Benefits of Peering for AI Workloads

William Sigmund
Director,
Sales Engineering

phoenixNAP

Brandon Ross
Senior Interconnection
Consultant

DE-CIX

Stay Tuned!

Poll Question

01

What kind of IT infrastructure are you using for your AI workloads?

a) Public cloud

b) Private cloud

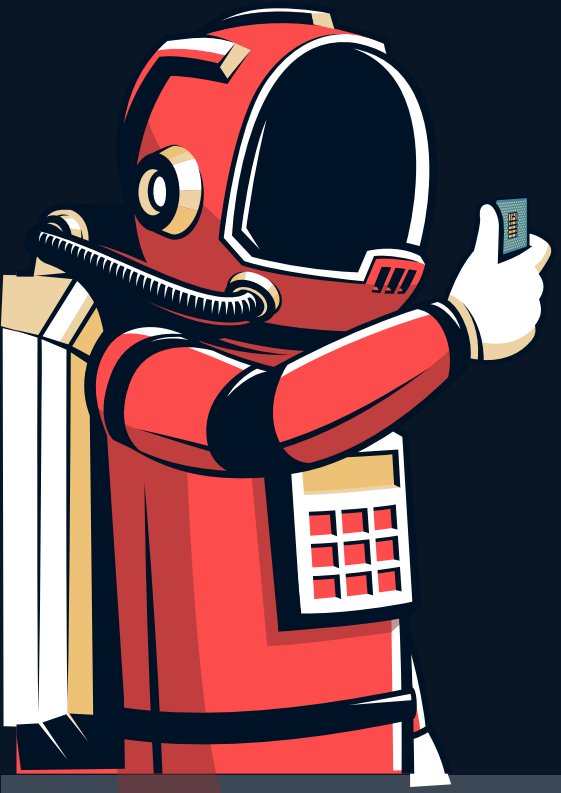
c) Multi-cloud

d) On-prem

e) Other (Hybrid)

Poll Question

02



What would you consider to be the top challenge of your AI workloads?

a) Latency

b) Availability

c) Security

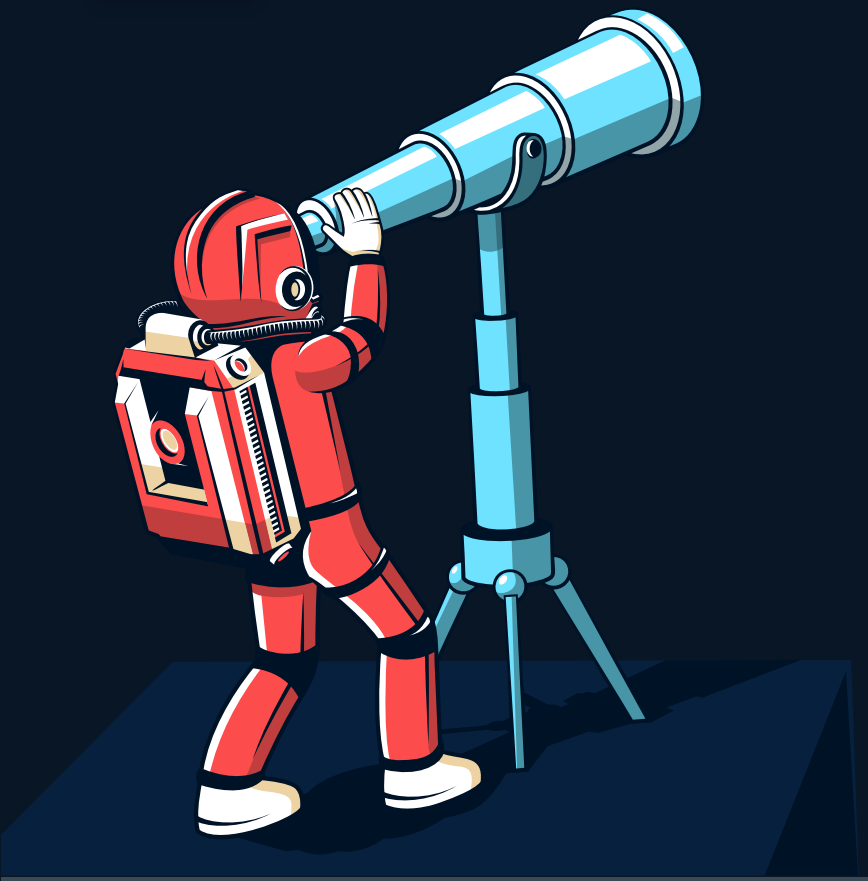
d) Compute

e) Storage

f) Other

Poll Question

03



Does your organization peer?

a) Yes, at several Internet Exchanges

b) Yes, at least one Internet Exchange

c) Yes, but not through Internet Exchanges

d) Maybe/not sure

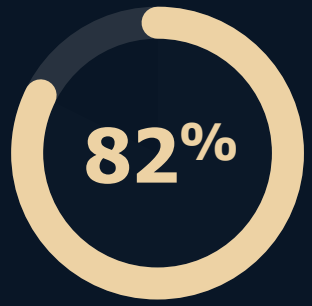
e) No

As AI Usage Surges, Infrastructure Sticking Points Persist

Organizations adopting AI still struggle to optimize ROI, and IT infra bottlenecks remain a big part of the problem.

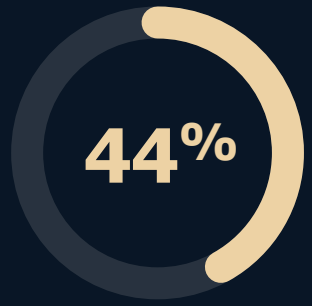
Fueling a market poised to hit $2.4 trillion in 2032, AI is seeing an unprecedented amount of investment. However, a lack of maturity and high infrastructure requirements pose a major challenge to adopters.

Combined with record-low data center vacancy rates (1.9% in 2025) despite a 34% supply increase, AI companies are facing a unique landscape of constraints.



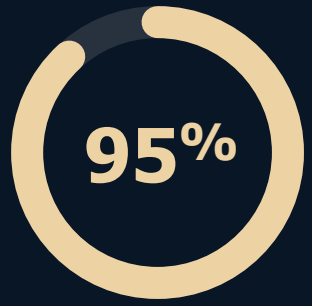
Performance Issues

The vast majority of organizations have run into at least one performance issue with their AI workloads in 12 months.



IT Infra Impedes Growth

44% of organizations point to IT infrastructure constraints as the top barrier to expanding AI initiatives.



More AI, More Threats

Almost all orgs believe that increased investment into AI has also grown their vulnerability to cyber threats.

Sources: Marketsandmarkets, Flexential

AI Is Mostly Hybrid, But Orgs Lack Confidence in Infra "Future-Proofness"

42%	On-Prem + Cloud
35%	Mostly Public Cloud
15%	Mostly On-Prem

Organizations tend to leverage hybrid deployments for their AI applications, committing their latency-sensitive workloads to on-prem while relying on the cloud to handle less performance-sensitive operations.

54%	Somewhat Confident
25%	Not Very Confident
13%	Very Confident

As much as 79% of organizations don't have full confidence that their current IT infrastructure can meet AI demands in the next two to three years, showing a gap between AI growth and adequate infra strategies.

Source: A10 Networks

Top Performance Challenges

What C-suites call the biggest performance bottlenecks:

- Compute limitations
- Software constraints
- Network latency or throughput
- Security

Compute

Interestingly, around **27% of organizations** are using compute solutions not optimized for AI workloads. Most solutions for fixing the compute gap include **faster networking & edge computing**.

Network

Bandwidth shortages and latency are massive constraints. **59% of orgs reported bandwidth problems** in 2025 compared to 43% in 2024, In the same period, **latency issues grew from 32% to 53%.**

Storage

Although storage roadblocks are less common, many organizations plan to increase their storage capabilities. Around **41% of orgs** have storage system upgrades on top of their priority list.

Sources: A10 Networks, Flexential

Latency & Availability on Top of Orgs' Priorities List

49%

Important

Almost half of all organizations consider low latency and high availability a high priority, with a bit of latency or downtime being acceptable.

23%

Extremely Important

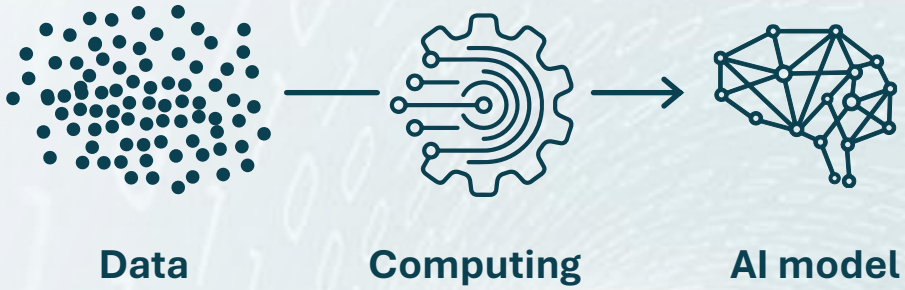
Nearly a quarter of orgs prioritize reducing latency and maximizing availability and require real-time responsiveness with no downtime.

22%

Moderately Important

Some companies handle a mix of real-time and batch workloads with varying levels of importance in terms of latency and availability.

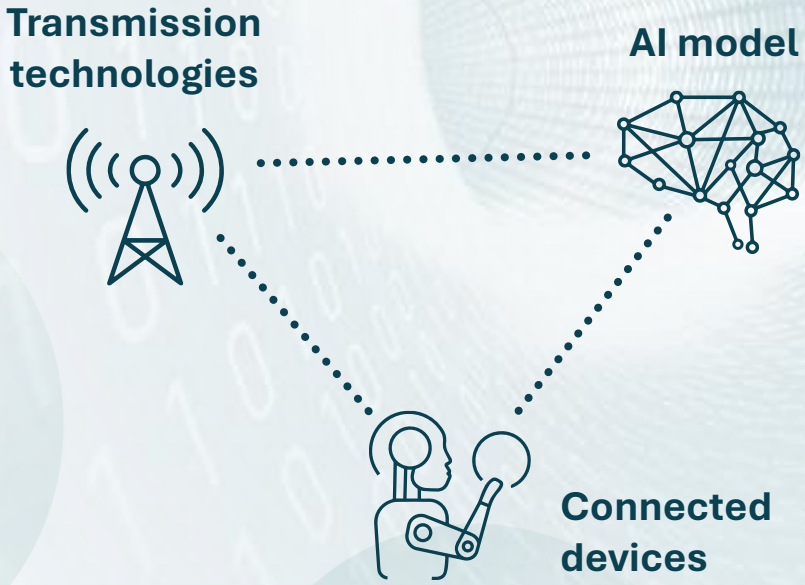
Source: A10 Networks



AI Training

Develop the model in the first place, periodic retraining of the model

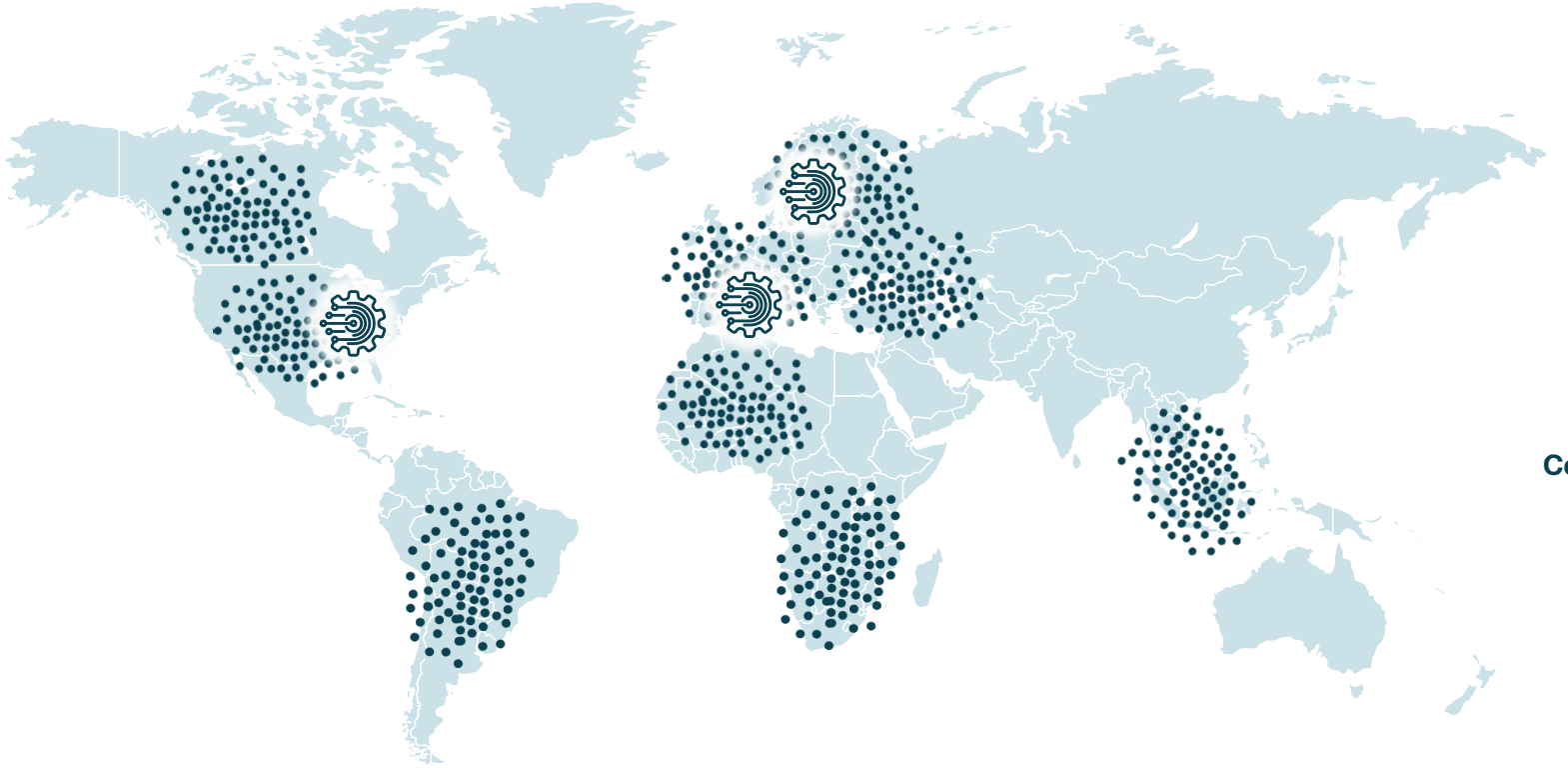
AI stages

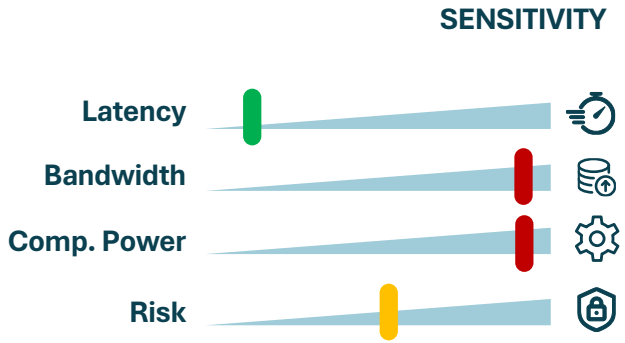


AI Inference ("AI in action")

AI-based diagnostics, simulate, predict, pilot

AI Training
requires manageability of multi-cloud data sourcing & computing





Data
everywhere

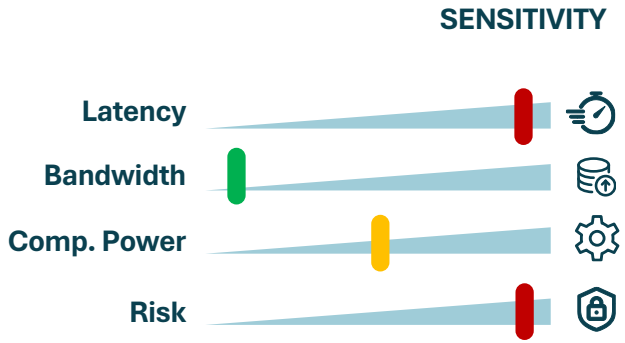
Computing power
somewhere

AI Inference Triangle

Transmission technologies

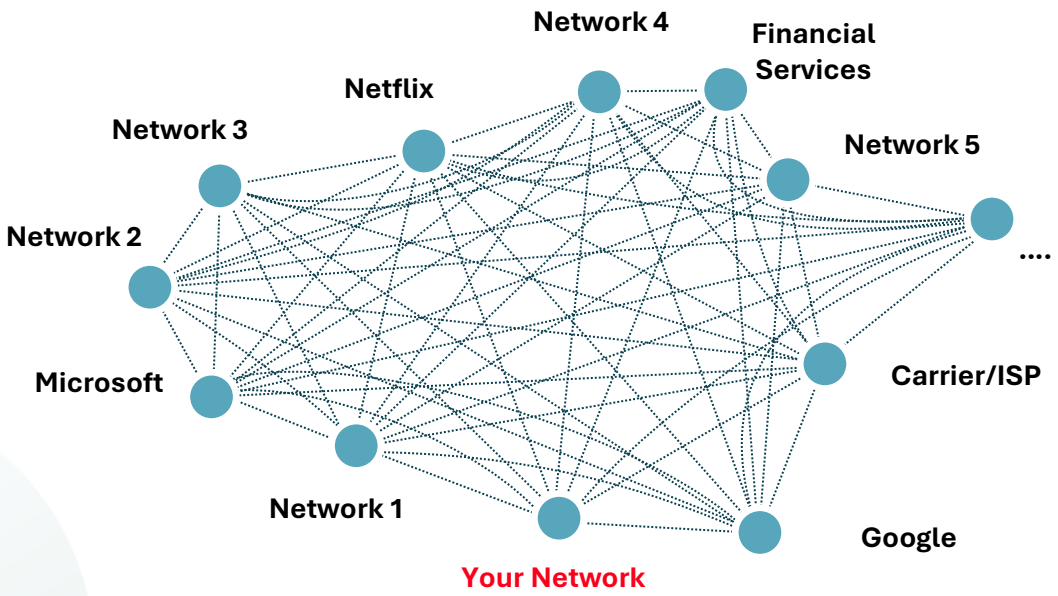
AI model

Connected devices

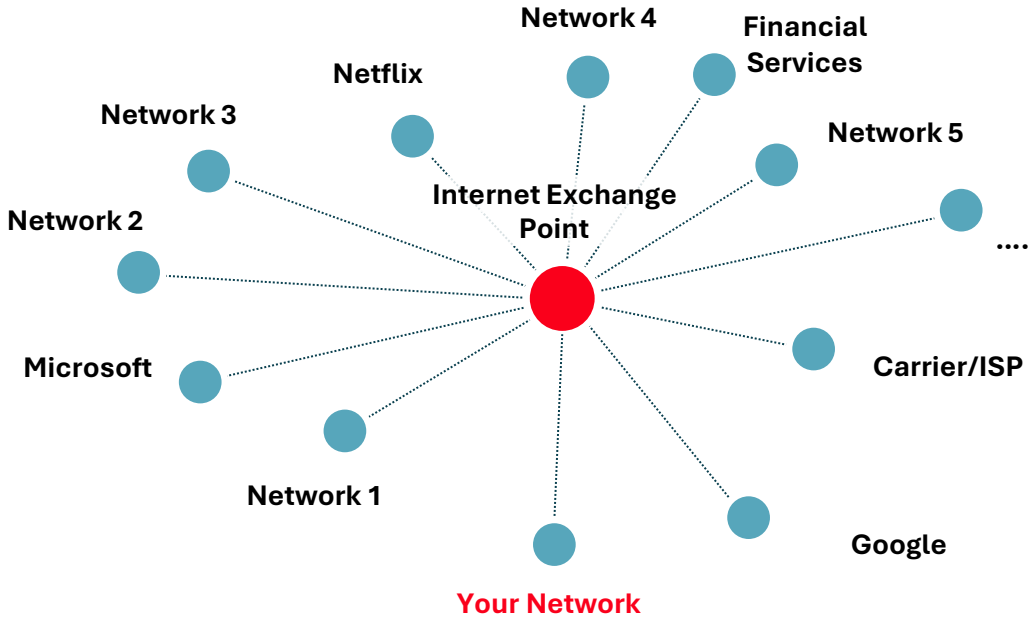


What is Peering?

Conventional Private Peering (PNI)



Scalable Internet Exchange Point (IXP or just IX)



How Peering Enhances AI Performance

Improved Security & Control

- Direct connection
- Increased route control
- Visibility

Performance boost

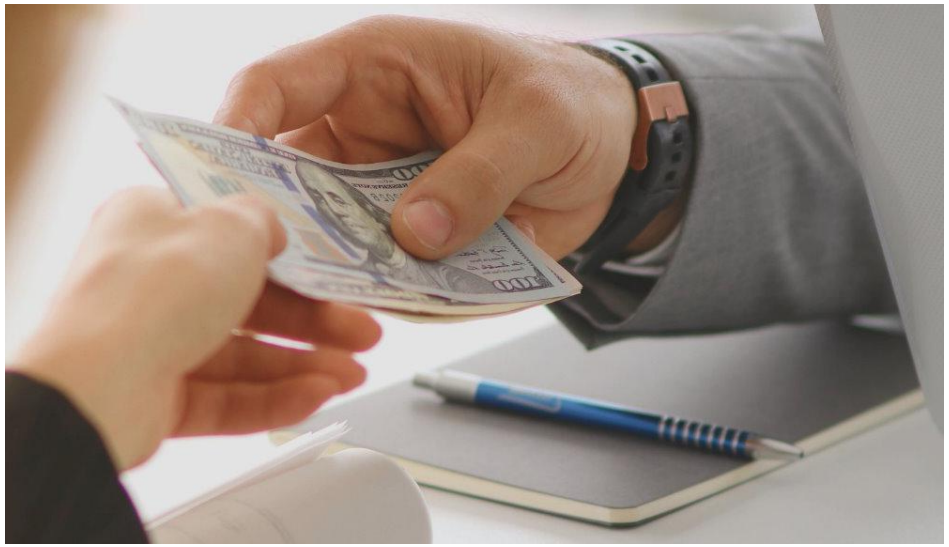
- Lowest latency path
- Increased capacity and flexibility

How Peering Enhances AI Performance

Enhanced Resilience

- Redundant connection to transit and other exchanges
- Diversified network paths
- Diversified players
- DDoS mitigation

Cost-Effectiveness



- No need for expensive transit agreements
- Slashes long-distance data transport costs

DE-CIX is the world's #1 Cloud & Internet Exchange operator, and the world's largest carrier-neutral interconnection ecosystem

60
Internet & Cloud Exchanges

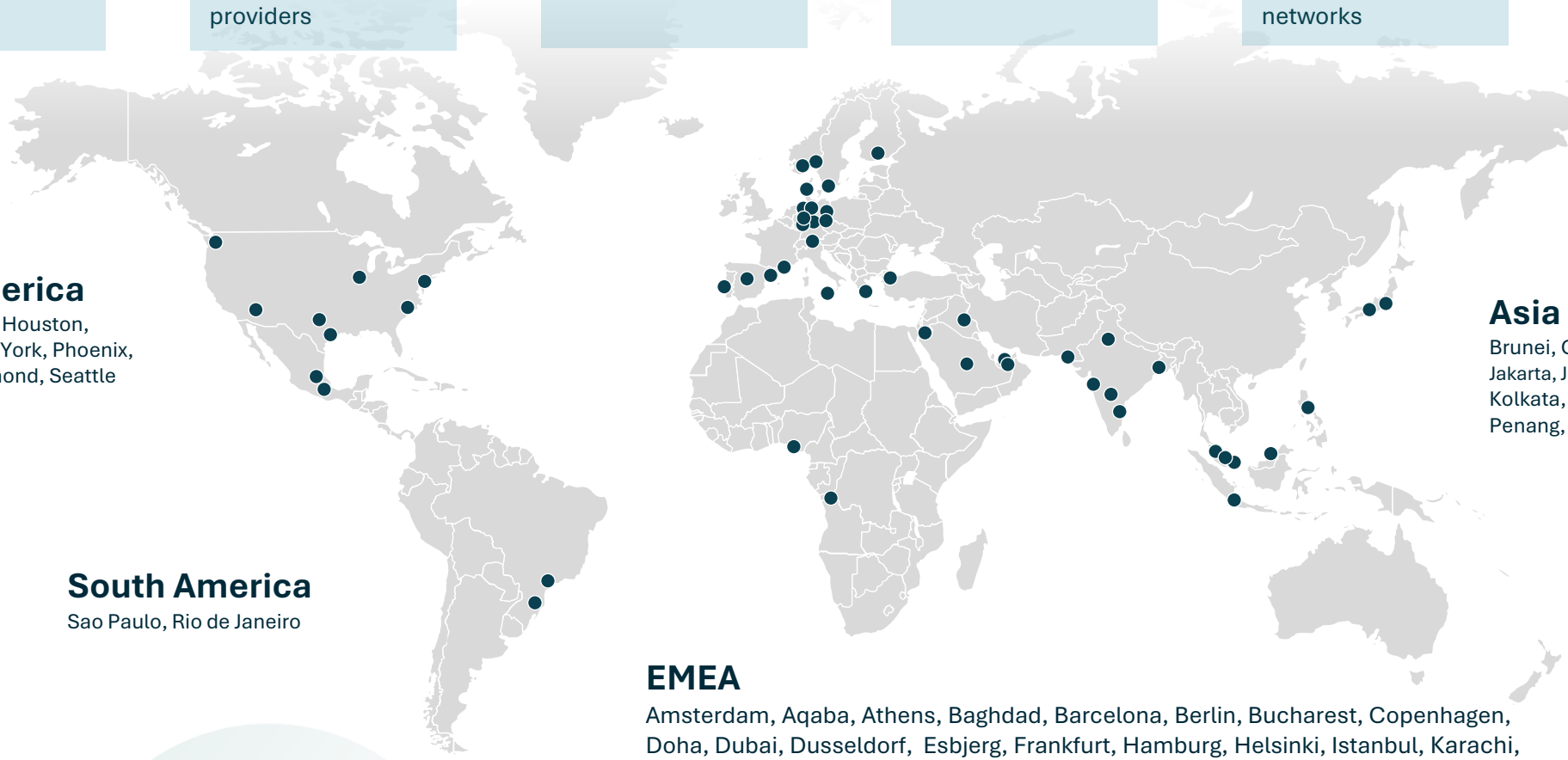
50+
Cloud service providers

600+
metros

1,000s
of data centers

4,000+
connected networks

185+
Tbits capacity



North America
Chicago, Dallas, Houston, Mexico City, New York, Phoenix, Queretaro, Richmond, Seattle

South America
Sao Paulo, Rio de Janeiro

Asia Pacific
Brunei, Chennai, Delhi, Hyderabad, Jakarta, Johor Bahru, Kuala Lumpur, Kolkata, Manila, Mumbai, Osaka, Penang, Singapore, Tokyo

EMEA
Amsterdam, Aqaba, Athens, Baghdad, Barcelona, Berlin, Bucharest, Copenhagen, Doha, Dubai, Dusseldorf, Esbjerg, Frankfurt, Hamburg, Helsinki, Istanbul, Karachi, Kinshasa, Kristiansand, Lagos, Leipzig, Lisbon, Madrid, Marseille, Munich, Oslo, Palermo, Prague, Riyadh*, Ruhr region, Sofia, Stockholm*, Warsaw

** Coming soon*

Phoenix Data Center

1

Enterprise-Grade Environment

2N power design with redundant configurations, 100,000 sq. ft. floor space **+ 550,000 sq. ft. expansion** underway.

2

Compliance

PCI DSS, SSAE 18, SOC 1, SOC 2, and SOC 3-compliant, and AZRamp-certified facility. Meets or exceeds requirements for HIPAA, GDPR, SOX, or GLBA.

3

Security

Tier 3+ data center design, trained guards on premises 24/7/365, biometric access control, video monitoring, VESDA.

4

Customization and Support

¼ cabinet to private cage configurations, hardware leasing, easy access to hybrid cloud-ready solutions and services. 24/7 Expert NOC and SOC support + Remote Hands option.

5

Connectivity

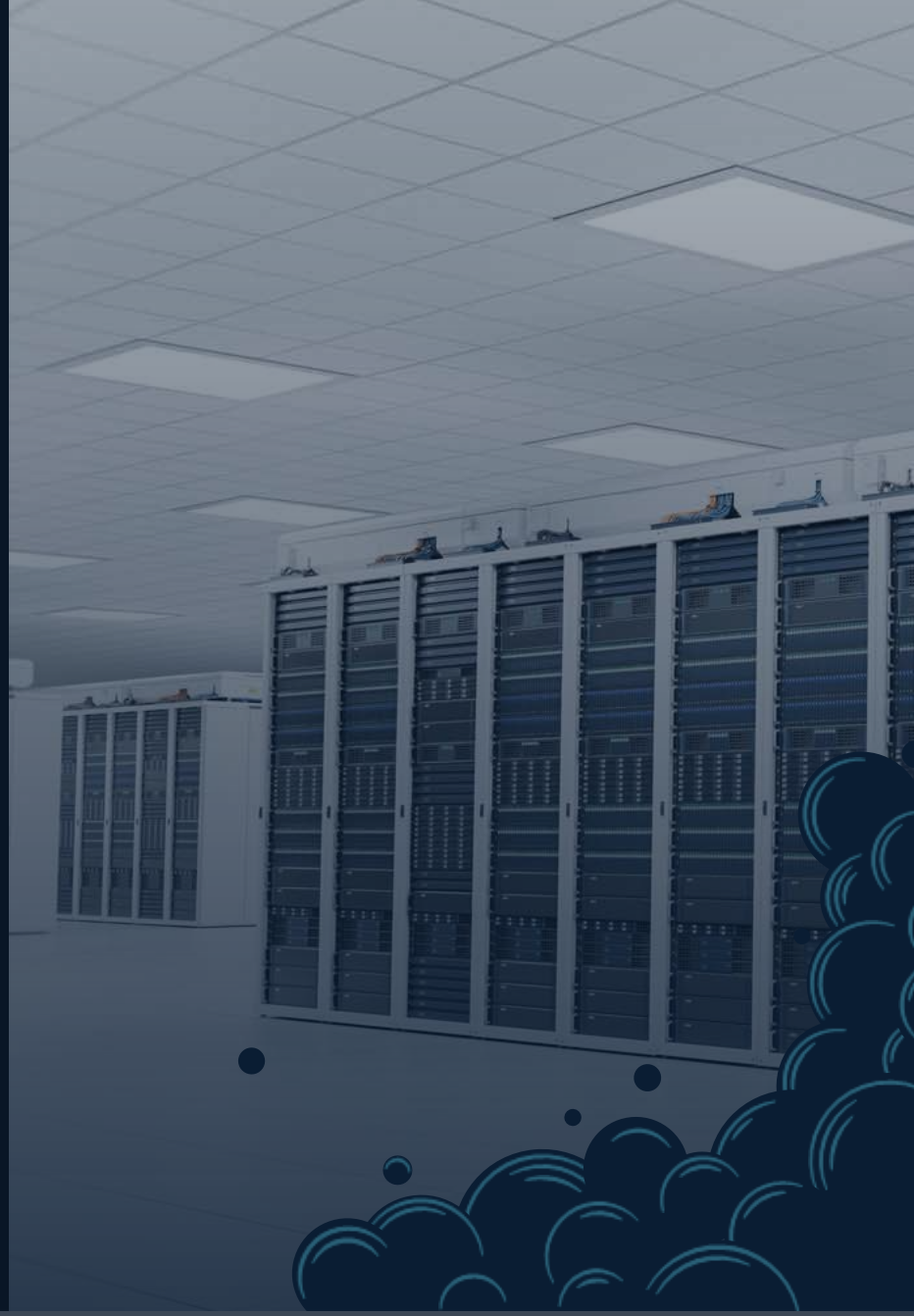
2 Meet-Me-Rooms with redundant fiber vaults, 40+ carriers + phoenixNAP's redundant Tier 1 carrier blend, Direct link to AWS, Google Cloud, and other popular hyperscalers.

phoenixNAP Data Center Expansion

- **First floor expansion**
  11,000 sq. ft. of new data center space

- **New Bare Metal Cloud cage**
  3,940 sq. ft. & 154 cabinets

- **Hyperscaler on-ramp expansion**
  1st & 2nd floor data halls (WIP)



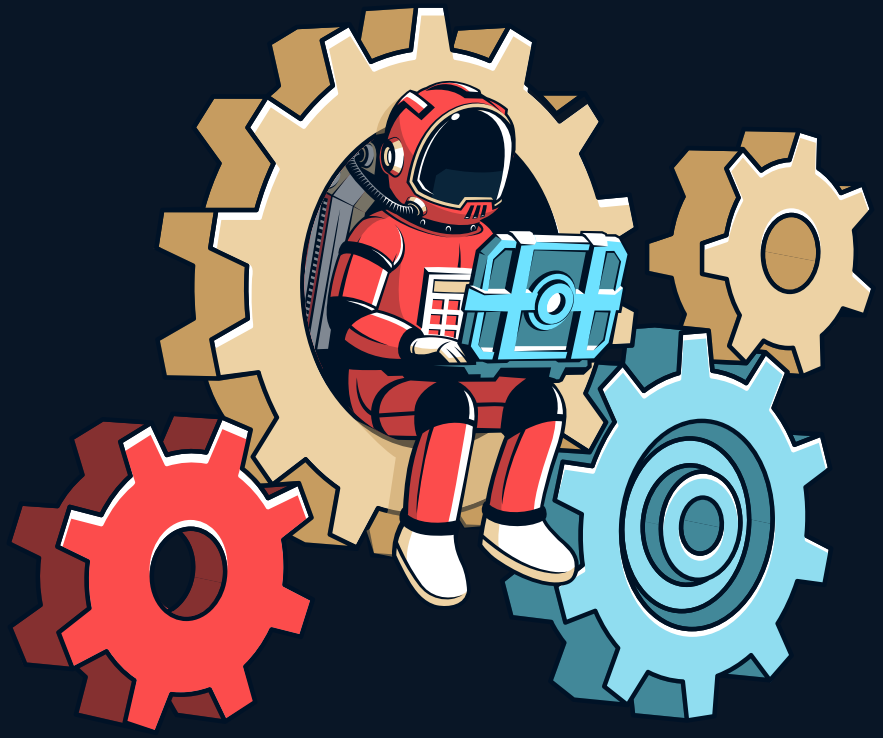
Introducing phoenixNAP's PHX2.0 Data Center – Breaking Ground Soon

- 530,000 sq. ft. of premium data center real-estate
- Redundant power and cooling
- 30 MW A and B power
- Over 44,000 racks total capacity
- Up to 44+kW power density per rack
- Heavy floor loading, perfect for AI and HPC racks
- Unique rooftop space for satellites, antennas, and specialized connectivity solutions
- 4 network vaults with direct connectivity to our current data center resources

Q and A

Poll Question

04



Would you Like to Get in Touch with phoenixNAP or DE-CIX Solution Experts?

a) Yes, phoenixNAP.

b) Yes, DE-CIX.

c) Yes, both phoenixNAP and DE-CIX.

d) No, not now.

Raffle Time!

THANK YOU!

phoenixNAP®
GLOBAL IT SERVICES

phoenixnap.com

facebook/phoenixnap

@phoenixnap

sales@phoenixnap.com

1.877.588.5918