Summary and Review

GCP Fundamentals: Core Infrastructure

Last modified 2018-08-12

Agenda

Course review

Next steps

Google Cloud

Comparing compute options

	Compute Engine	Kubernetes Engine	App Engine Flex	App Engine Standard	Cloud Functions
Service model	IaaS	Hybrid	PaaS	PaaS	Serverless
Use cases	General computing workloads	Container-based workloads	Web and mobile applications; container-based workloads	Web and mobile applications	Ephemeral functions responding to events

Toward managed infrastructure *Toward dynamic infrastructure*

Google Cloud

Remember the continuum that this course discussed at the very beginning: the continuum between managed infrastructure and dynamic infrastructure. GCP's compute services are arranged along this continuum, and you can choose where you want to be on it.

Choose Compute Engine if you want to deploy your application in virtual machines that run on Google's infrastructure. Choose Kubernetes Engine if you want instead to deploy your application in containers that run on Google's infrastructure, in a Kubernetes cluster you define and control.

Choose App Engine instead if you want to just focus on your code, leaving most infrastructure and provisioning to Google. App Engine Flexible Environment lets you use any runtime you want, and gives you full control of the environment in which your application run; App Engine Standard Environment lets you choose from a set of standard runtimes and offers finer-grained scaling and scale-to-zero.

To completely relieve yourself from the chore of managing infrastructure, build or extend your application using Cloud Functions. You supply chunks of code for business logic, and your code gets spun up on-demand in response to events.

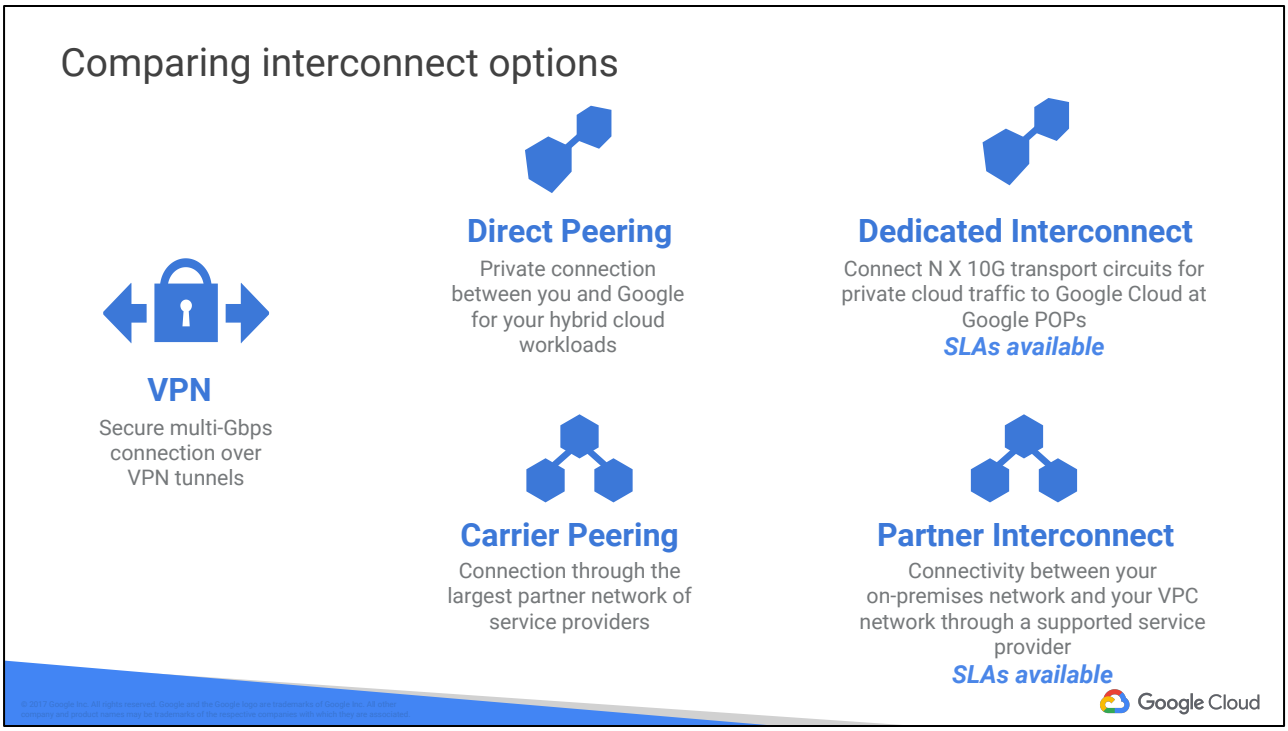
Comparing load-balancing options

Global HTTP(S)	Global SSL Proxy	Global TCP Proxy	Regional	Regional internal
Layer 7 load balancing based on load	Layer 4 load balancing of non-HTTPS SSL traffic based on load	Layer 4 load balancing of non-SSL TCP traffic	Load balancing of any traffic (TCP, UDP)	Load balancing of traffic inside a VPC
Can route different URLs to different back ends	Supported on specific port numbers	Supported on specific port numbers	Supported on any port number	Use for the internal tiers of multi-tier applications

Google Cloud

GCP offers a variety of ways to load-balance inbound traffic. Use Global HTTP(S) Load Balancing to put your web application behind a single anycast IP to the entire Internet; it load-balances traffic among all your backend instances in regions around the world, and it's integrated with GCP's Content Delivery Network.

If your traffic isn't HTTP or HTTPS, you can use the Global TCP or SSL Proxy for traffic on many ports. For other ports or for UDP traffic, use the Regional Load Balancer. Finally, to load-balance the internal tiers of a multi-tier application, use the Internal Load Balancer.



GCP also offers a variety of ways for you to interconnect your on-premises or other-cloud networks with your Google VPC. It's simple to set up a VPN, and you can use Cloud Router to make it dynamic. You can peer with Google at its many worldwide points of presence, either directly or through a carrier partner. Or, if you need a Service Level Agreement and can adopt one of the required network topologies, use Dedicated Interconnect. A Partner Interconnect connection is useful if your data center is in a physical location that can't reach a Dedicated Interconnect colocation facility or if your data needs don't warrant an entire 10 Gbps connection.

Comparing storage options

	Cloud Datastore	Cloud Bigtable	Cloud Storage	Cloud SQL	Cloud Spanner	BigQuery
Type	NoSQL document	NoSQL wide column	Blobstore	Relational SQL for OLTP	Relational SQL for OLTP	Relational SQL for OLAP
Best for	Getting started, App Engine applications	"Flat" data, Heavy read/write, events, analytical data	Structured and unstructured binary or object data	Web frameworks, existing applications	Large-scale database applications (> ~2 TB)	Interactive querying, offline analytics
Use cases	Getting started, App Engine applications	AdTech, Financial and IoT data	Images, large media files, backups	User credentials, customer orders	Whenever high I/O, global consistency is needed	Data warehousing

Google Cloud

Consider using Cloud Datastore if you need to store structured objects, or if you require support for transactions and SQL-like queries.

Consider using Cloud Bigtable if you need to store a large amount of single-keyed data, especially structured objects.

Consider using Cloud Storage if you need to store immutable binary objects.

Consider using Cloud SQL or Cloud Spanner if you need full SQL support for an online transaction processing system. Cloud SQL provides terabytes of capacity, while Cloud Spanner provides petabytes and horizontal scalability.

Consider BigQuery if you need interactive querying in an online analytical processing system with petabytes of scale.

Choosing among Google Cloud Storage classes

	Multi-regional	Regional	Nearline	Coldline
Intended for data that is...	Most frequently accessed	Accessed frequently within a region	Accessed less than once a month	Accessed less than once a year
Availability SLA	99.95%	99.90%	99.00%	99.00%
Access APIs	***Consistent APIs***			
Access time	***Millisecond access***			
Storage price	Price per GB stored per month			
Retrieval price	Total price per GB transferred			
Use cases	Content storage and delivery	In-region analytics, transcoding	Long-tail content, backups	Archiving, disaster recovery

Google Cloud

I'd like to zoom into one of those services we just discussed, Cloud Storage, and remind you of its four storage classes. Multi-regional and regional are the classes for warm and hot data. Use multi-regional especially for content that's served to a global web audience, and use regional for working storage for compute operations. Nearline and coldline are the classes for, as you'd guess, cooler data. Use nearline for backups and for infrequently accessed content, and use coldline for archiving and disaster recovery.

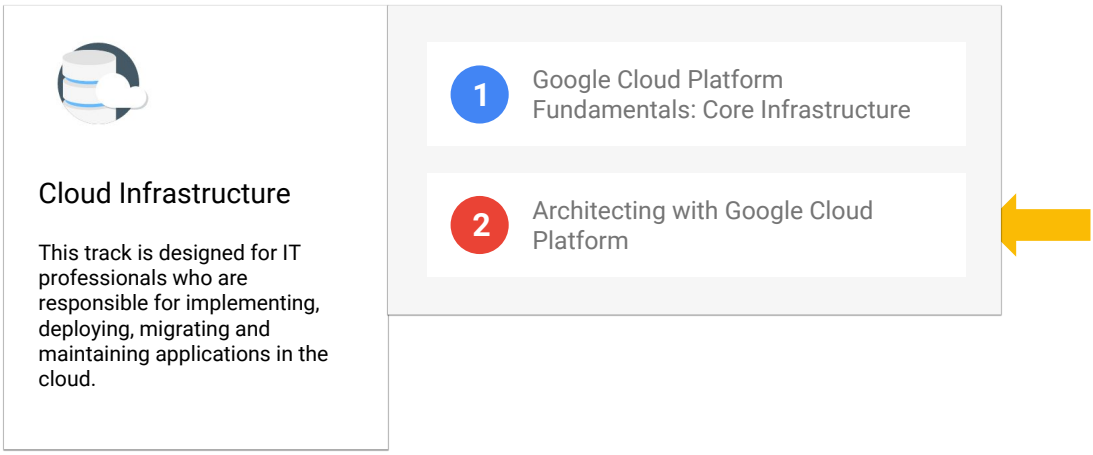
Agenda

Course review

Next steps

Google Cloud

What's next in the Cloud Infrastructure track?



Google Cloud

If you're a cloud architect, a DevOps person, or any other kind of IT professionals who deploys, migrates and maintains applications in the cloud, continue with the Coursera specialization: Architecting with Google Cloud Platform.

What's next in the Application Development track?

Application Development

This track is designed for application programmers and software engineers who develop software programs in the cloud.

1. Google Cloud Platform Fundamentals: Core Infrastructure
2. Developing Applications with Google Cloud Platform

Google Cloud

If you're an application programmer or any other kind of software engineer who writes code for the cloud, continue to the Coursera specialization: Developing Applications with Google Cloud Platform

If you're an application programmer or any other kind of software engineer who writes code for the cloud, continue to the Coursera specialization: Developing Applications with Google Cloud Platform