

Projeto Data Zaad

[Projeto Data Zaad](#)

[Proposta](#)

[Cenário atual](#)

[Diagrama Cenário Atual On-Premises](#)

[Solução Proposta](#)

[Fase 1: Carga de Dados](#)

[Fase 2: Replicação, Real Time e APIs](#)

[Fase 3: Data Viz & One Cloud](#)

[Arquitetura Final](#)

Proposta

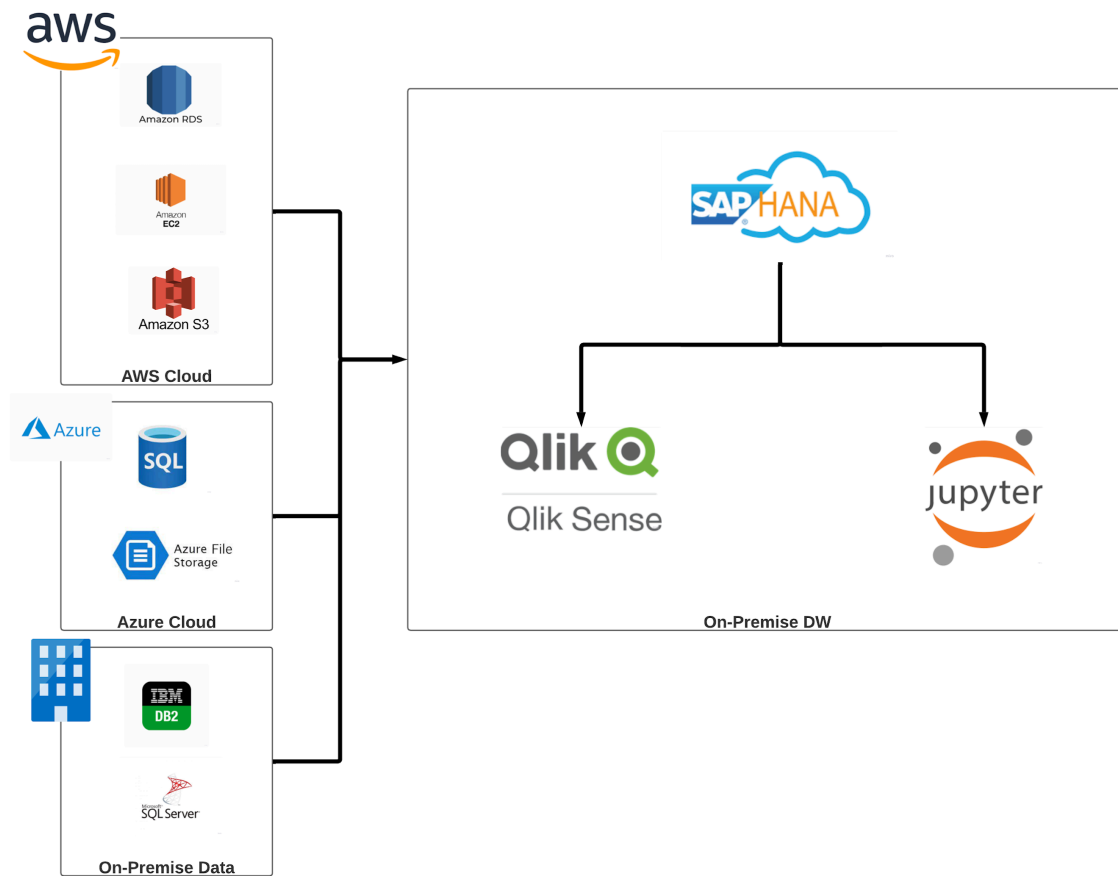
Esse documento tem por objetivo propor a migração do atual cenário de consolidação dos dados de um cenário On-Premises para um ambiente em nuvem dentro da Google Cloud.

Cenário atual

O atual ambiente consta com um ambiente On-Premise podendo conter diversos bancos de dados, além de outros alguns serviços de nuvem como Azure e AWS.

Todos esses ambientes são atualmente consolidados dentro de um estrutura On-Premise de uma solução SAP Hana

Diagrama Cenário Atual On-Premises



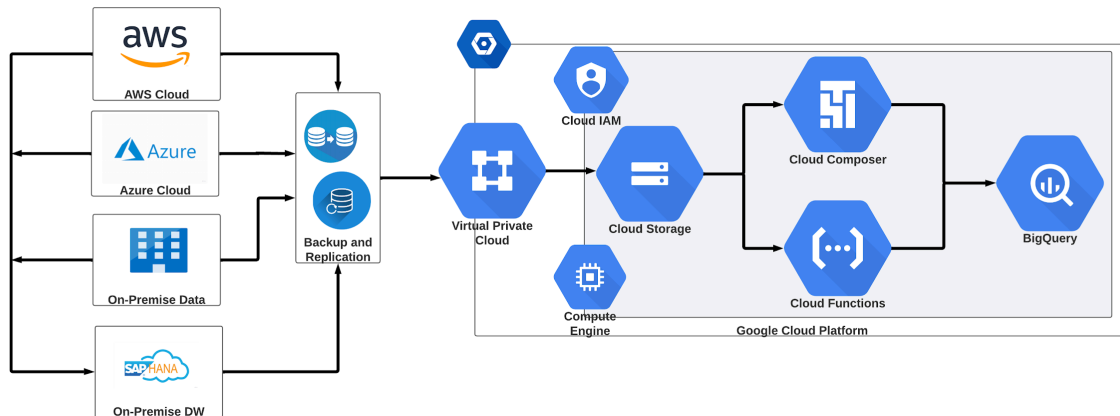
Solução Proposta

A solução proposta irá utilizar toda a estrutura do Google Cloud para realizar uma migração em 3 fases de todo o ambiente.

Abaixo o detalhamento de cada fase com os serviços a serem utilizados

Fase 1: Carga de Dados

O foco inicial dessa fase é ter todos os dados históricos carregados para dentro da estrutura de dados do Google Cloud

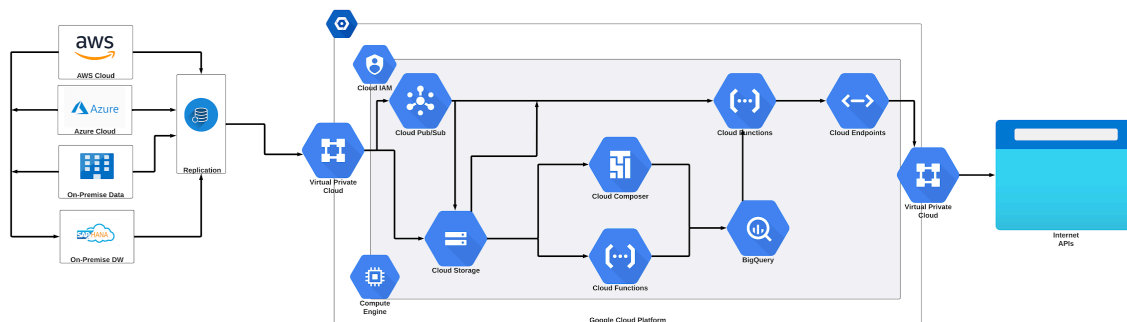


Nessa primeira fase serão realizados os seguintes passos:

- Criação de uma estrutura de replicação das bases de dados On-Premises, de outras clouds e do SAP Hana;
- Criação de backups das bases de dados clouds e do On-Premises, de outras clouds e do SAP Hana;
- Criação de uma estrutura de Virtual Private Cloud permitindo a conexão de forma segura entre o ambiente externo e a Google Cloud
- Criação de uma Compute Engine para orquestrar as conexões entre a leitura dos backup e das réplicas criadas
- Essa Compute Engine será responsável por ler das réplicas e backups e salvá-las em uma estrutura de arquivos dentro do Cloud Storage
- A criação de serviços de Cloud Composer e de Cloud Functions que serão responsáveis pela extração dos dados do Cloud Store e realizar todas as modificações e transformações necessárias
- A criação de um datawarehouse com a Big Query, onde os dados serão salvos e já estarão inicialmente disponíveis para acesso via consultas SQL
- Toda essa estrutura estará sob o controle da Cloud IAM, com todos os controles de acesso, podendo ser herdado de uma estrutura como LDAP. Dentro desse controle de acesso serão configurados desde o controle de acesso aos serviços até quais tabelas e dados um usuário poderá acessar no Big Query

Fase 2: Replicação, Real Time e APIs

O foco dessa fase é a disponibilização dos dados através dessa leitura das réplicas, via estruturas de API, além de uma possibilidade de dados em tempo real

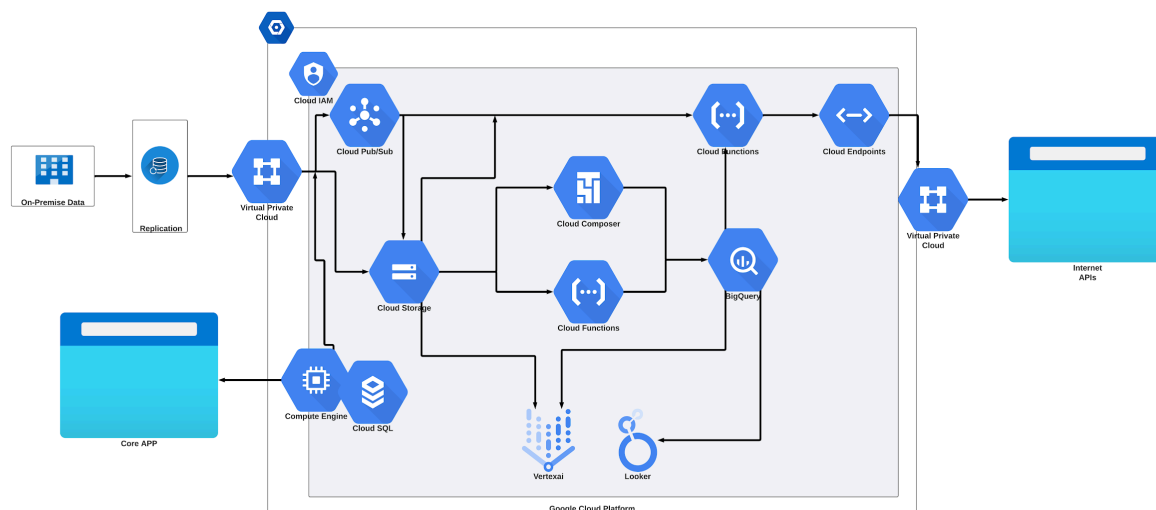


Nessa primeira fase serão realizados os seguintes passos:

- A adição de uma estrutura de Cloud Pub/Sub onde dados das réplicas dos bancos de dados serão lidos e por meios de controle de filas poderiam ser processados pelo Cloud Functions e disponibilizados para leitura no Big Query. Com isso temos a criação de uma estrutura de stream de dados, permitindo o acesso em tempo real dos mesmos.
- Dentro dessa mesma estrutura temos a adição de um Cloud Endpoint, que junto com a VPC e as Cloud Functions consegue facilmente disponibilizar os dados de forma externa em formato de APIs

Fase 3: Data Viz & One Cloud

A terceira fase vem para implementar as visualizações de dados e eliminar as estrutura legadas externa de SAP Hana e de outras nuvens, mantendo o acesso ainda as estruturas On-Premises



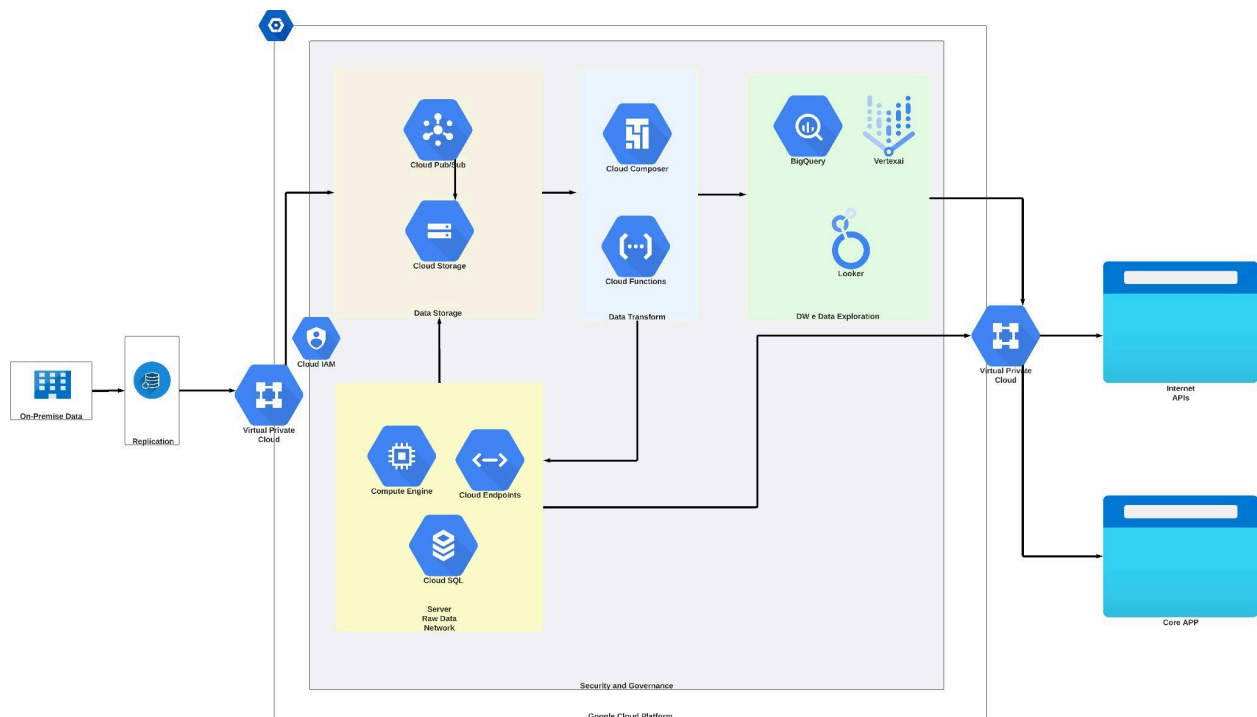
Nessa primeira fase serão realizados os seguintes passos:

- A migração das estruturas das outras nuvem para dentro de uma estrutura de Cloud SQL apoiada pela Cloud Engine, que junto com o Cloud Endpoint e VPC são externalizadas como o sistema transacional e se comunicando com a estrutura de consolidação previamente criada

- Juntamente com o Bigquery temos a adição do Looker e do Vertexai que serão utilizados para a visualização e exploração dos dados

Arquitetura Final

Organizando a arquitetura temos um cenário com a seguinte estrutura:



Essa solução final claramente permite um crescimento maior além de um total controle e governança dos dados e acessos.

Mantendo uma estrutura de cloud única e ainda permitindo a conexão com estruturas externas, sejam legadas ou novas, de forma segura e sólida a se manter as estrutura de fonte da verdade única