Other Google Cloud Platform Services

Pricing Calculator

- Estimating the cost of a Google Cloud solution is NOT easy
- You would need to take a number of factors into account
- How do you estimate the cost of your GCP solution?
- Use Google Cloud Pricing Calculator
- Estimates for 40+ Services:
- Compute Engine
- Google Kubernetes Engine
- Cloud Run
- App Engine
- Cloud Storage
- etc
- (REMEMBER) These are Estimates! (NOT binding on GCP)

Google Cloud Deployment Manager – Introduction

- Lets consider an example:
- I would want to create a new VPC and a subnet
- I want to provision a Load balancer, Instance groups with 5 Compute Engine instances and an Cloud SQL database in the subnet
- I would want to setup the right Firewall
- AND I would want to create 4 environments

- Dev, QA, Stage and Production!
- Deployment Manager can help you do all these with a simple (actually NOT so simple) script!

Google Cloud Deployment Manager – Advantages

- Automate deployment and modification of Google Cloud resources in a controlled, predictable way
- Deploy in multiple environments easily!
- Avoid configuration drift
- Avoid mistakes with manual configuration
- Think of it as version control for your environments
- Important Note - Always modify the resources created by Deployment Manager using Deployment Manager

Google Cloud Deployment Manager

- All configuration is defined in a simple text file - YAML
- I want a VPC, a subnet, a database and ...
- Deployment Manager understands dependencies
- Creates VPCs first, then subnets and then the database
- (Default) Automatic rollbacks on errors (Easier to retry)
- If creation of database fails, it would automatic delete the subnet and VPC

- Version control your configuration file and make changes to it over time
- Free to use - Pay only for the resources provisioned
- Get an automated estimate for your configuration

Cloud Deployment Manager – Example

```
- type: compute.v1.instance
  name: my-first-vm
  properties:
    zone: us-central1-a
    machineType: <<MACHINE_TYPE>>
    disks:
    - deviceName: boot
      type: PERSISTENT
      boot: true
      autoDelete: true
      initializeParams:
        sourceImage: <<SOURCE_IMAGE>>
    networkInterfaces:
    - network: <<NETWORK>>
      # Give instance a public IP Address
      accessConfigs:
      - name: External NAT
        type: ONE_TO_ONE_NAT
```

Cloud Deployment Manager – Terminology

- Configuration file: YAML file with resource definitions for a single deployment
- Templates: Reusable resource definitions that can be used in multiple configuration files
- Can be defined using:
- Python (preferred) OR
- JinJa2 (recommended only for very simple scripts)
- Deployment: Collection of resources that are deployed and managed together

- Manifests: Read-only object containing original deployment configuration (including imported templates)
- Generated by Deployment Manager
- Includes fully-expanded resource list
- Helpful for troubleshooting

Cloud Marketplace (Cloud Launcher)

- Installing custom software might involve setting up multiple resources:
- Example: Installing WordPress needs set up of compute engine and a relational database
- How do you simplify the set up of custom software solutions like Wordpress or even more complex things like SAP HANA suite on GCP?
- Cloud Marketplace: Central repo of easily deployable apps & datasets
- Similar to App Store/Play Store for mobile applications
- You can search and install a complete stack
- Commercial solutions - SAP HANA etc
- Open Source Packages - LAMP, WordPress, Cassandra, Jenkins etc
- OS Licenses: BYOL, Free, Paid
- Categories: Datasets/Developer tools/OS etc
- When selecting a solution, you can see:
- Components - Software, infrastructure needed etc

- Approximate price

Cloud DNS

- What would be the steps in setting up a website with a domain name (for example, in28minutes.com)?
- Step I : Buy the domain name in28minutes.com (Domain Registrar)
- Step II : Setup your website content (Website Hosting)
- Step III : Route requests to in28minutes.com to the my website host server (DNS)
- Cloud DNS = Global Domain Name System (Step III)
- Setup your DNS routing for your website (in28minutes.com)
- Route api.in28minutes.com to the IP address of api server
- Route static.in28minutes.com to the IP address of http
- server Route email (ranga@in28minutes.com) to the mail server(mail.in28minutes.com)
- Public and private managed DNS zones (container for records)

Cloud DNS – CLI

- gcloud dns managed-zones create ZONE_NAME -- description (REQUIRED - Short description for the managed-zone)

- --dns-name (REQUIRED - DNS name suffix that will be managed with the created zone)
- --visibility (private/public)
- --networks (List of networks that the zone should be visible in if the zone visibility is [private])
- Three Steps to add records to a managed zone:
- Start Transaction for Zone
- gcloud dns record-sets transaction start --zone
- Make Changes
- gcloud dns record-sets transaction add --name=REC_NAME --ttl --type A/CNAME --zone=ZONE_NAME
- End Transaction for Zone
- gcloud dns record-sets transaction execute –zone

Cloud Dataflow

- Cloud Dataflow is a difficult service to describe:
- Let's look at a few example pipelines you can build:
- Pub/Sub > Dataflow > BigQuery (Streaming)
- Pub/Sub > Dataflow > Cloud Storage (Streaming - files)
- Cloud Storage > Dataflow > Bigtable/CloudSpanner/Datastore/BigQuery (Batch - Load data into databases)
- Bulk compress files in Cloud Storage (Batch)
- Convert file formats between Avro, Parquet & csv (Batch)

- Streaming and Batch Usecases
- Realtime Fraud Detection, Sensor Data Processing, Log Data Processing, Batch Processing (Load data, convert formats etc)
- Use pre-built templates
- Based on Apache Beam (supports Java, Python, Go ...)
- Serverless (and Autoscaling)

Cloud Dataproc

- Managed Spark and Hadoop service:
- Variety of jobs are supported:
- Spark, PySpark, SparkR, Hive, SparkSQL, Pig, Hadoop
- Perform complex batch processing
- Multiple Cluster Modes:
- Single Node / Standard/ High Availability (3 masters)
- Use regular/preemptible VMs
- Use case: Move your Hadoop and Spark clusters to the cloud
- Perform your machine learning and AI development using open source frameworks
- (REMEMBER) Cloud Dataproc is a data analysis platform
- You can export cluster configuration but NOT data
- (ALTERNATIVE) BigQuery - When you run SQL queries on Petabytes

- Go for Cloud Dataproc when you need more than queries (Example: Complex batch processing Machine Learning and AI workloads)