



Top 5 Cloud Storage Struggles and How to Resolve Them



Cloud databases are the backbone of modern infrastructure, providing scalability, flexibility, and convenience. However, organizations face several critical challenges when managing cloud storage, particularly for relational workloads. This ebook explores the five most common cloud storage struggles and provides strategies to resolve them, ensuring optimized performance, cost savings, and improved data agility.

Cloud Storage Issue #1

Poor Storage Performance and Its Impact on User Experience

Poor storage performance is one of the most significant challenges that organizations face when managing data in the cloud, and it directly affects the user experience.

These performance issues manifest as slow-loading applications, long wait times for retrieving data, and systems that become unresponsive under heavy workloads. This sluggishness can stem from cloud configurations that weren't optimized for the complexity or volume of data they are expected to handle, or from mismanagement of the underlying infrastructure. The disconnect between expected and actual performance, often due to bottlenecks within the storage or compute layers, can severely hinder an organization's operations, particularly in data-heavy industries.



Industry Example

When a healthcare provider's database is unable to retrieve patient records quickly, it disrupts care delivery. Similarly, in retail environments, where real-time data is critical for tracking inventory and managing transactions, performance issues can lead to significant delays at checkout or prevent customers from completing their purchases altogether. For financial services, where high-frequency trading or real-time analytics is paramount, even small delays in data access can lead to substantial financial losses. Poor storage performance, therefore, is not just a technical challenge—it affects the organization's ability to deliver core services, resulting in decreased customer satisfaction and potential revenue loss.

The broader organizational impact is just as concerning. Slow database performance results in a steady increase in calls to IT help desks, with users frustrated over delays in application performance. This can lead to costly, time-consuming support efforts and an over-reliance on consultants and database administrators (DBAs) to optimize performance. Worse yet, organizations may end up investing in expensive yet ineffective solutions, such as repeatedly tuning the database without addressing the root cause of the problem—an underlying storage infrastructure that simply cannot handle the load.



Here's how to solve it!

Resolving poor storage performance requires a multifaceted approach, beginning with a comprehensive understanding of your system's baseline performance. Organizations should collect performance data before migrating to the cloud, which will allow them to measure whether performance is meeting expectations once the migration is complete. Often, problems arise because the performance requirements were not properly defined or accounted for during initial testing phases. Without these critical baselines, identifying the source of slowdowns becomes much more difficult, leaving teams in the dark about how to fix them.

Once performance bottlenecks are identified, there are several strategies available to improve cloud storage performance. First, organizations can consider upgrading their cloud-native storage options. Most cloud providers offer different tiers of storage solutions—ranging from standard to ultra-high-performance—each designed for different workloads. Upgrading to a higher-performance storage tier can often resolve latency issues and improve response times for critical applications. However, this approach typically requires downtime as new volumes are provisioned and data is migrated, so careful planning is necessary.

Another option is to adjust the virtual machine (VM) configurations. In many cases, the bottleneck is not the storage itself, but the VM shape being used. Some VMs are limited in the amount of data they can transfer per second or by the number of input/output operations per second (IOPS) they can handle. By choosing a VM with a higher bandwidth or more optimized for intensive workloads, organizations can improve performance without making drastic changes to the underlying storage.

Finally, more advanced solutions, such as Silk's software-defined cloud storage, offer organizations the ability to decouple storage performance from compute limitations. By virtualizing storage and using dedicated compute resources to handle storage I/O, solutions like Silk can dramatically reduce latency, improve data transfer speeds, and ensure consistent performance across even the most demanding workloads. This allows organizations to optimize their cloud configurations without the need for constant resizing of VMs or upgrading native storage tiers. Silk's ability to deliver up to two million IOPS and 20 GB/s with ultra-low latency ensures that applications perform reliably, even during peak usage times.

By employing these strategies, organizations can significantly enhance their cloud storage performance, ensuring that applications remain responsive, and users can access the data they need without delays. This not only improves user satisfaction but also reduces support costs and prevents the need for expensive, long-term optimization projects that may not even address the core issue. In the long run, by investing in the right solutions upfront, organizations can ensure the scalability and reliability of their cloud environments, even as their data needs grow.

Cloud Storage Issue #2

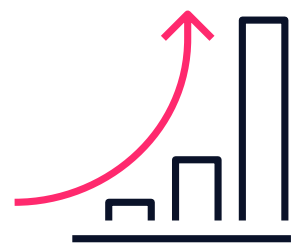
Long-Running Batch Jobs and Analytical Workloads

Long-running batch jobs and analytical workloads are critical to business operations, yet they often suffer from storage performance issues in the cloud. These workloads typically include tasks like data processing, analytics, reporting, and backups, which are designed to run during off-peak hours, such as overnight or during scheduled maintenance windows. In theory, this approach minimizes disruption to regular business operations, but in practice, cloud storage bottlenecks can cause these processes to take much longer than anticipated.

The underlying cause of these delays is often related to storage performance limits that cannot keep up with the demands of large-scale, data-intensive jobs. As organizations continue to grow and the amount of data they process increases, the IOPS and throughput of the cloud infrastructure may no longer suffice. Even if these workloads were running efficiently on-premises, the migration to a cloud environment with different storage capabilities can introduce new challenges. As a result, tasks that once fit neatly within predefined timeframes begin to exceed their allotted windows, creating a ripple effect that impacts subsequent business processes.

For example, an organization may schedule its daily batch jobs to run overnight, expecting them to complete before employees arrive in the morning. However, if storage bottlenecks slow down these processes, the batch jobs may still be running during the business day, disrupting normal operations. Similarly, a retail company may expect end-of-day reports to be ready for analysis at the start of the next business day, but if the storage infrastructure cannot handle the data throughput required, reports may not be available until hours later, delaying critical business decisions.

The delays caused by long-running batch jobs and analytical workloads can have a cascading effect across an organization. When these jobs don't complete on time, other systems that rely on the processed data can't function as expected, leading to broader operational inefficiencies.



Industry Examples

In financial services, for example, batch processing is often used for tasks such as end-of-quarter reporting, auditing, and transaction reconciliation. If these jobs run longer than expected, the finance team may be unable to close the books on time, leading to compliance risks and delays in financial reporting.



Insurance providers face similar challenges. During high-demand periods, such as after a natural disaster or during open enrollment, the volume of data being processed for claims, customer applications, and policy renewals can increase significantly. If the storage infrastructure cannot handle the surge in data, these critical processes are delayed, resulting in frustrated customers, bottlenecked workflows, and increased strain on support teams. Delays in processing insurance claims can erode customer trust and negatively impact the company's reputation.

The issue becomes even more pronounced during peak business cycles, such as the holiday season for retail or tax season for financial institutions. During these times, workloads are already elevated, and any additional delay in batch job completion or data processing can lead to substantial business losses. If an online retailer cannot process orders quickly due to lagging back-end batch jobs, it may miss shipping deadlines, resulting in unhappy customers and lost sales.



Here's how to solve it!

Addressing performance issues related to long-running batch jobs and analytical workloads requires a targeted approach to diagnose and resolve the root cause. The first step is to identify where the bottleneck is occurring. This can be done by tracking runtime inconsistencies, such as fluctuations in job completion times, and analyzing system logs to detect performance lags during peak workload periods. By gathering data over time, organizations can determine whether storage is the primary bottleneck, or if the issue lies elsewhere, such as in the compute resources allocated to the job.

Once the bottleneck has been identified, there are several strategies for improving performance. One approach is to resize the VM instances used to run these jobs. By selecting larger VMs with more compute resources, organizations can process jobs faster and handle a higher IOPS rate, reducing delays. Additionally, upgrading to the latest generation of CPUs can provide a significant performance boost, as newer CPUs are optimized for handling intensive data workloads and offer improved data transfer rates. Another solution is to optimize the storage infrastructure itself. For example, upgrading to faster storage options within the cloud provider's offerings can help resolve I/O bottlenecks. Switching to a higher tier can result in faster data access and improved job completion times. However, this approach may also increase costs, so it's important to evaluate whether the performance gains justify the added expense.

A more comprehensive solution involves leveraging Silk's consistent performance architecture, which can isolate and resolve storage bottlenecks in the cloud. Silk's architecture is designed to provide consistent performance even during peak business periods, preventing the runtime inconsistencies that often occur with native cloud storage options.

By implementing these solutions, organizations can ensure that their long-running batch jobs and analytical workloads complete within their designated windows, minimizing operational disruptions. In turn, this enables teams to access processed data on time, make informed decisions, and avoid costly delays. Optimizing batch job performance not only improves the efficiency of critical business processes but also enhances the overall reliability of the cloud infrastructure, making it easier to scale as data volumes grow.

Cloud Storage Issue #3



Escalating Cloud Costs

One of the most significant and ongoing challenges for organizations using cloud infrastructure is the escalation of cloud costs over time, often referred to as “cloud cost creep.” This occurs when the cost of maintaining cloud environments, particularly for storage and compute resources, steadily increases, sometimes without a clear or immediate reason. What starts as an affordable, scalable solution can quickly turn into a financial burden as organizations attempt to scale their resources to maintain performance. This is particularly problematic in environments with heavy data workloads or fluctuating traffic, where increasing demand often leads to overprovisioning of cloud resources.

As previously discussed, adding more VMs or upgrading to higher performance storage tiers can address performance issues. However, this approach, while effective in the short term, can result in runaway cloud bills as these resources add up. Unlike on-premises infrastructure, where scaling up might involve a one-time investment, cloud resources are billed on a pay-per-use basis. As a result, each additional compute instance or gigabyte of storage adds to the organization’s monthly expenses. If this isn’t carefully managed, cloud costs can spiral out of control, putting significant strain on the company’s budget.

For many organizations, the tipping point occurs when costs rise faster than projected, often driven by increased storage needs, unoptimized data workloads, or inefficient configurations. In some cases, organizations don’t even realize the full extent of the cost increase until it’s too late to make meaningful changes without major disruptions.

Escalating cloud costs can have far-reaching impacts across an organization, affecting both its financial health and its long-term strategic decisions. When cloud bills grow unpredictably, it becomes difficult to forecast IT spending accurately, leading to budget overruns that catch finance teams off guard. The more resources are added to maintain performance, the harder it becomes to justify the escalating cost. Often, this leaves organizations in a tough position: they either continue to bear the rising costs or are forced to consider drastic measures, such as scaling back operations, reducing service levels, or even migrating back to on-premises or hybrid solutions. Each of these options carries its own set of challenges and additional costs.

Overprovisioned VMs and storage are common culprits of rising costs. Many organizations, in their rush to ensure smooth operations and avoid performance bottlenecks, opt for larger or more powerful VMs than they truly need. These overprovisioned instances, often combined with underutilized storage resources, result in a cloud environment where the actual usage is much lower than the allocated capacity. This leads to the organization paying for more resources than they actually use, a common but costly mistake in cloud environments. Additionally, attached storage often scales automatically as data grows, creating even more cost without immediate visibility into the full extent of the growth.

Another less obvious impact of cloud cost creep is the strain it places on strategic planning. When IT leaders are constantly fighting to keep costs under control, it detracts from their ability to focus on innovation, growth, and long-term transformation projects. The financial pressure of cloud cost escalation can also prevent organizations from taking full advantage of the cloud’s potential, limiting their ability to experiment with new technologies like artificial intelligence (AI), machine learning, or advanced analytics due to budget constraints.

Unfortunately the problem is compounded because storage costs are the part of cloud infrastructure that benefits least from discounts and committed use or reservation-based cost efficiencies.



Here's how to solve it!

To combat cloud cost creep, organizations need to adopt a proactive approach that includes both better resource management and advanced tools that optimize cloud infrastructure usage. The first step is “right-sizing” cloud resources, particularly VMs, to ensure that the organization is only paying for what it needs. This involves closely analyzing workload requirements and adjusting the size and number of VMs to match actual demand. Rather than overprovisioning to avoid performance issues, organizations can use performance monitoring tools to dynamically allocate resources based on real-time needs. This way, they only scale up when necessary and scale down when demand decreases, which can lead to significant cost savings.

Optimizing storage resources into tiers is another key strategy. By classifying data according to its access needs, organizations can move less frequently accessed data to lower-cost storage tiers, while ensuring that high-demand data remains in faster, more expensive tiers. This approach balances performance with cost and helps prevent unnecessary expenses from accumulating over time.

Additionally, organizations should review and adjust their data lifecycle management policies. Often, cloud environments accumulate unnecessary data over time—logs, backups, or obsolete files that no longer need to be stored in high-performance (and high-cost) environments. By establishing automated data deletion policies or moving infrequently used data to cheaper archival storage, organizations can reduce the volume of data consuming high-cost resources.

A more advanced solution to controlling cloud costs is leveraging a software-defined storage platform like Silk, which can dramatically reduce the need for overprovisioned VMs and storage. By eliminating the need to rely on more expensive VM configurations for performance, Silk helps organizations keep cloud costs under control while still delivering the necessary levels of service.

The unique advantage of Silk is its ability to provide the same—or even better—performance with fewer compute resources, which directly translates into lower operational costs. By optimizing storage I/O and improving throughput, Silk allows organizations to reduce their reliance on oversized VMs and high-performance storage tiers. This not only helps reduce immediate costs but also decreases the ongoing expenses related to database and operating system licenses, both of which can increase significantly with larger cloud resource allocations.

Moreover, Silk’s ability to scale resources dynamically ensures that organizations only use what they need, when they need it, without the risk of long-term overprovisioning. This flexibility can lead to substantial savings, particularly for organizations that experience fluctuating workloads or have seasonal demand spikes, such as retailers during holiday shopping periods.

Managing escalating cloud costs requires a strategic, data-driven approach to resource allocation, combined with the right technology to optimize performance and reduce unnecessary expenses. By focusing on right-sizing VMs, optimizing storage resources, and leveraging solutions like Silk, organizations can rein in cloud spending, regain control of their budgets, and unlock more value from their cloud investments.

Cloud Storage Issue #4

Availability and Resiliency Issues

One of the most significant challenges organizations face when using cloud infrastructure is the risk of planned or unplanned downtime causing disruption to business-critical workloads. While the cloud is designed to provide flexibility and scalability, it is not immune to system downtime caused by maintenance events, hardware failures, or issues with cloud infrastructure providers. These events—whether planned, such as scheduled upgrades or hardware replacements, or unplanned, such as system crashes or zone failures—can severely disrupt operations, leading to database resets, system stalls, and extended periods of downtime.

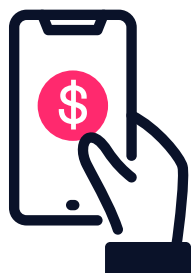
Cloud environments are often configured with automatic recovery mechanisms, which can mitigate minor issues but may fall short when it comes to more complex, large-scale failures. When database systems are reset or moved to another zone due to infrastructure issues, it can result in hours of downtime, where mission-critical applications and databases are either offline or inaccessible. The broader the cloud infrastructure or the more distributed the data, the more complex recovery becomes, especially for industries where continuous data access is non-negotiable.

For many organizations, the impact of these disruptions can be catastrophic. During downtime, essential business services grind to a halt, causing not only operational inefficiencies but also financial loss and damage to the organization's reputation. In sectors such as healthcare, financial services, or retail, where real-time data access is essential, even brief outages can have far-reaching consequences.



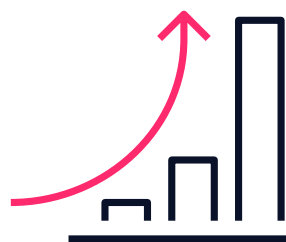
Industry Examples

Take, for example, a healthcare provider managing patient data during a natural disaster. Hospitals and clinics rely on real-time access to medical records, test results, and patient histories to provide timely and accurate care. If a cloud database goes offline due to an infrastructure failure, healthcare providers are left without the vital information they need to make critical decisions, potentially compromising patient outcomes.



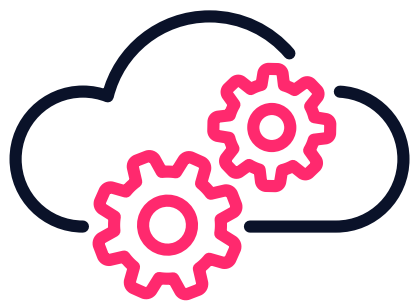
In the retail industry, businesses are particularly vulnerable during peak shopping seasons, such as Black Friday or Cyber Monday.

Retailers depend on seamless access to customer data, inventory systems, and payment processing to handle the surge in traffic and sales during these high-demand periods. If a retailer experiences a system stall or a zone failure in the cloud, it can result in missed transactions, lost sales, and dissatisfied customers, all of which negatively impact the bottom line. Beyond immediate financial loss, downtime during peak business periods can lead to long-term reputational damage, as customers may turn to competitors if they experience delays or service interruptions.



In financial services, downtime during critical periods such as end-of-quarter reporting, high-frequency trading, or processing payroll can disrupt business cycles and cause delays in reporting or compliance.

Financial organizations are highly dependent on real-time data processing to meet regulatory obligations, and downtime can cause cascading effects across other business units, putting the organization at risk for penalties and legal action. The ripple effect of downtime can also extend to internal teams, such as IT and support staff, who must scramble to diagnose and fix the problem while the rest of the business experiences delays and frustrations. Employees may be unable to access the systems they need to perform their jobs, resulting in wasted hours and decreased productivity.



Here's how to solve it!

To mitigate the risks of planned or unplanned downtime in the cloud, organizations need to implement robust, highly available configurations that ensure business continuity even during infrastructure failures. One of the most effective solutions is leveraging technologies like Oracle's DataGuard, which provides data protection and disaster recovery by maintaining standby databases that can take over in the event of a failure. DataGuard enables organizations to set up Active DataGuard configurations, allowing real-time synchronization between primary and standby databases across multiple cloud zones or regions. This ensures that, even in the event of a failure in one zone, the system can quickly switch to a secondary zone, minimizing downtime and maintaining business continuity. However, ensuring high availability goes beyond simply configuring backup databases. Organizations need to invest in a cloud architecture that can handle failures with minimal disruption. This is where Silk's resilient architecture can play a pivotal role. Silk is designed to provide continuous availability of data, even during infrastructure failures, making it an ideal solution for organizations that cannot afford to go offline.

Silk's architecture allows data to be spread across multiple nodes and duplicated between zones, ensuring that the loss of one node or zone doesn't impact the availability of critical applications. By providing near-instantaneous failover and recovery, Silk minimizes the downtime caused by planned maintenance events or unexpected failures. For example, if a database is reset due to an underlying hardware issue, Silk ensures that the data remains available without noticeable disruption to the end-user.

Silk always works with customers to ensure critical databases are deployed using configurations that are resilient to failures within a cloud zone or region. This is especially important for organizations with global operations or those that serve large, distributed customer bases. By providing real-time redundancy across different geographic locations, Silk ensures that applications and databases are always available, reducing the impact of outages or maintenance events.

In addition to high availability, Silk's architecture also simplifies disaster recovery efforts. Instead of relying on time-consuming backup restoration processes or complex failover configurations, organizations can leverage Silk's instantaneous snapshots and clones to quickly recover from failures. These snapshots allow businesses to revert to a previous state of the database in minutes, without needing to manually restore data from backups. By combining Silk's resilient storage architecture with technologies like Oracle's DataGuard, organizations can achieve true high availability in the cloud. This approach ensures that even during the most severe infrastructure failures, business-critical workloads remain online, preventing the financial loss, reputational damage, and operational inefficiencies that result from downtime.

Planned or unplanned cloud downtime poses a significant threat to business continuity, but organizations can minimize the risk through highly available database configurations and resilient cloud architectures. By leveraging tools like Silk, businesses can ensure continuous availability of data and maintain high levels of performance, even during the most challenging of cloud infrastructure failures. This not only protects against financial loss but also ensures that businesses can continue to operate smoothly, no matter what disruptions occur.

Cloud Storage Issue #5

Lack of Data Agility

Mission-critical data is no longer confined to production environments. Organizations need to leverage this data across multiple functions, including development, testing, analytics, and AI models. However, managing and distributing large volumes of data for these purposes can be both time-consuming and costly. The process of creating multiple copies of data often involves duplicating massive datasets across environments. This is a resource-intensive task, particularly as data volumes grow into the terabytes or even petabytes range.

For many organizations, the traditional approach of copying data to support non-production workloads introduces significant delays and expenses. Each copy requires additional storage capacity, network bandwidth, and time to complete, especially when dealing with complex datasets that span multiple databases or geographic locations. Moreover, these processes often require coordination between various teams (e.g., IT, development, operations), which adds to the complexity and length of time needed to complete data replication tasks. As organizations scale, the cumulative cost and operational burden of managing multiple copies of mission-critical data becomes unsustainable, leading to inefficiencies that stifle innovation.

In addition to the challenges of creating and managing these data copies, organizations also face security concerns. Sensitive data such as personally identifiable information (PII), financial records, or healthcare data must be masked or otherwise protected before it can be used in non-production environments. This adds another layer of complexity, making it difficult for teams to quickly access data without violating compliance requirements or exposing the organization to security risks. In industries like finance and healthcare, where strict data privacy regulations are in place, the inability to securely share data across non-production environments is a significant liability.

The inability to rapidly and efficiently spin up secure copies of mission-critical data for non-production uses can have far-reaching consequences across an organization. Delays in data availability hinder the ability of development teams to test new features, slowing down the software release cycle and delaying time-to-market for new products or services. Similarly, data scientists and analysts, who rely on fresh, up-to-date data for their models and reports, may be forced to work with outdated datasets, leading to inaccurate insights or missed opportunities for optimization.

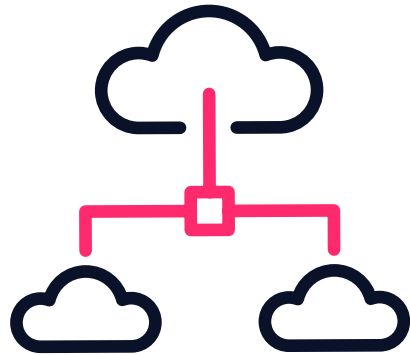


Industry Examples

A healthcare organization that is developing a new patient management system may need to test its application against real patient data. If the IT team is unable to quickly create secure, up-to-date copies of this data for testing purposes, the development process will be delayed, slowing down the implementation of the new system.



Similarly, an insurance company building AI models to predict claim fraud may not have access to the most recent data, resulting in less accurate models that fail to detect emerging patterns in fraudulent activity. These delays and inefficiencies ultimately hinder the organization's ability to remain competitive in a data-driven world.



Here's how to solve it!

To address these challenges, organizations need a solution that can streamline the process of creating and distributing data copies while maintaining security and compliance standards. One such solution is Silk's thin clone technology, which allows organizations to create instantaneous, zero-footprint copies of data without incurring the high storage costs traditionally associated with duplicating large datasets. This technology dramatically improves data agility by enabling teams to quickly generate copies of mission-critical data that can be used for Dev/Test, analytics, or AI, without having to wait for time-consuming data replication processes.

Silk's thin clone technology works by creating virtual copies of data that consume minimal additional storage space. Rather than physically duplicating the data, Silk leverages snapshots that reference the original dataset, allowing teams to access an up-to-date, full copy of the data without the need for additional storage capacity. These thin clones can be created and deployed across different zones or regions in a matter of minutes, allowing teams to access the data they need without delays. This means that whether an organization needs to support a development team testing a new feature, an analytics team running complex queries, or an AI team training new models, they can do so with the most recent and accurate data available, without overburdening their storage resources. Beyond improving speed and cost efficiency, Silk's thin-clone technology also enhances security and compliance by making it easy and quick to mask sensitive data in the cloned environments. This ensures that data is protected from unauthorized access or misuse. Organizations can mask PII or other sensitive data at the time of creating thin clones, allowing development teams to work with real-world data sets while remaining compliant with data privacy regulations such as GDPR, HIPAA, or CCPA. By integrating data masking into the cloning process, Silk helps organizations reduce the risk of data breaches and maintain compliance while still enabling rapid innovation.

The ability to quickly and securely create multiple copies of mission-critical data is essential for maintaining business agility and fostering innovation. Silk's thin-clone technology offers a powerful solution that enables organizations to overcome the traditional challenges of data replication, providing fast, cost-effective, and secure access to data across all functions of the business. By improving data agility, Silk empowers organizations to innovate faster, respond more effectively to changing market conditions, and ensure that their data-driven initiatives are always powered by the most accurate and up-to-date information.

Cloud storage offers numerous advantages, but without careful management, organizations can face performance, cost, and availability challenges. By addressing these major cloud storage struggles—poor performance, long-running workloads, escalating costs, resiliency issues, and lack of data agility—organizations can unlock the full potential of cloud environments. Solutions like Silk's software-defined cloud storage enable businesses to overcome these challenges, offering faster performance, lower costs, and increased agility across their cloud infrastructure.

Ready to see what Silk can do for you? Visit www.silk.us to learn more.