Uptime Institute® INTELLIGENCE

OPERATIONS & PROCEDURES

UI Intelligence report 31

How planning reduces the impact of outages

Author
Kevin Heslin, Chief Editor, Uptime Institute

Data center outages remain a major concern despite industry efforts to reduce their frequency, duration and cost. Maintaining and training staff on comprehensive, up-to-date procedures is a proven best way of reducing the likelihood of an outage and is key to restoring operations quickly afterward. This report examines the impact of outages and the relationship between operating procedures and outages.

30-45 minutes to read

UII-31 v1.0 published September 12, 2019 last updated September 12, 2019

This Uptime Institute Intelligence report includes:

OUTAGES: BE PREPARED

- **Understand the stakes. A severe outage can cripple an organization.**
- **Prioritize. Consider the effects of a possible outage. What systems are most important to the organization? Which are the most likely to cause an outage?**
- **Develop a robust set of emergency operating procedures, methods of procedure, site configuration policies and standard operating procedures. These procedures will include step-by-step instructions for restoring individual pieces of equipment and systems. Emergency operating procedures will guide staff response so an incident does not cascade into an outage.**
- **Train on procedures and keep them updated.**
- **Promote a culture of continuous improvement.**
- **Be prepared. Conditions that can lead to an outage can occur at any time. Management must plan so on-site personnel are qualified for the tasks they are asked to perform, and staff must be trained on procedures and understand the escalation process.**
- **Catch incidents before they cascade. A limited outage may not cause a service interruption.**
- **Know the limits of the facility. Procedures that work in a Tier IV facility may not work — or be necessary — in a Tier II facility.**

Introduction

Recent Uptime Institute research suggests that outages remain common, costing operators many millions of dollars in recovery, damage, compensation and reputational damage. The increased complexity of IT systems (sometimes due to efforts to make IT more resilient) has, in some cases, also increased the risks or the costs of an outage. Given this, it is unsurprising that almost all data center owners and operators take some steps to prepare for the eventuality of an outage, to limit its duration and impact. But this is not always done with sufficient diligence.

Many of the steps necessary to recover from an outage are the same as those needed to prevent failures: have a deep knowledge of the facility; identify single points of failure (e.g., single-corded servers); identify critical loads; develop and refresh procedures; and conduct scenario tests, then update the procedures.

The process for recovering from an outage can make extensive use of the operational procedures that reduce the risk of an outage. These procedures, including standard operating procedures (SOPs), methods of procedure (MOPs), site configuration policies (SCPs) and emergency operating procedures (EOPs), are most effective when they become part of an organization's everyday operations and are continually rehearsed and updated. Doing this prepares an organization to recover from an outage and limits the impact if an incident does occur.

MOPs, SOPs, EOPs and SCPs should be customized to the facility and the organization's business requirements, which makes facility knowledge a critical asset — data centers are very rarely alike. For example, one data center operator maintains a 60% power reserve in its Tier II facilities that allows for a smooth restart in the event of a data center outage, but they eliminate or reduce that power reserve in their Tier III or Tier IV facilities. Such a difference is critical — and is reflected in the procedures for these facilities, which are regularly reviewed and improved by on-site personnel.

Procedure development begins during commissioning. During this phase, data center owners/operators have the opportunity to assess the operation of mission-critical IT systems and minimize risk to them by developing procedures, defining adequate staffing levels, ensuring that on-site personnel will have the necessary skills to respond to an incident, and establishing that procedures be kept up to date and regularly rehearsed. The levels of investment in these areas can vary, according to the business requirements for the facility.[1]

The threat is real

The potential costs of a downtime incident make it imperative that organizations strive for high (continuous) availability. But failures are difficult to eliminate altogether. According to our **Annual Outage Analysis**, the number of downtime incidents remains stubbornly high, despite high levels of investment and new resiliency strategies. Size and prominence of an organization do not matter; organizations of all types have experienced downtime incidents in recent years.

Uptime Institute's longitudinal facility performance data indicates that establishing, maintaining and rehearsing a comprehensive set of procedures may reduce the likelihood of a downtime incident. In our 2017 global data center survey, members of the Uptime Institute Network (which emphasizes the importance of procedures) were half as likely to report a downtime incident in the preceding 12 months as nonmember respondents.

About outages

While all types of organization may suffer failures, impacts may vary with size, prominence or sector. For larger and more prominent companies or those with a particular business scope, a downtime incident may have a larger effect on the organization, its operations and its customers. Organizations that have significant consumer business or that support critical or complex logistical operations are particularly vulnerable. As documented in our report **2019 Uptime Institute global data center survey**, more than 10% of respondents said that their most recent significant outage cost more than $1 million ("most recent" could have been at any time in the past).

[1]Some facilities obtain Uptime Institute Management and Operations (M&O) Stamps of Approval or Tier Certifications of Operational Sustainability as independent third-party verification that they are adequately staffed and organized to operate their facilities to manage change over time.

Outages affecting large companies will generally affect more customers, have greater business impact and attract more media attention. Consider the costs that airlines and banks face in the aftermath of an outage, including lost revenue, customer loss, damage to reputation, possible fines and, of course, the labor and equipment needed to restore operations. In some cases, an outage can cause the share price and market capitalization of a company to drop, which can further affect a business's overall profitability.

In relative terms, smaller companies may face even greater consequences from downtime events. The absolute costs — and media coverage — may be less, but the loss of a data center may affect a larger portion of a small company's operations and customers. In some cases, smaller companies may lack the revenues to invest adequately in disaster recovery or contract for the same service level agreements (SLAs) as larger companies.

Frequency, duration and cost of outages

According to Uptime Institute's 2019 global data center survey of 1,126 IT and data center managers, just over a third (34%) of all respondents had an outage or severe IT service degradation in the past year, while half (50%) had an outage or severe IT service degradation in the past three years (see Figure 1).

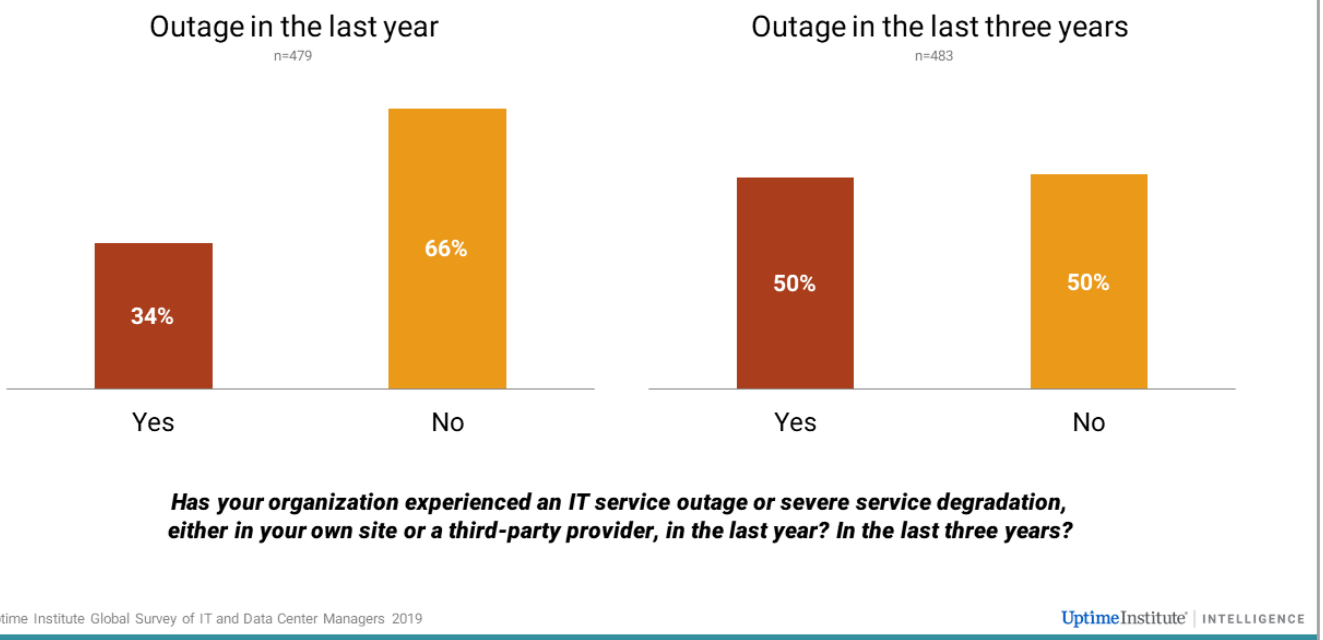


Figure 1. **The frequency of IT outages and service issues remains persistently high**

Duration is a further related measure of the impact of a downtime incident. Just over half of the publicly reported outages in our 2018 sample — more than in earlier years — lasted over four hours (see Figure 2). This aligns with 2019 survey data that suggests IT/network-related disruptions, a growing proportion of all incidents, can take longer to address than power/data center facilities disruptions.

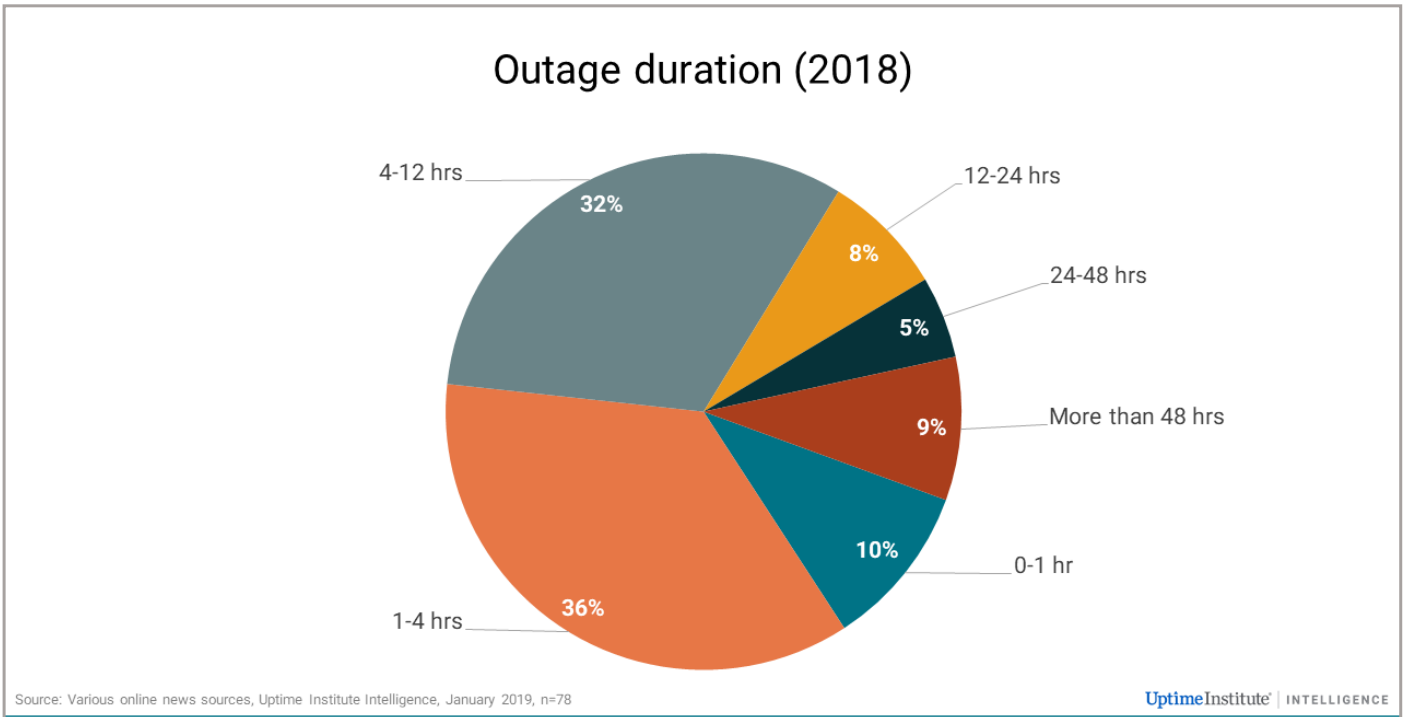


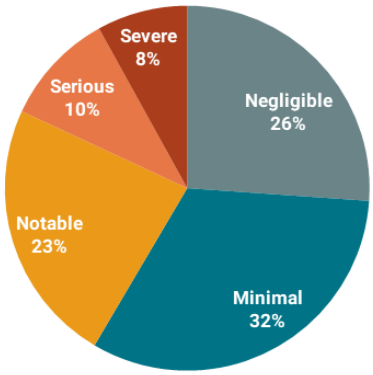
Figure 2. **Few outages can be addressed in less than an hour**

The same Uptime Institute 2019 survey documented the high cost of these outages. More than half the outage incidents reported by survey respondents cost under $100,000. Forty-one respondents reported costs over $1 million, and six additional incidents cost more than $40 million. These findings correspond to data gathered in 2018 and 2017.

Organizations do not always appreciate the full extent of their exposure until after an outage has occurred. For instance, a high percentage of the 2019 survey respondents — 61% — said that their organization did not formally calculate the business cost of downtime incidents. Fewer still, we believe, attempt to conduct a cost-benefit or business impact analysis when planning resiliency investments.

Outage severity

Discussions about downtime costs have long been impeded by the lack of a common vocabulary for describing the extent and severity of an outage. To address this void, Uptime Institute developed its Outage Severity Rating for classifying the impact of public outages (see Figure 3). The scale we developed is based on two main criteria, extent and customer impact, which we use to classify publicly reported outages into five categories.



Category 1 Negligible	Recordable outage, but little or no obvious impact on services.
Category 2 Minimal	Services disrupted. Minimal effect on users/customers/reputation.
Category 3 Notable	Customer/user service disruptions, mostly of limited scope, duration or effect. Minimal or no financial effect. Some reputational or compliance impact(s).
Category 4 Serious	Disruption of service and/or operation. Ramifications include some financial losses, compliance breaches, reputation damages, possibly safety concerns. Customer losses possible.
Category 5 Severe	Major and damaging disruption of services and/or operations with ramifications including large financial losses, possible safety issues, compliance breaches, customer losses, reputational damage.

Thinking of your most recent service outage, how would you categorize it? Choose one.

Source: Uptime Institute Global Survey of IT and Data Center Managers 2019, n=448

UptimeInstitute | INTELLIGENCE

Figure 3. **Uptime Institute developed the Outage Severity Rating to help organizations evaluate the impact of outages on their businesses**

When we applied this scale to three years of publicly reported outages (2016-2018), we found that the proportion of category 5 outages (severe, business-critical outages) is falling. At the same time, we found that IT-based outages led to four of the 11 category 5 outages we documented from 2016-2018.

A recent trend is that IT-based outages, which are now more common than full data center outages, may be more likely to be partial than network or facility-based outages, which may affect all applications and create cascading effects. However, failures of hybrid or distributed architectures will often affect many IT operations and the business units that depend on them. IT-based outages, while certainly disruptive, may have less impact than a complete data center outage.

Business impact analysis

Organizations that conduct a business impact analysis (BIA) to evaluate the impact of an outage and to help direct spending to prevent outages are more likely to avoid disruptions and their resultant costs. A BIA should be updated on a regular basis to keep current with changes. A refresh every two years is considered desirable for many organizations.

A BIA forces an organization to take a holistic look at its operations to evaluate how its business units and services interact, then evaluate which units or services if lost would have the most impact on the business. For instance, the loss of mission-critical IT will cause business units to suffer lost revenues and may affect customers. In finance

or online retailing, increased latency may have a business impact. In other cases, increased latency may not be a problem at all, just as the temporary loss of internal email or human resources functions may be simply an inconvenience.

A BIA normally focuses on one or more related business units — it is not limited to IT functions. Businesses such as retailers also need to evaluate how the loss of physical locations would affect them, and many businesses, especially manufacturers, need to analyze their supply chains. Mission-critical IT may play a role in all these functions.

A BIA also includes cost estimates for reducing a business's exposure to loss of critical facilities or services. These costs can include those incurred if, for example, physical data center redundancy was increased, procedures or processes were improved or if software were upgraded or rewritten.

How to conduct a BIA

There is no single BIA format that fits all business environments, but a complete BIA will often include all the steps below. Organizations should tailor the process to fit business requirements, strengths and weaknesses.

- Identify a team leader, chartered by the C-level.
- Assemble a project team representing all key stakeholders within the organization and schedule working sessions.
- Develop a matrix to capture the following information:
  - Business vulnerabilities
    - Most critical services.
    - Business entities affected in the event of an outage.
    - The extent of operational impact on each entity.
    - The likelihood of a business interruption, which should include considerations of physical redundancy in all IT facilities.
  - Business impacts
    - The cost of repairing or replacing damaged infrastructure.
    - Revenue loss.
    - Customer loss.
    - Legal expenses.
    - Potential fines.
    - Other, including marketing and advertising costs to recover business losses.

Many of the business consequences can vary greatly, depending on the duration and impact of the service interruption. Therefore, the BIA must include estimates that involve lengthy outages, such as might occur in

the event of a natural disaster or if orders for replacement equipment involve long lead times for delivery.

The BIA will take some time to complete, with work proceeding in phases. For instance, the project team may find that it needs outside help to facilitate the process or to develop tools such as surveys and interviews to gather information that the organization does not ordinarily collect.

BIA results

A BIA concludes with a final report that explains the impact of business interruptions, documents legal and regulatory requirements, and defines acceptable levels of downtime and losses (for example, detailed recovery time objectives, which establish maximum application downtimes, and recovery point objectives, which describe maximum data loss).

A BIA report typically includes:

- Executive summary with key findings.
- Methodology for data gathering and analysis.
- Detailed findings on how a service interruption would affect operations.
- Charts and diagrams to illustrate potential losses.
- Costs of remedial actions.
- Recommendations for recovery.

BIA follow-up

Follow-up to a BIA is the domain of executive leadership. Implementing the recommendations of a BIA ordinarily takes C-level authorization and organization-wide participation and support.

Senior management reviews the BIA report, which it may use to devise a business continuity plan (BCP) or disaster recovery (DR) strategy to dictate the organization's response to an incident. The two responses are similar, with the BCP focusing on restoration of the entire business and the DR strategy focusing on IT infrastructure and operations.

A BIA is not designed to identify all the possible risks and potential failures in the digital infrastructure. Rather, it is focused on quantifying the impacts should an outage occur. Uptime Institute's Hybrid Resiliency Assessment is an example of a methodology designed to identify risks and performance variables in a complex digital environment. Such assessments often include deeper analytical approaches, such as failure mode effects analysis.

Preventing outages

While outages remain common, according to our data, they are not inevitable. Many organizations successfully maintain continuous operations for years without experiencing a downtime incident.[2] These facilities, and others like them, rely on highly developed procedures that enable them to respond appropriately to incidents, whether they occur in the course of stable state operations, during routine maintenance, or as a result of a configuration or equipment change.

In our 2019 survey, we asked, "Would your organization's most recent significant downtime incident have been preventable with better management/processes or configuration?" Sixty percent (60%) said "Yes." This supports the view that even the best-designed data centers can be run less than optimally and suffer more incidents, while data centers with less-resilient designs can exceed expectations if they are well managed.

Human error

Data collected by Uptime Institute's Abnormal Incident Reports system suggests that approximately two-thirds of outages can be attributed to human error. Our data also reveals that downtime incidents are rarely caused by one failure, but instead are the result of cascading events. A facility designed around N+1 uninterruptible power systems (UPSs) should not experience disruption when one of the units unexpectedly fails or is taken offline – that is exactly the purpose of the +1. However, the loss of the unit will lead to dropped IT loads if the resulting power demands on the remaining units exceeds 100% or if the A&B sides are not properly configured.

Configuration problems like these are common and can result when SOPs and MOPs are incomplete or are not followed with precision. Further damage may result if an organization does not have appropriate EOPs or if on-site personnel do not follow them in the aftermath of an incident.

Scheduling

Uptime Institute data shows that abnormal incidents occur fairly uniformly across season, day of the week and hour of the day. Data centers with 24x7 shift coverage and cross training for periods when staff levels are low reduce the likelihood of downtime because these practices ensure the initial staff response to any incident is timely and escalation is appropriate.

[2]Uptime Institute Network presents its Data Center Uptime Performance Awards to facilities that report continuous operation of IT loads for a year, with many facilities reporting continuous availability for more than 10 or even 15 years.

Procedures

Data center owners and operators will find it far easier to recover from an outage in a properly run and maintained data center, further underscoring the significance of procedures. Uptime Institute has written extensively about the importance of procedures, including SOPs and MOPs, and their criticality to maintaining operations (see **Sustaining operational effectiveness for the long term** and **Top considerations for addressing data center facilities management risks**, among other Uptime Institute publications).

Procedures remain important when restoring operations in the aftermath of an outage. This is because the same MOPs, SOPs, EOPs and SCPs used in basic maintenance, configuration change or equipment replacement processes still hold and should be followed when working in a facility that has experienced an incident.

EOPs

EOPs play a particularly important role in downtime prevention and restoration of service after an outage. Unlike other procedures that tend to apply during normal operation of a facility, EOPs are the procedures to be employed during an unexpected incident, primarily to prevent a problem from intensifying. They include a detailed escalation path, which should always be up to date.

Organizations should develop their own EOP lists (usually including between 20 and 40 EOPs) based on a detailed single point of failure analysis of each facility's critical infrastructure. EOP lists can become excessive if they are extended to non-severe conditions, such as loss of communications to a single UPS — normal, non-emergency response is sufficient for these types of events. EOPs should align to the highest criticality in a loss-of-resiliency situation hierarchy, which in turn maps to real impact or loss of redundancy (i.e., a reduction to N).

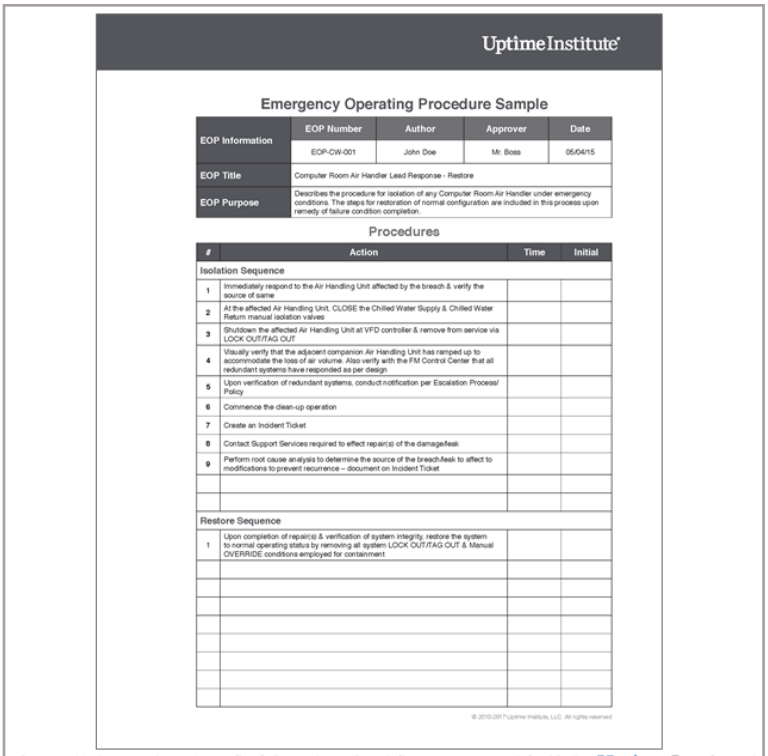
Most mission-critical facilities develop EOPs to address the following situations

- Loss of city power
- Loss of city water
- Generator fail to start
- Generator run fault (low coolant, ruptured block heater line, etc.)
- Activation of the fire suppression system
- Activation of the emergency power off system
- Loss of controls, mechanical or switchgear (response requires manual operation)
- Loss of building management system/electric power management system (particularly the loss of alarm notifications)
- UPS in static bypass
- UPS on battery
- Any loss of IT load
- High temperature alarm (data hall, UPS room, battery room, etc.)
- Loss of chilled water
- Chiller failure
- Loss of condenser water (chillers/package units high head lockout)
- Pump failure
- Mechanical, loss of:
  - Roof top unit
  - Air handling unit
  - Computer room air handler
  - Computer room air conditioner
- Power, related to loss of critical load:
  - Power distribution unit
  - Static transfer switch
  - Automatic transfer switch
- Breaker event
- Water or glycol leak detection
- Fuel leak detection alarm
- Battery high temperature alarm (thermal runaway hazard)

EOP response is usually alarm-driven but can also be triggered by physical identification of concerns during site rounds, etc.

The use of EOPs

EOPs should be invoked before an incident has caused an outage: they are intended to prevent downtime and service interruptions. Once a facility is in a stable state — operating without risk of a service interruption — an organization can concentrate on restoring normal operations by repairing equipment, restoring set points or transitioning from backup equipment. Figure 4 shows a sample EOP developed by Uptime Institute.



UptimeInstitute

Emergency Operating Procedure Sample

EOP Information	EOP Number	Author	Approver	Date
	EOP-CW-001	John Doe	Mr. Boss	05/04/15
EOP Title	Computer Room Air Handler Lead Response - Restore			
EOP Purpose	Describes the procedure for isolation of any Computer Room Air Handler under emergency conditions. The steps for restoration of normal configuration are included in this process upon remedy of failure condition completion.			

Procedures

#	Action	Time	Initial
Isolation Sequence			
1	Immediately respond to the Air Handling Unit affected by the breach & verify the source of same		
2	At the affected Air Handling Unit, CLOSE the Chilled Water Supply & Chilled Water Return manual isolation valves		
3	Shutdown the affected Air Handling Unit at VFD controller & remove from service via LOCK OUT/TAG OUT		
4	Visually verify that the adjacent companion Air Handling Unit has ramped up to accommodate the loss of air volume. Also verify with the FM Control Center that all redundant systems have responded as per design		
5	Upon verification of redundant systems, conduct notification per Escalation Process/ Policy		
6	Commence the clean-up operation		
7	Create an Incident Ticket		
8	Contact Support Services required to effect repair(s) of the damage/leak		
9	Perform root cause analysis to determine the source of the breach/leak to affect to modifications to prevent recurrence – document on Incident Ticket		
Restore Sequence			
1	Upon completion of repair(s) & verification of system integrity, restore the system to normal operating status by removing all system LOCK OUT/TAG OUT & Manual OVERRIDE conditions employed for containment		

Source: Uptime Institute Accredited Operations Specialist course materials 2017 UptimeInstitute

Figure 4. **A typical EOP**

Generally, EOPs are basic instructions intended not to restore normal operation (i.e., full redundancy) but to create a stable state so that IT load is not dropped as the result of an incident. The EOP is meant to prevent a cascade of events that might make matters worse. For example, the first steps in the sample EOP, which addresses concerns with an air handler, include isolating the unit, verifying redundant systems and following the escalation process.

Repairing the equipment comes later, allowing time for a root-cause analysis; restoring normal site configuration policies is the final step. EOPs should be geared toward system-condition verification and event notification and escalation, with physical intervention kept to the absolute minimum. All repairs conducted as part of an EOP should conform to a formal corrective maintenance procedure.

EOPs account for the differing configurations of different data centers, and related facility plans and site configuration policies should include up-to-date contact information and escalation plans, as well as an inventory of frequently replaced and hard-to-get equipment, even if SLAs require vendors to obtain parts. In this way, the process of restoring normal operations can begin as soon as practical after a root-cause analysis has been completed. Having up-to-date EOPs will help trained and qualified staff limit the duration of any incident, if and only if the EOPs are readily available to the staff and the staff is familiar with all the processes and equipment.

Automation

In Uptime Institute's 2019 supply-side industry survey, 42% of data center designers, consultants and suppliers said most of their customers are now deploying a data center infrastructure management (DCIM) system. The increasing use of DCIM and data center operations management (DCOM) software means that automation plays an important role in data center operations and a facility's procedures.

DCIM systems collect data, report on trends, analyze information, and assist with planning and forecasting. More pertinent to this report, they can also alert managers when problems arise and trigger automated or human-mediated corrective actions.

In the event of an outage, it is essential that staff is familiar with these notifications and automated processes. The DCIM system is designed to provide information about the current state of the IT operation even during an outage, which can help IT and facilities staffs identify and resolve the causes of the outage.

Facilities deploy DCOM software to manage their conventional day-to-day processes and SOPs. DCOM automates the kinds of functions and digitizes the kinds of data often found in computerized maintenance management systems. These include asset management information as well as maintenance scheduling and tracking, equipment maintenance policies, alerting and escalation procedures, description of interdependencies with other equipment, change management workflow and the root-cause analysis of incidents. Some of these functions are critical and may have to be performed under the pressure of imminent and sometimes catastrophic failure (such as restarting a cooling unit before overheating occurs); many are documented or referenced in an organization's MOPs, SOPs and EOPs.

The use of automation does not change the need to maintain comprehensive policies or train staff to follow procedures. For example, staff may find their DCOM system difficult to use, especially when security, multiple classes of operator and different asset databases are involved. In addition, users of DCIM and DCOM must be sure that alerting and notification systems will remain functional if the software should fail. For these reasons and more, maintaining and following appropriate procedures is fundamental to ensuring continuous availability.

Anticipating outage scenarios

As with all aspects of data center operation, even during an outage or service interruption, fortune favors the prepared. As discussed at a recent Inside Track roundtable, even data center operators who feel they have planned for an outage can still feel unprepared. As one participant asked, “What do we need to know that we don’t know?”

For some organizations, planning for an outage means having identified, considered and practiced a number of “what if” scenarios. Other methods may be considered, but the goals are the same: reduce the duration of an outage, restore essential services and minimize the cost of the disruption.

Typical “what if” scenarios include:

- What if we should experience an outage due to a power outage?
- What if should experience a network failure?
- What if our colocation provider should experience a service interruption?

These exercises, properly executed, provide a way to test the systems that will come into play during an outage, including escalation methodology, vendor responses, parts and fuel availability, staff experience and procedure viability. Most importantly, “what if” exercises will create visibility across the organization for the IT, facilities, network and management teams that will be working to restore services.

In many ways, these exercises test the organization in the same way as practicing EOPs, but with a major difference: IT and business units are simulating a race against time to restore services. “What if” scenarios help organizations anticipate accelerated spending levels for equipment replacement and labor, reduced on-site staff presence, SLA violations, and other worst-case scenarios. They extend an organization’s EOPs and test its level of preparedness.

Cost authorizations

EOPs normally describe steps to stabilize operations. Once stability has been achieved, operations will have a window for normal procurement, including budget submissions and approvals. In the event of an outage, though, emergency spending may be required. The cost of expensive items and the relevant purchase authorization chain should be researched in advance so the proper executive(s) can be quickly contacted for budget approvals during an outage (or the clearances organized in advance).

The escalation process must also anticipate increased labor costs and even service fees to vendors, all of which must be resolved quickly.

Other equipment replacement considerations

Operating staff must be careful not to view pre-approved spending as an opportunity to purchase wish list items. Despite pre-planned and approved spending, preferred equipment may not be available on short notice. In these instances, IT may have to purchase or rent the equipment that can be available on short notice, even if the equipment lacks some desirable features. Alternatively, an organization may choose to stock equipment because of concerns about availability or shipping time.

In the course of restoring normal site configuration, IT may also have to plan to uninstall or re-install rented equipment when preferred equipment becomes available.

Staff operations

The most critical consideration is communication, as even well thought-out plans do not work well if organizations do not communicate well. Most IT, networking and facilities teams have a good understanding of their roles in the event of an outage. However, without visibility across the organization, the different teams may struggle to restore availability.

Reliance on third-party providers can make the situation worse. “They promise a lot but don’t always deliver,” said one Uptime Institute Network member, who noted that their program team offered good ideas but that they “…work in a bubble,” isolated from the groups that had to implement the plans.

The involvement of third parties at times of failures can cause particular problems, especially where an outage is serious and costly. For reasons of their own, third parties, such as vendors or service companies, may instruct their employees to disregard procedures. These employees — and their employers — may also wish to avoid blame for any errors. Cooperation is best maintained by establishing clear responsibilities ahead of time and incorporating strict access and supervision policies within MOPs, SOPs and EOPs.

Returning to normal

Having restored service, much work remains to be done. Plans must be developed and executed to restore full redundancy, a root-cause analysis must be conducted to prevent a recurrence, and site configuration policies and procedures must be updated (especially if the replacement equipment includes new makes and models). Recommissioning may be desired as well.

At this phase, the root-cause analysis is especially critical — understanding the causes of the initial failure is a prerequisite to preventing its recurrence. Procedures should be refined to address the source of the outage. In addition, there may be lessons to be learned from mistakes made after the outage occurred. These lessons should be incorporated in the appropriate procedures.

The findings of these investigations may seem mundane but identifying vulnerabilities and addressing them are key to avoiding downtime. For example, one Network member learned the importance of creating a physical record of hot points in the facility only after experiencing an outage. These critical pieces of equipment are now marked in red.

If an outage is particularly serious (category 4 or 5), the outage is likely to attract media attention, and some parties affected may be seeking compensation or legal recourse. It is particularly important, therefore, that staff and third parties be trained in how to react should such a situation occur. It may advisable to instruct staff not to talk about the failure on social media or to journalists and the public.

About disaster recovery

Lessons gleaned from the BIA should be incorporated into a DR plan or overall business continuity or resiliency plan. A DR plan exists to help organizations restore full IT operations in the event of an extended outage, with the acceptable outage duration limits determined by the business needs of the organization. For some systems in some organizations, acceptable downtime may be measured in days. In other instances, there is no acceptable downtime and systems will require instantaneous failover.

Elements of a DR plan

The core elements of a DR plan are well known, all intended to meet an organization's infrastructure, operations and IT requirements in a minimal timeframe. Ideally, the DR plan should resemble facility procedures, comprising written step-by-step instructions for transferring IT load to a backup or secondary facility and then back again. In addition, staff — particularly IT staff — should be trained to these procedures, which should also be rehearsed on a regular basis and updated about every two years.

Significant resources need to be marshalled to implement a DR plan on almost any scale. For instantaneous failover, this can mean operating or leasing a dedicated backup site, using a mirroring architecture to ensure data reliability.

In other instances, organizations will lease or maintain dedicated infrastructure, fully provisioned with sufficient network, power and cooling to meet the anticipated IT load. Sometimes, but not always, these facilities may host some noncritical IT activities, but the DR plan will include provisions for purchase and installation of new servers to host critical applications, as well as procedures for installing software and deploying data backups.

Finally, there are DR vendors, whose services, approach, technology and capabilities vary widely. Some end users feel the cost of these services exceed their value, while others appreciate the experience these vendors provide. When considering this service, end users should establish that the DR vendor has adequate capabilities,

especially for situations in which the vendor may have to meet the needs of multiple clients.

The emergence of business continuity as a service and cloud-based disaster recovery as a service approaches, which involve failing over semi-automatically to remote cloud services, are a significant new development, as is the emergence of availability zones and distributed resiliency. But none of these advances change the requirement for detailed planning and development of step-by-step processes and emergency procedures.

DR readiness

In reality, the scale and difficulty of the task mean that organizations may not have a tested, proven DR plan to follow when it is needed. Organizations may develop a complete DR strategy but fear the consequences of a failed rehearsal — sometimes justifiably so, as an incident during testing could lead to a loss of availability. But without testing the DR plan, organizations risk deploying an unrehearsed or outdated plan in the midst of a true outage. In these instances, the DR plan may reveal itself to be simply an audit or checklist that includes provisions for infrastructure, utilities, staffing and space but no real process for ensuring that they can be readied for use in a timely fashion. Eventually, the organization may learn that its DR plan does not include adequate capacity, staff or space to meet its IT requirements.

The reluctance to test DR plans increases the chances that an outage will be prolonged. As with other procedures discussed in this report, DR tests are scheduled in ideal conditions, with necessary IT and facilities staff available and without the time pressure that may result from an imminent risk. By contrast, an untested DR plan is even less likely to work during an outage, when conditions will be less than ideal.

A new approach to availability

According to Uptime Institute research, approximately 87% of enterprises have IT assets distributed in and connected across more than one facility, in many cases spreading workloads across a mix of off-premises and privately owned on-premises capacity, in a distributed IT architecture.

The distributed IT architecture provides uninterrupted IT services, applications and data access, despite individual or multiple facilities or IT equipment disruptions or software failures.

More than half of the respondents (61%) to our 2018 survey said that that using a distributed architecture approach has made them more resilient and able to provide a higher level of availability. However,

about a third (31%) reported having suffered an outage in the past year, and almost one in ten (9%) stated that their hybrid approach has made them less resilient.

Failure redefined

Organizations utilizing a distributed resiliency environment may have a different mindset regarding equipment failure (and in theory, even the loss of an entire data center). These enterprises are effectively adopting a wide area active-active N+2 approach, with software components and data often replicated many times. Cushioned by an integrated hardware and software environment that doesn't fail, they may view individual failures as inevitable. In this environment, the failure of a server, rack, cluster or even entire data center is not necessarily a serious downtime incident. When resilient designs are utilized, IT services may be interrupted but service performance maintained, due to the inherent resiliency of the architecture.

Such architectures are likely be more resilient, especially when implemented at scale (enabling many nodes to put in place). Uptime Institute consultants have worked with a number of mature enterprises that operate in this way. In two such instances, the global enterprise operates three geographically distributed data centers, each of which can host all the organization's mission-critical IT functions without negatively impacting customer service availability or performance, even if the other two sites fail. In these instances, the loss of a single data center – even for a protracted period – is not considered an urgent problem.

However, experience also shows that systemwide failures do occur and a sudden loss of capacity, even if manageable, does often cause service problems. For this reason, most operators still strive for a high level of site availability and a low level of component failure.

Resiliency in distributed IT architectures should be regularly tested to verify that client services are unaffected when fail over or load balancing occurs between facilities or software platforms. Tests should be conducted in a controlled and recoverable manner and should include recovery of:

- Physical and logical components (data center and networking).
- Automated processes.
- Procedures.

Summary and conclusions

Much of the thinking about recovering from outages has centered around measures that should be taken after the outage, which evokes images of harried personnel working at high tempo, improvising when necessary.

Yet the hard work of recovering from an outage should take place long before an outage, with the necessary steps written into an organization's SOPs, MOPs and EOPs. Staff should be trained to these procedures, with regular practices. The procedures should anticipate incidents and incorporate steps to recovery, with regular updates as the IT environment changes. The procedures should also include an up-to-date escalation process, ensuring that the proper decisionmakers are aware of the problem and the correct technicians are on-site.

To augment these routine efforts, operations management should also conduct "what if" analyses, which simulate adverse conditions and probe for weakness in an organization's overall response.

Conducting a resiliency assessment, to determine where weaknesses may lie, and a BIA are both good ways for an organization to determine the operational importance of a facility and how to allocate resources. A BIA considers the likely cost of a downtime incident and the cost of restoring service, which provides a basis for matching resources to business requirements.

Uptime Institute® | INTELLIGENCE

ABOUT THE AUTHOR

Kevin Heslin is Chief Editor and Director of Ancillary Projects, Uptime Institute. Kevin served as editor of The Uptime Institute Journal, where he authored and co-authored numerous papers, including case studies involving Uptime Institute Tier-certified facilities. Kevin has had a more than 30-year career examining energy and sustainability issues from the end-user perspective. In that regard, he has worked in academic, nonprofit and standards organizations, in addition to serving as editor of various print publications. Contact: kheslin@uptimeinstitute.com

ABOUT UPTIME INSTITUTE INTELLIGENCE

Uptime Institute Intelligence is an independent unit of Uptime Institute dedicated to identifying, analyzing and explaining the trends, technologies, operational practices and changing business models of the mission-critical infrastructure industry. For more about Uptime Institute Intelligence, visit uptimeinstitute.com/ui-intelligence.

ABOUT UPTIME INSTITUTE

Uptime Institute is an unbiased advisory organization focused on improving the performance, efficiency and reliability of business critical infrastructure through innovation, collaboration and independent certifications. Uptime Institute serves all stakeholders responsible for IT service availability through industry leading standards, education, peer-to-peer networking, consulting and award programs delivered to enterprise organizations and third-party operators, manufacturers and providers. Uptime Institute is recognized globally for the creation and administration of the Tier Standards and Certifications for Data Center Design, Construction and Operations, along with its Management & Operations (M&O) Stamp of Approval, FORCSS® methodology and Efficient IT Stamp of Approval.

Uptime Institute – The Global Data Center Authority®, a division of The 451 Group, has office locations in the US, Mexico, Costa Rica, Brazil, UK, Spain, UAE, Russia, Taiwan, Singapore and Malaysia. Visit uptimeinstitute.com for more information.

All general queries:
Uptime Institute
5470 Shilshole Avenue NW, Suite 500
Seattle, WA 98107 USA
+1 206 783 0510
info@uptimeinstitute.com

