



# Hyperscale Compliance

10X Masking Speeds for Large Datasets

## With ramped up cyber attacks and stringent privacy regulations, protecting sensitive data and PII is a major priority for all businesses.

Enterprise security teams are laser focused on protecting production data in backend systems and customer facing applications, but often non-production environments are an afterthought. Non-production environments used for software development and cloud migration projects, as well as ML/AI, analytics, and reporting make up as much as 80% of an enterprise's potential attack surface. Test environments are copied over and over and accessed by many people for DevOps processes, making them vulnerable to manual errors, negligence, and malicious threats that lead to data exposure and loss.

When enterprises have big datasets and complicated data sources, it can take weeks to protect sensitive and PII data using manual processes or bulk masking to anonymize the data. This time-lag slows innovation and contributes to loss of revenue and lower customer satisfaction. Not protecting this data though can result in fines, penalties, and reputational damage if a breach occurs.

Hyperscale Compliance from Delphix provides incredibly fast masking speeds for continuous compliant data delivery for CI/CD and DevOps initiatives as well as analytics, ML/AI.

## Hyperscale Compliance In Action

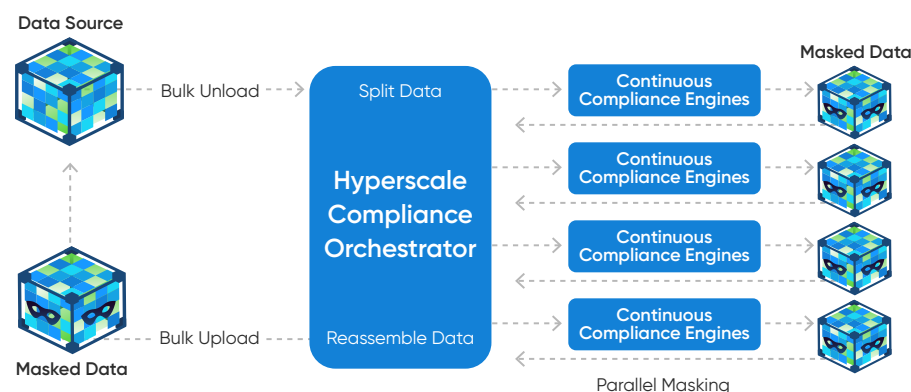
Traditional batch masking replaces the entire contents of a database table with masked data in bulk. It works well for smaller data sets, but when applied to multiple terabytes or billions of dataset table rows, the wait time for masked data is unacceptable for CI/CD speeds. Hyperscale Compliance reduces the time to mask data for large databases through increased scalability and efficiency.

Hyperscale Compliance can run on-premises or in the cloud. It distributes the masking workload for a single job across multiple virtual appliances called Continuous Compliance engines. Through automated processes, Hyperscale Compliance splits large amounts of data into digestible CSV file chunks that are then masked in parallel across multiple engines and re-assembled

into bulk upload artifacts. The engines support both on-premises storage and cloud native object storage, which provide storage efficiency and cost saving.

Hyperscale Compliance is ideally-suited for large amounts of test data used in software development, cloud migration, app modernization, and analytics initiatives. It provides masked or tokenized data that protects PII and other sensitive information to meet corporate compliance and security needs while preserving business value and referential integrity across databases.

What would have taken days or weeks to manually split data across multiple engines or to run billions of rows of data on one engine, takes only minutes to a few hours with Hyperscale Compliance.



# Hyperscale Compliance Capabilities

## Resource Efficient

Automated load balancing across engines for parallel database masking

## Faster

10X faster than traditional batch masking

## Scalable

Ideal for >10TBs of data or 1B+ dataset rows

## Flexible

On-premises storage or public cloud object stores for efficiency and cost savings

# Hyperscale Compliance Key Features

## Sensitive Data Discovery

Automatically discovers sensitive information, such as names, email addresses, and credit card numbers across a broad range of data sources.

## Automated Masking

Irreversibly masks data values with realistic data that remains fully functional, enabling the development of higher-quality code, more accurate analytics, and better AI/ML models.

## Tokenization

Tokenize data instead of masking for obfuscating sensitive information in data sent for analysis or processing, then the data can be reversed back when the processed data set is returned.

## Referential Integrity

Produces realistic values that stay consistent within and across sources.

## Algorithms/Frameworks

Preconfigured and customizable algorithms for data types and regulatory requirements.

## Support for Large and Complicated Data Sets

Enables masking of large data sets and data warehouses as well as specialized structures that require intense compute to achieve throughput.

## Support for High Volume Cloud Stores

Workflows can utilize on-premises storage or public cloud object stores (like Amazon S3, Azure Blob or Google Cloud Platform cloud storage) for intermediary storage.

## Customer Example

A leading insurance provider needed to reduce the time it took to mask ~ 4B rows of Oracle data from 36 hours to under 24.

Hyperscale Compliance, using three Continuous Compliance Engines, was able to reduce this to ~ 17.5 hours including:

- » Data unload = 3 hours
- » Orchestration = 2 hours
- » Data masking job = 8 hours
- » Data upload = 4.5 hours

They were able to beat their SLA with Hyperscale Compliance.



# Hyperscale Compliance



Delphix is the industry leader for DevOps test data management.

Businesses need to transform application delivery but struggle to balance speed with data security and compliance. Our DevOps Data Platform automates data security, while rapidly deploying test data to accelerate application releases. With Delphix, customers modernize applications, adopt multicloud, achieve CI/CD, and recover from downtime events such as ransomware up to 2x faster.

Leading companies, including Choice Hotels, Banco Carrefour, and Fannie Mae, use Delphix to accelerate digital transformation and enable zero trust data management. Visit us at [www.delphix.com](http://www.delphix.com). Follow us on [LinkedIn](#), [Twitter](#), and [Facebook](#).

©2022 Delphix

369