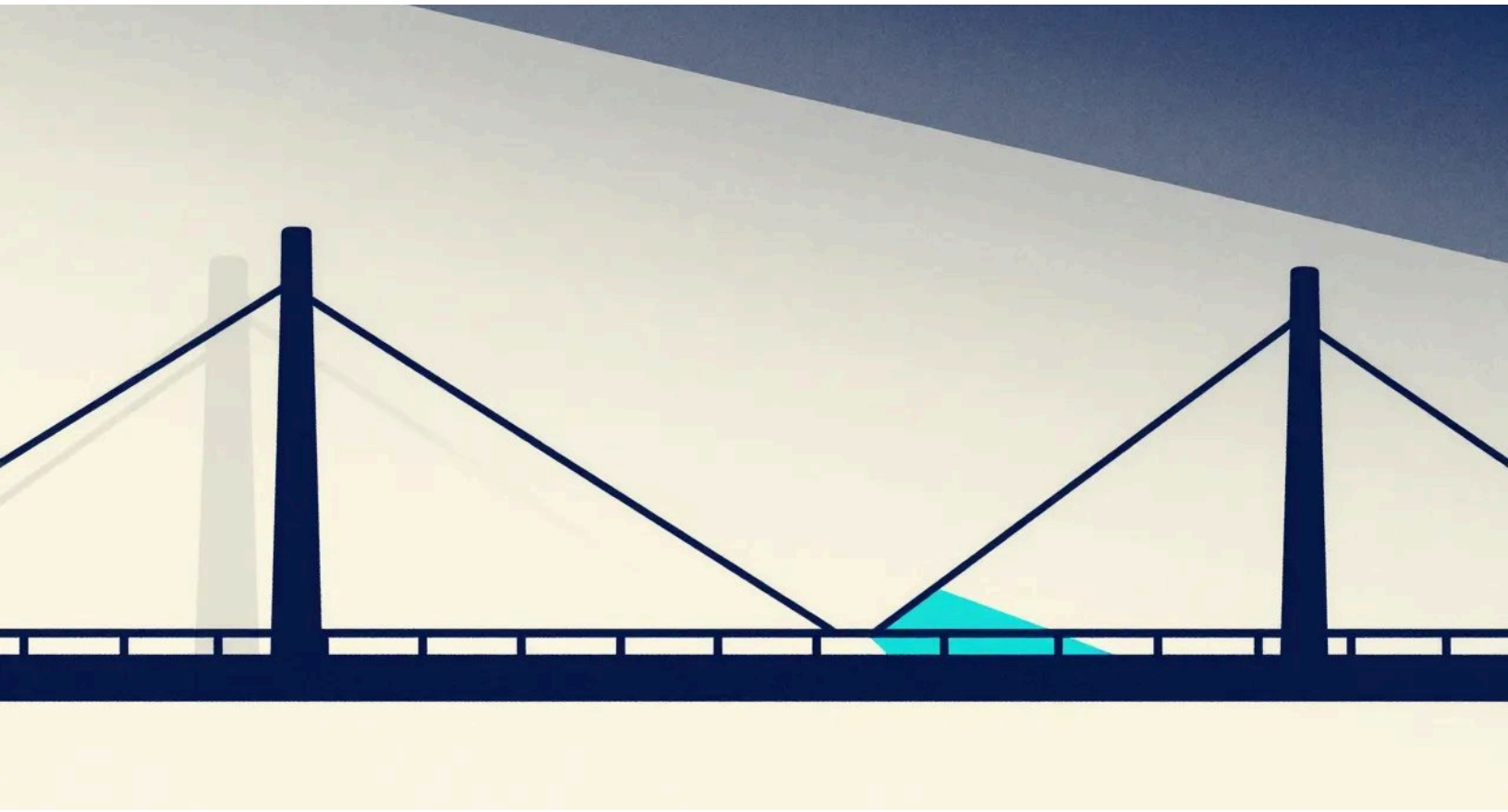


EXECUTIVE PAPER · DISASTER RECOVERY

Build High Availability Before Disaster Recovery

Most outages start with one failed component, one zone or one bad change. Fund the architecture that absorbs those every month, then size disaster recovery to the risk that is left.

AUTHOR

Chris Watkins

SERVICE

Disaster Recovery

FOR

CTO, CFO · 12 min read

AT A GLANCE

The problem. Companies often buy a disaster recovery site before they remove the single points of failure in their primary environment. The secondary site waits for a rare catastrophe while ordinary failures of a server, a database or a cloud zone keep taking revenue offline.

Why it persists. Auditors, insurers and boards ask whether a disaster recovery plan exists, and a plan is easy to show. High availability is harder to see. It lives in design choices that nobody reviews until the day they fail.

The recommendation. Rank failures by how often they happen and how far they spread. Fund high availability for the frequent, local ones first: redundant components, at least two zones, automatic failover and changes rolled out one zone at a time. Then size disaster recovery to the site-wide events that remain, with immutable backups in a separate location as the floor for every critical system.

The Decision in Front of You

Every resilience budget forces a choice about what to protect against first: losing the whole primary site, or losing one piece of it.

Protection against each pays back on a different schedule. High availability, meaning a system built with enough redundancy to keep serving through a component or zone failure without human intervention, earns its keep whenever a disk, host, switch or zone fails. Disaster recovery, meaning the ability to restore service in a different location after the primary one is lost, earns its keep in the rare event that the whole site or region goes.

The money at stake is material. In Uptime Institute's 2026 outage analysis, 57% of operators said their most recent major outage cost more than $100,000, and one in five put it above $1 million for the second year running [1].

What's Happening

Most outages are local

Uptime's 2026 report found power caused 45% of the most recent impactful outages in 2025, led by failures of uninterruptible power supplies, transfer switches and generators [2]. Network problems are the most frequent cause of IT service outages, and configuration and change management remain the main source of those [2]. Within IT systems, firmware and software faults now account for the largest share [2]. A year earlier, IT and networking issues had risen to 23% of impactful outages, with change management and misconfiguration the likely cause [10].

These are local failures: a UPS in one room, a misconfigured switch, a database host that loses a disk. Redundancy and automatic failover are designed to absorb exactly these.

Uptime also found human error contributed to nine of ten outages, and 87% of operators with an impactful outage said better management, processes or configuration could have prevented it [2]. In the 2025 analysis, 85% of human-error outages traced to staff not following procedures or to flawed procedures [10].

The standards already say so

NIST SP 800-34 Rev. 1, the federal contingency planning guide, scopes a disaster recovery plan narrowly: it covers information system disruptions that require relocation to an alternate site [3]. The same guide says strategies for high-impact systems should consider high availability and redundancy in their design, and that preventive measures are preferable to recovery actions where they are feasible and cost-effective [3].

ISO 22301:2019, the international standard for business continuity management systems, frames continuity as reducing the likelihood of disruption as well as recovering from it [4]. Amazon's Well-Architected guidance tells customers to run every production workload in at least two availability zones, separate data centers with independent power, cooling and networking within one region. It rates single-zone production as high risk and reserves multi-region designs for workloads that need extreme resilience [5].

Two 2026 cases show the difference

A single zone overheats, May 2026. On May 7, 2026, multiple cooling units failed in one AWS data hall in Northern Virginia, and racks shut down to protect themselves [6]. Coinbase lost trading for several hours, and full recovery took much of the following day [6]. Its postmortem found the matching engine ran as a five-node cluster packed into one placement group to cut latency. The thermal event took three of the five nodes, the cluster lost quorum, and the design had no automated failover to another zone [6]. Coverage of the postmortem gives no cost figure.

The lesson is about high availability. The failure was confined to one zone, and the customer's architecture had concentrated a critical system inside it. Coinbase's remediation list starts with automated cross-zone recovery and adds expanded disaster recovery testing [6].

A region is struck, March 2026. On March 1, 2026, objects struck an AWS facility in the United Arab Emirates during the conflict between the United States and Iran, causing a fire. Firefighters cut power to the building and its generators, and a second zone in the region also lost power [7][8]. Months later, AWS said it still could not restore resources and data held only in the struck zone, and it advised customers to move to other regions; most did so by restoring backups or copying data that remained reachable [8].

This is the event disaster recovery exists for. Multi-zone design could not save a workload when two zones in one region failed together, and data kept only in the struck zone was still out of reach months later. Copies in another region let customers resume.

Software can take a region down too: the AWS DynamoDB DNS failure of October 2025 hit most services in Northern Virginia for up to 15 hours [9].

SIDEBAR: THREE TERMS EXECUTIVES CONFUSE

High availability keeps a service running through a component or zone failure, with no human action and minutes of disruption at most. **Disaster recovery** restores service in another location after a declared event, usually within hours. **Backup** is a separate, preferably immutable copy of data that protects against corruption, deletion and ransomware. High availability replicates mistakes as fast as good data, so it never replaces backups.

Why the Usual Approach Fails

Disaster recovery is easy to buy and easy to audit. A secondary site contract and a binder of runbooks are evidence that a plan exists, and auditors and insurers ask for that evidence, so it gets funded.

High availability gets little scrutiny. Few audits check whether the order database has a replica in a second zone. Engineers make those choices under delivery pressure, and the cheap option is a single instance in a single zone.

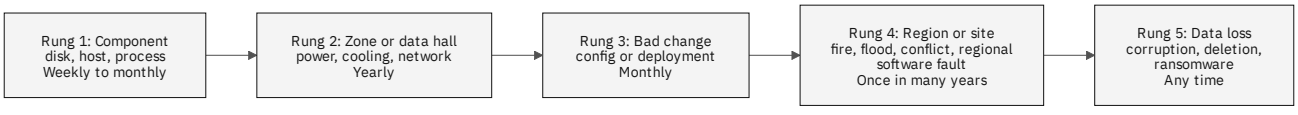
The company ends up owning a secondary site it has never failed over to under real conditions, while single points of failure in the primary environment cause most of its outages. When a local failure hits, the team will not declare a disaster for a problem that might be fixed in an hour, so the DR investment sits unused while the outage runs for three.

Failover that runs only during an annual test also decays between tests. Automatic failover within a region runs whenever something breaks, so the team learns quickly whether it works.

The Framework: The Blast Radius Ladder

Rank each failure by how often it happens and how far it spreads. Fund protection rung by rung, and stop climbing where the next rung costs more than the risk it removes.

Exhibit 1. The Blast Radius Ladder: the frequent, contained failures come first, where protection pays back fastest



Source: author's framework. Frequencies are indicative for a mid-size organization and should be replaced with your own incident history. Cause categories from Uptime Institute [2]; zone and region guidance from AWS [5].

Rung 1, components. Run at least two of everything that serves customers, behind a load balancer or a clustered database with automatic failover.

Rung 2, zones. Spread each critical service across at least two zones or two independent rooms, with data replicated synchronously [5] and failover tested monthly. Coinbase shows the failure mode: redundancy that lives inside one zone counts as rung 1 protection [6].

Rung 3, changes. Roll changes out one zone at a time with an automatic halt and rollback. Uptime's data puts configuration and change at the center of network outages [2].

Rung 4, regions and sites. Here disaster recovery begins. Choose the cheapest pattern that meets the recovery time objective, or RTO, the maximum tolerable time to restore service, which NIST ties to the business's maximum tolerable downtime [3]. Many systems need only a pilot light, a minimal copy of the environment that scales up when needed. A few need a warm standby, a smaller running copy ready to take traffic [5].

Rung 5, data. Immutable backups in a separate location apply to every critical system regardless of the other rungs. After the UAE strike, restoring backups is how most affected customers moved to other regions [8].

What the ladder is worth

The following is an illustrative model for a hypothetical distributor whose order platform brings in $120,000 per hour. Substitute your own figures.

Exhibit 2. Illustrative annual value: high availability pays for itself on frequent failures; DR value depends on how it is sized

Assumption or result	High availability (rungs 1 to 3)	Warm-standby DR	Pilot-light DR
Events covered per year	4 local outages	0.05 regional outages (one in 20 years)	0.05 regional outages
Hours down per event, without protection	3	72	72
Hours down per event, with protection	0.17 (10 minutes)	4	12
Expected annual loss without protection	$1,440,000	$432,000	$432,000
Expected annual loss with protection	$80,000	$24,000	$72,000
Annual loss avoided	$1,360,000	$408,000	$360,000
Annual cost of protection	$300,000	$400,000	$150,000
Net annual value	$1,060,000	$8,000	$210,000

Source: author's illustrative estimate. Assumes $120,000 of revenue lost per hour of outage, four local outages a year averaging three hours without automatic failover, and one regional loss every 20 years lasting 72 hours without a prepared recovery site. Costs include infrastructure and engineering time. Excludes data-loss, regulatory and reputational costs.

High availability returns its cost several times over on events that happen every year. Disaster recovery value swings with the probability of a regional event and the pattern chosen, so it has to be sized deliberately. Permanent data loss cannot be priced per hour, which is why rung 5 applies to every critical system.

Costs, Risks, and Tradeoffs

Exhibit 3. Tradeoffs of funding high availability first

Area	Cost or risk	Mitigation
Infrastructure spend	Running in two zones roughly doubles capacity for stateful tiers and adds cross-zone data charges	Apply it to services ranked critical by impact analysis; leave internal tools in one zone
Complexity	Distributed designs add synchronization complexity and can hide new failure modes, as Uptime notes [2]	Prefer managed multi-zone databases; test failover monthly in production hours
Correlated failure	HA within one region shares regional dependencies, as the October 2025 DynamoDB event showed [9]	Keep rung 4 and rung 5 sized to the remaining risk; hold backups outside the region
Replicated mistakes	HA copies corruption and bad deletes to every replica within seconds	Immutable, versioned backups with restore tests; keep them outside the reach of production credentials
Audit expectations	Some auditors look for a DR plan first	Document HA as a preventive control under NIST SP 800-34 [3] and ISO 22301 [4]
When this is wrong	A single on-premises site in a flood, hurricane or conflict zone, or a regulated firm with a mandated recovery time, may face site loss as its dominant risk	Fund rung 4 and rung 5 first, then add zone-level redundancy

Source: author's analysis.

How to Act

Exhibit 4. A one-year roadmap

Window	Actions	Owner	Milestone
First 30 days	List the ten systems with the highest hourly cost of downtime, with their single points of failure and zone count. Tag the last two years of incidents by rung.	CTO with the CFO for downtime costs	Ranked list approved; incident history tagged by rung
First 30 days	Confirm every critical system has an immutable backup held outside the primary region and outside production credentials.	Infrastructure lead	Restore test passed for the top five systems
First 90 days	Remove single points of failure in the top five systems. Move stateful tiers to multi-zone configurations with automatic failover. Stage all deployments one zone at a time with automatic rollback.	Engineering and platform leads	Top five systems survive a planned zone shutdown during business hours
First 90 days	Run a business impact analysis to set RTO and recovery point objective per system [3].	Business continuity owner with business unit leaders	Signed RTO and RPO for each of the ten systems
First 365 days	Choose a DR pattern per system sized to its RTO: backup and restore, pilot light or warm standby. Fail each one over at least once. Report failover results quarterly.	CTO with risk or compliance	Every critical system meets its RTO in a recorded test; local incident hours down 50% against the baseline

Source: author's recommendation.

Questions to Ask Your Team

1. Of our last ten outages, how many would a second zone or a redundant component have prevented, and how many needed a second region?
2. Which critical systems still run as a single instance or in a single zone?
3. When did we last shut down a zone on purpose, and what happened?
4. Do our deployments reach every zone at once, or one at a time?
5. What does an hour of downtime on each critical system cost, and who agreed to that number?
6. How many times have we used our DR capacity for a real event?
7. Could an attacker with production credentials delete our backups?

Closing

Disaster recovery covers the day the building or region is gone. High availability covers the far more common day when a server, a zone or a change fails. Spend first on the architecture that absorbs frequent local failures, because it pays back every year. Then buy disaster recovery sized to the regional risk that remains, and keep immutable backups under every critical system regardless.

Sources

[1] Uptime Institute. "Uptime Announces Annual Outage Analysis Report 2026." Press release via Business Wire. 2026. https://www.businesswire.com/news/home/20260513289344/en/Uptime-Announces-Annual-Outage-Analysis-Report-2026.

[2] Uptime Institute. "Annual Outage Analysis 2026." 2026. https://datacenter.uptimeinstitute.com/rs/711-RIA-145/images/2026.AnnualOutageAnalysis.pdf.

[3] National Institute of Standards and Technology. "Contingency Planning Guide for Federal Information Systems." NIST SP 800-34 Rev. 1. 2010. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-34r1.pdf.

[4] International Organization for Standardization. "ISO 22301:2019 Security and resilience: Business continuity management systems: Requirements." Edition 2. 2019. https://iso.org/standard/75106.html.

[5] Amazon Web Services. "REL10-BP01 Deploy the workload to multiple locations." AWS Well-Architected Framework, Reliability Pillar. https://docs.aws.amazon.com/wellarchitected/latest/framework/rel_fault_isolation_multiaz_region_system.html.

[6] InfoQ. "Coinbase AWS failure postmortem." Coverage of Coinbase's published postmortem. 2026. https://infoq.com/news/2026/06/coinbase-aws-failure-postmortem.

[7] DatacenterDynamics. "AWS UAE outage after objects struck the data center, cause fire amid Iran attacks." 2026. https://www.datacenterdynamics.com/en/news/aws-uae-outage-after-objects-struck-the-data-center-cause-fire-amid-iran-attacks/.

[8] W.Media. "AWS can't recover resources and data stored in UAE region hit during US-Iran conflict." 2026. https://w.media/aws-cant-recover-resources-and-data-stored-in-uae-region-hit-during-us-iran-conflict/.

[9] InfoQ. "AWS DynamoDB outage postmortem." Coverage of the AWS service disruption summary. 2025. https://www.infoq.com/news/2025/11/aws-dynamodb-outage-postmortem.

[10] Uptime Institute. "Uptime Announces Annual Outage Analysis Report 2025." Press release via Business Wire. 2025. https://www.businesswire.com/news/home/20250506733553/en/Uptime-Announces-Annual-Outage-Analysis-Report-2025.

The views in this paper are the author's own and do not represent any employer. Scenarios are hypothetical. This paper is not legal, financial, or compliance advice.