

Cloud Solutions, BigQuery

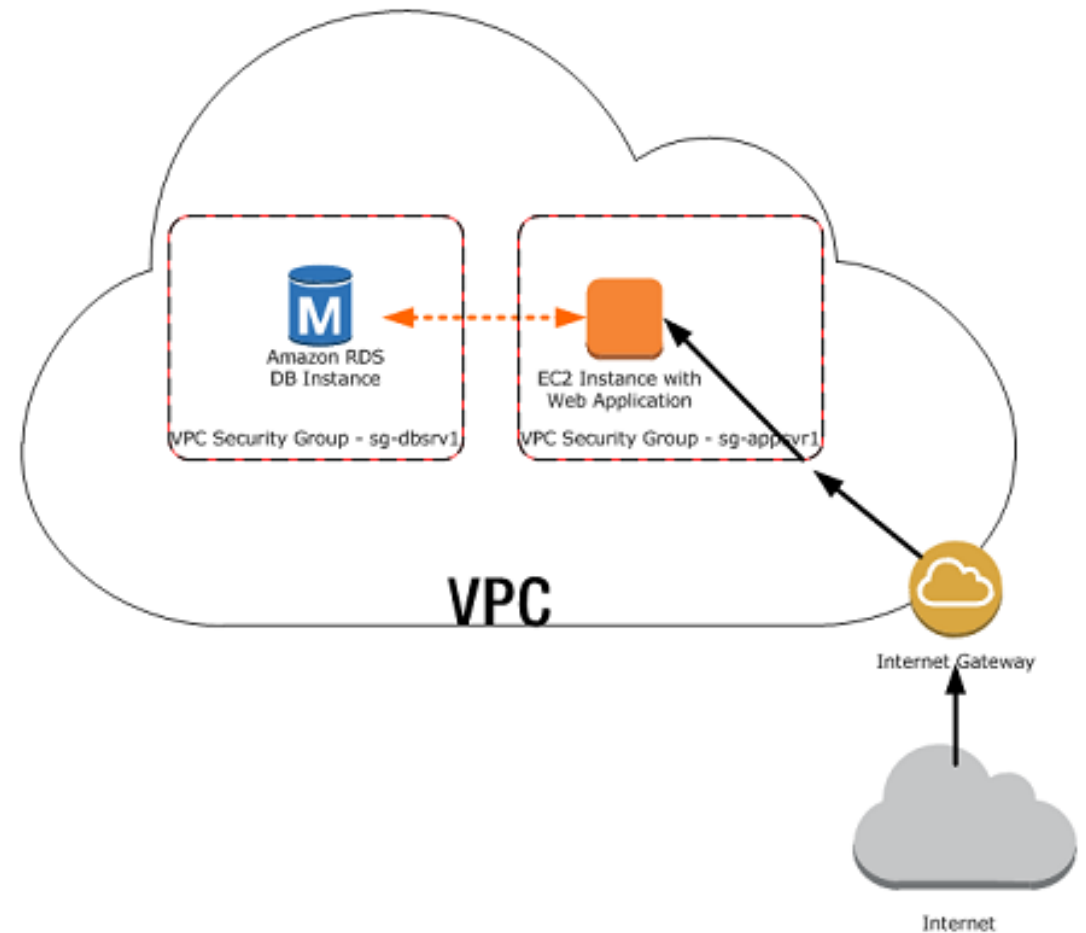
Akitaka Matsuo

Services

Database Type	AWS	GCP	Azure
Managed RDS	Amazon RDS	Cloud SQL	Azure SQL
Data Warehousing	Redshift	Bigquery	Snowflake
NoSQL (simple key-value)	DynamoDB	BigTable	Azure Tables
NoSQL (document)	MongoDB on EC2	MongoDB on GCE	DocumentDB

Connect to cloud RDS

- Setting up
 - EC2 machine for computation (installing computing environment, such as Jupyter)
 - Database like Amazon RDS
 - Placed in the same virtual private network
 - Allowing the access from the computation machine



Google Cloud Platform: Bigquery

- What is it: GCP's data warehousing
- Fully managed service
- Scalable: SQL-like syntax for datasets of *any* size
- Fast
- Affordable pricing and cost control
- API access allows integration with python

BigQuery: Structure

- A dataset is like a database
 - Tables in a dataset
- Jobs: tasks to perform on the data
 - Queries
 - Loading/exporting data

Bigquery pricing

Operation	Pricing	Details
Active logical storage	\$0.02 per GiB per month	The first 10 GiB is free each month.
Long-term logical storage	\$0.01 per GiB per month	The first 10 GiB is free each month.

Operation	Pricing	Details
Queries (on-demand)	\$6.25 per TiB	The first 1 TiB per month is free.

BigQuery Pricing

- You can use free-tier + 300 dollars credit
- How much you can do with that with 300 dollars?
 - 48 TB of queries
 - 4,800 queries for 10GB dataset