
SOLUTIONS

# Critical Capabilities Checklist for SAP HANA Backup, Recovery & Cloning

As businesses have become data-driven, their need to backup, retain, and recover data on-demand has become more critical than ever. Similarly, the demand for cloning data from many internal users such as Dev, QA, UAT for testing new features, business intelligence users for analytics, and production support to accelerate root cause analysis has increased exponentially.

With data management capabilities for SAP HANA databases being so critical, there is an urgent need to ensure that data management solutions adhere to best practices and can be evaluated objectively by enterprises. After working with the top Fortune 2000 enterprises, we have created the following checklist to provide a comprehensive list of desired features.

Additionally, we provide a spreadsheet to assist your vendor evaluation process. The sheet allows you to assign a weighted score to each capability to help you better assess vendor offerings. The spreadsheet helps you compare and contrast vendors and decide the best features and solution for you.

## Backup Features

1 **Block level incremental forever backup**

Traditional backup applications perform full backups weekly, if not daily, and incremental backups on a daily basis. Unfortunately, full backups impact production HANA database performance and backup window because they create a very high storage & network IO, CPU and memory utilization.

An ideal solution relies on block-level (not file-level) incremental forever backups thus backing up just the changed blocks and leverages the SAP HANA native snapshot API 'HDBSQL' to ensure database consistency is maintained with incremental backups. Even more important is the ability to synthesize a full image quickly and efficiently after each incremental backup.

2 **Protocol flexibility**

Many environments have invested in Fibre Channel, and they prefer to leverage native Fibre Channel to perform backup, recovery and cloning instead of impacting the IP network. Other enterprises prefer IP-based technologies like iSCSI and NFS.

It's essential for a backup solution to provide the flexibility to backup, recover, or clone using any of these underlying protocols.

3 **Storage reduction**

Many environments perform HANA database dumps, using backint APIs, to a staging disk and have backup software sweep the staging disk. With recurring full backups, this increases storage consumption. The TCO gets even worse in public cloud.

A better approach would be for the backup software to do SAP HANA native snapshot API 'HDBSQL' incremental forever backups with storage compression and/or deduplication to reduce storage consumption.

4 **DBA self-service, role-based access controls**

It is very important for DBAs to have control over when and how often to backup, recover and clone HANA databases.

Role-based access control is critical in ensuring security while delivering self-service capabilities via a very simple UI, API and CLI for DBAs.

5 **Flexible backup retention**

Based on an organization's data retention policies, an admin/DBA should be able to specify data retention for days, weeks, months or years at primary and DR sites.

Basic replication, CDP or snapshot tools do not support this because they generally do not offer cost-effective long-term retention options.

**6** **Store backups in cloud object storage**

For SAP HANA databases running on-premises or in the cloud, leveraging cost-effective cloud object storage such as AWS S3/S3-IAS, Azure Blob or Google Nearline/Coldline delivers cost-effective long-term data retention for weeks, months or even years.

Another benefit is that DBAs don't have to worry about storage capacity management as cloud object storage can grow and shrink on demand.

## Database Recovery Features

**7** **Instant Recovery in minutes from older points in time**

Traditional recovery mechanisms rely on a full recovery followed by the application of incremental restores and/or transaction logs. The older the point-in-time, the longer the recovery time.

An ideal solution would have the ability to enable DBAs to recover data, in minutes, from a previous point-in-time with full automation. For example, recovering a 5TB HANA database from a day ago should take the same amount of time as a recovery from two weeks ago -- minutes.

**8** **Log backup, purge, retention, and recovery**

A backup solution should not only have the ability to do block-level incremental forever backups but also provide the capability to backup logs in between these incremental backups.

Thus the ideal solution would allow DBAs to recover, in minutes, to any point-in-time and roll forward logs with full automation.

**9** **Recover a full machine with SAP HANA database in AWS/Azure/Google cloud**

Many enterprises are leveraging public cloud for on-demand DR. Many SAP HANA databases run on on-premises physical servers or in VMs.

An ideal solution would have the ability to recover the entire on-premises physical machine in the cloud and provide instant mount and recovery of the SAP HANA database to deliver low RTO.

**10** **Cloud first approach**

Many enterprises are beginning to run SAP HANA in a public cloud such as Google, Azure,or AWS. This requires that the backup software deliver HANA backups efficiently in the cloud to minimize impact on production HANA with incremental forever backup, minimize expensive block storage in the cloud by eliminating recurring full database dumps, and leverage cloud object storage such as AWS S3, Azure Blob, or Google Nearline/Coldline for long-term data retention.

More importantly, it's absolutely critical that users be able to recover/clone HANA databases instantly, in minutes, from any point-in-time using backups stored in any storage tier in the cloud, including the cloud object storage.

**11** **Inter-Region Cloud Recovery in minutes**

Enterprises running SAP HANA databases in one cloud region typically want to replicate to another region for disaster recovery purposes. It's important for the solution to backup in one region, enable rapid recovery in that region and replicate to another region where it can also deliver rapid recoveries.

**12** **Multi-Cloud Recovery in minutes.**

Some enterprises want to protect their SAP HANA databases in one cloud and replicate to another cloud provider for DR purposes. This is is similar to inter-region recovery except that the recovery option spans public clouds. (e.g., run in Amazon and have DR in Azure). Multi-cloud recovery enables users to avoid cloud vendor lock-in and hedge their risks and costs between multiple cloud vendors.

## Database Cloning Features

**13** **Reuse backups to provision database thin-clones in minutes**

Dev, QA and UAT users need copies of production HANA databases for testing new features; BI users need copies for analytics; Production support teams need copies for root-cause analysis. Creating physical copies for so many users consumes significant time and storage.

An ideal solution would enable users to reuse backups to provision database thin-clones to all users in minutes. Note that these clones are thin, i.e., they consume no extra storage. Thin database clones only consume storage when changes are written to them, thus keeping the storage consumption low.

**14** **Self-service database cloning in minutes**

Dev, QA, UAT and BI users want self-service so they can provision HANA DB clones, on demand, from any point-in-time. Moreover, different users would like to refresh the clones at different intervals. This flexibility with self-service is critical to eliminate wait cycles and inter-dependencies amongst users.

**15** **Role based access control**

Self-service and the ability to instantly provision database clones are very beneficial. However, the solution must have role-based access control to ensure that clones can only be mounted on authorized test machines and accessed by authorized users.

**16** **Automation for sensitive data masking**

Some environments may have sensitive data in the databases that needs to be masked before clones are provisioned to testers. The solution must have the ability to leverage any masking scripts or masking tools that DBAs may already have to automate the masking process.

**17** **Efficient incremental refreshes**

Once the database clones are provisioned to testers, the test database diverges from production over time. After each test cycle (in hours, days, or weeks) testers, typically, want to refresh their clones with changes from production so they can test on the latest data in the next test cycle.

The solution must offer the capability to refresh just the incremental blocks from production to test environments to accelerate the refresh process.

**18** **API integration to popular DevOps, CI/CD tools**

It's not uncommon for enterprises to use Ansible, Chef, Puppet, ServiceNow, Jenkins or other similar tools to automate infrastructure and code provisioning in test environments. The solution should have sample plugins so architects can integrate via APIs to their existing toolchain.

## Analyst Ratings

**19** **High Analyst Ranking for Databases**

Gartner does a thorough job of compiling critical capabilities for backup and disaster recovery. Their latest report analyzes many workloads including databases. It's a good checkpoint to consider whom they rank high for Databases, Self-service, Cloud integration. You can access their document here.

**20** **Customer Reviews for Databases**

Gartner also has a peer insights website where their enterprise customers have rated their experiences with various vendors. Check out which vendors get good reviews for databases.

Also, don't forget to download this spreadsheet so you can assign priority rank and capability rank to each feature while evaluating a data management solution for SAP HANA databases.

**We hope this was helpful. If you have any questions, please don't hesitate to contact us.**

## About Actifio

Actifio virtualizes the data that's the lifeblood of businesses in more than 35 countries around the world. Its Virtual Data Pipeline™ technology enables businesses to manage, access, and protect their data faster, more efficiently, and more simply by decoupling data from physical storage, much the same way a hypervisor decouples compute from physical servers. For enterprise-class backup modernization, self-serve instant data access, or service provider business transformation, Actifio is the first and only enterprise class copy data virtualization platform.

For additional information, visit Actifio at **www.actifio.com** or contact Actifio at **info@actifio.com** or **855.886.8997**.

actifio®