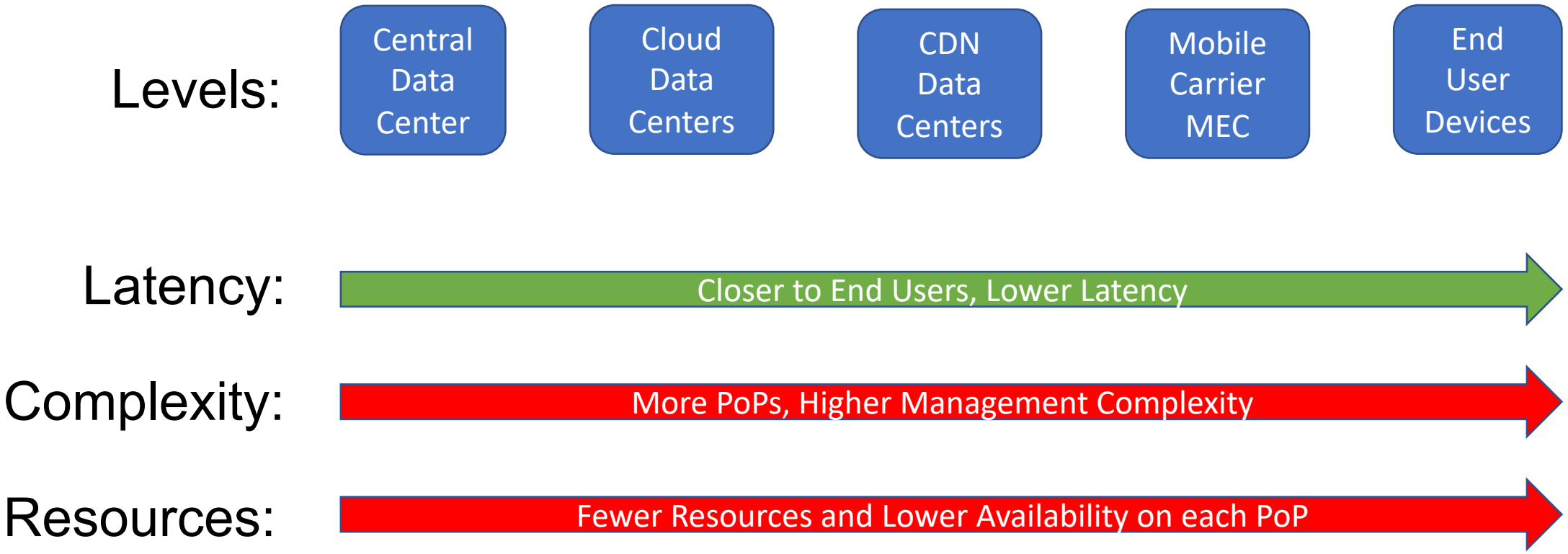
Building Edge as a Service

Dr. Bin Ni, CTO @ CDNetworks

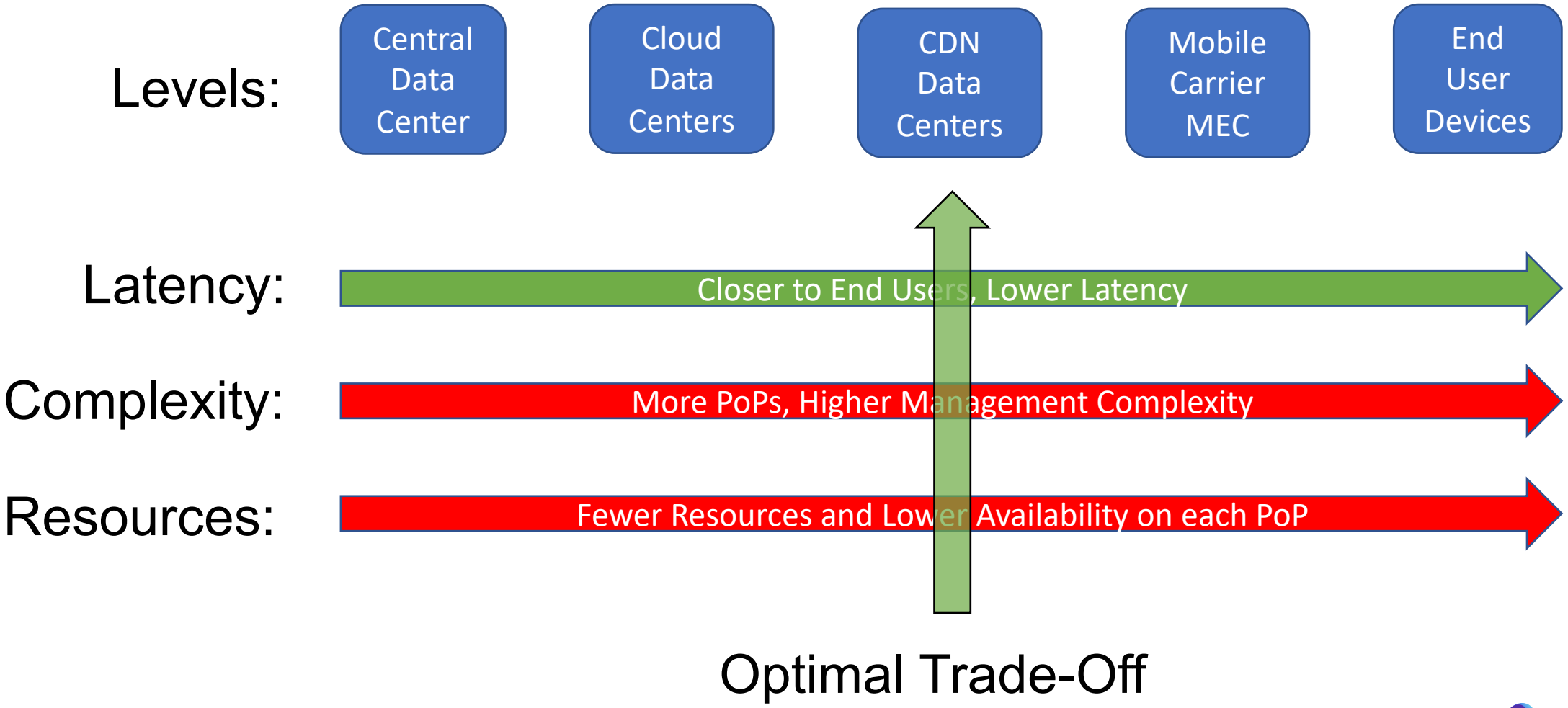
Nov 5, 2020

What is the “Edge”?
CDNetworks

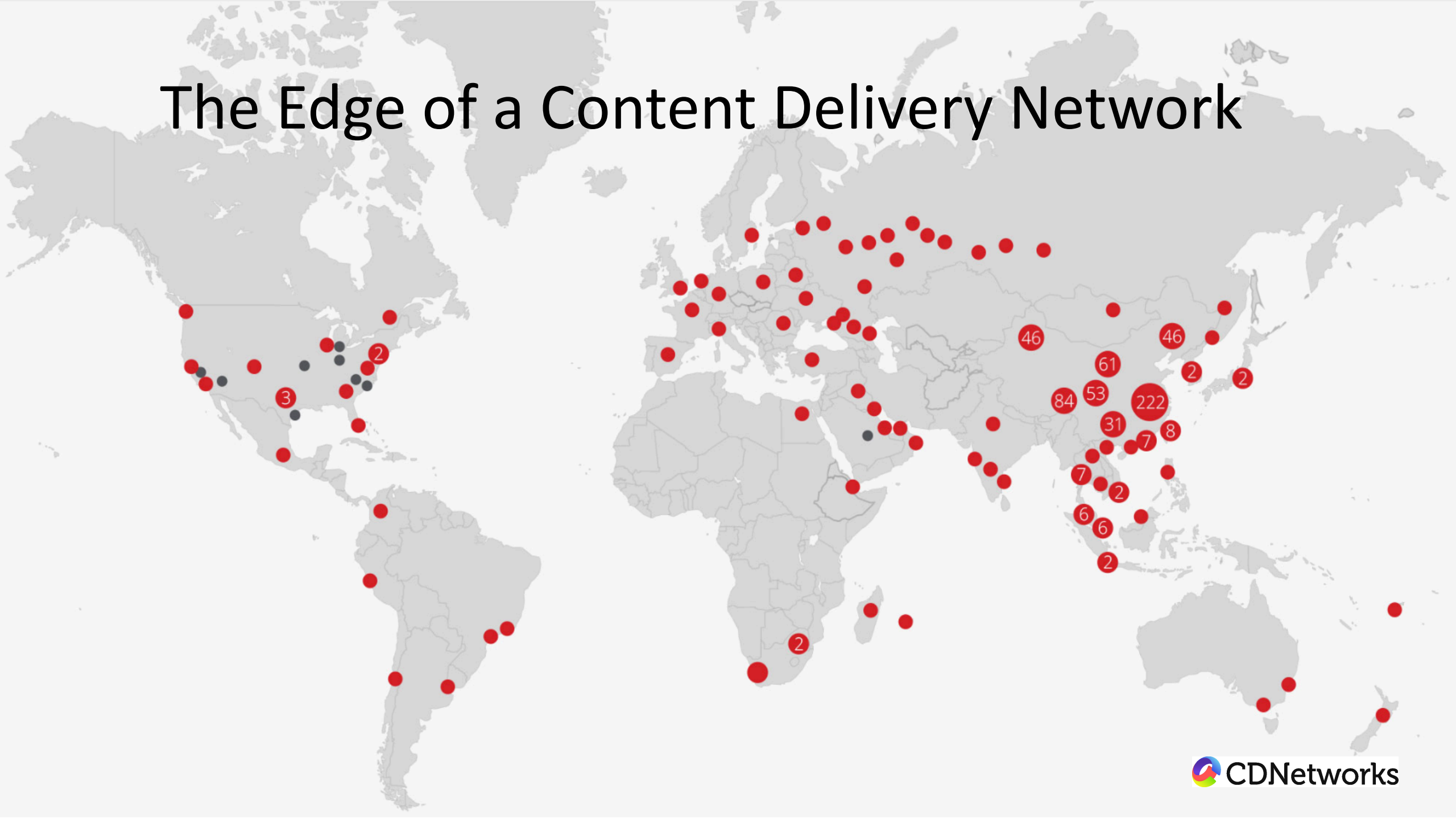
Different Levels of Edge



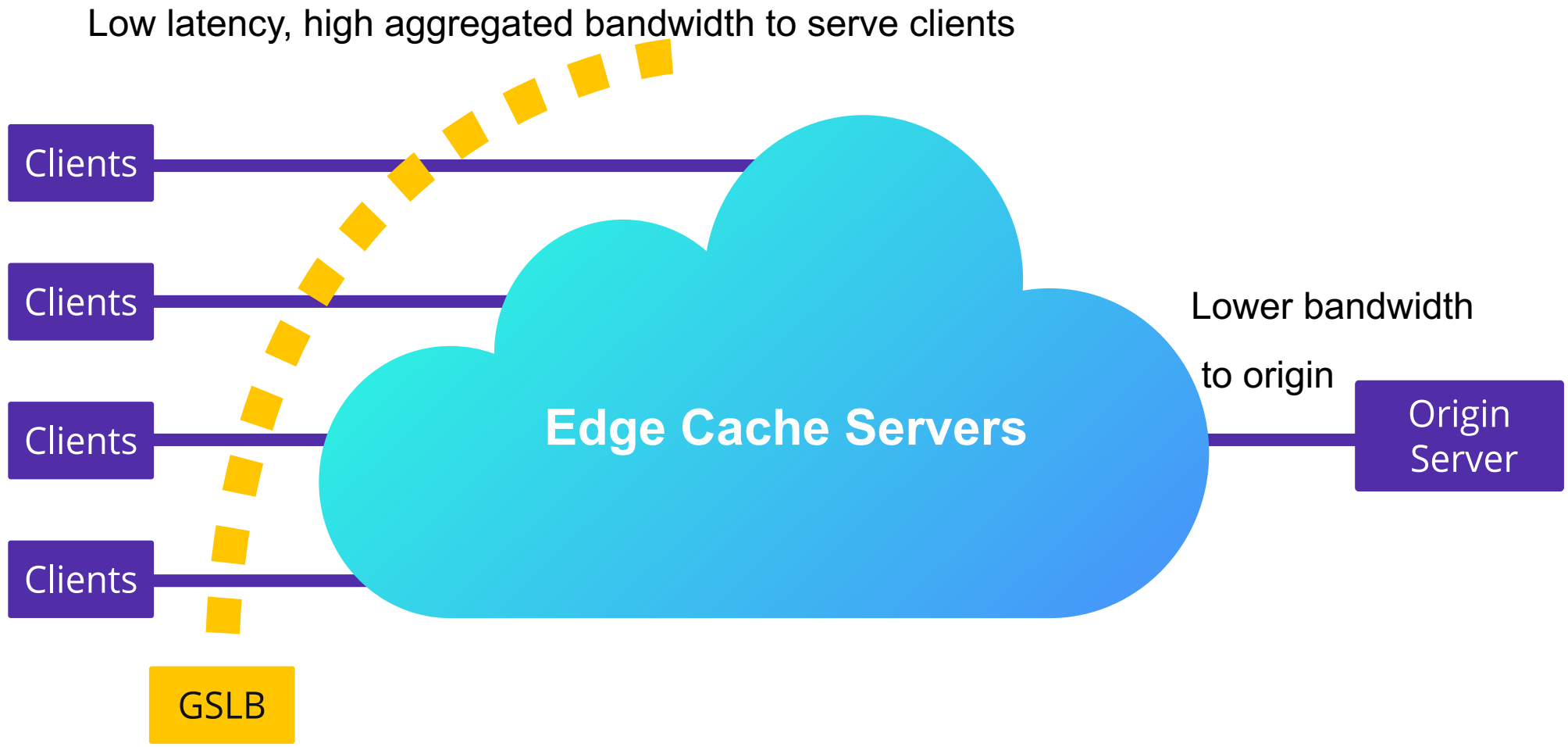
Different Levels of Edge



The Edge of a Content Delivery Network



How A CDN Works



CDN Service Latency in US and China

Platform	Median (ms)	Avail.
1. Citrix ITM Multi CDN Optimized **	19	99.98%
2. Fastly CDN	21	99.80%
3. CenturyLink CDN	21	99.76%
4. CacheFly CDN	21	99.74%
5. Akamai Object Delivery	21	99.73%
6. Azure CDN from Verizon	21	99.51%
7. Edgecast CDN	21	99.43%
8. Cl[illegible]DN	22	99.[illegible]
16. MaxCDN	24	99.75%
17. CDNetworks	24	99.75%
18. Quantil CDN	24	99.43%

Platform	Median (ms)	Avail.
1. Alibaba Cloud CDN 2	11	99.22%
2. BaishanCloud CDN (China)	12	99.20%
3. Quantil CDN	13	99.15%
4. CDNetworks	13	99.14%
5. ChinaCache CDN	13	99.09%
6. Kingsoft Cloud CDN	14	99.20%
7. Baidu Cloud CDN	15	99.23%
8. Tencent Cloud CDN	22	99.28%
9. Azure CDN from Microsoft	73	82.70%
10. Akamai Object Delivery	87	97.70%

Data Source: ”Citrix Country Report”

CDNetworks

Cloud vs CDN PoPs

	Cloud	CDN	Comments
Global PoPs	~10	~1000	~100x
PoP Size	~1M Cores	~1K Cores	Smaller PoP Size
Total BW	~100 Gbps	~10 Tbps	Higher BW
Latency	~100 ms	~10 ms	Lower latency, higher reachability
Single PoP Availability	High (>99.99..%)	Medium (~95%)	No guarantee for each single PoP
Overall Availability	High	Higher	GSLB guarantees overall availability
Number of Services	Too many	Caching, Security	

How Should Edge Computing Look Like ?

- Edge Computing should NOT be

  “Cloud Computing with lots of PoPs”

- Achieve high availability through GSLB
- The application should be “stateless”

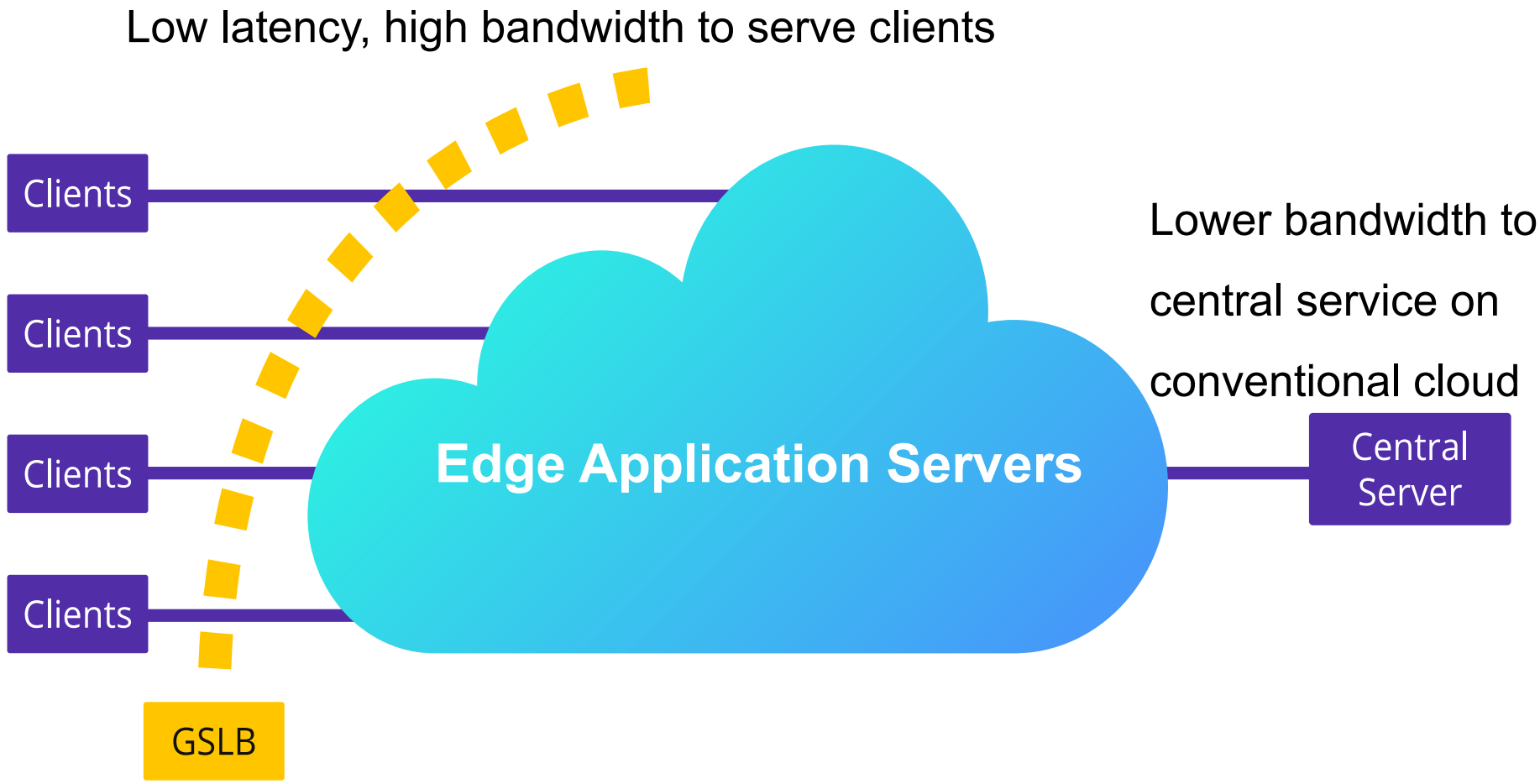
About “Stateless”

- The module running on the Edge should be stateless
- State can be stored on the conventional cloud
- “My application really have to have state on the edge !!”
  - That is fine, but you have to be aware of the consequence.

How to Offer Edge as a Service

- Deploy K8s on every edge PoP, managed by a console.
- The customer (app developer) build the pods in their own K8s.
- Upload the yaml file on the console, distribute to PoPs all over the world.
- Expose the service through a domain name, GSLB will route the end users to the fastest PoP.

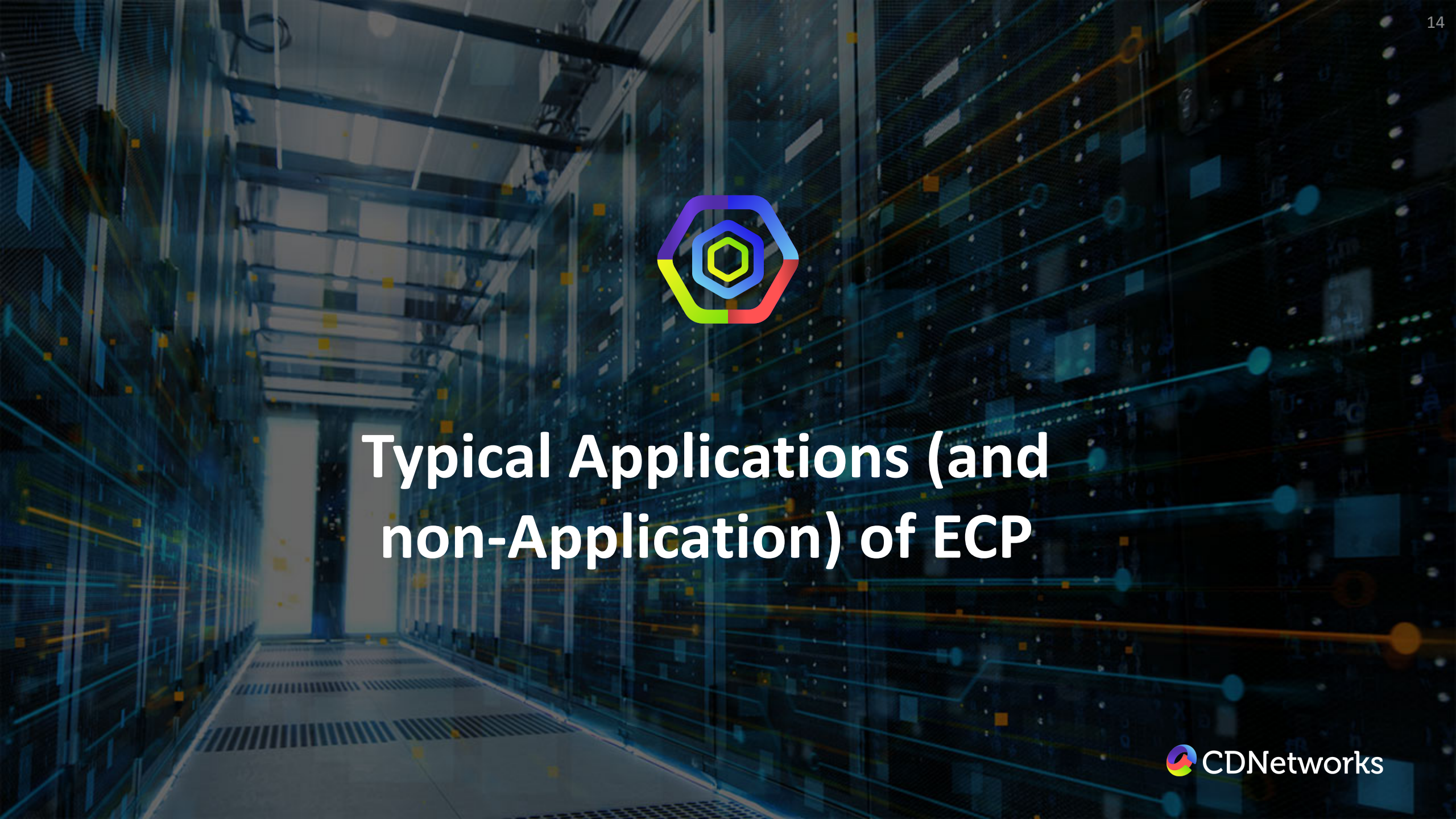
Live in Harmony with Conventional Cloud



CDNetworks

App Architecture for Edge Computing

- Module Partitioning
  - High BW, Latency-Sensitive, Low Computation, Stateless – Edge Computing Platform
  - Low BW, Latency-Insensitive, High Computation, Stateful – Cloud
  - UI, Low Computation, Stateful – Device
- Interaction Between Edge PoPs
  - Through a central service to discover each other, sync data etc.



Typical Applications (and non-Application) of ECP

Cloud (Edge) Gaming

- Module Partitioning
  - Process control signals, 3D rendering, video compression – ECP
  - Save progress, data analysis – Conventional Cloud
  - Video decompression, display, send control signals – Device
- Multi-Player Interactions
  - Only support players on the same or nearby PoPs

Artificial Intelligence (Google Lens, Siri …)

- Module Partitioning
  - Receive image/audio/video, AI inferencing, return results – ECP
  - Machine learning, update the neuro networks – Cloud
  - Collect image/audio/video, display results – Device

Augmented Reality

- Module Partitioning
  - Receive video, AI inferencing, superimpose and compress video – ECP
  - Machine learning, update NN – Conventional Cloud
  - Collect video, display results – Device

Real-time Data Analysis of 1000 Sensors in a Factory

- IoT does not automatically mean ECP should be used
  - For geographically concentrated applications, conventional cloud (either in data center or on-premises) is still a better fit

CDNetworks

Accelerate. Secure. Control.

Thank You!

info@cdnetworks.com | www.cdnetworks.com