

Dynamiczne wahania mocy w centrach przetwarzania danych

Jim Spitaels

White Paper 43

APC[®]
Legendary Reliability[®]

Wersja 2

Streszczenie

Zapotrzebowanie na moc w centrach przetwarzania danych zmienia się z minuty na minutę zależnie od obciążenia obliczeniowego. Amplituda tych wahań znacznie wzrasta wraz z wdrażaniem w serwerach i urządzeniach komunikacyjnych technologii zarządzania zasilaniem. Te wahania stanowią źródło nowych problemów związanych z dostępnością i zarządzaniem.

Wstęp

Centra przetwarzania danych pobierają całkowitą moc elektryczną, która jest sumą mocy zużywanej przez zainstalowane urządzenia komputerowe. Dawniej zużycie mocy przez te urządzenia utrzymywało się na mniej więcej stałym poziomie i zależało tylko nieznacznie od obciążenia obliczeniowego lub trybu pracy.

Wraz z komputerami przenośnymi pojawiło się wymaganie zarządzania mocą pobieraną przez procesor w celu wydłużenia czasu pracy akumulatorów. Technologia zarządzania zasilaniem umożliwiła obniżenie poboru mocy procesorów komputerów przenośnych nawet o 90% przy niskim obciążeniu. Wraz z osiągnięciem przez tę technologię odpowiedniego stopnia rozwoju zaczęto wprowadzać ją w serwerach. W rezultacie nowo konstruowane serwery mogą charakteryzować się poborem mocy o znacznych wahaniami zależnych od obciążenia pracą w czasie.

Czasowe wahania poboru mocy powodują pojawienie się nowych problemów w zakresie projektowania i zarządzania centrami przetwarzania danych. Jeszcze kilka lat temu problem ten był marginalny. Obecnie znaczenie tego problemu osiągnęło poziom, przy którym nie można go pominąć, a skala tego problemu wzrasta.

Wahania poboru mocy mogą wywoływać nieplanowane i niepożądane skutki w centrach przetwarzania danych, takie jak wyzwolenia automatycznych wyłączników, przegrzanie czy też utrata nadmiarowości w nadmiarowych systemach zasilania. Taka sytuacja stawia przed projektantami i operatorami centrów przetwarzania danych nowe wyzwania.

Amplituda dynamicznych wahań mocy

W latach dziewięćdziesiątych prawie wszystkie serwery charakteryzowały się niemal stałym poborem mocy. Główne przyczyny wahań mocy w serwerach były związane z rozpędzaniem się napędów dyskowych oraz zmianami prędkości wentylatorów z regulacją temperaturową. Obciążenie obliczeniowe nakładane na procesory oraz podsystemy pamięci powodowało pomijalne wahania całkowitego poboru mocy. W przypadku typowych serwerów przeznaczonych dla małych firm lub przedsiębiorstw całkowite wahania mocy miały wartość rzędu 5% i były całkowicie niezależne od obciążenia obliczeniowego.

Znaczące obniżenie poboru mocy wymaga współpracy systemu BIOS, układu chipset, procesora i systemu operacyjnego. W takim systemie z funkcją zarządzania zasilaniem, za każdym razem, gdy wykorzystanie procesorów spada poniżej 100%, system operacyjny wprowadza tryb bezczynny, co powoduje przejście procesorów w stan niskiego poboru mocy. Ilość czasu spędzonego w trybie niskiego poboru mocy jest odwrotnie proporcjonalna do obciążenia obliczeniowego systemu (tj. procesor pracujący przy wykorzystaniu 20% mocy obliczeniowej będzie przez 80% czasu znajdował się w stanie niskiego poboru mocy).

Rozwiązania stosowane w celu osiągnięcia stanów niskiego poboru mocy różnią się pomiędzy producentami i rodzinami procesorów, niemniej jednak te najczęściej spotykane polegają na obniżaniu częstotliwości lub zatrzymaniu zegara taktującego oraz wyłączeniu lub obniżeniu napięcia zasilającego różne elementy procesora, układów chipset i pamięci.

Ostatnio producenci procesorów wprowadzili techniki oszczędzania energii podczas czynnej pracy procesora. Metody te polegają na dostosowywaniu częstotliwości taktowania oraz poziomu napięcia zasilającego procesora do obciążenia procesora w trybie innym niż bezczynność.

Warto zaznaczyć, że każde rozwiązanie, które warunkowo obniża moc procesora, redukuje jedynie średnią moc pobieraną przez system — moc maksymalna pozostaje na niezmiennym poziomie i wzrasta wraz z pojawianiem się coraz nowszych generacji procesorów. Należy również zdawać sobie sprawę, że w przypadku, gdy udział mocy procesora w całkowitym poborze mocy serwera wzrasta, wahania całkowitego poboru mocy przez serwer powodowane obciążeniem obliczeniowym stają się procentowo odpowiednio większe. Serwery wieloprocessorowe oraz serwery o niewielkiej liczbie napędów dyskowych (np. serwery kasetowe) będą charakteryzowały się największymi procentowymi dynamicznymi wahaniami mocy.

W **tabeli 1** zebrano rzeczywiste dane zmierzone dla niektórych serwerów. Przedstawia ona wahania mocy prądu zmiennego zmierzone przy różnych obciążeniach obliczeniowych komputera.

Tabela 1 — *Dynamiczne wahania mocy rzeczywistych serwerów*

Platforma	Procesor	Pobór mocy przy niskim obciążeniu	Pobór mocy przy wysokim obciążeniu	Zmienność procentowa
Dell PowerEdge 1150	Dwa procesory Pentium III — 1000	110 W	160 W	45%
Intel Whitebox	Pentium 4 — 2000	69 W	142 W	106%
IBM BladeCenter HS20 Pełna obudowa — 14 kaset	Dwa procesory Xeon 3,4 GHz	2,16 kW	4,05 kW	88%
HP BladeSystem BL20pG2 Pełna obudowa — 8 kaset	Dwa procesory Xeon 3,06 GHz	1,55 kW	2,77 kW	79%

Problemy związane z dynamicznymi wahaniami mocy

Dynamiczne wahania mocy stanowią przyczynę następujących nowych problemów:

Przeciążenie obwodów zasilania

Większość serwerów przez długi czas pracuje przy niskich obciążeniach obliczeniowych. W przypadku serwerów z funkcją zarządzania zasilaniem oznacza to pobór mocy poniżej potencjalnego. Większość osób instalujących lub obsługujących centra przetwarzania danych nie zdaje sobie jednak sprawy, że typowy obserwowany pobór mocy serwerów może być znacznie niższy od potencjalnego poboru mocy w warunkach

wysokiego obciążenia obliczeniowego. Taka sytuacja może prowadzić do przypadkowego podłączenia przez operatora lub personel IT danego centrum przetwarzania danych nadmiernej liczby serwerów do jednego obwodu zasilania.

W przypadku, gdy suma maksymalnych poborów mocy serwerów podłączonych do jednego obwodu zasilania przekracza parametry znamionowe obwodu, pojawia się możliwość wystąpienia przeciążenia. Taka grupa serwerów będzie działać prawidłowo aż do momentu, w którym odpowiednia liczba serwerów będzie jednocześnie podlegać silnemu obciążeniu. Warunki obliczeniowe, które prowadzą do takiego przeciążenia mogą występować bardzo sporadycznie, przez co system może działać bezawaryjnie tygodniami, a nawet miesiącami.

W warunkach przeciążenia spowodowanego opisaną powyżej sytuacją w obwodzie zasilania będzie płynął prąd o natężeniu przekraczającym parametry znamionowe obwodu. W środowisku centrum przetwarzania danych najpoważniejszym następstwem takiej sytuacji będzie zadziałanie automatycznego wyłącznika obwodu i przerwanie zasilania sprzętu komputerowego. Jest to oczywiście wyjątkowo niepożądane zdarzenie. Co więcej, ponieważ występuje ono podczas wysokiego obciążenia obliczeniowego, jest bardzo prawdopodobne, że urządzenia komputerowe będą w tym czasie obsługiwać znaczną liczbę transakcji, a więc taka awaria będzie najprawdopodobniej miała miejsce w szczególnie niepożądanym momencie.

Przegrzanie

W centrum przetwarzania danych cała energia elektryczna pobierana przez sprzęt komputerowy jest rozpraszana w postaci ciepła (wyjątkiem są przełączniki PoE, które przesyłają znaczą część swojej mocy za pośrednictwem kabli sieci Ethernet do telefonów VOIP, punktów dostępowych WiFi oraz innych zasilanych urządzeń). Wraz z wahaniami poboru mocy urządzeń komputerowych spowodowanymi zmianami obciążenia obliczeniowego zmienia się również ilość wytwarzanego ciepła. Jeśli pobór mocy urządzeń znajdujących się w jednej części centrum przetwarzania danych nagle wzrośnie, może powstać lokalne miejsce o podwyższonej temperaturze. W przypadku, gdy system chłodzenia centrum przetwarzania danych został zbilansowany z wykorzystaniem typowej wartości rozpraszania energii, podwojenie mocy na lokalnym obszarze może prowadzić do niepożądanego wzrostu temperatury, którego system chłodzenia nie będzie w stanie zniwelować. Taka sytuacja może spowodować wyłączenie urządzeń z powodu zbyt wysokiej temperatury, nieprawidłowe działanie urządzeń lub utratę gwarancji na urządzenia.

Utrata nadmiarowości

Wiele serwerów jest wyposażonych w podwójne, nadmiarowe wejścia zasilania. Cecha ta jest wykorzystywana w większości centrów przetwarzania danych o wysokiej dostępności i do serwerów doprowadzane są podwójne tory zasilania. Takie systemy mogą przetrwać całkowitą awarię w dowolnym miejscu jednego z torów zasilania i kontynuować pracę. Komputery te są zaprojektowane w taki sposób, aby podczas normalnej pracy oba tory zasilania były równomiernie obciążone.

W przypadku awarii w jednym z torów zasilania pełne obciążenie serwera zostaje przeniesione na pozostałe źródło zasilania. Wiąże się to z podwojeniem obciążenia tego toru zasilania. Z tego powodu obciążenie obwodów zasilających urządzenia w systemie dwutorowym nie może nigdy przekraczać 50% znamionowej obciążalności prądowej, aby miały one możliwość przejęcia w razie potrzeby pełnego obciążenia.

Zapewnienie obciążenia obwodu zasilania poniżej 50% jego wartości znamionowych jest zadaniem trudniejszym w przypadku, gdy zasilane urządzenia charakteryzują się dynamicznym poborem mocy. Możliwa jest sytuacja, w której system podczas instalacji został poddany testom, w wyniku których stwierdzono, że obwody zasilania pracują bezpiecznie poniżej 50% ich parametrów znamionowych, ale w przyszłości na skutek wysokiego zapotrzebowania obliczeniowego obciążenie obwodów może przekroczyć 50%.

W przypadku, gdy obwód zasilania w układzie dwutorowym osiągnie stan, w którym obciążenie przekracza 50% jego obciążalności, nadmiarowość systemu zostaje utracona. W razie awarii jednego toru zasilania drugi zostanie natychmiast przeciążony i prawdopodobnie nastąpi zadziałanie jego wyłącznika, co opisano w poprzednim punkcie. I znów, ponieważ zdarza się to podczas wysokiego obciążenia obliczeniowego, jest bardzo prawdopodobne, że urządzenia komputerowe będą w tym czasie obsługiwać znaczną liczbę transakcji, a więc utrata nadmiarowości będzie najprawdopodobniej miała miejsce w szczególnie niepożądanym momencie.

Zamaskowanie problemu

Urządzenia, które charakteryzują się dynamicznym poborem mocy, mogą stanowić jedynie niewielki ułamek całkowitego poboru mocy centrum przetwarzania danych. Jeśli przykładowo 5% urządzeń w centrum przetwarzania danych wykazuje dynamiczne wahania mocy rzędu 2 do 1, a reszta urządzeń pobiera stałą moc, wówczas całkowita moc centrum przetwarzania danych mierzona na głównym doprowadzeniu zasilania lub na głównej listwie zasilającej może wykazywać wahania rzędu 2,5%. Taka sytuacja może upewniać operatora, że nie istnieje żaden istotny problem dynamicznych wahań mocy, podczas, gdy w rzeczywistości może występować znaczne ryzyko uruchomienia automatycznego wyłącznika, przegrzania lub utraty nadmiarowości. A zatem istnieje bardzo realna możliwość, że problem będzie występował, a jednak nie zostanie rozpoznany przez doświadczonych operatorów.

Postępowanie z dynamicznymi wahaniami mocy

Aby złagodzić problemy opisane w poprzednich punktach, projektanci i kierownicy centrów przetwarzania danych muszą przystosować się do nowych realiów dynamicznego poboru mocy. Można to osiągnąć za pomocą wielu różnych środków; niektóre z nich opisano poniżej:

Rozdzielenie obwodów zasilania dla każdego serwera

Jeśli każdy serwer jest zasilany z oddzielnego obwodu zasilania, to przeciążenie obwodu nie może mieć miejsca. Jest to prawda, ponieważ konstrukcja każdego serwera gwarantuje prawidłową pracę przy zasilaniu z dedykowanego obwodu zasilania. Rozwiązuje to problem przeciążenia obwodu zasilania oraz utraty

nadmiarowości. Nie rozwiązuje jednak problemów cieplnych; zazwyczaj jednak to nie one stanowią największe ryzyko. Takie rozwiązanie będzie jednak bardzo skomplikowane i kosztowne w sytuacji, gdy mamy do czynienia z niewielkimi serwerami o rozmiarze 1U lub 2U, ponieważ będzie ono wymagać bardzo dużej liczby obwodów zasilania na szafę. W skrajnym przypadku szafa wypełniona serwerami 1U o podwójnym zasilaniu wymagałaby 84 obwodów zasilania, co odpowiada dwóm dużym tablicom rozdzielczym wyłączników obwodów. Takie rozwiązanie jest bardziej praktyczne w przypadku dużych serwerów lub serwerów kasetowych.

Ustanowienie norm określających marginesy bezpieczeństwa dla najgorszego przypadku i pomiar zgodności przy instalacji

Większość operatorów centrów przetwarzania danych dysponuje normami określającymi marginesy obciążenia, które są zazwyczaj wyrażone w postaci ułamka wartości znamionowych dla w pełni obciążonego obwodu. Typowo dobierane wartości mieszczą się w zakresie od 60% do 80% parametrów znamionowych obwodu, przy czym poziom 75% uważa się za rozsądny kompromis pomiędzy wydajnością zasilania, kosztami oraz dostępnością. W celu weryfikacji zgodności z normą dokonuje się pomiaru faktycznych obciążeń danego obwodu zasilania. Warto zaznaczyć, że z podejściem tym wiąże się poważny problem w przypadku, gdy system charakteryzuje się dynamicznie zmieniającym się poborem mocy, ponieważ ustalenie obciążenia obliczeniowego w chwili pomiaru może być trudne. W idealnej sytuacji zabezpieczone urządzenia powinny się podczas pomiaru poddać wysokiemu obciążeniu obliczeniowemu w celu zapewnienia zgodności w najgorszym warunkach.

Ustanowienie norm określających marginesy bezpieczeństwa dla najgorszego przypadku i obliczenie zgodności

W innym przypadku prowadzi się szczegółowe spisy urządzeń podłączonych do każdego obwodu oraz maksymalnych opublikowanych lub zmierzonych wartości obciążeń powodowanych przez te urządzenia, które następnie sumuje się w celu sprawdzenia, czy dany obwód nie będzie obciążony. Informacje dotyczące maksymalnych obciążeń dla różnych urządzeń można uzyskać od producentów poszczególnych urządzeń lub za pomocą aplikacji selektorów zasilaczy UPS, takich jak te dostępne pod adresem www.apcc.com. Prowadzenie szczegółowych spisów obwodów zasilania jest praktyką powszechnie stosowaną w centrach przetwarzania danych o wysokiej dostępności. Takie rozwiązanie wymaga jednak od operatora posiadania stale dokładnej wiedzy na temat urządzeń podłączonych do każdego z obwodów zasilania. W przypadku większości pomieszczeń z urządzeniami sieciowymi oraz mniejszych centrów przetwarzania danych kontrola nad użytkownikami nie jest wystarczająca, aby zagwarantować, że dane urządzenie nie zostanie przeniesione, wymienione lub po prostu podłączone do innego gniazda. Z tego powodu takie podejście jest niepraktyczne w wielu instalacjach.

Te marginesy mogą zostać jeszcze bardziej zmniejszone w celu uwzględnienia dynamicznych wzrostów mocy. Przykładowo, specyfikacja marginesu bezpieczeństwa może określać, że mierzone obciążenie obwodu nie może przekraczać 35% wartości znamionowych obwodu, gdy urządzenia są w stanie bezczynności.

Ustanowienie norm określających marginesy bezpieczeństwa dla najgorszego przypadku i bieżące monitorowanie zgodności

W tym przypadku ustanowione zostają marginesy bezpieczeństwa, a następnie wszystkie obwody zasilania są na bieżąco monitorowane przez automatyczny system monitorowania. W przypadku, gdy obciążenie obwodu wykracza poza margines bezpieczeństwa, wysyłane są ostrzeżenia. Przykładowo, w przypadku normy obciążenia obwodu równej 60%, alarmy są wysyłane po przekroczeniu przez obciążenie poziomu 60%. Marginesy bezpieczeństwa są ustalane w taki sposób, aby operatorzy otrzymywali z wyprzedzeniem ostrzeżenia o problematycznych obszarach i mogli podjąć działania naprawcze przed wystąpieniem stanu przeciążenia. Tę metodę można stosować w połączeniu z innymi wyżej opisanymi metodami. Dużą zaletą tej metody jest to, że sprawdza się w sytuacjach, gdy istnieje prawdopodobieństwo, że użytkownicy będą instalować, przenosić lub podłączać urządzenia do innego gniazda bez wiedzy kierownika centrum przetwarzania danych. Z taką sytuacją mamy często do czynienia w pomieszczeniach z urządzeniami sieciowymi, w pomieszczeniach kolokacyjnych oraz w centrach przetwarzania danych o średnim poziomie zabezpieczeń. Takie rozwiązanie umożliwia również ostrzeganie o zbliżającej się utracie nadmiarowości. Jest to najwydajniejsze narzędzie, jakim może dysponować kierownik centrum przetwarzania danych w celu postępowania z dynamicznymi wahaniami mocy w stale zmieniającym się środowisku.

Wniosek

Odsetek urządzeń komputerowych w centrum przetwarzania danych, które charakteryzują się poborem mocy zmieniającym się znacznie wraz z obciążeniem, stale wzrasta. Taka sytuacja stanowi dla operatorów infrastruktury centrum przetwarzania danych przyczynę wielu nieoczekiwanych problemów. Procedury stosowane dawniej w celu minimalizacji ryzyka przeciążenia wymagają dostosowania do nowych realiów. Właściwe planowanie i monitorowanie mocy obwodów zasilania są kluczowe dla zapewnienia dostępności zarówno w nowych, jak i istniejących instalacjach, w których zostanie zainstalowana znaczna liczba serwerów.

O autorze:

Jim Spitaels jest inżynierem-konsultantem w firmie APC. Posiada stopnie licencjata oraz magistra elektrotechniki uzyskane w Worcester Polytechnic Institute. Podczas swojej 14-letniej pracy w firmie APC zajmował się projektowaniem zasilaczy UPS, urządzeń, architektur i protokołów komunikacyjnych, obudów urządzeń oraz urządzeń dystrybucji zasilania, a ponadto kierował wieloma zespołami ds. rozwoju produktów. Jim jest również posiadaczem 3 amerykańskich patentów związanych z zasilaczami UPS i układami zasilania.