Mice and Men

By Simon Campbell-Whyte, DCA Executive Director

I OFTEN HEAR PEOPLE quoting a percentage of failures within data centres due to human error as 75%, sometimes more. However, I can't help but think that it must be nearer to 100%. For example, think of something that you would normally consider beyond the control of the data centre such as a sub-station failure, surely the blame rests with the utility company and not me? But you could argue that I should have anticipated this risk and mitigated against it by installing the appropriate back up or recovery system? In a another example, a colleague of mine recently reported the discovery of the remains of an unfortunate rodent who met a grisly end not only for themselves but also the demise of an entire server room - an act of nature perhaps? Or could this event been prevented with an appropriately proactive anti-contamination strategy?

The point I'm getting to here is that as the population continues to inadvertently rely more and more on the digital services we now can't live without, are owners of data centres and server rooms which service this growing demand responding to these changing expectations and demands by regularly reviewing their data centre against their ICT requirements and risk profile?.

This involves regular stakeholder consultation alongside robust strategic thinking, planning and investment. Granted, this is easier said than done when faced with managing IT infrastructures that have exploded in scale, cost and complexity, alongside an increasingly demanding user base. But this is a job that if avoided will only increase the risk stakes and exponentially the scale of the task. Here in Europe we now have a suite of guidelines to help users navigate the problem. For example within the CENELEC EN50600 series of standards there are six parts, which if used in conjunction with best practices such as the EU Code of Conduct, are a good place to begin tackling these issues.

The EN50600 series addresses the design, construction and operation of data centre facilities. It also deals with defining resilience levels or "classes" and each of the major systems and subsystems of the data centre with a great deal of informative information, on power distribution, environmental control, telecommunications & cabling, management & operational information and security.

At this moment in time, the bulk of the EN50600 series is already complete with some remaining parts close to release, so at this early stage it remains to be seen how long it takes for EN50600 to "stick" as the definitive guide within Europe.

However long this takes, one thing is clear, EN50600 is a Standard with a capital "S" which represents a significant step forward for the sector, and that alone will give it the necessary legal gravitas. So perhaps this should be the catalyst to see if your data centre and the people charged with running it are properly equipped to deliver your business's needs now and in the future. I therefore hope you find the articles from DCA members in this month's journal informative and helpful.

The vital elements of an end-to-end disaster recovery plan

Mike Bennett, VP, global data centre acquisition and expansion, CenturyLink EMEA.

Today's data centre must maintain continuous operations in all but the most dire of circumstances. Achieving 100 per cent uptime is difficult in any situation but achieving it nowadays, during times of continuous high-velocity IT change only increases the need for thorough planning.

Data centre resiliency has two distinct components: IT stack resiliency (server, network, storage, etc.) and facility resiliency (building security, power, cooling, etc.) At the highest level, IT leaders must develop a comprehensive plan that identifies business processes that could be affected by adverse conditions and details how to maintain continuous operations through such events. The first step is mapping business procedures to the systems that support them and determining how critical each process is to the health of the business. Based on those priorities, Recovery Time Objectives (RTOs) and Recovery Point Objectives (RPOs) need to be established for each critical application. RTOs represent the maximum time the business can be without a key service before significant impact whereas RPOs represent the amount, or extent of, data loss that can be tolerated.

With this information in hand, appropriate disaster recovery and business continuity provisioning can be developed to align with the business requirements of each stage.

The next priority should be the creation of a detailed recovery or 'failover' strategy. This must include crisis and emergency management actions and a formal communication plan. It must also prioritise which applications to restore, in what order and how to activate them.

The strategy must include:

- The activities needed to 'activate' a disaster recovery site, including: communication, connectivity, security, database and application services
- Predetermined disaster recovery mode operations and an overview of active controls and decisions on non-standard or manual procedures
- Steps for restoring to the primary site and resuming normal operations

- A disaster recovery invocation process, with roles and responsibilities for each functional area and all staff contact details
- Defined escalation procedures for all processes related to the recovery of the application services

Recovery instructions for getting back to the original application environment

Underpinning all this should be a disaster recovery and failover testing programme that continuously evaluates how each participant reacts and executes his or her role, as per the plan. It should compare targeted RTOs and RPOs with what is actually achieved and identify ways to improve — even if the initial targets have been met.

There are varying degrees of validation testing for operations and recovery plans, from table top walkthroughs of the process to the extreme case of blind failover testing. In blind failover testing, specialised systems automatically terminate a process, typically during relatively slack periods for the process involved, and without prior knowledge of data centre personnel. This scenario tests data centre managers' alertness and ability to respond to the failures generated by the termination. The approach not only maintains management skills at a high level, it often reveals unexpected issues that can then be isolated and addressed.

Conclusion

The creation of a detailed and functional uptime plan is not the end point. Companies often find that they must revisit these plans regularly, simply because technology is a dynamic process that's constantly evolving. Therefore, procedures, processes and training needs to evolve with it.

As business dependency on IT systems increases and the cost of downtime skyrockets (the average cost of IT downtime to an organisation has risen 65 per cent over a two year period) IT leaders are being called upon to increase the resiliency of their IT infrastructure.

This requirement includes both continuity of normal data centre operations and disaster recovery scenarios. Achieving 100% uptime amid continuous IT change requires a culture of rigorous planning, testing, and continuous improvement.

Providing an outlet for thought sharing, The DCA community is an ideal place to discuss disaster recovery with fellow industry professionals and its members' portal has further reading, webinars and seminars on the subject.

Service availability & resilience

By Lexie Gower, Marketing Manager at Datum Data Centres.

IF YOU ASK MOST LINE of business personnel their basic requirement from their organisation's data centre, availability is almost certainly going to be towards the top of their list. And as we all become more connected, and more and more business is transacted over the web, the potential impact of downtime increases exponentially.

When downtime is not an option
In today's always-on world, customers increasingly expect 24x7x365 service. The local department store, the high street bank, the local library may close their doors on high days and holidays, staff may even be allowed to go home at night, but regardless of time or day, we still expect to be able to access goods, information, money, entertainment… Whatever it is we want, the eternal availability of the internet is almost a divine right – and if we can't get what we want when we want it, we will throw a tantrum or go elsewhere.

Organisations get their own specific availability profile wrong at a high cost. Assessing what is acceptable in their market and to their customers and suppliers is critical in planning the accessibility of their systems. If they misjudge the expectation, even a short period of downtime can impact corporate revenue, reputation, customer retention, market confidence and even employee satisfaction.

It isn't just about the facility

In the fast moving world of IT, service availability can be impacted by anything from the data centre to the equipment to the people. No matter how resilient the fundamental infrastructure, change is the norm, and when change occurs, human and technical errors have the potential to create havoc. Proper procedures and training are vital to ensure that no one person having a bad day impacts the whole business. And if the unforeseeable occurs, proper protocols have to be in place to identify and rectify the route cause as soon as possible.

Getting the basics right

As in most things, not every solution is the answer to everyone's question, but for organisations where high availability is crucial, the fundamental check box has to be the level of resilience offered by the data

centre itself. If you need a non-stop service you need a non-stop facility, designed and managed for 100% availability.

Consider firstly data centre security. Both the access protection and security provided by the location should be able to defend your IT from malicious attacks and vandalism that could interrupt your service.

Then consider data centre reliability. A failure in power or cooling can all too quickly bring down a data centre. Dual redundancy should be built in, cooling and power systems doubled up, with proper maintenance and service regimes, providing sufficient well serviced redundant electrical and mechanical systems to reliably eliminate single points of failure. The higher the resilience, the greater the level of redundancy and as a result, proportionately higher investment costs, including multiple UPS systems and generators.

And don't forget the impact of the equipment itself. High-density computer equipment needs greater power. Overheated equipment will have a shorter life and is more likely to fail at peak times unless the air conditioning systems are sophisticated enough to manage the incremental heat load.

Telling the whole story

Because not all clients have the same needs and not all data centres are the same, levels of resilience vary across data centres. This is where the value of a program such as the DCA's Certification Scheme comes into its own.

Not only does the scheme provide meaningful insight into a specific data centre's fit for purpose for a desired business role, it also provides clear identification of the facility's effective resilience, incorporating a combined assessment of design, deployment strategy and accompanying maintenance processes. Using the scheme, potential data centre clients can review business needs and their availability profile against the differing levels of resilience offered by different providers.

For a facility such as Datum's FRN1, experience has shown us that our story resonates very strongly with organisations for whom high availability and security are paramount. When it truly matters, resilience has to be designed in to the build along with sufficient levels of redundancy and accompanied by proper procedures in order to deliver the assurance that non-stop really is non-stop.

The transition from data centre cooling to smart thermal management

Emerson Network Power has been a partner with the Data Centre Alliance (DCA) since 2013, contributing to forums, standards and joint industry initiatives via the DCA's collaborative platform, Data Central. Roberto Felisi, product marketing director Thermal Management for Emerson Network Power in EMEA shares his insights on the future of data centre cooling.

THE LAST FIVE YEARS have been an interesting turning point for the data centre cooling industry. Data centre managers have come under increased pressure to manage and reduce their energy consumption and cooling has been targeted as one area for drastic improvement.

Data centre cooling, a task once given little attention by managers, has evolved into a highly skilled and specialised art as an increasing number of organisations are innovating and investing in the technology. Today, even the term "cooling" has become outdated and replaced by a phrase describing the science and process of managing heat within the data centre; "thermal management".

Looking forward, the Emerson Network Power Data Center 2025 report has shown that, while thermal management broadly appears to be meeting today's market needs, many experts have predicted that current thermal management technologies are likely to stay the same for the foreseeable future.

With this in mind, questions are being raised around whether or not existing technologies are meeting the demands of future data centres; or whether more innovative cooling methods will be required to meet these new demands?

Future trends suggest that the data centre eco-system will continue to evolve, as will the need for thermal management to process and monitor additional power in warmer environments. The data centre environment is getting hotter and data centre management has evolved to incorporate these higher temperatures. As little as ten years ago the standard working temperature of a data centre would have been around 22 degrees Celsius. Today, it is quite normal to have data centres running in temperatures of 24-30 degrees Celsius or higher, as confirmed by the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) that expanded its data centre operating temperature guidelines.
As a result of this changing trend, data centre operators and manufacturers have new scope to be creative with cooling solutions.

The most critical challenge that data centre managers' face is pressures to effectively manage the impact that their operations have on the environment. To combat this, managers will need to adopt more innovative solutions and processes such as evaporative and adiabatic cooling. These methods are now starting to be applied here in the UK and in the rest of Europe, and with their cost saving potential; it's easy to see why.

Evaporative cooling is a process of evaporating water to lower air temperature. This method was first used in the ancient city of Pompeii over 2,000 years ago to lower the temperature of a hot room by spraying water on the floor. This then has a cooling effect on the air as the water evaporates. One interesting use of this technology is the enhancement of a freecooling chiller's efficiency by adding adiabatic cooling, which is obtained by humidifying ambient air going into the heat rejection coils (both the condensing and freecooling ones). By combining adiabatic, freecooling and mechanical cooling technology in one single unit, data centre managers can rely on full redundancy and new levels of efficiency as well as reduced overheads.

A second application of this age-old technology is the combination of indirect, evaporative freecooling with an air-to-air heat exchanger within a single unit. Through the water evaporation process, this technology is capable of achieving enhanced partial Power Usage Effectiveness (pPUE) levels and reducing CO2 emissions, thus minimizing operating costs.

There are many companies in the cooling market who have created solutions and products to address this technology; however Emerson Network Power has created products such as Liebert® EFC, the indirect evaporative freecooling unit and Liebert® AFC, the adiabatic freecooling chiller, which are the most advanced thermal management solutions in the evaporative and adiabatic cooling market. As a result of these latest applications; new generation data centres will have the ability to provide their customers with unmatched benefits in terms of cost savings, reliability, and cooling availability.

The industry is also moving towards data centres that will be able to autonomously adjust to specific requirements and to external conditions in order to run the optimal, most efficient cooling system – calculating the perfect balance between the local water and electricity costs, while always guaranteeing the optimal condition at which servers operate.

As a result, the evaporative and adiabatic cooling approaches will be a significant cost saver, but may not be suitable to all kinds of data centre installations. These methods of cooling, therefore, are unlikely to make traditional systems obsolete any time soon. The probable approach is that existing systems will be able to run traditional cooling in tandem with new solutions.

What seems clear above all else, is that over the next decade, the data centre will become increasingly "smart". While reducing cooling costs will continue to be the primary focus for data centre managers, so will the desire for investment in smart thermal management technologies. We are therefore likely to see further innovations in this field as the industry fights to reduce high energy costs and more effectively manage environmental impact.

Considerations for hardware maintenance & resilience

By Mat Jordan, Head Of IT Asset Recovery at Procurri.

I SUSPECT, as I write this that the majority of readers would be focussed around hosting , and datacenter services, redundant power , guarantees of power etc..; Procurri offers many services surrounding the physical hardware held within the datacenter environment. It is a critical piece, and who maintains and manages the front line is key.

Hardware maintenance solutions for servers and comms within the datacenter are key. Who monitors the machines, and how do they instigate a call should maintenance on a machine be required. Do they operate a fix first policy, rather than rejecting a call because a serial number is not recognised in the early hours? Important considerations to help you have a restful night's sleep knowing that action will happen without your intervention.
Are multiple manufacturer platforms operating within your datacenter environment? Do you hive off different hardware to multiple maintenance providers, or is a single point of contact preferred to manage your estate? Perhaps to manage your global estate? One point of contact, one number to call in the event of a failure within the estate?

We encourage our customers to consider the resilience of their hardware service partners , breaking the solution down into three key areas often aids this process when reviewing maintenance providers -

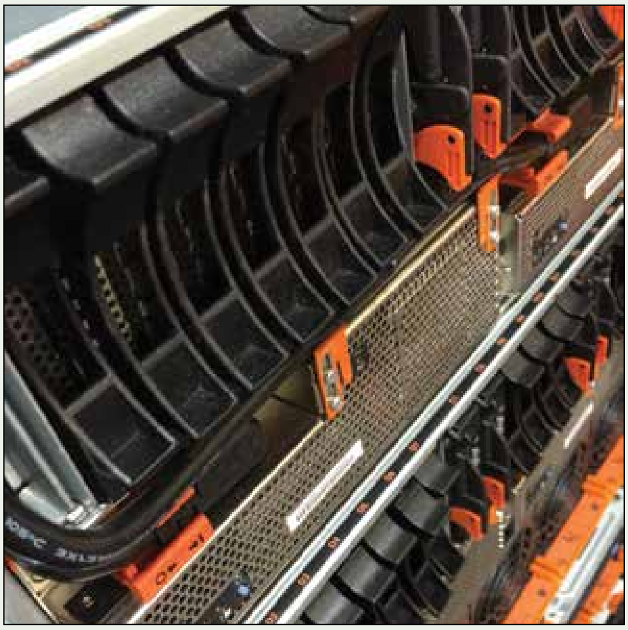
- **Parts -** availability of parts, location and proximity of those parts in relation to the site. Variety of hardware supported. Parts sparing methodology, are the right parts provided for? How are end of service life machines catered for?
- **Process** - is there a clearly defined process to instigate / log a call, how is it managed and delivered; what is the ticketing and escalation process. Is it 24 x 7?
- **People** - the last mile, are the skill sets held within the organisation to install the parts, who is likely to attend site, are they qualified to work on the type and model of machine you are running. What and where the Level 2 or 3 support is should it be required.

You need to feel confident that your chosen partner:
Understands your business and the importance that ICT plays in enabling your business to operate effectively. Possesses the experience your business requires to help you design and build a

resilient platform which meets your immediate business needs and works with you to agree a road map and strategy which will enable you to cope and respond to what may lay ahead. Sleeps with one eye open so you don't have too and can work with you to proactively manage and monitor your systems and perform regular health checks and physical system checks within the data centre.

Is responsive and "on the ball" when it comes to fault resolution - after all nothing is ever 100% so when (not if) a problem does occur your provider needs to posses the experience to quickly understand and interpret the issue and the data they see and can respond with the right people and parts needed to fix the problem which meets the SLA's in place with you and between you and your customers.
Is reactive to any issues that arise and can work with you to improve and or modify processes and systems to prevent a reoccurrence of the issue again.

Thank you for taking the time to read this article and I hope you found it informative, you can be confident that Procurri tick all these boxes and would welcome the opportunity to discuss your specific business needs to see if we can help you too.



The importance of UPS maintenance

All UPSs and associated system components need periodic maintenance and, on occasion, parts replaced to ensure optimal reliability. They also need the security of an emergency call-out facility, as failures can occur in any electrical equipment, no matter how well-maintained it is. In this article, Alan Luscombe, Marketing Director at Uninterruptible Power Supplies Ltd, a Kohler company, talks to PBSI about these elements of effective maintenance and how they contribute to an overall plan that assures the UPS system's uninterrupted availability over periods of many years.

MOST ELECTRONIC and electrical equipment, no matter how well-designed and built it is, will fail eventually unless it is maintained according to its manufacturer's instructions. On rare occasions, maybe after suffering adverse environmental conditions, even well-maintained equipment can fail.

Failure-induced downtime is unacceptable for any commercially-used equipment, but it's particularly serious for UPS systems; if they go offline, they compromise all the equipment they have been installed to protect. Fortunately, there are a number of steps, at several levels, that UPS operators can take to minimize this exposure.

The highest-level steps to ensure availability include keeping the UPS in optimum condition within a clean, dry, tidy and well-lit area, and checking that all alarms and indications are recorded, logged and reported correctly. Responsibility for the UPS and its associated equipment should be assigned to a named member of staff, and a suitable maintenance contract negotiated with a reputable UPS specialist.

This maintenance contract should cover three key components: A regular testing and preventative maintenance schedule; emergency call-out facilities with a guaranteed response time; and a policy for end-of-life component replacement.

The maintenance schedule should be designed to include all the UPS installation's major parts, which are the UPS itself, the battery and the generator if present. The schedule will be based on a planned programme of visits in which service technicians perform inspection and maintenance as required.

Regular UPS maintenance

For the UPS, meters and instrumentation should be checked for correct operation, and any meter readings checked, recorded and verified for accuracy. Local and remote monitoring panels and communications channels, as well as indicator lamps, should be checked for correct status indication. Switchgear and circuit breaker operation should also be proved.

Environmental conditions should then be reviewed, starting with actions as simple as removing any material and obstructions from around the equipment, and checking for any abnormal conditions. Ensure that airflow in and around the installation is unimpeded while changing any air filters that are due for replacement.

The equipment itself should be checked for evidence of excessive heat, noise or vibration, as well as signs of damage to components or to power and control wiring and connections. This includes looking for swelling and leakage on ac and dc capacitors. Power supply voltages and waveforms, together with the overall UPS operation, should be checked.

The service plan should also identify the degradation of any critical components, allowing repairs or replacement to be performed before a fault occurs. Backup spares should be available immediately on demand.

Battery inspection and maintenance

Traditional batteries such as open-vented lead-acid types could be checked using low technology methods such as visual inspection and voltage testing. These are not possible or effective with the Valve Regulated Lead Acid (VRLA) batteries found in most UPSs today. However, methods appropriate and effective to VRLA technology have been available for many years and are now widely used. The most important of these is battery impedance testing.

This approach depends on age-related changes in battery impedance. The internal impedance of a new battery is low, and measured in milliohms. As the battery ages its impedance increases marginally due to normal internal corrosion. For healthy batteries, this rate of increase should be uniform across the single and parallel arrays of serial battery strings that comprise them. Any battery, or group of batteries that shows an impedance increase higher than the others becomes suspect.

Battery impedance testing is useful and convenient, because it is easy to perform, and almost all battery problems manifest themselves as a rise in internal impedance. These include a loss of electrolyte due to venting through overcharging, leakage through seals, or sometimes migration of electrolyte between cells. Occasionally, excessive corrosion can cause loss of metal area or even disintegration within a battery. To perform the test, pass an AC current at a level related to the battery Ah rating through the battery; then use this value together with the millivolt reading across the battery terminals to calculate the internal impedance.

Regular impedance testing will track battery condition and allow accurate prediction of the end of the battery's reliable working life. Results can be computer-generated, as shown in Fig.1. This chart, recorded near the end of a battery's life, shows how impedance readings provide a better indication of problems than float voltage readings do. Impedance readings for blocks 8 and 9 are high, revealing problems not made visible by voltage readings.

Impedance testing has a useful complement in load bank testing; connecting an artificial full load to a battery and monitoring its performance during discharge. This is the only way of measuring battery capacity at the time of the test, and proving the integrity of all components and interconnections. It is a valuable tool, for example, to prove after commissioning that the battery will support the specified load for the specified time. However it should be used sparingly as it completely discharges the battery; the load is denied of full support until recharging is complete. Also, completely discharging VRLA batteries in this way reduces their life and carries some risk of permanent damage.

Permanently-fitted fixed battery monitoring systems provide another diagnostics solution. These allow continuous monitoring of battery voltages, current, temperature and impedance, and can detect differences between individual battery blocks during any charge/discharge cycles. As such systems are relatively expensive, they are used mainly in battery installations that are large or considered highly critical.

Generator and overall system checking

Generators, which are frequently part of a site power protection plan, should also have a regular service program. This should cover all major generator components, including the cooling, fuel and lubrication systems, the starting system and mechanical parts, and control panel. Insulation and load testing should also be performed.

As well as planned maintenance of the UPS, battery and generator components, occasional testing of the entire system is also worth considering. This test could ensure that the automatic mains failure (AMF) equipment functions correctly, the generator starts properly, the UPS accepts output from the generator successfully and the batteries recharge normally.

Emergency call-out resource

Even good-quality, well maintained equipment can fail occasionally, so any scheduled maintenance plan must be complemented by an emergency call-out service. Different options should exist to allow UPS users to balance response levels against cost. Typically, telephone support will be available 24/7/365, while a guaranteed time within which a service engineer arrives on site following a call-out can be negotiated. Similarly, levels of spareholding can be discussed and negotiated.

Uninterrupted availability now and into the future

UPSs exist entirely to protect other equipment, often on a large scale and running business-critical applications. Accordingly, providing protection for the UPSs themselves becomes essential. Fortunately, this can be effective provided that good quality, well-specified equipment is installed and maintained regularly, with provision for a rapid, robust response if a failure should ever occur. By following these guidelines, UPSs can offer the protected equipment's owners and users the level of uninterrupted availability they have come to expect, now and for several years of continuous operation.

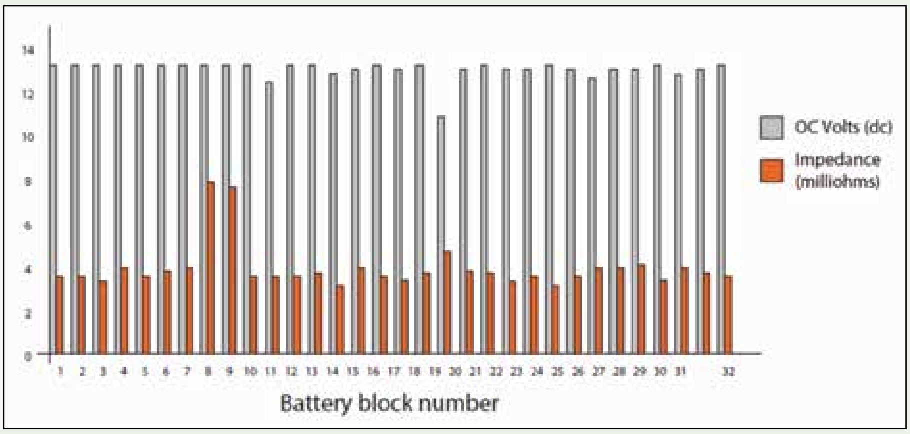


Fig.1: Impedance graph for battery at end of its working life