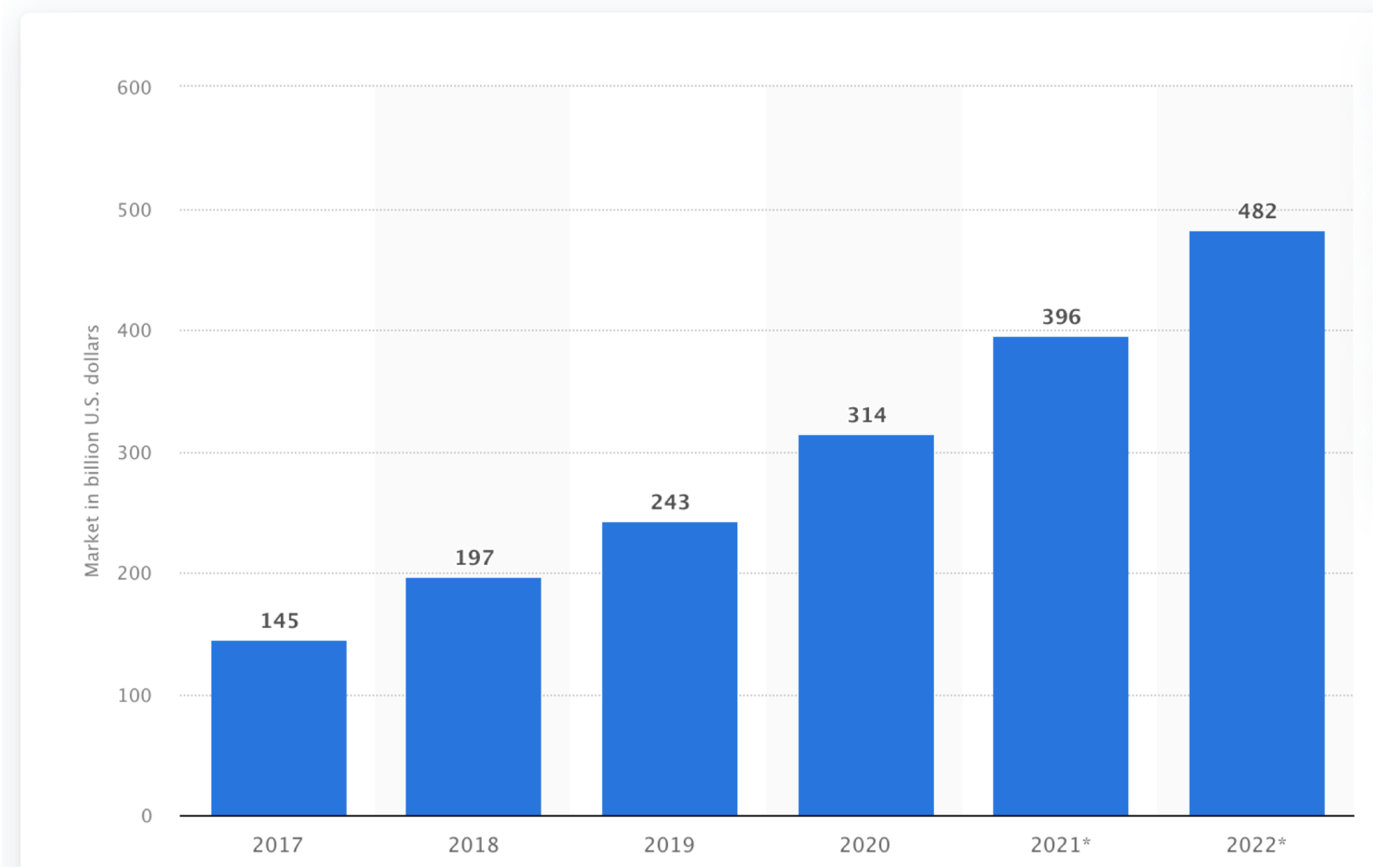
CS 5472 - Advanced Topics in Computer Security

Topic 2: Security in Cloud Computing (1)

Spring 2023 Semester
Instructor: Bo Chen
bchen@mtu.edu
https://cs.mtu.edu/~bchen
https://snp.cs.mtu.edu

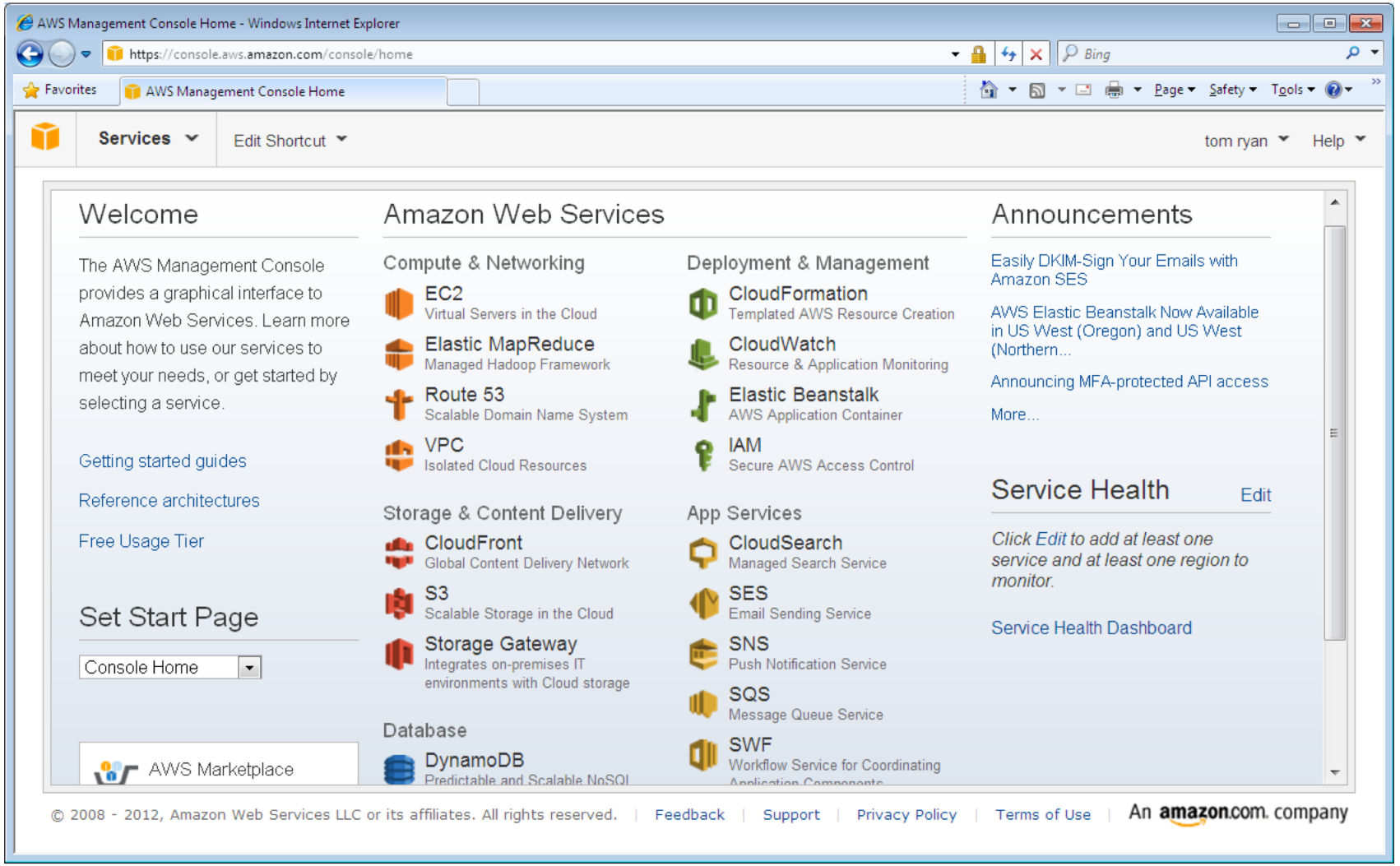
Clouds are Everywhere Today



Public cloud computing market worldwide 2017-2022

Major cloud service providers

Amazon AWS Cloud

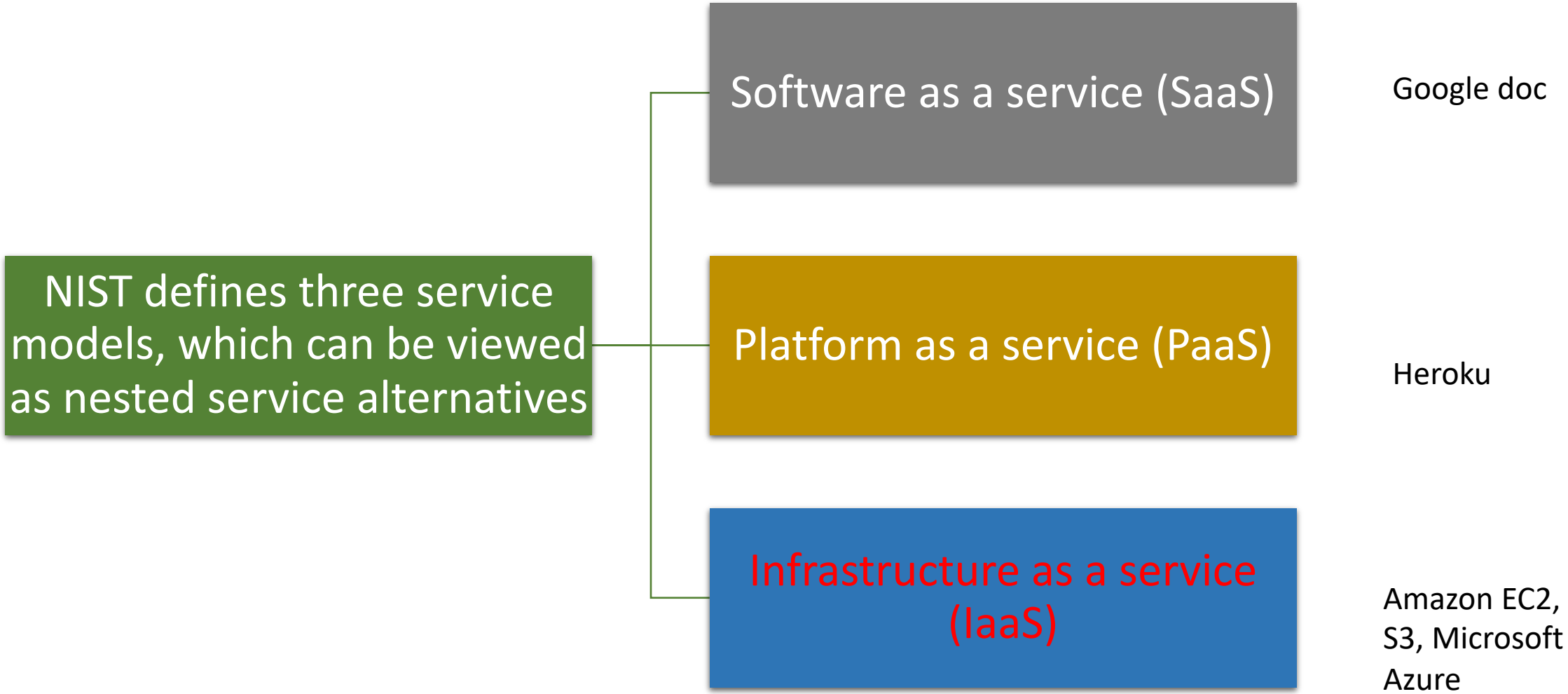


What is Cloud Computing?

A model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models."

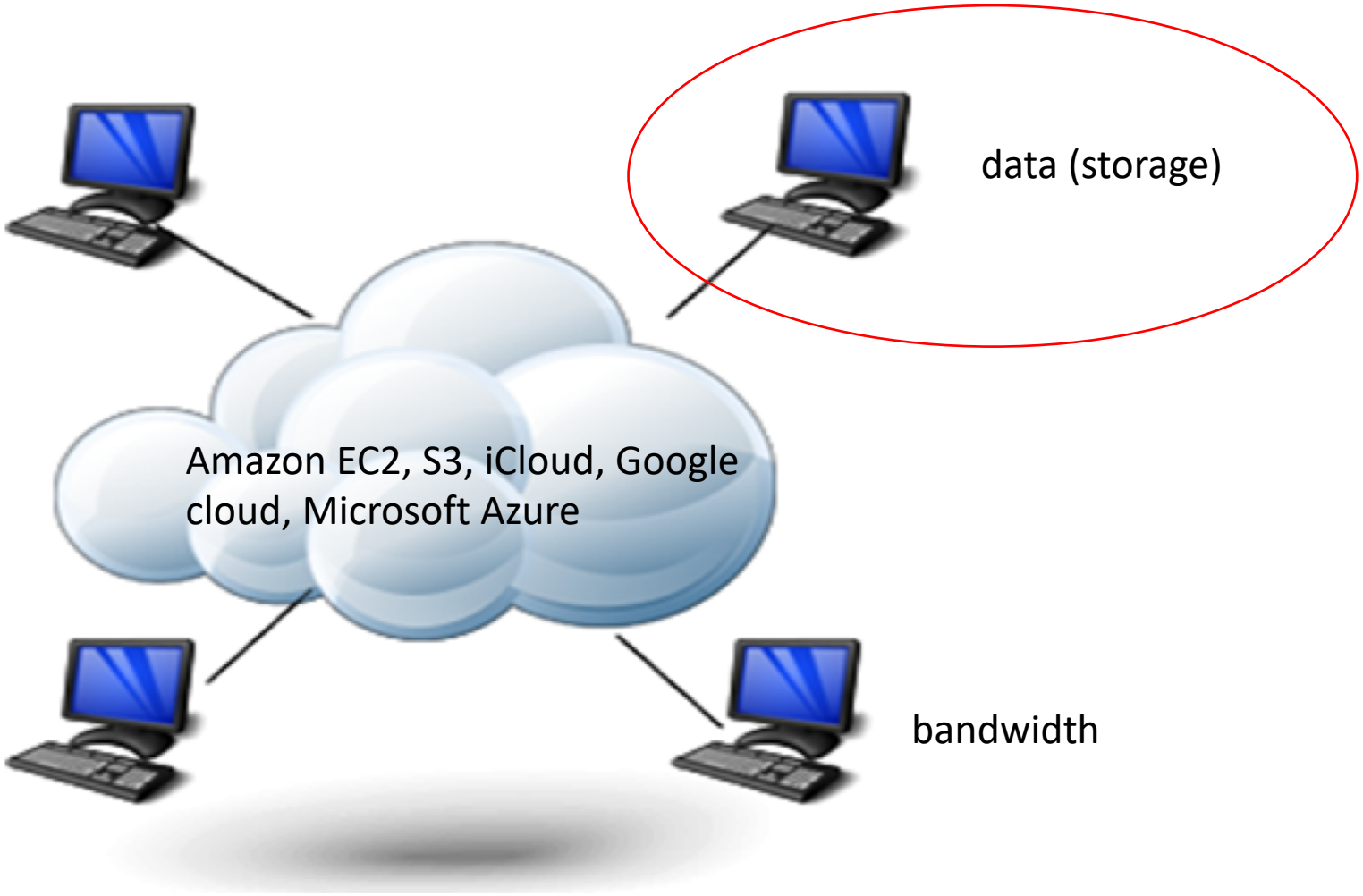
--NIST SP-800-145

Cloud Service Models



Cloud's Basic Model - Outsourcing

computing (virtual machines, instances)



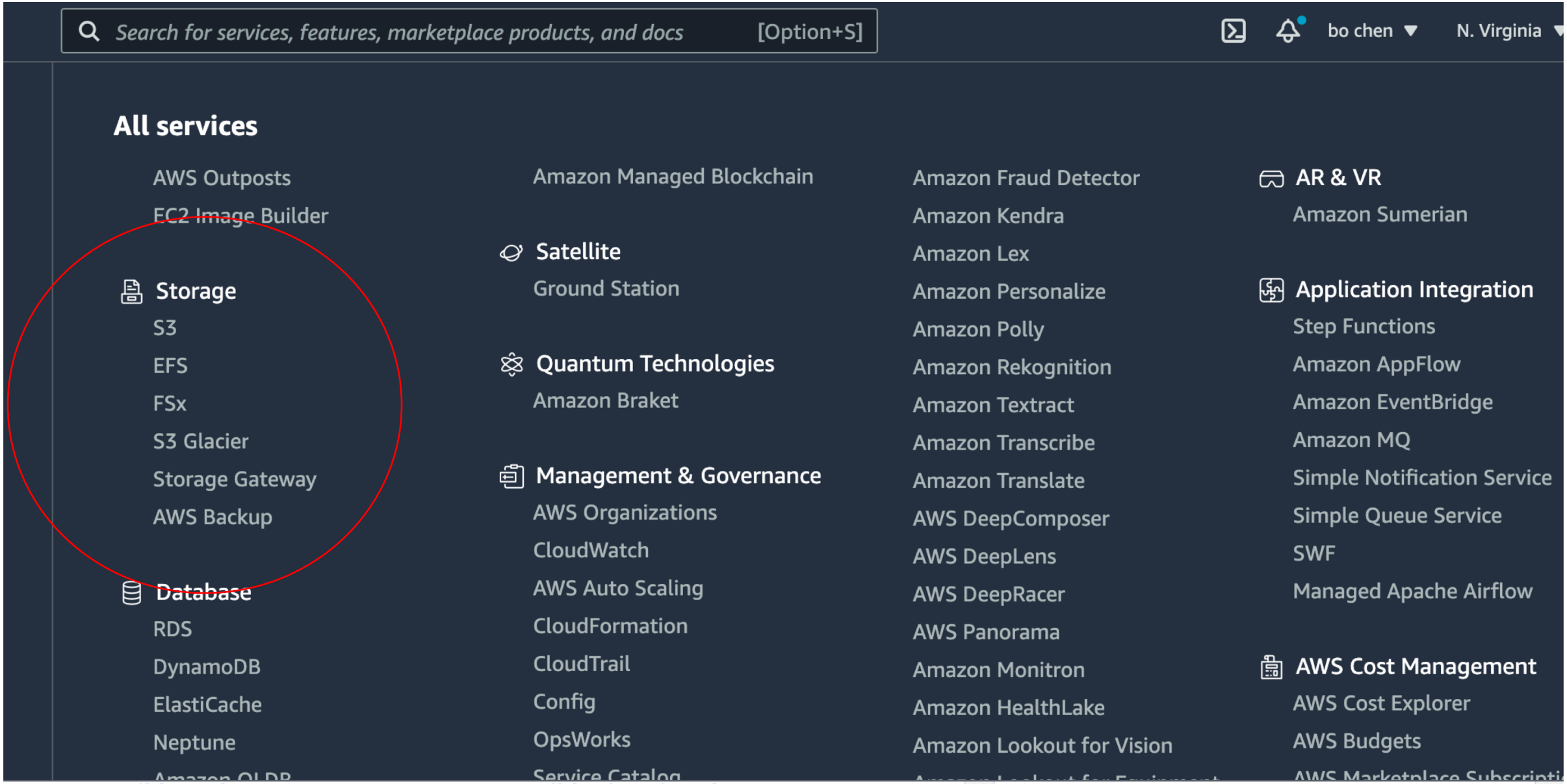
Infrastructure as a service

Cloud Storage Outsourcing

- Very popular
  - Dropbox, Google Drive, Microsoft OneDrive, Box, iCloud, Amazon S3, …

- Very useful and convenient
  - All the data can be stored remotely
  - Access them as you want in any devices
  - No need to maintain a large local storage
    - Good for mobile devices, IoT devices

A Cloud Storage Provider - Amazon AWS Storage



Traditional Cloud Storage Is Fully Centralized

- The cloud storage provider (CSP) creates, manages, and maintains dedicated IT infrastructures/data centers
  - Users outsource their data to the CSPs' data centers

https://aws.amazon.com/about-aws/global-infrastructure/

Traditional Cloud Storage Is Fully Centralized (cont.)

- Pros and cons:
  - Pros:
    - easy deployment, easy management
  - Cons:
    - dedicating computing infrastructure, leading to high cost of creating the cloud and hence high price of cloud usage
    - vulnerable to unexpected instances like power outage, flooding
    - do not scale well for the large number of IoT devices

Transitioning Centralized Cloud Storage to Decentralized Cloud Storage

- Decentralized cloud storage: connect users who need file storage with hosts worldwide offering underutilized hard drive capacity
  - The idea is similar to the sharing economies like Airbnb
  - Users from the network form virtual data centers

Filecoin

STORJ.IO

Siacoin

Internxt

IPFS

Transitioning Centralized Cloud Storage to Decentralized Cloud Storage (cont.)

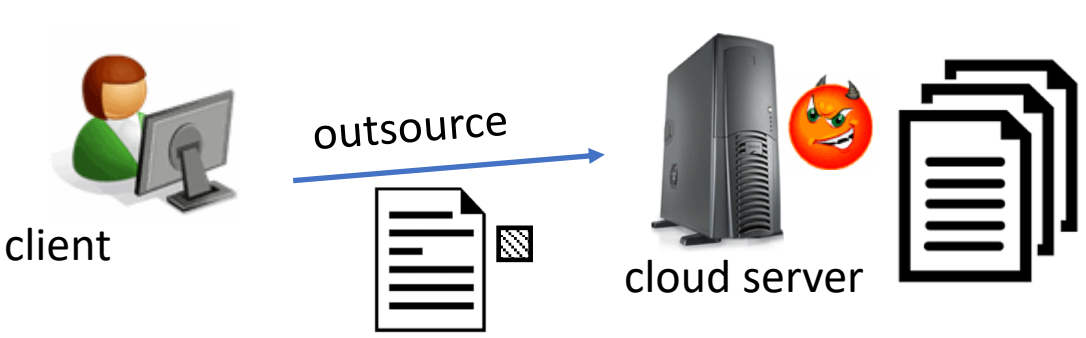
- Benefits:
  - No need to maintain dedicated computing infrastructures, fully utilize the spare disk space from peers. Price is much cheaper
    - Sia cloud ($0.002 per GB per month) vs. Amazon S3 ($0.023 per GB per month)
  - Much more robust by distributing data shares to multiple peers across the globally distributed storage network
  - Can be easily scaled up to support a huge number of computing devices in the coming IoT era
  - Users outsource data to the storage peers nearby, and storing/retrieving data would be much faster

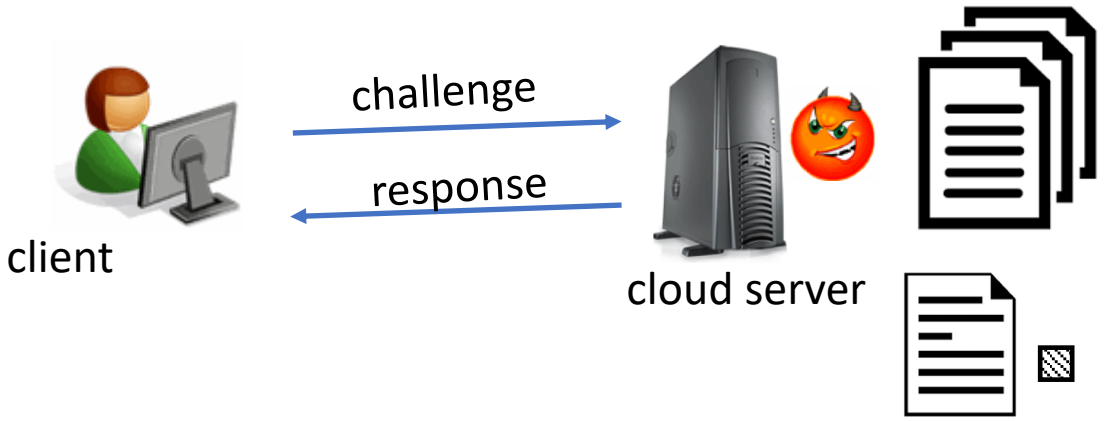
Constructing a Decentralized Cloud Storage Network is Challenging

- How can we **incentivize** the peers to participate
  - Peers (farmers or miners) who will provide storage services
  - Peers (users) who will use storage services
- How to ensure **security** in a purely decentralized storage network in which all peers are untrusted and there is no trusted entity
  - How we ensure the peers will function correctly
  - How to ensure confidentiality of the data stored
  - How to ensure integrity of the data stored
  - How to ensure reliability/replication of the data stored

Proofs of Storage (PoS)

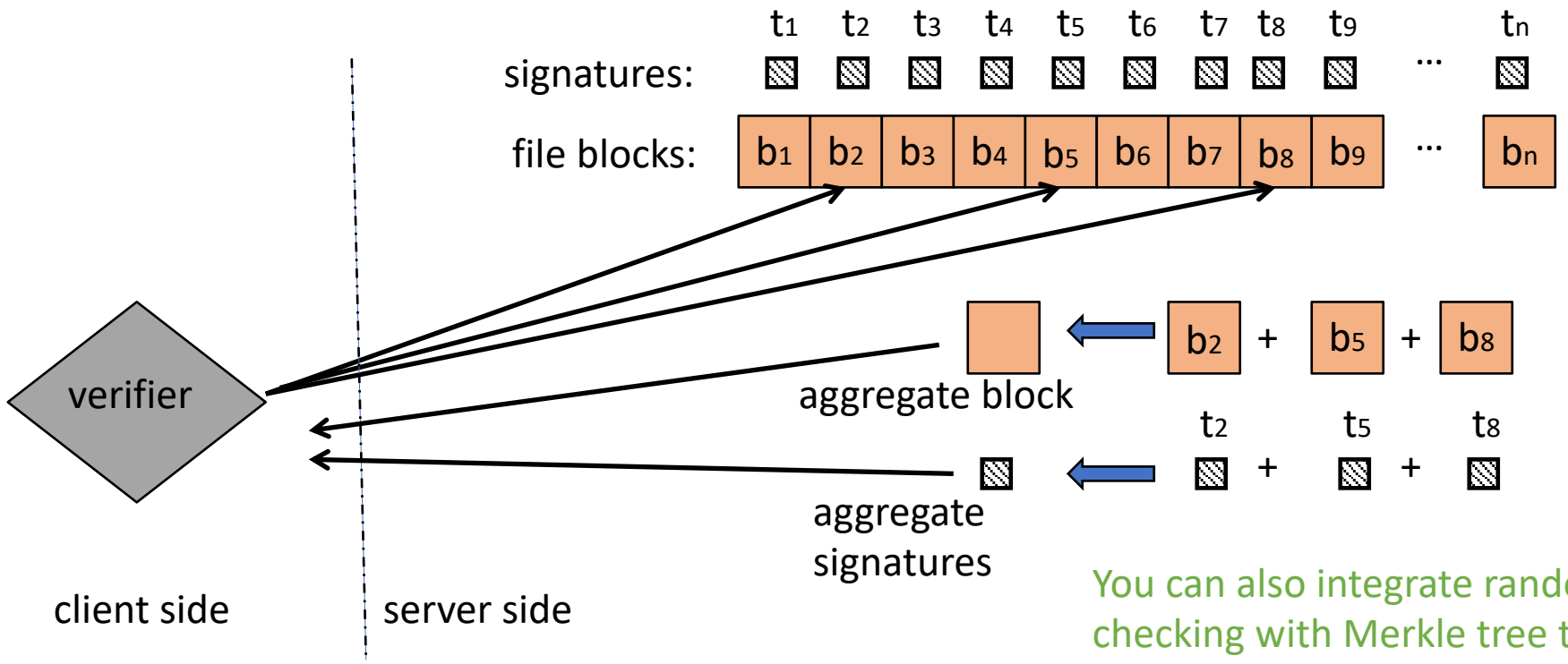
- A top security concern in the cloud storage outsourcing is: how can the data owner obtain proofs that the outsourced data in the cloud are stored correctly (i.e., proofs of storage, or PoS)
  - Provable data possession (PDP)
  - Proofs of Retrievability (PoR)





Proofs of Storage (cont.)

- A random checking technique for efficiency: the client randomly samples a certain number of blocks for checking (random challenge)
  - Rather than check the entire outsourced data



You can also integrate random checking with Merkle tree to support efficient data dynamics

Proof of Replication and Proof of Spacetime

- Proof of replication
  - How can the data owner obtain a guarantee that the outsourced data are indeed stored redundantly in a few different peers
  - The challenge is: even though the cloud storage may claim that 3 copies of the data have been stored, but the storage peers can easily collude and only store 1 copy and it is hard to detect this cheating.
- Proof of spacetime
  - PoS can allow to obtain a proof that the data are stored correctly at the time upon checking, but cannot ensure that the data can be stored correctly for a certain amount of time
  - Proof of spacetime enables this new guarantee

Others

- A current project of the SnP lab is about the security and privacy in decentralized cloud storage.
  - Let me know if you would like to get involved
  - Currently supported by national science foundation
- Any interested students feel free to use this topic for your term project (would be a great project experience)
  - Decentralized cloud storage is projected as the future cloud storage for IoTs and big data
  - Decentralized cloud storage needs to integrate the popular blockchain technologies

Paper Presentation

- Filecoin: A Decentralized Storage Network

- Presented by Trevor Hornsby