palais des congrès de paris

From Mobile Networks to AI Infrastructure

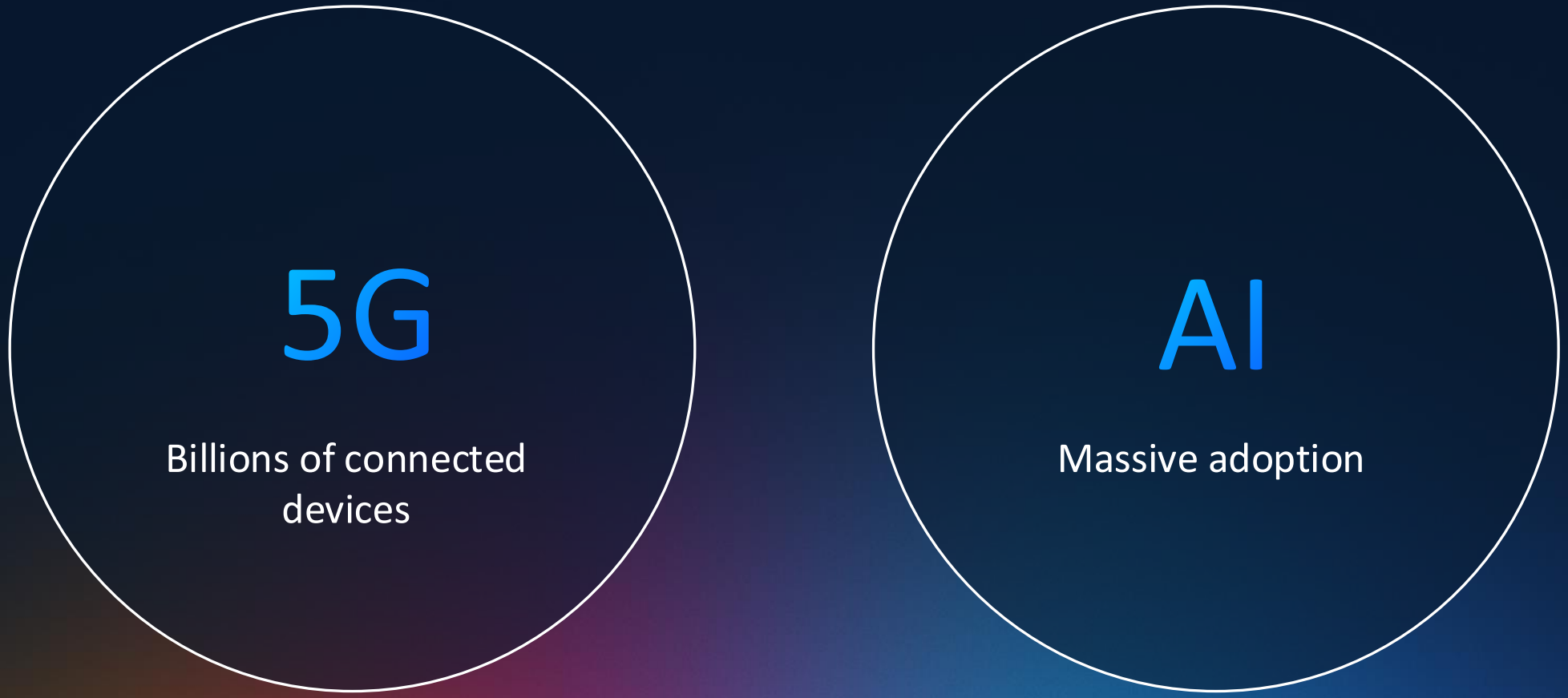
SRv6 as the unifying technology

Michael Beesley
Chief Technology Officer, SP Networking
Cisco

UPPERSIDE WORLD PARIS 2026

palais des congrès de paris

Current state



Networking is fundamental

Two industries. One protocol.

Mobile and AI are rewriting network demands simultaneously

Mobile

Mass scale at every level

Native programmability requires simplicity

All apps are IP

Future-proofing architecture is a necessity

AI

82% CAGR in AI network traffic through 2029

10,000+ GPUs must synchronize in milliseconds

Every dropped packet degrades entire training run

East-west traffic at a scale traditional WAN never anticipated

The answer isn’t another protocol. It’s IP — back in control.

Programmable. Scalable. Universal.

cisco

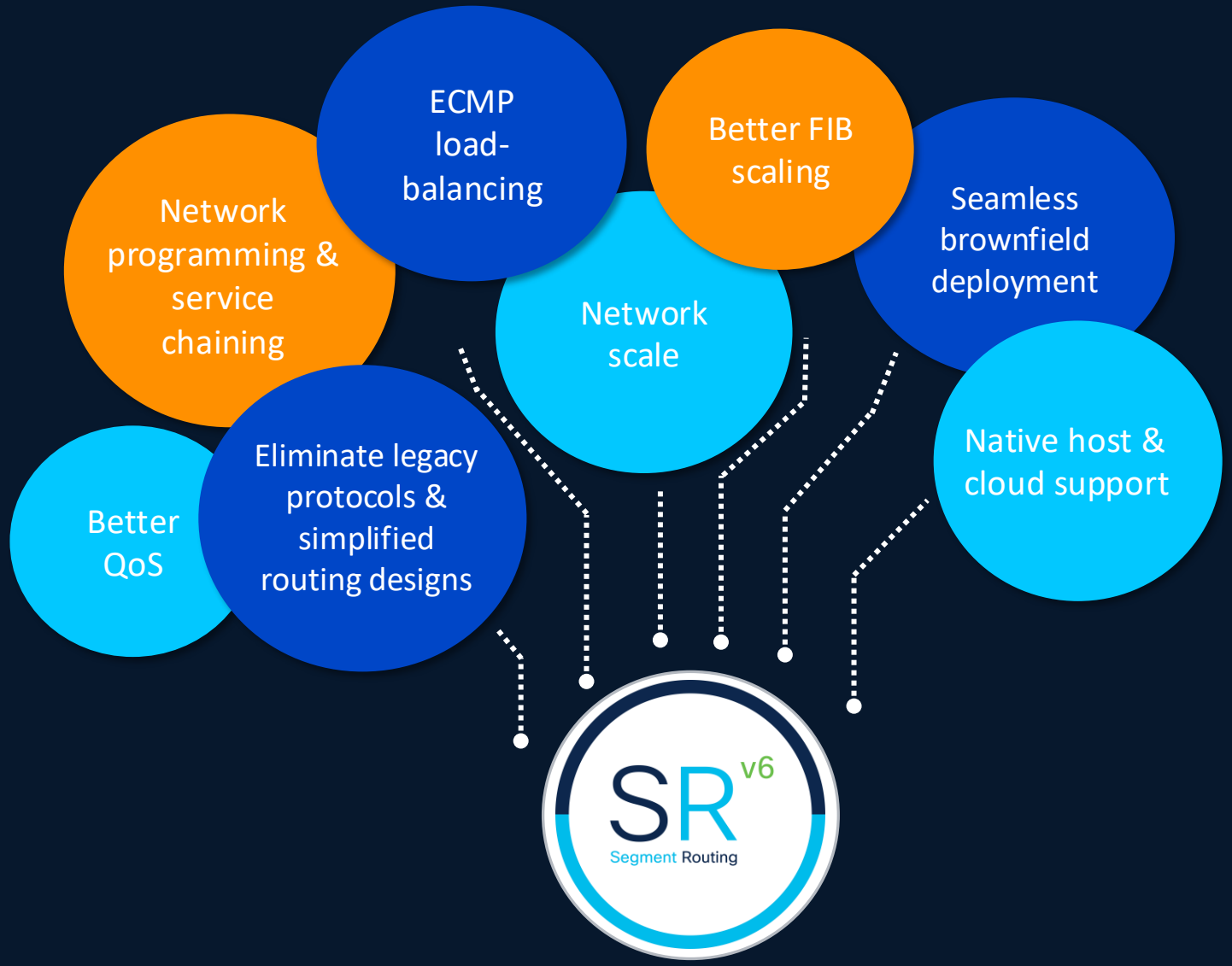
SRv6: A global standard at scale

Adopted by major CSPs worldwide across mobile, enterprise, and cloud domains

2017	uSID	3GPP	3 RFCs
IETF debut	Microsegment technology	UPF	Core standards ratified

500+ customers have adopted Segment Routing

SRv6 is the tool for network evolution



SRv6 capabilities

1

Simplified VPN services

L3VPNs across networks with only edge routers requiring SRv6, core just forwards IPv6

2

Service function chaining

Firewalls and load balancers embedded directly in routing paths - no extra protocols

3

Traffic engineering and FRR

Fine-grained control over traffic routes for low latency and bandwidth guarantees

4

Operational simplicity

Single IPv6 framework clarifies overlay protocols - easier troubleshooting and lower OPEX

5

Enhanced scalability

IPv6 address space manages large networks with fewer prefixes, simplifying routing

Mobility evolution with SRv6

5G

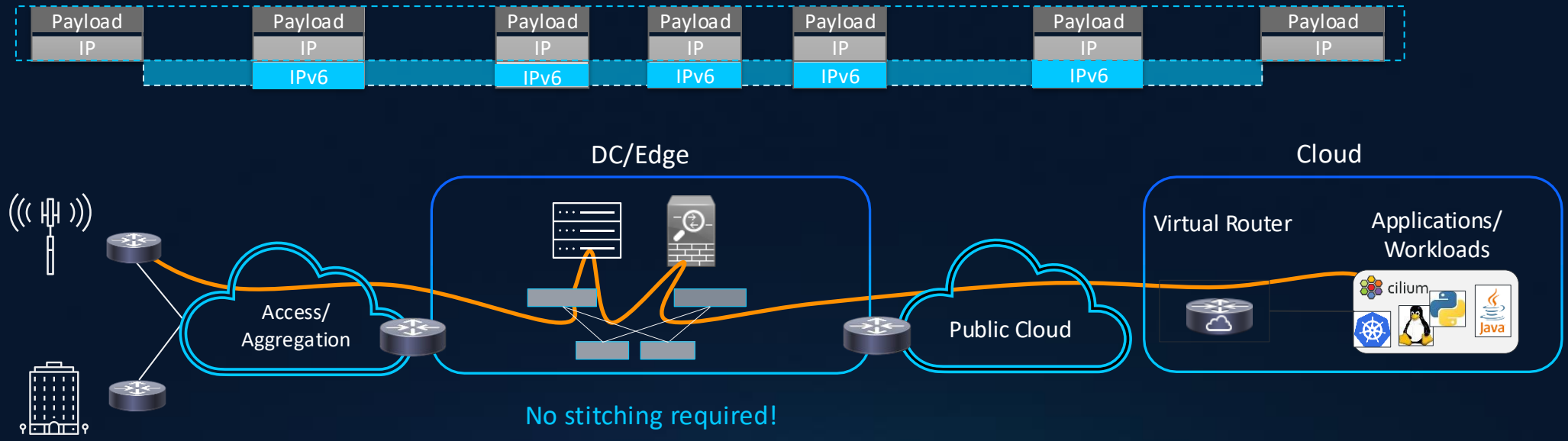
- Complex overlay tunnel architectures
- Multiple disconnected protocol domains add cost and friction
- Ultra-low latency hard to achieve
- Network slicing support limited
- Mobile edge computing bottlenecks

6G

- End-to-end IPv6 path composition
- Integrated TE, VPNs, and service chaining
- No per-session tunnel state needed
- Network resources → IPv6 addresses
- Even DWDM wavelengths addressable

cisco

Unified mobility data plane



Enables native routing (IPv6) to cloud/virtual data center providers

Works on any IPv6 network

Optimizes and simplifies service-chaining provisioning, customization, and operations

AI is rewriting the network

Efficient creation and delivery of tokens will be a critical differentiator

33%

CAGR increase in demand for AI DC capacity through 2030[1]

82%

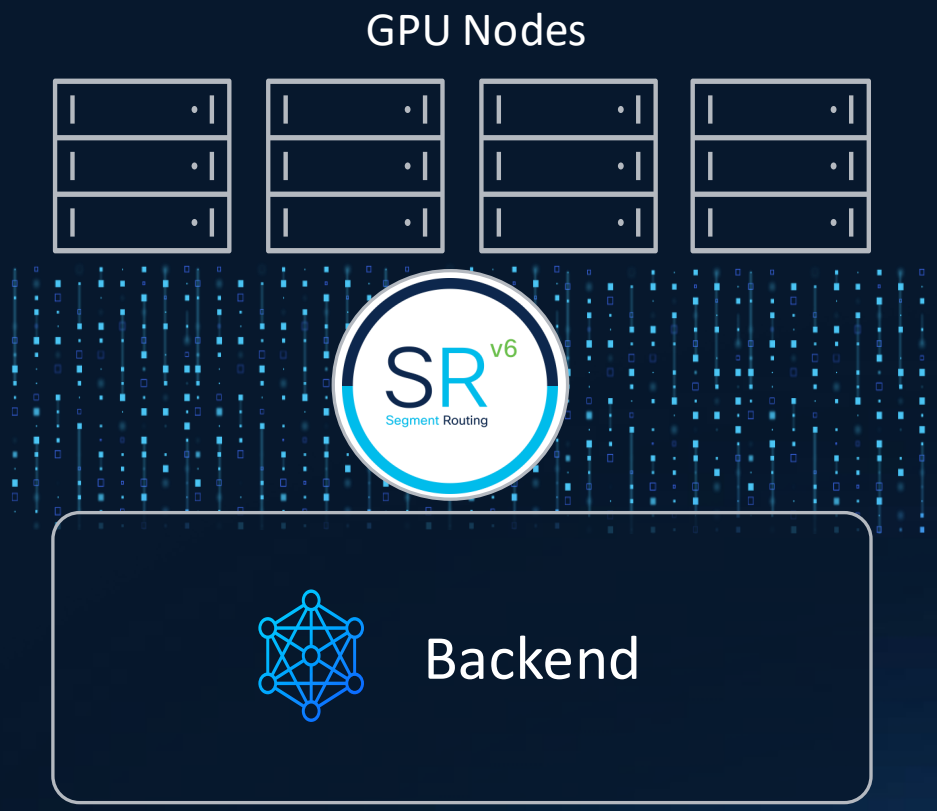
CAGR for AI network traffic through 2029[2]

75%+

Inferencing/processing moves to edge by 2030[3]

[1]McKinsey & Company
[2]650 Group
[3]KPMG

SRv6 in AI backend networks



SRv6 uSID Approach

Deterministic path placement for RDMA
NIC-level source routing via uSID
Stateless, predictable distribution
Proactive — reduces congestion frequency

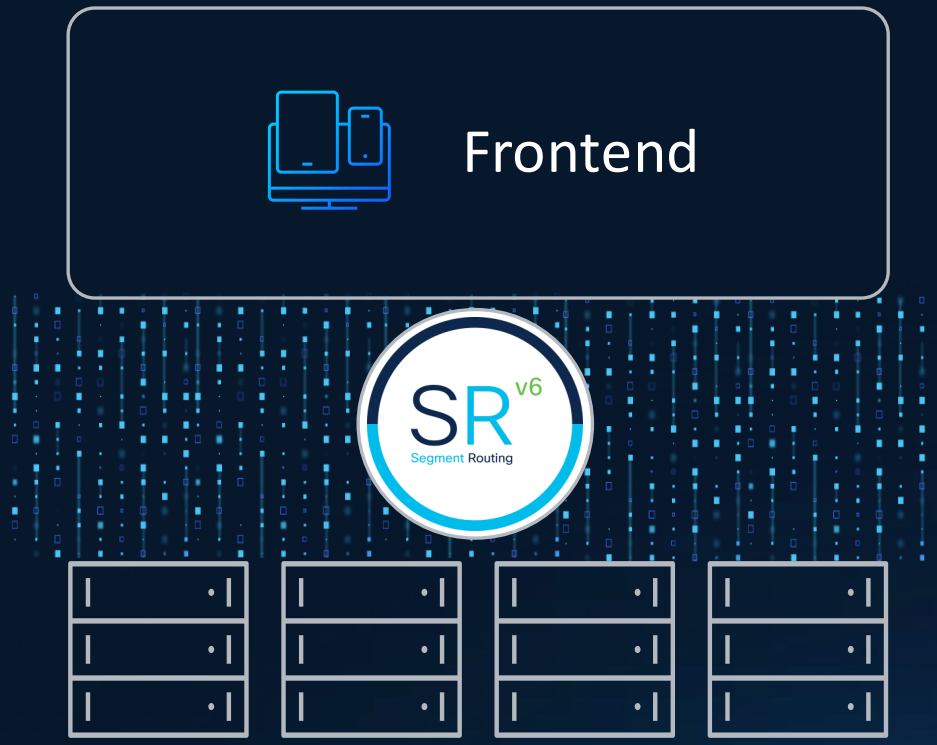
+

SRv6 Fast Failure Recovery

SRv6's explicit path encoding enables rapid rerouting. GPU jobs recover in microseconds rather than seconds.

GPU fabric optimization through deterministic path placement

SRv6 in AI frontend networks



Variety of traffic types

Checkpoint Writes	Large model checkpoints to distributed storage systems
Telemetry Streams	Real-time monitoring, metrics and observability data
Control Plane	Job scheduling, coordination and orchestration traffic
User Access	API calls, inference requests and user-facing traffic

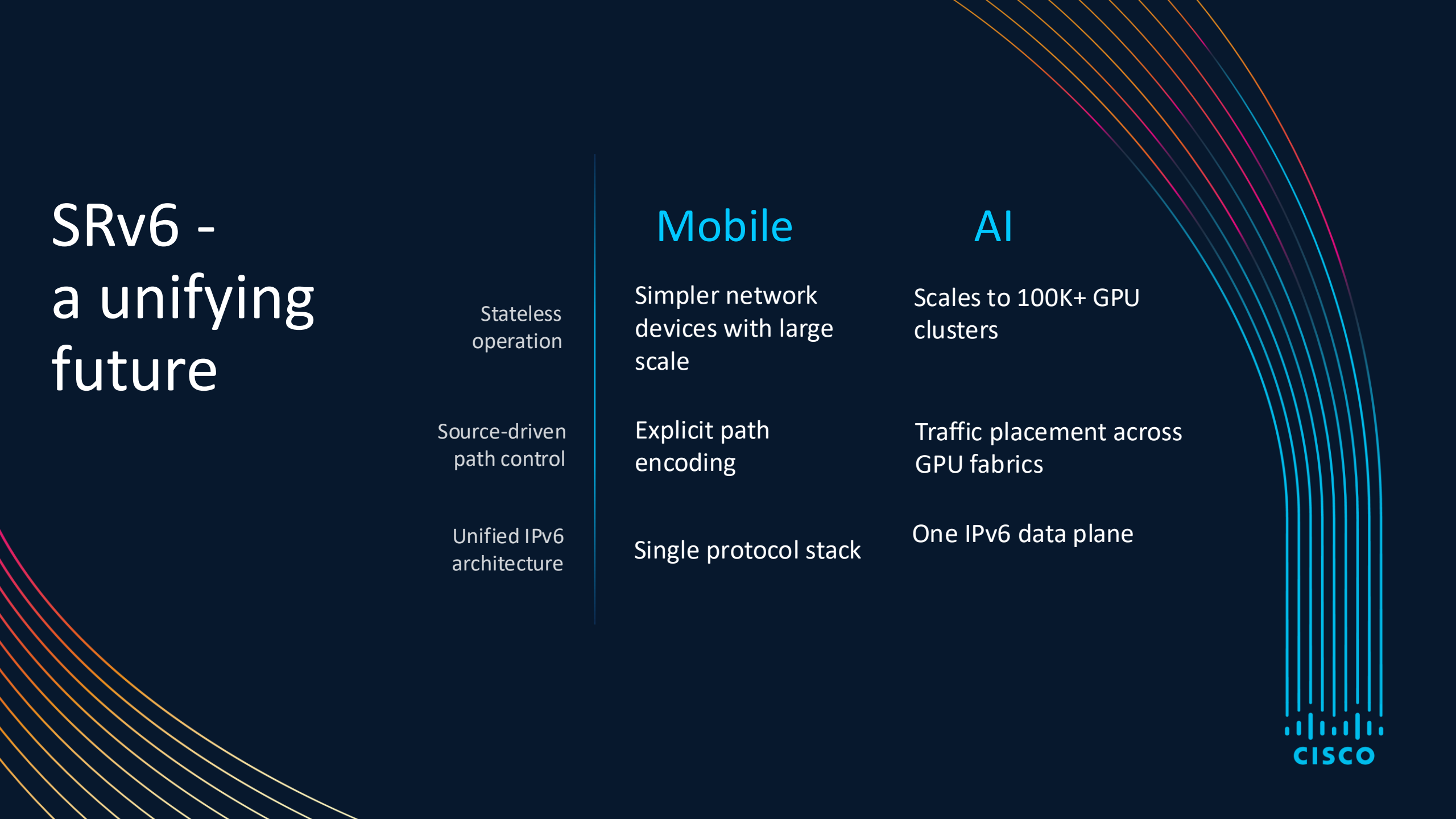
SRv6 Unified Framework

One framework for QoS policies, security policies, traffic steering — across both backend and frontend domains

Scale-across DC to WAN unified architecture

SRv6 - a unifying future

	Mobile	AI
Stateless operation	Simpler network devices with large scale	Scales to 100K+ GPU clusters
Source-driven path control	Explicit path encoding	Traffic placement across GPU fabrics
Unified IPv6 architecture	Single protocol stack	One IPv6 data plane




CISCO