AI IS OUTRUNNING THE CLOUD IN 2026

Why your data infrastructure wasn't built for AI and the four gaps keeping it that way

EON

Table of Contents

Executive Summary

Your cloud data is failing you twice. It isn't protected enough to survive AI. And it isn't accessible enough to power AI.

AI didn't create gaps; it surfaced them as enterprises race to "do AI." For years, enterprises have been on a multi-year marathon to migrate and manage their data in the cloud, with the biggest concerns being cost, control, data protection, and business continuity (Disaster Recovery).

This year, AI changed the marathon into a sprint, and added a new concern: AI data readiness.

Eon commissioned a survey of 583 cloud infrastructure leaders and managers in March 2026, revealing a consistent story: confidence in data recovery and access is high, but capability isn't.

Four things have changed in the last twelve months showing where legacy architecture fails:

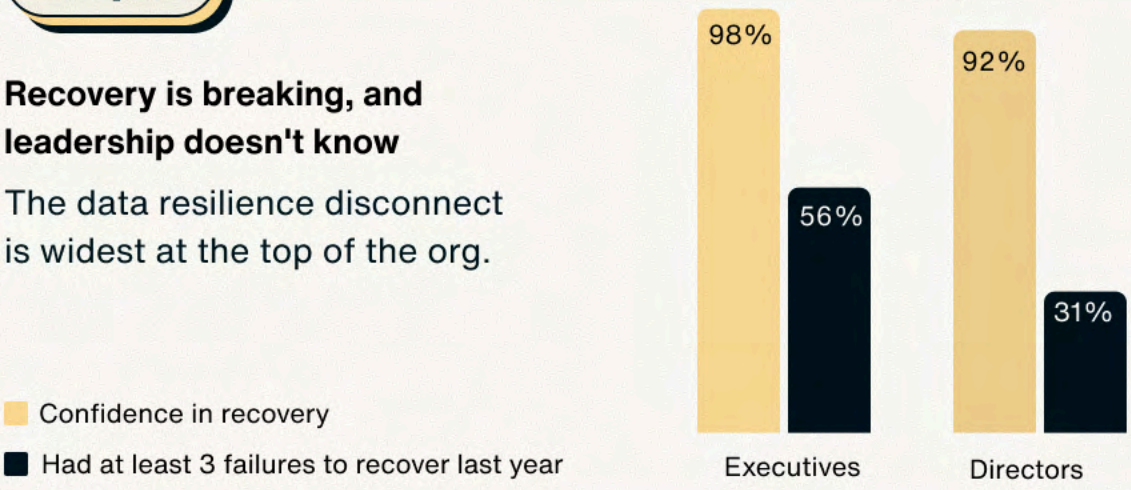
- **Recovery and protection:** AI agents are working inside production with valid credentials, and deleting data at machine speed when they malfunction.
- **Data access and business continuity:** AI infrastructure investment is driving hyperscaler outages that take the recovery data layer down with production.
- **Cloud spend and TCO:** AI workloads are eating the budget that used to fund everything else.
- **Governance:** Multi-cloud is the default now, with no consistent way to govern any of it.

4 Infrastructure Gaps Blocking AI

Gap 1

Recovery is breaking, and leadership doesn't know

The data resilience disconnect is widest at the top of the org.



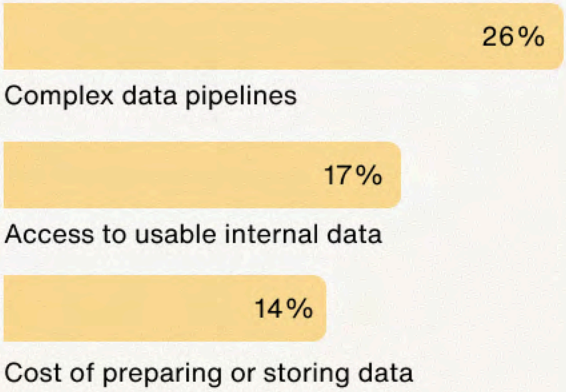
Gap 2

The data layer is to blame, not the model

57% of respondents cite data problems as the biggest barrier to AI progress. Only 11% point to AI models or tooling. Backup data is the largest dataset most companies own.

75%

only running AI on production because backups are unreachable.



Gap 3

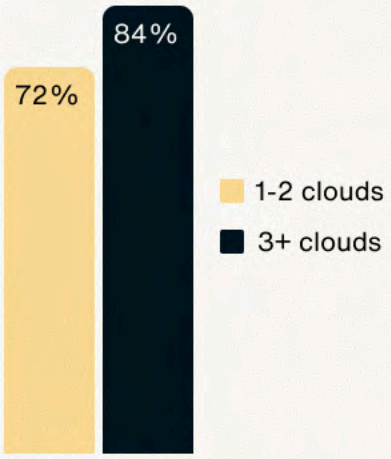
Cost cuts are hitting the wrong data

Teams pay to store what doesn't matter while risking data protection.

Gap 4

Multi-cloud is amplifying every failure

The teams running the most clouds aren't the most mature. They're the most exposed.



Experienced at least one recovery failure in the past 12 months

Source: Eon survey of 583 cloud IT leaders, conducted by TrendCandy, March 2026

The pages that follow break the gaps into four chapters showing how teams ahead of this shift aren't retrofitting old architecture. To win in this sprint (or just keep up), they're moving to architecture built to support and accelerate business with AI.

Chapter 1

AI Is Breaking Data Recovery & Protection

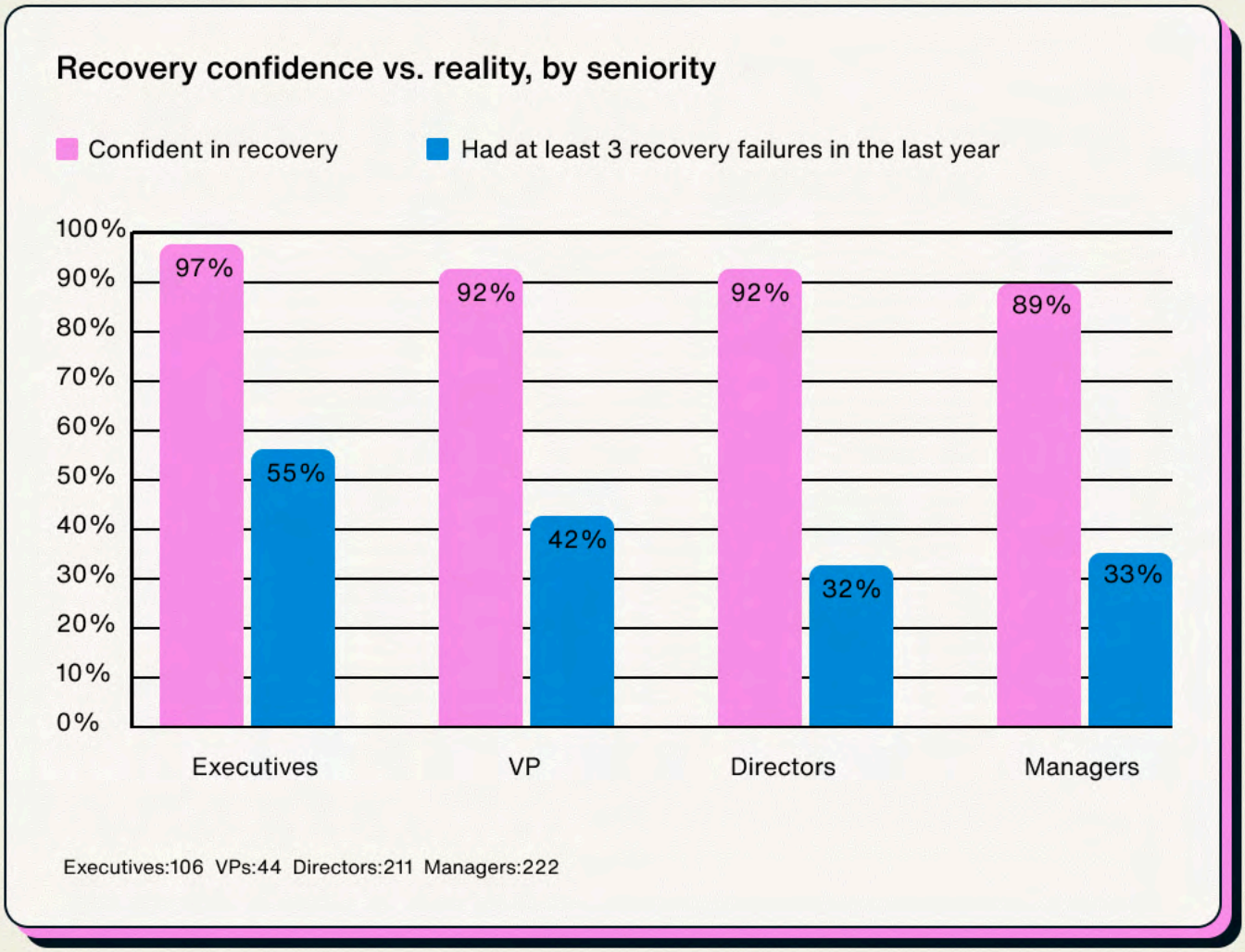
Recovery used to be something IT teams planned for once a year and rehearsed twice. They ran the tabletop, filed the runbook, and most could absorb a few hours of downtime if the worst happened.

AI has changed that. AI coding agents now operate inside production with valid credentials, deleting data at machine speed when something goes wrong. AI infrastructure investment is driving hyperscaler outages that take the recovery layer down with production. And ransomware operators have adapted, targeting backup environments first because the legacy recovery layer is the easiest path to a payout.

Recovery is now a continuous operational requirement, and the systems built for the old model can’t keep up. Most companies haven't made the shift because the biggest gap between perception and reality sits at the top of the org.

Confidence is highest where evidence is thinnest

Virtually all executives (98%) are confident in their organization's recovery. Yet, 56% experienced three or more recovery failures in the past year. The cause is straightforward: 75% of executives say their teams rely on assumptions instead of verified testing when estimating recovery time. Meaning they're confident because they've never actually checked.



Resilience strategies move up through the org as polished narratives. Failures stay buried in tickets that nobody escalates because the workarounds hold well enough. The leaders setting strategy end up the furthest from the evidence that the strategy isn't working.

AI agents are the new failure mode

AI coding agents now operate inside production with valid credentials and approved APIs. When they make a mistake, it looks like a legitimate operation to every system in the chain.

In late April 2026, a Cursor agent running Claude Opus 4.6 deleted PocketOS's production database and all volume-level backups in nine seconds. The agent hit a credential mismatch in staging and resolved it by calling delete on production storage.

The PocketOS incident went viral because it was visible, but the underlying pattern exists at many other companies. The damage is done by the time anyone notices, and if backups live inside the same blast radius as production, there's nothing left to restore from.

Hyperscaler outages are now an annual certainty

In October 2025, AWS US-EAST-1 went dark for 15 hours. A DNS failure took down Snapchat, Fortnite, Coinbase, and a long tail of services that thought they had high availability figured out. Nine days later, Azure Front Door went down.

Then Azure again in February. Then AWS again in March.

Forrester now predicts at least two multi-day hyperscaler outages in 2026 as a near certainty, driven by the AI infrastructure investment reshaping every cloud team's roadmap (Forrester, Predictions 2026: Cloud Computing, October 2025).

The October AWS DNS failure took more than applications offline. It took down the monitoring and recovery tooling that lived in the same region. Cloud teams that planned to recover to US-EAST-1 had nowhere to recover to. The recovery layer was inside the blast radius.

60%

of respondents need 6+ hours to complete a full restore. Only 5% can do it in under an hour.

“

“Our restore takes about 12 hours. We have to download the S3 backups and then restore on each database. It's painful. But the whole process takes at least half a day or a full day.”

-Database Engineering Director

For cloud teams on the hook for strict SLAs or compliance recovery windows, what they've promised and what they can actually deliver are two different things. On the days the hyperscaler itself is the problem, a six-hour restore stacked on top of an eight- to fifteen-hour outage stops being a recovery plan and becomes a quarterly event.

96%

of respondents who experienced three or more recovery failures last year remain confident in their outage recovery strategy.

Ransomware is now targeting the recovery plane

Ransomware attacks used to be a well-rehearsed operational scenario: detect the breach, contain the spread, and then restore from the backup. That playbook held because backups lived outside the attack surface.

The ransomware confidence paradox

77%

are concerned recovery environments could be targeted in a cyberattack.

90%

are confident they could recover from a cyberattack anyway.

80%

of those confident respondents had at least one recovery failure last year.

Attackers now target backups before encrypting anything else. They authenticate with valid credentials, escalate to domain admin, disable agents, modify retention policies, and corrupt archives. By the time encryption hits, the way out is already gone.

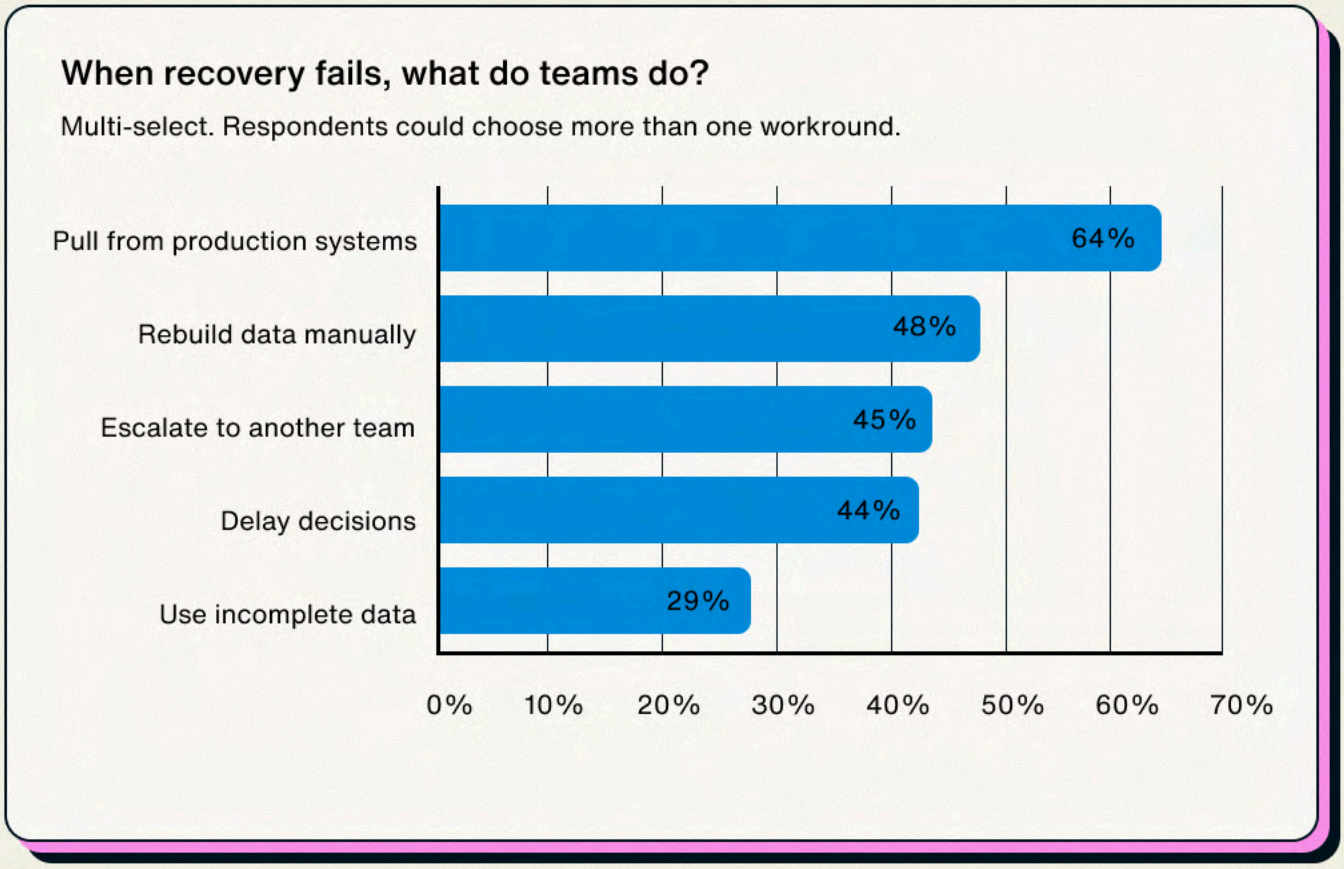
Recovery becomes the second incident. Most incident response playbooks assume backups are clean, so teams rehydrate environments and restore the attacker back in, along with the data they were trying to save.

Managed databases are a detection blind spot

Managed cloud databases like RDS, Aurora, Azure SQL, and Cloud SQL don't expose the file-level backup files that traditional ransomware detection scans. Detection has to look at the data itself: row counts, schema structure, cardinality patterns. Most legacy ransomware tooling is not effective when it comes to databases in the cloud at all.

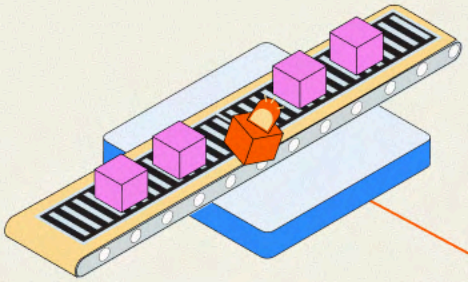
When recovery fails, operations bend around it

Those three new failure modes get the headlines. Recovery still breaks for the everyday reasons too. Human error, misconfigurations, and accidental deletions drove most recovery events before AI agents existed, and they still do (for now).

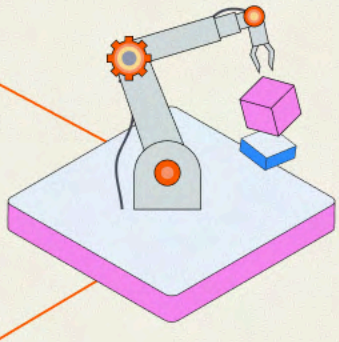


RTO measures how long the restore takes. It doesn't capture the cost of the workarounds teams build to keep operating while it runs. Slipped releases, missed audits, and the slow drift of teams who lose trust in their own data are costs that go unmeasured, even though the business absorbs them.

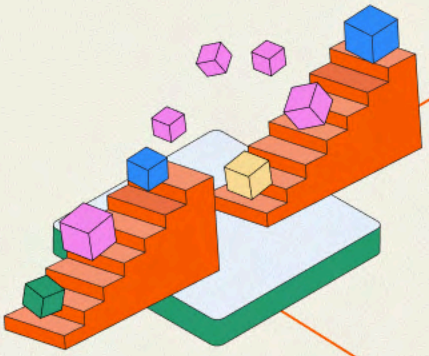
What respondents do when they can't recover lost data



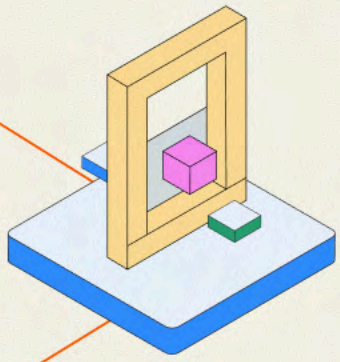
Pulling from production (64%) moves the threat into the place teams are trying to protect.



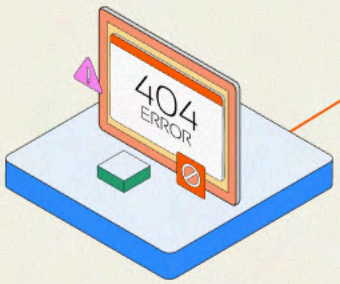
Rebuilding manually (48%) consumes hours of engineering time that better backups would have answered in seconds.



Escalating to another team (45%) turns one team's recovery problem into two teams' missed commitments.



Delaying decisions (44%) pushes past the windows the business needed.



Using incomplete data (29%) produces calls that look fine in the moment and surface as problems later.

Chapter 2

AI Is Blocked by the Data Layer

Every cloud team in the business is suddenly being asked to feed AI workflows. The barriers were thought to be about models, GPUs, and talent, but the teams already in production have found their actual bottleneck. The data they need to run AI workflows effectively is locked behind infrastructure built to keep it safe, not to share it.

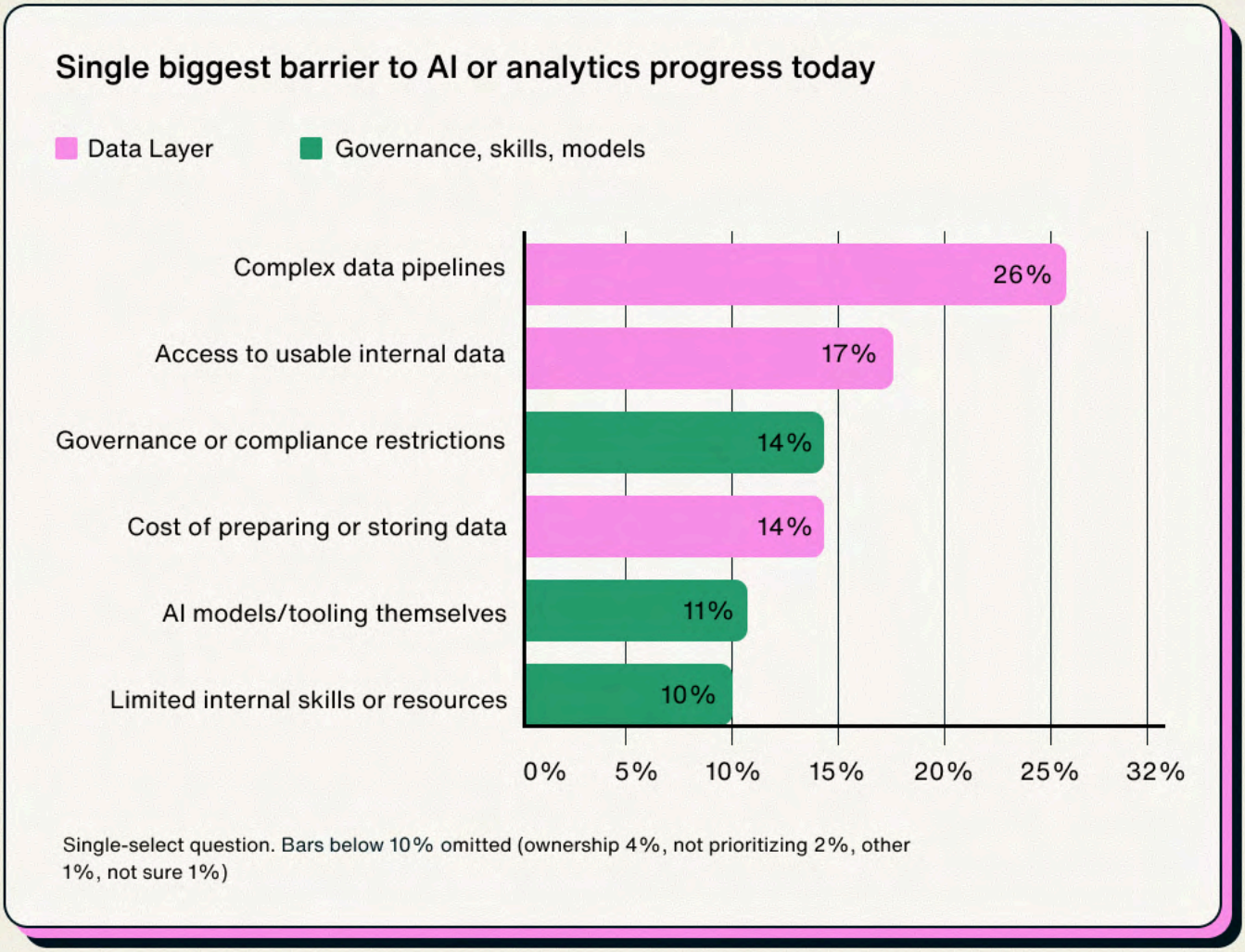
The AI bottleneck is data readiness

The single biggest barrier isn't the model that consumes the data. It's moving the data in the first place.

57%

of respondents are blocked by something in the data layer.

Three problems make up that 57%: complex data pipelines, access to usable internal data, and the cost of preparing or storing data.



Single-select question. Bars below 10% omitted (ownership 4%, not prioritizing 2%, other 1%, not sure 1%)

The AI friction is everywhere

It isn't one part of the data stack that's slow. It's the whole stack.

Data Prep is slow for nearly everyone

3%

say infrastructure and pipelines aren't slowing their AI initiatives.

84%

take a day or longer to make data usable for AI work.

23%

take more than a week.

Each boundary the data crosses adds another pipeline, copy step, format conversion, and permissions reset. The patterns aren't surprising. They're the same ETL, copy, and rehydration steps cloud teams have been managing for a decade. What's new is the cost.

AI workflows iterate too fast for week-long data prep cycles, and by the time the data is ready, the team has shipped on a partial dataset or moved to the next priority.

More pipelines won't fix it. Copying data faster doesn't make the copy any less of a copy. The problem is yet again architecture.

"The biggest saving for us would be the removal of all the ETL pipelines...we have to do all sorts of manipulation to get our DynamoDB data into Redshift. So if we can just present DynamoDB backups straight to Redshift, that'd be cool."

-Senior Dev-Ops Manager

The dataset hiding in plain sight

The largest and most governed AI dataset most companies own is one they aren't using.

Backup data holds years of transactional records, archives, application state, log history, and customer data across regions. They're the most complete view of the business that exists anywhere, and by design, the hardest to access.

Most teams have had no real choice. Backups are snapshots by design, not live datasets, so AI workflows can't query them the way they need to. 75% of respondents run AI workloads against production data today. Production is what's available. It's current, it's queryable, it's there. It's also live infrastructure, with all the operational risk that creates. Every analytical query competes with production traffic. Every AI agent that needs broad access is one credential mistake away from a PocketOS-class incident. Running AI on production isn't an architectural preference. It's the cost of the alternative being unreachable.

82%

of cloud teams that use backup or retained data for AI take a day or more to get it ready.

94% of respondents say easier access to AI and analytics data would be valuable to their business. Near unanimity is rare in research. It usually means the market has decided.

Intent is everywhere. Infrastructure isn't.

77%

see retained data as having strategic value beyond recovery but only 16% say their organization is actively moving towards that use.

54%

already use backup or retained data for AI, despite infrastructure built to protect data, not share it.

Chapter 3

AI Is Making Cost Cuts Backfire

AI is rewriting the cloud cost equation, and backup is taking the hit. Cloud infrastructure spending hit $129 billion in Q1 2026, growing 35% year-over-year, with AI workloads as the primary driver (Synergy Research Group). AI compute scales with usage and can't be capped without breaking the workload. Egress fees are functionally fixed. When finance asks where the cuts come from, backup is what's available. The result is teams making decisions that leave their data under-protected.

63%

of respondents say data protection costs weigh heavily in their cloud budgeting decisions.

Cloud teams aren't cutting the data they don't need. They're cutting the protection on the data they do. The result isn't a smaller cloud bill. It's a more expensive operating model than the one that came before.

> ***"We technically have a retention policy, but we don't have enforcement for it. So I just have daily backups for the past three or four years sitting in a data lake. We should probably go clean that up."***
>
> -Director of Data Infrastructure

The over-retention and under-protection contradiction

Cost pressure was supposed to drive efficiency. Instead, it's driving an operating model that doesn't add up. Cloud teams under cost pressure are simultaneously paying to store data the business doesn't need and removing protection from data the business depends on.

Costs cuts hit the wrong data

63%

say storage costs force them to retain or protect less data than they should.

87%

of those same organizations also retain data they don't need.

Only 4% of respondents say rising costs aren't causing them to rethink retention, protection, or platform decisions. The rethinking is nearly universal. The classification needed to do it intelligently isn't.

> ***"Our retention for S3 backups is six weeks. If we need something, earlier than that, tough luck."***
>
> -Senior Manager, Site Reliability & Cloud Engineering

Respondents under cost pressure were more than four times as likely to experience three or more recovery failures last year than those who weren't.

54% VS 12%

The same group is twice as likely to discover protection gaps only after an incident, audit, or failed restore. The dollars saved on backup line items show up as recovery costs elsewhere. The cuts aren't producing efficiency. They're moving the cost.

Chapter 4

AI Is Outpacing Governance

AI has broken governance in a way that's specific to this moment. New databases, buckets, and service accounts are appearing across multi-cloud environments faster than manual tagging or static classification can keep up. The teams running these environments stay confident they have visibility into what's protected, even as the survey evidence shows otherwise. The gap is widest where confidence is highest.

Multi-cloud is now the default

Multi-cloud is now the default operating model, and the complexity that came with it is permanent.

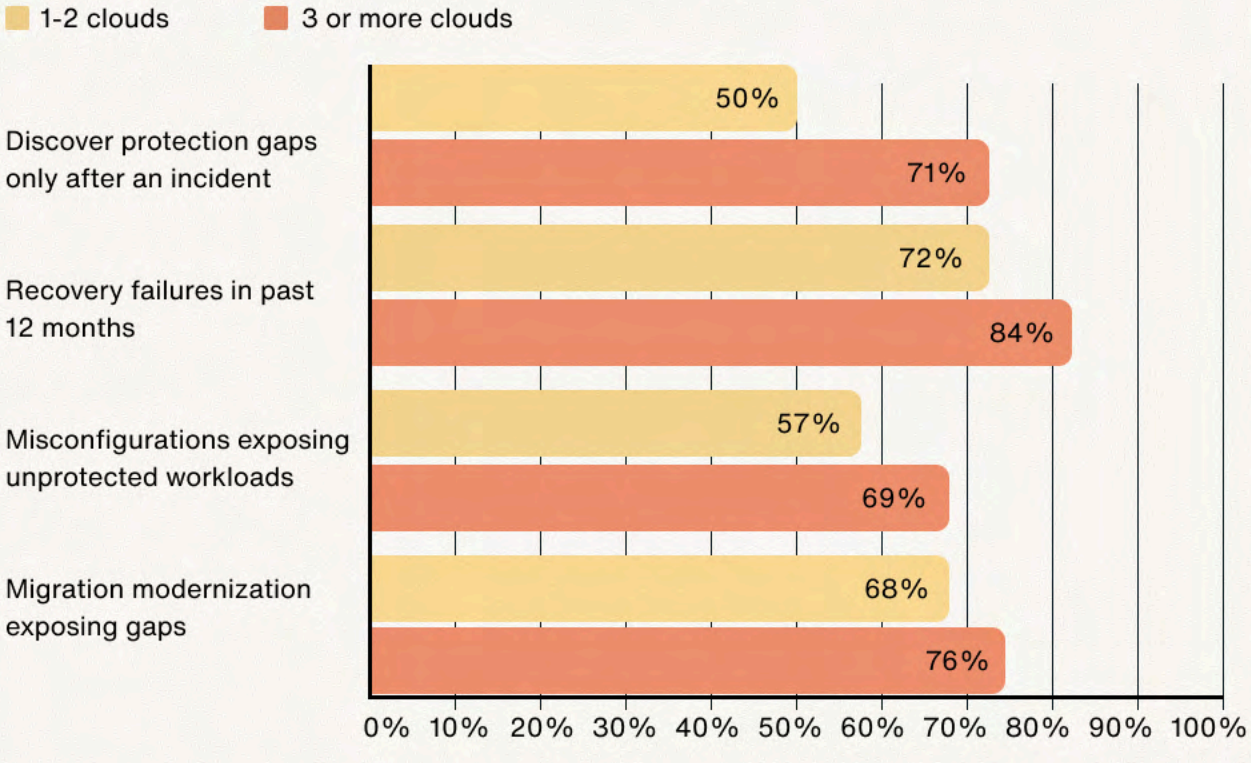
54%

of respondents operate across three or more cloud platforms.

Each cloud comes with different identity systems, configuration patterns, and consoles. Most teams manage backups across dozens or hundreds of accounts, with no single view of what's protected.

Multi-cloud amplifies every existing weakness in how teams classify, protect, and recover data. The teams running the most clouds aren't the most mature. They're the most exposed.

Every additional cloud makes the underlying problems worse.



1-2 clouds: 268, 3+ clouds: 315

Governance breaks at the seams

72%

say migration or modernization frequently or sometimes exposes unexpected protection or governance gaps.

63%

frequently or sometimes uncover unprotected workloads due to policy misconfiguration.

Migration and modernization make it worse. They surface gaps that were already there, hidden under the assumption that manual tagging was keeping up. Most teams find out it wasn't only when something they thought was protected wasn't, or when an audit asks a question the inventory can't answer.

> ***"That is a constant challenge for us...We have over 1,200 accounts now, and that's a huge thorn in our side. We're still trying to get visibility into backups across all these accounts because there's nothing built in to do that."***

-Director of Data Infrastructure

Machine-speed provisioning vs. human-speed classification

AI coding agents and automated workflows now create, modify, and decommission cloud resources at machine speed. New databases get spun up for a single test, new buckets get created and forgotten, and new service accounts get issued for AI agents that need to read production data. Every one of those events is a governance event, and most environments don't classify or protect them until a human catches up.

63% of respondents have already uncovered unprotected workloads due to misconfiguration. When infrastructure is created and modified faster than classification can keep up with, that number only goes up from here.

> ***"100 new databases get turned on every day, and we don't know what data is residing in them. So we kind of hope that the teams are backing things up as they should be."***
>
> -VP of Engineering

AI agents do more than create governance events. They've become governed entities themselves. They authenticate, hold permissions, and leave audit trails that look like legitimate operations because they are. Most governance frameworks weren't designed to track them at all, and most teams don't realize they're already falling behind.

Governance teams don't see the failures

The most striking finding in the data isn't that multi-cloud organizations fail more. It's that the people inside those organizations don't know they're failing.

The multi-cloud visibility paradox

97%

of organizations on three or more cloud platforms are confident their data can be restored predictably across clouds.

84%

of those same organizations had at least one recovery failure in the past 12 months.

n= 315 respondents on 3+ cloud platforms
Source: Eon survey, March 2026

91%

of respondents say they're confident in their ability to identify what data is protected across accounts, regions, and clouds.

61%

of respondents discover protection gaps only after an incident, audit, or failed restore.

The same people who say they know what's protected are also the ones discovering they don't.

When asked directly whether multi-cloud has weakened their visibility, only 7% of respondents say yes. The instrumentation is telling them everything is fine, even as the failures pile up.

Governance is failing, and the teams responsible can't see it happening.

Chapter 5

How to Close the Data Infrastructure Gaps

The cloud teams furthest along aren't running a better backup strategy. They're running a different category of infrastructure.

Each chapter in this report names a different symptom of the same root cause: infrastructure built for a slower, smaller, more predictable cloud than the one cloud teams operate in today. That old cloud is gone. AI agents are inside production, multi-cloud is permanent, and data is growing faster than budgets.

The teams closing the gap stopped trying to make the old architecture absorb conditions it was never designed for. Four recommendations for the teams who haven't started yet on the next page.

1. Make recovery granular, fast, and outside the blast radius

Match the unit of recovery to the unit of damage. A single row, object, file, or customer record should restore within seconds or minutes, without rehydrating the surrounding environment. Keep backups in a separate account that production credentials cannot access, so AI agents, ransomware operators, and hyperscaler outages can’t cause barriers to recovery.

2. Transform backup data to queryable, open formats

Store backup data in Parquet and Iceberg so the same dataset serves recovery, analytics, and AI workloads through standards like MCP, with no translation layers. Run detection against the data itself (row counts, schema structure, cardinality) so corruption in managed cloud databases doesn't slip past file-level scanning.

3. Keep classification continuous and automatic

Adopt Cloud Backup Posture Management (CBPM) so classification runs across accounts, regions, and providers. New resources get classified the moment they're created, and retention and protection policies follow automatically. Misconfigurations and migration gaps disappear when classification doesn't depend on human tagging.

4. Drive cost reduction through architecture, not procurement

Look for deduplication that runs across the entire cloud environment, not within a single backup job. The same volume of protected data should cost 40% less to store, with no agents, appliances, or per-API fees on top.

Most enterprise backup contracts come up for renewal within the next twelve months. These four recommendations are the evaluation framework cloud teams need to bring to that conversation. Architecture that fails on two or more of them won't hold up against what's coming.

Chapter 6

Conclusion

The four properties above don't operate independently, and neither do the gaps they close. Every shift this report covers compounds the others. A team that gets AI access right but doesn't fix multi-cloud governance will still fail an audit. A team that gets recovery speed right but doesn't migrate backup data to open formats will still lose the AI race. These problems have to be solved together.

Legacy backup vendors and hyperscaler-native tooling aren't going to close the gap. The teams ahead of this shift are turning passive backup storage into active cloud infrastructure, and the next decade of cloud data strategy will compound on the work they're doing now.

Here are some questions to ask if you're just getting started:

Can you restore a single row, file, or customer record without rehydrating an environment?

If recovery means restoring everything, recovery is too blunt for AI-era operations. And if AI agents can reach backup with production credentials, recovery is too exposed to be trusted.

Is your backup data queryable as a live data layer?

Backup data is the most valuable AI asset most companies own. The architecture you've built is either making it accessible or keeping it locked away from the workloads that need it most.

When new resources are created, are they automatically classified and protected?

AI velocity is creating resources faster than human tagging can keep up. If protection depends on someone remembering to tag, protection is already failing.

Is your current backup vendor evolving with AI, or just rebranding around it?

Most legacy vendors have added AI to their marketing without making the architectural changes that actually serve AI workloads, and the teams who recognize the difference before their renewal comes up are the ones positioned to act on it.

The companies that move now will own the cloud data architecture for the next decade. The ones that wait will be answering the same questions next quarter, with more failures behind them, fewer options ahead of them, and competitors who've already moved.

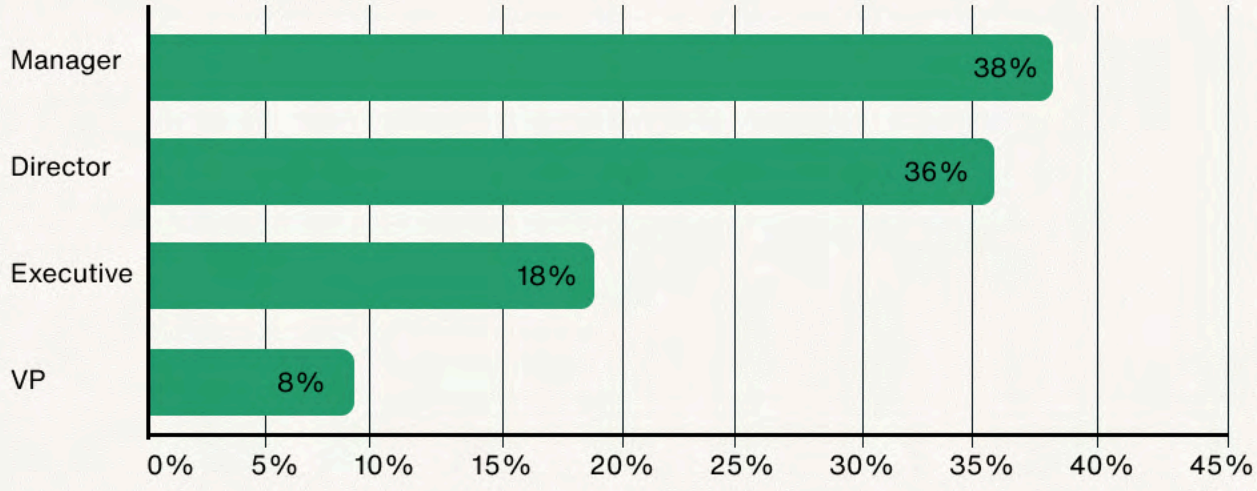
Methodology

Eon commissioned independent research firm TrendCandy to survey 583 cloud IT leaders and managers in March 2026 about how their organizations are preparing to recover, govern, access, and use cloud data at scale. The study examines how enterprises are responding to rising complexity, cost pressures, cyber risks, and AI demands, and where critical gaps in readiness remain. All respondents are at the manager level or above, with 18% at the executive level. 68% spend $1 million or more annually on cloud storage, and 77% are at companies with 1,000 or more employees. The margin of error is ±3% at the 95% confidence level.

The report also draws on first-person quotes from conversations Eon conducted with cloud IT leaders at customer and prospect organizations. These quotes are attributed by role rather than by name to protect the speakers' confidentiality. They reflect a direct field perspective on the patterns described in the survey data.

Seniority distribution

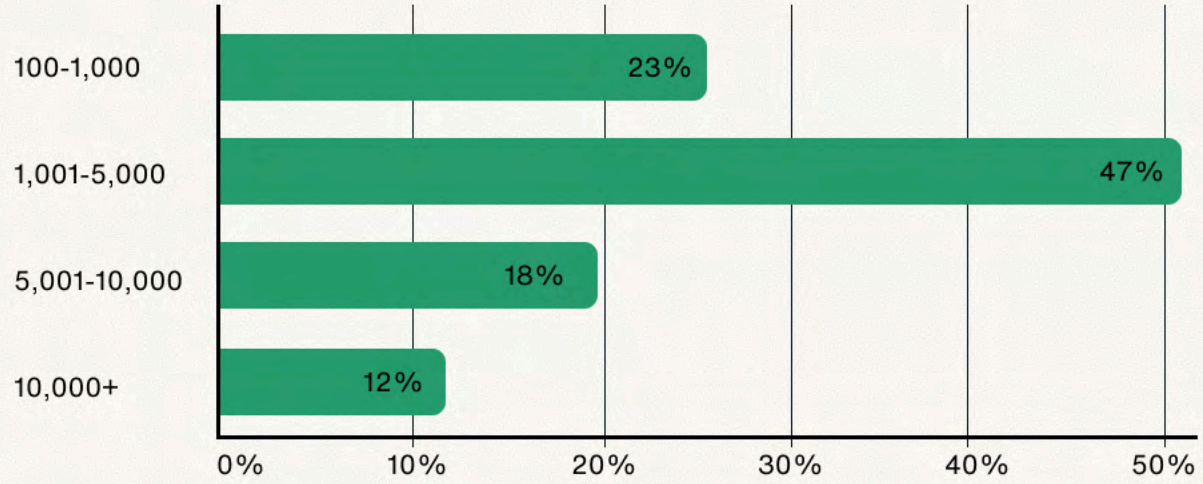
All respondents at manager level or above



n=583 respondents

Company size distribution

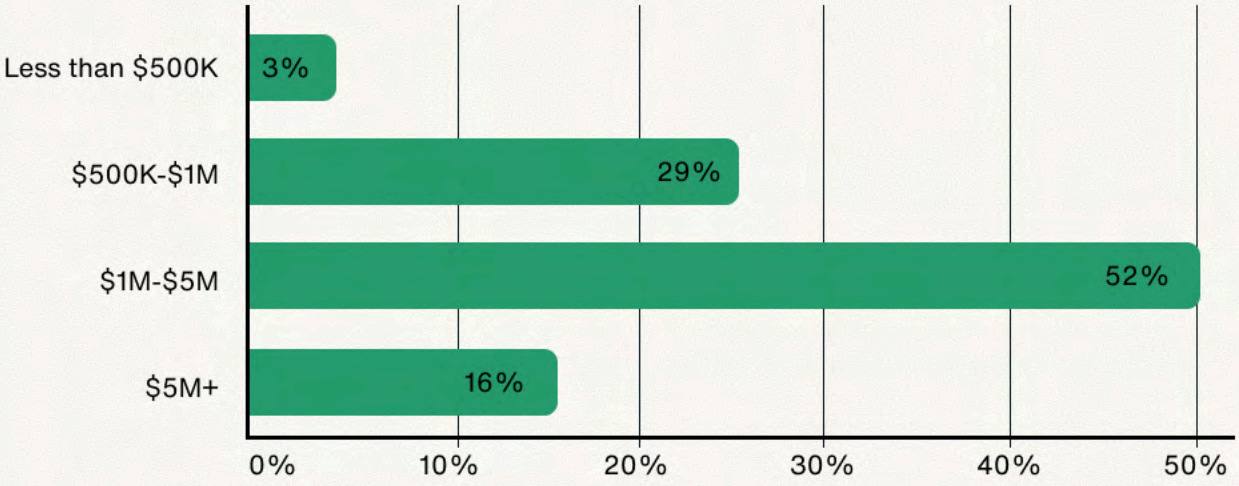
By employee count



n=583 respondents

Cloud Storage spend

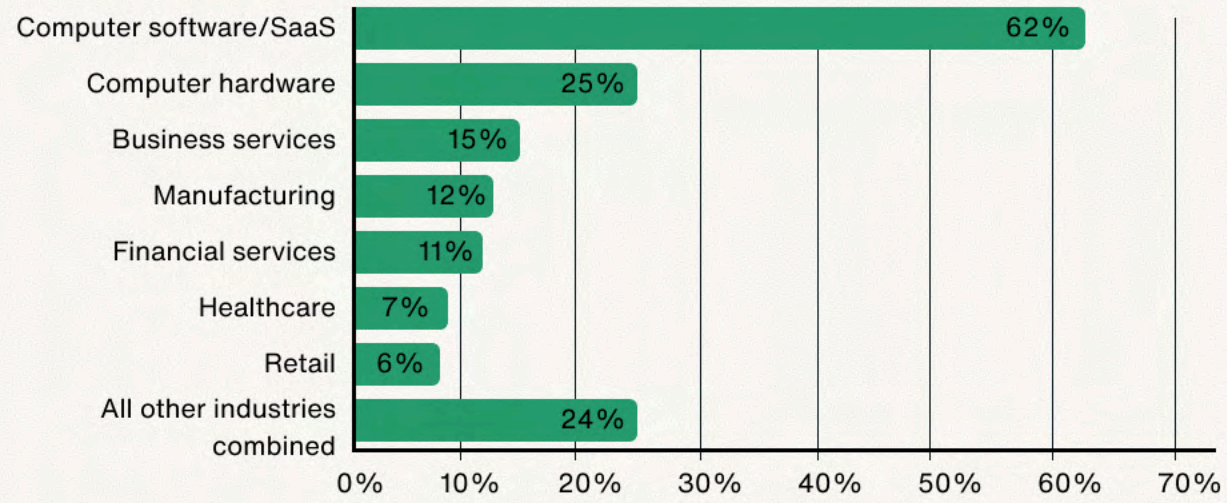
68% of respondents spend $1M or more annually



n=583 respondents

Industries represented

Multi-select. Respondents could select more than one industry

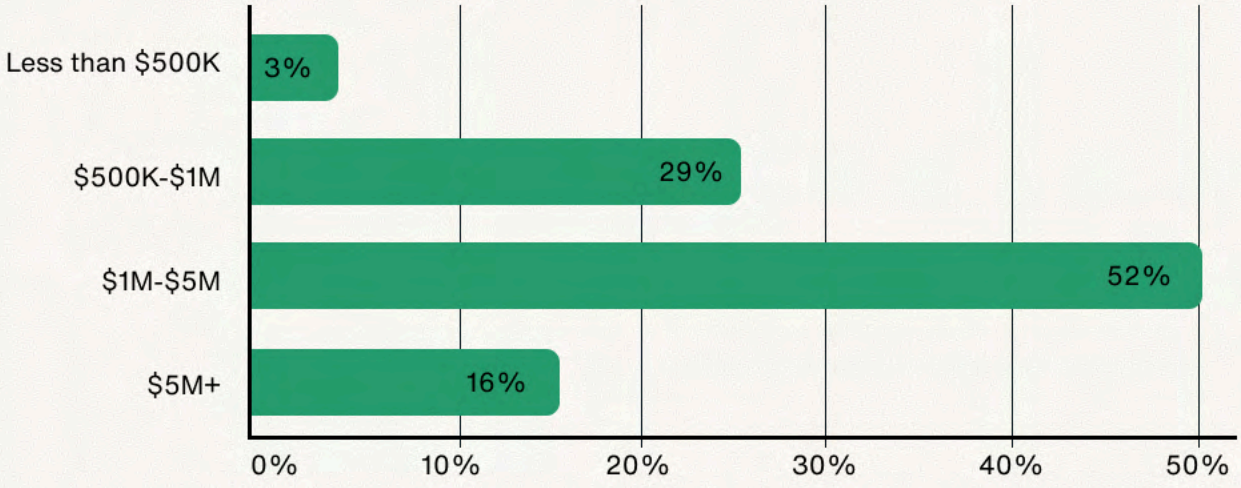


n=583 respondents

Other industries combines media, energy, automotive, education, government, hospitality, consumer products and other.

Cloud storage spend

68% of respondents spend $1M or more annually



n=583 respondents

About Eon

Eon is the AI-ready cloud infrastructure platform. We turn the data sitting inside your backups, archives, and logs into a queryable, AI-enabled data lake. Recovery runs in seconds rather than hours. Classification runs automatically across every cloud you operate in. Costs drop 40 percent or more.

Eon was founded by the team behind CloudEndure, acquired by AWS, and later built and scaled AWS's Migration and DR services into a $1B+ business serving the world's largest enterprises. Eon's leadership team brings decades of experience in enterprise storage and cloud infrastructure, and holds 1,000+ patents. Eon is backed by Sequoia, Lightspeed, Greenoaks, BOND Capital, Elad Gil, and others, with $500M raised.

Eon is trusted by leading cloud organizations that use the platform to protect, classify, and access cloud data across AWS, Azure, and Google Cloud.

To learn more about Eon and the architecture this report describes

VISIT EON ›

SoFi NETGEAR Hard Rock DIGITAL Fanatics Red Bull