

# 数据中心和 网络机房的 动态功率变化

作者: Jim Spitaels

## 43 号白皮书



## 摘要

数据中心和网络机房的功率要求随计算负载的变化而瞬息万变。这种变化的幅度已经在增长，并且随着电源管理技术在服务器和通信设备中的部署会继续急剧增长。这种变化带来了可用性和管理方面的新问题。

## 简介

数据中心和网络机房消耗的总电源功率是已安装的 IT 设备所消耗功率的总和。在以前，这类设备的功耗随计算负载或工作模式的不同，只会发生极小的变化。

为延长笔记本电脑的电池供电时间，需要对处理器的电源消耗进行管理。电源管理技术的应用使笔记本电脑处理器在负载较小的情况下可节电高达 90%。随着此项技术的逐渐成熟，人们已开始将其移植到服务器的设计中。其结果是当新开发的服务器的工作负载随时间发生变化时，其功耗可能会随之发生显著变化。

当功率随时间发生变化时，随之而来的是数据中心和网络机房设计和管理方面的各种新问题。在几年以前，这种问题是可以忽略的。现在，问题已到达不容忽略的程度，并且问题的严重性还在不断加剧。

这种功耗波动会导致在数据中心和网络机房环境中发生意外的不良后果，包括断路器跳闸、过热和冗余电源系统中的冗余功能丧失。这种情况给数据中心和网络机房的设计和操作人员带来了新的挑战。

## 动态功率变化的幅度

在整个 20 世纪 90 年代，几乎所有服务器所消耗的功率都接近于恒定。造成服务器功率变化的主要因素是磁盘驱动器的旋转以及温控风扇的速度变化。处理器和内存子系统上的计算负载所导致的功率变化很小，在总功耗中可以忽略不计。在典型的小型公司或企业服务器中，总功率变化一般在 5% 左右，并且这种变化几乎与计算运行状态无关。

要大幅度降低功耗，需要 BIOS、芯片组、处理器和操作系统之间协调配合。在这样一个电源受到管理的系统中，每当处理器的使用率低于 100% 时，操作系统就会执行空闲线程，使处理器进入低功率状态。处于低功率状态的时间量与系统上的计算负载成反比（例如，当 CPU 的使用率为 20% 时，处理器将有 80% 的时间处于低功率状态）。

不同的供应商和不同种类的处理器采用不同的方法来实现低功率状态。但是，最常见的方法包括减少或停止时钟和减少或停止对处理器、芯片组和内存的各个部件供电。

最近，处理器供应商开始引入可在 CPU 执行任务时节省电力的方法。这些方法包括改变处理器的时钟频率和电压大小，以便更好地匹配处理器在非空闲状态下的工作负载。

需要注意的是，任何在一定条件下减小处理器功耗的方法，所减小的都是系统平均功耗。最大功耗不会改变，并且每一代新 CPU 都有功耗升高的趋势。另外，您还必须认识到，当处理器功耗在服务器总功耗中所占比例较大时，由计算负载造成的服务器总功耗的变化也会相应变大（按百分比计）。因此，具有多处理器的服务器和磁盘驱动器很少的服务器（如，刀片服务器），其动态功率变化百分比最高。

表 1 给出了几种服务器的实际测量值，并显示出在计算机上加载不同的计算负载时，测量到的 AC 电源变化。

表1 — 实际服务器的动态功率变化

| 平台                                        | 处理器                     | 小负载消耗   | 大负载消耗   | 变化百分比 |
|-------------------------------------------|-------------------------|---------|---------|-------|
| Dell PowerEdge 1150                       | Dual Pentium III - 1000 | 110 W   | 160 W   | 45%   |
| Intel Whitebox                            | Pentium 4 - 2000        | 69 W    | 142 W   | 106%  |
| IBM BladeCenter HS20<br>整个机箱 — 14 个刀片服务器  | Dual Xeon 3.4 GHz       | 2.16 kW | 4.05 kW | 88%   |
| HP BladeSystem BL20pG2<br>整个机箱 — 8 个刀片服务器 | Dual Xeon 3.06 GHz      | 1.55 kW | 2.77 kW | 79%   |

## 与动态功率变化相关的问题

动态功率变化导致以下新问题的产生：

### 分支电路过载

在多数时间下，大部分服务器都在小计算负载下运行。对于具有电源管理功能的服务器而言，这意味着服务器将消耗少于潜在功耗的功率。但是，大多数数据中心和网络机房的安装或维护人员并没有意识到其通常观察到的服务器功耗可能远小于高计算负载下的潜在功耗。这种情况可能会导致数据中心或网络机房的操作人员或 IT 工作人员无意地将过多的服务器连接到分支电路中。

当分支电路中服务器的最大功耗总和超过分支电路的额定值时，就有可能发生过载。在这种情况下，这些服务器将会正常运行，直到条件发生变化，即足够多的服务器同时在大负载下运行。导致此类过载的计算条件很少发生，因此系统可能会连续数周甚至数月无故障地正常运转。

在由于上述情况而导致的过载条件发生期间，分支电路将在高于电路额定值的电流下工作。在数据中心或网络机房环境中，此情况造成的最严重后果是分支电路断路器可能跳闸并中断对计算设备的供电。毫无疑问，这是我们最不希望发生的情况。此外，由于这种供电中断是发生在计算负载很高的时间段，因此计算设备有可能正在处理大量事务，这意味着故障很有可能发生在最不希望发生的时间点上。

## 过热

在数据中心和网络机房中，计算设备所消耗的所有电能都会变成热量散发出来（PoE 交换机是个例外，它将大部分功率通过以太网电缆传输到 VOIP 电话、Wi-Fi 接入点和其他受电设备）。当计算设备的功耗因计算负载而变化时，其热量输出也会变化。如果数据中心某一处设备的功耗突然增加，就会在数据中心出现局部热点。数据中心制冷系统的制冷能力可能已根据典型功率耗散情况进行了分配，因此局部区域的功率翻倍可能导致不期望的温度上升，而这种温度上升在设计制冷系统时并未考虑。这可能导致设备在过热时关机，工作异常或者使设备的保修失效。

## 冗余丧失

很多服务器都具有双冗余电源输入，而具有最高可用性的数据中心和网络机房正是利用这一特性为服务器提供双路供电。当其中一条供电线路上的任何一点彻底无法供电时，这些系统仍然可获得电能并继续运行。在正常运行条件下，计算机设计为让两条供电线路平均分担负载。

当其中一条供电线路发生故障时，服务器的全部负载就会转移到剩下的另一条供电线路上。这会导致供电线路上的负载翻倍。因此，对于为双路供电系统中的设备供电的 AC 电源分支电路而言，其负载必须小于额定载流量的 50%。这样，它才有足够的剩余能力在必要时承担全部负载。

当负载呈现出动态变化的功耗时，确保分支电路的负载小于其额定值的 50% 变得更加困难。系统在安装时已经过测试并且分支电路确实是在其额定值的 50% 之下运行。但是，当未来某个时刻出现高计算负载时，系统就有可能开始在大于额定值的 50% 的条件下运行。

如果双路供电系统的分支电路出现负载超过自身能力的 50% 的情况，那么系统就丧失了冗余功能。如果一条供电线路发生故障，第二条供电线路会立即过载并可能发生上一节所述的断路器跳闸情况。同样，由于这种供电中断是发生在计算负载很高的时间段，因此计算设备有可能正在处理大量事务，这意味着冗余功能丧失很有可能发生在最不希望发生的时间点上。

## 掩盖问题

具有动态功耗的设备可能只占数据中心或网络机房总功耗的很小部分。如果数据中心的 5% 设备具有 2 比 1 的动态功率变化，其他设备消耗恒定的功率，那么在主供电线路或配电单元处测量到的数据中心总功率只会变化 2.5%。这种测量结果会误导操作人员，使其确信不会有重大的动态功率变化问题发生。实际上，发生断路器跳闸、过热或冗余功能丧失的风险是非常大的。因此，很有可能问题已存在但经验丰富的操作人员并未察觉。

# 管理动态功率变化

为减少上节所述问题的发生，数据中心和网络机房的设计、管理人员必须针对动态功耗这一新情况采取必要的应对措施。解决这一问题的方法有很多，下面评述了其中的几种：

## 为每台服务器提供专用分支电路

如果每台服务器都有独立的分支电路，分支电路过载就不会发生。这是因为设计方案假定每一台服务器都通过专用分支电路供电运行。此方法解决了分支电路过载问题和冗余功能丧失问题。虽然它没有解决过热问题，但问题通常不是最大的风险所在。但是，对于部署了小型服务器（如 1U 或 2U 服务器）的工作环境，由于每个机架所需分支电路的数量极其巨大，因此这是一个非常复杂而又昂贵的解决方案。在极端的情况下，装满了双线 1U 服务器的机架可能需要 84 个分支电路，这意味着需要两个大型电路断路器配线板。当使用较大型的服务器或刀片服务器时，此解决方案更加实用。

## 建立最坏情况下的安全容限标准并在安装时测量兼容性

多数数据中心和网络机房操作人员都有负载容限标准，通常以占全部负载分支电路额定值的百分数表示。所选择的典型值一般在分支电路额定值的 60% 到 80% 之间，75% 被认为是兼顾电源容量、成本和可用性的最理想折衷值。要确认与标准的兼容性，应对实际分支电路进行测量并确保其符合标准。注意，当系统具有动态变化的功耗时，由于在测量时很难知道计算负载的情况，使用此方法会面临严峻的问题。最好的方法是在测量时将大计算负载加在有保护的设备上，以确保在最坏的情况下也能符合标准。

## 建立最坏情况下的安全容限标准并计算兼容性

在另一种情况中，保留一份连接到每个分支电路的设备的详细清单，以及公布的或测量出的设备最大负载数据，然后将负载相加以确保特定分支电路没有过载。各设备的最大负载信息可从其设备制造商处获得（通常会将负载夸大），或从 UPS 选择器应用程序获得。这类应用程序可在 [www.apcc.com](http://www.apcc.com) 网站上找到。保留详细的分支电路清单是大型高可用性数据中心的工作惯例。但是，这需要操作员能做到时刻准确掌握每个分支电路中所插入的设备。对于大多数网络机房和较小的数据中心，由于对用户没有足够的控制，因此不能确保用户不会移动设备、交换设备的位置，或将设备插入其他的插座中。因此，对于很多安装场所，此方法是不切实际的。

可以将这些容限值进一步降低以便为动态功率的增加做好准备。例如，安全容限规格可以定为当设备在空闲状态下运行时，测量到的分支电路负载不能超过分支电路额定值的 35%。

## 建立最坏情况下的安全容限标准并实时监控兼容情况

在此情况中，建立安全容限并通过自动监控系统对所有分支电路进行实时连续监控。当分支电路负载开始进入安全容限区域时，发出警告。例如，如果使用 60% 的分支电路负载标准，则当负载超过 60% 时就会发出警告。所建立的安全容限应该确保操作人员能够提前获得有关问题区域的警告，并且在电流过载情况发生前有足够的时间来采取纠正措施。此方法可与前面所述的其他方法配合使用。它的最大优点在于其适用于用户可能在数据中心管理员不知情的情况下安装、移动设备或将设备插入其他插座的情况。这种情况在网络机房、配置室和中等安全的数据中心时有发生。此方法还可对即将发生的冗余功能丧失发出警告。它是数据中心管理员管理多变环境中的动态功率变化的最有力工具。

## 结论

随着时间的推移，功耗随负载变化而显著变化的 IT 负载在网络机房或数据中心所占的百分比正在不断增加。这种情况给数据中心基础设施操作人员带了许多不曾预料到的问题。因此，需要对以前用于减小过载风险的措施加以改进以便适应新的情况。对于要在其中安装大量服务器的新设施和现有设施而言，正确的规划和监控分支电路功率是确保其可用性的关键所在。

### 关于作者：

**吉姆·斯柏塔斯 (Jim Spitaels)** 是 APC 的咨询工程师，曾获得 Worcester Polytechnic Institute 的电子工程学士学位和硕士学位。作者在 APC 的 14 年工作生涯中，曾参与开发 UPS、通信产品、架构和协议、设备机柜和配电产品，并管理着多个产品开发团队。作者还拥有 3 个与 UPS 和电源系统有关的美国专利。