Move2Cloud

WHITE PAPER — 2027 EDITION

RESILIENCE & BUSINESS CONTINUITY

Cloud Business Continuity
Designing for the Loss of a Region

The physical loss of a cloud region is no longer a purely theoretical scenario. This guide details the resilience tiers, architectural patterns and cost trade-offs needed to prepare for it concretely.

11 pages · Comparison of recovery patterns, self-assessment checklist and roadmap included

$5,600
average cost of one minute of downtime for a company (IBM)

37%
of organisations cannot meet their target recovery time

1/5
of backups turn out to be unusable once tested

PREPARED BY
Move2Cloud — Cloud & Infrastructure Consulting, Paris

PUBLICATION
September 2026

CONTENTS

What you will find in this document

A six-stage journey, from a real-world case study to an operational business continuity roadmap.

01 **Executive summary — the essentials on one page** p.3

02 **The case study: physically losing a cloud region** p.4

03 **RTO, RPO and the four continuity tiers** p.5

04 **Four architectural recovery patterns** p.6

05 **The cost / resilience trade-off** p.7

06 **Designing beyond a single region: best practices** p.8

07 **Self-assessment grid** p.9

08 **A 5-step roadmap** p.10

09 **Conclusion & about Move2Cloud** p.11

Who is this document for? CIOs, CISOs and infrastructure managers whose critical operations depend on a single cloud region, and who want an objective view of their real level of resilience.

Copyright Move2Cloud — move2cloud.fr

EXECUTIVE SUMMARY

The essentials on one page

For a long time, the total loss of a cloud region was treated as a textbook scenario — unlikely, and therefore secondary in continuity priorities.

The events of March 2026 changed that equation. In March 2026, physical damage inflicted on several data centres of a major cloud provider in the Middle East caused the loss of several availability zones and the degradation of **more than a hundred services** across an entire region (details in section 02). This event is a reminder that a region loss can have far more varied causes than a simple software failure: natural disaster, fire, prolonged power outage, or — as here — direct physical damage to the infrastructure.

Yet, according to the available industry data, preparedness remains largely insufficient: **37% of organisations state that they cannot meet their target recovery time** for lack of complete or tested backups, and **one backup in five proves unusable** at the moment it is actually tested.

The cost of inaction, on the other hand, is measurable: downtime costs a company an average of **5,600 dollars per minute**, across all sectors.

$5,600
average cost of one minute of downtime for a company (IBM)

37%
of organisations do not meet their target recovery time (RTO)

1/5
of backups are unusable when they are actually tested

For any organisation whose business depends on a single cloud region — still the most common configuration — the central question is no longer "can a region be lost?" but "how long could we operate without it, and with how much data loss?". This document offers a framework for answering it.

SECTION 02

The case study: physically losing a cloud region

In March 2026, several AWS data centres in the United Arab Emirates and Bahrain suffered direct physical damage in the context of the ongoing regional conflict — a scenario few continuity plans had anticipated in this form.

1. 1 March, ~4:30 a.m. (PST)

A first strike damages an availability zone in the ME-CENTRAL-1 region (United Arab Emirates), causing structural damage and cutting off the power supply.

2. 2 March, 4:19 p.m. (PST)

AWS confirms damage to several facilities, affecting both ME-CENTRAL-1 and the neighbouring ME-SOUTH-1 region (Bahrain). The triggering of the fire protection systems causes additional water damage.

3. 3 March, 8:14 a.m. (PST)

Recovery operations continue; storage services show gradual improvement, but full restoration remains dependent on the physical repair of the infrastructure.

“ No single-region architecture, however well distributed internally, would have withstood the simultaneous loss of two neighbouring regions. ”

At the peak of the incident, 109 services were affected in the ME-CENTRAL-1 region alone — 25 completely down, 34 degraded, 50 impacted to a lesser extent — including building blocks as central as S3, DynamoDB, EC2, Lambda and RDS.

Copyright Move2Cloud — move2cloud.fr

SECTION 03

RTO, RPO and the four continuity tiers

Two metrics frame any continuity strategy: the RTO, the maximum acceptable delay before service recovery, and the RPO, the maximum volume of data it is acceptable to lose — neither is chosen in the abstract, both follow directly from the business impact of downtime, measured notably in cost per minute.

TIER	TYPICAL RTO / RPO AND PRINCIPLE
Backup / restore	RTO of hours to days, RPO of hours: regular backups stored outside the region, restored manually onto infrastructure rebuilt on demand.
Pilot light	RTO of tens of minutes, RPO of minutes: the minimal core (replicated databases) runs continuously in a backup region; the rest is started at failover time.
Warm standby	RTO of minutes, near-zero RPO: a reduced but functional version of the environment runs permanently in the backup region, ready to scale up.

> **Key takeaway:** An RTO and an RPO are only worth something once they have been verified by a real failover exercise. A plan that is never tested — however well documented — statistically has a one-in-five chance of relying on a backup that turns out to be unusable the moment it is needed.

Multi-site active-active, the top tier

Once target RTO and RPO have been defined, four continuity tiers — from the simplest to the most expensive — make it possible to reach them, and the fourth is multi-site active-active. Both regions handle real traffic continuously, so RTO and RPO are both close to zero and the loss of one region is absorbed with no failover visible to the user. It is the most demanding tier, and it only makes sense where the other three cannot deliver the targets the business has set.

Copyright Move2Cloud — move2cloud.fr

SECTION 04

Four architectural recovery patterns

Each continuity tier corresponds to a concrete architectural pattern, with its own technical requirements.

PATTERN	TIER	HOW IT WORKS	BEST SUITED TO
Backup / restore	The minimal baseline	Automated, encrypted backups replicated to a region distinct from the primary region, with a documented retention policy.	Every architecture, including the more advanced ones — but insufficient on its own for a critical activity.
Pilot light	Replicate the data, not the capacity	Databases are replicated continuously to the backup region, while application servers are provisioned at failover time via infrastructure-as-code.	Services that can tolerate a recovery time measured in tens of minutes.
Warm standby	A reduced but real capacity	A version sized to absorb a fraction of the traffic runs continuously in the backup region; at failover, autoscaling takes over to absorb the full load.	Services that must recover within minutes, as a trade-off between running cost and recovery speed.

Common pitfall: Applying your infrastructure-as-code for the very first time in the middle of a crisis: pilot light provisioning automation has to be tested regularly, otherwise the announced RTO collapses on the day the failover is real.

Multi-site active-active: maximum resilience

Both regions handle real traffic simultaneously, with bidirectional data replication and automatic DNS failover. It is the only pattern that guarantees continuity that is all but invisible to the end user — at the price of significantly higher operational complexity and cost.

SECTION 05

The cost / resilience trade-off

Each additional continuity tier lowers the RTO and the RPO — but raises infrastructure cost, often in a non-linear way.

From extra storage to fully duplicated capacity

Backup/restore adds additional storage cost only; pilot light, continuously replicated databases; warm standby, dedicated application capacity; active-active, full capacity duplicated continuously.

Not every service deserves active-active

Not every service in an organisation justifies the same continuity tier: classify services by real business criticality — not by intuition — and reserve multi-site active-active for those whose downtime has a direct, measurable financial or regulatory impact.

A simple classification exercise

For each service, estimate the cost of one hour of downtime and compare it with the annual cost premium of the tier being considered. When the latter far exceeds the former over a reasonable horizon, the tier is probably disproportionate — and vice versa.

TIER	COST OVERHEAD AND OPERATIONAL EFFORT
Warm standby	Significant — dedicated application capacity running continuously; regular monitoring and failover tests on two live environments.
Multi-site active-active	High — full capacity duplicated continuously; ongoing management of data consistency and of the failover logic.

Copyright Move2Cloud — move2cloud.fr

SECTION 06

Designing beyond a single region: best practices

Beyond the choice of a pattern, several cross-cutting practices determine whether a continuity strategy will actually work on the day it is called upon.

Multi-AZ does not replace multi-region

Spreading a workload across several availability zones within the same region protects against the failure of an isolated data centre — not against an event that affects the region as a whole, be it natural, electrical or physical. The two levels of resilience address different risks and are not substitutes for one another.

Test the failover, don't just document it

A real failover exercise — actually switching traffic over to the backup region, not merely checking that replication works — is the only reliable way to know whether a target RTO is genuinely achievable. Organisations that only test on paper tend to discover the gap at the worst possible moment.

> **Key takeaway:** The choice of a backup region must factor in regulatory data-residency constraints (GDPR and sector equivalents), not just geographical proximity or cost — a non-compliant backup region is not a viable option, even if it works technically.

A failover that depends on manual human intervention in the middle of a crisis adds a reaction delay straight onto the real RTO. Automated health checks coupled with a DNS failover driven by those checks bring that delay down to what the architecture genuinely allows.

Copyright Move2Cloud — move2cloud.fr

SECTION 07

Self-assessment grid

Before considering your organisation ready to absorb the loss of a region, check these six points.

- [ ] **Do you know the RTO and RPO actually achievable for your critical services?** Without real measurement, an RTO stated in a document is nothing more than an unverified assumption.
- [ ] **Are your critical backups replicated to a geographically distinct region?** A backup stored in the same region as the data it protects offers no protection against the loss of that region.
- [ ] **Have you run a real failover exercise within the last twelve months?** A failover never tested under real conditions remains an assumption, not an operational guarantee.
- [ ] **Does the continuity level of each service match its real business criticality?** An oversized level wastes budget; an undersized one exposes the organisation to uncontrolled risk.
- [ ] **Is failover to the backup region automated, or does it depend on manual intervention?** Manual intervention in the middle of a crisis adds reaction time directly to the real recovery time.
- [ ] **Does your backup region comply with your data residency constraints?** A region that is technically functional but non-compliant is not a viable option in a real failover.

> **How to read your answers:** Three "no" answers or more indicate that a structured review of the continuity strategy should precede any further growth of critical workloads hosted in a single region.

Copyright Move2Cloud — move2cloud.fr

SECTION 08

A 5-step roadmap

A progressive approach, designed to strengthen business continuity without interrupting existing services.

1. **Classify**
   Establish, for each service, target RTO and RPO based on the real business impact of downtime — not on an intuitive estimate.

2. **Secure the backups**
   Verify that all critical backups are replicated to a separate region and encrypted, and test that they can actually be restored.

3. **Choose a pattern per service**
   Map each service to the continuity level (backup, pilot light, warm standby, active-active) genuinely justified by its criticality.

4. **Automate failover**
   Put in place automated health checks and a DNS failover driven by those checks, for the services that still depend on manual intervention.

5. **Test continuously**
   Schedule real, regular failover exercises, and adjust the strategy based on their results rather than on documentation alone.

Copyright Move2Cloud — move2cloud.fr

CONCLUSION

Resilience is built before the incident, never during it

The Middle East incident of March 2026 is a reminder that a cloud region remains, ultimately, a physical infrastructure — exposed to the same risks as any other critical site. The best-prepared organisations are not those that avoided the incident, but those whose architecture had already anticipated its possibility.

Classifying services by real criticality, choosing a proportionate continuity level, and testing failover regularly remain, whatever the available budget, the three levers that determine whether an organisation goes through the loss of a region as a managed incident — or as a crisis.

Move2Cloud	is a consulting firm specialised in infrastructure, cloud and IT transformation. We support organisations in the audit, the design and testing of their cloud environment.
Contact	Move2Cloud — Paris · move2cloud.fr
Next step	A continuity audit makes it possible to objectively establish, service by service, the RTO and RPO actually achievable today — and the gap with your business targets.

Sources

Cybersecuritynews.com, "AWS Middle East (UAE) Region Hit by Drone Strikes, 109 Services Disrupted" (March 2026) — The Register, "AWS says drones hit two of its datacenters in UAE" (2 March 2026) — Cost and recovery-testing statistics, gitnux.org overview "Disaster Recovery Statistics" (IBM estimate of the cost of downtime; industry data on restore failure rates) — AWS documentation on disaster recovery strategies (backup/restore, pilot light, warm standby, multi-site active-active)

Document written by Move2Cloud as an example and illustration, for educational purposes. The incident cited in section 02 is a real event documented by the specialised press, presented here solely from the angle of its lessons for infrastructure resilience. This document does not constitute a contractual commitment to results.