby. Per evitare il noto fenomeno delle false partenze in cui i motori ricevono il segnale di avviamento anche qualora l'alimentazione venga ripristinata pochi secondi dopo l'interruzione, il sistema deve essere configurato appropriatamente. Secondo l'EPRI (Electric Power Research Institute), quasi il 99 per cento di tutte le interruzioni di rete durano meno di 10 secondi, e ciò comporta di dover impostare il ritardo su un valore di 5-10 secondi. Nei gruppi di continuità ad alta disponibilità progettati correttamente, per definizione il motore diesel si avvia e, ricevuto il comando, assume (avviamento progressivo) il carico in un arco di 5-6 secondi.

Il tempo di sincronizzazione delle configurazioni a più motori impiegate per aumentare capacità o ridondanza può incidere, aumentandolo, sul tempo di avviamento progressivo. Tuttavia, l'uso di regolatori computerizzati può contribuire a ridurre e a controllare il tempo di avviamento progressivo, limitandone la fascia di variazione e avvicinandolo maggiormente a quello di configurazioni a motore singolo. Un regolatore computerizzato del motore velocizza i tempi d'avviamento monitorando l'angolazione dell'albero a gomiti durante la fase di riduzione dei giri fino al completo arresto della macchina. Conoscendo l'angolazione dell'albero a gomiti all'avviamento è possibile ottimizzare il controllo di iniezione del carburante al prossimo avvio, riducendo i tempi di sincronizzazione e di avviamento progressivo.



FIGURA 15: ESEMPIO DI UN REGOLATORE COMPUTERIZZATO DEL MOTORE.

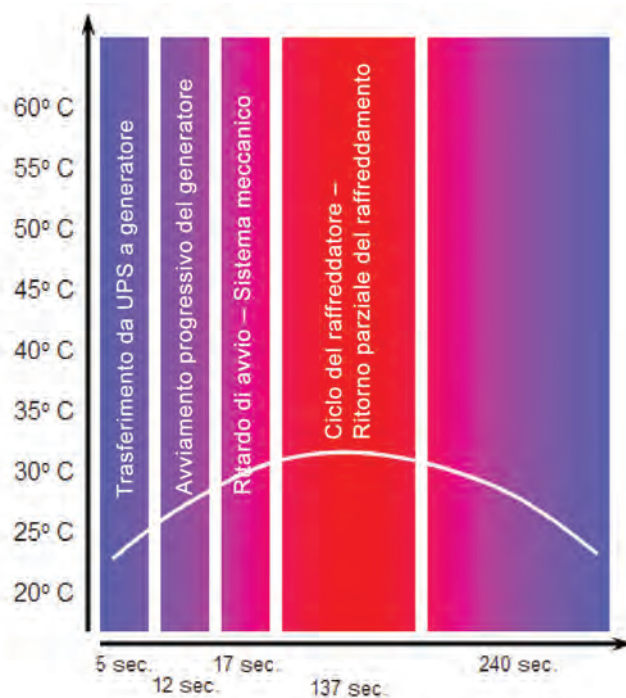


FIGURA 16: DIAGRAMMA SEMPLIFICATO CAUSA-EFFETTO DURANTE UN'INTERRUZIONE DI RETE CON L'AVVIAMENTO RAPIDO DEL MOTORE.

Tempi di 10-15 secondi dal momento della mancanza di alimentazione al ripristino dell'alimentazione al sistema meccanico sono in marcato contrasto con l'intervallo di tempo comunemente specificato di 100-120 secondi dei gruppi di continuità di progettazione attuale. Riducendo questa differenza è possibile diminuire sostanzialmente la temperatura ambiente dei locali e pertanto il rischio di fughe termiche o di arresti autoprotettivi dei server. Ad esempio, utilizzando l'equazione nella figura 9, la temperatura di un armadio da 10.000 watt pieno di server blade aumenterà di 0,4619 gradi Celsius al secondo durante un evento che comporta la perdita di raffreddamento. Se l'intervallo di tempo comunemente specificato di 120 secondi venisse impiegato nelle architetture dei gruppi di continuità, la temperatura raggiungerebbe i 55 gradi Celsius, una soglia molto superiore alla temperatura massima propria dei server. Tuttavia, se si adottasse l'intervallo di tempo di 15 secondi, molto più breve, nella stessa architettura del gruppo di continuità l'aumento della temperatura subirebbe un calo di 6,9 gradi Celsius, con un miglioramento dell'87 per cento sulle architetture comunemente specificate per i gruppi di continuità.

Questo sarebbe senza dubbio un cambiamento semplice da ottenere anche nelle strutture preesistenti; pertanto, come si giustifica a questo punto l'esigenza di 15 minuti di tempo di superamento comunemente consigliato? Con la costante crescita delle densità dei data center, l'utilità di una batteria in grado di concedere 15 minuti sembra essere considerevolmente sopravvalutata e ciò da molteplici punti di vista:

1. Un intervallo di 15 minuti per garantire l'“arresto dolce” di computer e altri carichi è realmente irrilevante, poiché per definizione, trascorso questo arco di tempo l'arresto è già diventato un evento inaccettabile per la maggior parte delle aziende.
2. Un intervallo di 15 minuti per una “seconda messa in moto” del motore in standby è anch'essa un'argomentazione priva di merito, poiché qualora il motore non si avvii nei primi cinque o sei secondi, fatto di per sé rarissimo, esso non si avvierà neanche nei successivi 15 minuti. Per comprendere meglio, è possibile esaminare l'interruzione verificatasi il 24 luglio 2007 presso 365 Main, Inc., un servizio di collocazione di San Francisco. Secondo quanto riportato, i motori non si sono avviati rendendo necessari 37 minuti per il riavvio manuale.
3. Un'autonomia di 15 minuti avrebbe avuto effetti devastanti sulla temperatura del data center. Secondo la figura 9, anche un data center con carico leggero pari a 2.000 watt/armadio avrebbe raggiunto i 55 gradi Celsius. Il 12 novembre 2007 fu riportato che lo sviluppatore e operatore del data center, Rackspace, Inc., aveva subito un'interruzione di rete presso il proprio data center di Dallas/Fort Worth. Durante il riavvio dei raffreddatori, il data center aveva registrato temperature in rapido aumento che costrinsero a disattivare il carico per mitigare la deriva termica.

CONCLUSIONE

In base ai dati provenienti dall'esperienza diretta e dai modelli di laboratorio condotti a supporto del presente documento, è possibile concludere che è preferibile un avviamento rapido del motore entro 10-15 secondi, invece dei 100-120 secondi comuni alla maggior parte delle architetture dei gruppi di continuità odierni, e ciò per contribuire a mitigare la deriva termica nei data center. Inoltre, un controllo più serrato delle fluttuazioni delle temperature nei locali può mitigare notevolmente il degrado prestazionale a lungo termine delle unità di archiviazione rotanti e dei server di calcolo provocato dall'eccessiva espansione e contrazione dei componenti mobili. La pratica diffusa di un'autonomia di 15 minuti è confutata dagli standard di oggi, specie alla luce della completa inutilità di un "arresto dolce" dei computer durante un'interruzione di rete e del tempo necessario per una "seconda messa in moto" dei motori in standby.

RIFERIMENTI

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", giugno 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/>