Powering Business Worldwide

Thought Leadership White Paper

Powering Converged Infrastructure

By Tatu Valjakka, Software & Connectivity Product Manager
Power Quality Division - EMEA
Eaton

Executive summary

IT environments today are typically characterised by unpredictable data growth accompanied by equally hard to predict performance demands from users. Data centre operators faced with responding rapidly yet cost-effectively to these challenges are finding that legacy silo-based infrastructure no longer provides them with the flexibility that they need at a price that allows profitable operation. As a result, they are increasingly turning to converged infrastructure.

Essentially modular, converged infrastructure provides those who adopt it with pre-integrated "building blocks" from which they can construct their data centres. Expansion is, therefore, reduced to adding one or more of these building blocks, which is a fast, straightforward and relatively inexpensive task.

Converged hardware infrastructure is the perfect complement to virtualisation, with the combination of the two providing exceptional levels of flexibility, economy and, if properly structured and managed, resilience. If converged and virtualised infrastructures are to achieve their full potential, however, careful attention must be given to the way they are powered, bearing in mind that traditional power architectures are unlikely to provide the most suitable solution.

This white paper looks at converged infrastructure and the opportunities it creates for novel approaches to optimising resilience. It then debunks the dangerous delusion that power protection is no longer a crucial issue, before moving on consider the best approaches to powering converged systems and, in particular, looking at UPS requirements in relation to these systems.

Table of contents

What is converged infrastructure?

As mentioned in the introduction, converged infrastructure is essentially a modular approach to data centre design and construction. According to one of the major proponents of the converged approach, converged infrastructure "is a systematic approach that brings all server, storage and networking resources together into pools of resources. It brings together management tools, policies and processes so resources are managed in a holistic integrated manner."

The basic building block of converged infrastructure is, therefore, likely to comprise CPU and memory, storage, network connection devices and, crucially for today's systems, provision for virtualisation. Physically, this basic building block can be realised in several different ways. It could be, for example, a bundled pod, a pre-populated rack or even a container. In many cases, the choices of server hardware, storage array, networking and hypervisor will be fixed by the vendor, which means that the assembly conforms to a reference design that has been fully tested and proved in typical operation scenarios.
Other vendors will offer a choice of components and, while this means moving away from a proven reference design, it does allow the configuration to be precisely tailored to suit specific requirements. Some data centre operators may even want to adopt a do-it-yourself approach to their converged hardware, purchasing the individual hardware and software elements separately and assembling them in-house to produce a fully bespoke basic module.

All of the approaches described are equally valid and the one chosen will depend on user preferences and requirements. Whatever the choice, however, decisions have to be made about the provision of power for the modules and, as always, these decisions will be influenced by the overall size and power consumption of the installation, and by availability/resilience requirements. To take full advantage of converged architecture, however, an intelligent UPS system used in conjunction with an intelligent power distribution system and comprehensive power management software will be needed.

In addition, the power system must be network connected, it must communicate and integrate with the IT management systems and it must link to the virtualisation layer. Ideally, it should also be capable of monitoring and reporting the power used by each server or outlet, and of reacting to external threats like over-temperature or water leaks. We'll return to many of these topics later in the paper but first we need to examine an issue that has enormous implications for the implementation, powering and operation of converged architectures. That issue is resilience.

New approaches to resilience

Traditionally, resilience has been considered an issue relating primarily to the hardware and operating system, and much time and money has been spent on making these as reliable as is realistically and technically possible. The result is a foundation that is strong and stable, providing a platform that makes for straightforward application design and testing.

In a system based on this philosophy, the power infrastructure will be configured to support maximum hardware availability. It will often be built to a very high specification with redundant UPS systems, redundant generator sets, ample fuel reserves, a dual power bus and a generously sized static transfer switch (STS).

This traditional approach to maximising resiliency, which involves providing a strong foundation on which application, operating system and other layers can rely, is perfectly valid. Indeed, for some users, such as those who must have guaranteed availability 24 hours a day, 365 days a year, and cannot tolerate even short-term performance degradation, it remains the best approach. It does, however, have a number of drawbacks.

(The pyramid diagram needs to be included here)

The first of these is a system that relies entirely or almost entirely on the hardware layer to provide resiliency is very static and possibly hard to modify. This makes it a poor match for the world of converged and virtualised computing, which as we've already seen, is very much concerned with responding

dynamically to changing requirements. The second drawback of hardware-centric resilience is cost; to achieve and maintain the required level of hardware reliability is unavoidably expensive.

Fortunately, the adoption of converged and virtualised infrastructures makes it realistically possible to implement resilience in layers other than the physical hardware layer. Let's start by considering the layer that may, at first seem the least likely source of resilience – that is, the user layer. The key question here is whether the IT resource really does have to be available continuously 24/7/365 under all conditions. Or is it possible that small data losses and/or momentary outages could be tolerated? The answer to this deceptively simple question has a big bearing on costs – as can be seen only too readily when purchasing cloud services with defined service levels, guaranteed up time comes at a cost.

Another possibility is to build in resilience at the application layer, by developing applications with fault tolerance, working on the basis that failures will inevitably occur. This approach has been adopted with considerable success by some of the largest service providers including, for example, Google. For it to work, however, a massively distributed architecture is essential and the fault tolerance must be incorporated into the very core of application software.

Finally – and in many cases this is the most attractive option – it is possible to build resilience into the cloud/virtualisation layer by adopting a cluster approach that is prepared to deal with failures on the lower layers. It can do this, for example, by moving virtual machines to hardware that is unaffected by the failure, by restarting virtual machines or even by using public cloud services as a backup site. In some cases, this type of strategy could lead to a short-term reduction in the capacity or speed of the supported IT services, but in many applications this can be tolerated.

Achieving resilience at the cloud/virtualisation layer has two big benefits. The first is that the costs are almost invariably lower than those associated with building a rock-steady practical-pig hardware layer. The second is that the application software doesn't have to be specially developed to be failure tolerant, as the resiliency is provided below the application layer, which should therefore be unaware of and unaffected by failures.

A dangerous delusion

Providing resilience at layers other than the hardware layer has many attractions, as we have seen but, unfortunately, it has also led to the development of a dangerous delusion. This is the growing belief that if the software layers of an IT system can handle equipment failures in the physical layer, power protection and power management become optional or even completely unnecessary. In reality, nothing could be further from the truth!

Power protection remains essential in every type of system, irrespective of where resiliency is achieved, for many reasons. Properly implemented power protection will, for example, condition the power received from supply source – whether that be the public supply or a standby generator – and will filter out transients and other fluctuations, thereby providing the IT hardware with invaluable protection against damage. Power protection also provides for smooth sequential start up and shutdown of devices, functions that are essential in systems where it is accepted that hardware will, from time to time, go down.

Well-designed power management can help in guarding against another potentially serious problem – that of zombie servers. These are machines that are functioning intermittently or erratically, but have not quite failed completely. Just because their behaviour is so erratic and unpredictable, zombie servers are a potent source of instability and data loss in IT installations. A good power management system can be used to "fence" zombie servers, isolating them from the rest of the installation and turning them off. In this way, the overall functioning of the installation remains deterministic and manageable.

And there's another crucial reason why power protection and power management are essential in systems where resilience is provided at levels above the hardware level. For these higher level resilience strategies to work, the level or levels at which the resilience is provided must always be power aware – that is, they must always know the current status of the power being supplied by the principal source, which is usually the grid.

It's not hard to see why the higher levels need to know about the power status. Consider, for example, an installation in which resilience is achieved at the cloud/virtualisation layer, with a strategy that involves migrating virtual servers to remote hardware if a problem occurs. If there is a failure of the mains supply, the UPS will continue to support the local servers for a predetermined time, and will inform the upper layers of how much time they have for the migration to take place. But if the power system had not informed the virtualisation manager that the installation was running on battery power, how would the migration have been initiated? Relying on manual intervention in such cases is a strategy fraught with danger!

Options for resilience

There is no doubt that today's complex and dynamic virtualised computing environments, where services are no longer tied to physical servers, require comprehensive power management if their full potential for flexibility and resilience is to be realised. When appropriate power management is in place, however, it has the additional benefit of making possible a range of options for achieving the required level of resilience.

The first of these options is to arrange for the power management system to initiate the transfer of applications from a server or site facing imminent power disruption to another server or site where the power is good. Subsequently, the power management system can perform a graceful shutdown of the hypervisor and a controlled power down of the physical servers at the site where power is disrupted. With this option, the full service provided by the IT system remains available at all times, although users are likely to experience a slow down for a short period while the transfer of applications between sites is taking place.

A second option for achieving resilience in the event of power problems is to suspend non-critical virtual machines and migrate the critical ones to a backup site or server, then to shut down the physical servers affected by the power problems. Part of the service provided by the IT system – the part provided by the virtual machines identified as non-critical – will, of course, be lost. This is, however, acceptable in some situations, such as at manufacturing sites where power loss will mean that the production plant is itself shut down, so there is no reason to maintain the IT services directly associated with it.

The third and final option we will examine here does not involve the transfer of virtual machines between sites. Instead, in the event of a mains power problem, non-critical virtual machines are shut down, and the remaining essential virtual machines are consolidated on a small number of physical servers on the same site. Unused servers are then powered down. The system can then continue to operate on standby power until the batteries are almost fully discharged, when a graceful shutdown can be executed. By re-grouping the essential virtual machines and powering down unused servers, long runtimes prior to the final shutdown can be achieved with UPS installations of only modest size. And runtime can be extended even further by, for example, capping the power demand of the physical servers, although this will, of course, impact their performance.

Whichever of these or, indeed, the many other possible options is chosen, close integration between the power management software and the virtual infrastructure management software is essential. This is not only to provide administrators with all key information in a single window on the management console, but also to ensure that the virtualisation and cloud stacks are aware in real time of power conditions and available run time. Finally, as has already been mentioned, the link between the power and infrastructure management systems is essential if mitigating actions are to be triggered automatically.

Implementing power protection

Three basic approaches are available for the hardware implementation of power protection in IT systems with converged architectures: centralised provision, end-of-row and rack/container based.

The centralised approach, where a single UPS installation powers the whole site, is tried and tested, but it is worth noting that when higher layer resiliency is in place, it is often possible to specify UPSs with a shorter runtime than would otherwise be needed, particularly if the resiliency strategy involves quickly migrating virtual machines away from the affected site. When adopting a centralised approach to power protection, it

is usual to plan from the outset for maximum capacity, but to add power modules only as the load grows. In well-designed systems, the modules can be added without dropping the load.

End-of-row power protection systems can have a single or double power bus, and can be modular or provide full capacity from the outset. They allow redundancy schemes of all types to be implemented and provide links to virtualistion, but they offer no special advantages or disadvantages in converged architecture applications.

Rack/container based power protection systems are, however, particularly well suited for use with converged architecture. With this implementation, power protection is an integral part of, and optimised for each basic unit – typically each computing pod comes with its own power system. This means that the power system has the same modularity as the IT equipment and it therefore scales automatically. Further, it is readily possible to provide 1+1 redundancy with dual power supplies in the servers.

Optimising power efficiency with fluctuating load

Irrespective of the approach adopted in their implementation, power protection systems for converged and virtualised IT installations must take into account fluctuations in the application load. In some cases, the load can change from zero to almost 100% over a short time period, especially where strategies such as “following the moon” are in place. These changes in application load are reflected by similar changes in the power demand of the installation and, unless the power protection system has been designed with this in mind, the result is almost certain to be poor overall energy efficiency.

The reason for this is simple: UPS systems operate most efficiently when they are running at high load levels. For example, an 1100 kVA UPS operating close to full load may well deliver an efficiency of around 95%, but the same UPS operating at, say, 10% load, will probably struggle to achieve 85%. Fortunately, there is a solution to this problem, and that is to specify a system made up of several UPSs, which share the load, and which are complemented by intelligent multi-UPS management.

The power management software is designed in such a way that, at any given instant, it ensures that only those UPSs needed to meet the current power demand are operational. The other modules are held in a standby state where they consume almost no power, but the management software can bring them back into service almost instantaneously – typically in less than five milliseconds – when the load increases.
This arrangement means that the load is concentrated on the minimum number of UPSs needed to meet it and that, as a consequence, these UPSs are well loaded and will therefore operate efficiently. With the right hardware it is, in fact, possible to take this multi-UPS approach even further, by specifying UPSs that are themselves made up of modules that can be instantly transitioned between standby and operational mode.

Such an arrangement, which is usually described as variable module management (VMM) allows the UPS power capacity to be accurately matched to the power demand over a very wide range of loadings, ensuring that high operating efficiency is achieved under all operating conditions, irrespective of load fluctuations.

Converged power in a nutshell

This paper has discussed in detail a wide range of issues relating to the provision of power for converged infrastructure. It is, however, worth summarising some of the main points.

The first is that, with converged infrastructure, the pod is the basic unit and it makes good sense therefore to provide it with its own power system, as this ensures compatibility and easy scaling. The second point is that effective power protection and management provisions are absolutely vital – even when resiliency is built into layers above the hardware layer – to guard against possible hardware damage by power transients, to eliminate the stability problems associated with zombie servers that are neither fully operational nor totally off-line, and to keep the higher layers informed, in real time, about power status.

The power management system must also be tightly integrated with virtualisation management systems, so that administrators are fully aware of the power status at all times, and so that power-fail routines – such as

the migration of virtual servers to unaffected physical servers – can be automated. Such automation is essential because many of today's IT systems are too complex for effective and timely power failure response to be managed manually.

Finally, the power systems must be designed from the outset to deal not only with fluctuating demand, but also future changes in demand, and with the clear understanding that failures can never be totally eliminated. When all of these factors have been fully taken into account, the ultimate goal of combining maximum application availability with optimum cost will be achieved.

About Eaton

Eaton is a diversified power management company providing energy-efficient solutions that help our customers effectively manage electrical, hydraulic and mechanical power. With 2012 sales of $16.3 billion, Eaton is a global technology leader in electrical products, systems and services for power quality, distribution and control, power transmission, lighting and wiring products; hydraulics components, systems and services for industrial and mobile equipment; aerospace fuel, hydraulics and pneumatic systems for commercial and military use; and truck and automotive drivetrain and powertrain systems for performance, fuel economy and safety. Eaton acquired Cooper Industries plc in 2012. Eaton has approximately 103,000 employees and sells products to customers in more than 175 countries. For more information, visit www.eaton.eu

About the author

Tatu Valjakka is a Product Manager for Eaton's power quality business and first started working for the company, a global leader in power protection, distribution and management solutions, in 1998. In his current role, Tatu is responsible for the management and marketing of Eaton's power management software in the EMEA region. Prior to this, Tatu was a Software Project Manager at Reiter Engineering for 6 years before becoming a Project Engineer developing IT systems for the Finnish Civil Aviation Authority. Tatu studied Computer and Information Science at Aalto University School of Science and Technology in Helsinki and, outside of work, enjoys gardening and fishing as well as sports.