Infinitely Scale Graph Analysis with TigerGraph Cloud

Speaker

Vikram Vunduru
Startup Solutions Architect, AWS

Shuo Yang
TigerGraph -- Head of Cloud

Agenda

- AWS Global Infrastructure
- Architect your SaaS Applications on AWS
- Infinitely scale graph analysis with TigerGraph - quick architecture overview
- TigerGraph on AWS - from database to service
- Demo - Launch an instance to get started

Global Infrastructure

Why customers choose AWS

Most experience

15

years helping millions
of customers

Global reach & high availability

81

Availability zones spanning 25
geographic regions
Private network backbone

Security & compliance

230+

security features

Platform breadth / depth & innovation

200+
Service Offering

2,757
New Features & Services in 2020

Ecosystem

10,000

software listings from 1,600 ISVs

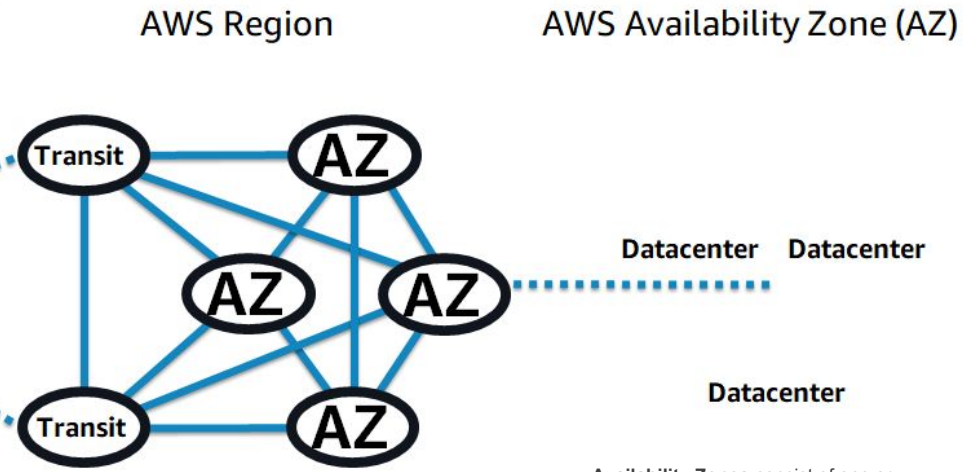
Improve TCO

107

price reductions since 2006
granular billing

Scale globally with resilience in every region

The largest global foot print consistently built with a multi-AZ and multi-datacenter design



Region & Number of Availability Zones

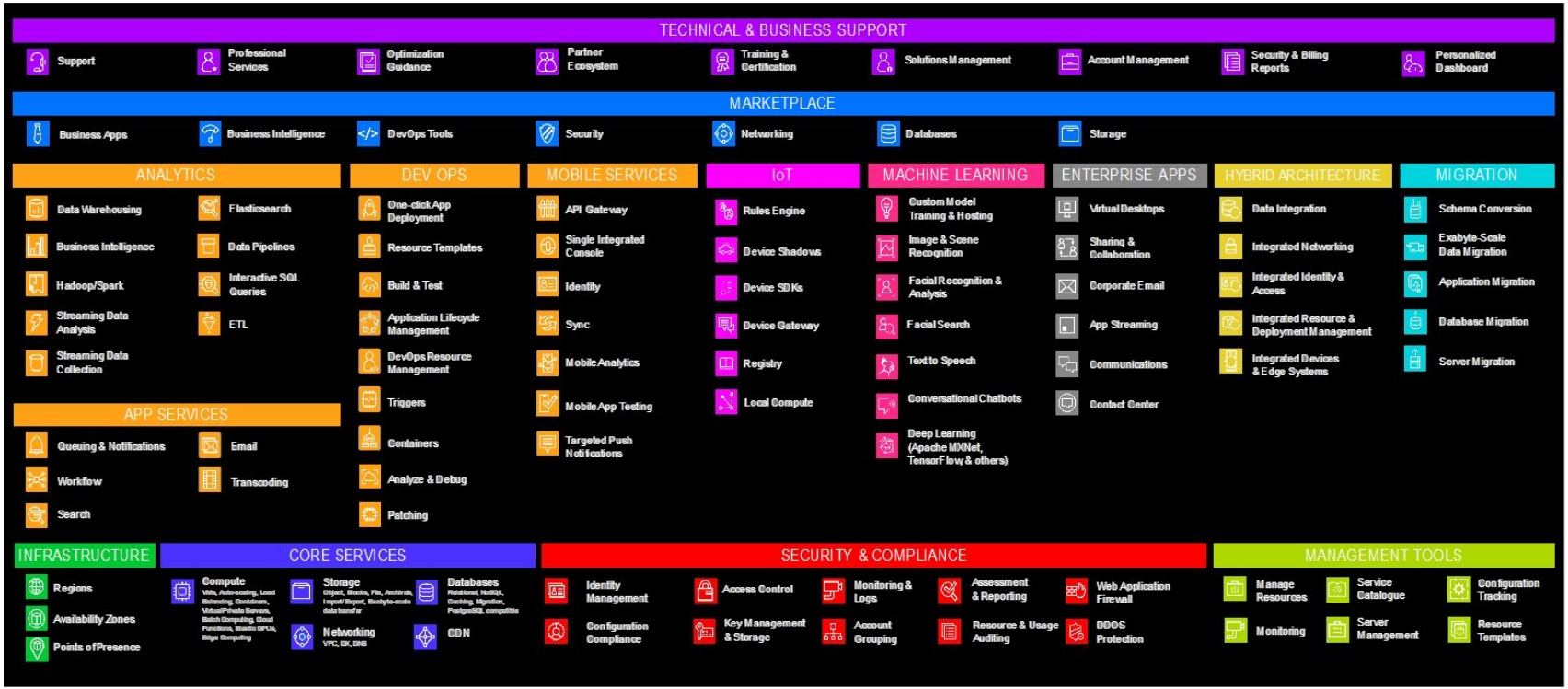
Announced Regions
Cape Town, Jakarta, Milan, Spain and Osaka

A Region is a physical location in the world where we have multiple **Availability Zones.**

Availability Zones consist of one or more discrete data centers, each with redundant power, networking, and connectivity, housed in separate facilities.

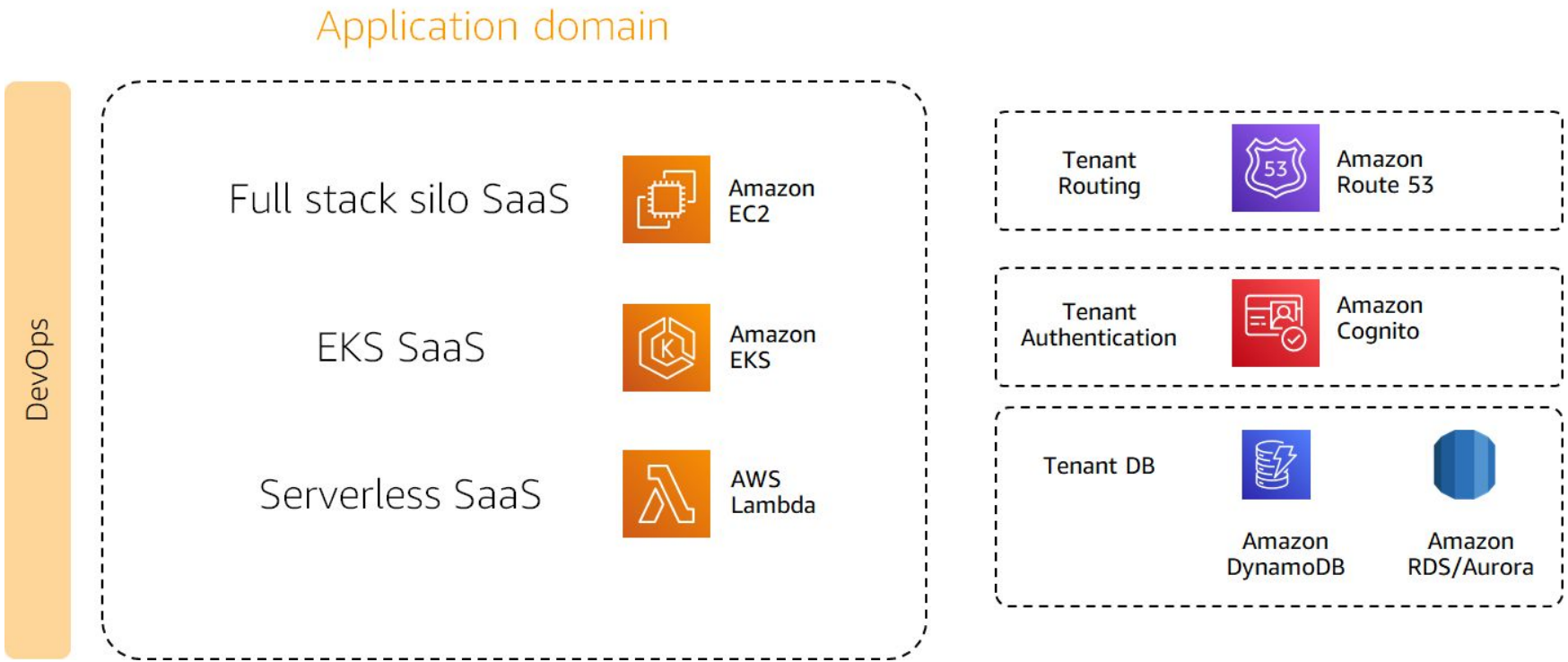
24 Regions
76 Availability Zones
216 Points of Presence

Broad and Deep Functionality



Architect your SaaS Applications on AWS

Architect your SaaS Applications on AWS



AWS PrivateLink

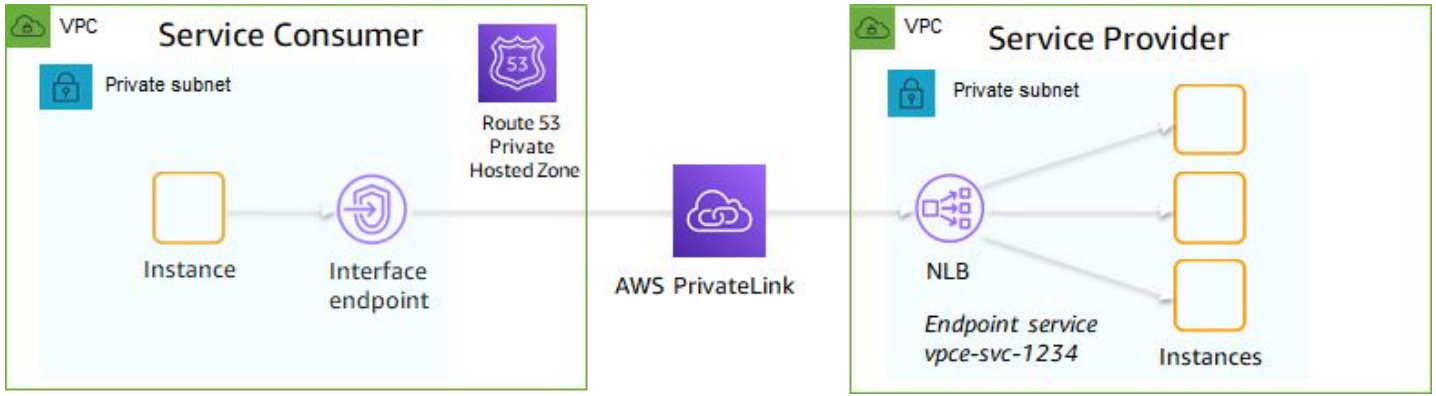
Establish private connectivity without exposing data to the Internet

Interface endpoint (in consumer VPC)

- Entry point for traffic to a PrivateLink-powered service. One or more ENIs1 created by AWS that uses private IP
- Associate a security group with the ENI to control access
- Apps use the endpoint-specific DNS host name or default DNS name2 (for AWS and AWS Marketplace Partner services)

Endpoint service (in provider VPC)

- Only needed if you are offering a PrivateLink-powered service to other consumers
- Network Load Balancer used as service front-end
- Create a VPC Endpoint Service configuration and specify your NLB



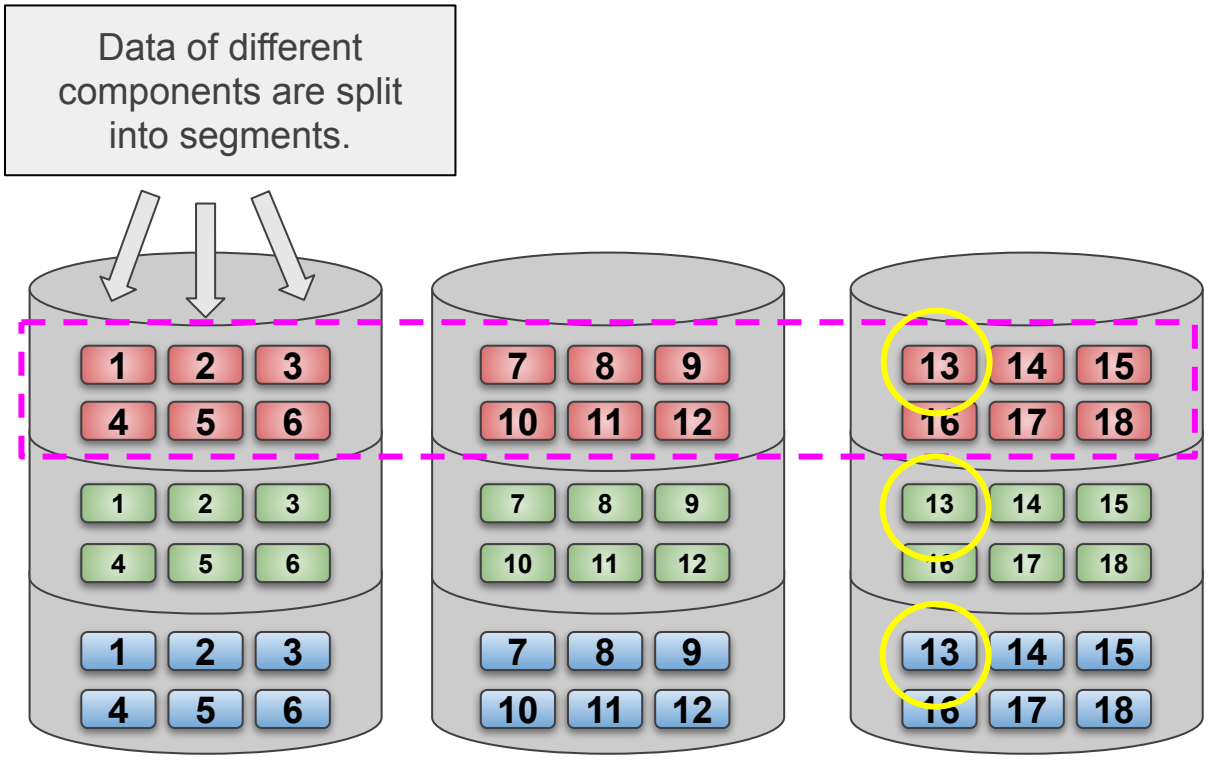
[1] ENI: Elastic Network Interface

[2] Requires Private DNS to be enabled

Infinitely scale graph analysis with Tigergraph - quick architecture overview

Partition: the scale-out graph storage engine

Data of different components are split into segments.



IDS

VERTEX

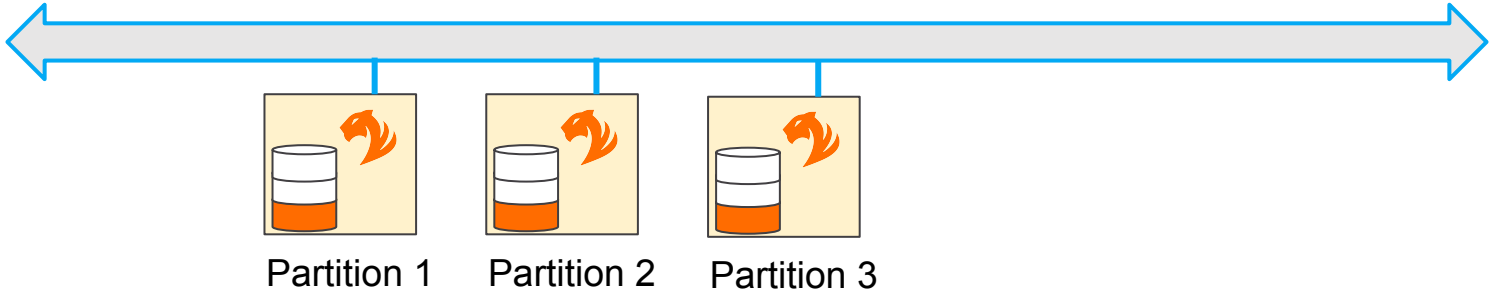
EDGE

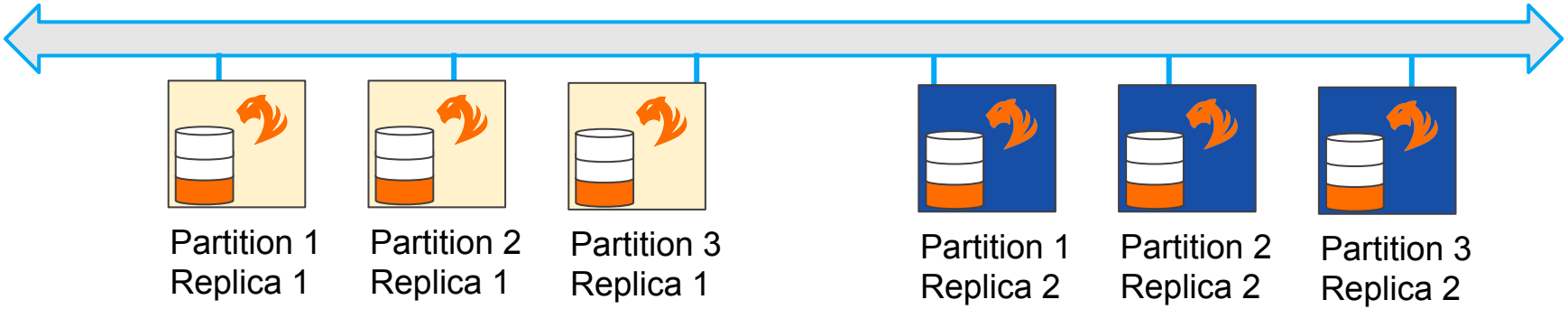
The segments are stored distributedly across the cluster.

The segments of different components with same ID stores data for the same set of vertices under the same vertex type.

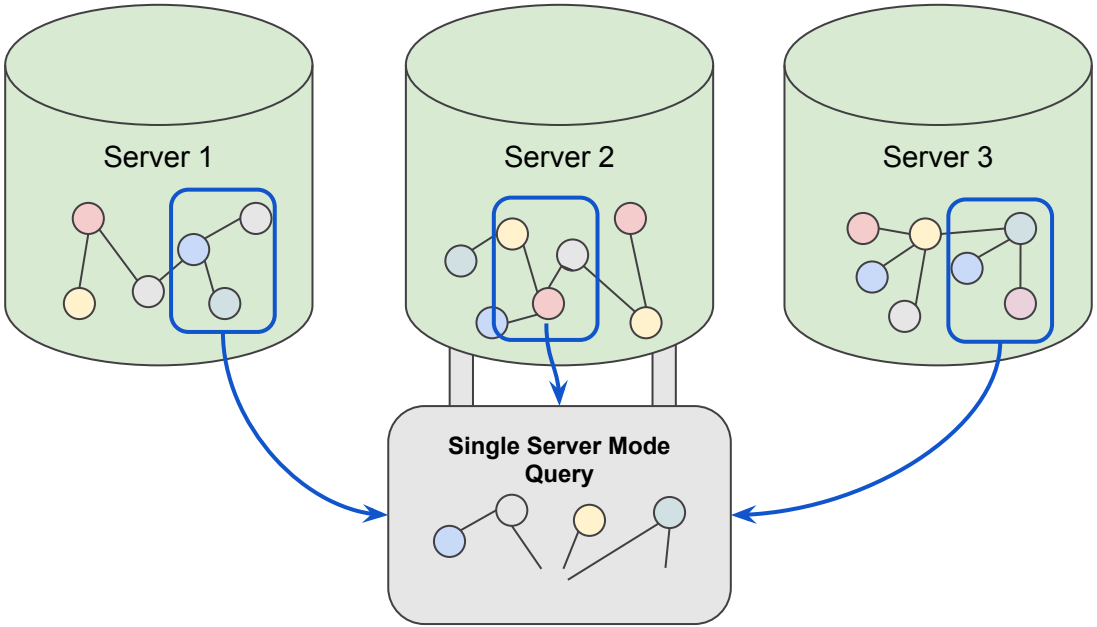
The location of a vertex can be calculated based on its internal ID

Replication: the highly available architecture





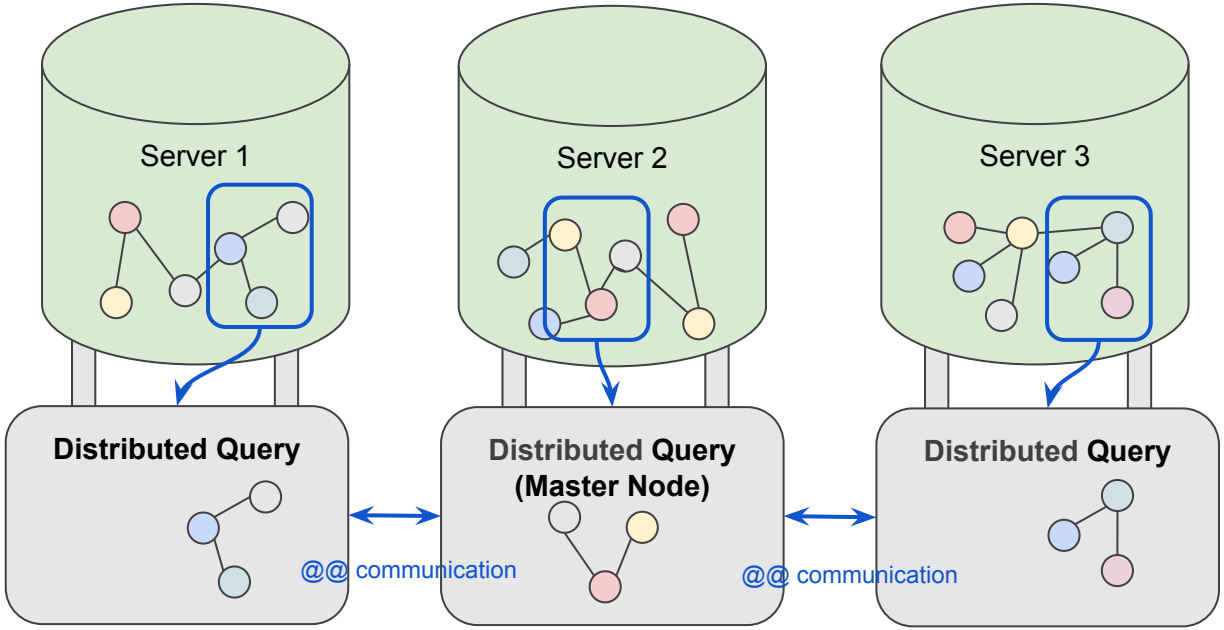
Query - Single Server mode



Single Server Mode

- The cluster elects one server to be master for that query.
- All query computation takes place on query master.
- Vertex and edge data are copied to the query master as needed.
- **Best for queries with one or a few starting vertices.**
- **If your query starts from all vertices, don't use this mode.**

Query - Distributed mode

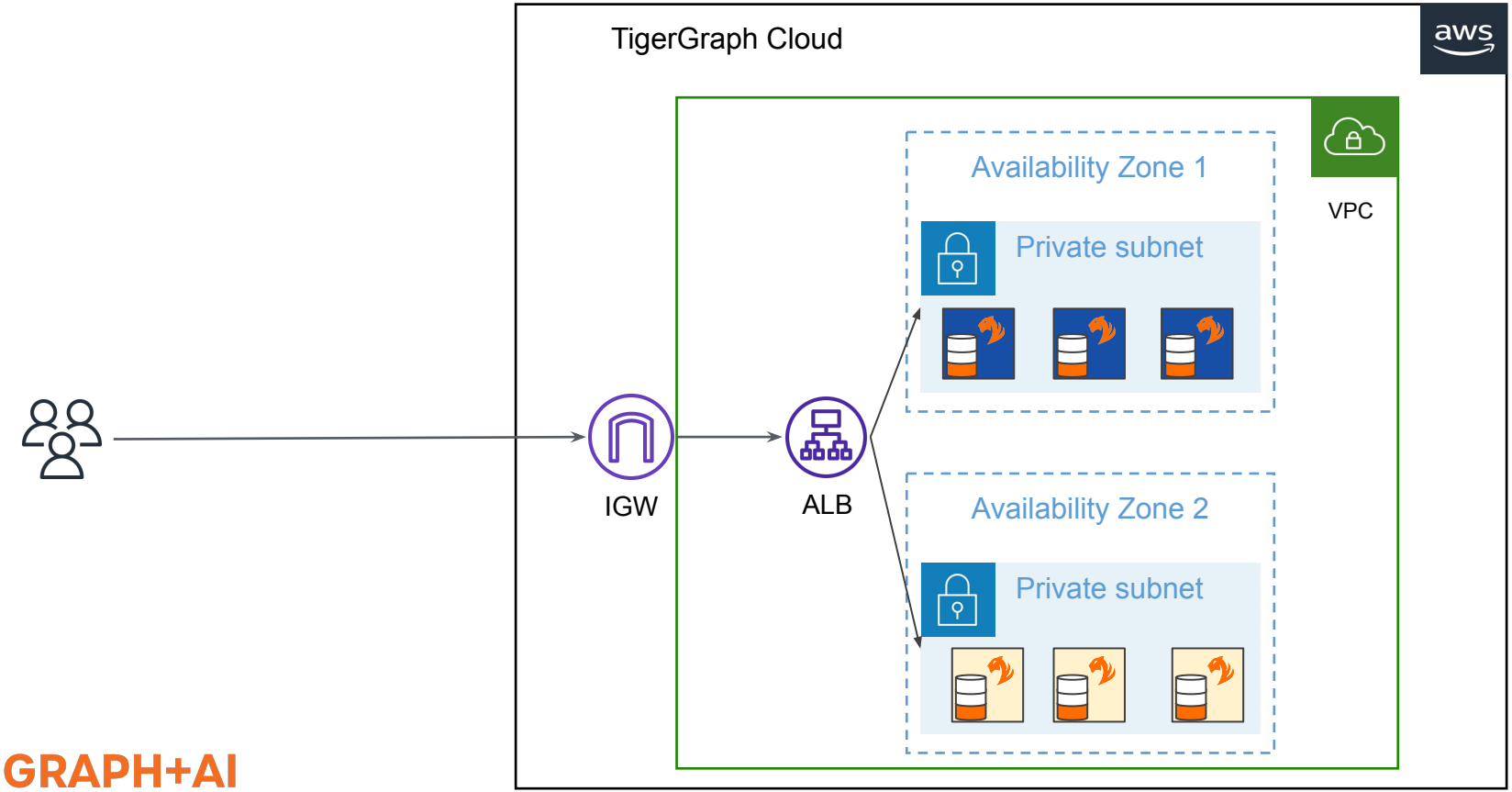


Distributed Mode

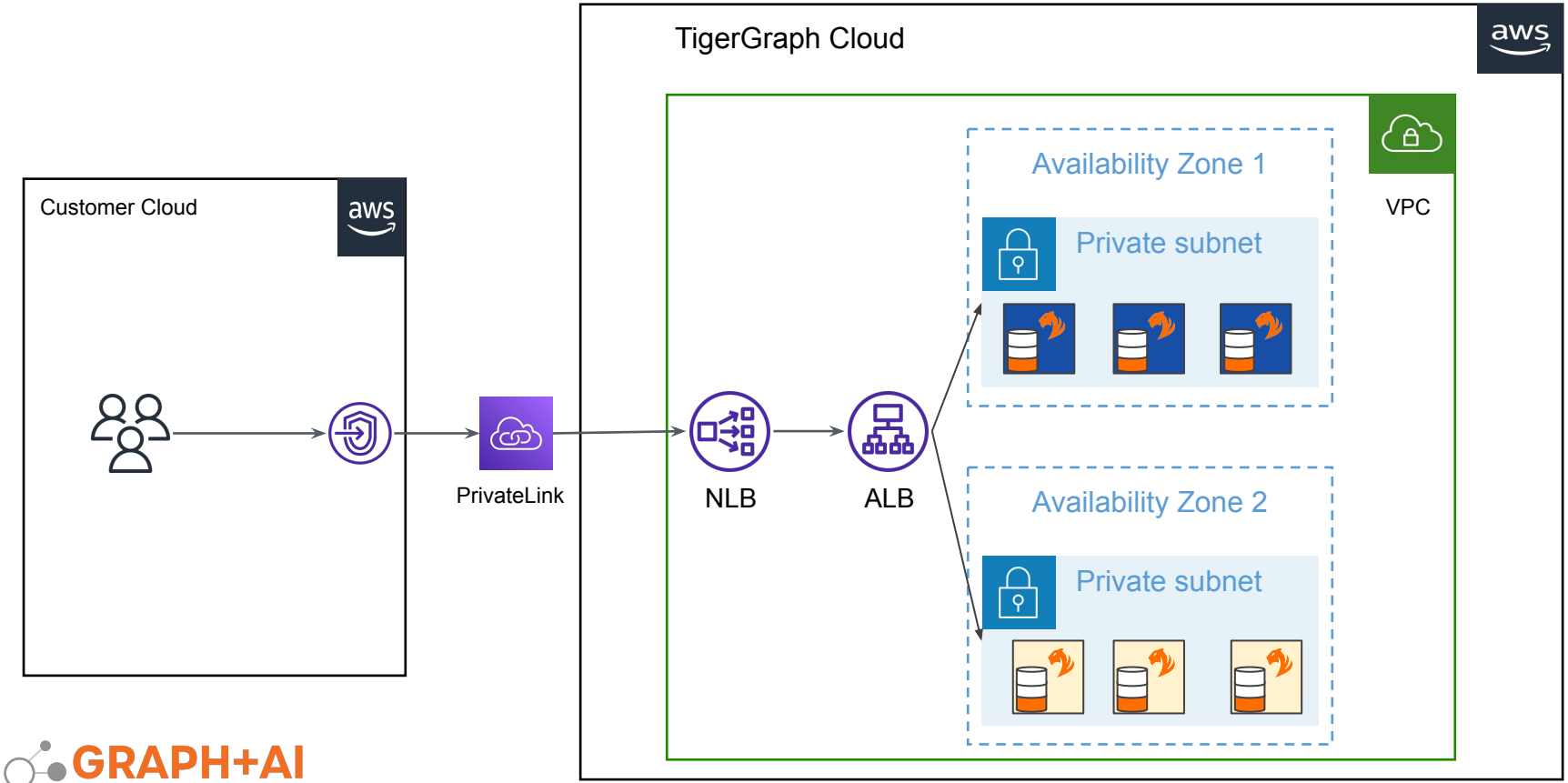
- The server that received the query becomes the master.
- Computation executes on **all** servers in parallel.
- Accumulators are transferred across the cluster.
- **If your query starts from all or most vertices, use this mode.**

Tigergraph on AWS - from database to service

Managed Service on Cloud



End to end private link



Live Demo and Q&A

Please visit and try out cloud service

www.tgcloud.io