

State of Cloud Reliability 2026

Analysis of 47 major public incidents across AWS, Microsoft Azure, and Google Cloud, 2017–2025.

Key findings

3.4h

median duration of a major cloud incident — recovery is measured in hours, not minutes

CloudDowntime, State of Cloud Reliability 2026 — median major-incident duration (204 minutes).

45%

of major incidents trace to configuration or change error — humans and automation, not hardware

CloudDowntime, State of Cloud Reliability 2026 — root-cause distribution.

63%

of major AWS incidents involve us-east-1 — flagship-region concentration is a measurable risk

CloudDowntime, State of Cloud Reliability 2026 — regional concentration analysis.

68%

of major incidents breached at least one monthly SLA target — most of the credits went unclaimed

CloudDowntime, State of Cloud Reliability 2026 — SLA breach and credit-gap analysis.

Major incidents by provider (2018–2025)

Provider	Major incidents	Median duration
AWS	19	3h 37m median
Microsoft Azure	16	3h 46m median
Google Cloud	12	2h 48m median

Root causes of major incidents

Configuration / change error	45%
Software defect	21%
Capacity / overload	13%
Network / hardware failure	11%
External attack (DDoS)	6%
Other / undisclosed	4%

CHAPTERS

1. Executive summary

Cloud reliability is high and improving on paper, yet the incidents that do occur are longer, wider, and more correlated than most architectures assume. Four findings define 2026: multi-hour recovery norms, change error as the dominant cause, flagship-region concentration, and a persistent gap between SLA credits owed and claimed.

2. Incident frequency by provider

The corpus counts 47 major public incidents 2017–2025. Frequency is remarkably stable year over year — five to seven majors annually across the big three — suggesting a floor set by complexity growth offsetting reliability investment.

3. Outage duration and recovery

Median major-incident duration is 204 minutes (~3.4 hours); the 90th percentile reaches 11 hours. The longest event in the corpus — AWS us-east-1 DynamoDB DNS event, Oct 2025 — ran 15 hours, most of it recovery cascade rather than the original fault.

4. What causes major outages

Configuration and change errors cause 45% of major incidents — more than software defects (21%) and hardware/network failure (11%) combined. The industry's biggest reliability lever is change safety, not redundancy.

5. Regional and global concentration

Failures concentrate where control planes live: 63% of major AWS incidents involve us-east-1, and roughly half of Azure and GCP majors involve global layers (identity, networking, API management) that cross every region at once.

6. The SLA credit gap

68% of corpus incidents breached at least one monthly SLA target, yet practitioner estimates put claim rates below 10%. With claim windows of 30–60 days, most owed credits expire quietly.

7. Outlook for 2026–2027

Expect fewer but wider incidents: consolidation of internal platforms means single faults reach more services (the 2025 DynamoDB DNS and Service Control events are the template). Architectural independence from provider control planes becomes the differentiating resilience investment.

Cite as: CloudDowntime, “State of Cloud Reliability 2026,” clouddowntime.com/report/2026. Full method: clouddowntime.com/methodology.