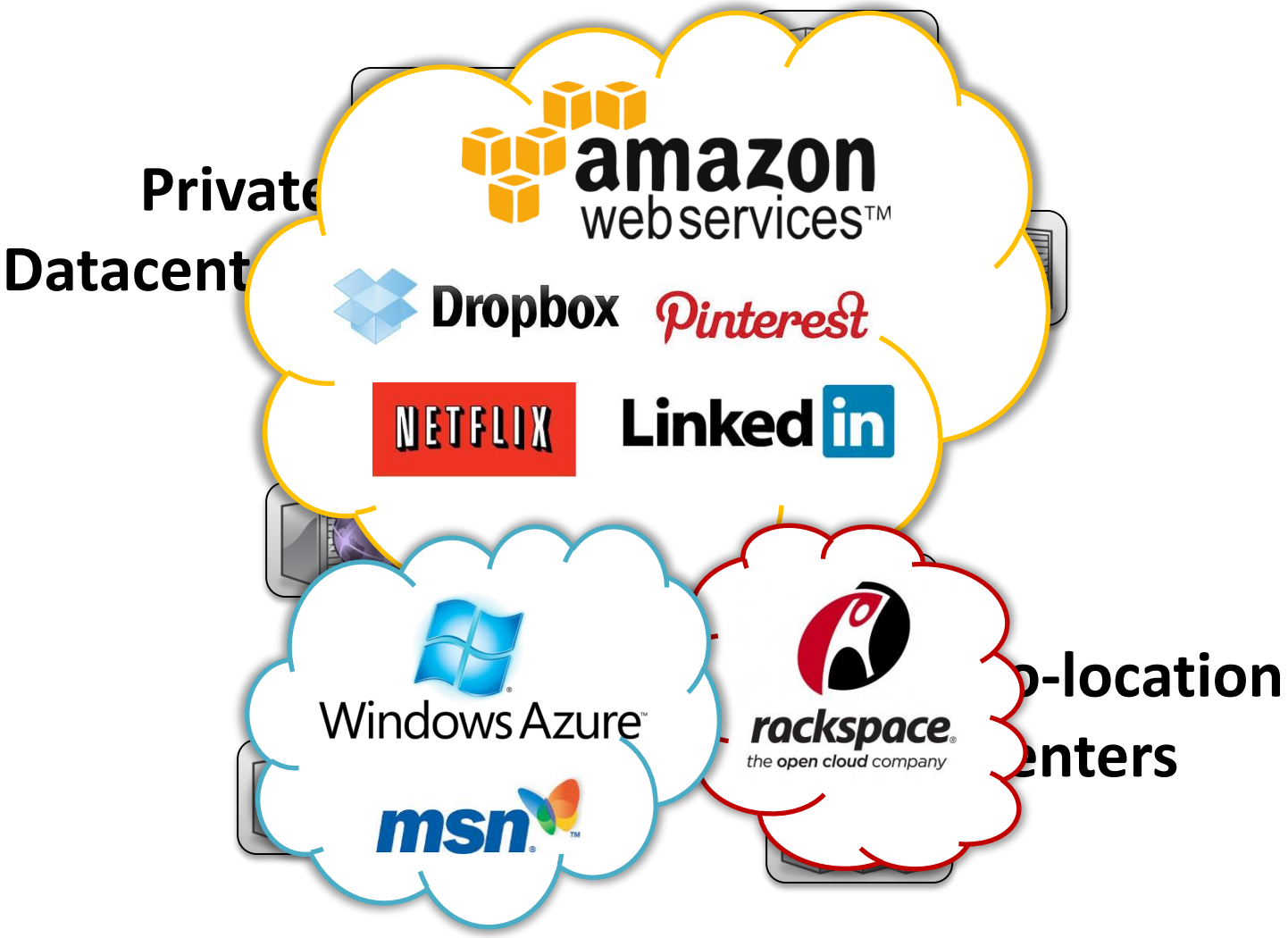
Next Stop, the Cloud: Understanding Modern Web Service Deployment in EC2 and Azure

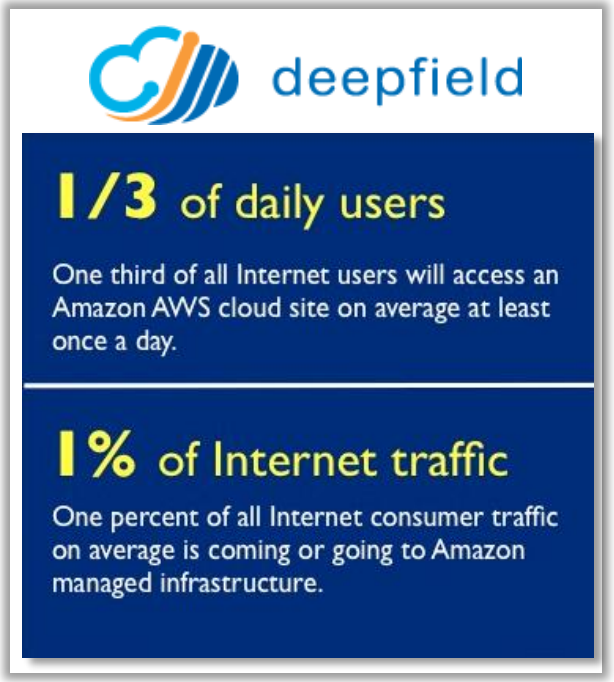
Keqiang He, Alexis Fisher, Liang Wang,
Aaron Gember, Aditya Akella, Thomas Ristenpart

University of Wisconsin-Madison

Move to IaaS clouds



Incomplete view of cloud use



Want a global, in-depth understanding of IaaS cloud usage patterns

Fundamental questions

1. ***Who is using public clouds?***

   → Percentage of cloud-using domains, traffic profile

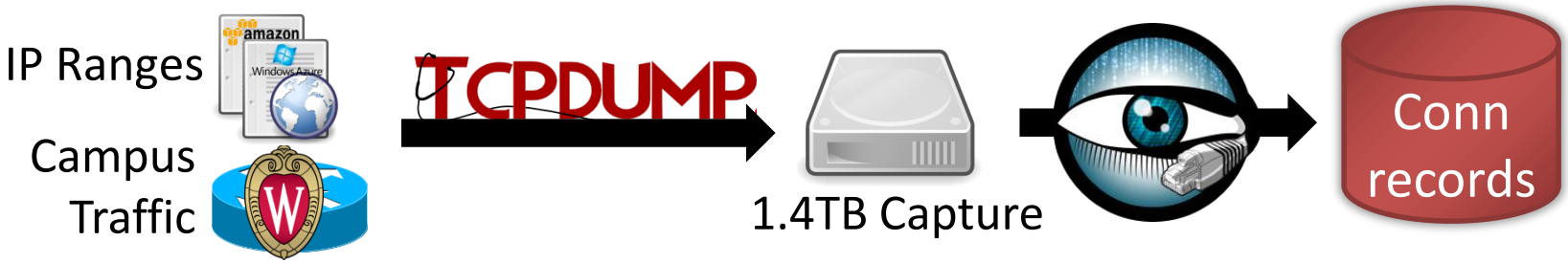
2. ***How are Web services using the cloud?***

   → Impact of failures; ways to improve availability

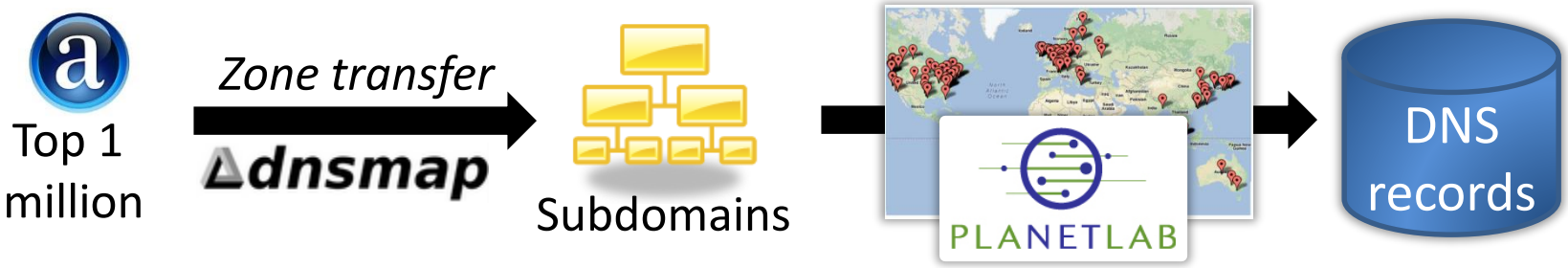
   → Design of new systems/services

Datasets

- University packet capture
  - Deep, but local perspective



- Alexa subdomains DNS records
  - Broad, high-level perspective

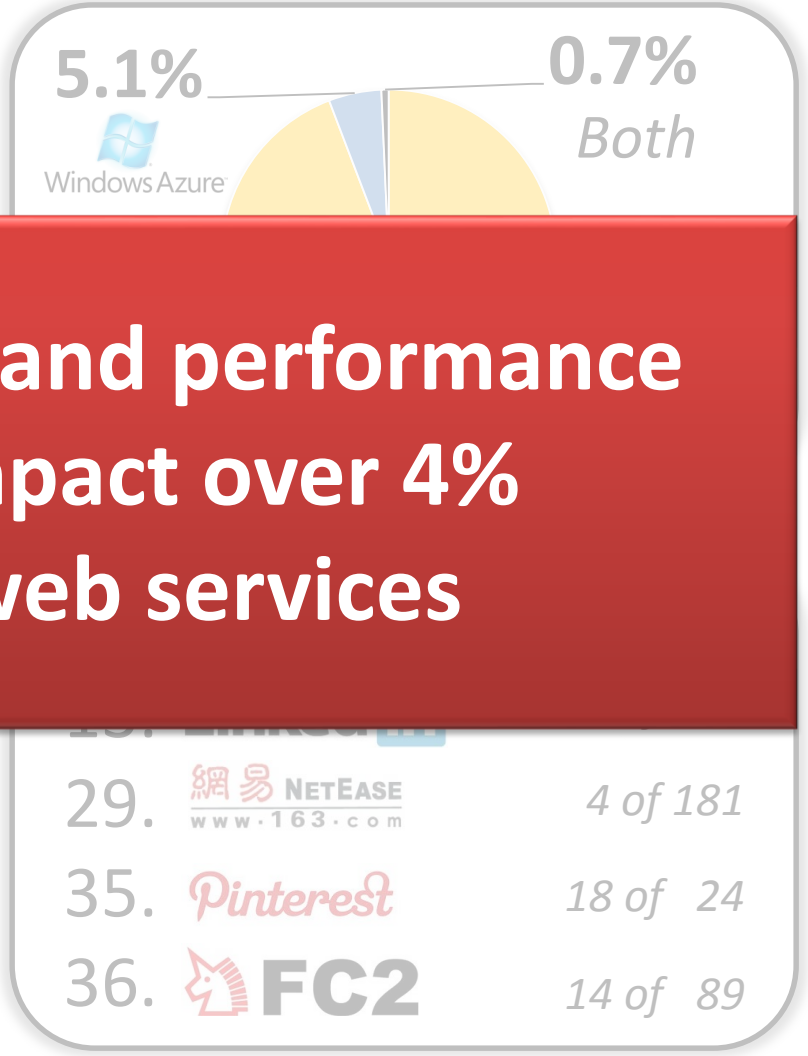


Fundamental questions

1. ***Who is using public clouds?***

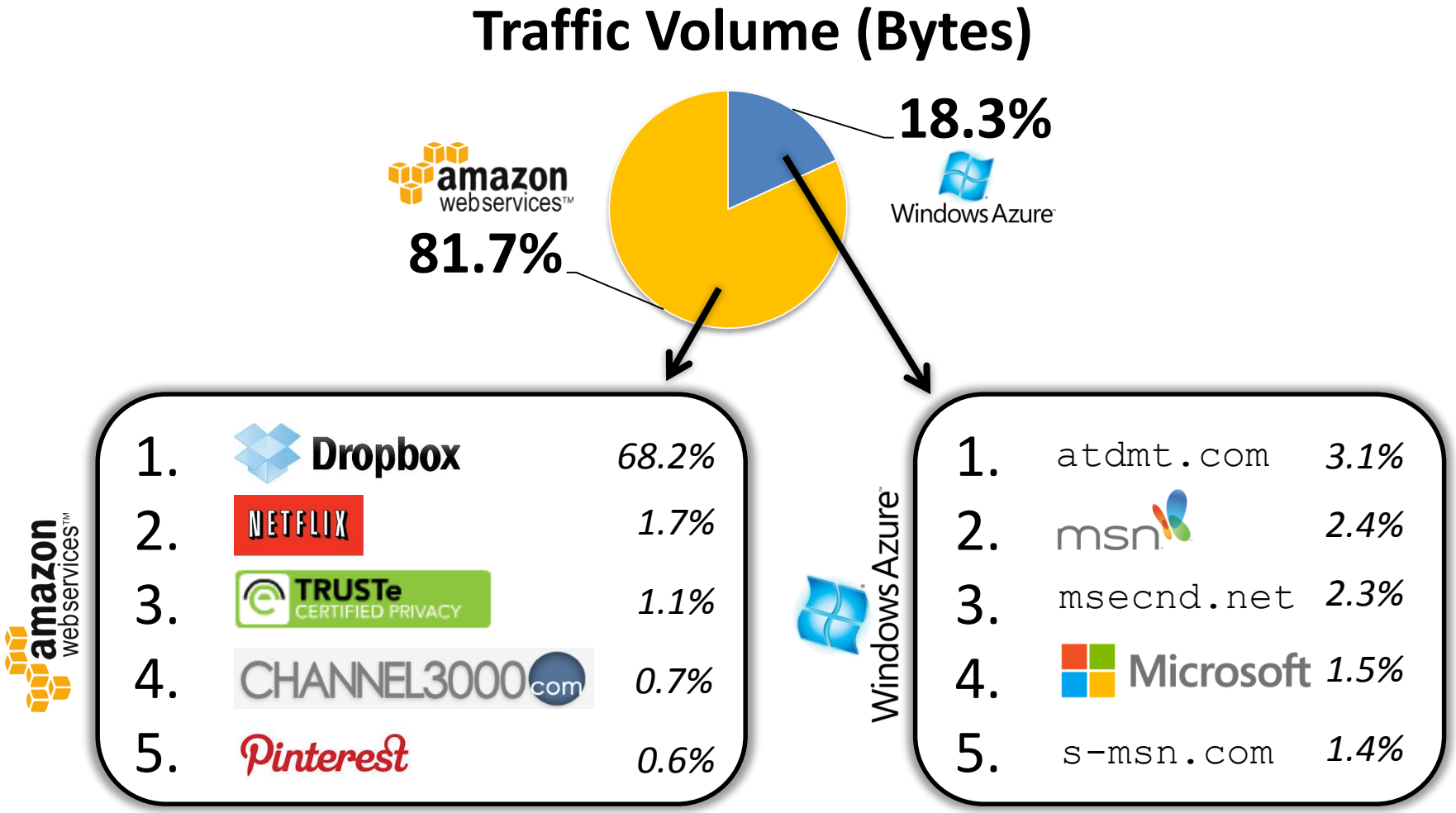
   → Percentage of cloud-using domains, traffic profile

How many popular web services use the cloud?



Cloud availability and performance issues may impact over 4% of popular web services

Which (locally) popular web services use the cloud?



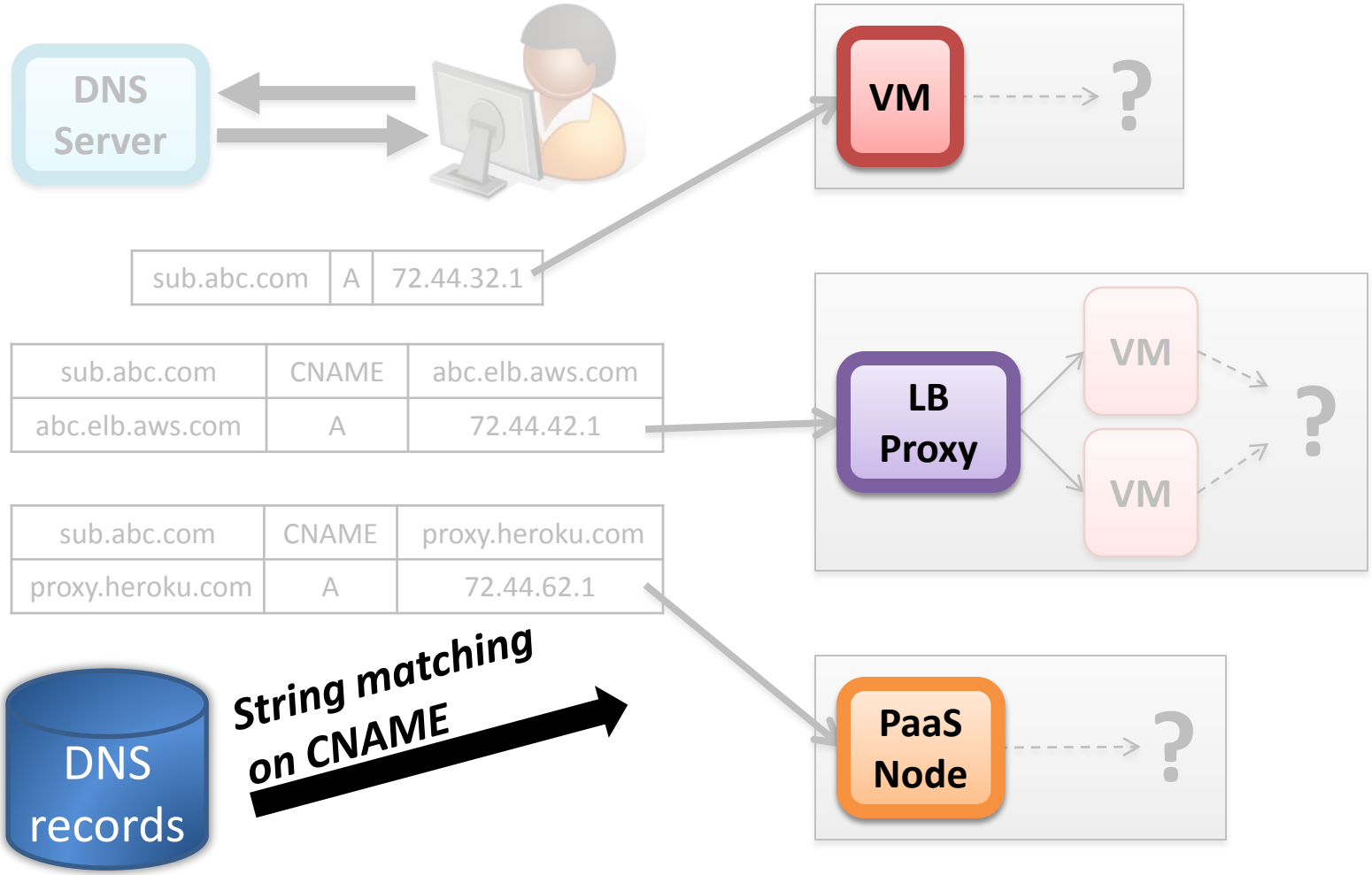
Fundamental questions

2. How are Web services using the cloud?

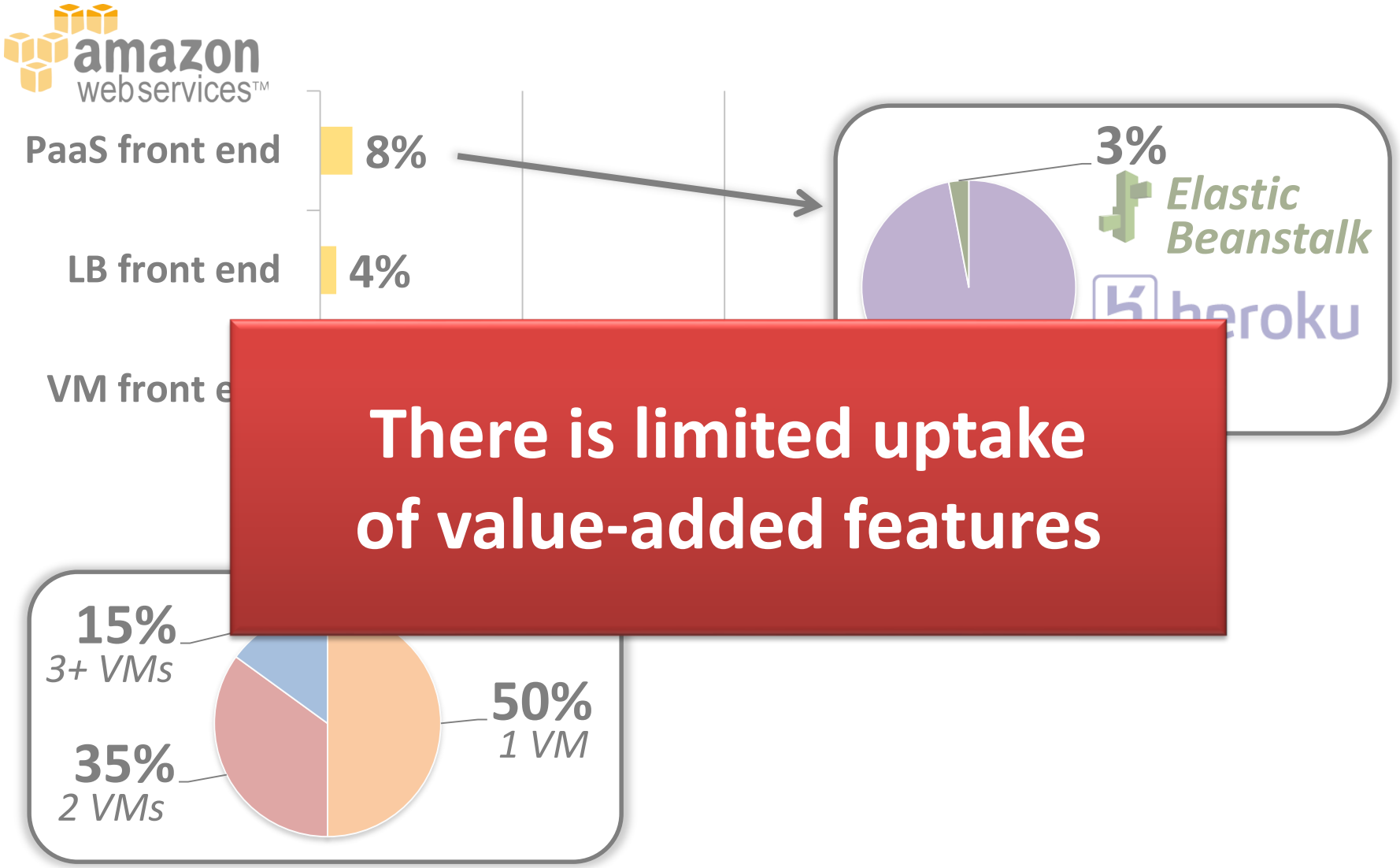
→ Impact of failures; ways to improve availability

→ Design of new systems/services

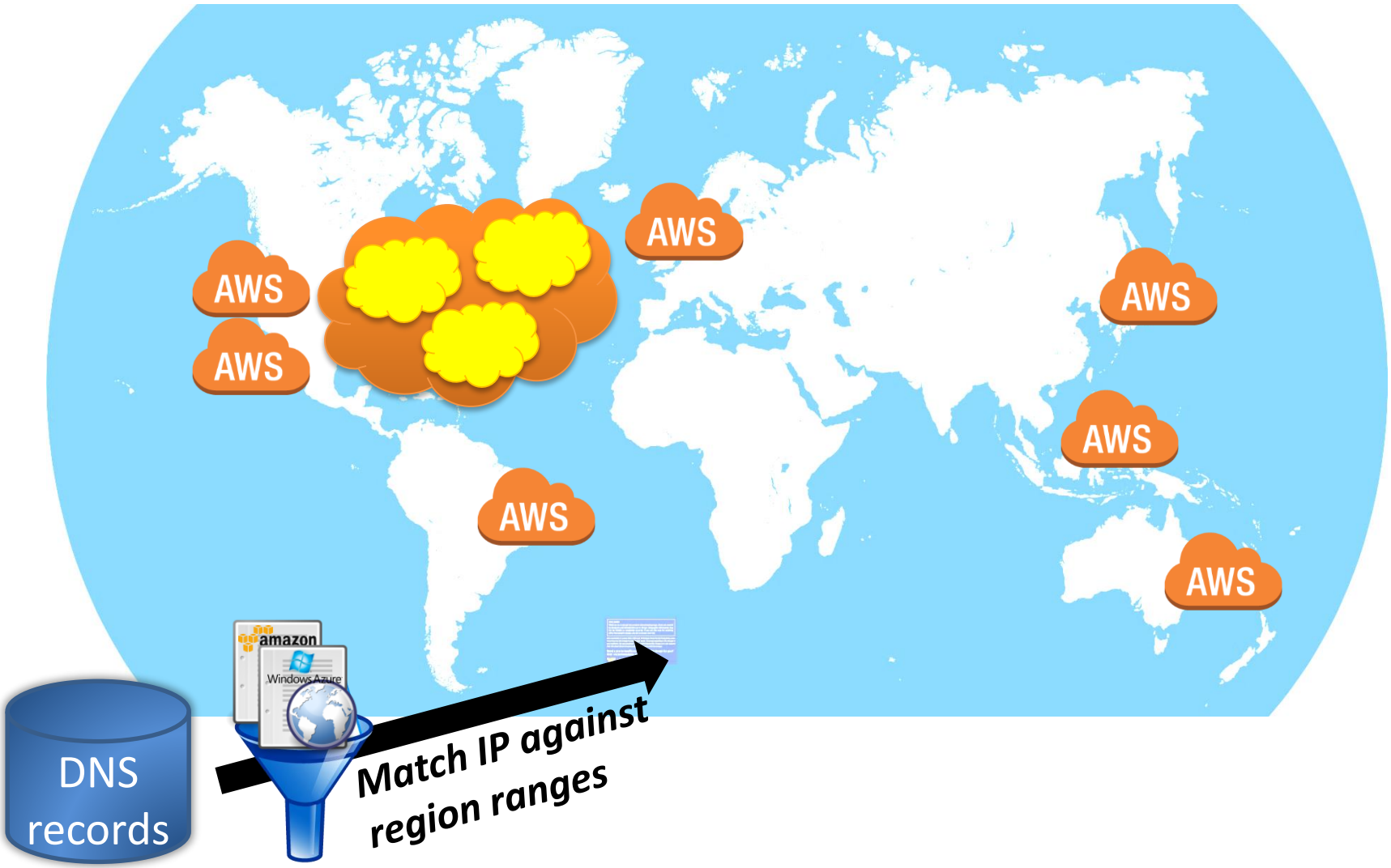
Which front ends are used?



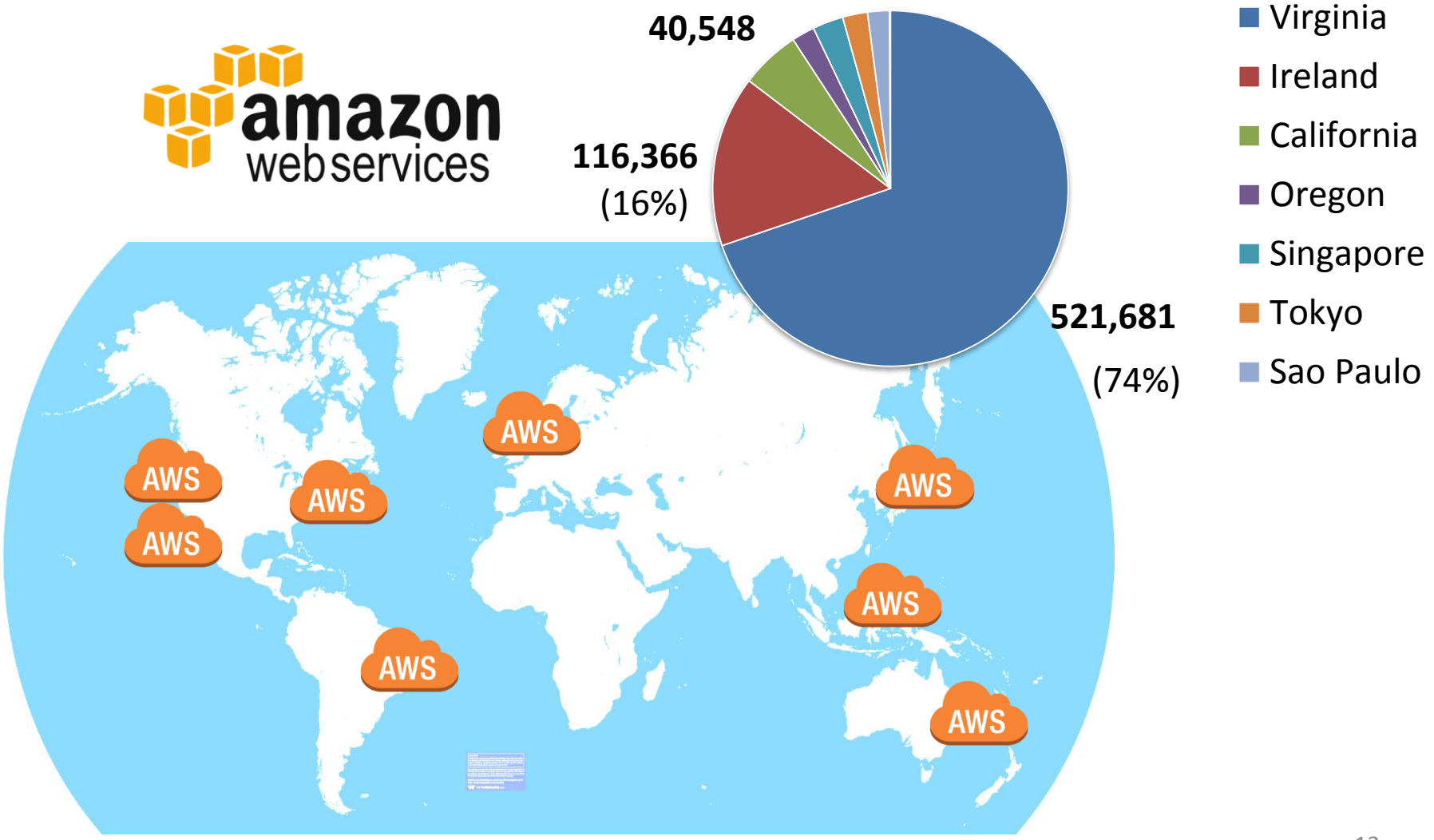
Which front ends are used?



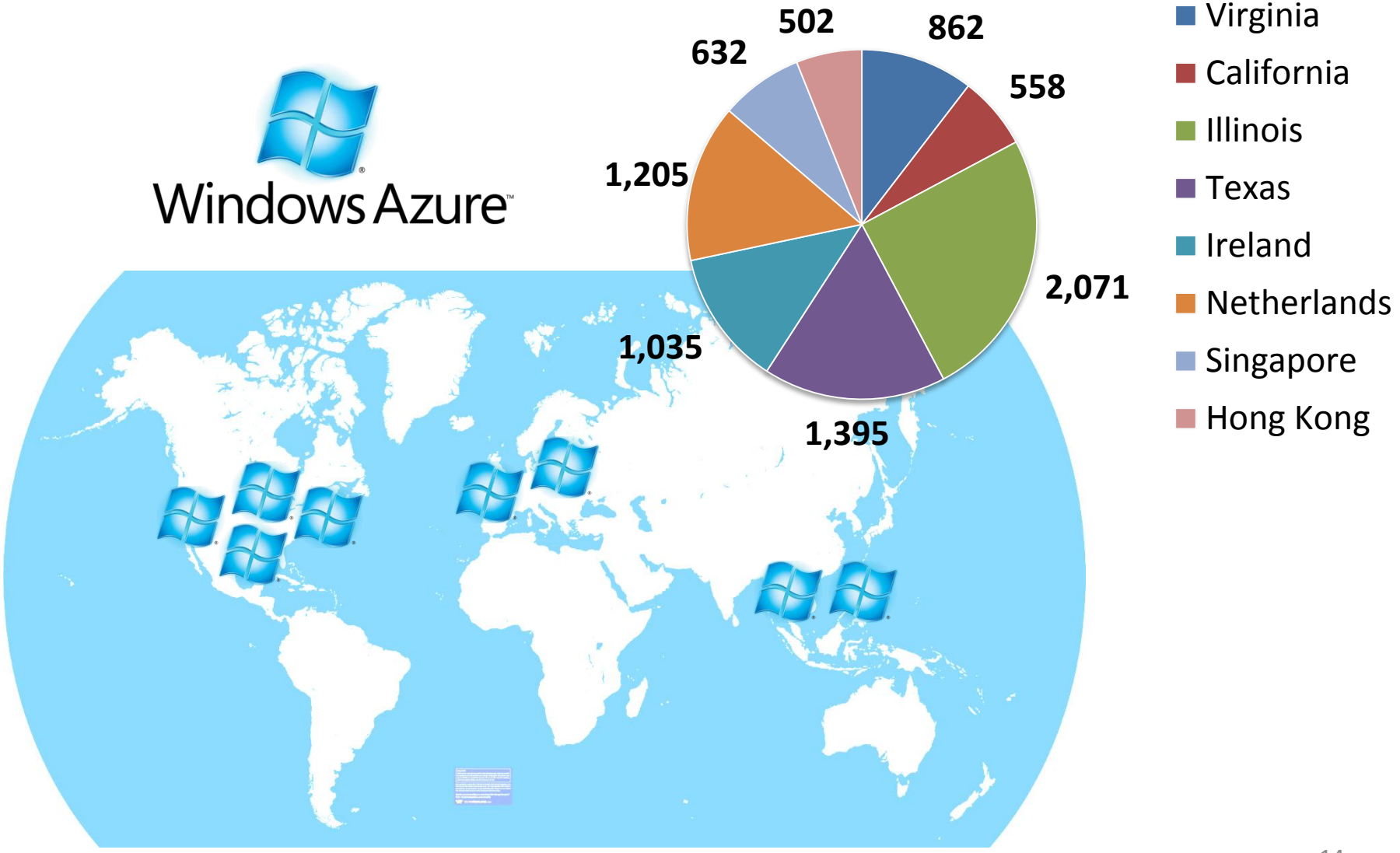
How many and which regions are used?



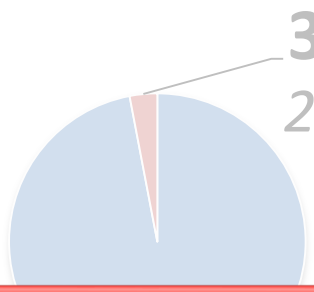
How many and which regions are used?



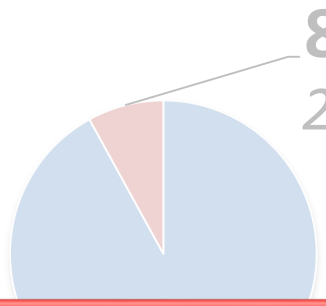
How many and which regions are used?



How many and which regions are used?



amazon web services™



Windows Azure

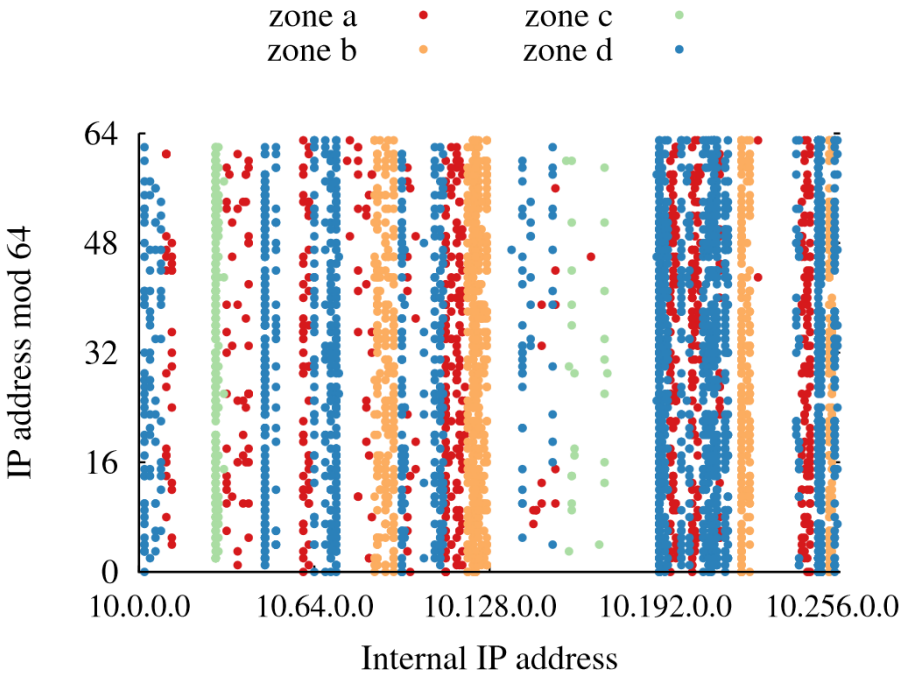
Single region failures can take down a large fraction of cloud-using subdomains

29. NETEASE www.163.com 1
35. Pinterest 1
36. FC2 2

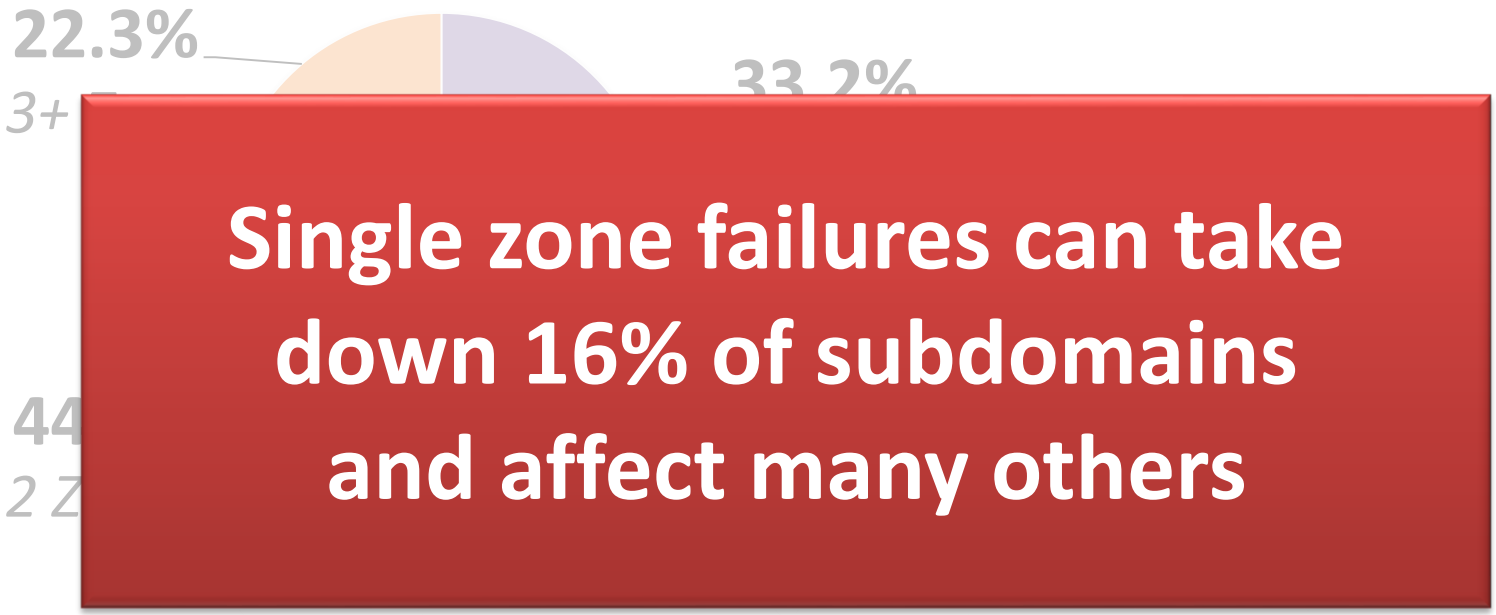
20. bing 1
31. Microsoft 5

How many and which availability zones are used?

- Latency measurements
  - Low latency => instances are in same zone
  - High coverage, low accuracy (noisy)
- IP address correlation
  - IPs are in the same /16 subnet => instances are in the same zone
  - Higher accuracy, low coverage



How many and which availability zones are used?



Single zone failures can take down 16% of subdomains and affect many others

Highlights

- The first comprehensive study of Web service deployment in public clouds
- 4% of popular web services use EC2/Azure
- VMs are the most popular EC2 front end, but value-added features are used by top domains
- Limited region and zone usage (and diversity) makes web services vulnerable to failures