



Truly Diverse Power in a Modern Datacentre

Robert Williams
(Robert@CustodianDC.com)

TABLE OF CONTENTS

The problem: Dual power isn't always Diverse..... 3

Dual Vs. Diverse 3

Know What You've Got 4

 SINGLE feed from single UPS systems 4

 DUAL feed from parallel UPS systems 5

 DIVERSE feeds from independent UPS systems..... 6

Things aren't always what they seem..... 7

The Solution: Ask the questions (and Deploy Correctly)..... 7

You have a single feed 7

You have two feeds from the same (or paralleled) UPS system 7

You've found out your diverse feeds are truly diverse? Great! But... 8

Where do we go from here? 9

Testing, Testing. 9

Caveats: The 110% Rule 9

Not all ATS/STS are alike (20ms makes a big difference) 9

So, what do we do?.....10

Thanks and related reading.....10

About us10

THE PROBLEM: DUAL POWER ISN'T ALWAYS DIVERSE

DUAL VS. DIVERSE

As the capabilities, designs and technologies used to deliver electrical power to the modern-day data centre rack have constantly evolved; some of the fundamentals have become somewhat blurred. This can inevitably lead to the incorrect provision of services and mistakes in the configuration of equipment. The result of which, is a sub-optimal usage of power resilience and a solution which, when a failure occurs, fails to meet the expectations of the operator. This then results in downtime and loss of key 'mission-critical' services when in reality the impact of the failure was totally avoidable.

This situation is made slightly worse by the wide variety of modern datacentre designs currently available to host your equipment with. Although there are datacentre 'Tier' standards out there, often the implementation used by a facility will not strictly follow the standard; especially when there is a particular market force in place pulling the facility along a certain path.

Efficiency in particular, in both the use of tangible resources (as well as financial ones) results in previously technical keywords like 'Diverse' and terms like N+1, 2N and so on, becoming more commonly used as marketing terms. As such, these terms are now often used to outline generic models by which a data centre aims to operate and no longer guarantee a direct reflection in their actual infrastructure. This is as much an issue for power delivery systems as it is for cooling and connectivity, but this paper will focus specifically on power delivery.

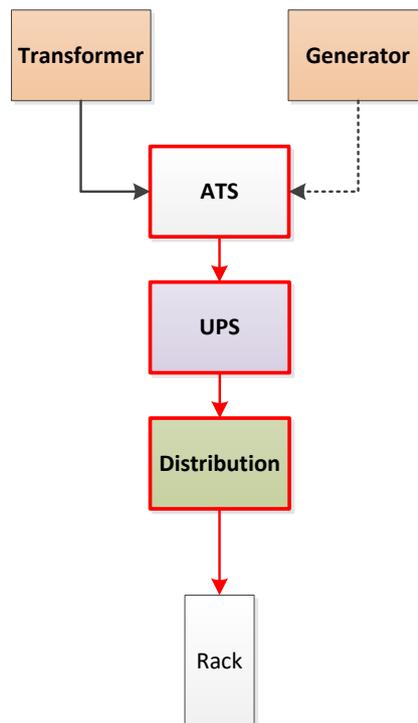
KNOW WHAT YOU'VE GOT

There are several ways in which power can be delivered to the modern colocation rack, the most common of these include:

SINGLE FEED FROM SINGLE UPS SYSTEMS

In this system, each rack has a single supply from a single UPS which will in most cases be able to switch between a mains supply and a generator backed supply. For years this was the staple diet of most data centres, giving a constant, clean supply regardless of the condition of the national grid.

A typical SINGLE (Tier 1) Power System:



Of course, the downside to this system is that any event at any point along the supply line can interrupt all the power to the rack, leaving it offline without warning. Possible failures include the loss of an upstream breaker, a local power bar fault or even a single UPS fault. All in, this solution (although still extremely common today) is far from being able to guarantee a 100% uptime of the rack's supply.

The Uptime Institute specifies this as either a Tier 1 or Tier 2 system, depending on the amount of resilience and spare capacity within the UPS system itself (i.e. N or N+1). If for example the UPS system is made up of several paralleled systems and one can be lost whilst the others are taking the full remaining load then this would be classically considered as a Tier 2 power distribution system.

With the above in mind, such systems do have substantial installation and on-going cost savings - and when used as part of a multi-geographic-site deployment with active/active systems at all locations can actually deliver a highly resilient end solution. However, this system is not ideally suited to a single site or mission-critical service delivery. For that you should as a minimum consider a dual feed from a parallel UPS system.

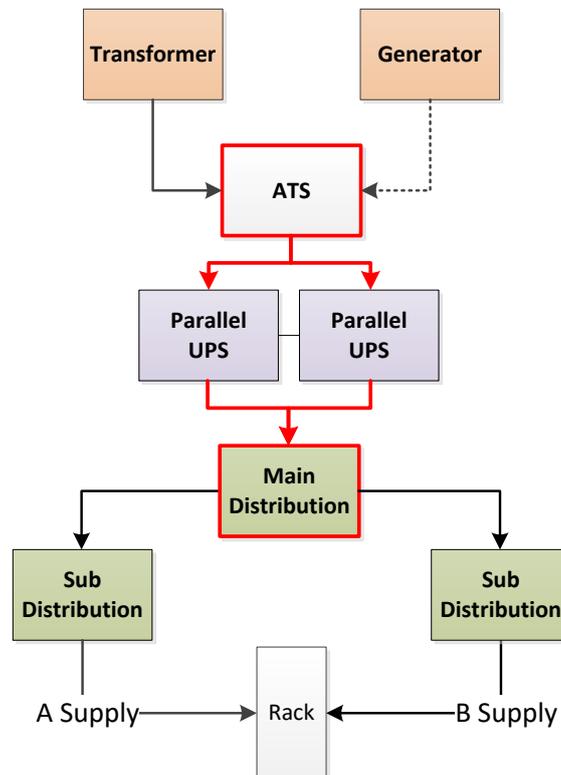
DUAL FEED FROM PARALLEL UPS SYSTEMS

The next logical step from a single feed is to take two feeds into each rack from two separate supply breakers. This way, any individual equipment failure within the rack itself will only affect one of the two feeds available to the rack. Our own research suggests that this is the most common method currently in use across UK facilities. With this system a facility operator can utilise all UPS systems to a reasonably high percentage of load.

However, this system is still at risk from many of the single points of failure which affect the single feed. What is often not considered is that the majority of these potential single failure points will result in a loss of supply to BOTH feeds to the rack at the same time.

For example, with a parallel UPS system all the individual units are synchronised onto a common bus. Two issues can arise here; firstly, they can become unsynchronised and remove themselves from the bus for safety. Whilst this does not cause an outage directly, if the remaining UPS systems cannot supply the load then the system will switch directly to raw mains supply. This also means that any UPS on the bus which develops a fault has the potential to affect all the systems.

A Typical DUAL Supply (Tier 2/3) Power System:



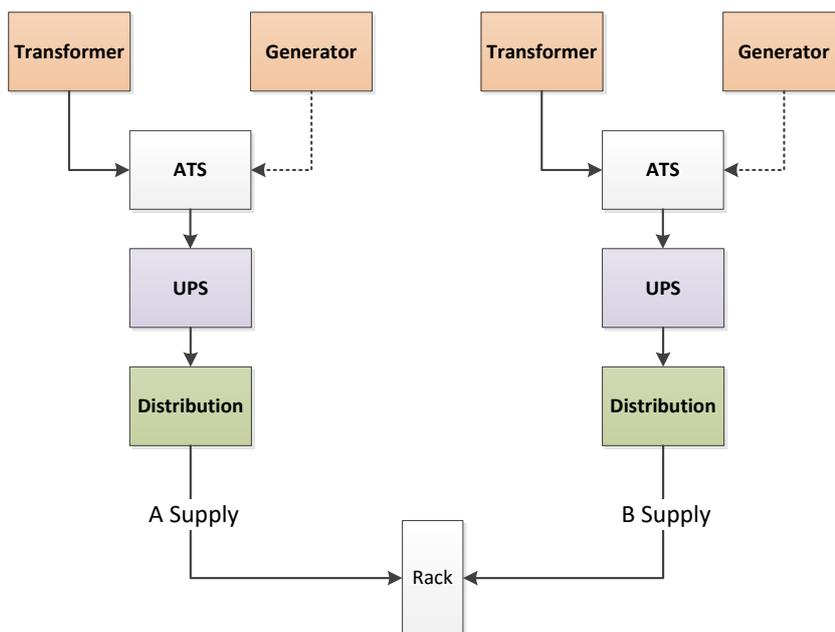
The second issue is that any kind of pre-UPS or post-UPS supply fault, such as a main distribution level breaker trip or ATS changeover failure will all have the same impact on both the supplies from the UPS cluster. So, although this method is an improvement over the single feed method, it still not ideal and actually carries more component level points of failure than a simple single feed system by having more interconnections.

According to the Uptime Institute, a Tier 3 datacentre should have 1 active and 1 alternate distribution path. As well as supporting concurrent maintenance; which means that each and every capacity or distribution component necessary to support the load can be maintained on a planned basis without impact to the IT environment. Therefore the above system is not actually Tier 3 compliant (you cannot maintain the ATS or Main Distribution systems without taking both supplies offline). Also there are several single points of failure (red) within the system at all times. As a result, many still refer to it as a Tier 2/3 power distribution system.

DIVERSE FEEDS FROM INDEPENDENT UPS SYSTEMS

Now we come to the first truly diverse configuration. In this system, two separate UPS units are used to supply the two feeds to each rack. The UPS systems are not linked and operate autonomously from separate ATS panels which can each provide separate mains and generator supplies. Each system also has seamless maintenance bypass capabilities enabling either UPS to be worked on without interruption. This design meets with the Uptime Institute Tier 4 requirements of having 2 simultaneously active paths.

A Typical DIVERSE Supply (Tier 4) System:



These two isolated UPS systems feed separate downstream distribution systems, which in turn provide separate power to the two feeds in the rack itself. This takes the diversity all the way to the rack footprint itself.

The downside of this system is that it requires twice the financial investment by the facility, as well as significantly higher operating and maintenance costs. The reason for this is that it is effectively delivered in the form of two separate datacentre power systems, requiring twice the cable, twice the number of distribution boards and twice the amount of UPS capacity to be available.

Initially the reasons for this are not always apparent, so let's look at an example. Consider a facility configured in this manner, using just two separate UPS systems. Now let's assume a total failure of a one of the UPS systems. At this point, the remaining power infrastructure must continue to hold its own load as well as taking the load which has transferred from the failed UPS supply.

This transfer has occurred as dual-PSU servers, or customers with single PSU equipment who are using ATS (automatic transfer switches) have had their equipment automatically begin pulling their full load from the remaining live feed. This means that every distribution board, UPS system, generator, cable etc. must be suitably sized to enable it to safely provide 200% of its normal operating load.

This 'spare' capacity must exist but cannot be used. So UPS systems must not exceed 50% load and breakers and cables that must be carefully oversized. All this is to ensure that a sudden load transfer (resulting from a failure or maintenance) does not result in a double-failure or overloading of systems. This is the price which must be paid for a truly diverse power delivery system.

THINGS AREN'T ALWAYS WHAT THEY SEEM

We did a snapshot survey among customers who also host some of their equipment at other datacentres. We were interested to learn that many could not identify which power delivery model their racks hosted at other facilities were using beyond knowing they had either one or two feeds. Most clients with two feeds were uncertain which category of diversity their racks were supplied by. With such a large resilience gap between **Dual Feed** and **Diverse Feed** it is critical to understand which of these solutions is being delivered to your equipment.

This question needs to be extended beyond the single-rack level also. In short, consider that you have two separate racks within the same facility, each with dual feeds, one acting as a primary and one as a standby. So you have 4 separate feeds, but there is a distinct chance that unless you have confirmation otherwise they may well all come from the same upstream distribution board, or even the same UPS chain.

THE SOLUTION: ASK THE QUESTIONS (AND DEPLOY CORRECTLY)

Now that you know which type of power system you are using, you will probably fit into one of these three categories:

YOU HAVE A SINGLE FEED

Ok, so maybe you hadn't considered dual-feeds before, or use entirely single-PSU equipment. If that's the case, we hope that this document may have given you some ideas or insight and you may want to investigate the possibilities of an additional feed or an ATS for your single-PSU equipment.

Alternatively, perhaps you have multiple sites in full active/standby or active/active configurations. In this case, we agree that having an additional feed to either location may well be unnecessary as you can lose an entire 'site' and still remain operational.

Single feed does offer cost savings in both the rack deployment and the equipment used within the rack and there are scenarios where this is a preferred (or necessary) solution.

YOU HAVE TWO FEEDS FROM THE SAME (OR PARALLELED) UPS SYSTEM

If you have found out that what you thought was 'Diverse' is actually just 'Dual', then don't despair just yet. Speak to your facility and ask if they can provide you with a feed from a different UPS system. Many facilities often offer this at an additional cost as they can provide you with power from a different zone (or floor).

Alternatively you may want to look at moving some of your equipment to another facility, or splitting it up within your existing facility in order to benefit from physical supply diversity if it's available.

YOU'VE FOUND OUT YOUR DIVERSE FEEDS ARE TRULY DIVERSE? GREAT! BUT...

Are you utilising them correctly? Earlier this year we did a technical survey of power connectivity across a selection of customer racks within our Maidstone hosting facility and were genuinely surprised by the results. Here are the two most common misconfigurations we found:

- **All equipment dual-PSU, but some devices with both supplies in the same feed.**
This is an easily avoidable fundamental error and will of course go unnoticed until such time as a fault occurs. One of the best (and simplest) systems we have seen for getting this right is just to use alternating colours or labels for 'A' and 'B' power cables. Also clearly label any extension bars you use so that you know where their power is sourced from upstream.
- **All servers with dual-PSU, correctly connected, but a network edge device with a single PSU.**
This was the most common issue we found. The client had invested heavily in highly resilient dual-PSU servers and storage equipment, but at the network edge, where their rack meets the internet, they had a single PSU device. Not only is this a risk in case the supply within the device itself fails, but it also negates much of the usefulness of a truly diversely fed rack as services will inevitably be affected if they lose all communications.

The ideal setup is of course perfectly configured with a mixture of dual-PSU and single-PSU equipment on ATS (or STS) fed bars. It goes without saying that single-PSU equipment will always have a place within most racks, but when connected to a small inexpensive ATS unit, it gets full access to both of the Datacentre floor diverse supplies. This and the combination of either dual or 2 x single PSU devices operating in active/standby helps you get the most out of your diverse supplies.

As you can see, there are still many caveats and design considerations which must be investigated and clarified in order to correctly utilise a diverse supply. These and several other restrictions are covered in more detail below, many of which apply regardless of whether your supply is **Dual** or truly **Diverse**.

WHERE DO WE GO FROM HERE?

TESTING, TESTING.

Whether your setup is Diverse or just a basic Dual feed - we recommend testing and simulating the loss of power at least once every 6 months during a scheduled at-risk window.

This basic testing detects many issues such as:

- A potentially faulty power supply which may fail and reveal itself under full load.
- Incorrect cabling where a device has been moved or re-tasked but not had its supply moved.
- Undersized power bars or breakers.
- Unexpected failures of active/standby networking hardware that occur during a state change.

We all accept that faults do happen and as with all issues like this, it's better to deal with them when you are expecting them than when the worst happens and you lose one of your feeds unexpectedly. Many of our clients test their deployments regularly by shutting down one of the supplies to their racks during a maintenance window.

CAVEATS: THE 110% RULE

One final important fact to consider is that for many kinds of dual or multi-PSU equipment, when they experience a loss of one of their supplies, they don't necessarily just 'transfer' their load to the other power supply. They often add to it by increasing internal fan speeds to compensate for the 'lost' PSU fan. Effectively entering a 'panic' state of cooling to ensure the system remains available.

On a full rack of dual-PSU equipment we have observed that this can lead to an additional 2-10% of power consumption in some cases. So make sure that both your feeds are suitably sized to take the full 110% load if one of them fails.

With that in mind, it is also worth checking with your provider that you are able to pull full load of your rack(s) from just one of their feeds. For example, if you have a 16A rack, and you are using both feeds concurrently, check that either one of your supplies is capable of delivering the full 16A to your rack in the event that the other supply has failed and your equipment is now pulling more 'total' power than it was when it had both supplies.

Specifically you need to confirm cable and breaker sizing, as well as your own infrastructure within your rack such as your power bars fuse-ratings.

NOT ALL ATS/STS ARE ALIKE (20MS MAKES A BIG DIFFERENCE)

Where you are using an ATS or STS unit to deliver two supplies to a single-PSU piece of equipment, or to a whole rack of equipment there are several design considerations which you need to cover off. Firstly, not all load transfer units are capable of switching their supplies in the same time. Static Transfer Switches (STS) are always extremely fast (typically around 4ms) but there are many mechanical ATS solutions on the market at the moment which can take up to (and over) 20ms to perform their switchover.

As a result, even though the load is correctly transferred to the alternate supply, a 'slow' transfer can result in unpredictable equipment behaviour and in some cases cause an unexpected reboot or reload of some devices.

Recommended Reading: APC provide an excellent whitepaper titled "Powering Single-Corded Equipment in a Dual Path Environment" which we highly recommend.

SO, WHAT DO WE DO?

At Custodian, we only offer racks with truly diverse feeds from separate UPS systems, each of which has its own grid supply and its own fully independent ATS system. We also have 3 separate 11kV transformers with three separate grid feeds to ensure we don't run on our generators for a minute longer than we have to.

We genuinely believe this is the only real way to achieve total availability and truly diverse power, as demonstrated by our 100% uptime record for power to all our racks since we launched. We've had our share of component failures of course, everybody does, but due to our truly diverse systems we have never failed to deliver power to a customer's rack during a power delivery system fault or maintenance.

To ensure we can deliver on our 100% SLA, we employ detailed power system monitoring at every level of our power distribution system. This is then modelled, in real-time, to ensure that at any given time a failure and the subsequent 'load shift and grow' can be handled by the counterpart UPS, distribution and generator systems.

THANKS AND RELATED READING

Thank you for taking the time to read this whitepaper. You may also want to see other upcoming papers on True Diversity of Network Connectivity and Resilient Dynamic Fresh Air Cooling systems; both of which tackle similar aspects of achieving true diversity within the networking and cooling areas of a modern datacentre.

ABOUT US

Our group has been involved in technical service and infrastructure provision for over 30 years. Custodian was born out of the growing demand for our data centre services, which have been an integral part of our facility for many years. Our latest dedicated Data Centre provision has been created utilising contemporary, forward-looking technologies, offering our customers the most robust, cost-effective solutions in today's demanding marketplace. Our total commitment to the people we do business with is the reason a number of major international companies have placed their business in our care.

As you will no doubt see splashed all over our web site, we won the Data Centre Leaders Awards 'Green' category. The same passionate, considered, and outside the box approach that won us that award runs right through our business. We aim to be the most efficient, resilient data centre facility provider in the UK, whilst maintaining a flexible, enthusiastic and easy to work with approach to doing business. We welcome visitors, whether you're a potential client, partner or another data centre provider we are happy to hear about your business and talk about ours.

Our culture towards service provision is very simple – we recognise every customer has individual requirements and we pride ourselves on being able to respond to those individual needs with non-prescriptive solutions.

This whitepaper was written by Robert Williams, the Technical Director and lead infrastructure designer at Custodian. He has designed and developed many of the efficiency improving and high resilience systems which have helped to keep Custodian ahead of the game in the marketplace. Feel free to contact him directly or speak to anyone from our team here. We all love talking about our infrastructure and we will gladly discuss any details or answer any questions you may have.