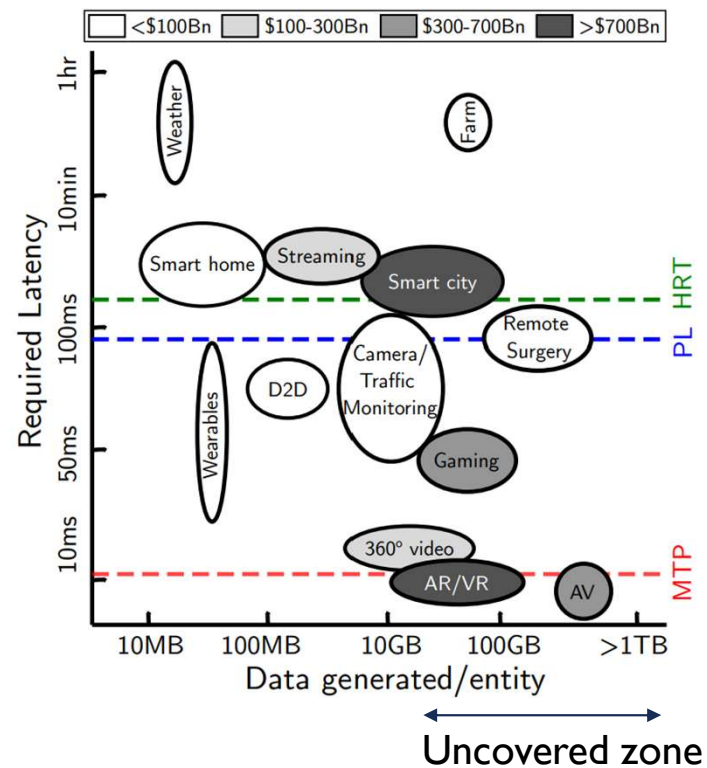




BEYOND CLOUD

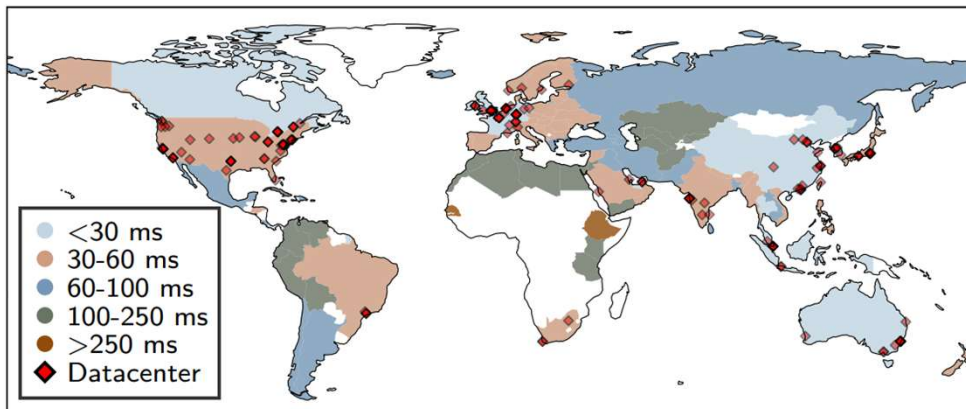


APPLICATION LANDSCAPE

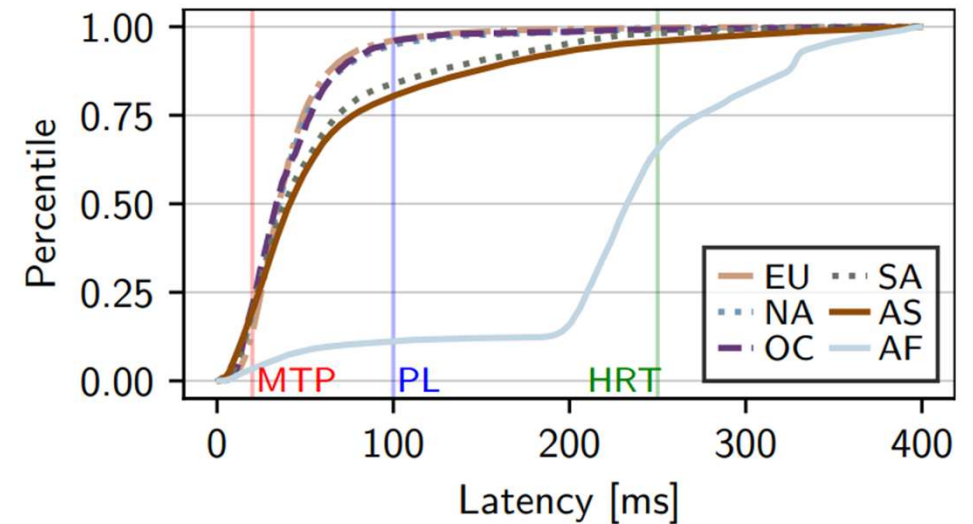


- MTP=Motion-to-Photon
- PL=Perceivable latency
- HRT= Human Response Time

CLOUD DC DISTRIBUTION AND LATENCY



	Datacenters per continent						Backbone N/W
	EU	NA	SA	AS	AF	OC	
Amazon EC2 (AMZN)	6	6	1	6	1	1	Private
Google (GCP)	6	10	1	8	-	1	Private
Microsoft (MSFT)	14	10	1	15	2	4	Private
Digital Ocean (DO)	4	6	-	1	-	-	Semi
Alibaba (BABA)	2	2	-	16	-	1	Semi
Vultr (VLTR)	4	9	-	1	-	1	Public
Linode (LIN)	2	5	-	3	-	1	Public
Amazon Lightsail (LTSL)	4	4	-	4	-	1	Private
Oracle (ORCL)	4	4	1	7	-	2	Private
IBM (IBM)	6	6	-	1	-	-	Semi
Total	52	62	4	62	3	12	



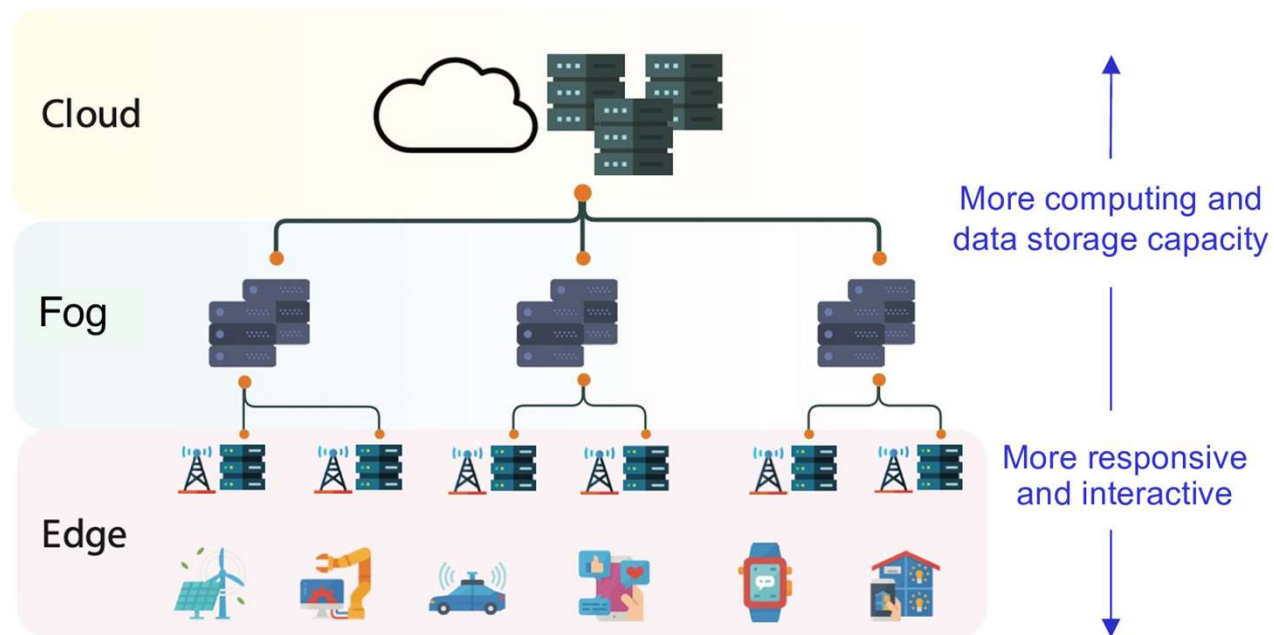
115,000 probes were distributed in over **140 countries** worldwide

Cloudy with a chance of short RTTs: analyzing cloud connectivity in the internet, Khang Dang et al. *MC '21*:
 Proceedings of the 21st ACM Internet Measurement Conference

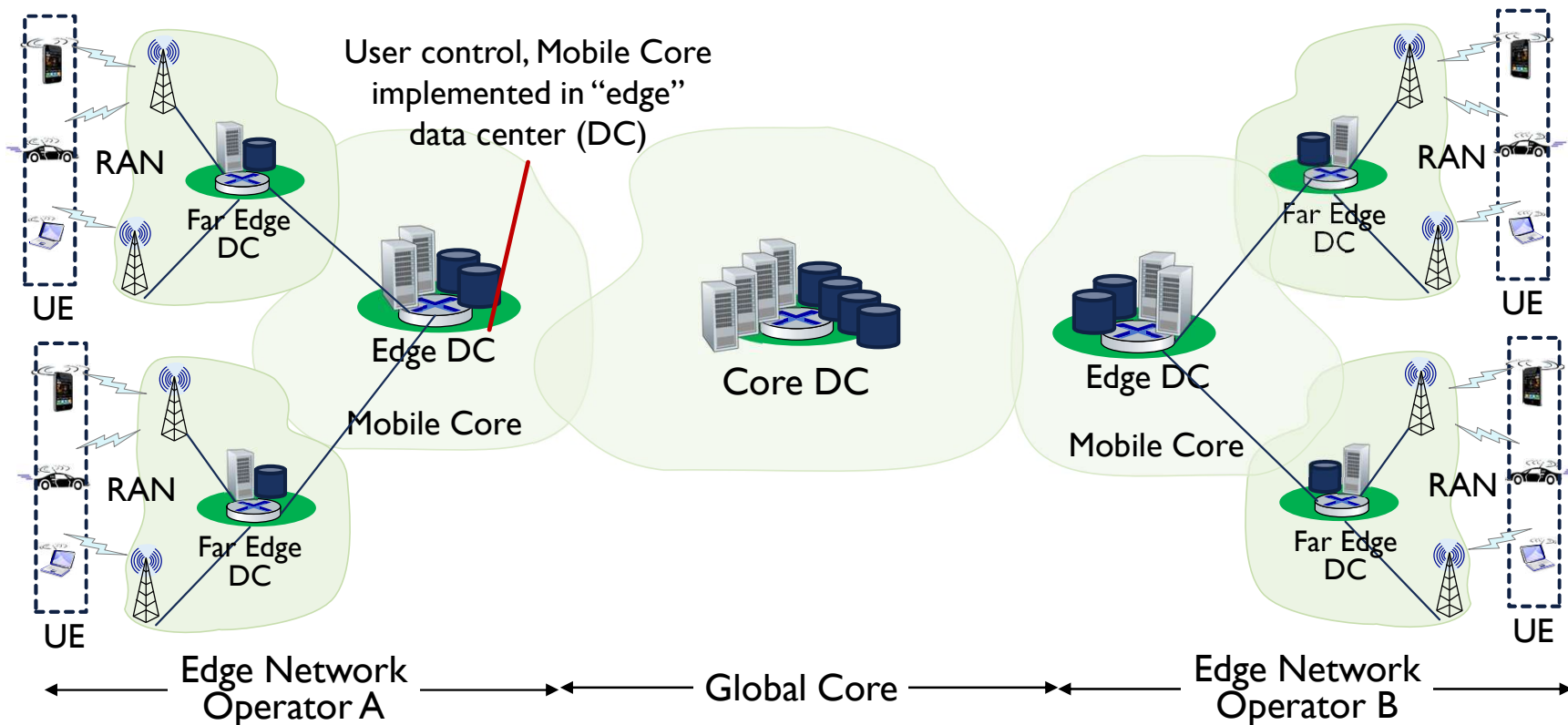
CLOUD LIMITATION

- Relying on cloud for some application can be impractical due to high network latency
- Apps like real-time manufacturing, self-driving vehicles, drones, cognitive assistants generate massive data volumes
 - Challenge with latency-sensitive and location-aware applications
- How to cope with these requirement?
- Move computation and storage to nearby resources located at the edges of Internet

CLOUD-FOG-EDGE COMPUTING



EDGE COMPUTING IN 5G



EDGE COMPUTING IN 5G

Real-time ($< \text{ms}$)

Near real-time ($> \text{ms}$)

Non real-time ($>> \text{ms}$)

