



Summer Meetup BDWSCL

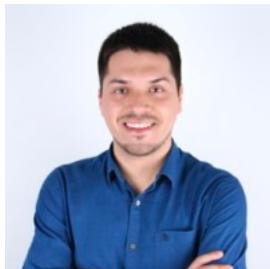
Data Warehouse Architecture: **Traditional vs Cloud**

Nicolás Lecaros

<https://www.linkedin.com/in/nicolaslecaros/>

¿Quién soy?

Nicolás Lecaros () {



Corporate IT Project Manager;



Professor, School of Computing and Telecommunications;

/* Eterno estudiante, profesor y profesional TI con +11 años de experiencia en análisis, diseño, desarrollo, implementación, soporte y gestión de proyectos TI. En donde me he especializado en Business Intelligence (BI), Business Analytics (BA), Big Data y Cloud Computing. */

}

Mi trayectoria:



Nicolás Lecaros

<https://www.linkedin.com/in/nicolaslecaros/>

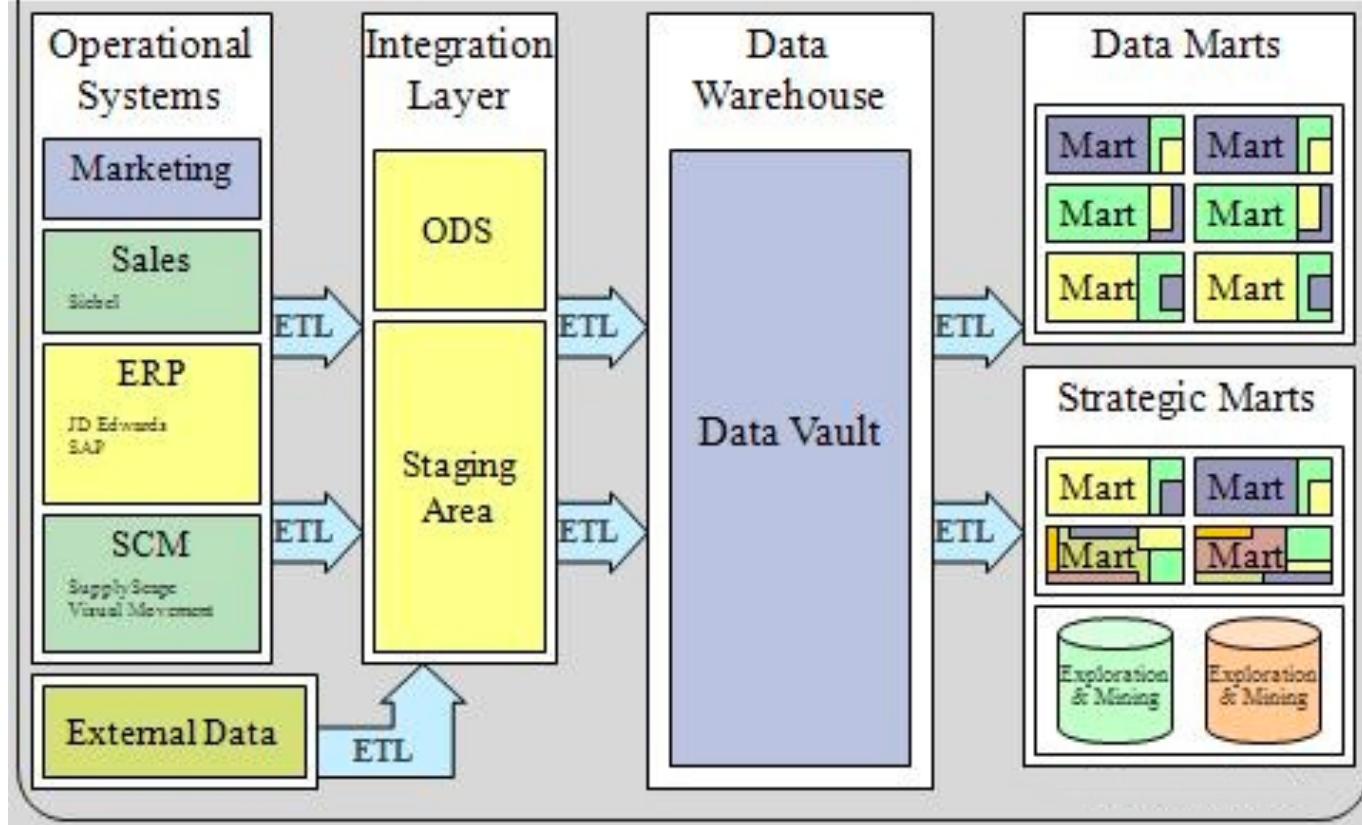
¿Qué es Data Warehouse?

Según el sitio wikipedia.org, Data Warehouse es:

Almacén de datos

En el contexto de la informática, un **almacén de datos** (del inglés *data warehouse*) es una colección de **datos** orientada a un determinado ámbito (empresa, organización, etc.), integrado, no volátil y variable en el tiempo, que ayuda a la toma de decisiones en la entidad en la que se utiliza. Se usa para realizar informes (reports) y **análisis de datos**¹ y se considera un componente fundamental de la **inteligencia empresarial**.² Se trata, sobre todo, de un expediente completo de una organización, más allá de la información transaccional y operacional, almacenado en una base de datos diseñada para favorecer el análisis y la divulgación eficiente de datos (especialmente **OLAP**, *procesamiento analítico en línea*). El almacenamiento de los datos no debe usarse con datos de uso actual. Los almacenes de datos contienen a menudo grandes cantidades de información que se subdividen a veces en unidades lógicas más pequeñas dependiendo del subsistema de la entidad del que procedan o para el que sea necesario.

Data Warehouse



¿Data Warehouse On-Premise o Cloud?



La Infraestructura On-Premise rápidamente nos queda pequeña



YOUR DATA WAREHOUSE IS BAD



AND YOU SHOULD FEEL BAD

www.meme-generator.net

Las empresas están llevando su trabajo a la nube.



¿Por qué Cloud, por qué ahora?

“Hubo un tiempo en que cada hogar, pueblo o aldea tenía su propio pozo de agua. Hoy en día, los servicios públicos compartidos nos dan acceso a agua limpia simplemente abriendo la llave.

Al igual que en la compañía de agua, *hay un equipo de profesionales dedicados que se aseguran de que el servicio proporcionado sea seguro y esté disponible 24/7*.

Cuando la llave no está abierta, no solo está ahorrando agua, sino que no está pagando por los recursos que no necesita actualmente.”



Existen muchas opciones, y sigue creciendo el mercado...



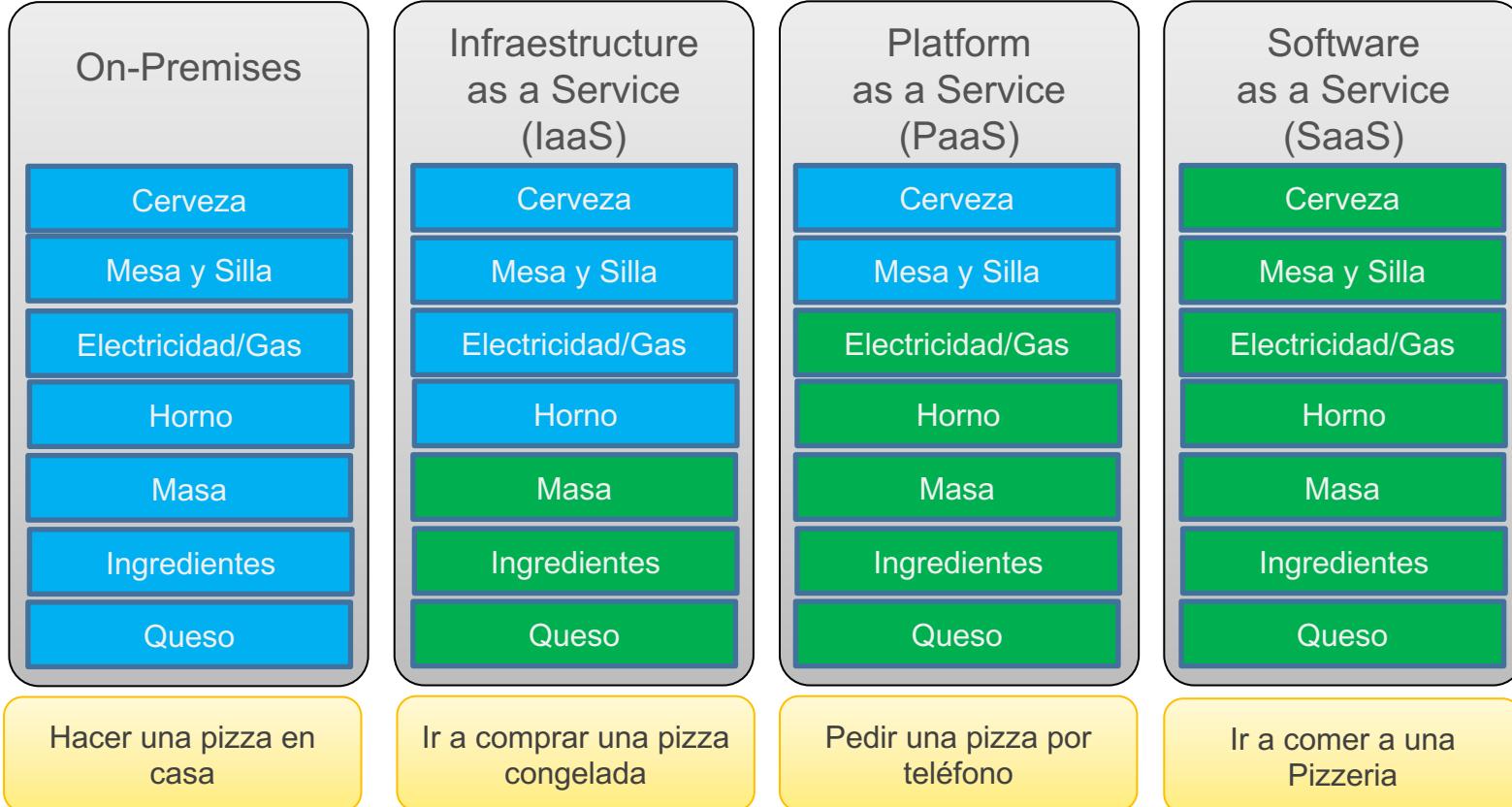


PIZZA AS A SERVICE

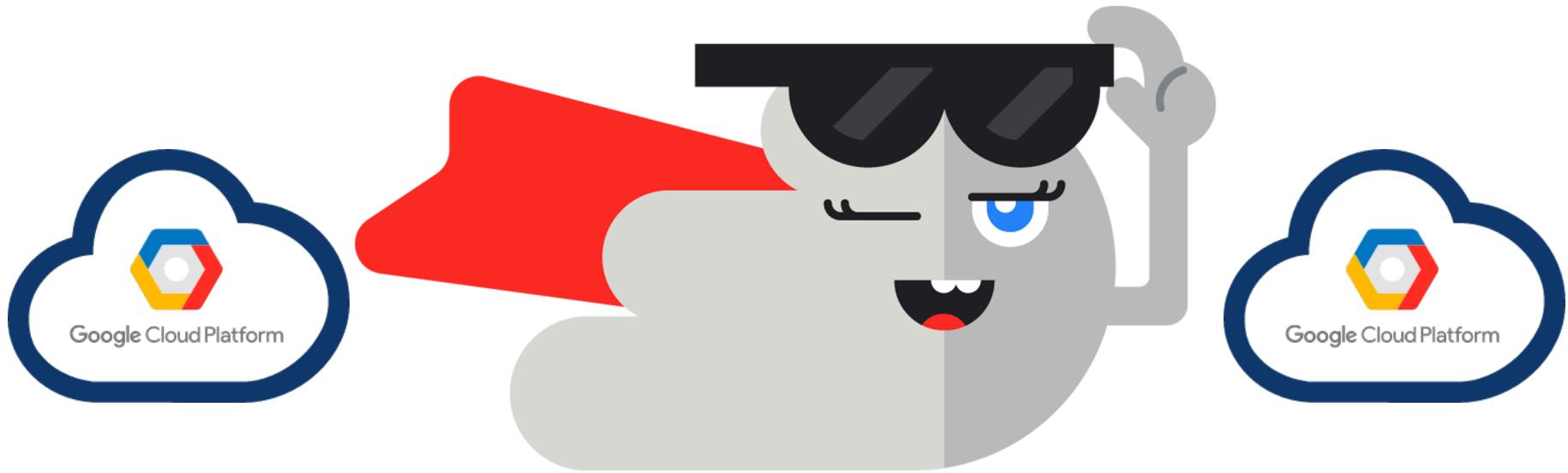




PIZZA AS A SERVICE



Acá Google Cloud nos viene a dar una ayuda.



Nicolás Lecaros

<https://www.linkedin.com/in/nicolaslecaros/>



Google
BigQuery

Nicolás Lecaros

<https://www.linkedin.com/in/nicolaslecaros/>

Comencemos hablando de **BigQuery**...

Nicolás Lecaros

<https://www.linkedin.com/in/nicolaslecaros/>

¿Qué sabemos de BigQuery?



Nicolás Lecaros

<https://www.linkedin.com/in/nicolaslecaros/>

¿Qué es BigQuery?

- Es el almacen de datos corporativos en la nube de Google.
- Escala de Petabytes.
- Facilidad de uso al usar SQL.
- Datos Cifrados, Durables y Áltamente disponibles.
- Servicio completamente gestionado.
- No hay infraestructura para administrar.
- Está diseñado para hacer que los analistas de datos sean más productivos.

Beneficios de BigQuery



Escalable

Escalabilidad horizontal real, con alto rendimiento para Petabytes de datos

En aplicaciones de producción en Google durante más de una década



Simple

Servicio gestionado: escalado automático de almacenamiento y computación

Analiza tus datos en la plataforma de Google Cloud usando SQL



Compatibile

Comparte el acceso a datos y resultados a un grupo mayor de usuarios

Conjuntos de datos públicos y comerciales para enriquecer los análisis



Seguro

Datos cifrados en movimiento y almacenamiento
Gestión de acceso granular con Google Cloud IAM



Ahorras

Servicio completamente gestionado, con un Coste de Propiedad bajo para organizaciones de todo tipo

Facturación flexible, con detalle de todos los costes por recurso



Google BigQuery

permite
redefinir los
límites de tu
Data
Warehouse



Re-imagina con BigQuery



Automatiza la entrega de datos

Automatiza la extracción de datos desde tus repositorios y consultarlos desde toda la Plataforma de Google



Democratiza la toma de decisiones basada en datos

Rompe los silos de datos y de las power apps, y genera datasets de datos de lectura para todos tus usuarios



Construye la base para la IA

Trabaja con las capacidades de la Inteligencia Artificial, con la simplicidad del lenguaje SQL



Decisiones en tiempo real

Analiza eventos de negocio en tiempo real, con la ingestión de datos automática, que permite hacer consultas inmediatas



Protege tus datos de negocio

Precio flexible y asequible
Fiabilidad con un 99,9% de SLA
Escalabilidad horizontal a demanda
Cumple con las regulaciones de compliance



Simplifica las operaciones con datos

Replicación de datos automática
Identidad y gestión de acceso (IAM) para controlar el acceso de usuarios y los permisos para ejecutar consultas

Mito



Costos en BigQuery



Controle los costos y reduzca el costo total de propiedad

Solo pague por los recursos de almacenamiento y cálculo que usa, gracias a la arquitectura sin servidor de BigQuery. La separación de BigQuery de almacenamiento y cómputo hace que sea fácil de escalar de forma independiente e interminable bajo demanda, lo que resulta en un almacenamiento económico y de bajo costo. BigQuery puede reducir el [costo total de propiedad](#) en un 56% –88%.

Costos en BigQuery

ÍT

PRECIO

Almacenamiento

\$ 0.02 por GB, por mes

\$ 0.01 por GB, por mes para [almacenamiento a largo plazo](#)

Inserciones de transmisión

\$ 0.01 por 200 MB

Cargando, copiando o exportando datos; Operaciones de metadatos

Gratis

Costos en BigQuery

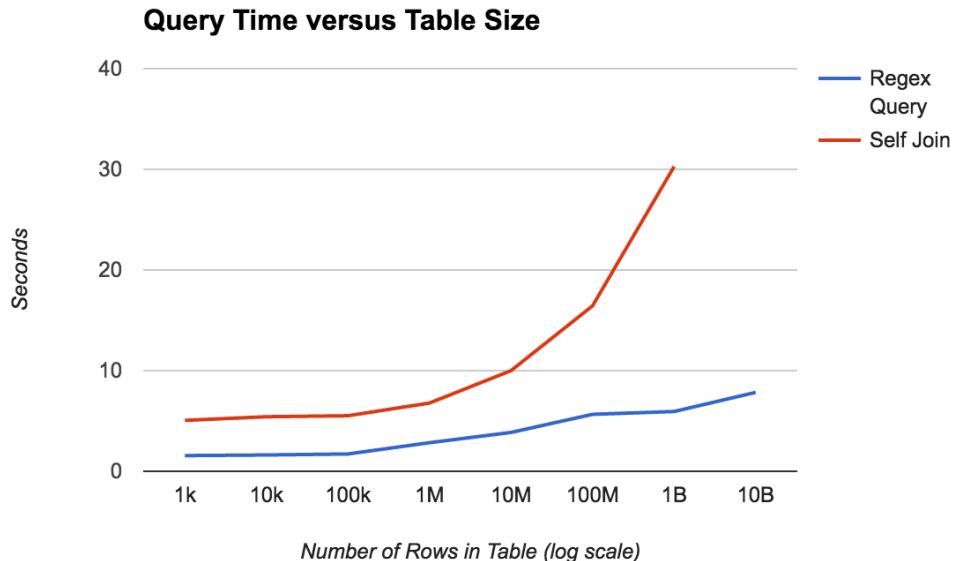
TIPO DE SUSCRIPCIÓN	PRECIO
Paga lo que consumas	\$ 5 por TB El primer terabyte (1 TB) por mes es gratis *
Precios a tanto alzado	A partir de \$ 10,000 / mes para una reserva dedicada de 500 máquinas tragamonedas. Para obtener más información, consulte precios de tarifa plana .

Realidad



Modelos Desnormalizados

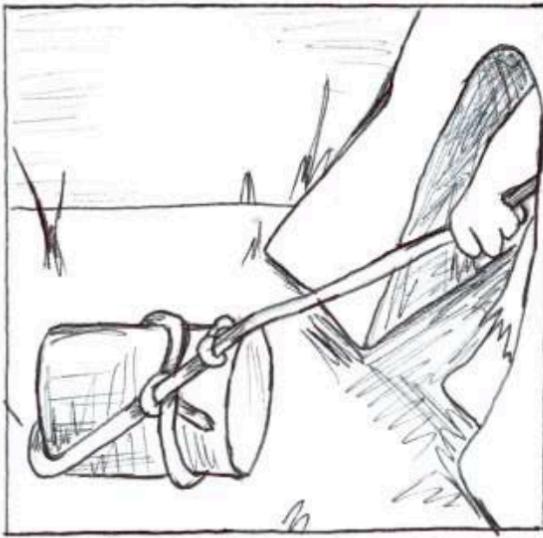
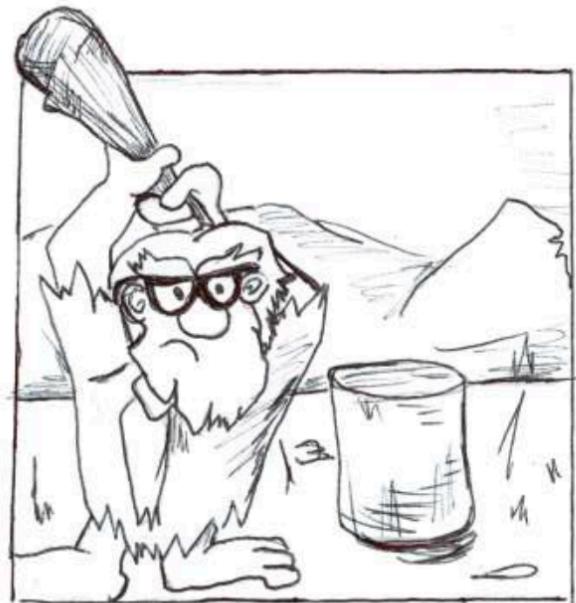
El siguiente gráfico compara el rendimiento de las consultas mediante JOIN con filtros simples en relación con el tamaño de la tabla. El rendimiento de la consulta muestra una disminución mucho más pronunciada en presencia de JOINs.



Malas prácticas en **BigQuery**...



Mala práctica: individuo canónico



“Recolectar los datos y crear nuestro dataset en la nube.”

“Arrastrar datos de vuelta al notebook.”

“Generar el modelo en nuestro Notebook.”

Buenas prácticas en **BigQuery**...



Buena práctica: trabajar como un equipo

- Trabajar en equipo significa *trabajar juntos*.
- Calcule sus datos en la nube.
- Automatice los cálculos recurrentes.
- Comparta sus datasets si es necesario incorporando políticas.



Migrar un Data Warehouses to BigQuery

- ✓ Introduction and overview
- ✓ Schema and data transfer overview
- ✓ Data governance.
- ✓ Data pipelines.
- ✓ Reporting and analysis.
- ✓ Performance optimization.

Migrar un Data Warehouses to BigQuery

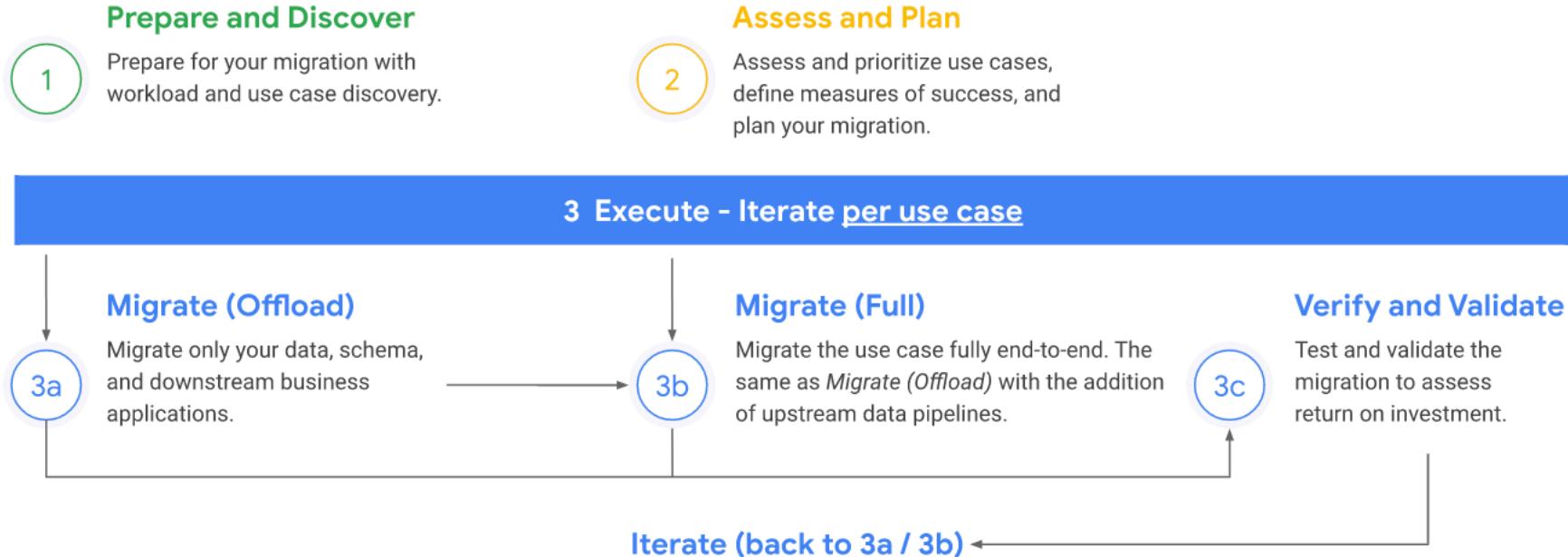
1. Prepárese y descubra : prepárese para su migración con [carga de trabajo](#) y descubra [casos de uso](#).

2. Evalúe y planifique : evalúe y priorice los casos de uso, defina medidas de éxito y planifique su migración.

3. Ejecutar : itere los siguientes pasos para cada caso de uso:

1. **Migrar (descargar)** : migre solo sus datos, esquemas y [aplicaciones comerciales posteriores](#) .
2. **Migrar (completo)** : como alternativa, migre el caso de uso completamente de extremo a extremo. Lo mismo que *Migrate (descarga)* , con la adición de los canales de datos ascendentes.
3. **Verificar y validar** : probar y validar la migración para evaluar el retorno de la inversión.

Migrar un Data Warehouses to BigQuery





Summer Meetup BDWSCL

Muchas gracias.

Nicolás Lecaros

<https://www.linkedin.com/in/nicolaslecaros/>

¡Nunca dejes de aprender!

