

IBM Big Replicate

The *Live Data* Platform

IBM Big Replicate is an unparalleled replication technology that gives you **Live Data** — consistent data everywhere, spanning platforms and locations, even for changing data at petabyte scale.

Compare the benefits

Meet stringent data SLAs

Dependable data consistency to help meet demanding enterprise SLAs across any combination of Hadoop distributions and cloud environment.

Help remove risks and complexity of traditional data movement

Managed in real time — continuous replication — without business disruption. Helps eliminate manual scripting and snapshotting.

Accommodate changes with flexible architecture

Enable multi-vendor or Hadoop strategies — no data gravity. Extensible plugins for multiple storage systems (cloud object storage, big data, NFS, physical devices, security such as Ranger, Sentry and others).

Get the most out of your compute power

Utilize passive backup clusters and convert these into active production assets (schedule different jobs across freed up capacity and ingest jobs on all sites).

Seize cost savings opportunities

Help reduce overhead associated with copies of data sitting idle and used only for backups.

Use cases

Data movement and data lake migration

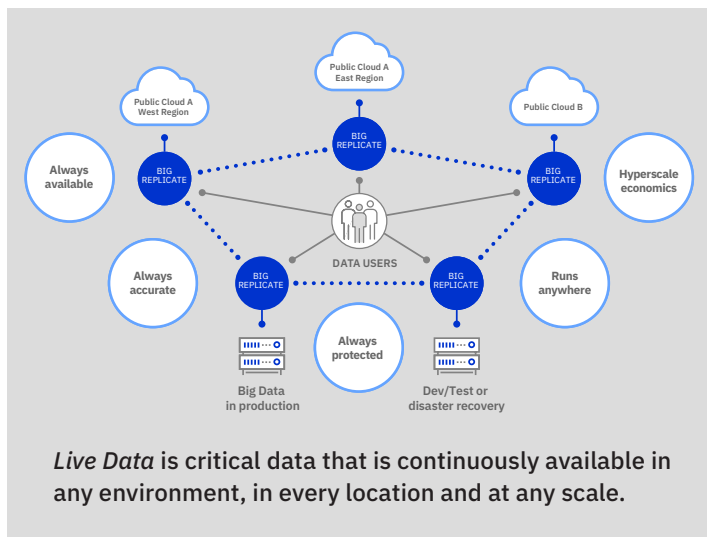
- Migrate petabyte-scale data to the cloud with virtually no downtime or data loss.
- Near real-time data consistency across multi-region environments.
- Connect any version of Hadoop to multiple clouds to help unify your data lake.

Disaster recovery, high availability and data governance

- Maintain business continuity with near-zero RPO/RTO replication and self-healing capabilities regardless of platform or geography.
- Automatic repair of missing data in the event of a failure or network outage.
- Complete control of data consistency and location to satisfy regulatory requirements.

Analytics, data discovery and experimentation

- Facilitate multi-location, multi-platform, and multi-data source analytics and development — all parties have access to the same data.
- Enable experimentation and exploration using new cloud services without breaking existing workflows.
- Help ensure data consistency for traditional BI, analytic environments and experimental ML models.



Is DistCP replication failing your SLAs?

Other replication technologies such as DistCP, CDC, dual load, master and replica and snapshot copy cannot ensure data consistency and business continuity for hyperscale infrastructure.

How it works

IBM® Big Replicate enables separate systems to coordinate and achieve consensus at scale, across wide area networks, with a common administration and management. Plugins for popular data sources help ensure compatibility with multiple systems and a robust SDK helps you build and customize your own plugins, providing the flexibility to design your optimal solution.

Supported environments

Hadoop

- Amazon EMR
- Cloudera CDH
- Google Cloud Dataproc
- Hortonworks (HDP)
- IBM BigInsights®
- MapR
- Microsoft® Azure HDInsight
- Oracle Big Data Cloud
- Oracle BDA and BDCS

Operating systems

- RHEL
- Centos
- Oracle Linux®
- Ubuntu
- SLES

Cloud

- Amazon
- Alibaba Cloud
- Google Cloud
- Microsoft Azure
- Oracle Cloud

Other

- Amazon S3
- IBM Cloud Object Storage
- Local and NFS mounted file systems
- NetApp ONTAP
- OpenStack Swift
- Oracle Object Storage (OCI and Classic)
- Virtustream Storage Cloud

IBM Big Replicate + plugins = unlimited possibilities

