Designing cloud workloads for LEO satellite internet users

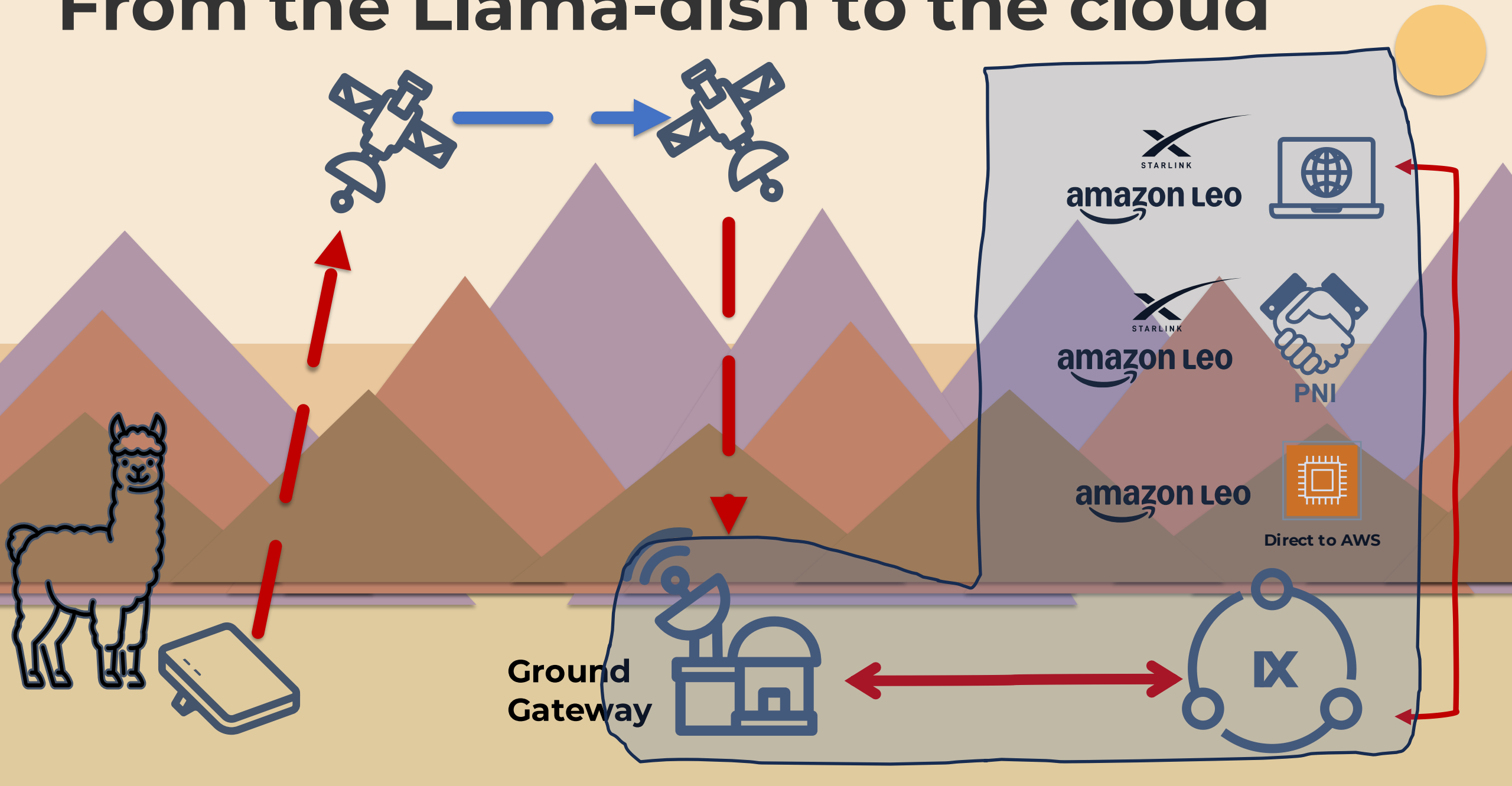
Sid Mathur, Fastah Inc.

NANOG 97

Today…

- The big picture – how does Low Earth Orbit (LEO) satellite work, for both Starlink and Amazon LEO
- Static user v/s In-motion users – aviation and maritime
- Hyperscale cloud deployments for LEO use cases
- Application (network) design patterns

From the Llama-dish to the cloud



Ground Gateways

- Satellite communication
- Carry data to/from the nearest Internet Exchange (**IX**)/ Point-of-Presence (**PoP**)
- Starlink: ~200
- Amazon LEO : 300+ (planned)

IX points or PoP – Starlink

Data: https://geoip.starlinkisp.net/pops.csv

NANOG™

IX points or PoP – Amazon LEO

amazon LEO

Data: aws directconnect --region=us-west-2 describe-locations

AWS Direct Connect
(peer here with AWS)

NANOG

Interactive Maps - IX / PoPs

Starlink

Amazon LEO /
AWS Direct Connect

Attribution: https://github.com/chriselsen/dx-location-details

Part II – Static Users vs. In-motion users

The unique struggles (privileges!) of airborne and marine internet

Static users – the many persona(s)

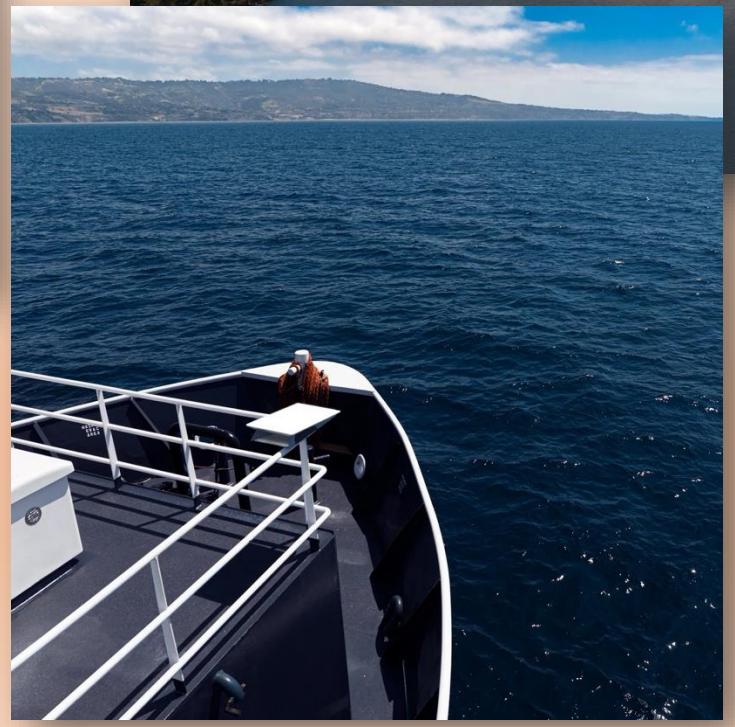
For an application :

- IP geolocation works
- Stable internet egress PoP
- Predictable RTT, jitter
- Just like cable / fiber

NANOG™

In-motion users – more personas

Travel fast and/or across international boundaries



For an application :

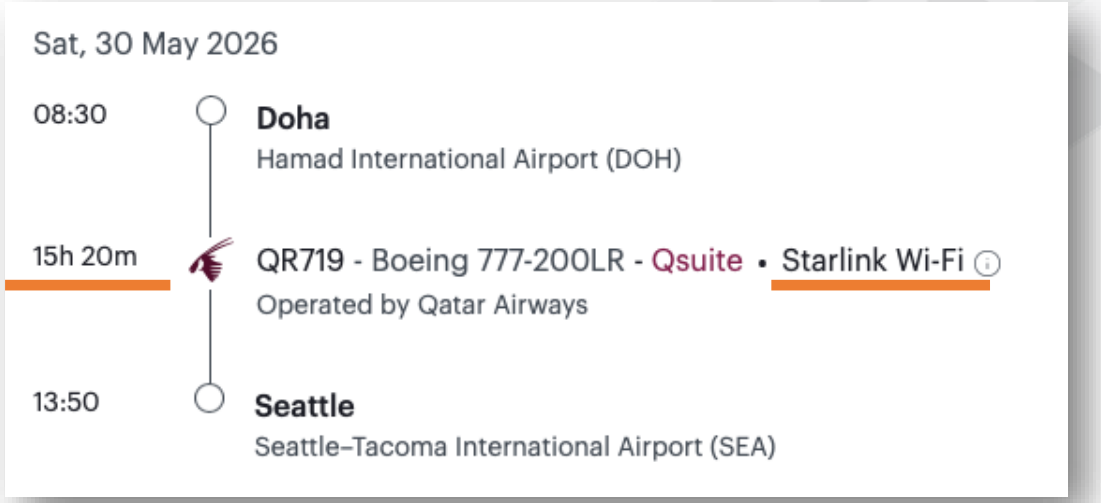
- internet egress changes
- higher jitter, TCP resets
- IP geolocation may change

NANOG™


IP Geo #7
France
IP Geo #6
Germany
IP Geo #5
Hungary
IP Geo #4
Bulgaria
IP Geo #3
Jordan
IP Geo #2
Kuwait
IP Geo #1
Qatar
01:35
Doha
Hamad International Airport (DOH)
7h 10m
QR41 - Boeing 777-300ER - Qsuite • Starlink Wi-Fi
Operated by Qatar Airways
07:45
Paris
Charles de Gaulle International Airport (CDG)

In-motion : Limitations of IP-based geo

- Should Starlink (or LEO) in a moving aircraft or ship change its advertised country of IP geolocation?
- Should your app dynamically localize the language?
- Which media or banking jurisdiction?
- EU laws and GDPR?



Part III
Hyperscaler Clouds for apps and workloads

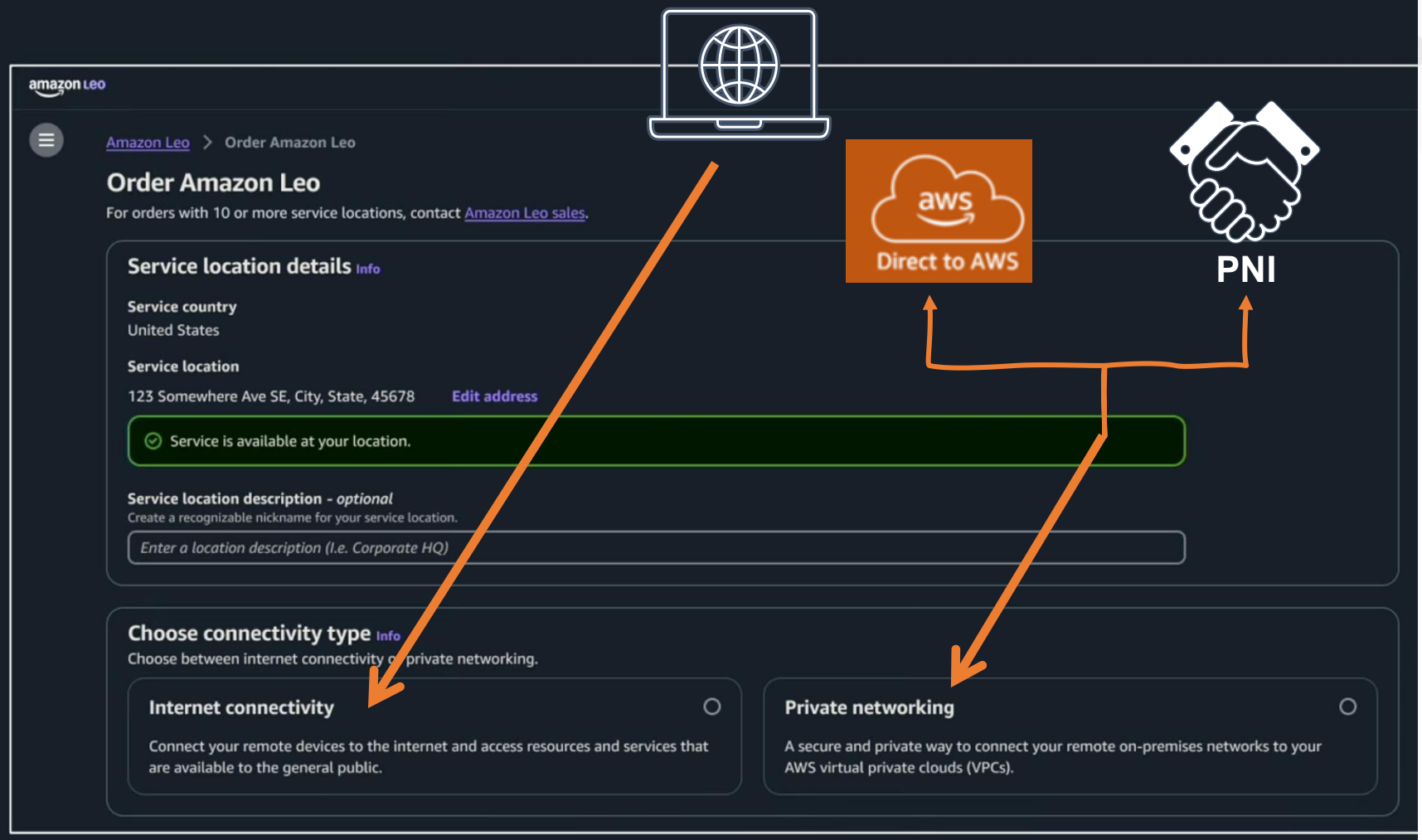
Making cloud apps play nice with Starlink and Amazon LEO

Amazon LEO leads straight to AWS cloud

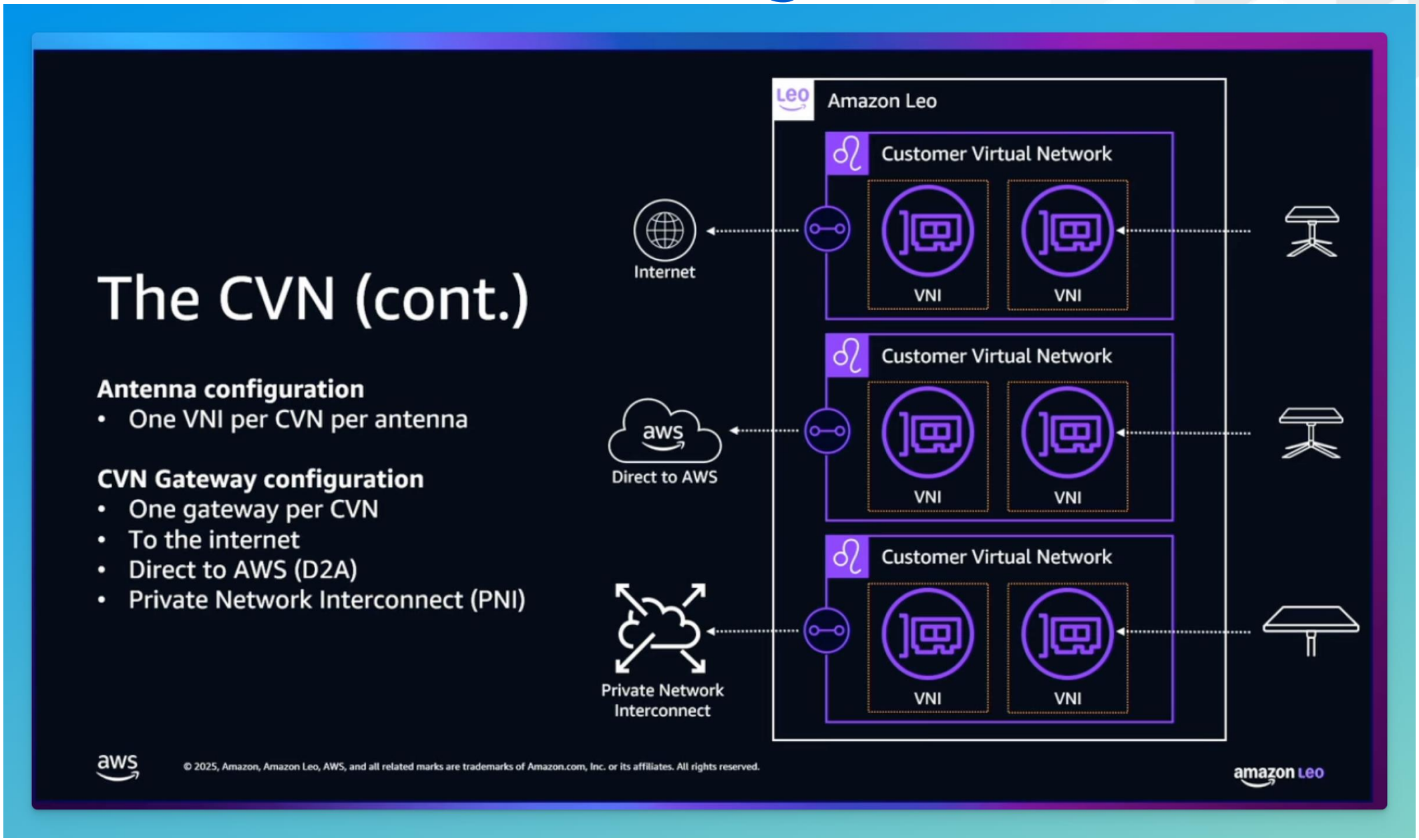
amazon Leo

Order

- Check availability
- Choose antennas
- Choose mount
- Choose a plan
  - Internet
  - Private networking

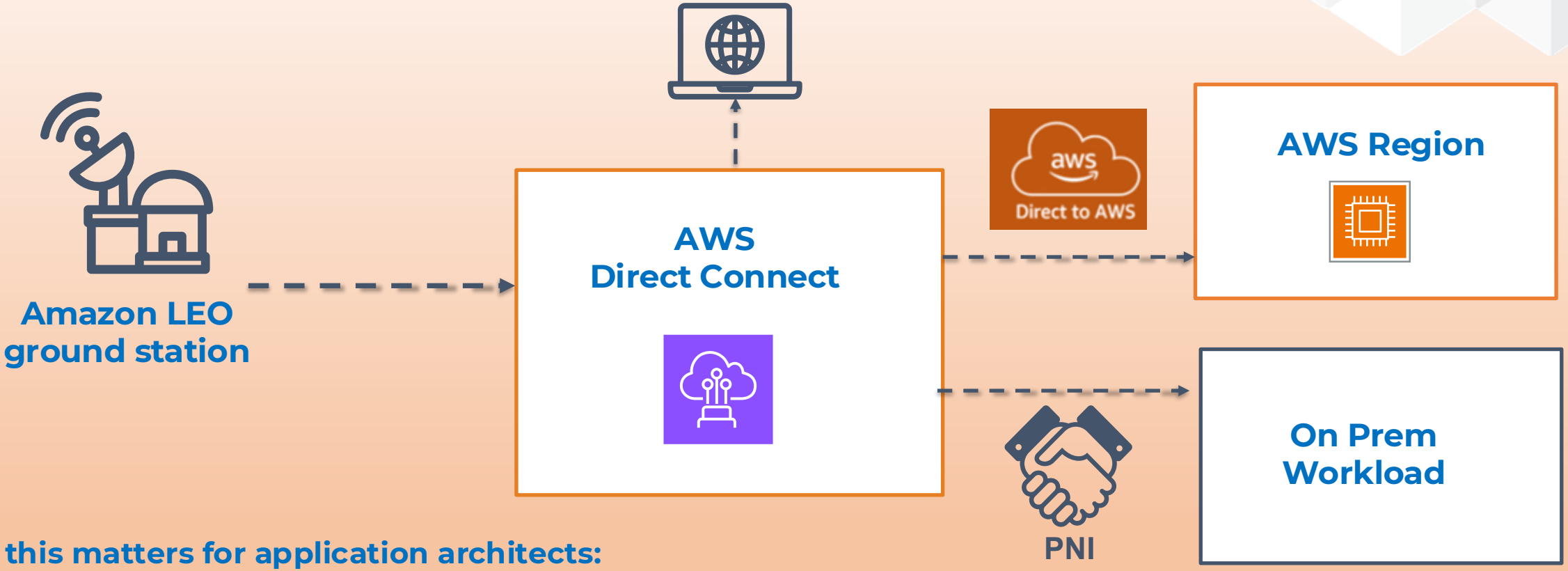


Amazon LEO leads straight to AWS cloud



AWS Direct Connect

The architectural backbone of Amazon LEO B2B service

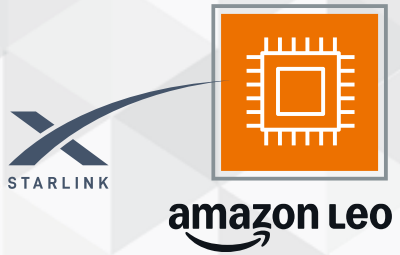


Why this matters for application architects:

Once the traffic reaches the AWS region, it behaves like any other VPC traffic.

Region selection considerations

Five inputs to the where-do-I-host question



User population geography

Where are your subscribers concentrated, and where do their satellite ground stations land?

Data sovereignty

Which countries' regulators care which workloads, and which regions are in-jurisdiction?

Egress cost

Inter-region and inter-cloud egress quietly compounds; model it before you commit.

Latency budget

How tight is the round-trip target for your application class?

Failover topology

What does your second region (or second cloud) look like when the first one is gone?

NANOG™

Hyperscalers have cloud-cloud peering now

Allows high MTU, high QoS, low jitter, and private peering for 1Gbps+ needs

AWS Interconnect - Multicloud

The natural for Amazon LEO B2B workloads

- us-east-1 (N. Virginia) with GCP, and Oracle
- us-west-1 (N. California) with GCP
- us-west-2 (Oregon) with GCP
- eu-west-2 (Europe – London) with GCP
- eu-central-1 (Europe – Frankfurt) with GCP
- ap-southeast-1 (Asia – SG) with GCP

GCP Cross-cloud Interconnect

Cross-cloud peer with AWS, Azure, Alibaba, and Oracle

- us-west1 (Oregon)
- us-east1 (South Carolina)
- europe-west4 (Netherlands)
- europe-north1 (Finland)
- asia-east1 (Taiwan)

When to use Private Network Interconnect (PNI)

PNI

“White Glove” peering - offered by both Starlink and Amazon LEO

- Lower jitter - no internet weather
- More throughput - higher MTU sizes and QoS
- Security - traffic never traverses the public internet

Equivalent products at each cloud+ LEO provider:

AWS Direct Connect · Starlink PNI

NANOG™

Hosting decision framework

User type × workload type → where to deploy

	Latency-sensitive	Throughput / batch
Static	Pin to nearest cloud region. Use PNI for predictable RTT.	Pick lowest egress cost. Co-locate with data.
in-motion	Multi-region active-active. Most components stateless. Stateful components are migration aware.	Stage at edge / CDN. Batch-upload to home region.

Part IV
Application design patterns

Cloud network and application architecture guidelines

Pattern 1 — Static low-latency

Pin to nearest cloud region

WHEN TO USE

Subscriber sits at a fixed location. Round-trip budget is tight (< 50 ms ideal).

Building blocks

- Single AWS or Google Cloud region nearest the subscriber's ground station
- PNI for select B2B workloads
- Edge caching (CloudFront / Cloud CDN) for read-heavy traffic

Examples: On-demand video, Most online commerce, Most apps

Pattern 2 — in-motion resilient

Multi-region, active-active

WHEN TO USE

Subscriber is moving (fast) – primarily aviation, less so for maritime -> cross regions/IX points mid-session.

Building blocks

- Active-active across at least two regions in the user's likely transit corridor
- Stateless services highly desirable
- Account for TCP layer (HTTP/3 with QUIC?), and TLS resets

Examples: in-flight Wi-Fi services

Pattern 3 — Cross-cloud failover

Primary in most-favored cloud, backup in another Region, or another cloud

WHEN TO USE

B2B workload cannot tolerate a single-cloud regional outage.

Building blocks

- Dual cloud region designs with automatic failover
- Cross-cloud interconnects
- Replicated state — database streaming, object-store mirroring

Examples: mission-critical fleet operations, live sports events, financial services back-ends

Three takeaways

Building performant and resilient apps and workloads

1

Routing path is the only thing that's truly different

Once LEO traffic lands in your cloud region, it behaves like any other VPC traffic. AWS Direct Connect makes Amazon LEO B2B look like a familiar architecture.

2

Static and in-motion users need different patterns

Pin static workloads to the nearest region. Build in-motion workloads as multi-region, with active-active.

3

Multicloud peering opens new failover designs

AWS Interconnect Multicloud and similar cloud-cloud interconnects let you fail over – useful for niche B2B high-value (“too expensive to fail”) workloads like live-event video broadcasts.

Thank you

Sid Mathur, Fastah Inc.
sid@getfastah.com
617-500-7576

NANOG™