TEMENOS
THE BANKING SOFTWARE COMPANY

Google Cloud

GUIDE TO:

Deploying Temenos Transact on Google Cloud Anthos for Bursting and Disaster Recovery

Financial services firms are increasingly turning to the cloud to enable digital banking capabilities while simultaneously building resilience and agility into their core business infrastructure. This trend has only been further accelerated by the upheavals of 2020, which are calling for banks to fast forward their digital transformation efforts in order to maintain business continuity through uncertain times and ensure their core applications and infrastructure are set up to handle whatever comes next in the future.

In addition to the very immediate use case of maintaining business continuity, banks also need to be able to move to the cloud at their own pace and require an application management platform flexible enough to support hybrid cloud environments and multi-cloud environments. For this reason, we conducted a Proof of Concept (PoC) for deploying Temenos Transact, our flagship core banking application on Google Cloud Anthos to explore how the solution can support bursting and failover use cases as part of a bank's business continuity plan.

In this report, we will discuss:

1. The key considerations for hybrid cloud and multi-cloud deployment models
2. How Temenos and Google Cloud support hybrid and multi-cloud deployments
3. Overview of how we set up Temenos Transact on Anthos for bursting and failover use cases
4. Key results of the PoC

Key considerations for hybrid cloud and multi-cloud

Banks expect to be able to build, manage and run their applications across on-premises, cloud, and multiple cloud environments consistently. They are asking for hybrid cloud and multi-cloud solutions, which offers a number of benefits:

Hybrid cloud

- **Maintain control over your sensitive workloads and data**. Hybrid cloud may be ideal for financial institutions that need to store sensitive data on-premises, but could still leverage the scalability and agility of the public cloud for other enterprise applications.

- **Supports ability to move to the cloud incrementally, at your own pace.** Institutions may start by putting some workloads like dev/test in the cloud or migrate customer-facing, front-office, and mobile banking applications first. Banks can also modernize core banking systems in place, containerizing them first on-premises so that it may be easier to move to the cloud in the future.

- **Flexibility to scale as needed to handle temporary capacity needs.** For very dynamic workloads, some banks may set up a cloud environment for bursting to handle unforeseen spikes in computing needs or for critical workloads they may failover to the cloud without needing to manage a physical disaster recovery site and hardware.

Google Cloud

Anthos

Multi-cloud benefits

- **Immunity from specific public cloud provider outages.** If the public cloud provider has service interruptions, the customer can continue to operate their business through a secondary public cloud provider. However, we are yet to witness a service interruption of this kind.

- **A stronger negotiation position by avoiding vendor lock-in.** A permanent lock-in becomes harder to avoid the more a customer uses the vendor's proprietary services. Even if changing vendors is theoretically possible, and the underlying architecture is flexible enough to do so, the complexity of moving out of proprietary services is considerable, from both the operational and the maintenance viewpoint. In a multi-cloud solution, the customer is no longer reliant on a single cloud provider. This empowers them to negotiate better rates with multiple cloud providers.

- **Access more innovations.** Customers can get access to a wider range of cloud innovations across all the hyperscale cloud vendors.

- **Reduced latency for end users.** With the help of content delivery networks, users can be redirected to the region that is closest to them.

- **Access to more geographies.** Not every cloud is available in every part of the world. Supporting multiple clouds provides the ability to expand business to more geographies.

Now, when it comes to managing applications across heterogeneous environments like hybrid cloud and multi-cloud, it then becomes crucial to have a single platform that allows banking IT teams to not only build banking applications, but also be able to manage security configurations, enforce policies, and monitor them across the different hybrid or multi-cloud environments. This is where the Temenos and Google Cloud partnership comes in.

How Temenos and Google Cloud support hybrid and multi-cloud deployment use cases

Temenos and Google Cloud have been working together to help banks achieve the agility and resilience required while enabling the flexibility to support hybrid and multi-cloud deployments.

Temenos became the first global banking software provider certified for Google Cloud Anthos. This open hybrid and multi-cloud application modernization platform gives financial institutions a platform for migrating and modernizing their applications, deploying easily and securely across any combination of hybrid cloud, on-premises, or multi-cloud environments.

Temenos on Google Cloud Anthos gives you the flexibility to modernize and securely run Temenos Transact core banking, solutions at scale anywhere – in data centres, in the cloud or across multiple public clouds, all with simple, unified management. It offers the following features:

- ✓ **Single pane of glass monitoring.** Viewing and managing Google Kubernetes Engine (GKE) clusters is made easy through a unified 'single pane of glass' view that also works across on-premises (bare metal or VM), hybrid, and multi-cloud environments.
- ✓ **Configuration management.** Anthos provides common configuration with policies-as-code enforced across all infrastructure, both on-premises and across clouds.
- ✓ **Centralised multi-cluster management.** Anthos manages and enforces policies across all the GKE clusters, whether on premise or in the cloud, from the Google Cloud Console.
- ✓ **Containerized applications under Anthos GKE.** Anthos runs custom-built or packaged software, including applications, containers, or functions, whether that's on premise or in the cloud.
- ✓ **Single service mesh.** Through open-source Istio, all deployments make use of the same service mesh transparently. Anthos supports dynamic request routing, canary deployments, gradual rollouts, the easy implementation of security policies, fault injection tools, as well as defining and enforcing policies as code.

Google Cloud Anthos enables consistency between on-premises and cloud environments that helps reduce complexity and accelerate application development.

The tecāology platform underpinning Temenos products enables banks to utilise the best use of new tecāologies and approaches to enable banks to function and respond optimally in all market situations.

Temenos Transact is a leading core banking application used by top global banks.

Temenos commissioned a Proof of Concept (PoC) which proved the ability to deploy Temenos Transact on Google Cloud Anthos and gain the flexibility to burst and failover to Google Cloud in cases of unforeseen spikes and outages.

More specifically, it was set up to demonstrate the hybrid deployment of Transact both on-premises and in the Google Cloud, and to validate the following scenarios:

1. **Burst into cloud.** In this scenario, an increasing workload exhausts the on premise resources and the system scales dynamically into Google Cloud. When the load subsides, the system scales back to on premise.

2. **Unplanned failover.** This scenario demonstrates the resilience of the system by introducing failure to the on premise site. We decided to test this under maximum load, during the burst-into-cloud scenario described above.

Deploy Transact on Google Cloud Anthos and gain the flexibility to burst and failover to Google Cloud in cases of unforeseen spikes and outages

Setting up the environment

In order to test the scenarios, a test environment was set up to simulate a fictional bank who has deployed their Temenos Transact core banking application running on NuoDB distributed database on Anthos and Google Kubernetes Engine (GKE). This scenario gave the fictional bank flexibility to run production workloads on-premises in their own data centre, whilst also having a secondary site on Google Cloud set up for bursting and disaster recovery.

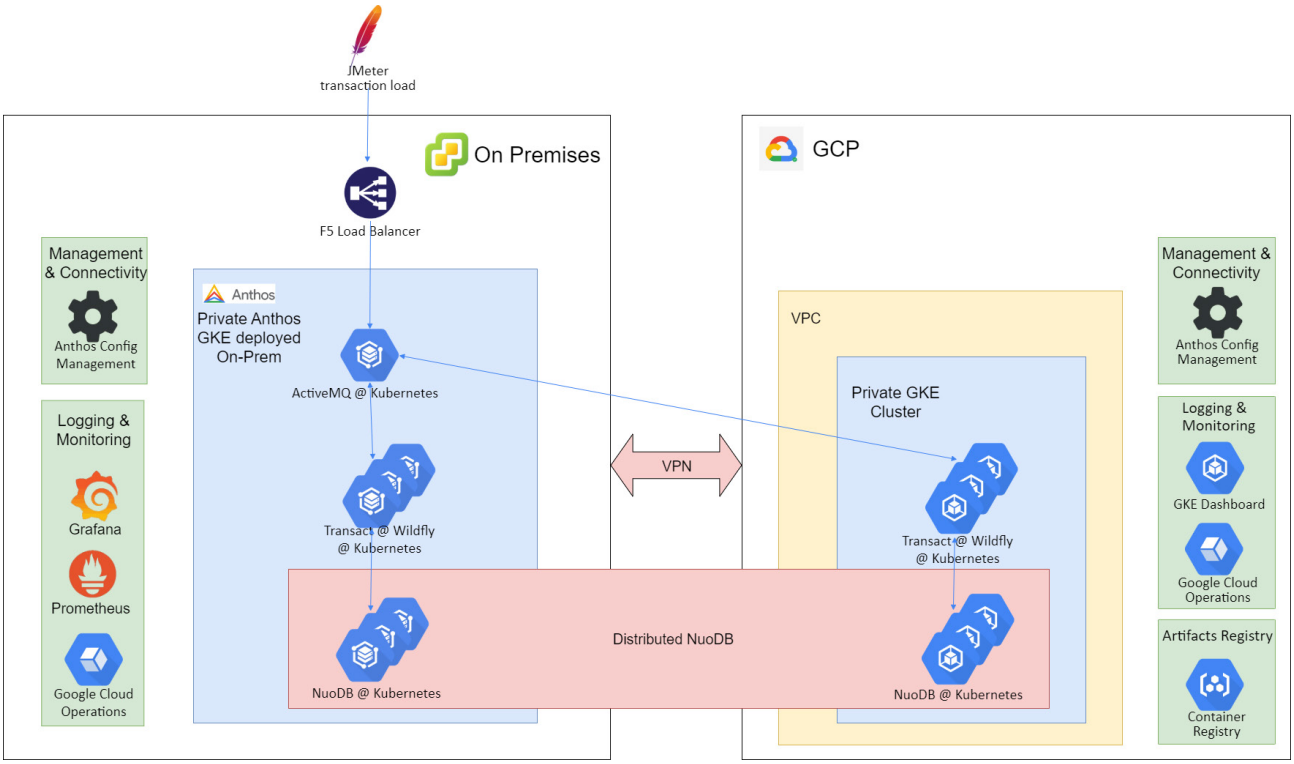


Figure 1: Google Anthos PoC

On-premises environment

Our on-premises environment deployed Google Anthos (GKE on-premises) on top of vSphere VMs. We set up the application tier of Transact in the on-premises environment, following a proven GKE-based reference architecture. The container images were stored in Google Container Registry (GCR). This allowed us to pump transactions directly into the Apache ActiveMQ message broker for Transact to consume and process.

Google Cloud environment

We set up the PoC in a very similar way in Google Cloud, except for the message broker. In both sites, Transact was configured to consume the load from the same message broker.

Configuring the load balancer

We installed and configured an on-premises F5 load balancer to forward requests to the Apache ActiveMQ deployed in the on premise environment. The two sites were connected by a VPN provided by the load balancer.

Configuring JMeter

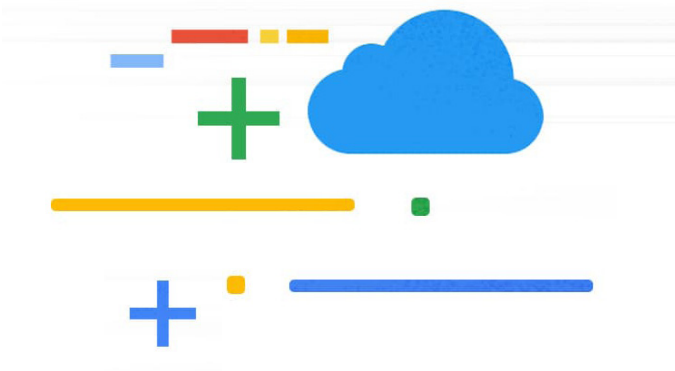
We configured JMeter to load transactions into the ActiveMQ (AMQ) message broker, based on a specific transactions per second (TPS) plan, which will be described below in the "Executing the test" section.

Setting up communications for NuoDB

NuoDB communicated with all its distributed components, across both sites, through a VPN. We achieved this through a sidecar which was deployed in all NuoDB pods, and which acted as a transparent proxy.

Setting up the Anthos GKE cluster on-premises

We deployed Anthos GKE on-premises, using Cloud Interconnect to communicate with Google Cloud. We used a single management console in Google Cloud to inspect both Anthos and GKE clusters deployed on-premises and in Google Cloud as well as all their deployments and services. We configured Anthos GKE, deployed on premise, as a 9 node cluster. Each node had 4 cores of CPU, 20GB for a total of 36 CPUs, and about 151 GBs of allocatable memory.



Setting up the GKE cluster in Google Cloud

The Google Cloud environment was a GKE cluster, deployed on two availability zones. Our testing showed that the latency between the Google Cloud environment and the on premise site was about 2 milliseconds. The GKE cluster was configured with dynamic node scalability from a minimum of 1 to a maximum 5 nodes per availability zone (that is, from 2 to 10). Each node had 8 cores of CPU and 32 GB of memory. This meant that the total capacity of the cluster could fluctuate from 16 to 80 cores and from 64 GB to 320 GB of memory.

Scaling the deployments

For the PoC scenarios that we wanted to cover, we set the scalability of each deployment as follows:

DEPLOYMENT	SCALABILITY
ActiveMQ	1 replica
Transact in Google Cloud	Minimum 1, maximum 30 replicas
Transact on-premises	Minimum 1, maximum 3 replicas
Transaction Engine of NuoDB in Google Cloud	Minimum 1, maximum 6 replicas
Transaction Engine of NuoDB on-premises	1 replica
Storage Manager of NuoDB in Google Cloud	1 replica
Storage Manager of NuoDB on-premises	1 replica

The GKE cluster in Google Cloud used a custom metric for scaling Transact. We based this custom metric on the number of pending messages on the ActiveMQ queue, divided by the number of Transact pods consumed from the queue.

We chose a different scaling metric for Transact in Anthos GKE on-premises because we needed a mechanism that would allow scaling on premise to happen preferentially. We wanted to make sure that Transact on-premises would scale based on CPU activity and that the cloud would start scaling only when it was necessary. We used the ActiveMQ queue depth as a metric because this would provide a good indicator that on premise was struggling.

Monitoring

In the PoC, we established a dashboard that would allow us to monitor both the on-premises and cloud sites.

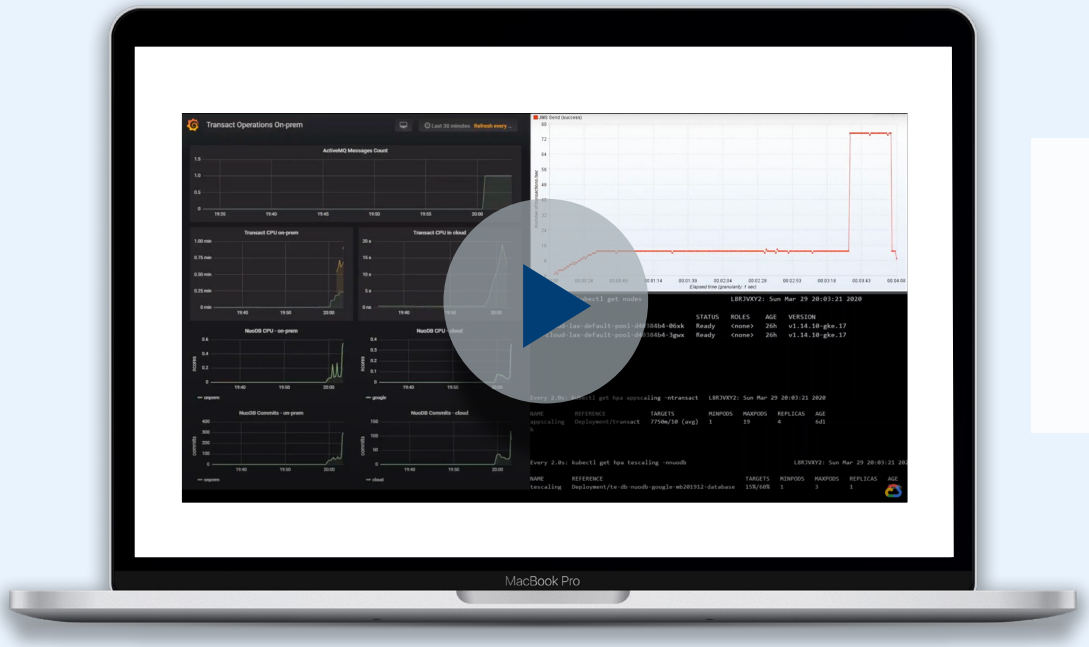
For monitoring in Google Cloud, we used the out-of-the-box functionality from Google Cloud Operations. We used Prometheus on-premises to collect all metrics from the Anthos GKE cluster on-premises and ActiveMQ. In addition, as part of the default functionality of the Anthos cluster, on-premises metrics were forwarded to provide a single view. We exposed the ActiveMQ queue depth to Google Cloud Operations in Google Cloud using a Prometheus-to-sd sidecar.

From the NuoDB installation, data was collected by default in an InfluxDB instance which was not part of either the on-premises or Google Cloud sites. Finally, Grafana consumed all of the data from Prometheus, Google Cloud Operations, and InfluxDB.

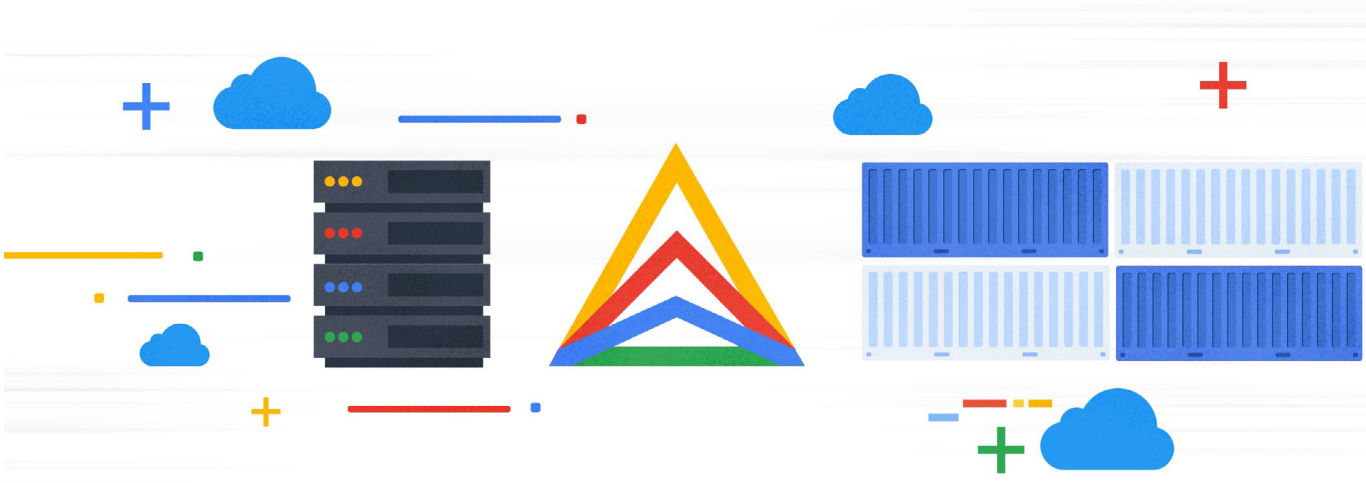
Executing the test

After we set up the system, we started to load transactions from JMeter into the ActiveMQ (AMQ) message broker, based on a specific TPS plan, to test the following phases; normal operation, burst out, failover, and then return to normal operation.

A breakdown of the executed test can be viewed in the following demo.



WATCH THE DEMO



Normal operation

To begin with, we applied an increasing load up to a normal point that the on-premises environment is designed to handle. As a result of the increasing load, the on-premises environment started to scale.

Burst out

To test burst out, we increased the load suddenly from 15 to 75 TPS and then settled to 40 TPS. This was designed to be a sudden and unforeseen event.

The three replicas on-premises were not enough to handle the load, so the system reacted correctly by requesting 6 more Transact replicas in the cloud. When this metric goes above 10 replicas, the Horizontal Pod Autoscaler (HPA) requests that more replicas be added. Entering the next phase of the plan, the load is further increased, resulting in requests for more replicas. The nodes are increased accordingly, and the system also requests a new Transaction Engine (TE) for the database.

The load reaches its peak value at 270 TPS, which is close to the maximum that the system can handle. The maximum number of TE replicas are ordered. This results in an order for one more node to increase GKE capacity.

Failover

At this point, we simulate an on-premises failure and the on-premises site stops responding. Google Cloud was shown to handle the entire load without problems.

Return to normal operation

The load gradually returns to normal levels. Transact in Google Cloud is the first to scale down to minimum, followed closely by NuoDB.

Results and Conclusion

In conclusion, we found that the hybrid cloud deployment with Google Cloud Anthos responded in a way that enabled the agility and resiliency expected.

In the **burst-into-cloud** scenario we were able to successfully simulate an environment where an increasing workload exhausts the on-premises resources. As a result, we found that the system will scale dynamically into Google Cloud only when the on-premises site was no longer able to cope alone. Once the load returns to normal levels, Google Cloud scales back to a minimum.

In the **unplanned failover** scenario we proved the resilience of the system. When the on-premises went offline, Google Cloud carried out all operations in a seamless transition. When the on-premises site came back online, it resumed normal operations after the necessary database synchronisation and Google Cloud then scaled back down.

Therefore, this test was successful in confirming that Temenos Transact can be set up to burst seamlessly from on-premises to Google Cloud to handle unforeseen spikes in load. Additionally, it can failover to the cloud, and also auto scale back down again to maintain business as usual.

Temenos and Google Cloud are working together to show that by embracing cloud solutions, financial institutions are in a position to create digital banking solutions suitable for building resiliency and agility into core business infrastructure and thereby ensuring business continuity through a rapidly evolving market.

About Temenos

Temenos AG (SIX: TEMN) is the world's leader in banking software. Over 3,000 banks across the globe, including 41 of the top 50 banks, rely on Temenos to process both the daily transactions and client interactions of more than 500 million banking customers. Temenos offers cloud-native, cloud-agnostic and AI-driven front office, core banking, payments and fund administration software enabling banks to deliver frictionless, omnichannel customer experiences and gain operational excellence.

Temenos software is proven to enable its top-performing clients to achieve cost-income ratios of 26.8% half the industry average and returns on equity of 29%, three times the industry average. These clients also invest 51% of their IT budget on growth and innovation versus maintenance, which is double the industry average, proving the banks' IT investment is adding tangible value to their business.

For more information, please visit www.temenos.com.

TEMENOS

THE BANKING SOFTWARE COMPANY

temenos.com