

Running SAP S/4 HANA on Google Cloud Platform

August 2019

written by Remi Locherer

Senior Network Engineer

Traditionally, enterprises that operate SAP S/4 HANA on-premise buy servers that are massively oversized. Why is it like that?

When new servers are purchased, they typically should last for three to five years. Estimating the sizing to include future growth during this timeframe is very difficult. We can look back and try to predict the future based on history. But the business might develop very differently compared to this prediction. In addition, we are used to doing the sizing for the estimated peaks plus a 20% safety margin. As a result, the sizing never fits the actual needs.

With the possibility to only pay the actual use, the whole process can change. When setting up the infrastructure in the cloud the resources can be sized according to current needs. Should the resource usage change, it is comparatively easy to adapt the infrastructure (i.e. spin up beefier virtual machines or decommission not needed resources). This also reduces Capex and shifts costs to Opex, but only for the actually used resources.

Also, systems that are not needed during night or weekends can be stopped automatically which can result in savings up to 50% or even higher as you pay your compute resources “on-demand”.

Another important point is also the high amount of agility you gain when migrating to GCP. Resources are “endless” and there is no need for capacity planning for hardware resources anymore. You get what you need - anytime and everywhere. Using pre-configured templates, you can also accelerate the deployment of SAP systems to save time and money.

GCP also built an integration for SAP’s Cloud Appliance Library - an “easy to use” tool, to deploy SAP systems for training purposes as well as proof of concepts to kick-start your SAP projects.

In this article, we outline how an SAP S/4 HANA set-up on Google’s cloud platform can be created. It is a collaboration between ZeroFault Solutions which contributed the SAP expertise and ngworx.ag which is a GCP partner.

Cloud Infrastructure

Running on Google's cloud enables enterprises to benefit from Google's experience running a highly distributed and performant infrastructure very reliable and secure. In order to get support from SAP, the hardware used needs to be tested and certified by SAP. Several GCP VM instances are certified. The biggest one available today is the m2-ultramem-416 with 416 vCPUs and 11.5 TB of memory. The smallest one has 32vCPUs and 208 GB of memory. There are many options in between these extremes. It is even possible to use custom-sized VM instances which is unique on GCP among all the big cloud providers.

Related Service See how we help businesses with our cloud services:

Cloud Integration We support companies in the analysis, planning, and implementation of migrations, as well as the creation of an operational concept after the migration to the cloud. See more

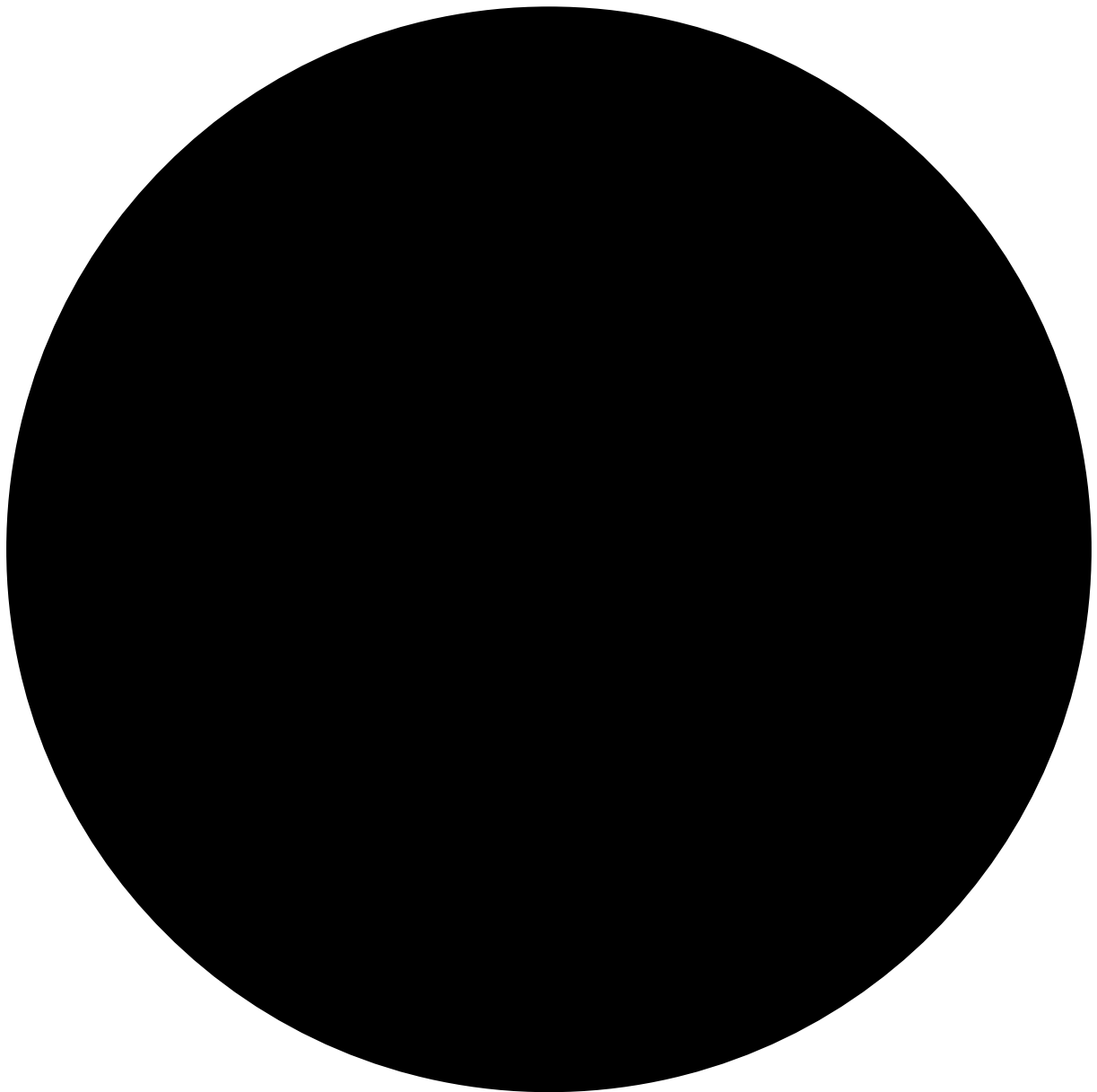
Deployment Scenarios

Depending on the load or availability requirements a deployment may look different. Three deployments for different scenarios are outlined below.

Scenario 1: Centralized Deployment

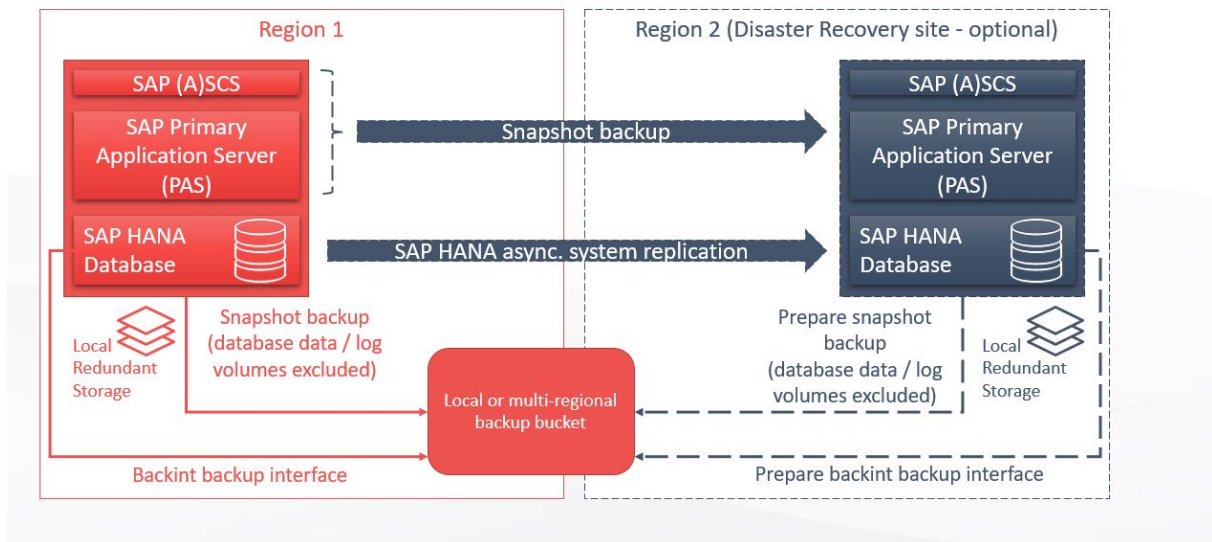
For a moderate user count and lower system loads, it is adequate to install all functions on a single virtual machine.

Even for non-cluster deployments, GCP offers good availability. Live migration of virtual servers protects from downtime caused by hardware or hypervisor maintenance. Automatic instance restart minimizes the downtime, in the rare case of a physical server failure.



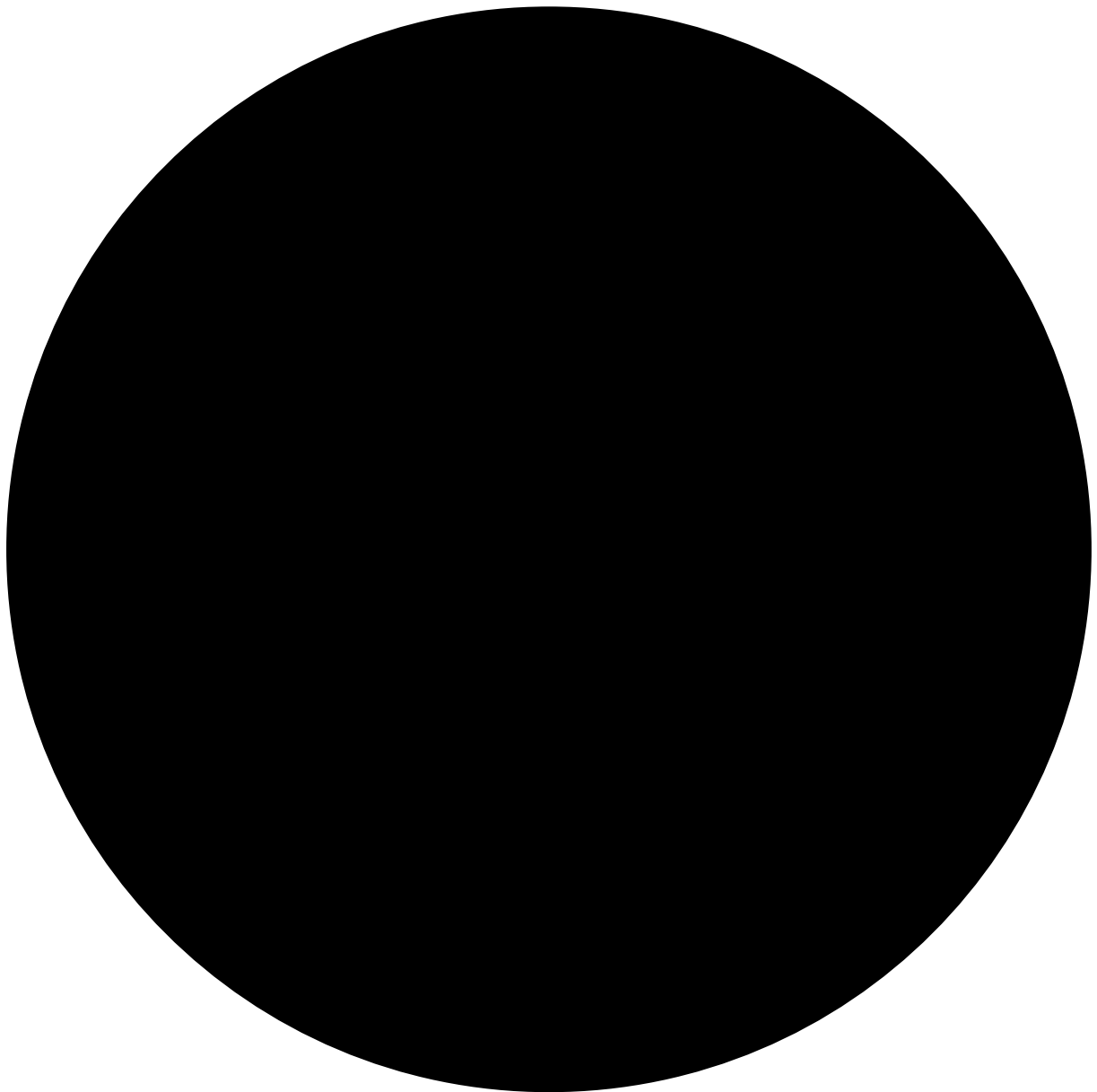
Deployment

Central System (low-load systems)



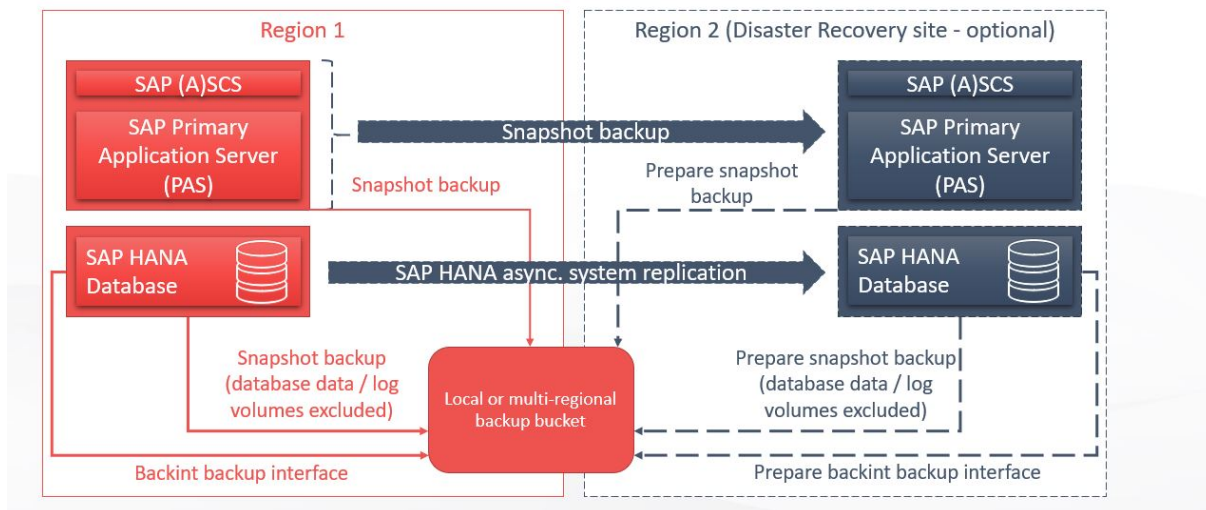
Scenario 2: Distributed Deployment

If the user count is higher and more performance is required, then a distributed deployment is recommended. Different functions get deployed to dedicated virtual machines.



Deployment

Distributed System (high-load systems)

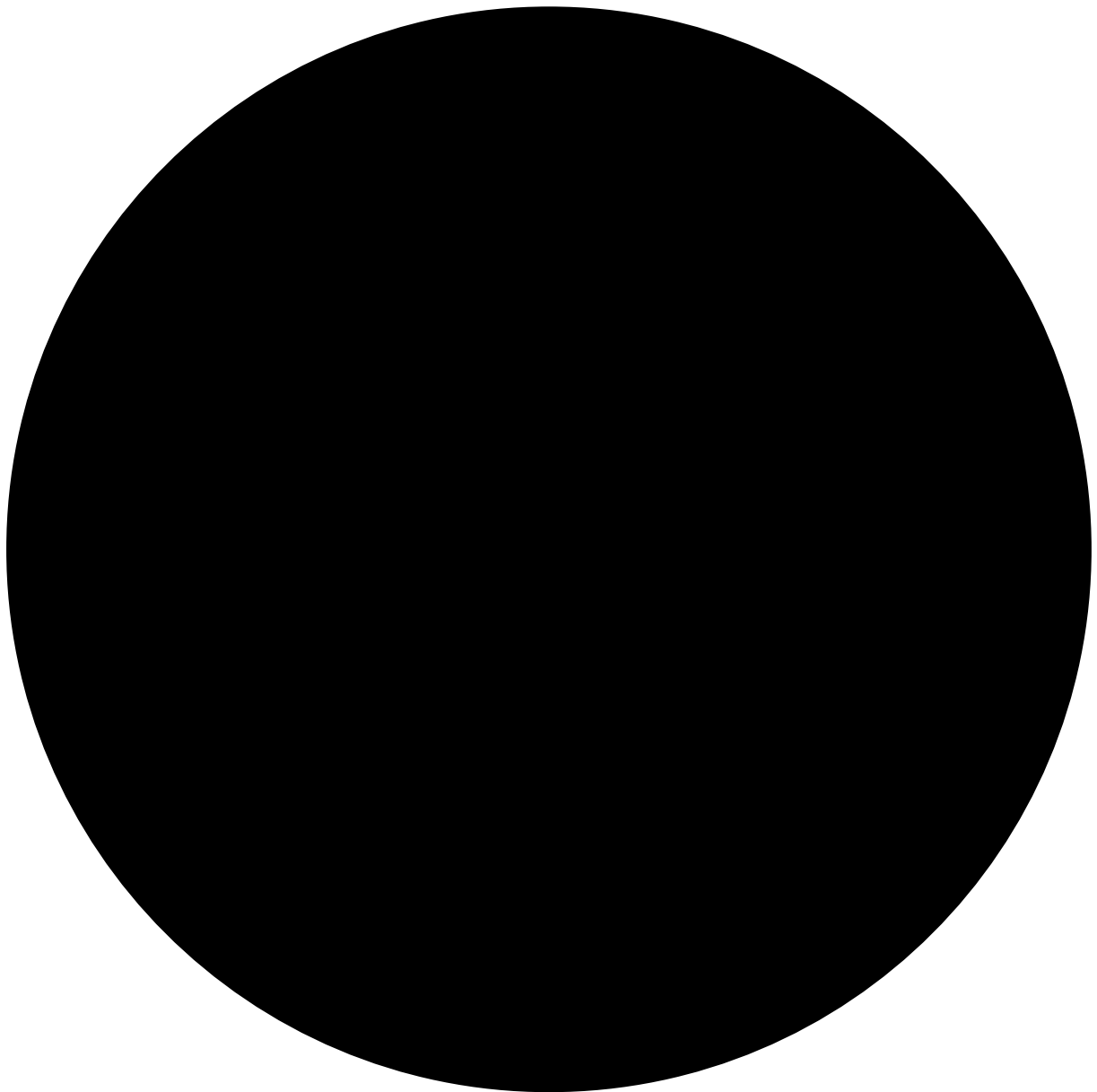


Scenario 3: High Availability Deployment

When availability is very important a high availability (HA) deployment is the right choice. Google provides a Monthly Uptime of at least 99.99% (Service Level Objective) for the infrastructure when VMs are distributed over at least two availability zones within a region.

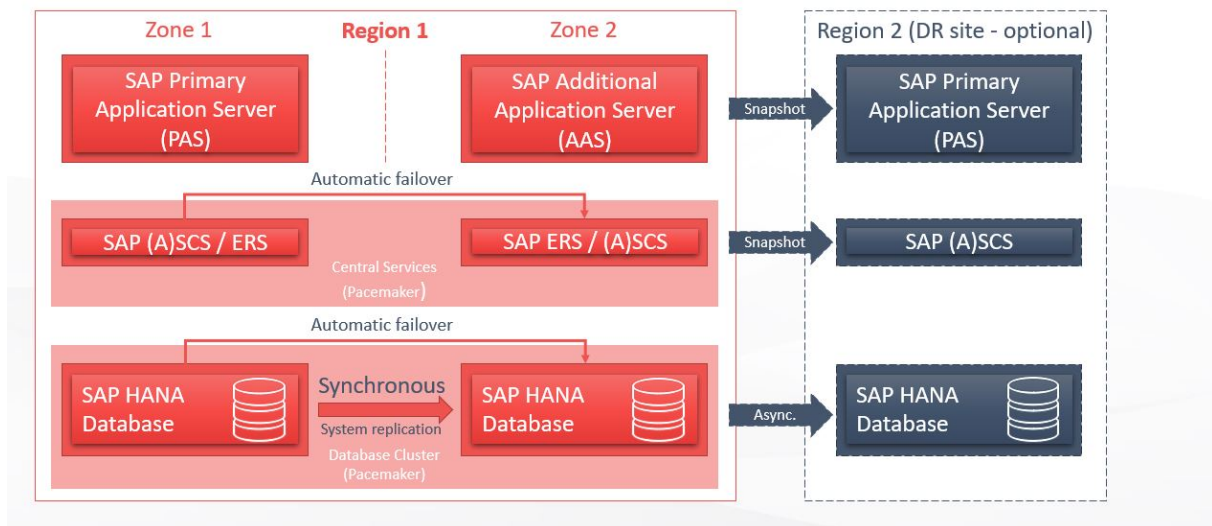
HANA databases are synchronized with the built-in system replication.

Cluster failover is done with the Pacemaker cluster tool. Since traditional clustering protocols like VRRP do not work in the GCP network Pacemaker uses GCP API calls to route traffic to the current master of the cluster.



Deployment

High Availability System



Network and Security

Regions and Availability Zones

When deploying on GCP a region and a zone needs to be selected. Regions are the specific geographic locations where Google operates its cloud data centers. For example, the region Europe-west6 is Zurich (or Lupfig to be more precise).

Each region is divided into multiple availability zones. These zones are physically separated. Between zones within a region, the network connectivity has very high bandwidth and low latency.

When deploying clustered applications, nodes should be distributed across multiple zones for higher resiliency.

Network Security Zones

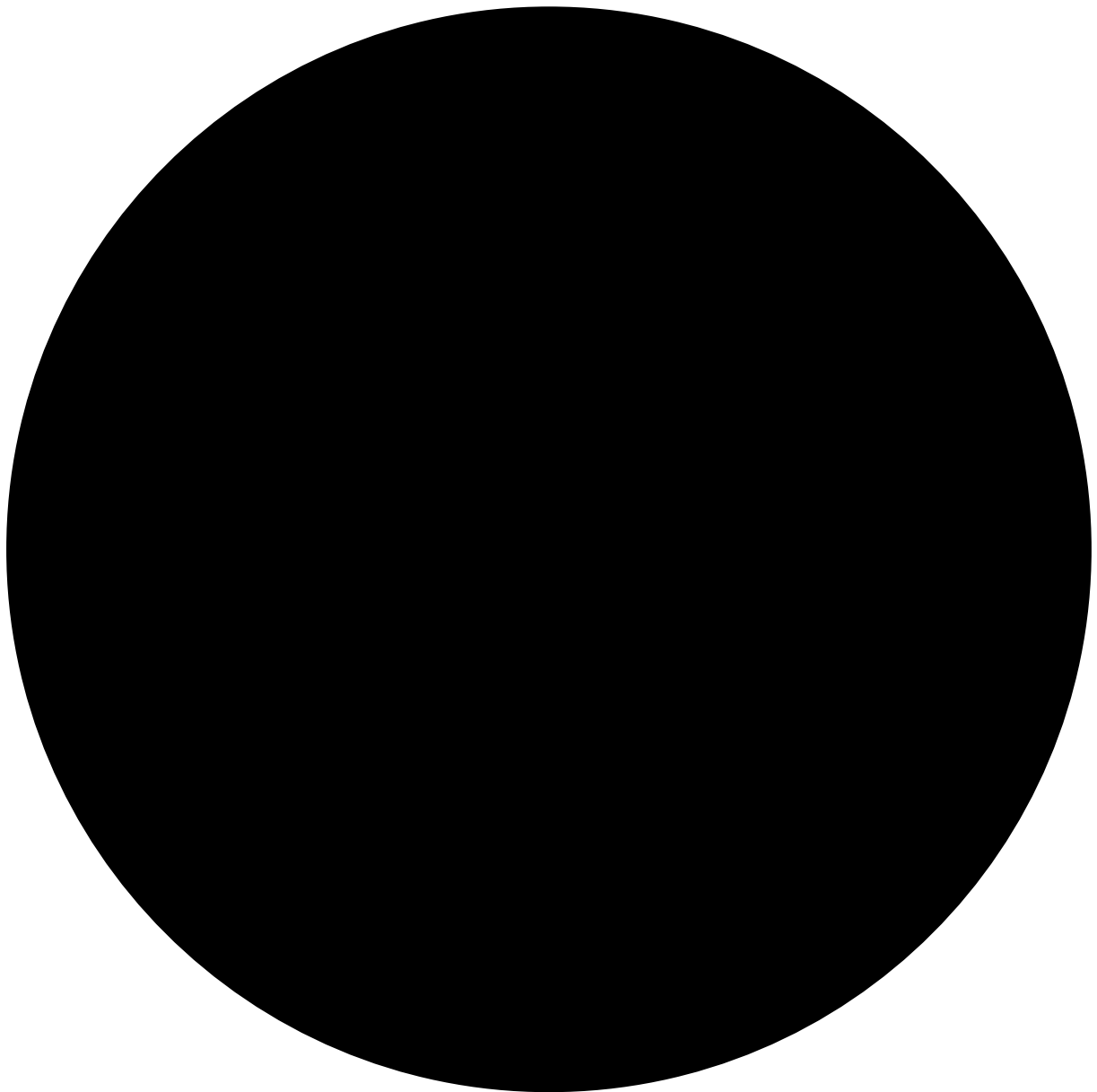
It is best practice to shield application servers from direct Internet access. On GCP, by default, virtual machines get a private and a public IP address assigned. It is strongly recommended to configure the VMs without a public address. This way their services are not directly reachable from the Internet.

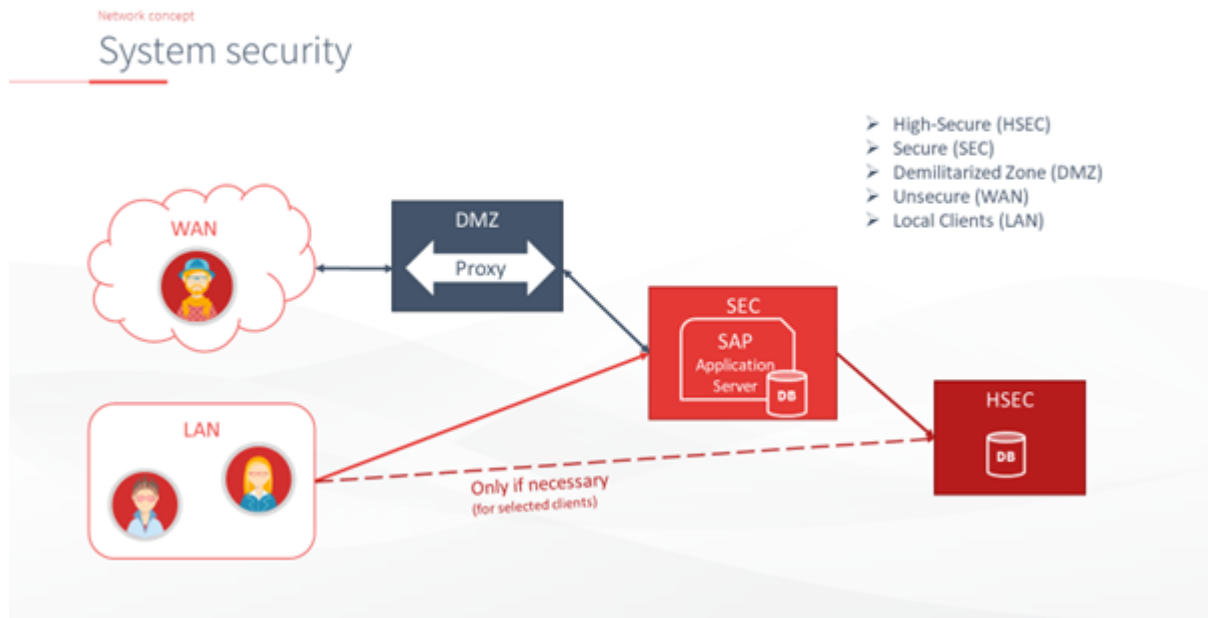
For the S/4 HANA setup we create two security zones:

- Public Zone with direct internet access (also known as DMZ).
- Private Zone without direct internet access.

Depending on the size and complexity it may make sense to introduce more zones (e.g. a dedicated zone for the database servers).

Segregation of the zones is warranted by firewall rules that are part of the standard GCP networking.





Internet Access for Servers

The application and database servers need connectivity outside of the GCP virtual network. E.g. it is necessary to download OS updates from the internet or to register licenses with vendors.

One solution for this is to use the GCPs NAT gateway. With that Internet access through this NAT can be granted to the virtual servers. The NAT function is not implemented on a virtual machine that we need to look after but is part of Andromeda, Google's software-defined networking stack.

If more control over outbound traffic is required, a proxy that scans the traffic can be deployed as a VM.

For support connections to SAP, the "SAProuter" needs to be deployed as a virtual machine in the public zone.

Secure Access

To enable secure access to S/4 HANA for the users an IPsec based VPN between the customer premises and Google can be setup. Through this VPN the users can get direct access to the private security zone. All user traffic traveling over the internet is strongly encrypted and authenticated.

The needed bandwidth of the on-premise internet connection needs to be considered case by case. But if Google's Network is designed to be highly available but depending on the required availability of the S/4 HANA system, a redundant internet connection is recommended for the on-premises site.

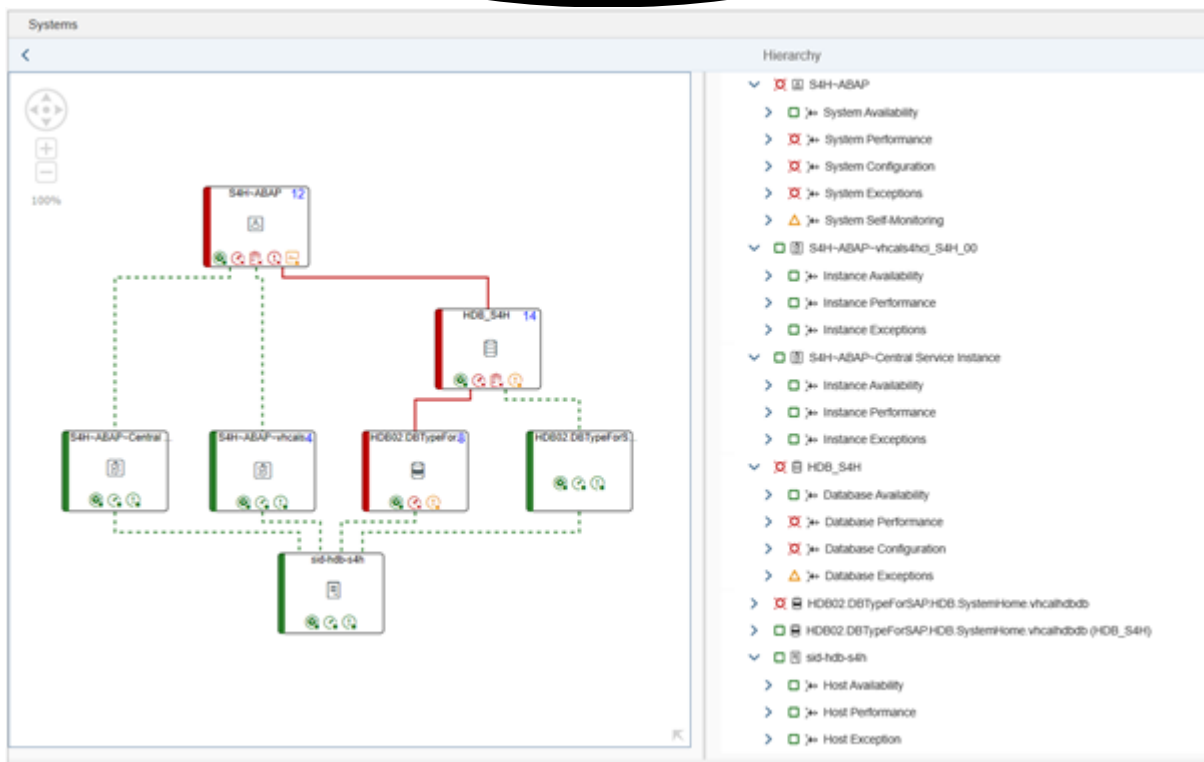
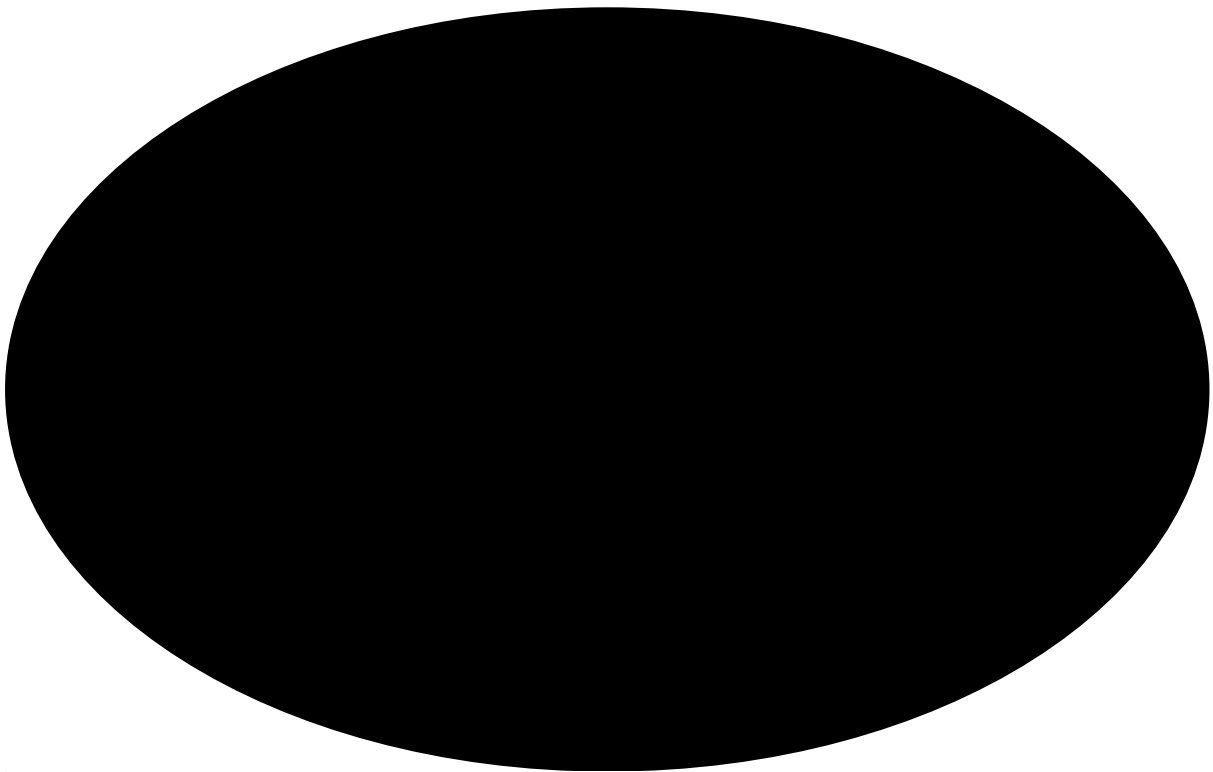
Web Application Firewall

Some services of S/4 HANA require access from untrusted Internet sources. For example, the HCM Employee Self-Service system accepts form data and uploads from job applicants. As with classical on-prem setups, such interfaces should be protected by a Web Application Firewall (WAF). A WAF terminates the encrypted connection, inspects the data and blocks malicious actions like "SQL injections" or "Cross-Site Scripting" (XSS). The WAF functionality can be implemented e.g. with A10's Lightning-ADC or with a Nginx reverse proxy with the naxsi module.

Monitoring

One of GCP's integrated solutions is Stackdriver Monitoring. Through this full visibility into performance, uptime and overall health of all services running in the cloud can be gained. It also supports basic metrics of SAP HANA.

More insights specific to the SAP software stack can be gained by using the **SAP Solution Manager**. The SAP Solution Manager offers a wide range of monitoring checks for SAP systems as well as all databases that are supported by SAP.



Also, the SAP Solution Manager delivers a lot of Application Lifecycle Management functionalities to keep your business up and running to successfully complete projects:

- Application Operations
- Business Process Operations
- Data Volume Management
- Change Control Management
- Custom Code Management
- IT Service Management
- Landscape Management
- Process Management
- Project Management
- Test Suite
- Cross Topics
- Focused Solutions

Backup

Database Backup

GCP has support for SAP's **Backint**. This is an interface defined by SAP for vendors of backup solutions to integrate with the HANA database. Backup and restore tasks can be configured and triggered through the SAP HANA cockpit or by SQL commands triggered through SQL clients (like the SAP HANA studio) or scripts. Log backups can be performed frequently and automatically by a defined interval to match your RPO needs. The backup data is securely stored in "Google Cloud Storage" buckets which is Google's object storage. These buckets offer availability in multiple regions. For the backup data, the storage class coldline can be chosen, which offers an attractive price per GB/month.

Virtual Machine Backup

To back up the disks from VMs, GCP offers a snapshot functionality. A first snapshot clones the whole disk. Subsequent snapshots only store the data that changed. This makes regular snapshots faster and much lower cost compared to full creating full images every time. To restore data a snapshot can be converted to an image which can then be attached to a VM in the same region. However, if you have complex landscapes with different databases and limited scripting knowhow or if you are running your systems in a hybrid cloud setup, a traditional backup software like Veeam Backup & Replication may be the better choice.

Disaster Recovery

While Google guarantees high availability values in its SLAs when using several availability zones within one region, enterprises might still wish to prepare a DR site for the worst case.

For this, a deployment can be prepared in a different region (e.g. Ireland or Frankfurt if the primary region is Switzerland). The use of Google's deployment manager or Terraform allows to quickly bring up all the needed infrastructure components for the DR deployment. New application servers can be started with the data from these snapshots within minutes when needed. Since volume snapshots can be stored in multi-region storage buckets, they are still available in the unlikely case the primary region is completely gone.

To minimize data loss a database instance should be up and running in the DR deployment all the time. The main database constantly replicates the data to the DR database instance. Asynchronous replication is used for this to not impact the performance of transactions.

The outlined setup provides an up and running DR set up within minutes or few hours with minimal resource usage during normal operations.

Conclusion

Using GCP to host SAP S/4 HANA results in cost-effective service with very good availability. It allows an organization to quickly react to changes and scale the infrastructure according to current needs.

A partner that is familiar with the specifics of S/4 HANA and GCP ensures a smooth and successful transition to the cloud.