NAMED DATA NETWORKING IN CLIMATE RESEARCH AND HEP APPLICATIONS
CHEP2015, APRIL 16 2015, OKINAWA, JAPAN

Susmit Shannigrahi*, Artur Barczyk^, Christos Papadopoulos*, Alex Sim**, Inder Monga**, Harvey Newman^, John Wu**, Edmund Yeh#

*Colorado State University, ** LBNL, ^California Institute of Technology, # Northeastern University

The NDN Project

- Part of the NSF Future Internet Architecture initiative

- One of four multi-institution projects funded in 2010-13, ~$8M

- Now in the second round of funding, 2014-16, ~$6M

- Goal: design the next generation Internet Architecture

NDN Institutions – Past and Present

http://named-data.net • http://github.com/named-data

Today's Internet

- To find content in the network

- ..you have to learn where the content is

- ..and then ask the network to take you there

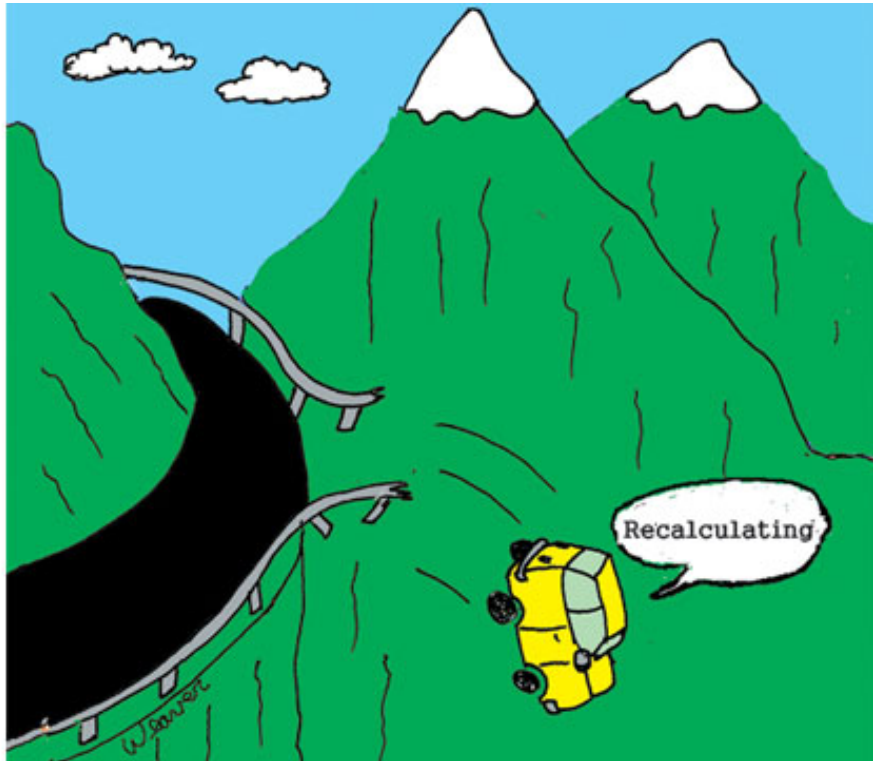- ..so you can tell the server what you want

- In other words, the Internet is like an **old GPS!**

Latitude:26.212424°
Longitude:127.680932°

But Things Often Fail

Often the content is not there..
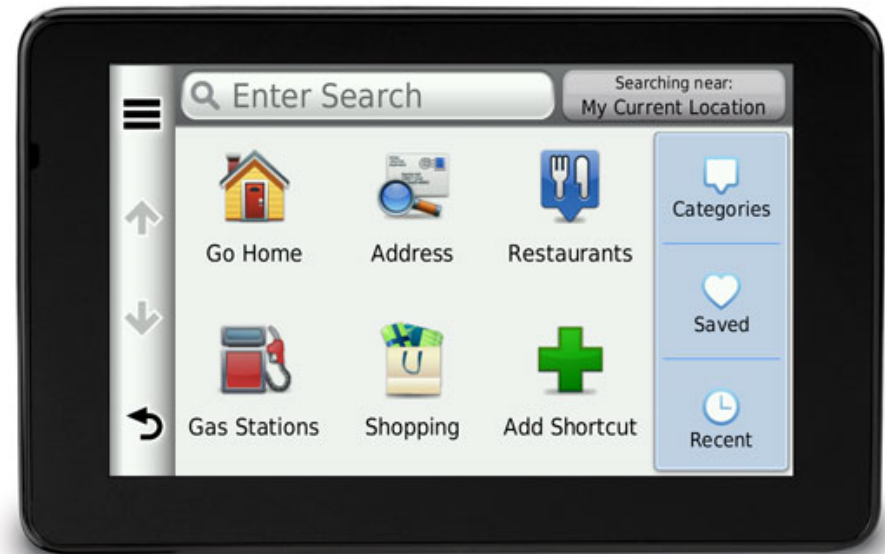
..or the path is too congested!

What if the Network was Smarter

And you could tell it what you want!

How do we make the network smarter?

- Instead of taking you where you **think** the content is..

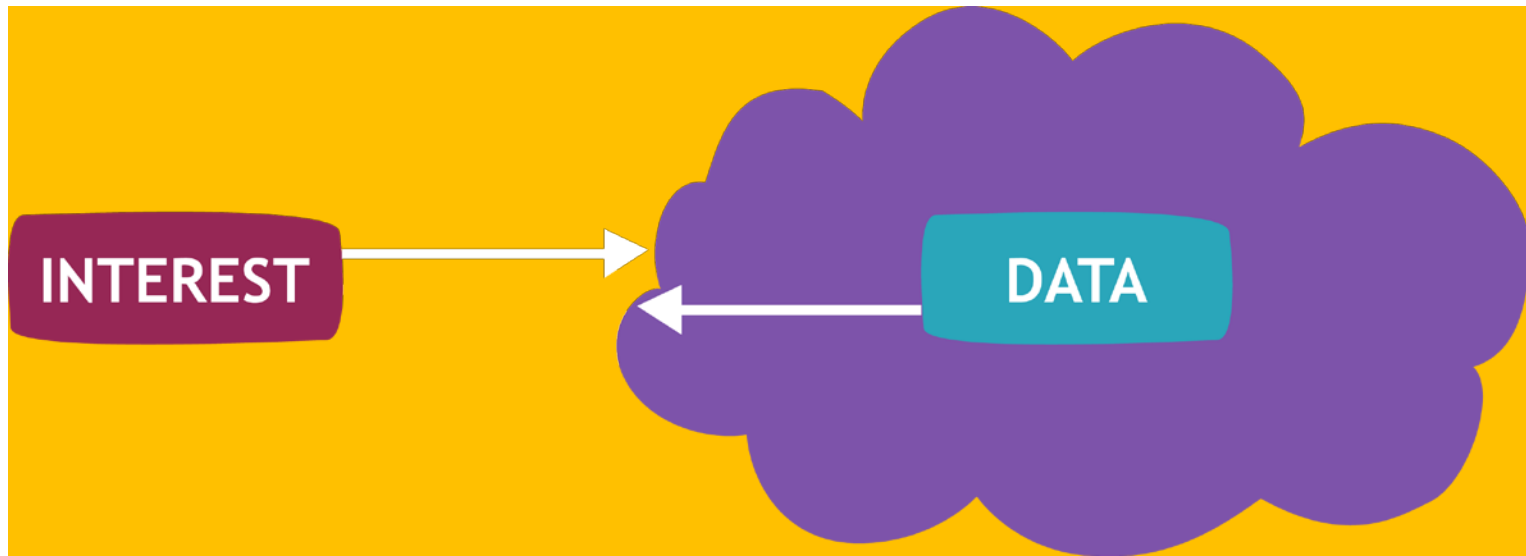- ..the network could **get** the content for you!

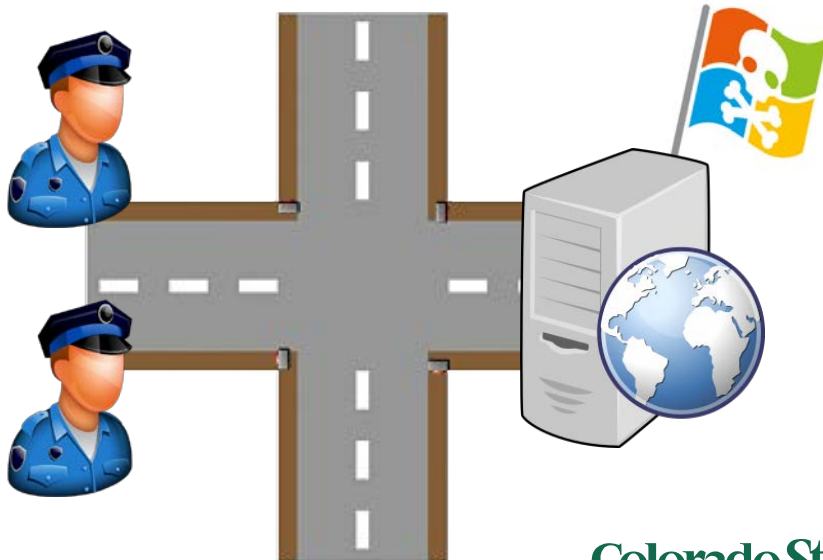- ..in the most efficient way!

Ask what you want by its name, not address

Named Data Network (NDN)

- The new idea: **Name the data, not the hosts!**
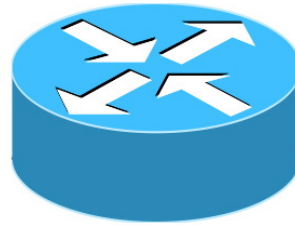- ..so you can ask the network directly for the data!

Named Data is Easy to Secure

- In the Internet you secure your path..

- ..but the server may still be hacked!

- In NDN you **sign** the data with a **digital signature..**

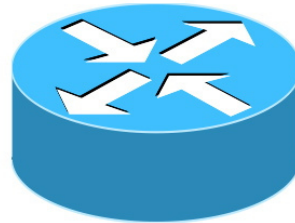- ..so the users know when they get bad data!
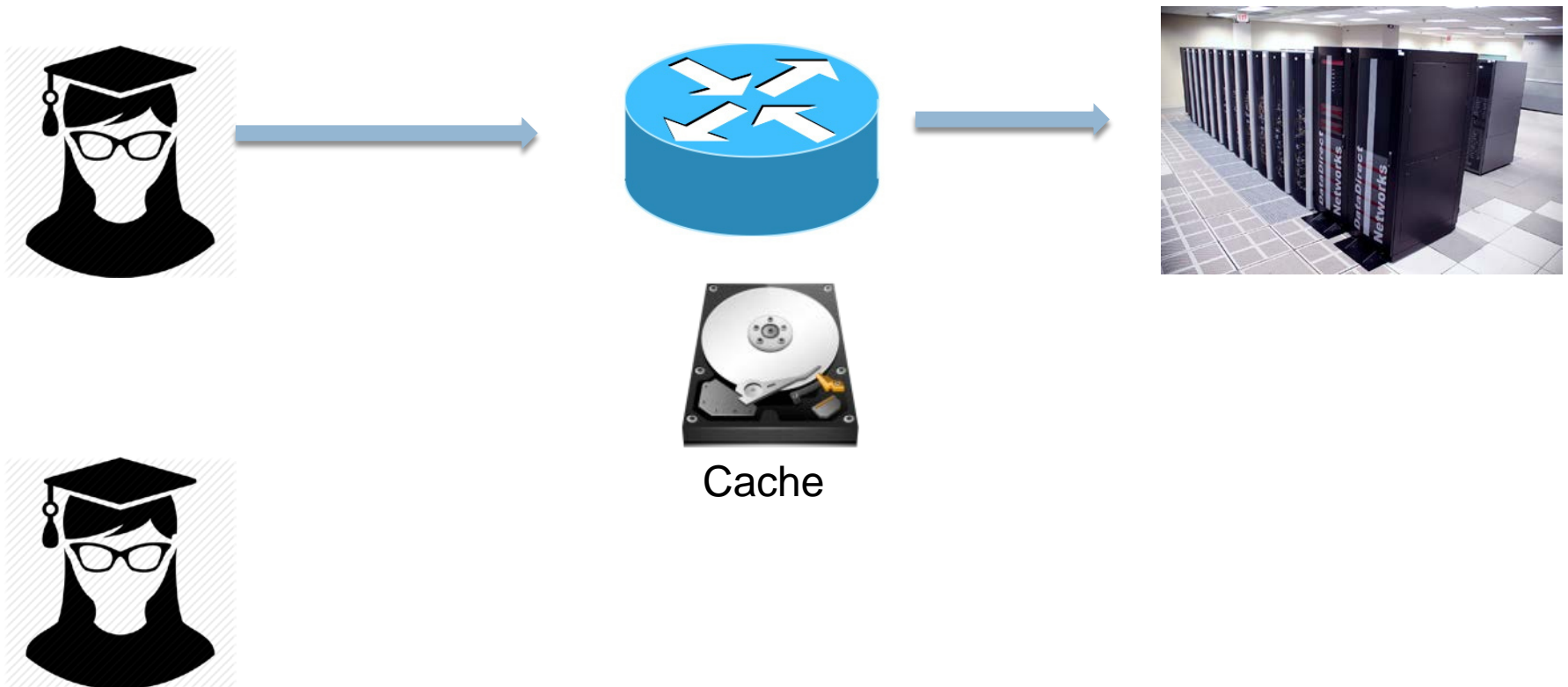
Named Data Can be Cached
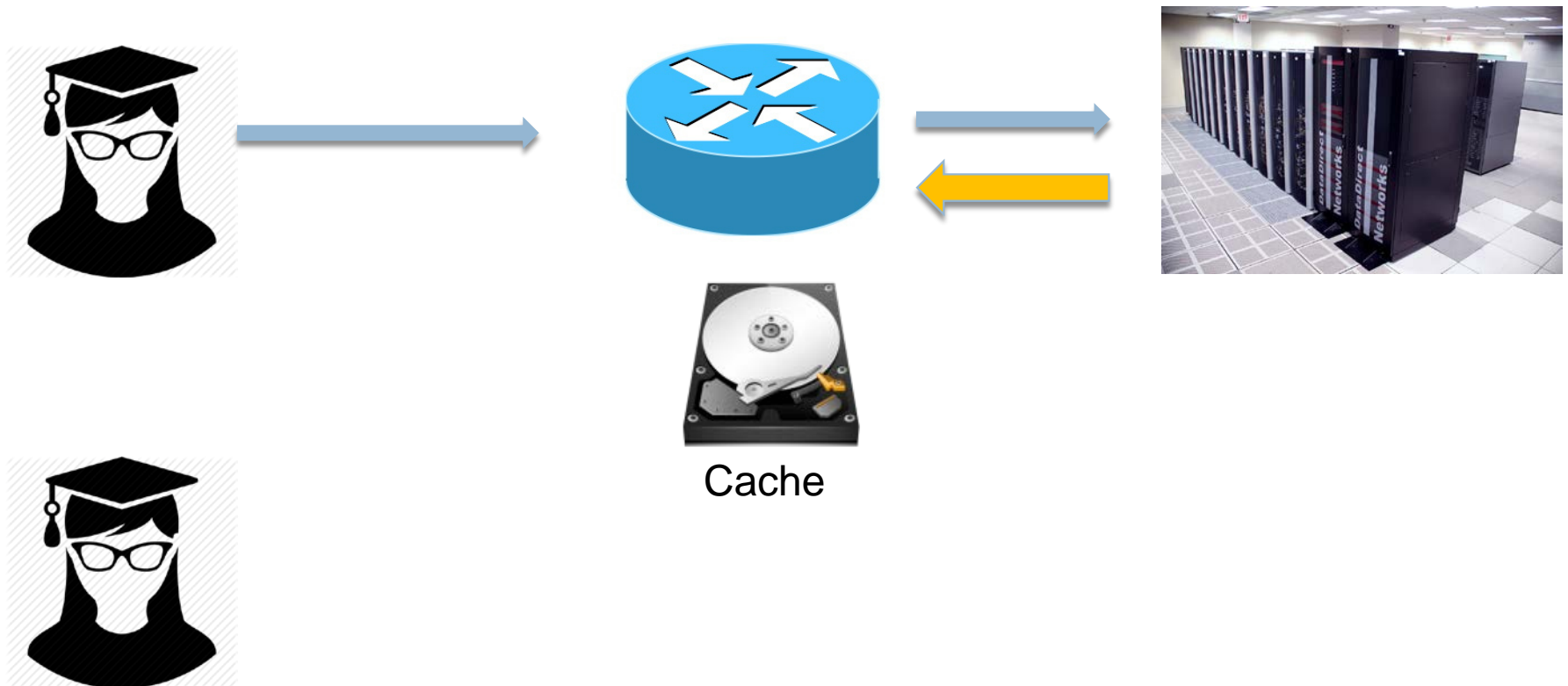

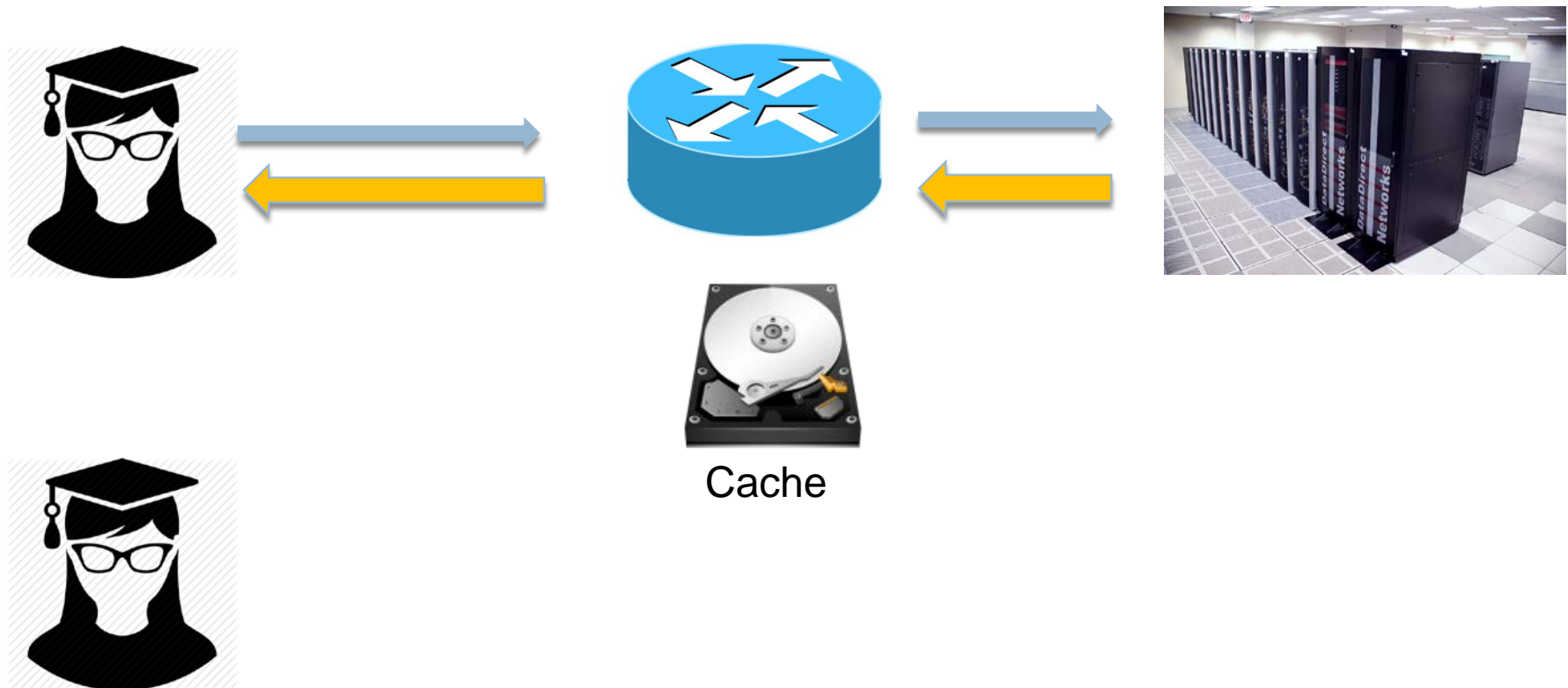Cache

Named Data Can be Cached
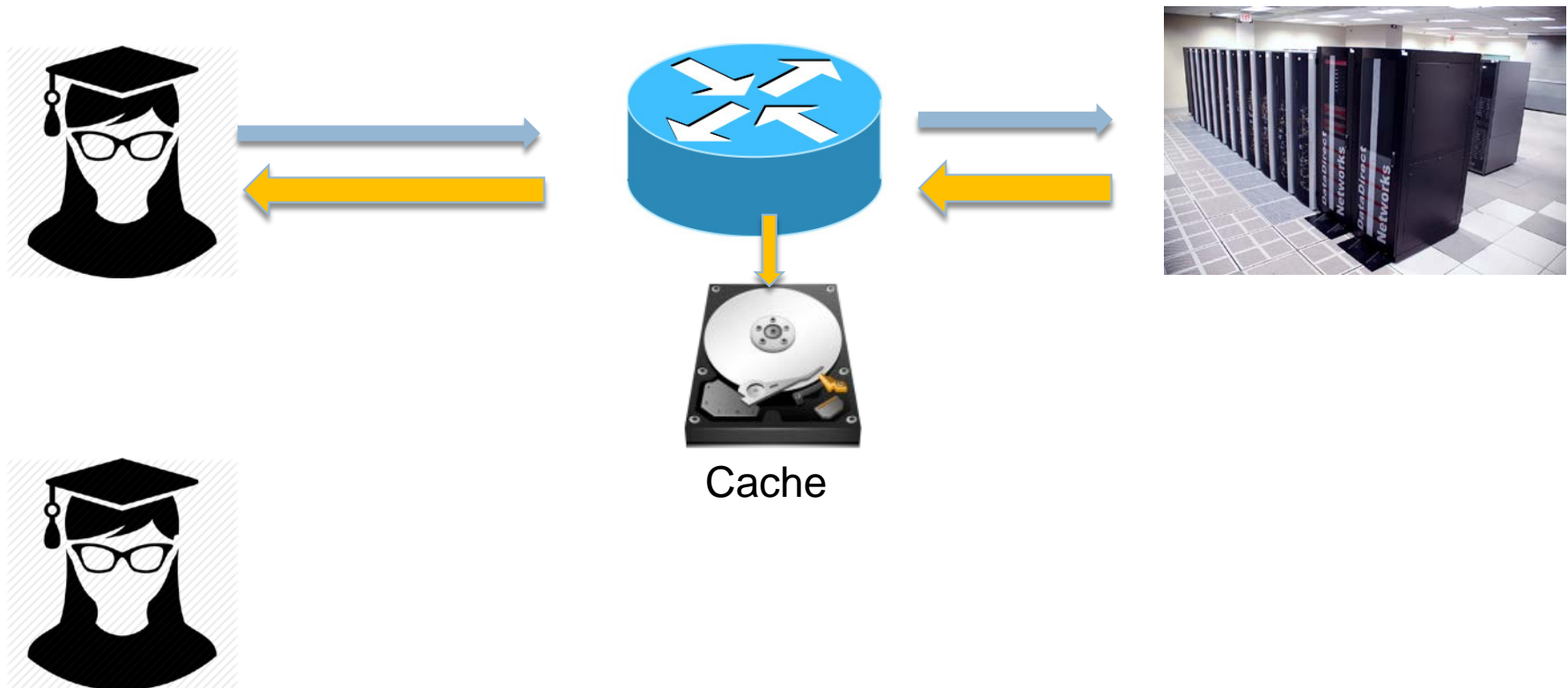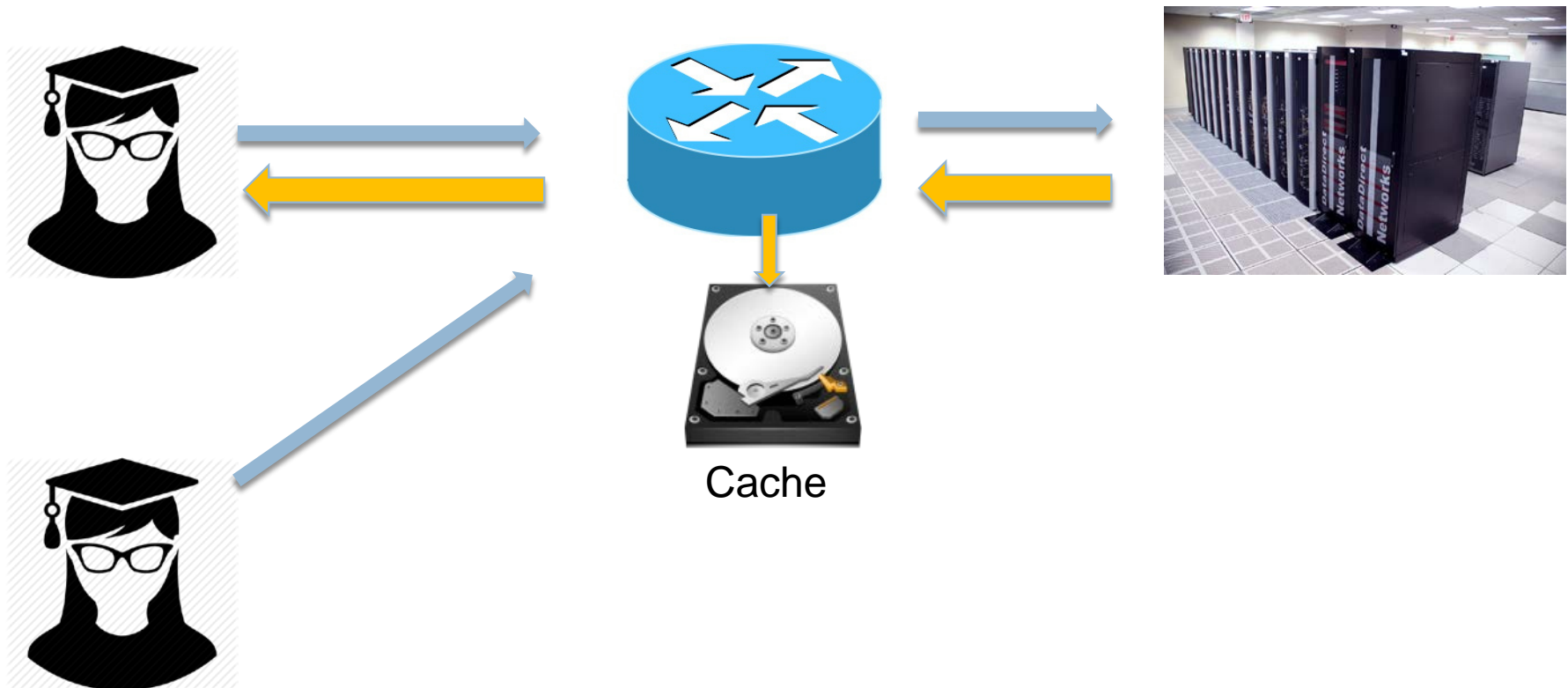

Cache

Named Data Can be Cached

Cache

Named Data Can be Cached

Cache

Named Data Can be Cached

Cache

Named Data Can be Cached

Cache

Named Data Can be Cached

Cache

Named Data Can be Cached

Cache

Named Data Can be Cached

Cache

The network caches the data for you
because it can answer similar questions later
Data is naturally cached to locations of high demand
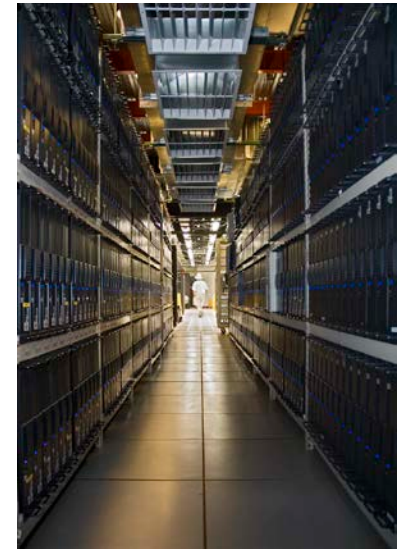**Caching becomes the network operator's problem!**
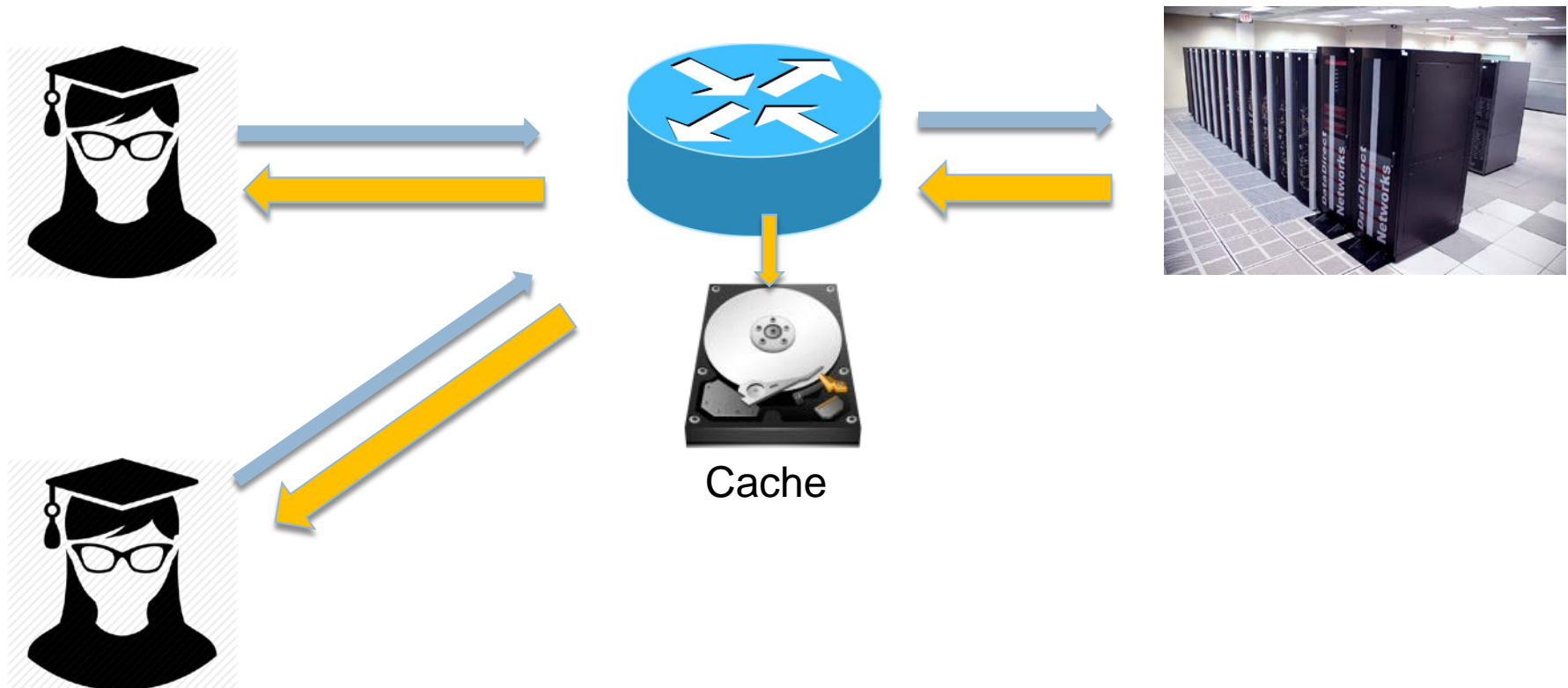
Robustness to Site Failure
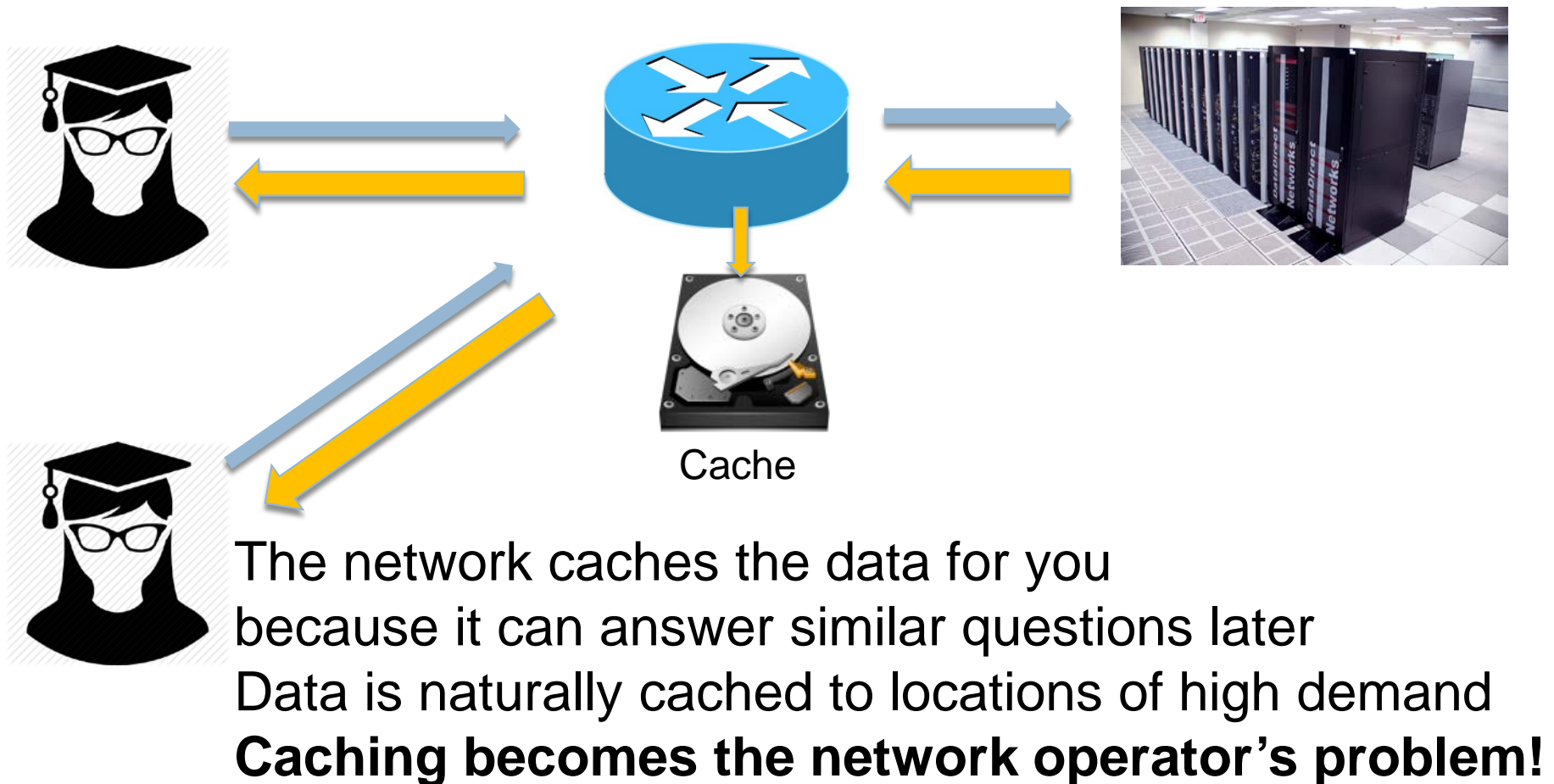
Cache

Robustness to Site Failure
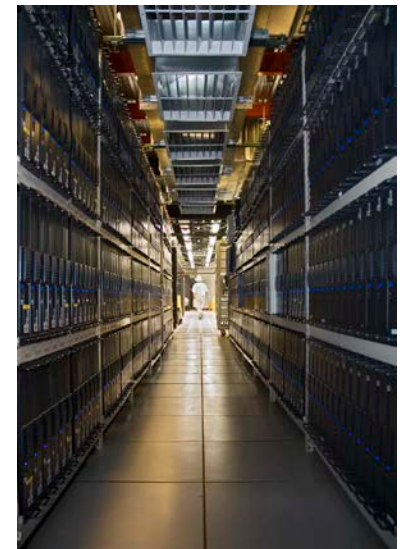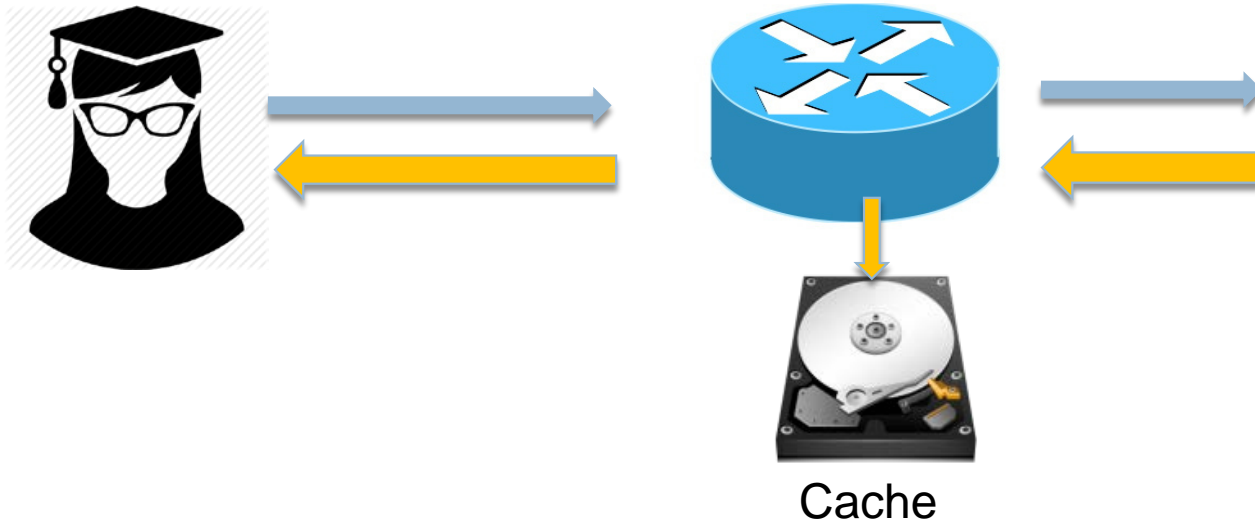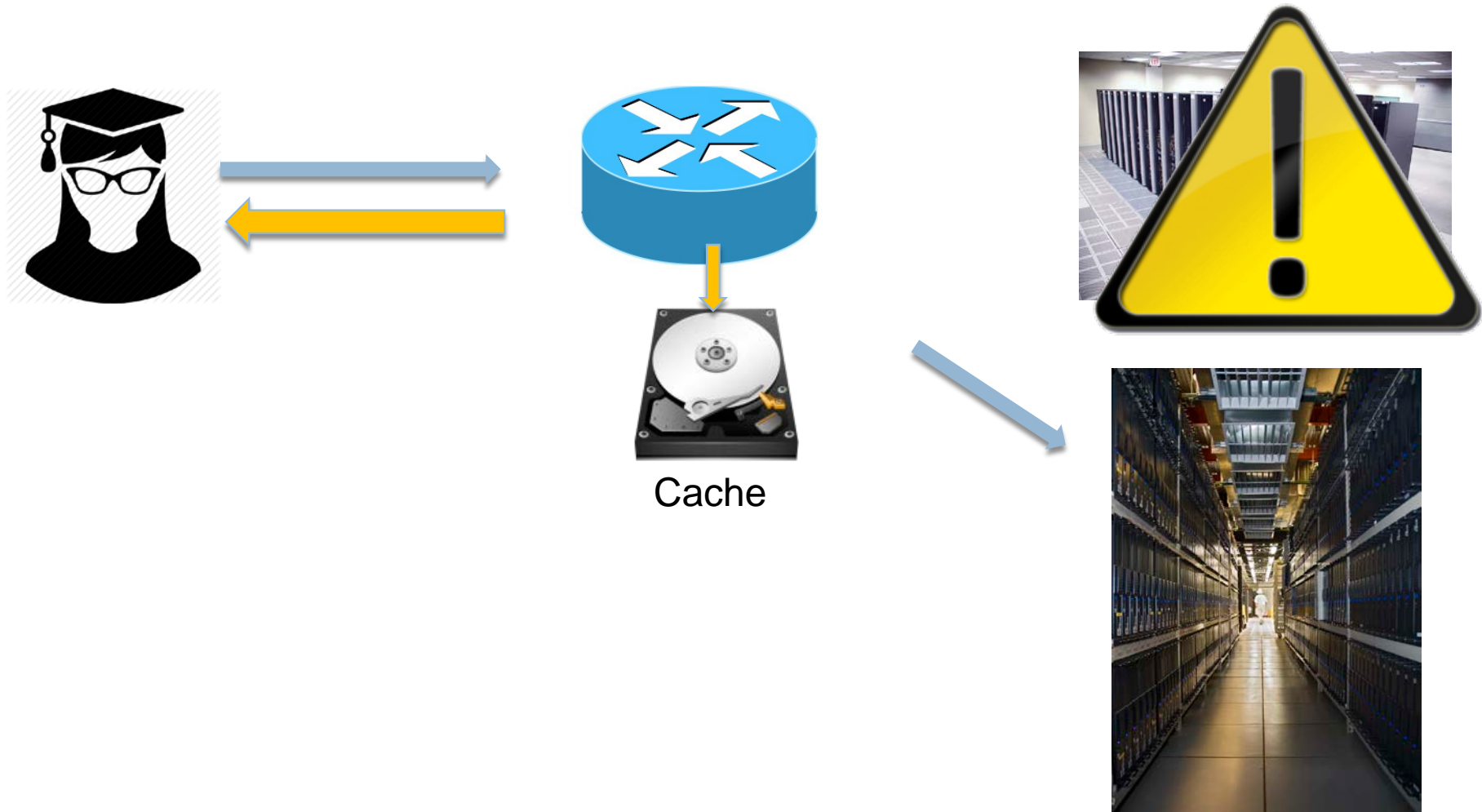
Cache
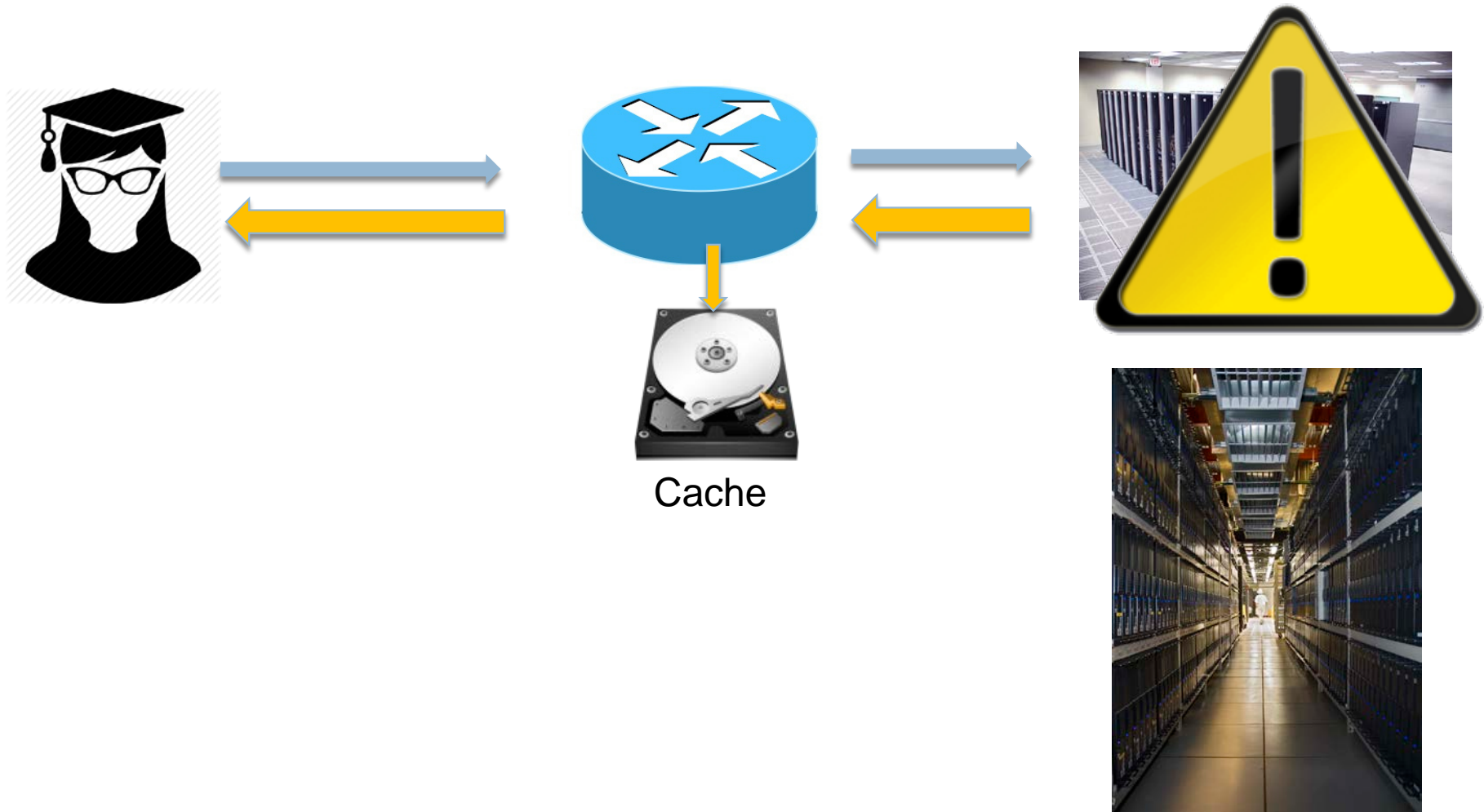
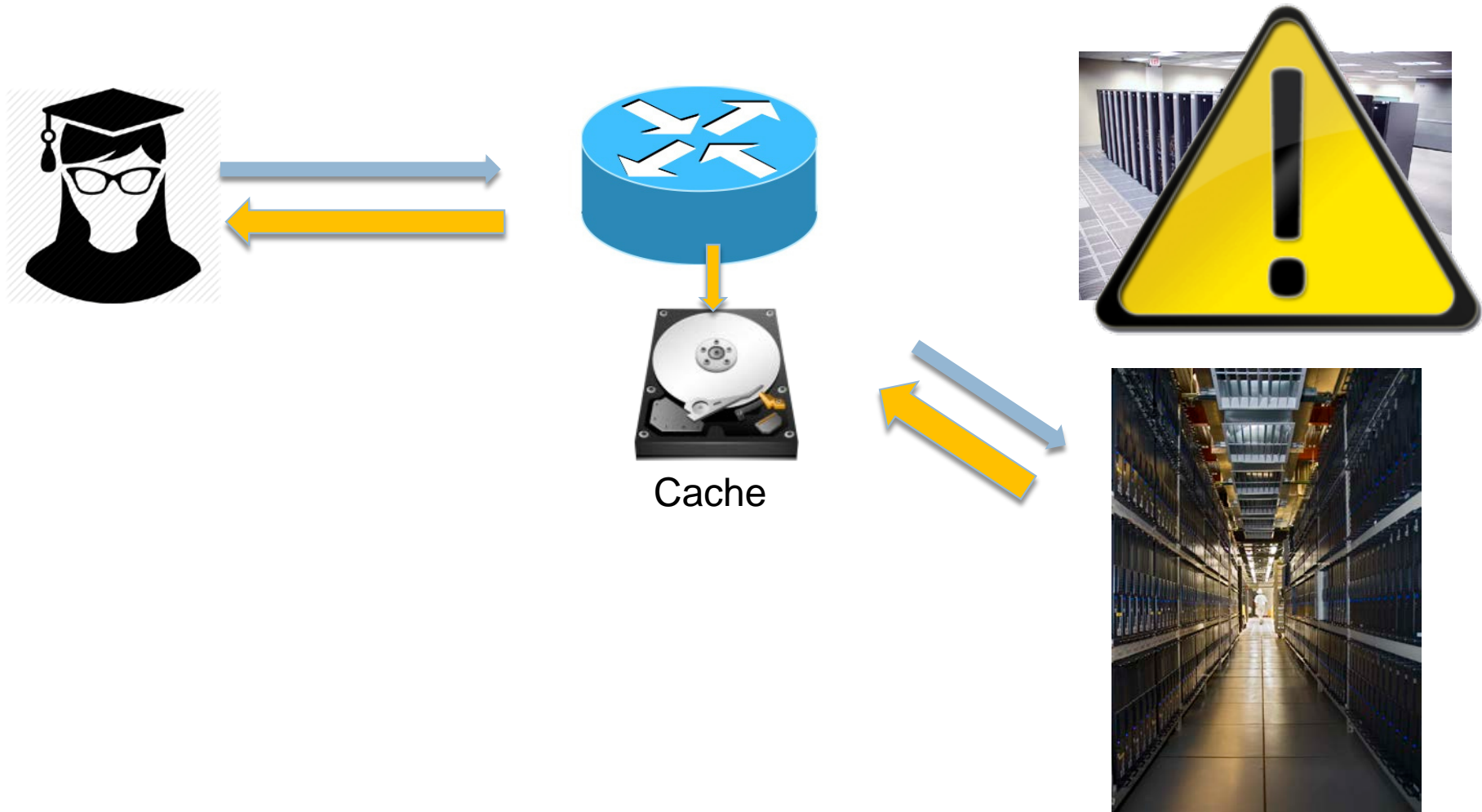# Robustness to Site Failure
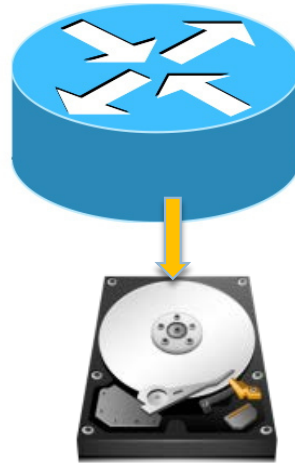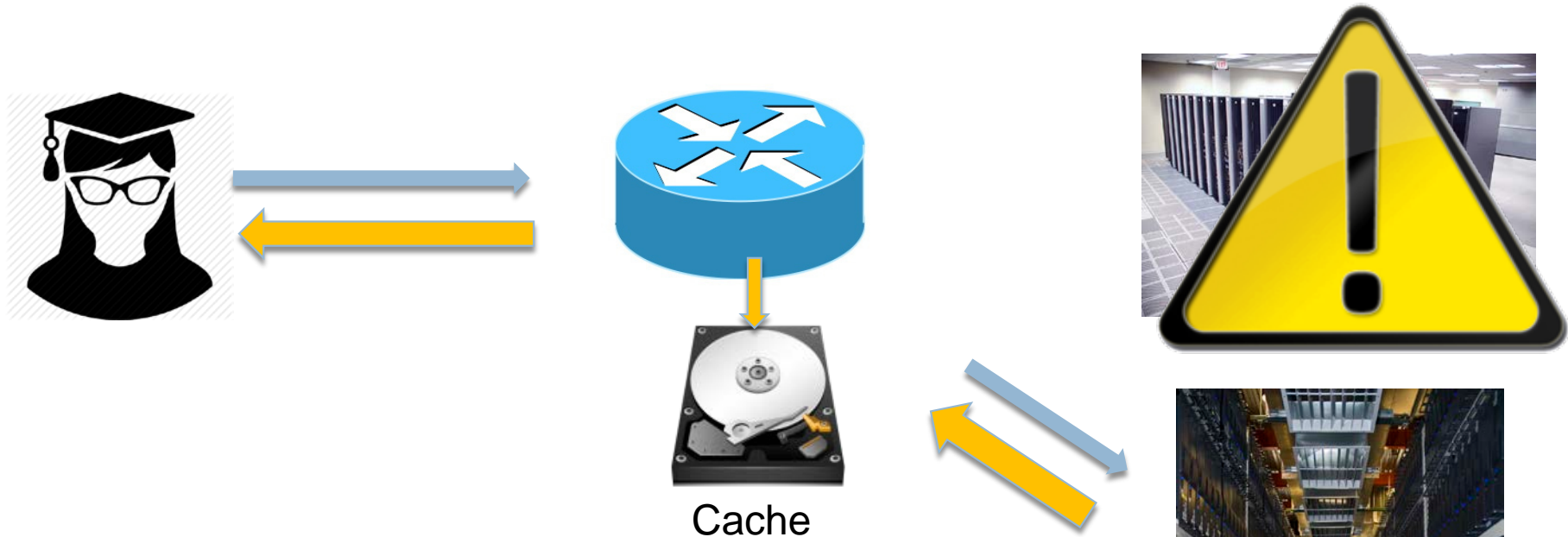

Cache

Robustness to Site Failure

Cache

Robustness to Site Failure

Cache

Robustness to Site Failure

Cache

Failover in a named data network happens **automatically** without application involvement

This Sounds Awfully Complex..

It's not! It's actually quite simple:

- First, name your datasets with a hierarchical, community agreed name:

  - /store/mc/fall13/BprimeBprime_M_3000/GEN-SIM/POSTLS162_v1-v2/10000 /<UUID.root>

- Then, advertise the *prefix* to the network:

  - I can answer any questions starting with:

  - /store/mc/fall13/BprimeBprime_M_3000/GEN-SIM/POSTLS162_v1-v2/*

- Finally, let users issue interests with the appropriate name or name prefix
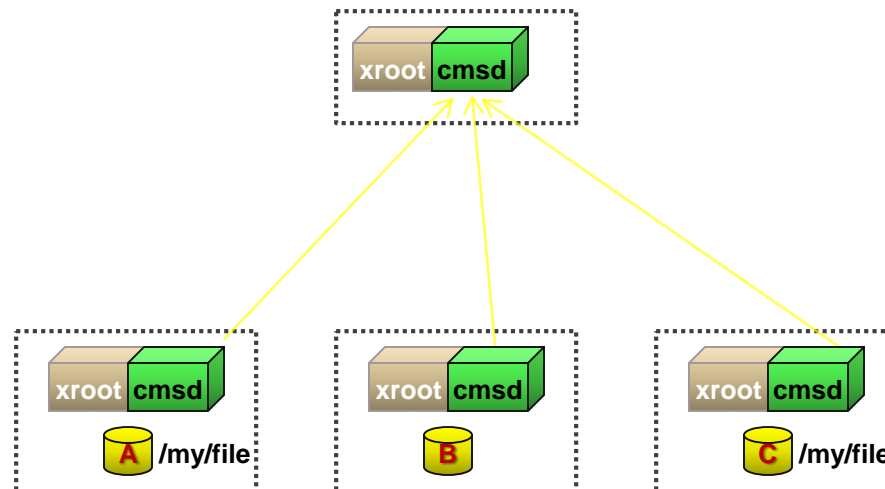
Example: xrootd Cluster

Here is how xrootd works today:
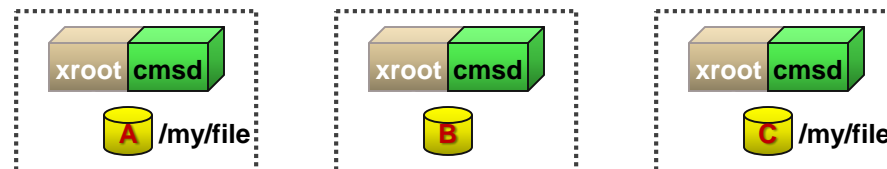
Data Servers
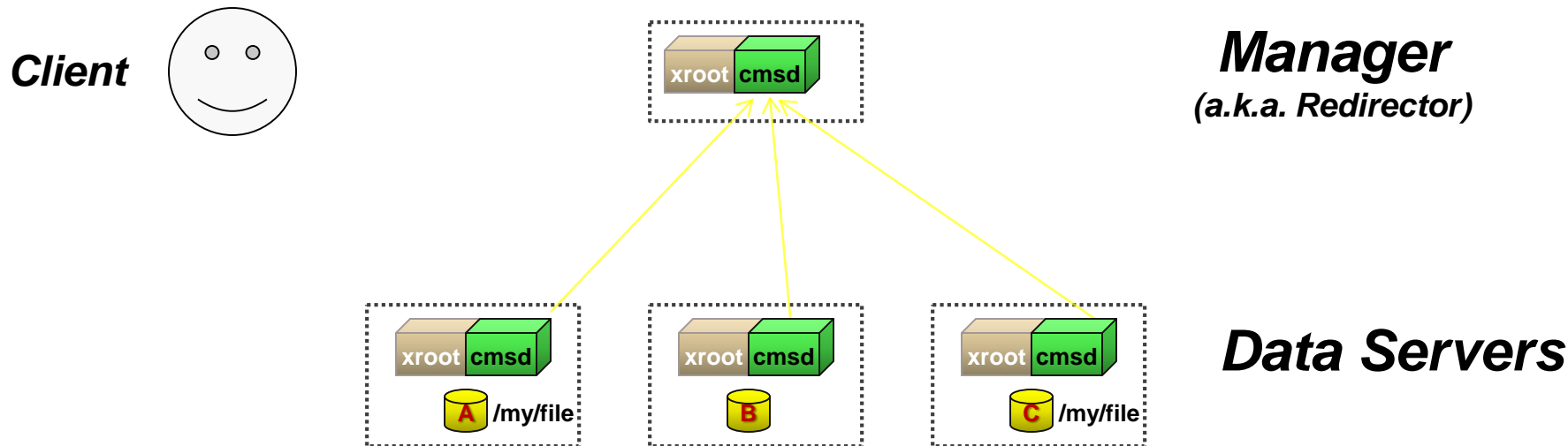
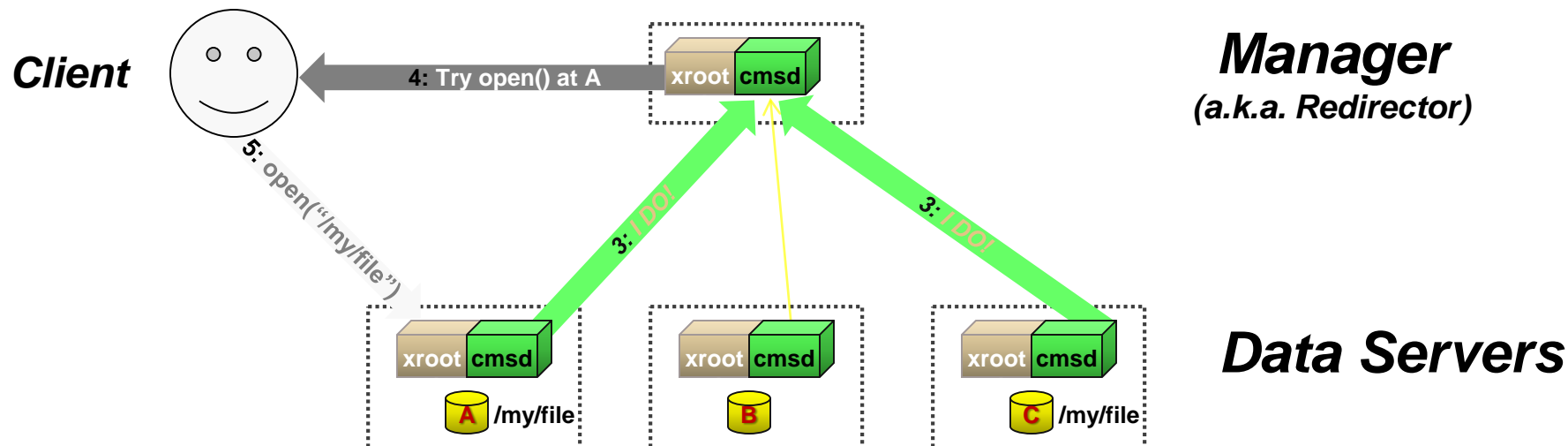# Example: xrootd Cluster

Here is how xrootd works today:
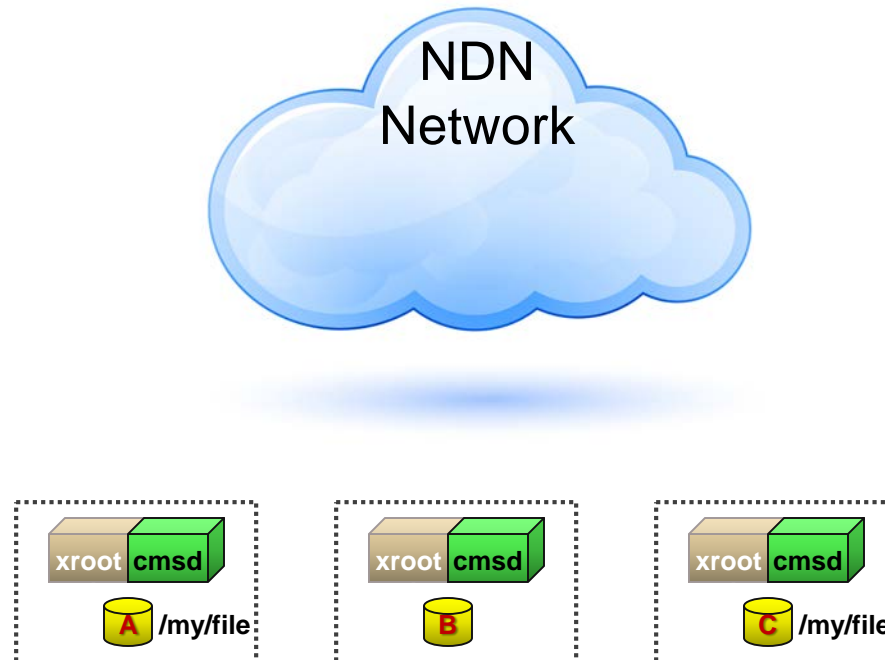
Example: xrootd Cluster

Here is how xrootd works today:

Example: xrootd Cluster

Here is how xrootd works today:

xrootd under NDN
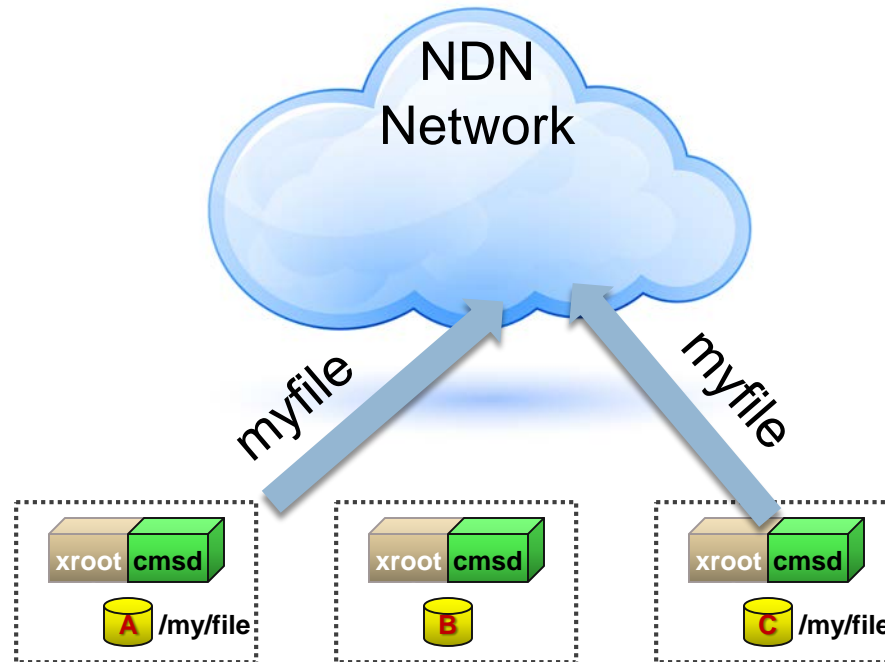
xrootd under NDN

NDN
Network

myfile   myfile

| xroot cmsd | | xroot cmsd | | xroot cmsd |

A /my/file   B   C /my/file

Data Servers

xrootd under NDN

xrootd under NDN

xrootd under NDN

NDN Network

myfile?

Client

myfile?

xroot cmsd
A /my/file

xroot cmsd
B

xroot cmsd
C /my/file

Data Servers
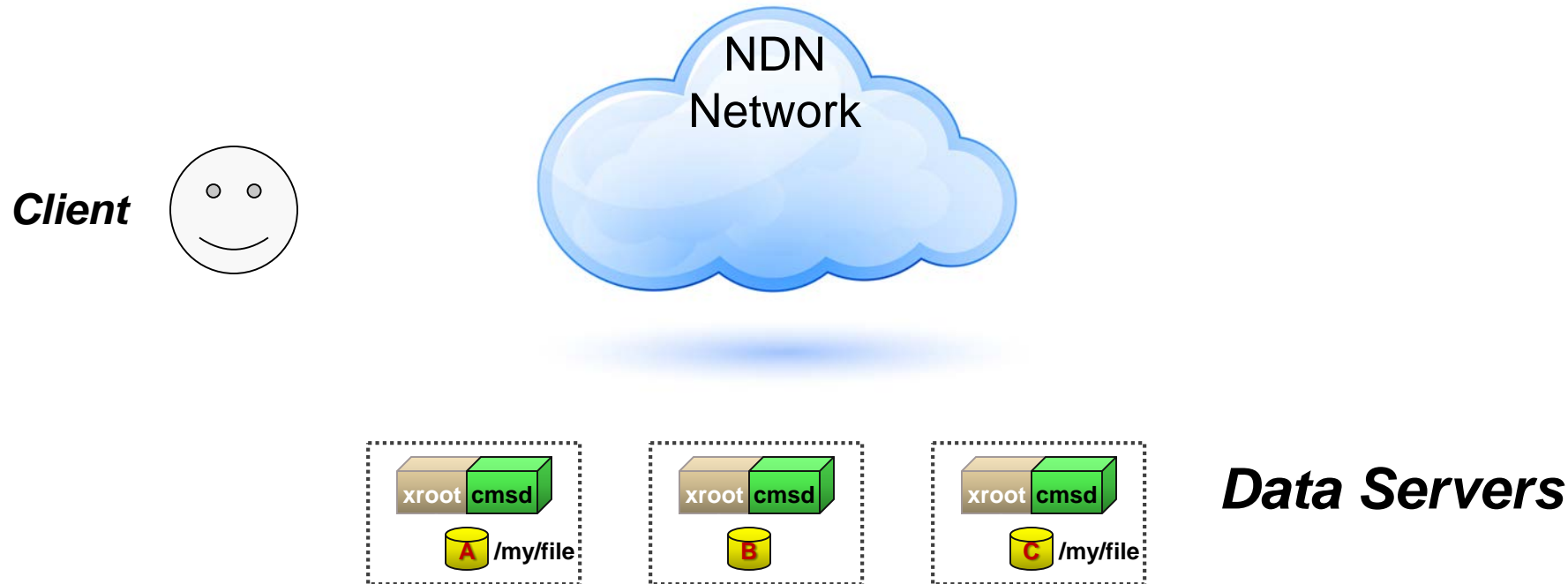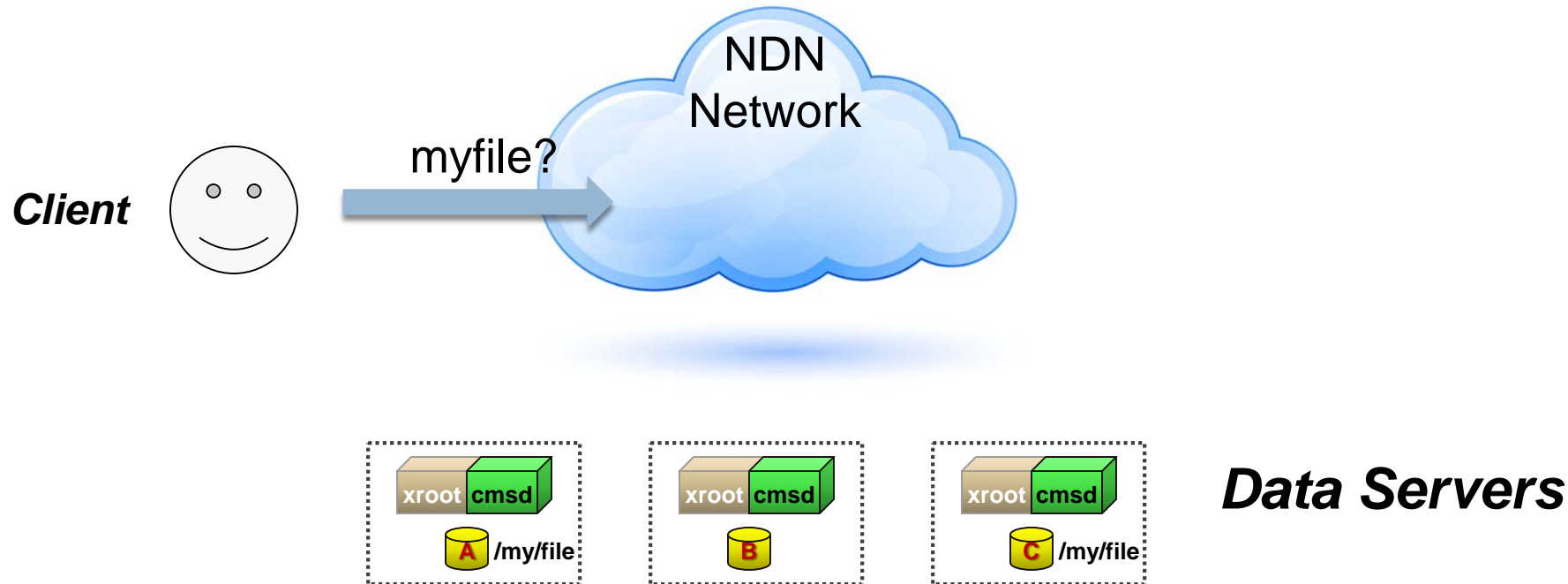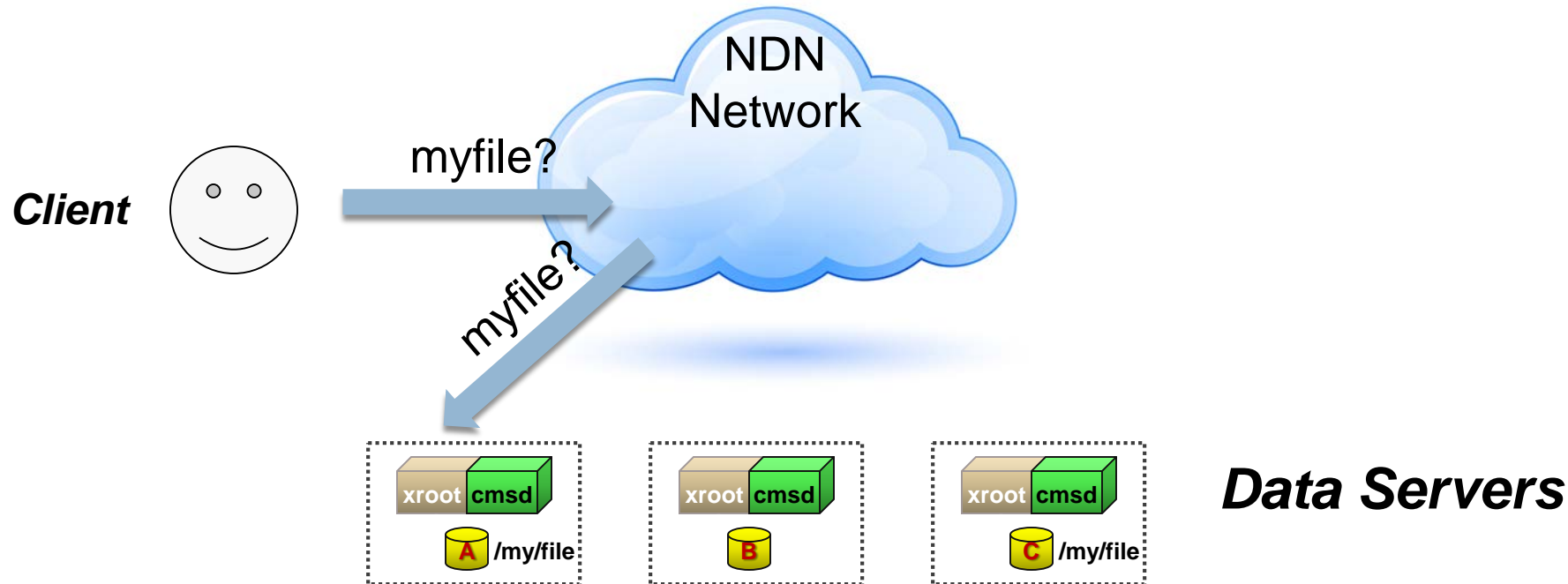
xrootd under NDN

No manager, fewer steps, more robust

NDN Simplifies Networking

IP

NDN

DNS Caching Load Balancing

Failover Security

Addressing Discovery

Naming

NDN Simplifies Networking

IP

NDN

DNS **Caching** **Load Balancing**

Naming
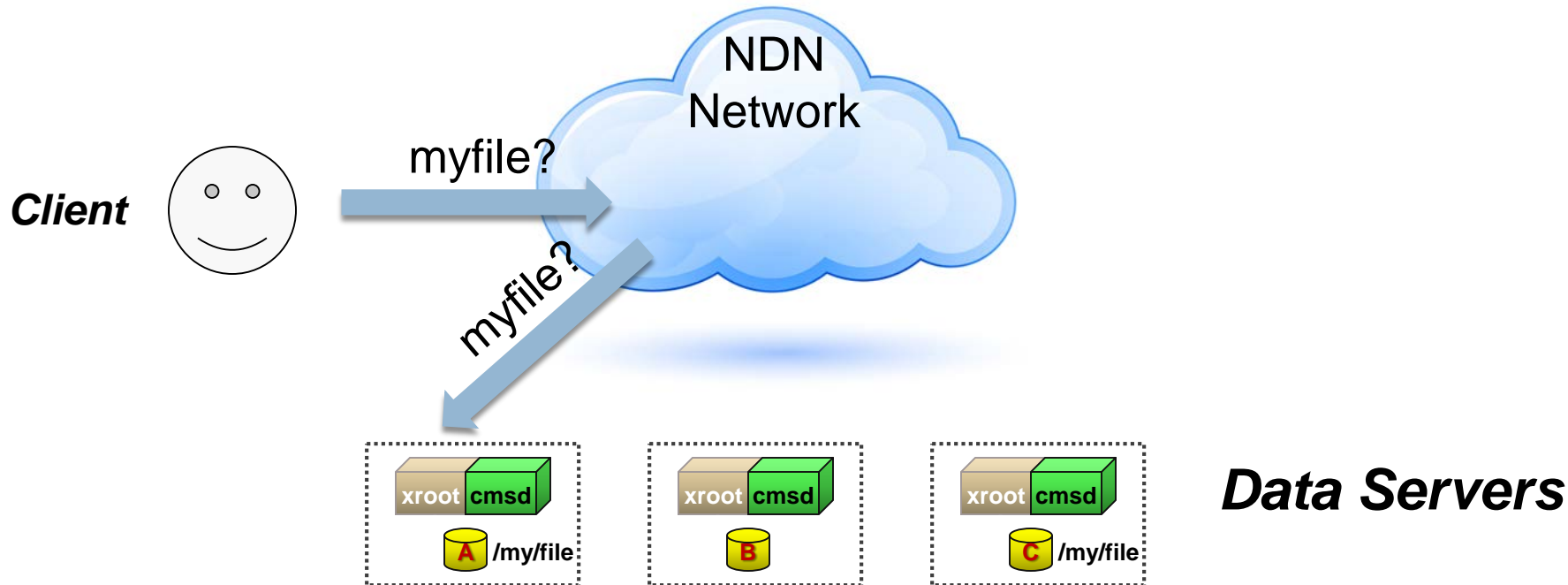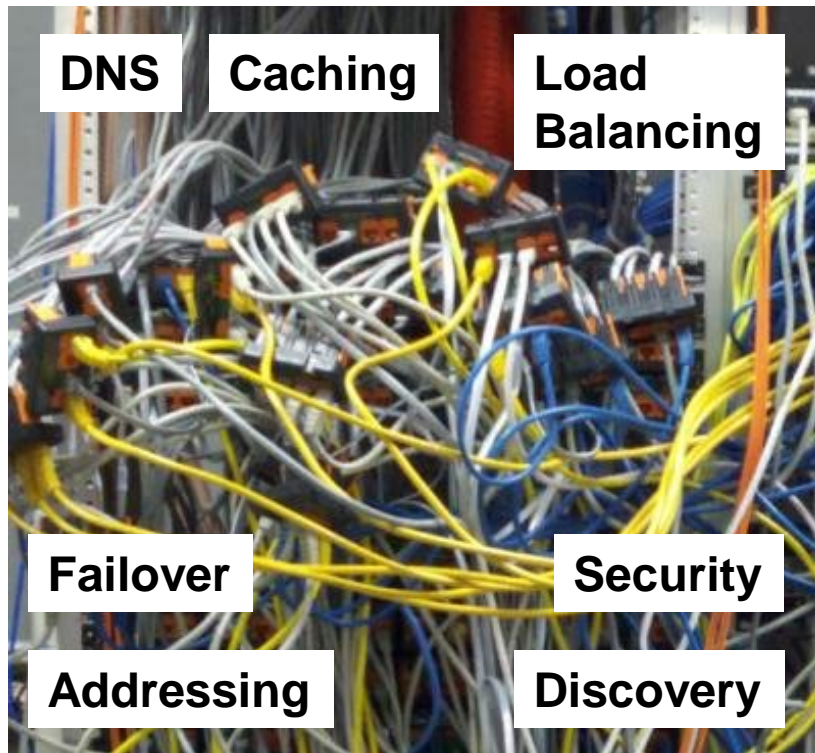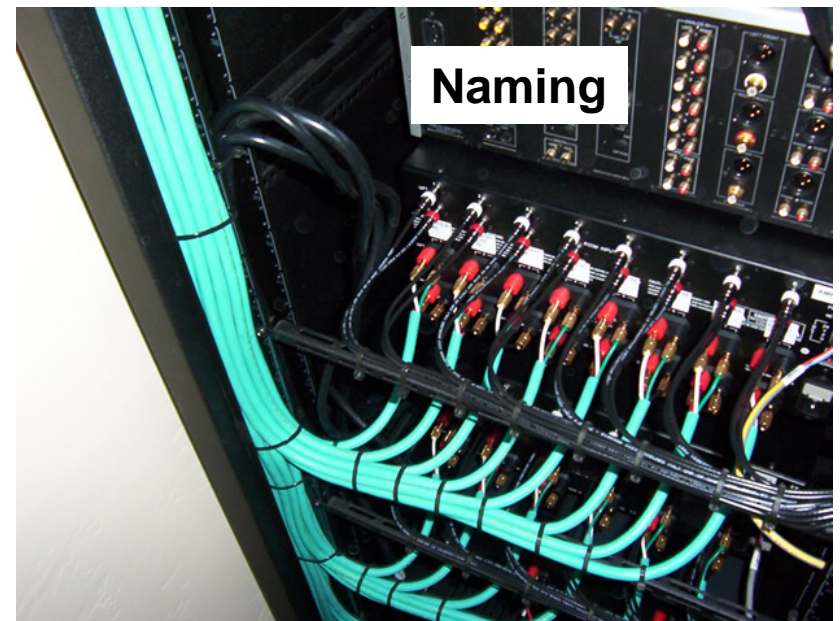
NDN let's you focus on the science not the plumbing!

Failover **Security**

**Addressing** **Discovery**

NDN in Our Climate Testbed
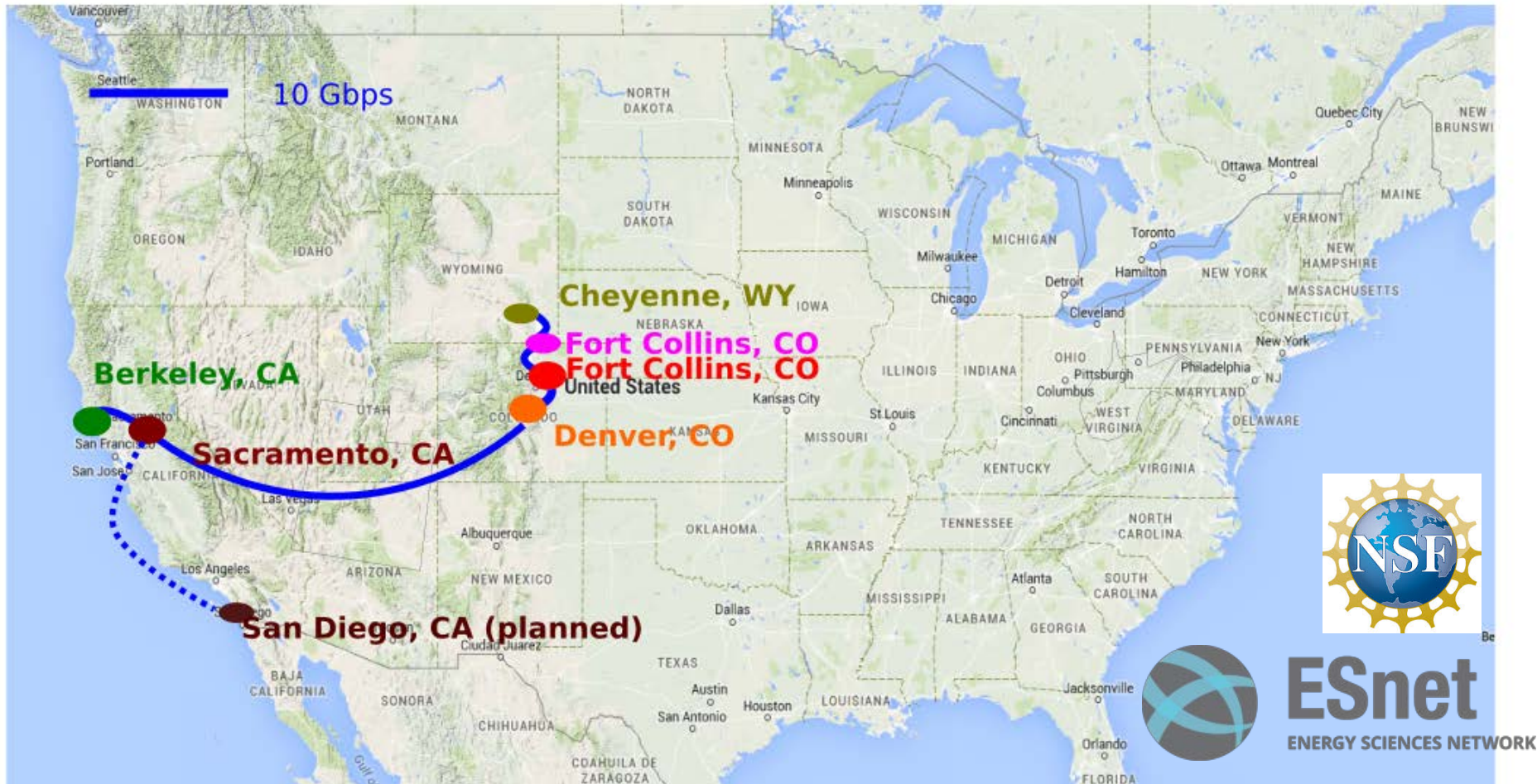
Current State of NDN

- Climate testbed hardware: six state of the art Dell PE720xd servers with 10G fiber connections throughout

  - 20 cores, 128GB RAM, 48TB local storage

- Cisco and others currently building specialized routers

- Washington University implementing forwarding algorithms that support lookups at 1B prefixes at line speed

Optimized NDN for LHC Network

- NDN can automatically and dynamically cache "hot" files near users for lower delay

- VIP algorithm: optimal interest forwarding and cache placement/replacement

- 76% lower latency over no caching
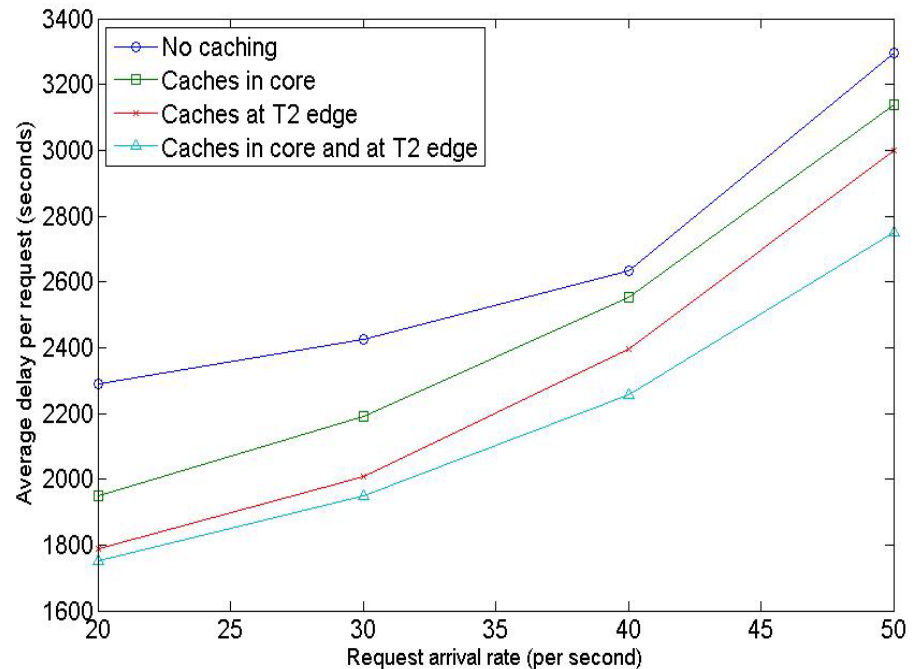
Yeh et al. 2014

Conclusions

- In HEP, mistakes are expensive..

- ..hard to do course correction midstream..

- ..so better get it right the first time!

- NDN frees up resources for more intelligent networks..

- ..to reduce waste

contact: christos @cs.colostate.edu