Introduction to Google Cloud Platform

Memi Lavi
www.memilavi.com

- Google's public cloud
- Announced in April 2008
- Released in November 2011
- The 3rd largest public cloud
- Often called GCP (we'll do that too…)

- First focused on PaaS services (App Engine)
    - To counter AWS's IaaS focus
- Later added IaaS
- Currently offers a large variety of cloud services

- Major clients:

TIME

Goldman Sachs

Regions

- Google built a lot of datacenters for GCP
- Datacenters are grouped in geographic locations
- Each location is called Region
- There are ~40 GCP Regions
- Almost every new resource in the cloud should be allocated to a region

Zones

- Regions have more than one physical datacenter
- Usually three
- Great for availability in case a datacenter fails
- Each datacenter is called Zone
- Some cloud services take advantage of multiple zones in a region

Source: https://cloud.google.com/about/locations#lightbox-regions-map/

Region Selection Consideration

- When selecting a region for your cloud services consider:

Service availability	Not all services are available in all regions
Proximity	Select the region closest to your audience
Price	Services pricing varies between regions

Low Emission Regions

- Google strives to become a carbon-free cloud
- Minimum CO_2 emission
- Depends on the electricity in a specific region
- Regions with less carbon emissions are marked with 

- You might want to consider that too when selecting the region

Zones Naming

- With most services you'll need to select the region and zone for deployment
- Regions and zones have consistent names:

  `geography-direction-zone`

- Note that region name does not include country name (except in US)

Zones Naming

- Examples:

Region + Zone	Location
`europe-west2-a`	Turin, Italy
`europe-west4-b`	Eemshaven, Netherlands
`us-central1-a`	Council Bluffs, Iowa, US
`us-east5-b`	Columbus, Ohio, US
`asia-east1-a`	Changhua County, Taiwan

See the full list of regions and zones here:
https://cloud.google.com/compute/docs/regions-zones#available

Global, Regional and Zonal Services

- Cloud services are placed in a region, zone or globally
- Determine their level of isolation and availability
- For example:
  - Regional services can failover between zones in the same region
  - Zonal services depend on their zone, and can access only services in the same zone

Global, Regional and Zonal Services

Global

- VPC – Deployed globally and can access resources in all regions
- Images – Can be used by any image anywhere

Regional

- App Engine – Deployed in a region and replicated across zones
- Managed Instance Group – Deployed in a region and balances load across zones

Zonal

- Instance (VM)– Deployed in a zone and depends on it
- Persistent Disk– Deployed in a zone

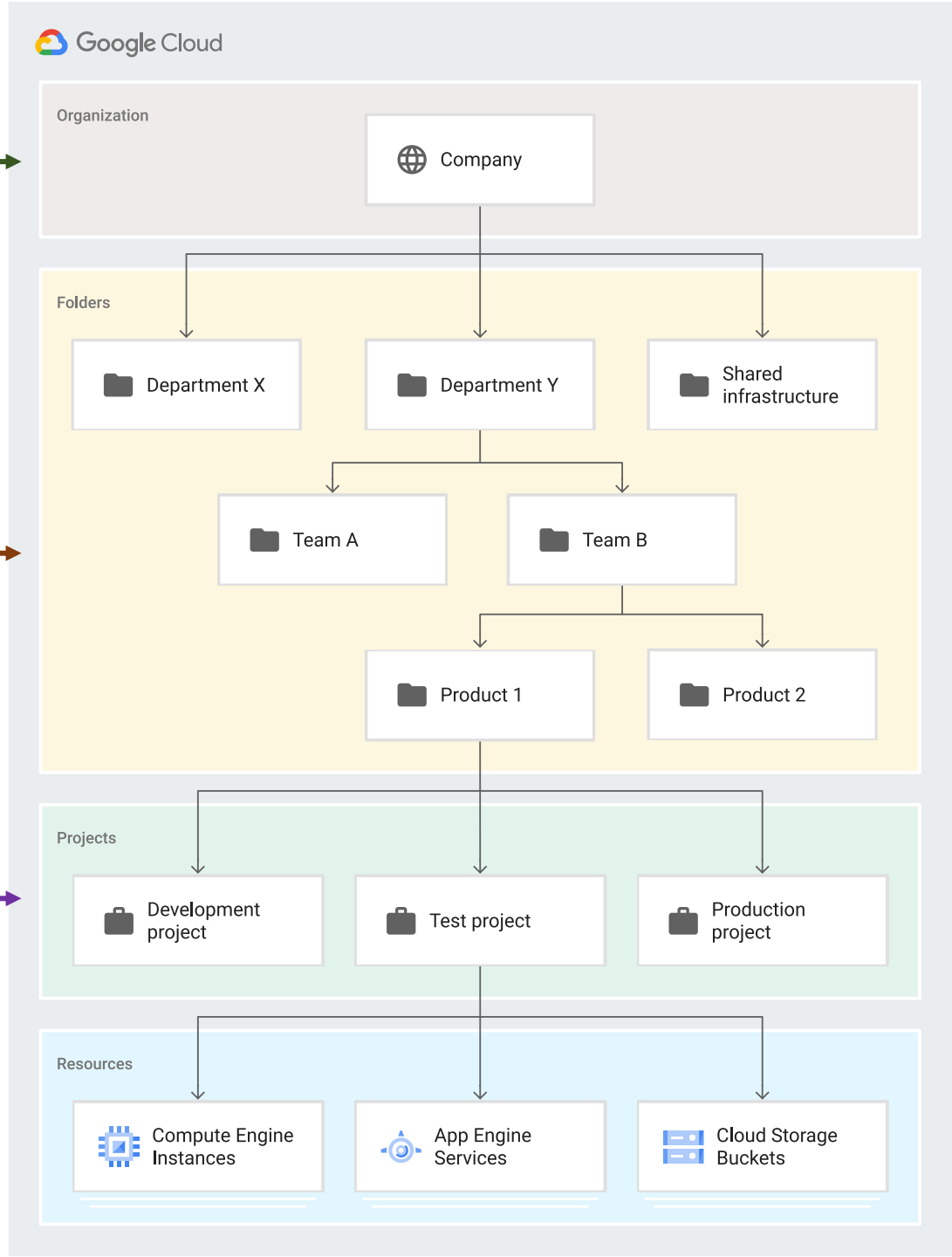
Resources Hierarchy

- Resources in GCP are organized in hierarchy
- Each resource is grouped in its parent resource
  - In most cases can have only a single parent
- Resources inherit policies of their parent

- Top level of the hierarchy
- Usually represents the company using the cloud
- Have access to all underlying resources

- Additional grouping mechanism
- Usually model legal entities, departments, teams etc.
- Can contain other folders
- Optional

- Contain the actual resources
- The most important level in the hierarchy
- Resources must be created in a project



Hierarchy and Free Tier

- We're going to work with the free tier of GCP
- No Organization and Folders
- Some features are unavailable
  - ie. Groups

GCP Services

- Everything that is deployed in the cloud is called

Cloud Service

- ie. VMs, databases, networks, AI, user management etc.

GCP Services

- View full list of cloud services in GCP:

  https://cloud.google.com/products

Using the Free Tier

- We'll use the free tier of GCP
- Grants 300$ credit for 90 days
- Some limitations
  - No Windows VMs, no GPU, and more…
- Some services have their own free tier
  - Cloud Storage, VMs, App Engine and more