

# **NAMED DATA NETWORKING: AN INTERNET ARCHITECTURE FOR THE FUTURE**

Christos Papadopoulos  
Colorado State University

**November 18, 2016**

**Work supported by NSF #1345236 and #13410999**





# The NDN Project

- Part of the NSF Future Internet Architecture FIA initiative
- Goal: design the next generation Internet Architecture
- NDN is one of four multi-institution teams funded in 2010-13, and 2014-16, ~\$15M

# NDN Institutions

---



THE UNIVERSITY OF  
MEMPHIS®



<http://named-data.net>

<http://github.com/named-data>

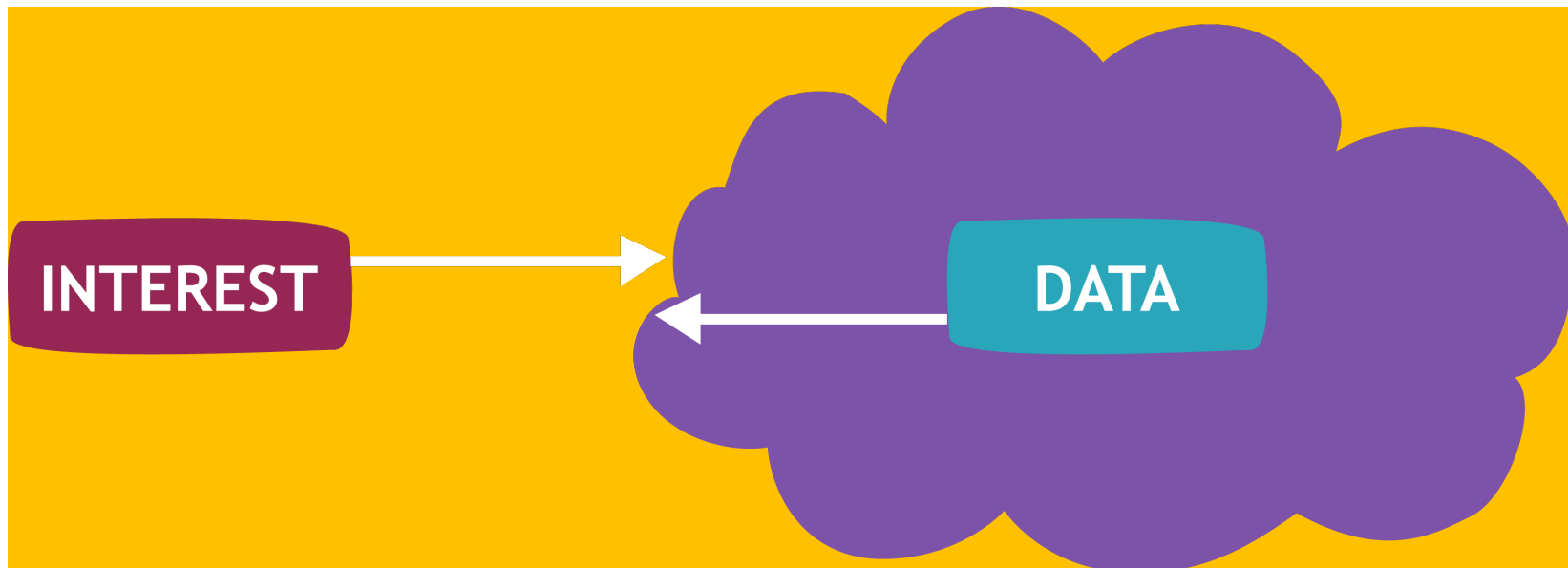
Colorado State University

# The Problem with Today's Internet

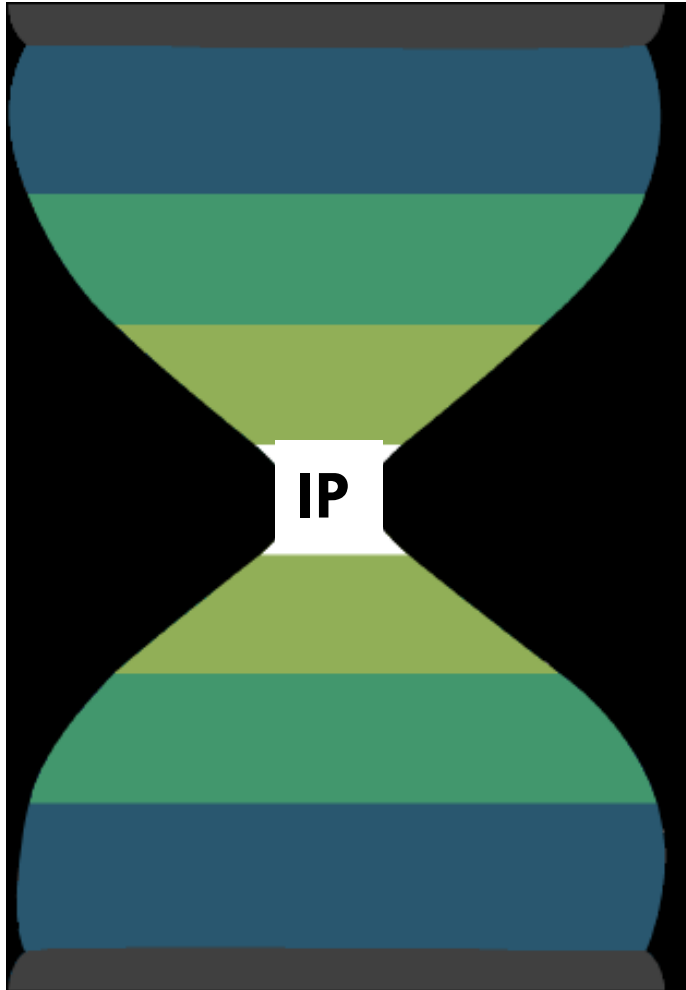
- To find content in the network
- ..you have to learn **where** the content is
- ..and then ask the network to take you there
- ..so you can ask the server for what you want
- But no-one cares about servers anymore..
- ..we want the Data!
- **Service model mismatch**

# Named Data Network (NDN)

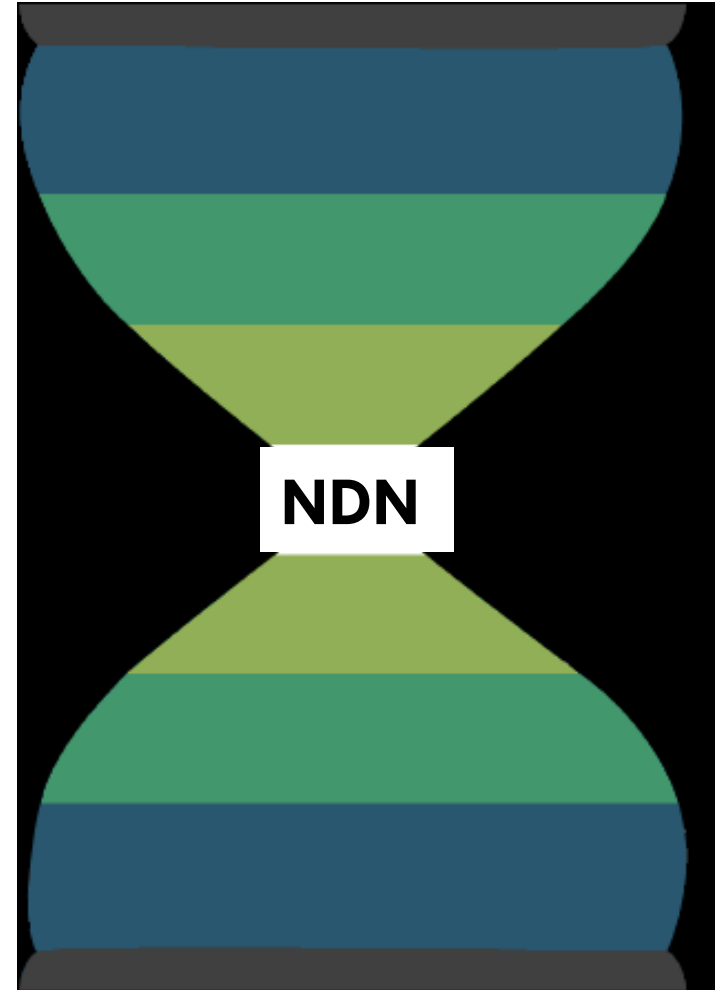
- The main idea: **Name the data, not the hosts!**
- ..so you just give the network the name of the data you want..
- ..and let the network find it and get it to you



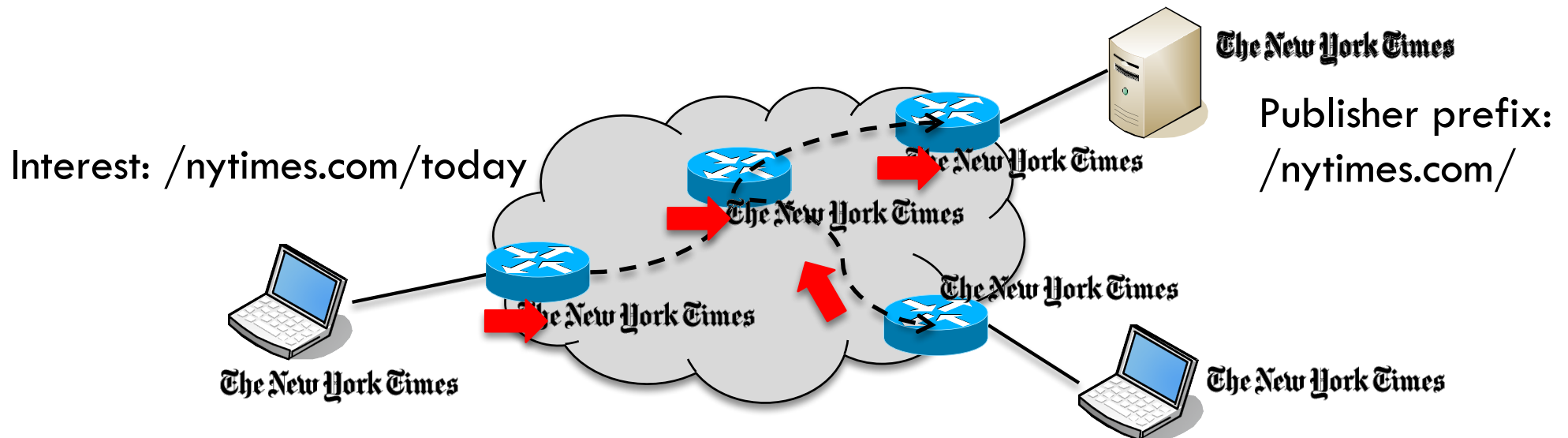
## Host-centric addressing



## Data-centric addressing

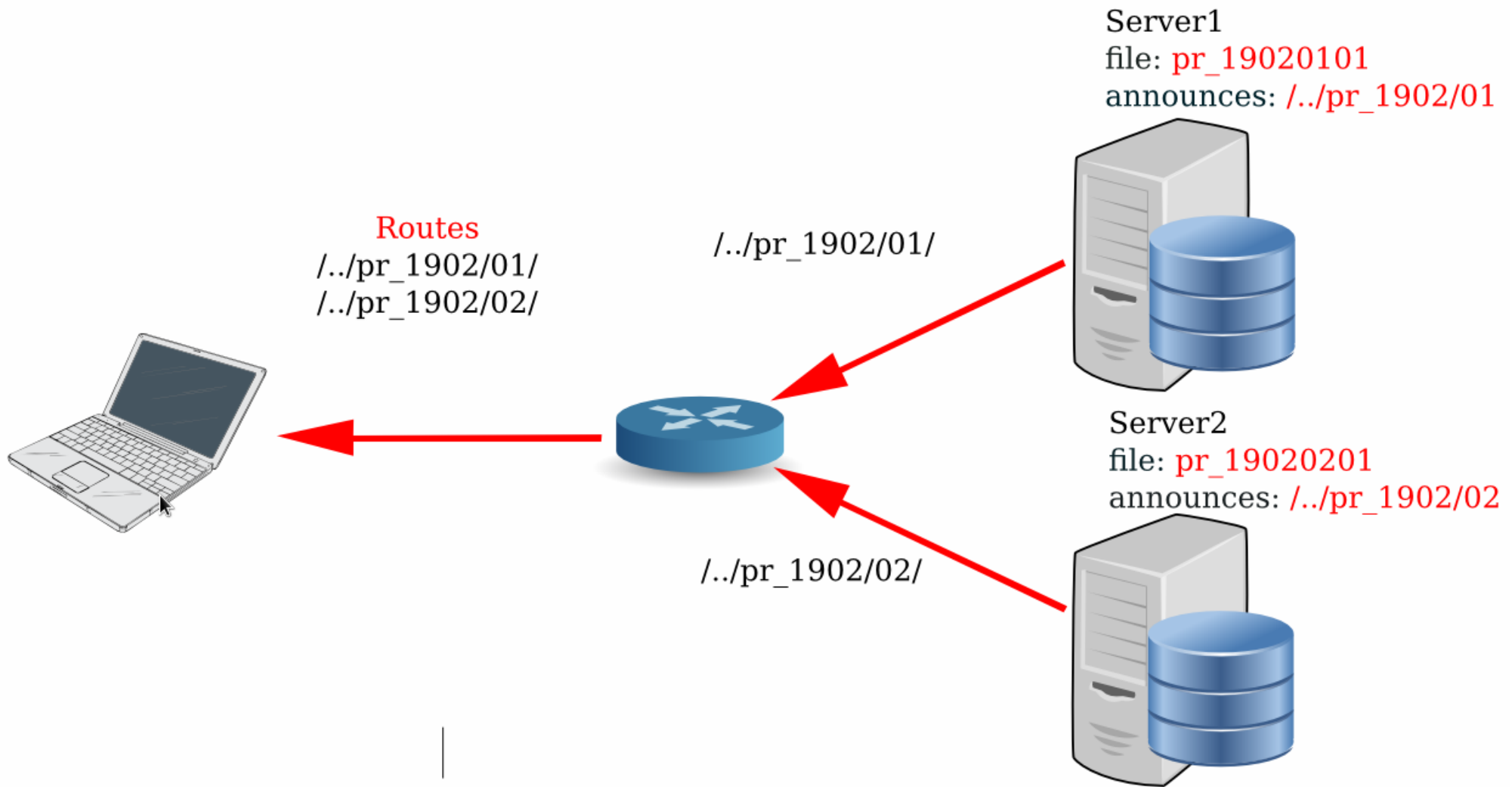


# NDN Operation



- ❑ Publishers push **hierarchical** name prefixes into the network
- ❑ Users send **Interests** that follow path to published prefix
- ❑ “Breadcrumbs” direct **data** back to the user
- ❑ Data is **cached** into the network

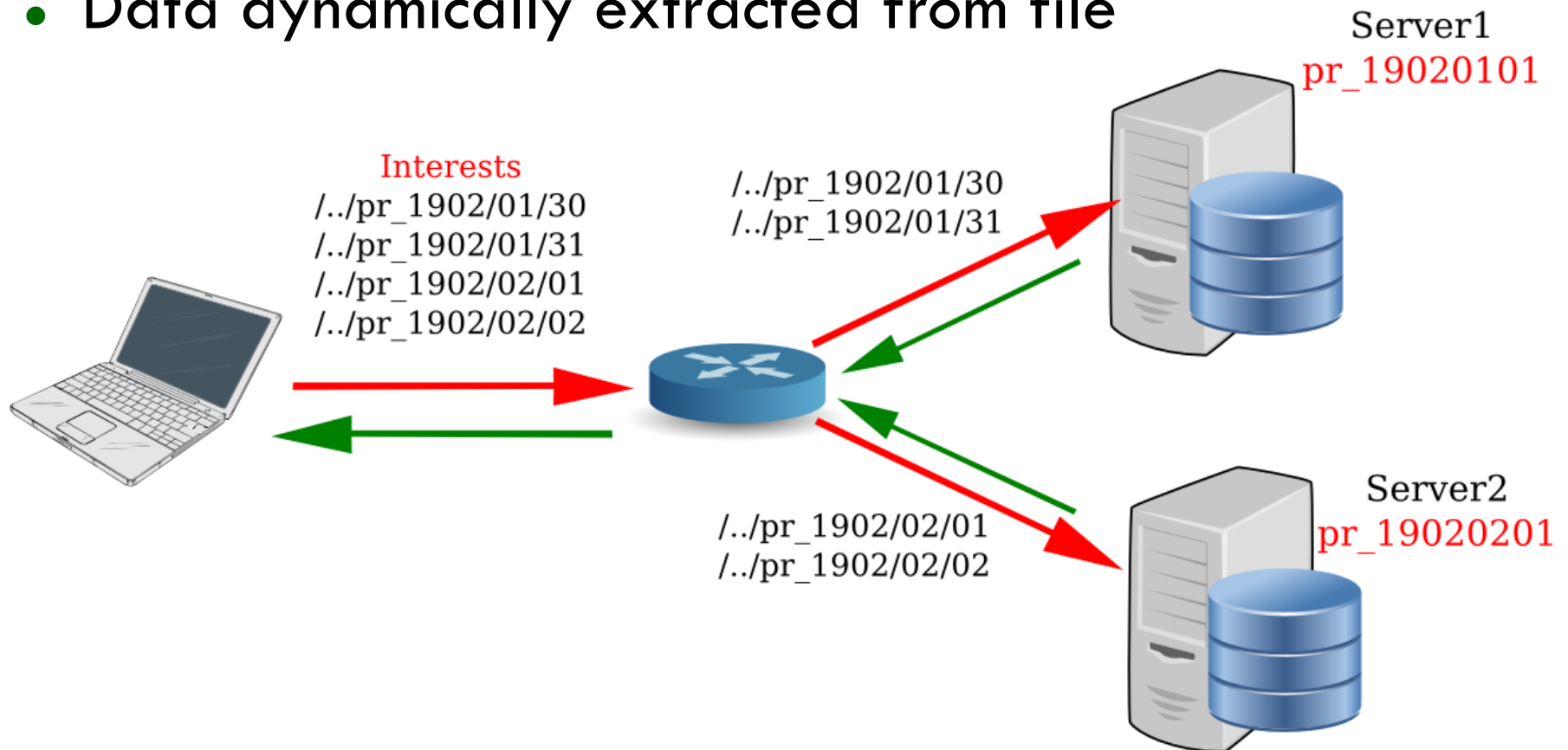
# Content Publishing





# Data Request

- Interests for Jan 30-31 go to server1
- Interests for Feb 01-02 go to server2
- Data dynamically extracted from file



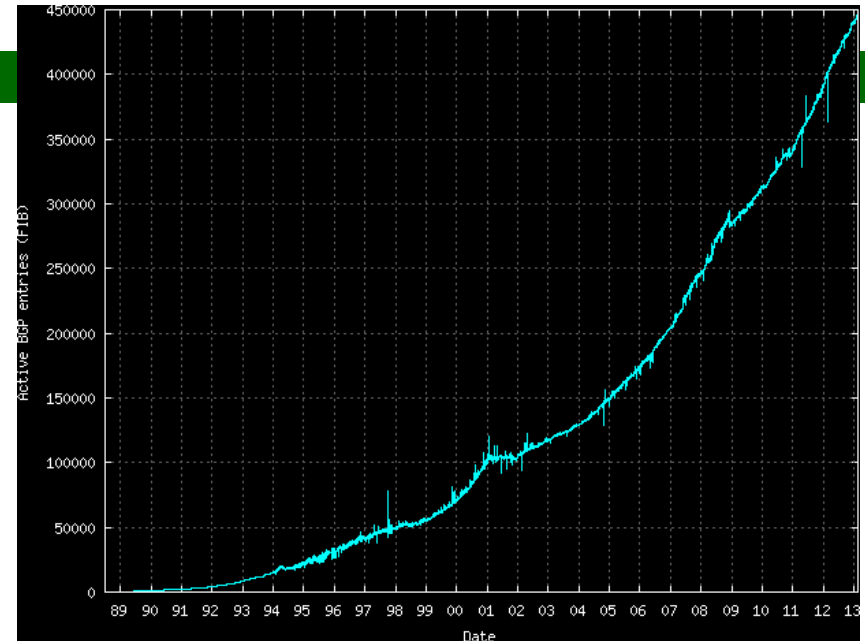
# This Sounds Awfully Complex..

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

# Can it Scale?

- ❑ WUSTL Results for NDN Forwarding
- ❑ Software router prototype
- ❑ Preliminary hardware design
- ❑ Multi-gigabit forwarding rates for:
  - Name-based FIBs, based on real world URLs, of 1-3M entries;
  - Synthetic FIBs, based on model of future namespace, of up to 1B entries.

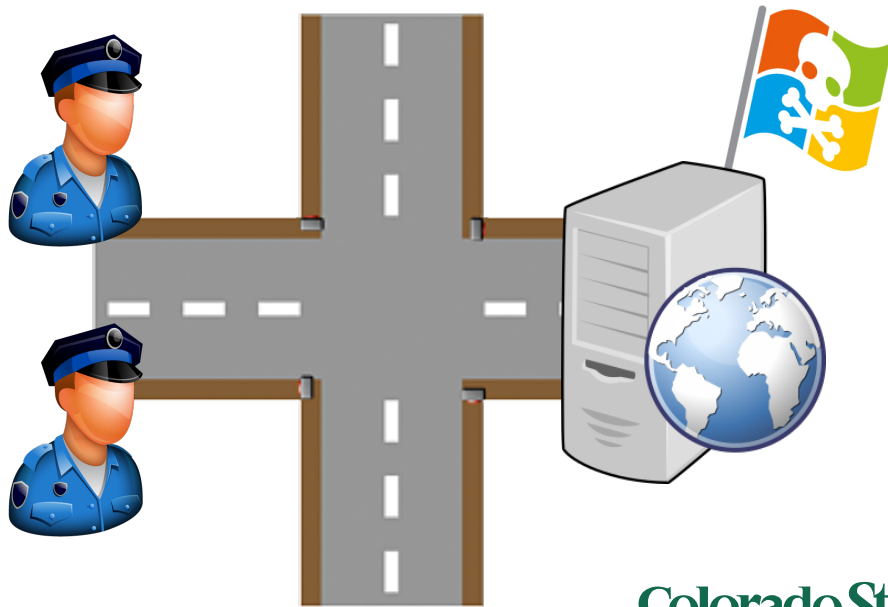


<http://www.cidr-report.org>

All	New	Deleted	Transferred	TLD
144,040,087	155,151	136,956	239,097	All TLDs
107,508,504	115,331	100,189	194,269	.COM
15,033,351	16,353	13,437	20,649	.NET
10,204,641	9,829	7,165	10,224	.ORG
7,185,246	8,227	12,882	8,916	.INFO
2,305,965	3,715	1,875	2,895	.BIZ
1,802,380	1,696	1,408	2,144	.US

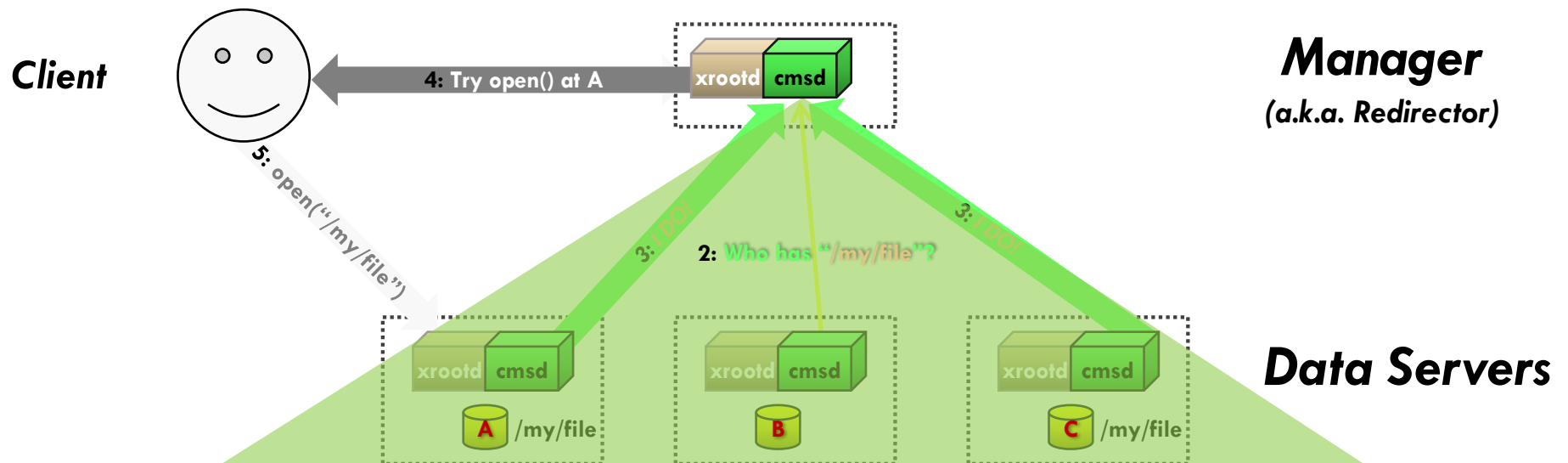
# Named Data is Easy to Secure

- 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!



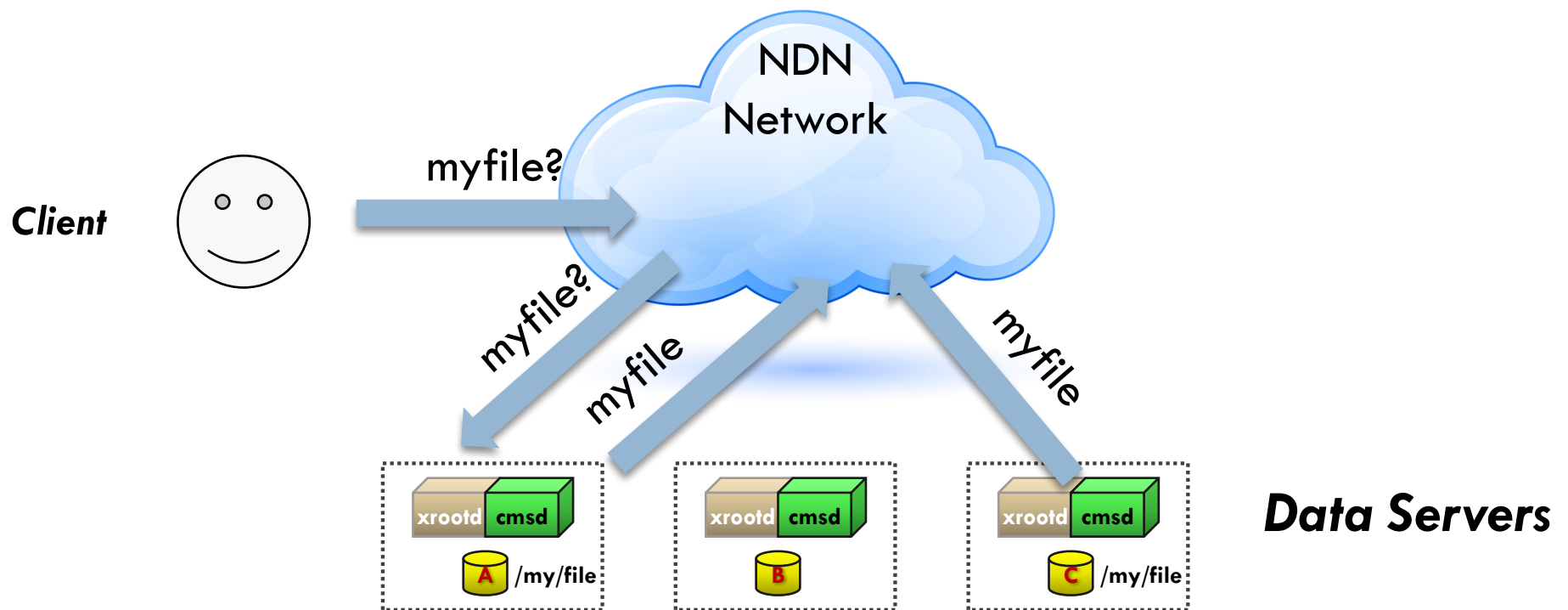
# Simplifying a Complex System: xrootd Cluster

Here is how xrootd works today:



# xrootd under NDN

No manager, fewer steps, more robust



# Supporting Science Applications

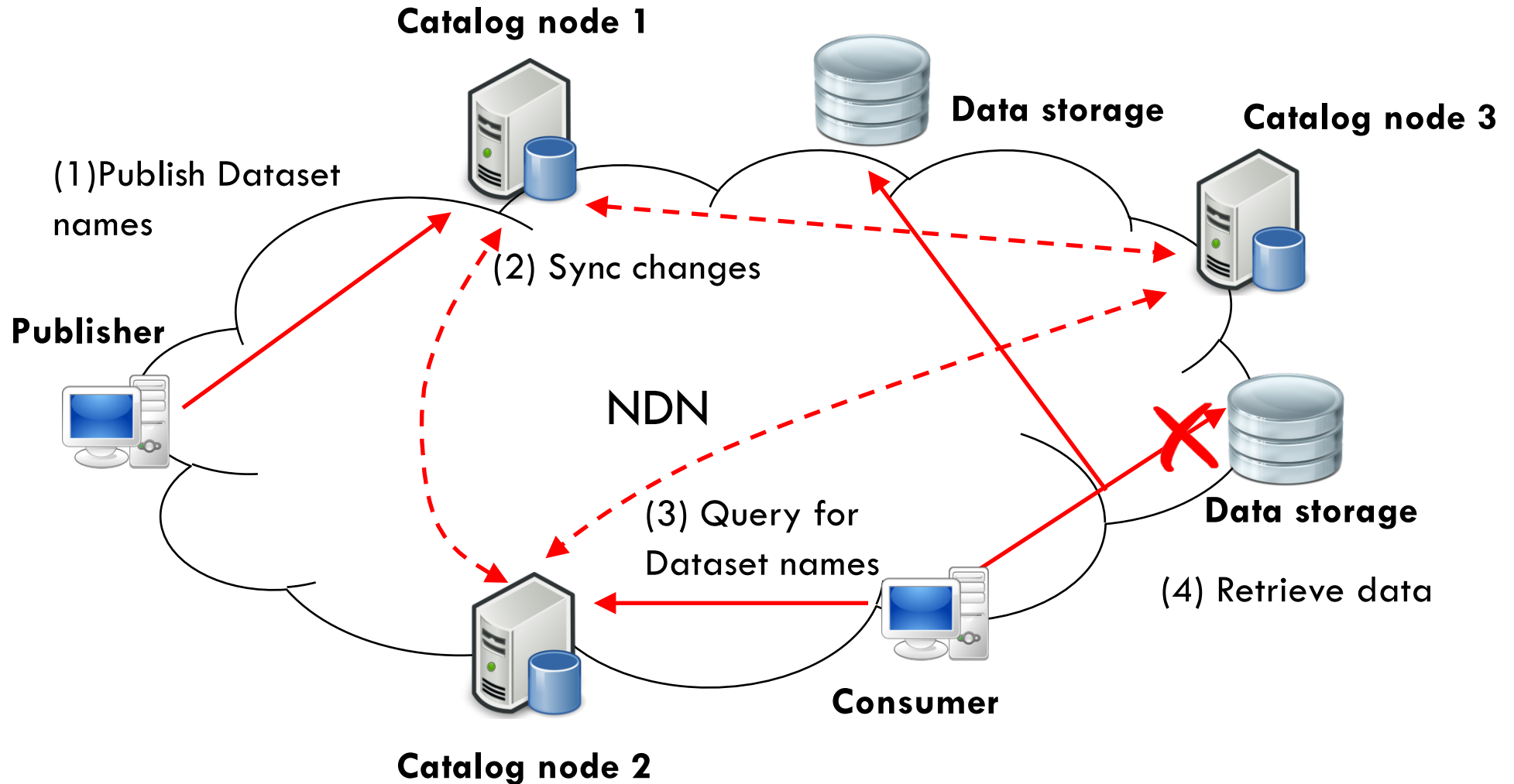
- Scientific apps generate tremendous amounts of data and face challenging management issues
  - Climate science CMIP5 dataset: 3.5 PB
  - High Energy Physics (HEP): 1 PB/s raw data, ATLAS project filters to 4 PB/yr
  - Data distributed to various local repositories
  - Variety of data naming schemes
    - E.g. different units and user defined parameters
  - Data provenance
- Existing, mature, software for dataset discovery, publishing, and retrieval
  - E.g. ESGF, xrootd, etc.
  - Lots of effort to overcome fragility of IP's host-centric paradigm

# First Step – Build a Catalog

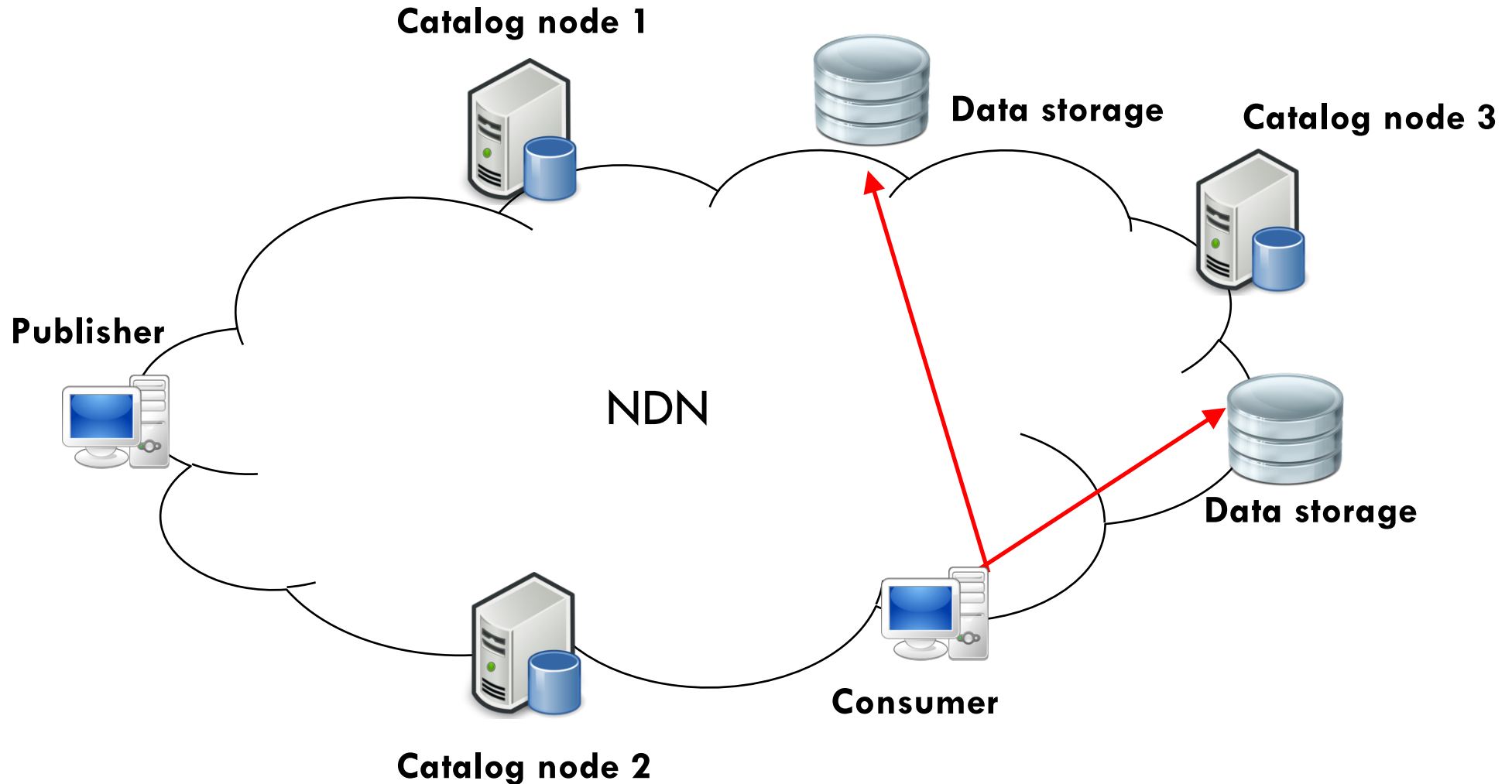
- Create a **shared resource** – a distributed, synchronized catalog of names over NDN
  - Provide common operations such as publishing, discovery, access control
  - Catalog only deals with name management, not dataset retrieval
  - Platform for further research and experimentation
- Research questions:
  - Namespace construction, distributed publishing, key management, UI design, failover, etc.
  - Functional services such as subsetting
  - Mapping of name-based routing to tunneling services (VPN, OSCARS, MPLS)



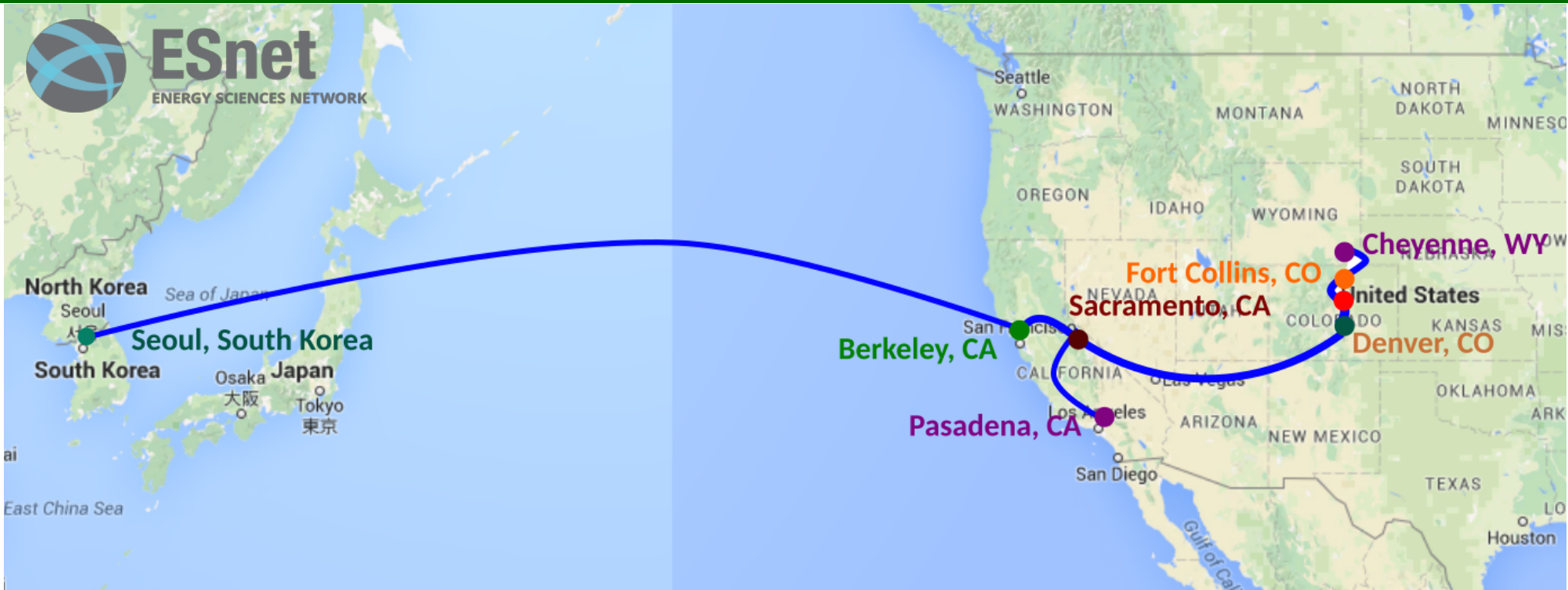
# NDN Catalog



# Forwarding Strategies



# Science NDN Testbed



- NSF CC-NIE campus infrastructure award
  - 10G testbed (courtesy of ESnet, UCAR, and CSU Research LAN)
- Currently ~50TB of CMIP5, ~20TB of HEP data

# Conclusions

---

- NDN encourages common **data** access methods where IP encourages common **host** access methods
  - NDN encourages interoperability at the content level
- Many playgrounds for you to play
  - Science, multimedia, IoT, Android, wireless and more
  - Ready-to-try catalog, supports a variety of applications
  - UI for data search and retrieval.

# For More Info

---

christos@colostate.edu

<http://named-data.net>

<http://github.com/named-data>