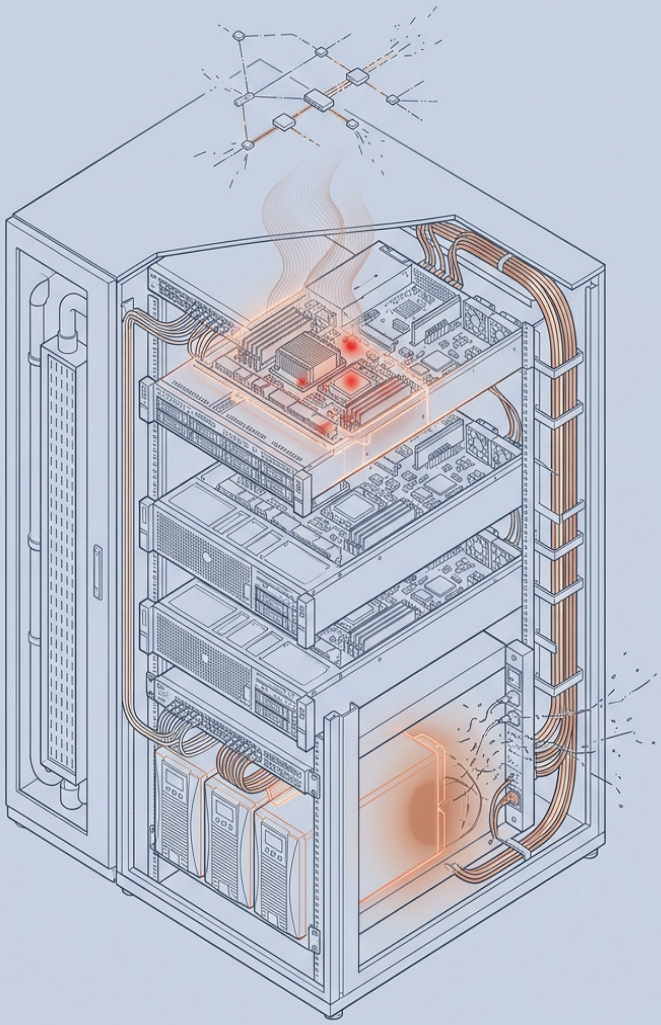
Bjumper

When the Data Center Stops

Anatomy of an outage: causes, impact chain, and how to prevent human error, fire, or a security breach from halting your critical infrastructure



Introduction

According to data collected in the industry, 25% of unplanned outages in data centers are caused by human error. Another 25% by UPS system failures. 12% by cyberattacks. And only 6% by weather-related causes. The message is clear: the majority of outages originate from factors that can be anticipated, designed for, and managed.

This e-book analyzes the anatomy of an outage from its root causes to its impact chain in the critical environment. It is aimed at professionals who already know data center infrastructure and want to understand why systems that were designed not to fail, on paper, do fail.

What you will find in this e-book

- What defines a mission-critical system.
- The 6 main causes of unplanned outages.
- The impact chain: from component to critical environment.
- Human error as a risk vector.
- Fire protection: detection, suppression and hypoxic systems.
- Physical and operational security.
- Disaster recovery: cold, warm and hot site.

What ‘mission critical’ really means

The term ‘mission critical’ is used too loosely in the IT industry. From the operational standpoint of a data center, it has a precise definition: a mission-critical system is one whose failure directly impacts life, safety, or the continuous operation of the business.

This definition has a direct implication for how design and operational decisions must be made: mission-critical operations are driven by the level of availability and the level of control and security, never by cost. If a data center is cost-driven, by definition it is not mission critical..

Real-world impact examples

To understand the magnitude of the concept, it is enough to consider specific infrastructure failure scenarios:

- A hospital without access to its electronic health records system or radiology data may make incorrect medical decisions, with potentially lethal consequences.
- A bank without active systems during operating hours can generate chaos at ATMs and branches, with direct financial losses and severe reputational damage.
- An air traffic control tower without real-time radar data can trigger a disaster of incalculable proportions.
- A telecommunications operator without an active mobile network incurs contractual and operational losses that accumulate by the minute of interruption.

The real cost of an outage

The recovery time following an infrastructure failure at a single site is estimated at between 1.25 and 5.75 hours in total, including site restoration, hardware restart, data/software reconstruction and telecommunications restoration. Each segment of that time carries a specific operational, contractual and reputational cost.

Availability: the indicator that measures everything

A data center's availability is not just a marketing figure from the Uptime Institute. It is the result of design decisions, investment in redundancy and, above all, the quality of daily operations. Design can be defeated by operations: a system designed for Tier IV can operate like Tier II if procedures are deficient or staff are not trained.

Tier	Availability	Max. downtime/ year	Redundancy
Tier I	99,671%	28,8 hours	No redundancy
Tier II	99,741%	22,7 hours	Partial redundancy
Tier III	99,982%	1,6 hours	Concurrent maintenance
Tier IV	99,995%	26,3 minutes	Fault tolerant

Bjumper

Anatomy of an outage: the 6 main causes

Understanding why data centers fail is the first step to preventing it. Industry data consistently identifies six main categories of causes of unplanned outages. None of them is inevitable: all can be significantly reduced with proper design, training and procedures.

Cause	% de outages	Primary origin
UPS system failures	25%	Design, maintenance, age
Human error	25%	Procedures, training, fatigue
Cooling system failures	22%	Design, maintenance
Cyberattacks / Cybersecurity	12%	Vulnerabilities, ransomware
Weather conditions	10%	Location, civil design
On-site generation failures	6%	Maintenance, fuel

1.- UPS System Failures (25%)

The UPS is the last line of defense for the critical load against any electrical disturbance. Its failure means that, at the moment the power grid fails, the critical load is left unprotected. The most frequent causes of UPS system failure are:

- Batteries at end of useful life without having been replaced in time..
- Overload beyond design capacity.•
- Bypass system failure during maintenance.
- Incorrectly implemented redundancy configurations.
- Lack of periodic transfer tests under real load.

Operational principle

A UPS that is not periodically tested under real load is a UPS that cannot be trusted. Simulated transfer tests do not replace full real-load tests.

2.- Human Error (25%), the most avoidable risk

Human error is statistically the most ‘democratic’ cause of outages: it affects high and low technical sophistication facilities equally. It is not a matter of incompetence, but of the absence of formalized procedures, insufficient training or operational pressure that leads to skipping critical steps.

The most common vectors of human error in data centers are:

- Accidental activation of the EPO (Emergency Power Off): a historical cause of complete shutdowns due to accidental sabotage or error during maintenance.

- Disconnection of wrong circuits during interventions.
- Configuration changes without a change management procedure.
- Maintenance carried out without coordination with operations.
- Working without proper protective equipment (ESD, insulated tools).

The EPO: the most dangerous button in the data center

The Emergency Power Off system is one of the main causes of downtime in data centers, due to employee sabotage, maintenance errors or accidental activations. Best practices require three-stage EPO systems (Off / Test / Armed), minimum dual simultaneous button activation, security cameras at each EPO station and audible and visible pre-alarm when the cover is lifted.

3.- Cooling Failures (22%)

Heat is the second operational enemy of the data center. A cooling system failure does not cause an immediate shutdown, but rather a progressive degradation that ends in equipment shutdown due to temperature or permanent hardware damage. The failure chain is particularly dangerous because it can be slow and invisible until it is too late.

The most frequent causes include failure of CRAC/CRAH units without active redundancy, loss of chilled water in Chiller systems, blockage of perforated panels in the raised floor and mixing of hot and cold air due to poor airflow management.

4.- Cybersecurity (12%)

Cyberattacks as a cause of outage are a steadily growing category. Ransomware has proven capable of paralyzing entire infrastructures, including the control and monitoring systems of the data center itself (BMS, DCIM). The convergence of IT and OT in critical infrastructure management significantly expands the attack surface.

IT/OT Convergence: the new risk front

Building Management System (BMS) and DCIM systems that control power, cooling and physical access are increasingly accessible from corporate networks. An attack that compromises these systems can shut down a data center without touching a single server.

5.- Weather Conditions (10%)

Natural phenomena are the cause of outage with the greatest uncertainty. Unlike human error or UPS failure, meteorological and seismic events belong to the category of unknown events: they cannot be predicted with accuracy, and their impact can affect a large area simultaneously. Disaster preparedness is, for this reason, a critical necessity in any Data Center operation.

The most frequent outage causes related to environmental conditions are:

- Flooding affecting electrical rooms, battery rooms or generator rooms.
- Lightning storms generating surges in the external power supply.
- Extreme temperatures that saturate the capacity of cooling systems.
- Seismic activity that damages civil infrastructure or critical systems.
- Strong winds that compromise outdoor equipment such as generators or cooling towers.

Bjumper

6.- On-site Generation Failures (6%)

The lowest percentage in the table should not be misleading: when the generator fails, there is no margin for error. Diesel generators are the controlled local power source that covers the entire Data Center load during an external supply outage. They are currently considered the most reliable power source for the continuous operation of the facility, precisely because they do not depend on the external electricity grid. But they are mechanical systems, and as such, they can fail.

The most frequent causes of on-site generation failure are:

- The generator fails to start at the critical moment.
- Lack of fuel or deterioration of stored diesel.
- Insufficient maintenance of the engine or the generator's own cooling system.
- Failure of the Automatic Transfer Switch (ATS) between the grid and the generator.
- Use of Standby-type generators, designed for a maximum of 200 hours per year, in facilities that require continuous operation.

Data Centers require Prime or Continuous type generators, both prepared to operate for an unlimited number of hours per year and to support a variable electrical load. Using the wrong type is a decision that gets paid for at the worst possible moment.

The impact chain: from component to critical environment

One of the most important and least intuitive concepts in data center operations is that the degradation of an individual component rarely causes an immediate failure of the critical environment. What causes the failure is the chain reaction triggered by that first failure, especially when redundancy systems are not properly sized or maintained.

The four levels of impact

The impact model in a data center operates at four hierarchical levels, from least to most severe:

Level	Scope	Service impact	Severity
Level 1	Component or distribution path	None if active redundancy	Low
Level 2	Complete system (A or B)	Reduction of safety margin	Medium
Level 3	Partial critical environment	Service degradation	High
Level 4	Complete critical environment	Total outage	Critical

The key is that redundant systems act as firewalls between levels.
A Level 1 failure only escalates to Level 2 if the redundant system also fails or is unavailable.
A Level 2 failure only escalates to Level 3 or 4 if there are no operational backup systems.

The weakest link in the chain

In any availability analysis, the point that determines the real resilience of the system is not the most robust component, but the weakest one. In practice, the site infrastructure (building, external power supply, telecommunications connections) is systematically the weakest link in the end-to-end chain, above the servers themselves or the software..

Murphy's Law applied to data centers

'If there is more than one way to do a job, and one of those ways will result in disaster, then someone will do it that way.' — E. Murphy. High availability is not a state achieved through design: it is a state maintained by actively eliminating Single Points of Failure (SPOF) in operations, not just on the design blueprint.

Essential systems vs. redundant systems

The distinction between essential systems and redundant systems is fundamental to understanding the logic of availability:

- **Essential systems:** their function is to keep operations active continuously. If an essential system fails without the redundant system taking over, the critical environment shuts down IMMEDIATELY. Examples: the active UPS, the CRAC unit in service, the generator in operation.

- **Redundant systems:** their function is to absorb the failure of essential systems without impact on service. They do not perform operational work under normal conditions, but they are as critical as the essential ones because they determine whether a Level 1 failure escalates or not.

The most frequent management error

Treating redundant systems as secondary in maintenance programs. A redundant system that is not tested and maintained with the same rigor as the essential system is, at best, a false guarantee. At worst, it is the cause of the next catastrophic outage.

Fire protection: the system that should never have to activate

The fire protection system is, by definition, the system we hope never to have to use. However, its design, maintenance and correct implementation are as critical as the electrical or cooling system. An uncontrolled fire in a data center can be an irreversible cause of outage.

The fire triangle and its application in the DC

For a fire to occur, three factors must coincide simultaneously: combustible material, oxygen (concentration equal to or greater than 16%) and sufficient ignition energy. Eliminating any one of the three prevents or extinguishes the fire. This logic defines the two protection strategies:

- Suppression: act on an already declared fire, removing oxygen or heat.

- Hypoxic prevention: preventively eliminate the possibility of ignition by reducing the oxygen level below the combustion threshold.

The three components of the protection system

Component	Type	Function	Typical technology in DC
Detection	Active	Identify fire at incipient stage	Aspirating early detection, ionization/photoelectric
Passive protection	Passive	Prevent fire spread	FR walls 1-2h, fire seals, FR doors
Supresión Suppression	Active	Combat and extinguish fire	Clean gaseous agents, pre-action

Gaseous agents: why not every gas is valid in a DC

Water-based (wet) suppression systems are not suitable for server rooms. Water can permanently damage hardware that the fire did not affect. Pre-action systems (dry pipes until activation) are the minimum acceptable for mission-critical areas. But the preferred option is clean gaseous agents under NFPA 2001.

Two families of agents are prohibited in data centers:

- Halons: prohibited due to their ozone-depleting potential (Montreal Protocol).
- CO_2: must not be used in areas occupied by personnel; it is also a greenhouse gas.

Agent	Type	ODP	Suppression mechanism
HFC-227ea (FM-200)	Chemical	Zero	Removes heat and interrupts chemical reaction
FK-5-1-12 (Novec 1230)	Chemical	Zero	Removes heat, extremely low GWP
HFC-125 (ECARO-25)	Chemical	Zero	Removes heat, compatible with halon installations
IG-541 (Inergen)	Inert	Zero	Reduces oxygen to hypoxic level (< 15%)
IG-55 (Pro-Inert)	Inert	Zero	Reduces oxygen, natural atmospheric gases

Hypoxic systems: prevention before suppression

The most advanced technology in fire protection for data centers is not suppression but prevention. Hypoxic systems continuously maintain the oxygen level in the room below 15%, the threshold below which ignition is impossible. Normal atmosphere contains 21% oxygen; the system continuously injects nitrogen to reduce and maintain that level.

Safety limit for personnel

The preventive oxygen level at 15% is equivalent to an altitude of 2,450 meters above sea level, with no adverse effects for personnel. Below 12%, fatigue begins. Below 10%, there is a risk of unconsciousness and death. Hypoxic systems operating in data centers are maintained in the 15%-16% range, above the medical safety threshold, but below the ignition threshold.

Physical and operational security: the three layers of protection

Data center security is not a single system: it is the integration of three layers that must work in a coordinated manner. A failure in any of them can compromise the availability of the critical environment, both through unauthorized access and through involuntary or intentional sabotage.

Layer 1: Physical architecture

The architectural design of the data center must implement two security principles from the blueprint:

- **Natural access control:** the flow of people must be designed so that transit to high-security areas is always deliberate, never accidental.

- **Natural surveillance:** all entrances to high-security areas must be visible both physically and through video surveillance systems.

The server room must be the most protected room in the facility. Access to it must require passing through at least two independent security zones: the lobby/vestibule and the secure corridor or mantrap. The mantrap (person airlock) ensures that only one person can enter per cycle, eliminating tailgating.

Layer 2: Electronic systems

System	Function	Minimum requirement
Electronic access control	Identificar y registrar cada acceso	ID único por persona, registro de hora/identidad (mecanismo AAA)
Security doors	Physical passage control	Fail-safe lock, panic bar, door-open-for->3-min sensor
Intrusion alarms	Detect unauthorized access	Electrical circuit, light beam, vibration, capacitance
24/7 Video surveillance	Monitor all areas	Server room, corridors, doors, plenum, telecommunications
Security lighting	Eliminate blind spots	Perimeter, doors, windows, stairs, public areas

Layer 3: Operational procedures

The most critical and most frequently neglected layer is the operational one. A state-of-the-art electronic security system is useless if procedures are not formalized, updated and followed by all stakeholders.

The Data Center Security Plan must cover at minimum:

- General security policies (alarms, surveillance, personnel).
- Asset protection policies: personnel, physical assets and data.
- Continuous review process for new threats.

Disaster Recovery Plan (DRP)

The DRP is the document that determines how the organization responds when design and prevention have not been sufficient. It must cover four sequential phases: pre-disaster activities (testing and preparation), emergency activities (immediate response), response activities (containment and stabilization) and recovery activities (return to normal operation).

Backup model	Operational capacity	Activation time	Relative cost
Cold Site	Data and software storage. No DC operations	Hours to days	Low
Warm Site	Partial capacity. Does not support all systems	Hours	Medium
Hot Site	Full operation. Can assume entire load	Minutes	High

Conclusion: The outage that didn't happen

The success metric of a well-operated data center is not the speed of recovery from an outage: it is the absence of outages. Every unplanned interruption that does not occur is the invisible result of correct decisions in design, maintenance, training and procedures.

Industry data is clear: 50% of outages originate in the UPS and human error. Both are completely manageable with the right tools. The question is not whether your data center can fail, but whether you have actively eliminated the most probable causes of it doing so.

The five principles of the resilient data center

1. Design does not guarantee availability: operations do. A Tier IV operated with deficient procedures is less reliable than a well-managed Tier III.

2. Redundant systems must be maintained with the same rigor as essential systems.

3. The EPO is the most dangerous button in the facility: it must have three activation stages and cameras at each station.

4. The DRP must be tested, not just documented. An untested plan is a plan that will fail.

5. Physical security and cybersecurity are the same discipline: unauthorized access to the BMS can shut down the data center without touching the hardware.