

SOFTWARE DEFINED POWER

The Cost-Effective Way to Increase Data Center Capacity

By: Mark Adams



CUI INC[®]

Global compute capability is changing: small, medium and large enterprises are shifting workloads off their own infrastructure and into the cloud, attracted by the opex cost model, flexibility and near-limitless room to grow. Consumers, too, are making increased use of the cloud (sometimes without realizing it), storing everything from their email to their documents, photos and health data on shared infrastructure.

THE CLOUD POWER CHALLENGE

The 'cloud' is effectively a network of vast data centers. And with growing worldwide demand, these facilities are being pushed to their limits. Many operators are continually juggling resources to ensure they can provide for every customer's needs. In many cases, the scarcest resource is not server or storage capacity, but power.

When you think about it, this makes perfect sense: servers and storage are commodity items that can be purchased relatively easily. But increasing the amount of power in a data center can involve complex, expensive and time-consuming infrastructure upgrades.

This is why data center operators are looking to optimize the way they use power. More-efficient cooling and humidity-control systems are appearing, using outdoor air and rainwater, rather than traditional air conditioning. Significant efforts have also gone into making modern servers more efficient and reducing their idle power consumption. As a result, large data centers are now able to achieve power usage effectiveness (PUE) of less than 1.2.

In the specific case of 2N data centers, there is a new way to get more from the existing power topology, thereby reliably creating headroom for additional racks (and hence revenue). To understand it, let us look briefly at 2N redundancy.

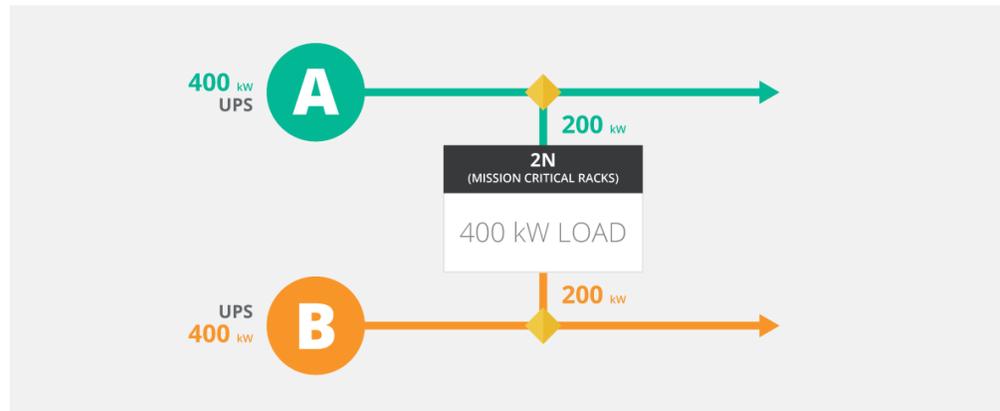
2N REDUNDANCY IN DATA CENTERS

A traditional 2N data center utilizes a pair of uninterruptible power supplies (UPSes) (Figure 1). Each UPS must be capable of powering all the workloads in the data center on its own. However, for much of the time, both UPSes will be operational, meaning neither runs at more than 50% capacity. This leaves up to 50% of the remaining capacity on both UPSes for redundancy purposes. With traditional power architectures, this redundant power is not available for other uses, and is only called upon very rarely – in emergencies or during planned maintenance.

2N redundancy also assumes that everything in the data center is mission-critical, and that it is mission-critical 24/7. In reality, this usually is not the case: test, development and other non-production environments, for example, do not generally require high-availability, or even

to be running at all times. They therefore may not need 2N power redundancy. Similarly, production systems might only require high-availability at certain times. Providing full 2N power redundancy all the time and across the board, as part of a one-size-fits-all approach, prevents significant portions of the data center's power from being used elsewhere.

Figure 1:
A traditional 2N
data center leaves
up to 50% capacity
on UPSes for
redundancy



Consequently, even if there is the physical space to install additional servers, a data center may not be able to power them.

SMARTER POWER-MANAGEMENT

This is where a pioneering technique called Software Defined Power® (SDP) comes in. SDP brings together intelligent software with specialized power-control hardware, turning power into a resource that can be pooled dynamically across the whole data center – using peak shaving and dynamic redundancy to unlock greater value from the existing architecture.

For data center operators, this is incredibly significant. It means they can reliably tap into power that was previously locked away for redundancy purposes, thereby creating headroom for additional, non-critical workloads. Crucially, they can do this without compromising the availability of mission-critical 2N workloads, even when one UPS is unavailable.

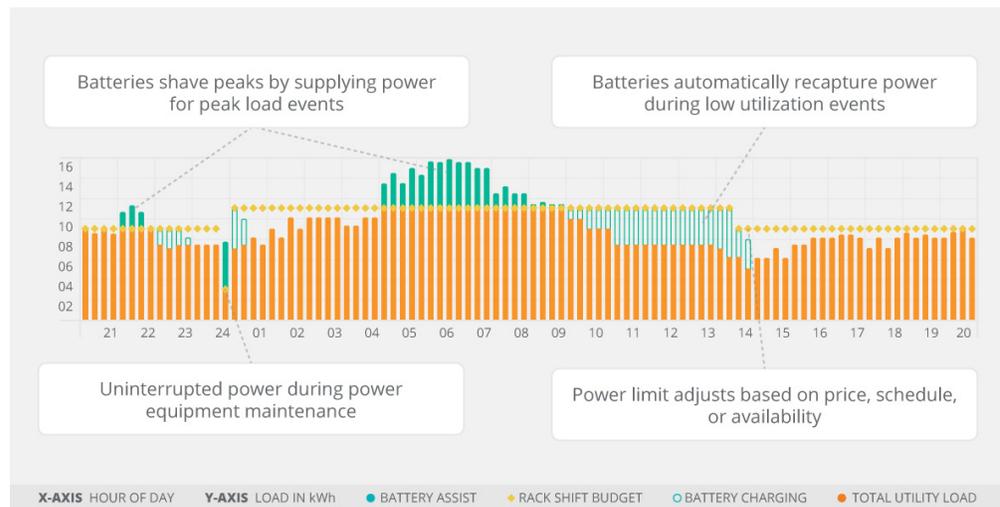
IN A NUTSHELL: HOW SOFTWARE DEFINED POWER WORKS

The SDP software collects data from the power-control hardware in each rack every second. It processes this data using predictive analytics and machine learning and, with its holistic view of the data center's overall power requirements, sends out device-specific power policies for each control unit.

These policies are sent to the control hardware every 10 seconds and contain instructions on what to do if one of the UPSes becomes unavailable. Should this happen, the power-control hardware will automatically take action to ensure the 2N racks remain operational, and the non-critical ones are shut down. This shutdown can either be immediate, or following a pre-defined hold-up period, to enable the workloads to be closed or migrated correctly.

Let us look at how this works in practice, by exploring both peak shaving and dynamic redundancy.

Figure 2:
Peak shaving
using the ICE Block
solution from CUI



PEAK SHAVING

Peak shaving works by charging batteries during times of low power usage, then drawing on them to power peak loads. In this context, peak shaving can protect the second UPS when the first is out of action, by ensuring the former is never pushed beyond capacity. Peak shaving temporarily provides extra power to the system, either to allow for initial hold-up time, or to cover temporary increases in demand from the 2N racks (Figure 2).

DYNAMIC REDUNDANCY

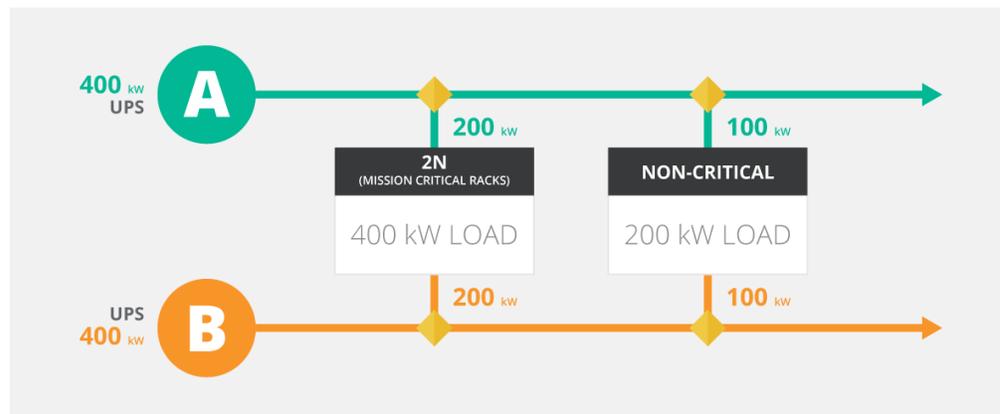
Dynamic redundancy, meanwhile, tackles the assumption that everything in the data center requires high-availability, by differentiating between critical (2N) and non-critical workloads. With dynamic redundancy, when both UPSes are operational, a large portion of the redundant capacity can be made available for non-critical purposes.

In this situation, as soon as one UPS becomes unavailable, the power-control hardware looks at the latest policy it has received from the control software and takes action to ensure the 2N workloads remain active.

Here is a specific example to illustrate this. Imagine a 2N data center with mission-critical racks requiring 400 kW of power under peak loads. This facility has two 400 kW UPSes to provide the required high-availability, but with its traditional power architectures, the data center is at capacity. Even though under normal operation each UPS is only running at up to 200 kW, no further equipment can be added because this would compromise the 2N requirement of the mission-critical racks. However, by using dynamic redundancy, this data center could add racks running non-critical environments, without affecting the availability of the 2N racks.

Say you add 200 kW-worth of non-critical racks, giving a total peak load of 600 kW, or 300 kW per UPS (Figure 3). Under normal circumstances, this is well within the 400 kW limit of each UPS, so the system is completely stable and no special control is needed.

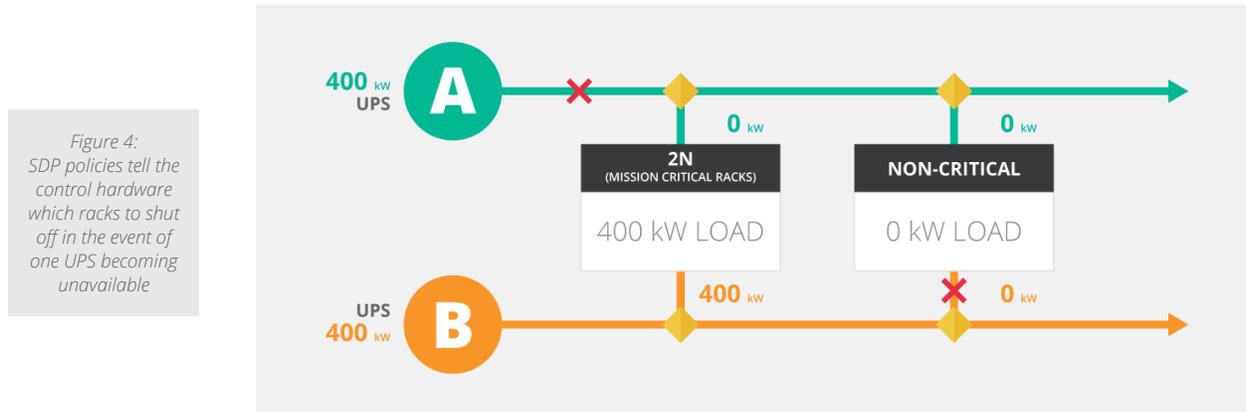
Figure 3:
Dynamic redundancy enables a 2N data center to create capacity for additional non-critical racks



However, if one UPS becomes unavailable, the smart power-control system kicks in, using the policies that have already been distributed. This means what happens next occurs at the speed of the local processors, with no need for each device to query the central power-management software.

The policy tells each device to do one of three things: keep running on the remaining UPS if it is a 2N rack or, if it is a non-critical rack, shut down immediately, or shut down after a defined hold-up period (in which case, it temporarily uses peak shaving, drawing on a battery to protect the remaining UPS).

By immediately shedding the 200 kW of non-critical workloads, the data center's power requirement instantly drops from 600 kW to the 400 kW required by the 2N racks (Figure 4). This is within the capacity of the remaining UPS, meaning it can continue to power the high-availability racks. Peak shaving will remain active, to ensure the draw on the UPS never exceeds its 400 kW capacity.



This example illustrates how software defined power, using peak shaving and dynamic redundancy, can create up to 50% additional power headroom in a data center for non-critical workloads, without the cost and complexity of upgrading the power infrastructure.

THE HARDWARE AND SOFTWARE THAT MAKES SDP POSSIBLE

Making it possible is a pioneering combination of smart, policy-based software from Virtual Power Systems (VPS), and hardware capable of being controlled by the VPS policies from CUI, collectively known as the Intelligent Control of Energy® (ICE) platform. The two hardware solutions from CUI include the ICE Block (which provides the peak shaving capability) and the ICE Switch (for dynamic redundancy). The ICE Switch can slot into a standard 19-inch rack, taking up just 1U. It comprises two identical, hot-swappable modules, each capable of delivering up to 50 A of current, and can be set up in single or three-phase configurations. The ICE Block is a rack-mount, Li-ion battery storage device delivering 12 kW of output power in 4U of space. This modular hardware provides operators with a plug and play way of bringing software defined power to data centers.

UNLEASHING THE BENEFITS

For data center operators, policy-based SDP is a significant development. It creates, for the first time, a truly robust way of adding capacity to a data center, without needing to expand the existing power infrastructure. This opens the door to increased revenue, with considerably less outlay than would be required to upgrade the facility's power infrastructure in a traditional way.

It also provides greater flexibility in the services data centers can offer to their customers. Racks can dynamically be assigned 2N or non-critical status, meaning that as workload priorities change throughout the day, week or year, the data center can adapt. This ultimately ensures maximum use of its power architecture, while offering customers an effective and economical service.

With flexibility being one of the key selling points for the cloud, smart power management techniques like SDP are needed to ensure the infrastructure behind it can deliver on its promise in a cost-effective way.



*Find out more about CUI's Software Defined Power
technology and capabilities*

LEARN MORE ►