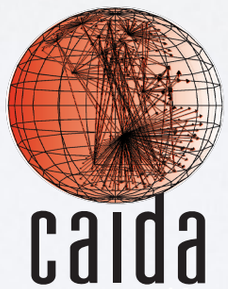
A Brief History of a Future Internet: the Named Data Networking Architecture

kc claffy
(w slides borrowed from NDN team)



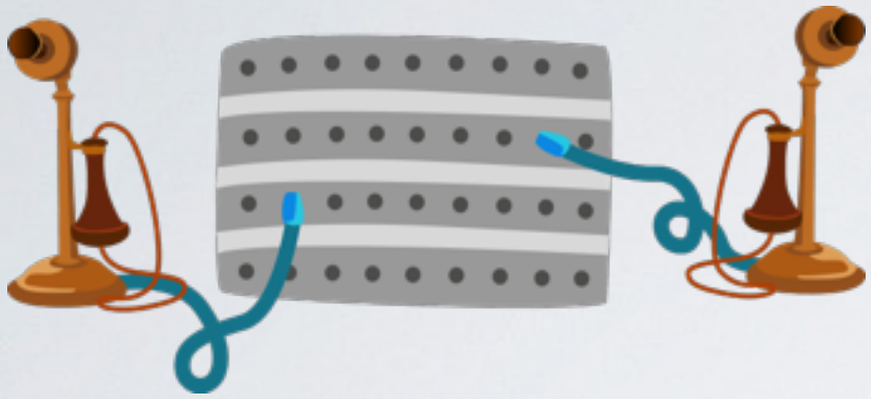
Center for Applied Internet Data Analysis
University of California, San Diego

OUTLINE OF TALK

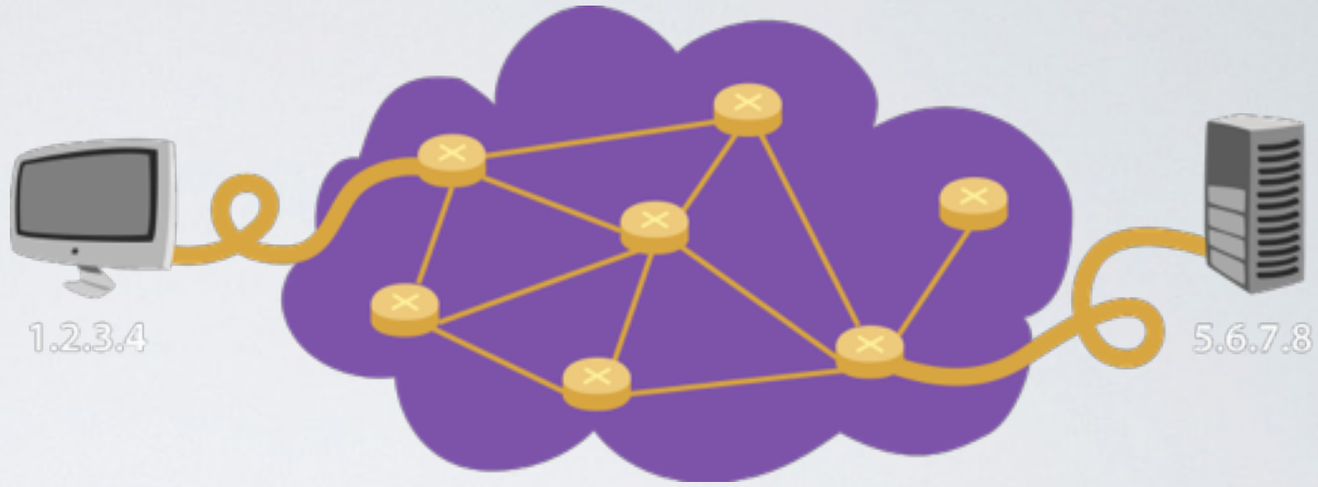
- Motivation
- Evolution of networking communications architecture(s) for last 100 years
- IP architecture matches its use less and less
- "New" (7-year old) research project
  - design a global Internet architecture
    - using what we have learned about the Internet

Big data, small data: exponentials abound

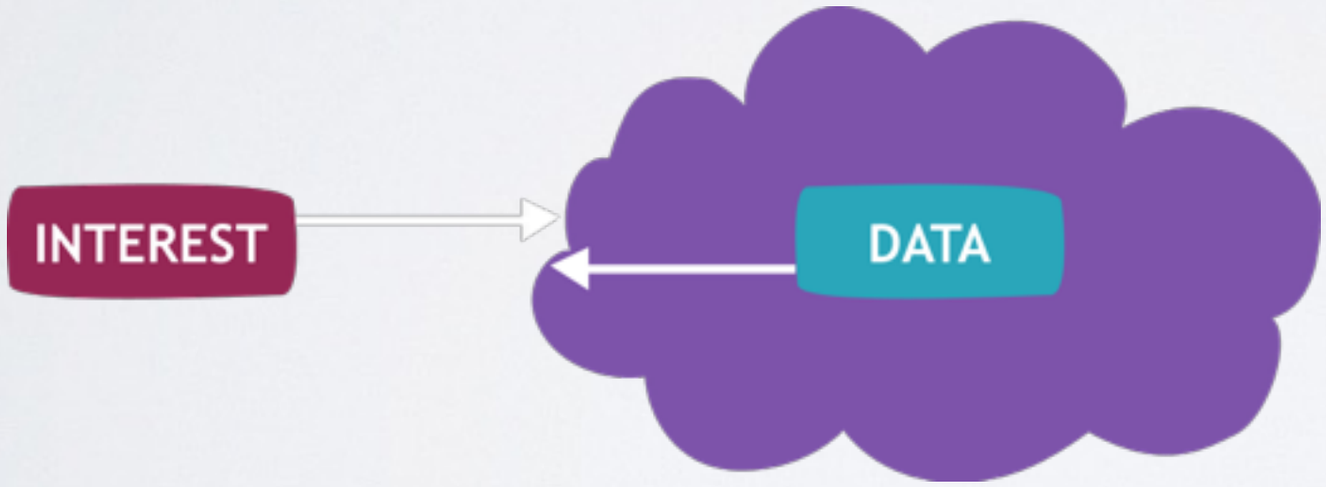
EVOLUTION OF COMMUNICATIONS



Telephone Network:
Focus: building the wires



Internet Protocol (RFC791): Focus:
deliver packets to destination node



NDN: Focusing on retrieving data from the "cloud"

Learn from how the network is used today

Superset of node-to-node communication model

WHY RETHINK? IS THE NET BROKEN?

Hugely successful, but core protocols are decades old.
(And not designed to support the global Internet..)

Stimulate innovation by addressing pain points:

Improve trust and security.

Reduce complexity (and cost).

Enhance "fit" with applications.

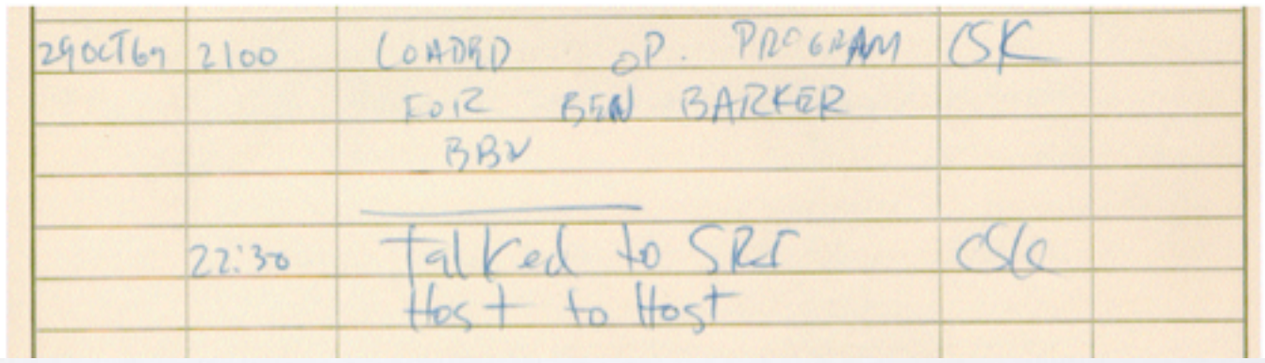
(and make it backward-compatible! think IP over leased lines, not 6to4…)

FIRST PACKET OVER THE ARPANET SENT FROM UCLA

In the Press | About | Publications | History | Twitter | Students

The Day the Infant Internet Uttered its First Words

Below is a record of the first message ever sent over the ARPANET. It took place at 22:30 hours on October 29, 1969. This record is an excerpt from the "IMP Log" that was kept at UCLA. Professor Kleinrock was supervising his student/programmer Charley Kline (CSK) and they set up a message transmission to go from the UCLA SDS Sigma 7 Host computer to another programmer, Bill Duvall, at the SRI SDS 940 Host computer. The transmission itself was simply to "login" to SRI from UCLA. They succeeded in transmitting the "l" and the "o" and then the system crashed! Hence, the first message on the Internet was **"lo",** as in "lo and behold! They were able to do the full login about an hour later.

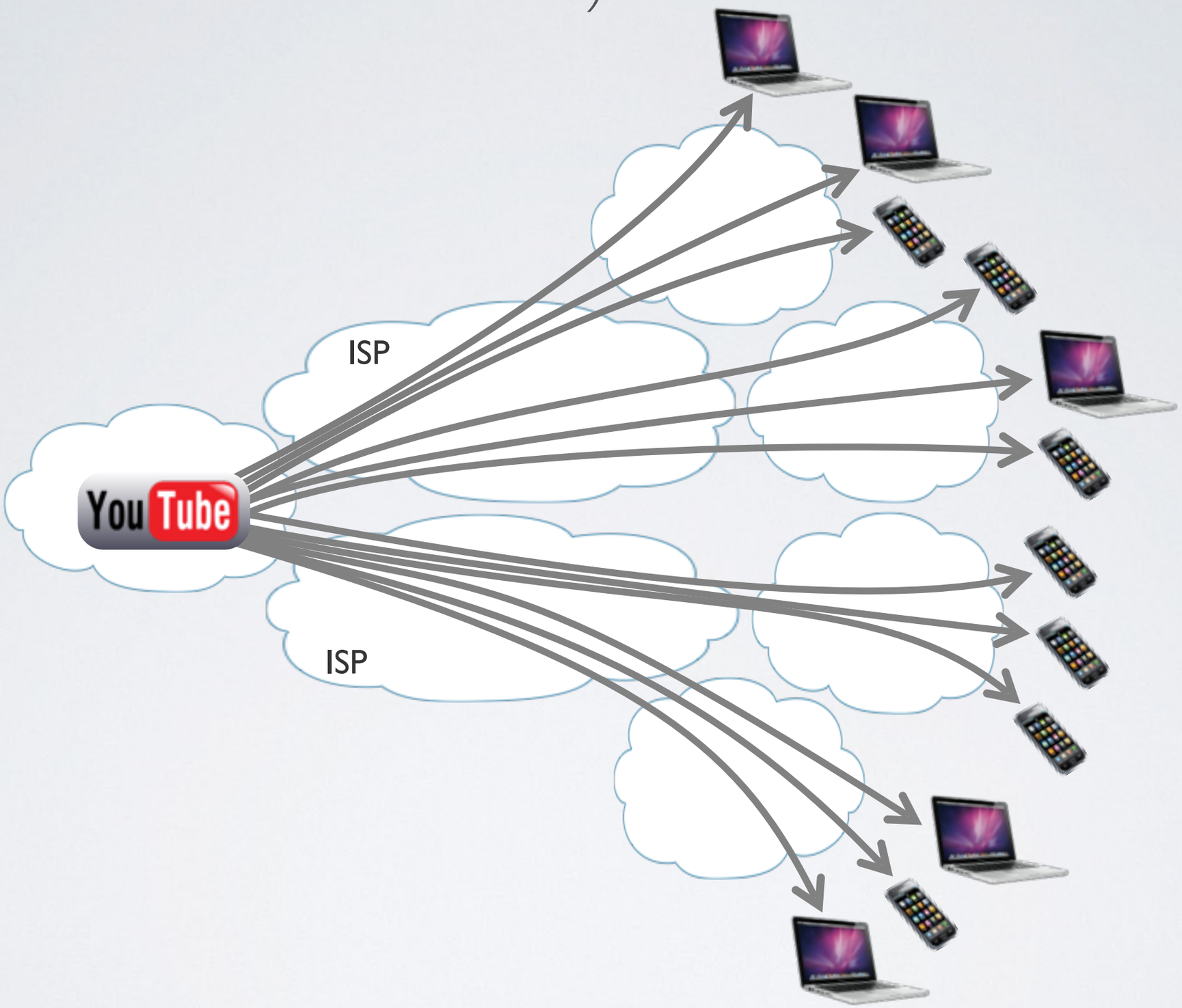


Professor Leonard Kleinrock, Department of Computer Science - http://www.lk.cs.ucla.edu/internet_first_words.html

40 YEARS LATER

178M Views

(Cost->pressure for consolidation)

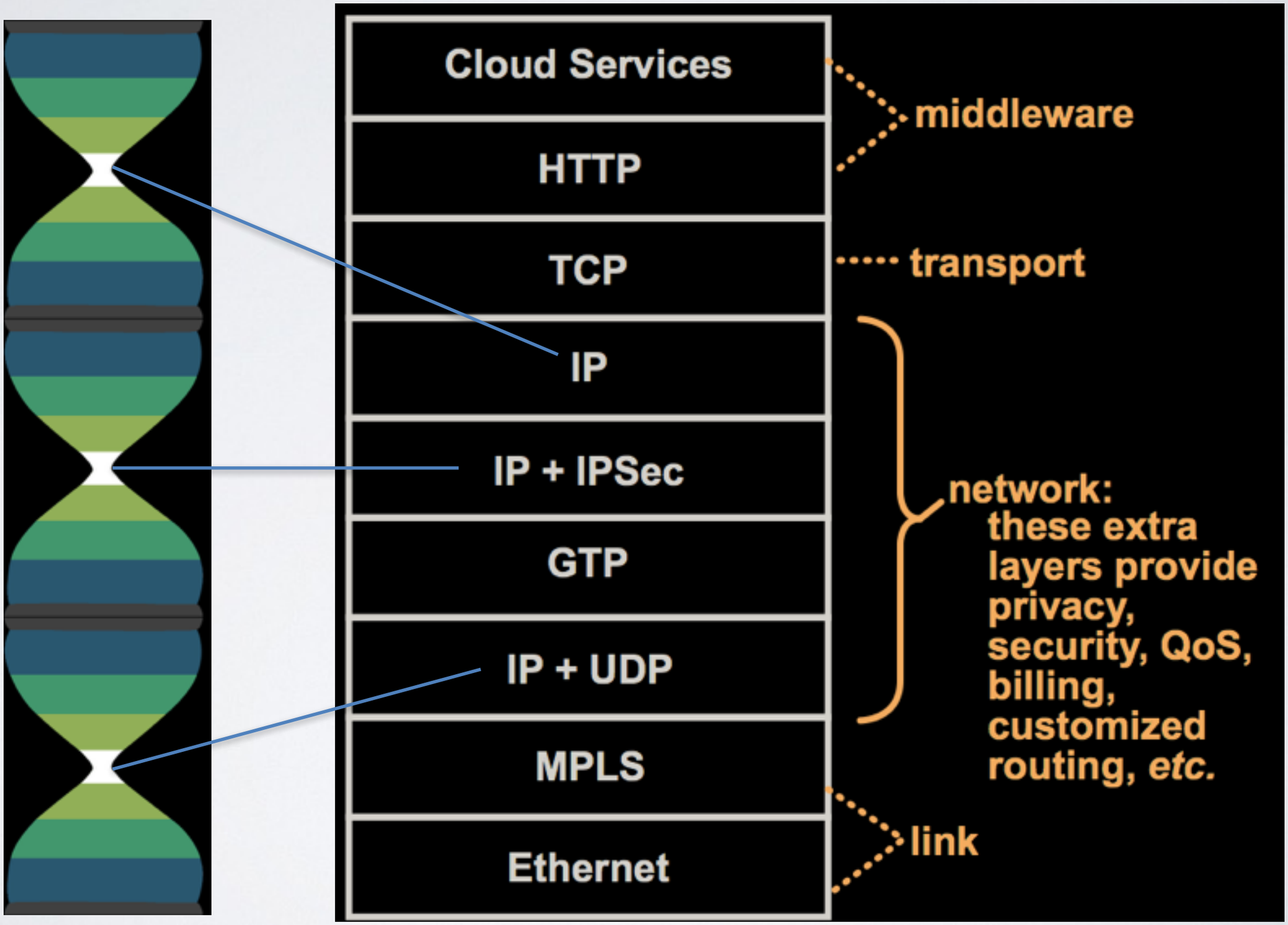


“edgy” data: IoT

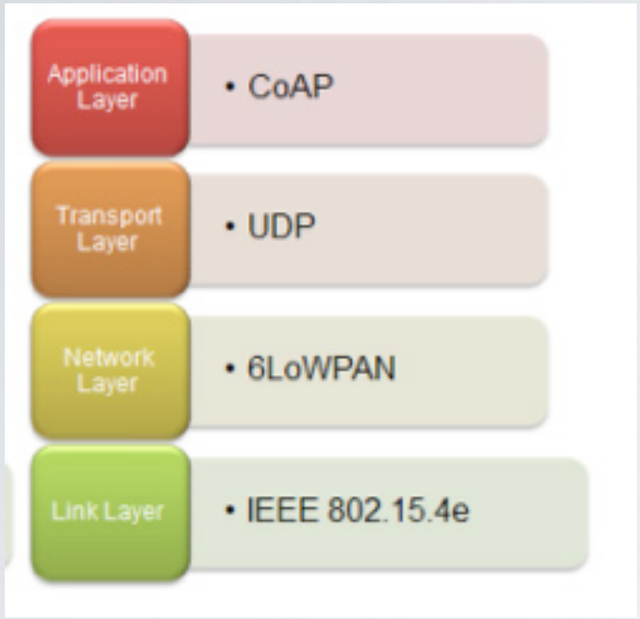
every electrical device
in your home/person..

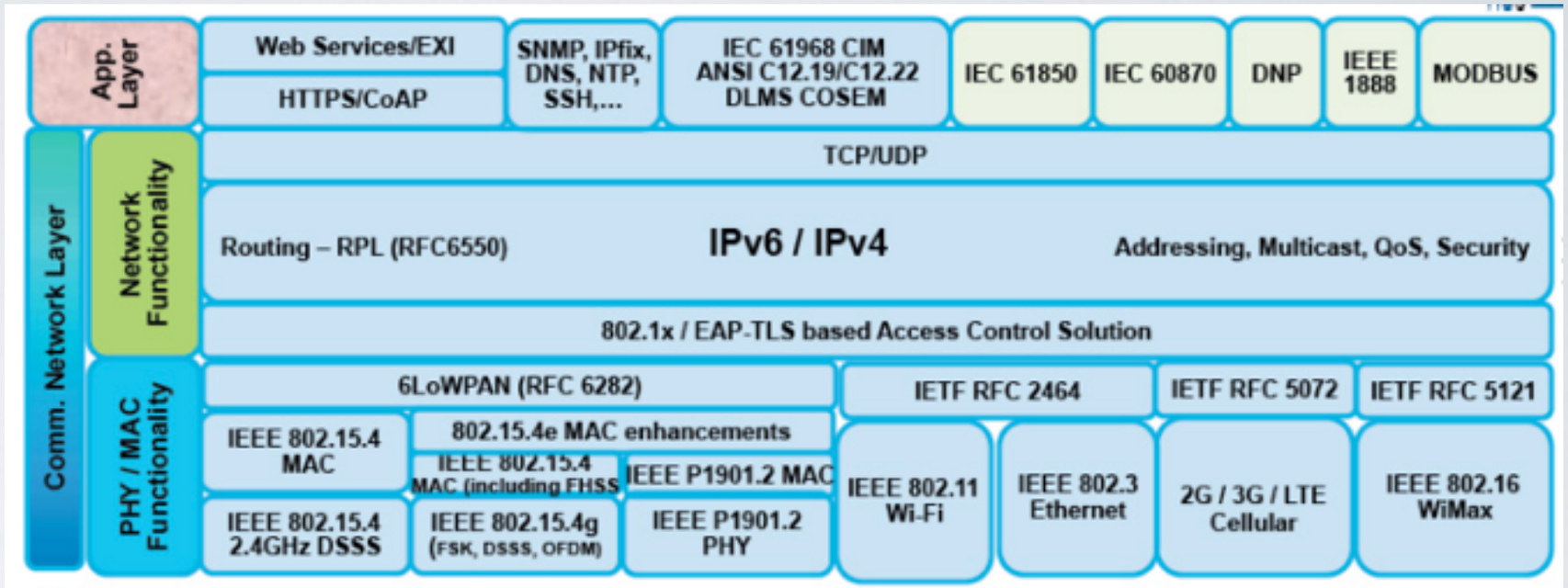
IP STACK IN THE WILD

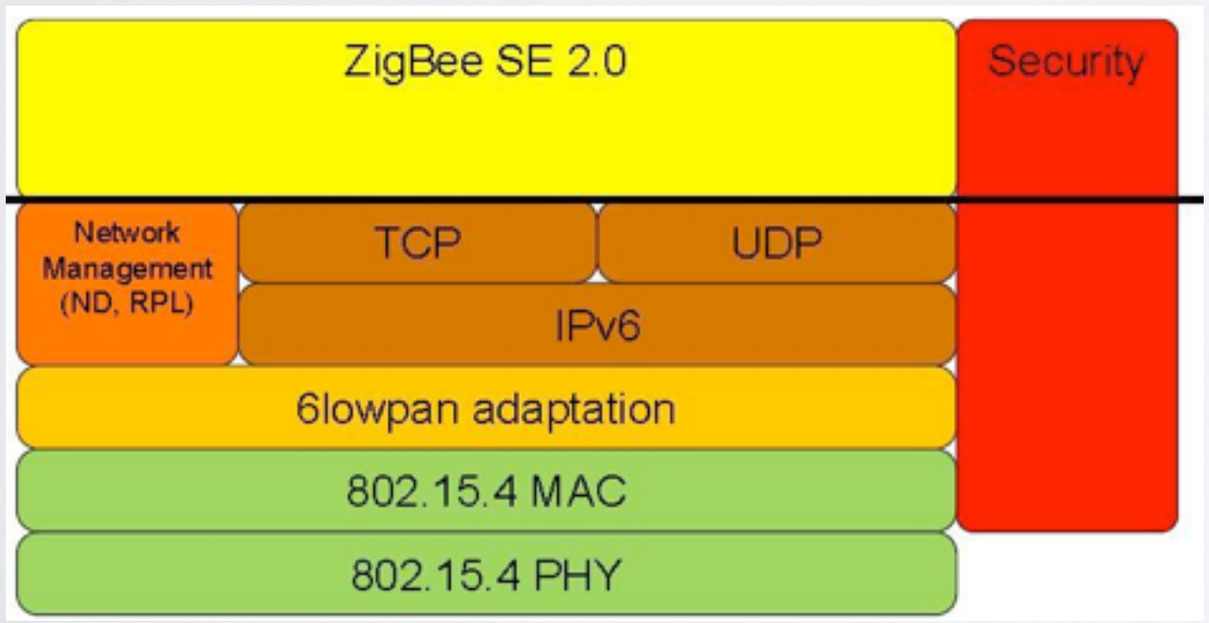
"A typical real packet (simplified)" – Pamela Zave, ATT Research, 2012

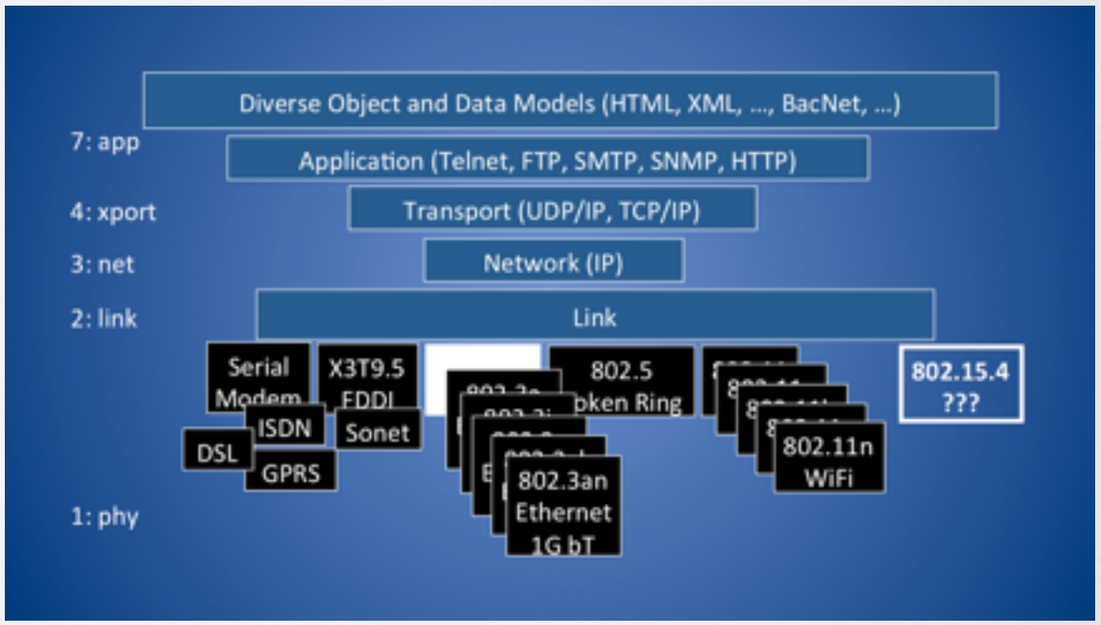


INTERNET OF THINGS "STACKS"









THE "MIDDLE" (CLOUD, CDN, ACCESS PROVIDERS)

- communication requires connectivity to centralized infrastructure
- hostile to ad hoc, DTN, P2P, intermittency
- 50%+ of population has no infrastructure
- other issues: energy consumption, privacy, vulnerability, delay, etc.

WHAT ARE OUR OPTIONS?

Continue status quo (i.e. incremental patches to TCP/IP)

- Number and scale of problems escalate
- Number of patches grows accordingly
- Ever-increasing complexity breeds problems, impedes innovation

Consider a new architecture, based on lessons learned

- New communication model: data distribution
- New security model: secure data not channel
- As a result: new application development model

ARCHITECTURAL MISMATCH

	Communication	Distribution
Naming	Endpoints	Stuff
Memory	Invisible, Limited	Explicit; Storage and wires equivalent
Security	Secure the process	Secure the stuff

(What would an architecture that supports
end-to-end communication
as a *special case* of distribution look like?)

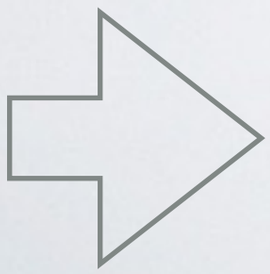
NEW COMMUNICATION MODEL

Network ships **data**, focal point of the architecture.

Network ships bits it knows are are needed.

In-network storage = bandwidth in serving content

Multicast delivery: move from point-to-point connection to multipoint synchronization



Yields efficiency and resiliency

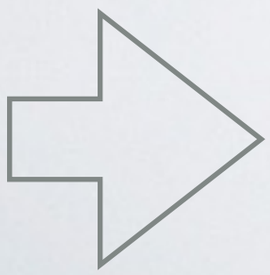
NEW SECURITY MODEL

Move security from container/channel to data itself.

Every piece of data contains the signature generated by the data producer to bind the content and the name

(Sensitive content are encrypted, can be stored in untrusted storage & delivered over unsecured channel.)

Hierarchical name provides context for trust management



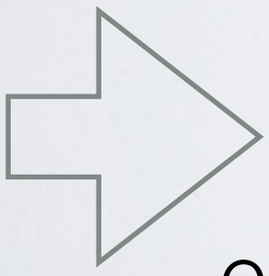
Ultimate end-to-end security: between data producer and consumer (not of channels)

NEW APPLICATION DEV MODEL

Focus on managing your data
Security model built in from beginning

- Developers select (or create) security model for trust management (key & confidentiality management)

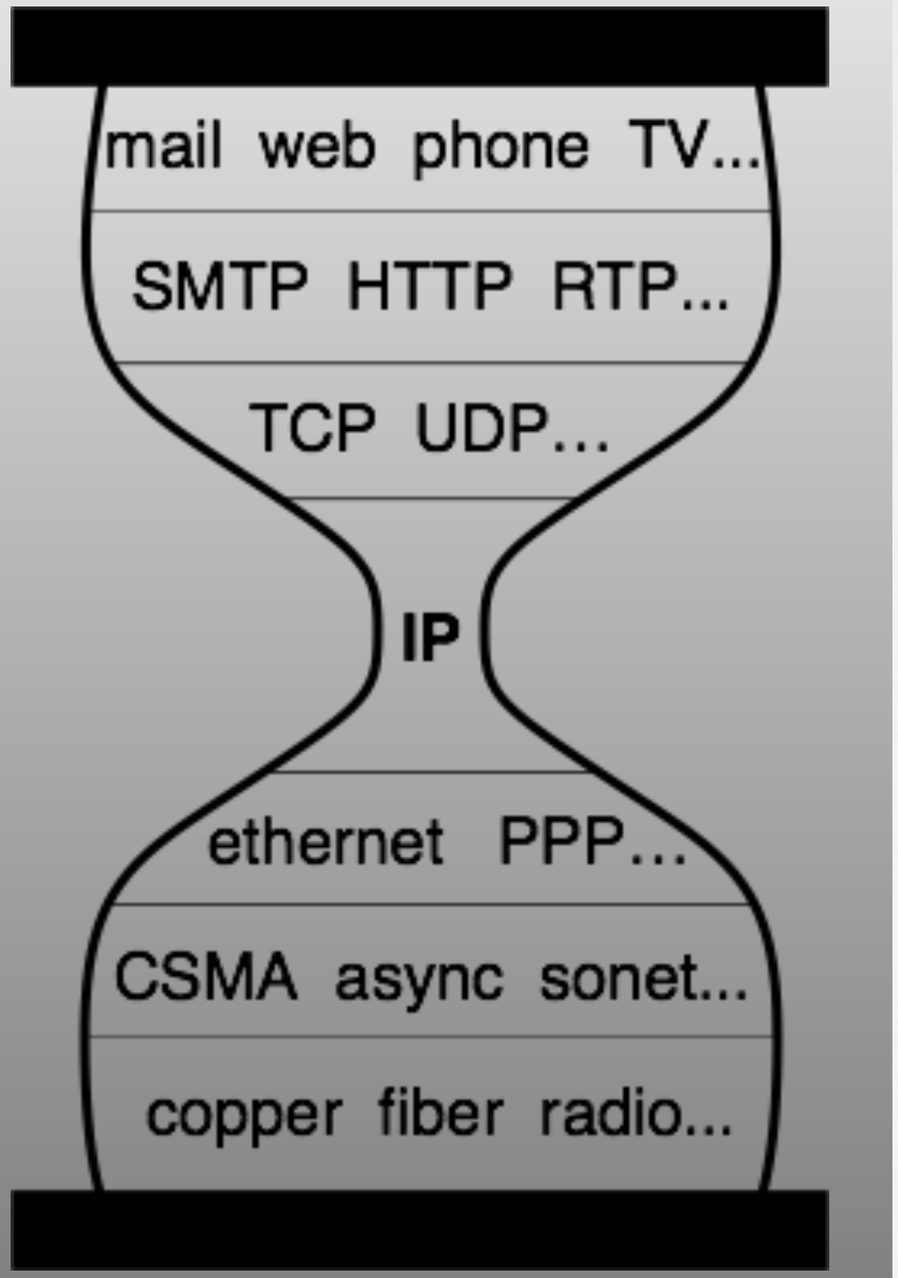
NDN is developing security tools and conventional models from pilot applications

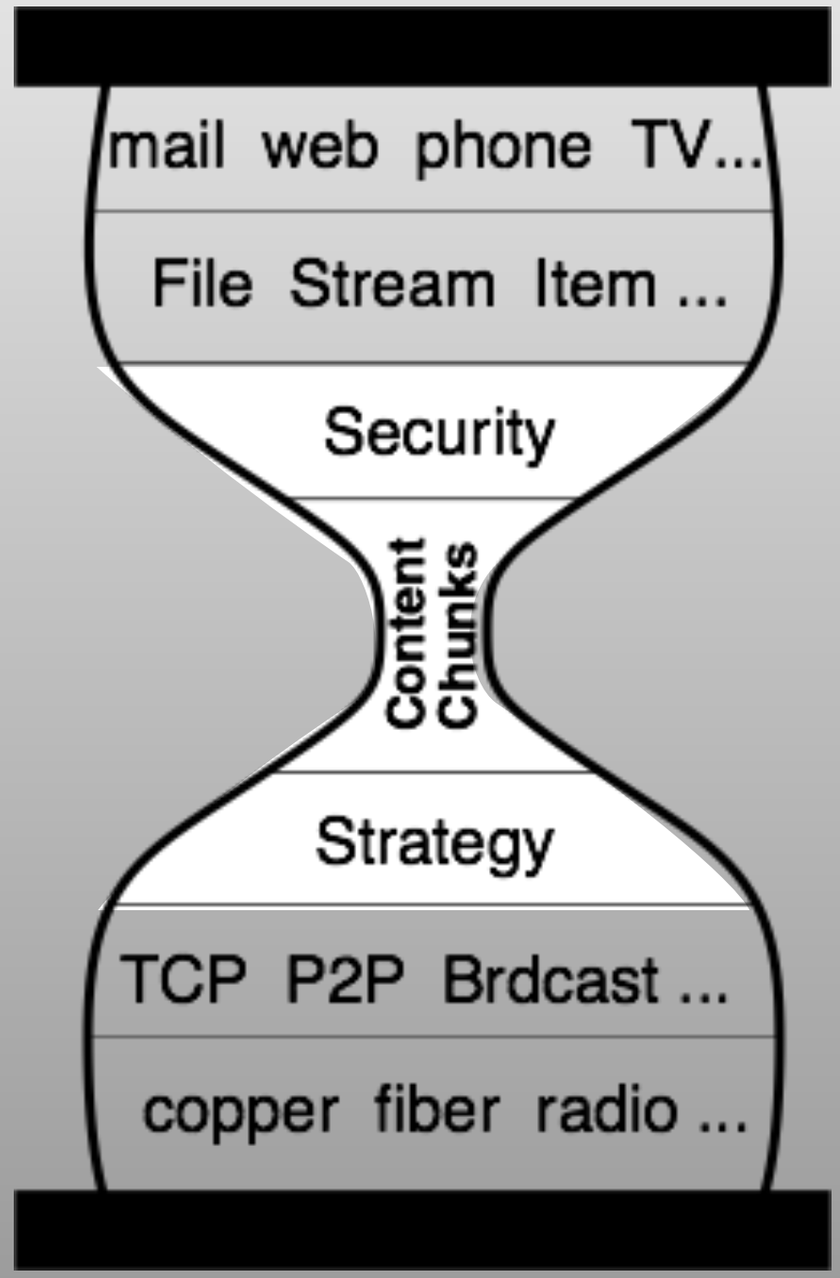


No longer worry about network details

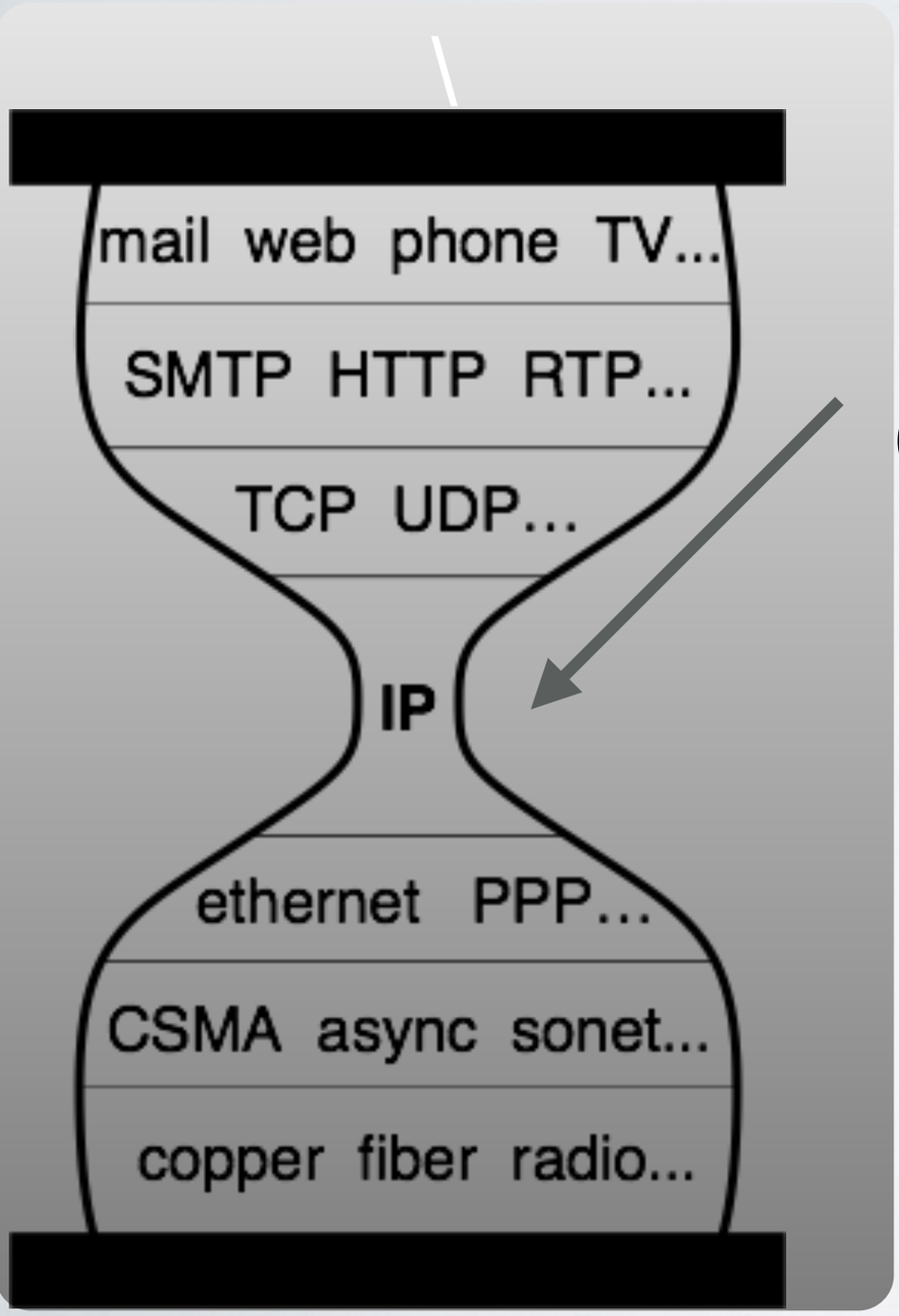
e.g., which server to use, which servers are overloaded.

TCP/IP VS NDN STACK



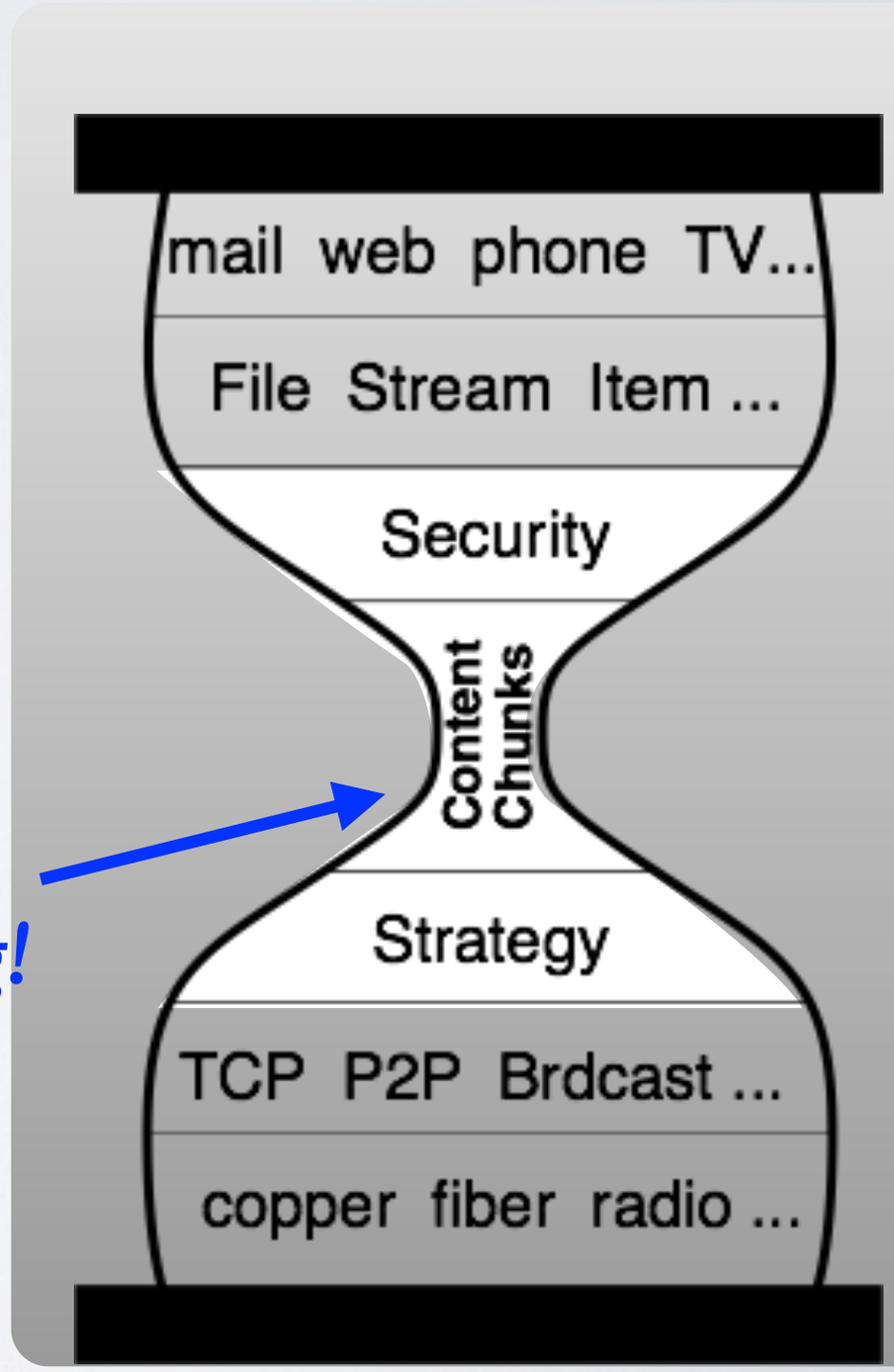


A GENERALIZATION OF IP

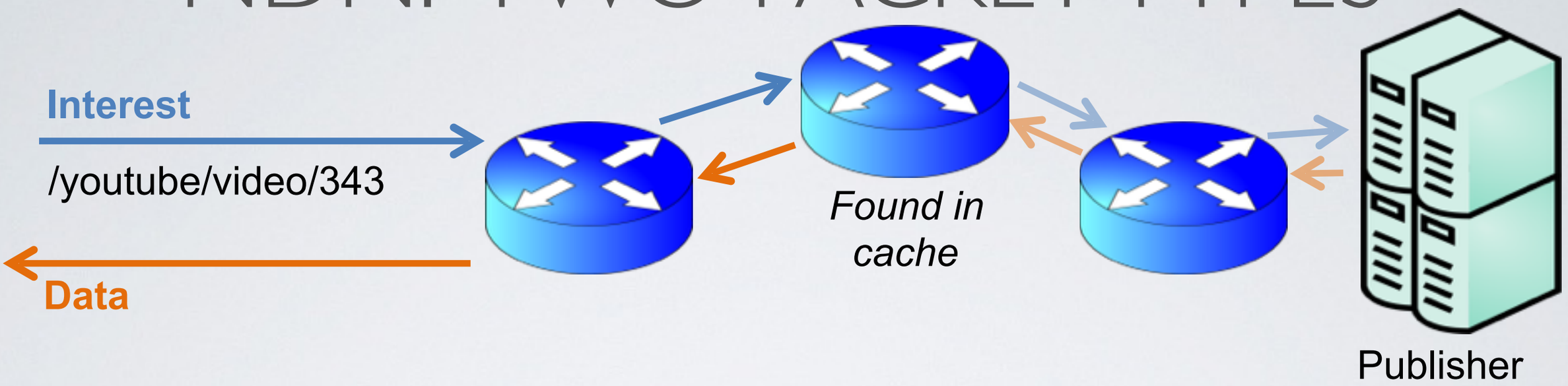


names
endpoints
(IP address)

names anything!



NDN: TWO PACKET TYPES

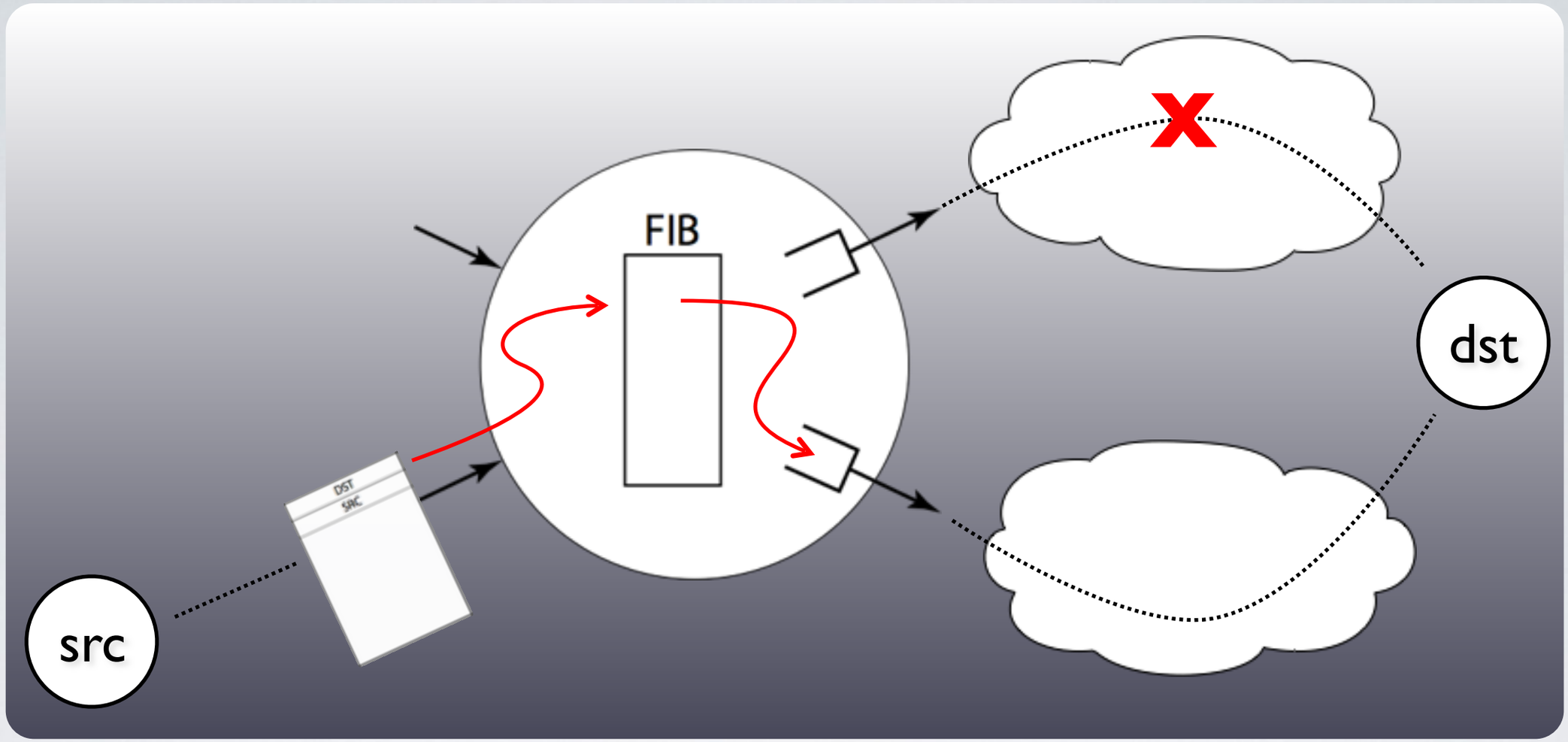


Interest Packet:

Content Name
Selectors (order preference, publisher filter, exclude filter, …)
Nonce
Guiders (scope, Interest lifetime)

Data Packet:

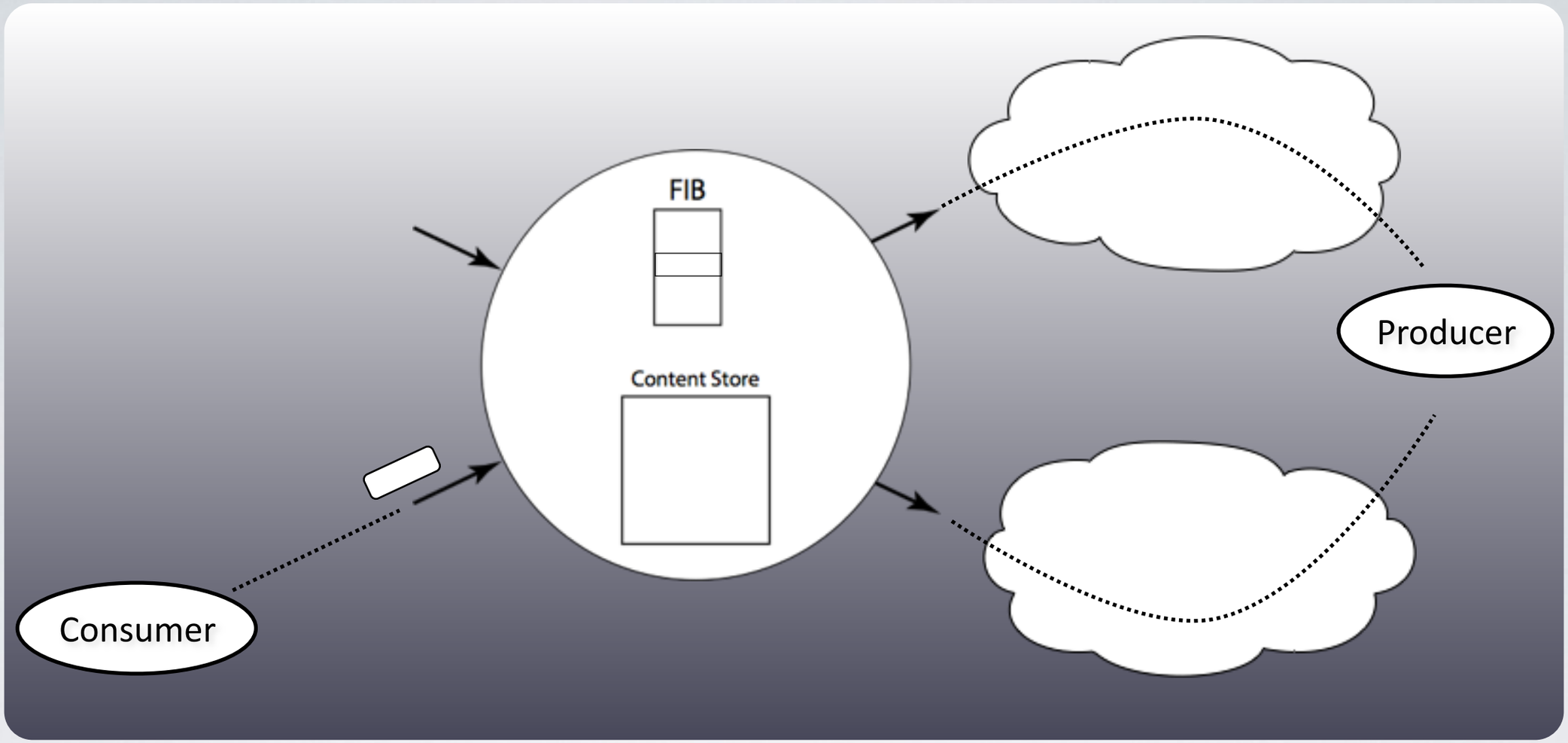
Name
MetaInfo (content type, freshness period, …)
Content
Signature



moving content around in a
TCP/IP architecture

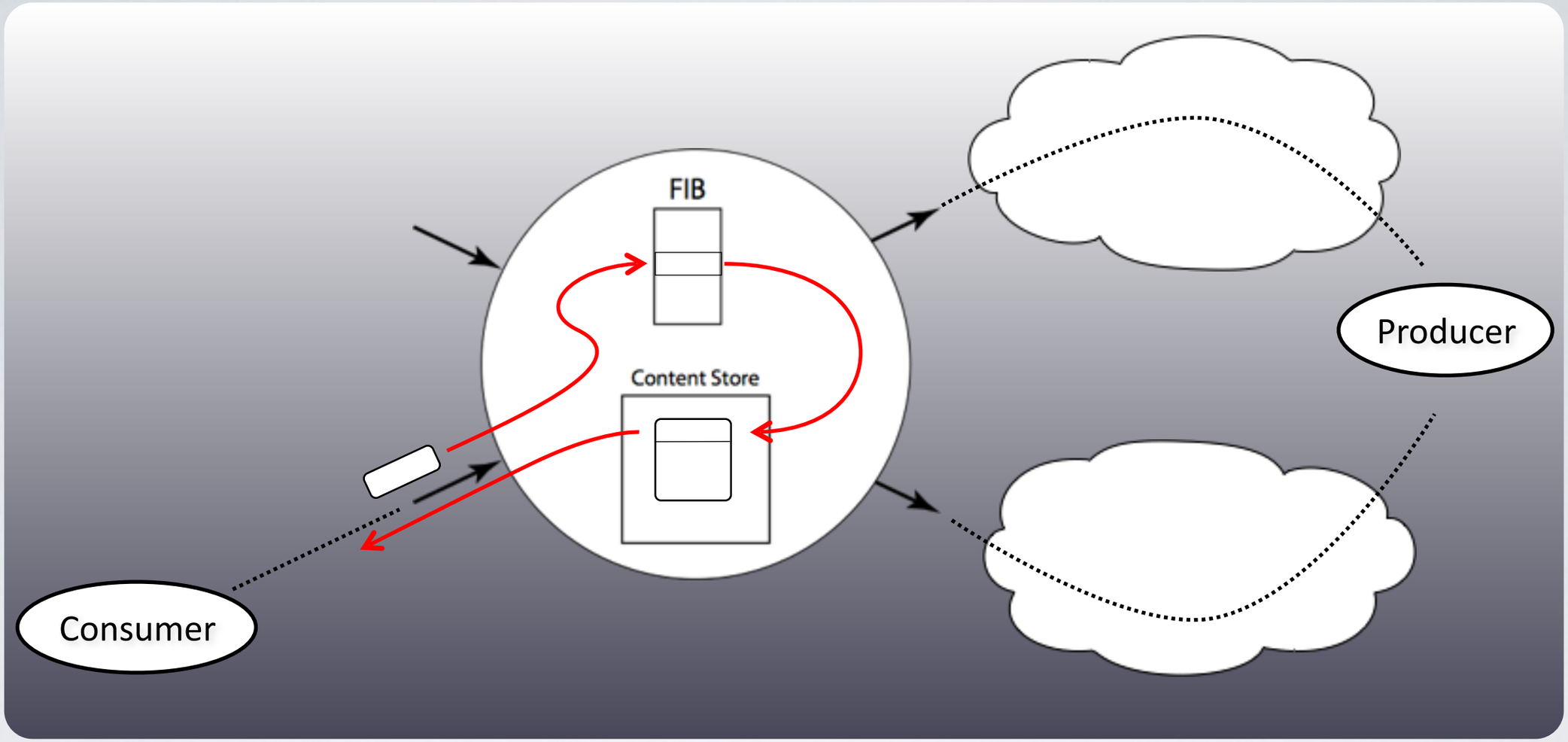
Path determined by global routing, not local choice.

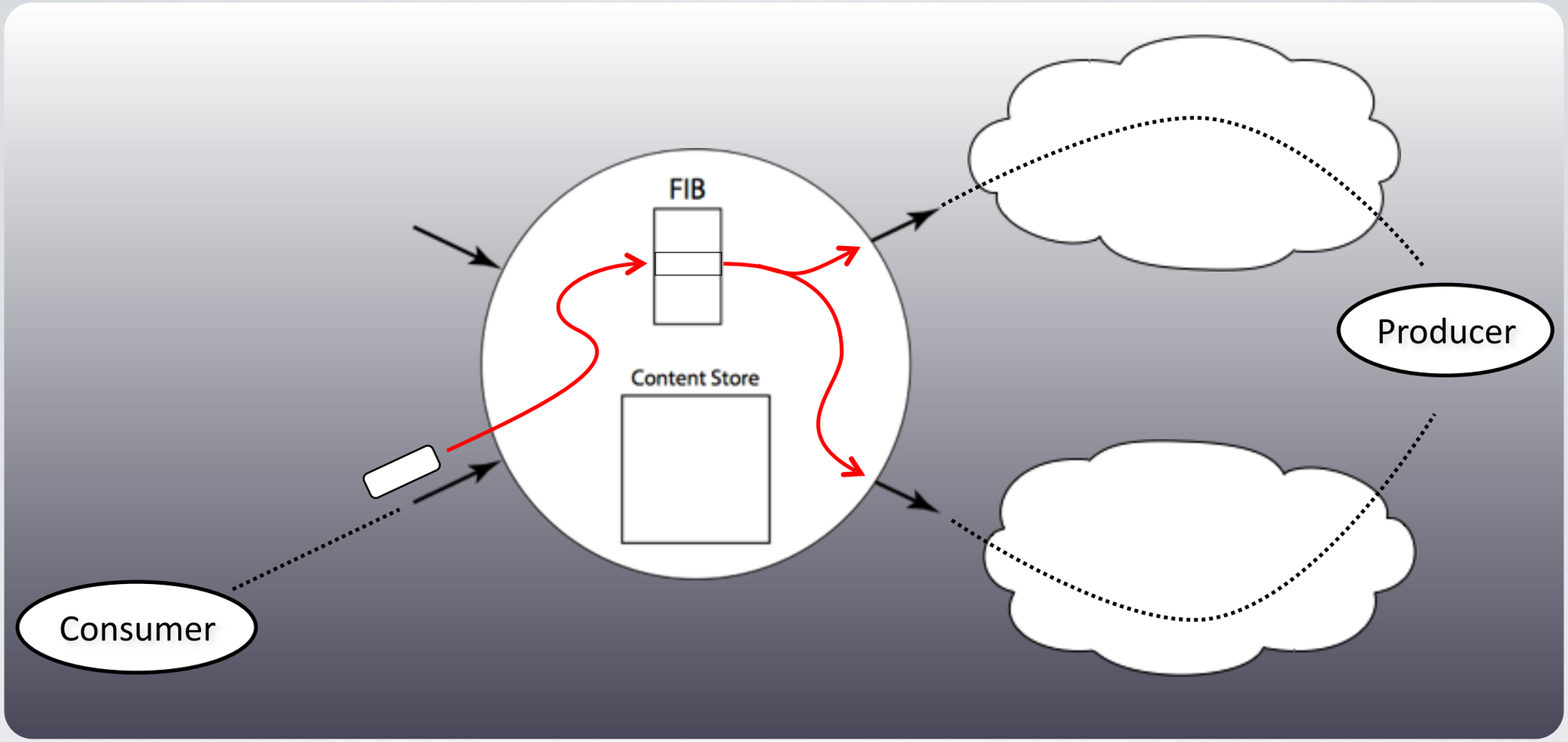
Structural asymmetry precludes market mechanisms
and encourages monopoly formation.

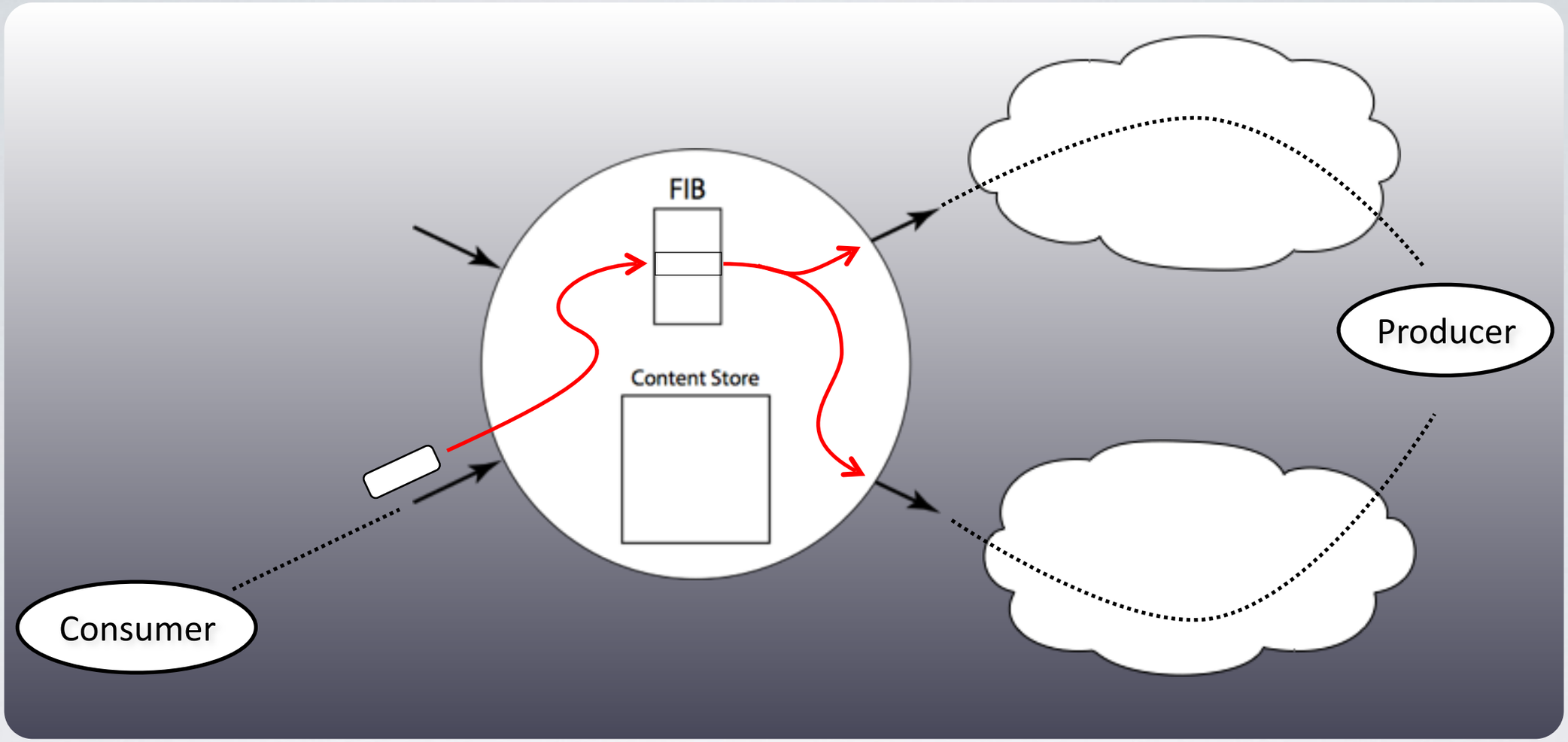


moving content around in an information-centric architecture

- requires Interest to trigger data transmission (no unsolicited)
- data flows over reverse path as Interest (flow control)
- all data packets cryptographically signed (security)

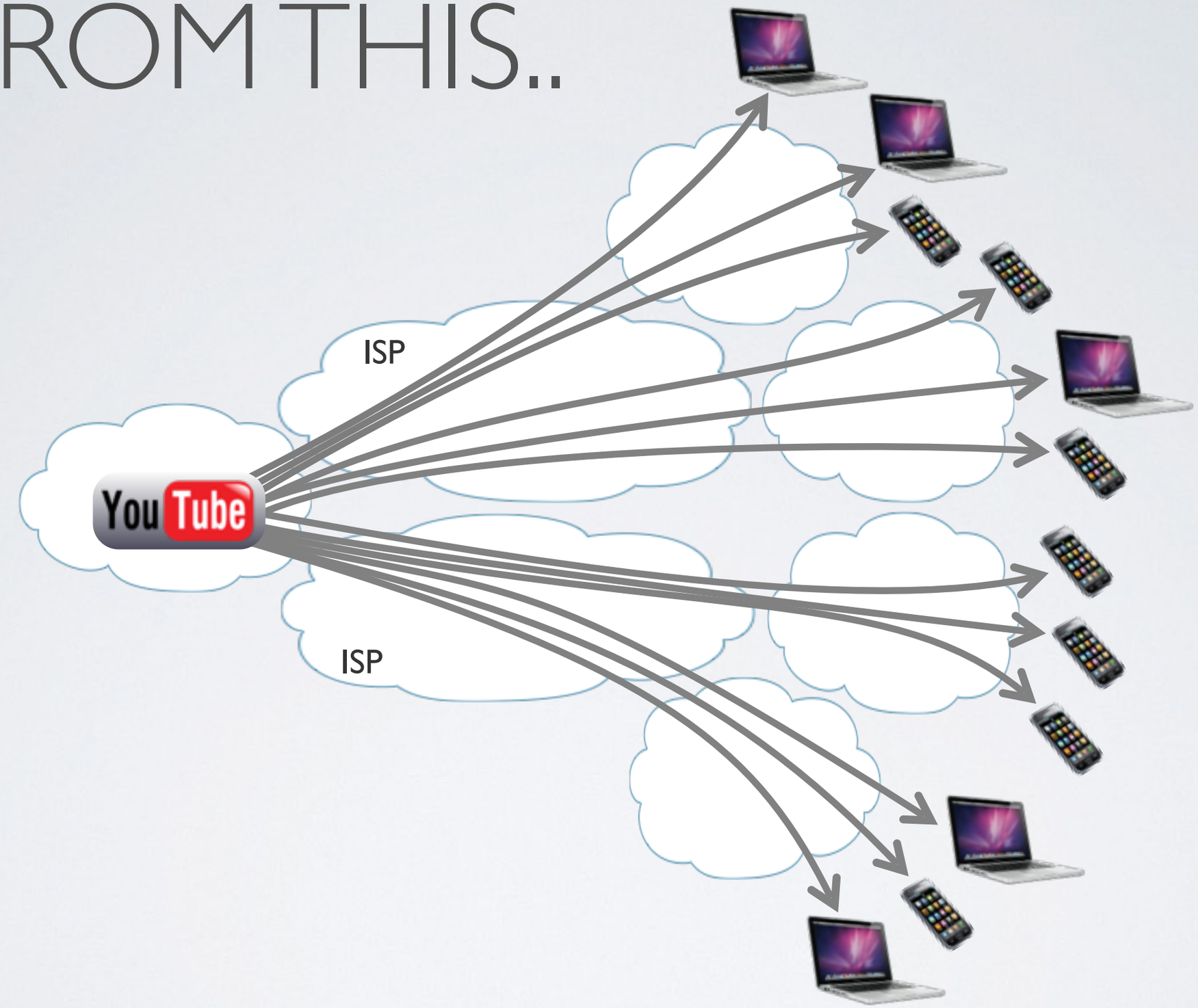

FIB
Content Store
Producer
Consumer


FIB
Content Store
Consumer
Producer

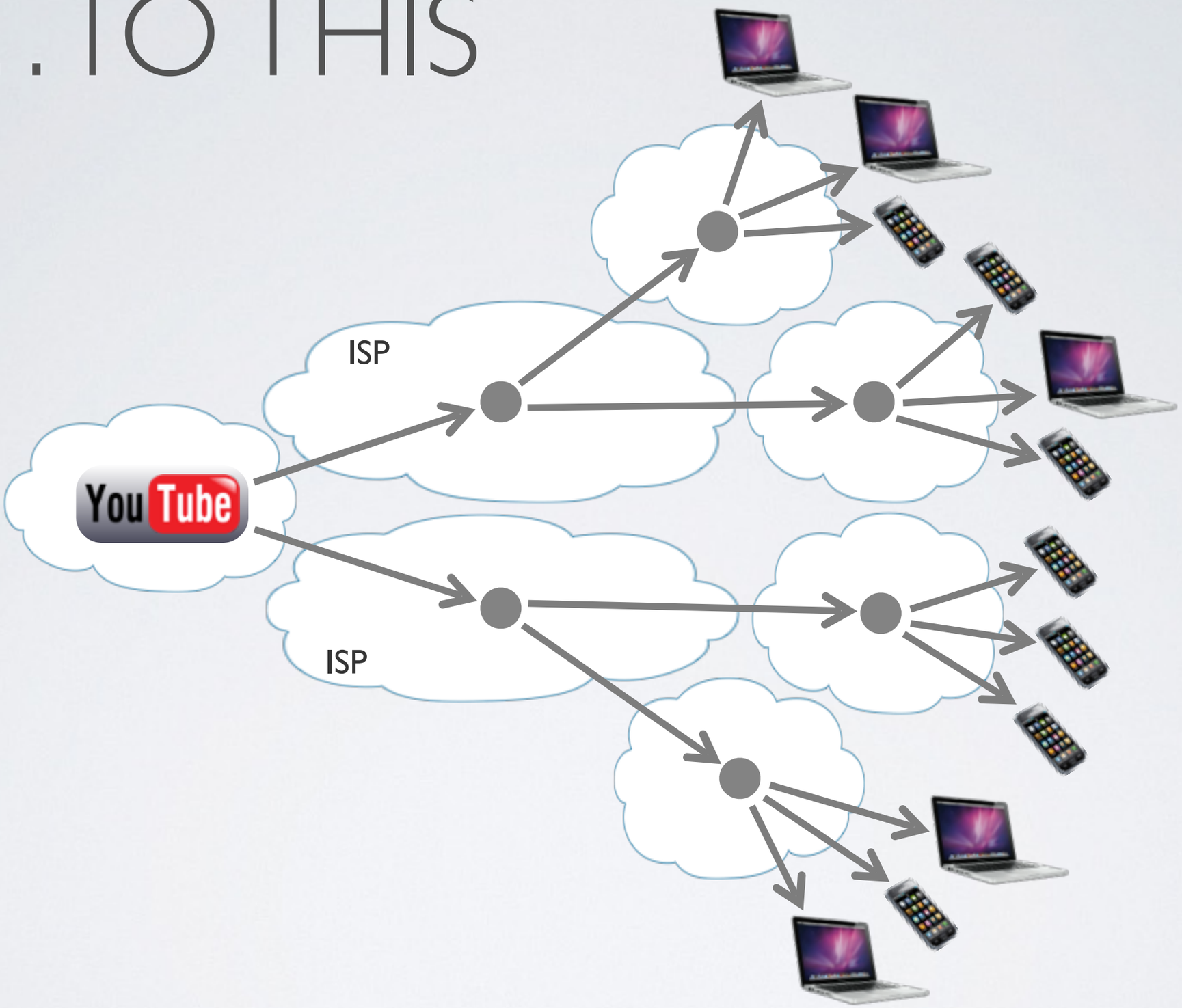


- Packets say what not who (no src or dst)
- Forwarding decision is local
- Upstream performance is measurable

FROM THIS..



…TO THIS



BUT!..

NDN » CONTENT DISTRIBUTION

There are persistent problems with Internet routing, transport and security that we have been unable to solve within IP's framework.

Recently, NDN (ICN) research efforts have begun to demonstrate credible solutions to these problems.

These solutions could make a big difference to the Internet & the World

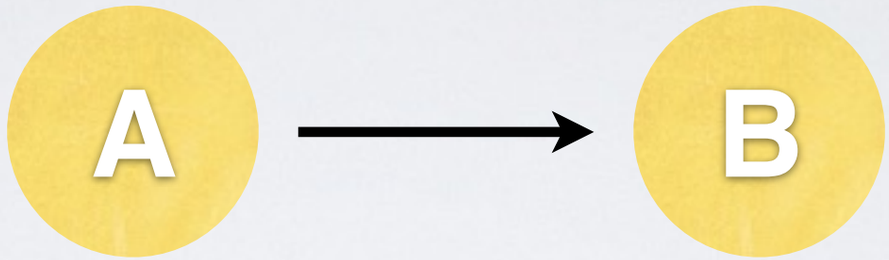
EXAMPLES (FROM NDN)

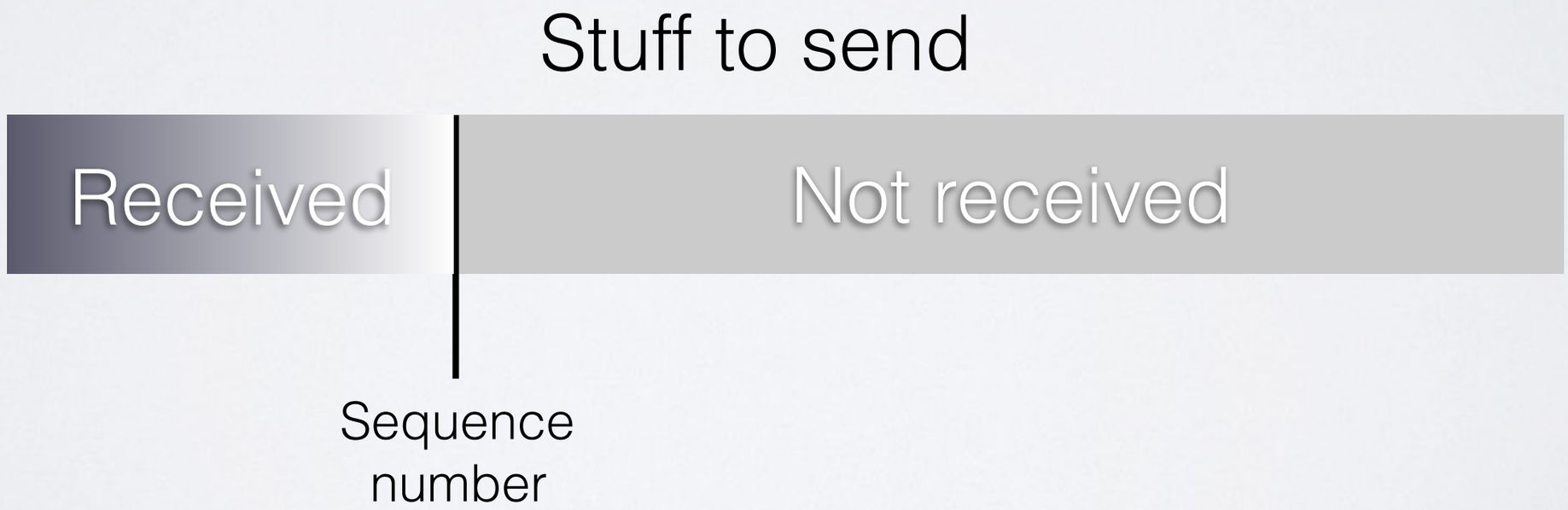
Transport via Set Reconciliation ('Sync')

Greedy Hyperbolic Geometric

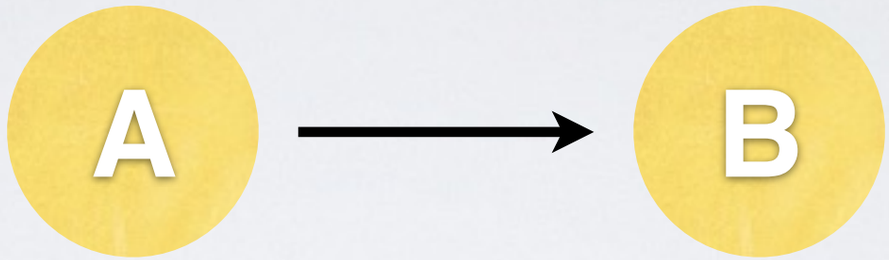
Schematized Trust Models

TRANSPORT THRU THE AGES

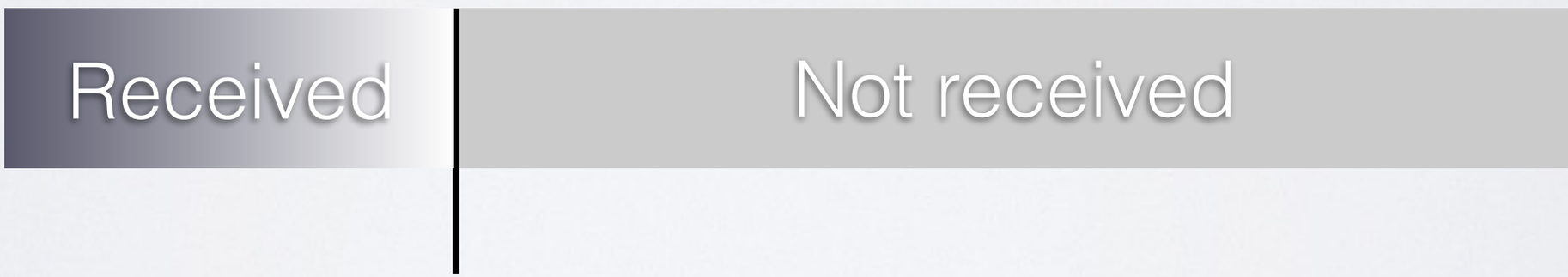




TRANSPORT THRU THE AGES

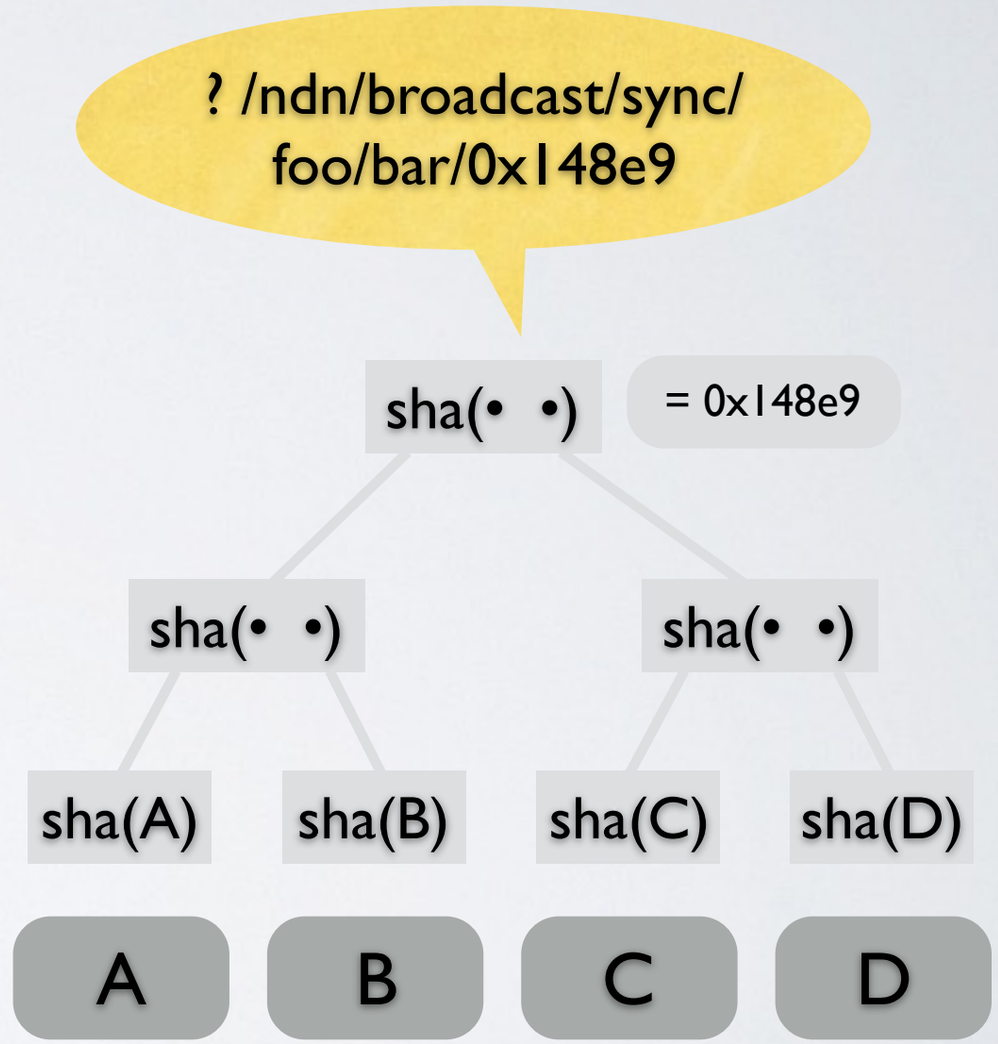


Stuff to send



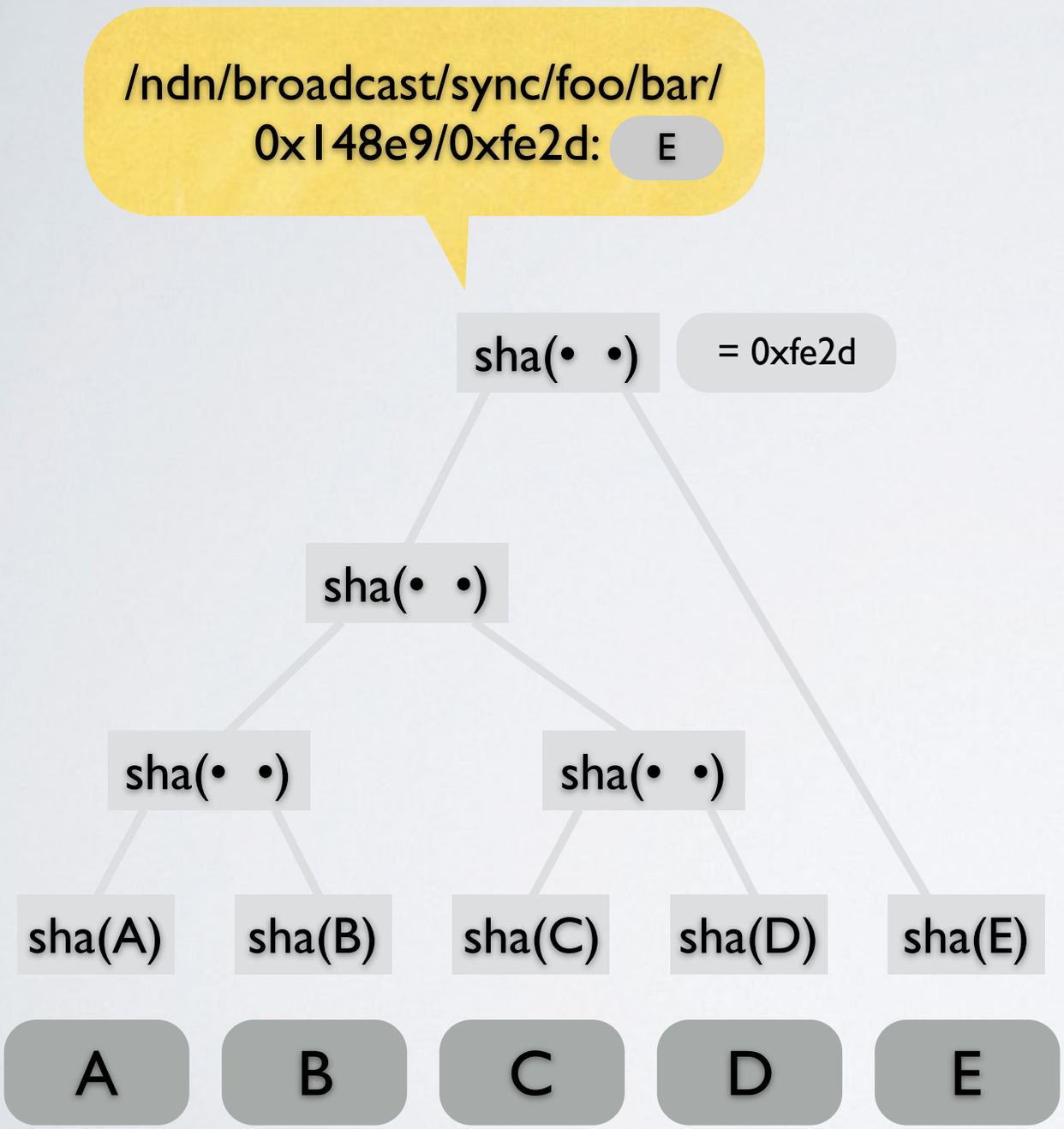
This models the process, not the outcome
(data movement is a side-effect)

A BETTER WAY

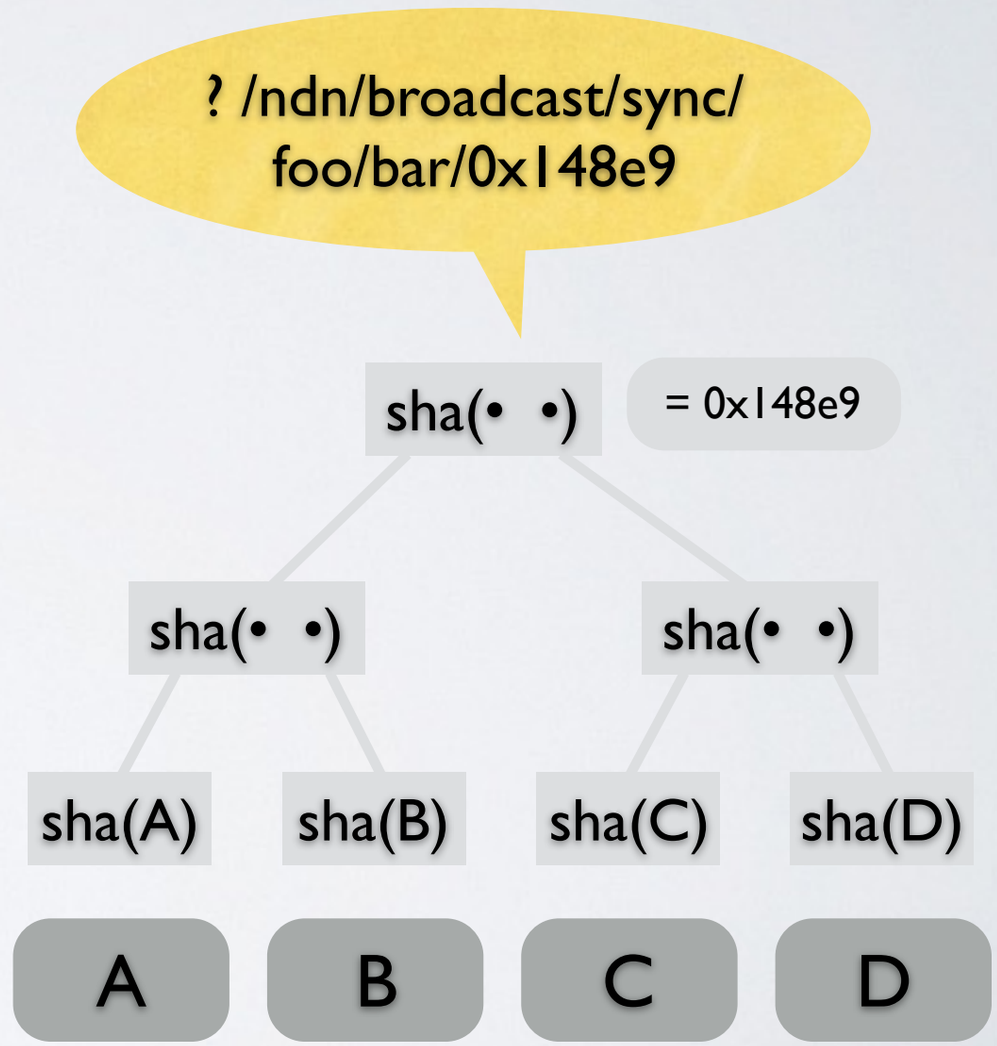


Bob's /foo/bar collection

A BETTER WAY



Alice's /foo/bar collection



Bob's /foo/bar collection

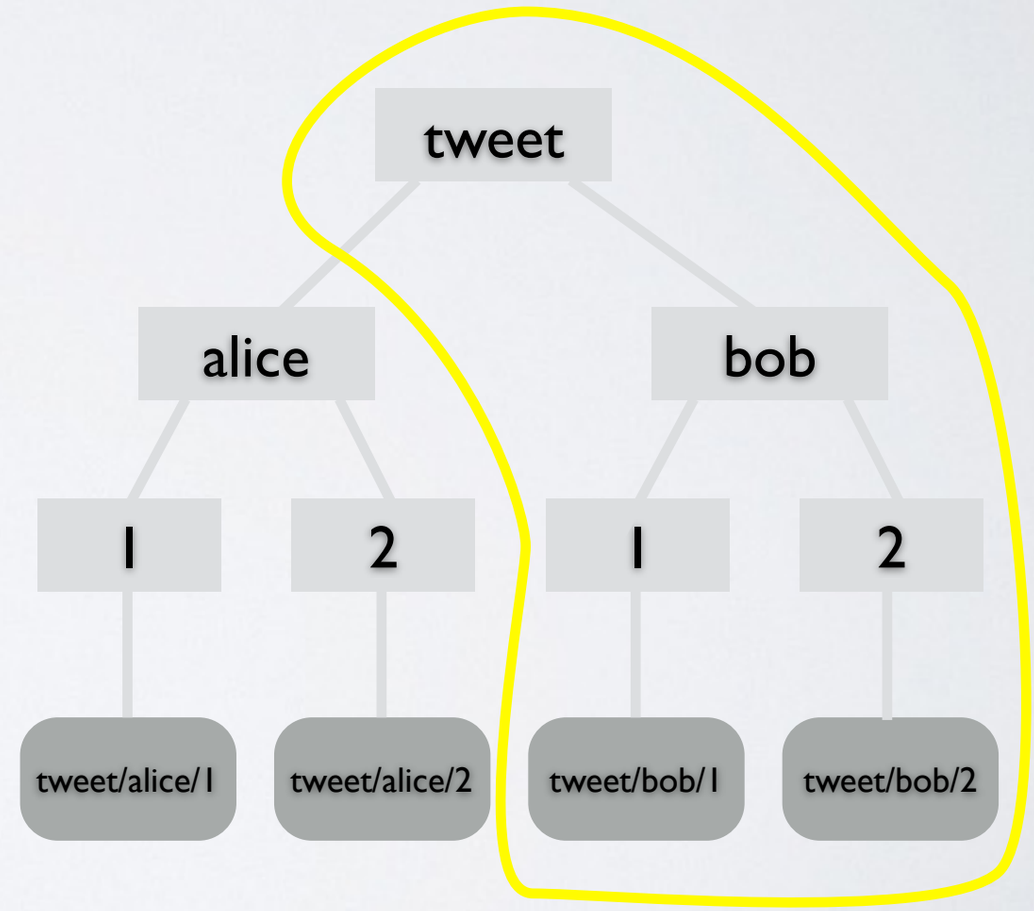
A BETTER WAY

? /broadcast/sync/
tweet/bob/0x0

[same communication cost as TCP
but much more general and robust]

Reconciliation of any two sets can be done with a communication cost proportional to their difference.

– Y.Minsky & A.Trachtenberg, IEEE Trans. on Information Theory, 49(9) 2003



Bob's tweet collection

EXAMPLES (FROM NDN):

Transport via Set Reconciliation (‘Sync’)

Greedy Hyperbolic Geometric

Schematized Trust

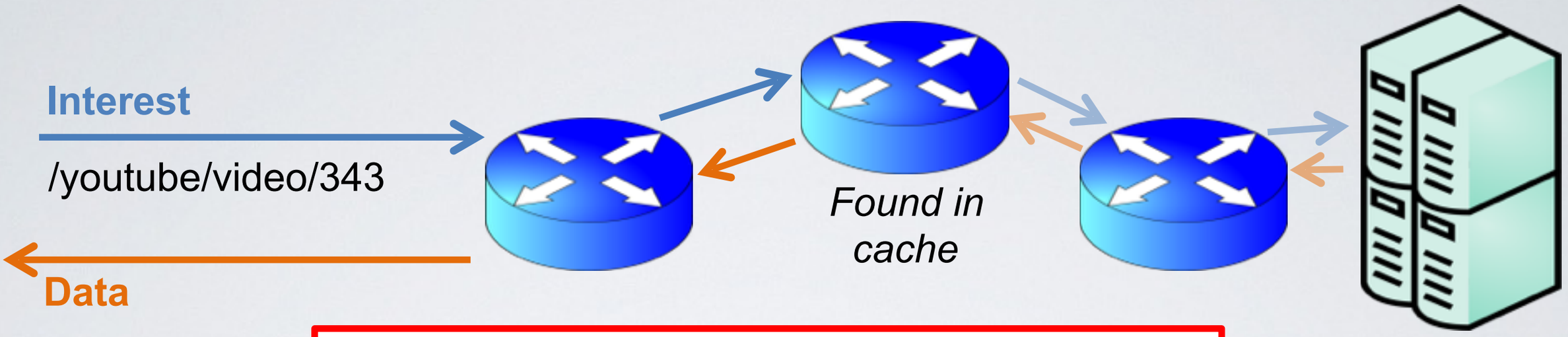
A BETTER WAY

Packet = ⟨ *name, data, signature* ⟩

Any consumer can assess *solely from the data:*

- Integrity (is data intact and complete?)
- Pertinence (is this an answer to my question?)
- Provenance (who asserts this is an answer?)

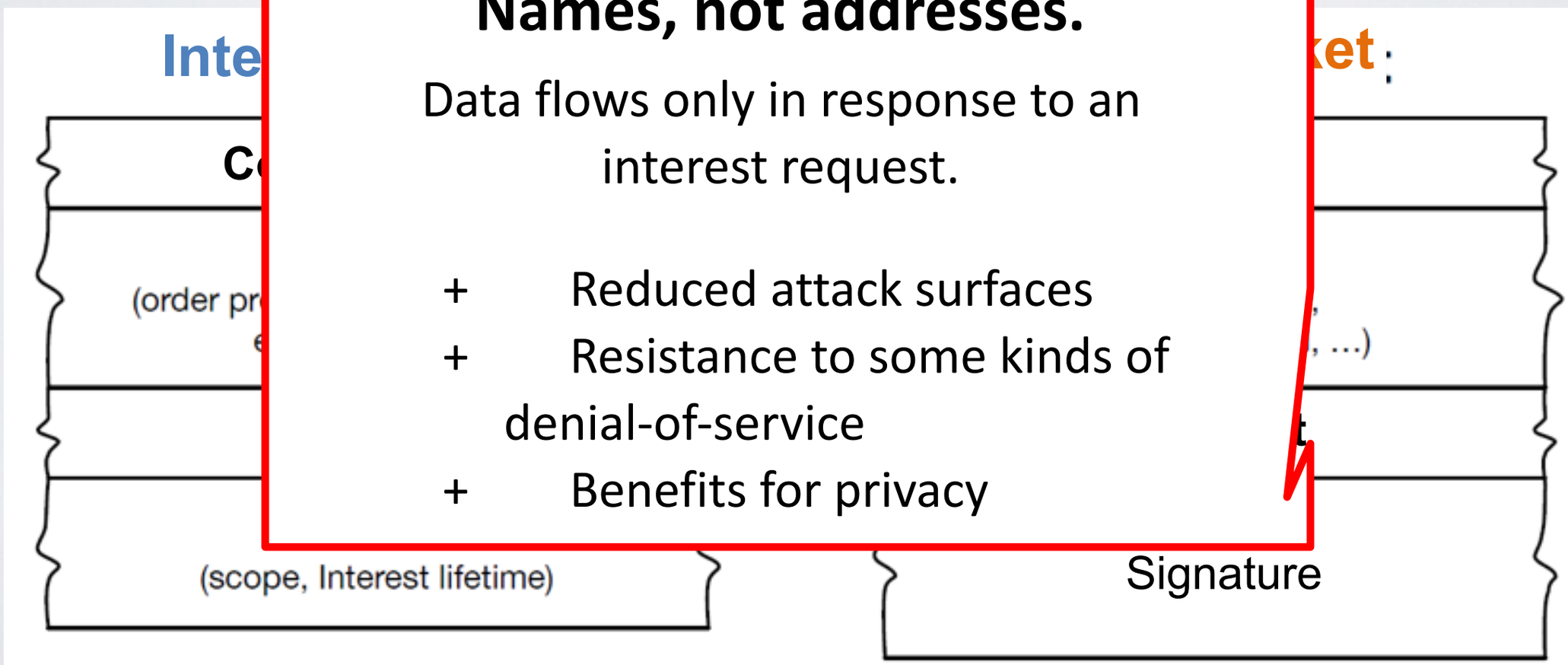
DATA-CENTRIC SECURITY



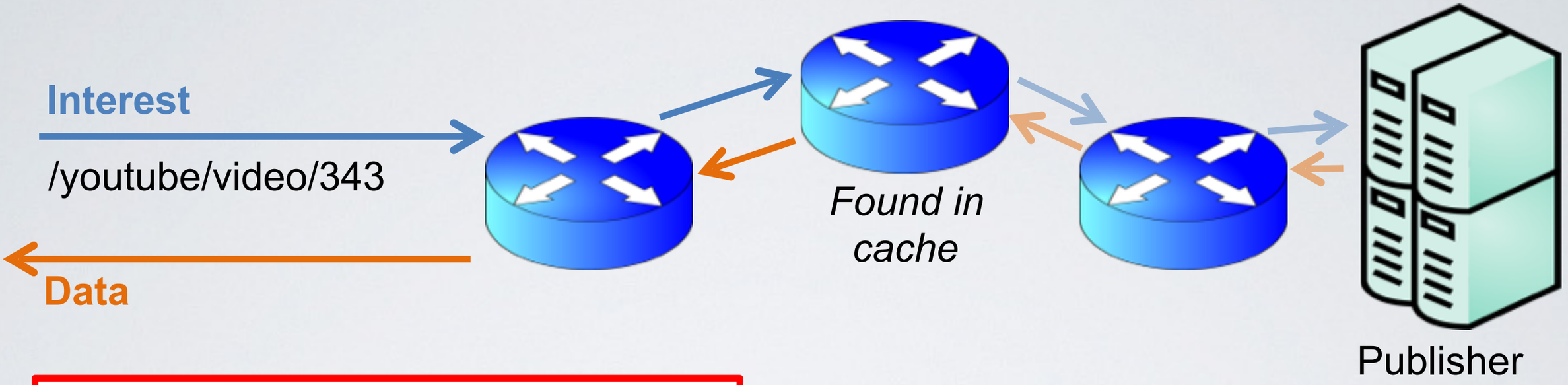
Names, not addresses.

Data flows only in response to an interest request.

+ Reduced attack surfaces
+ Resistance to some kinds of denial-of-service
+ Benefits for privacy



DATA-CENTRIC SECURITY



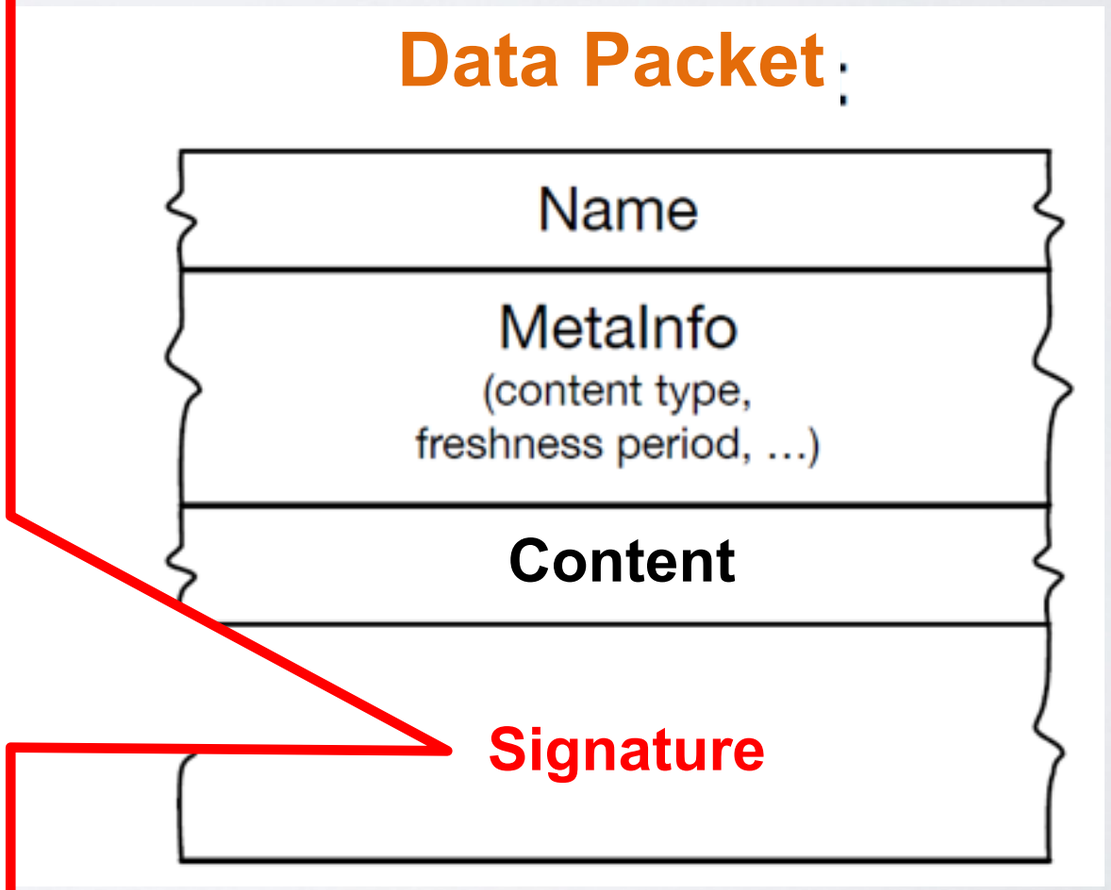
All content must be signed.

Routers may, clients shall, verify.

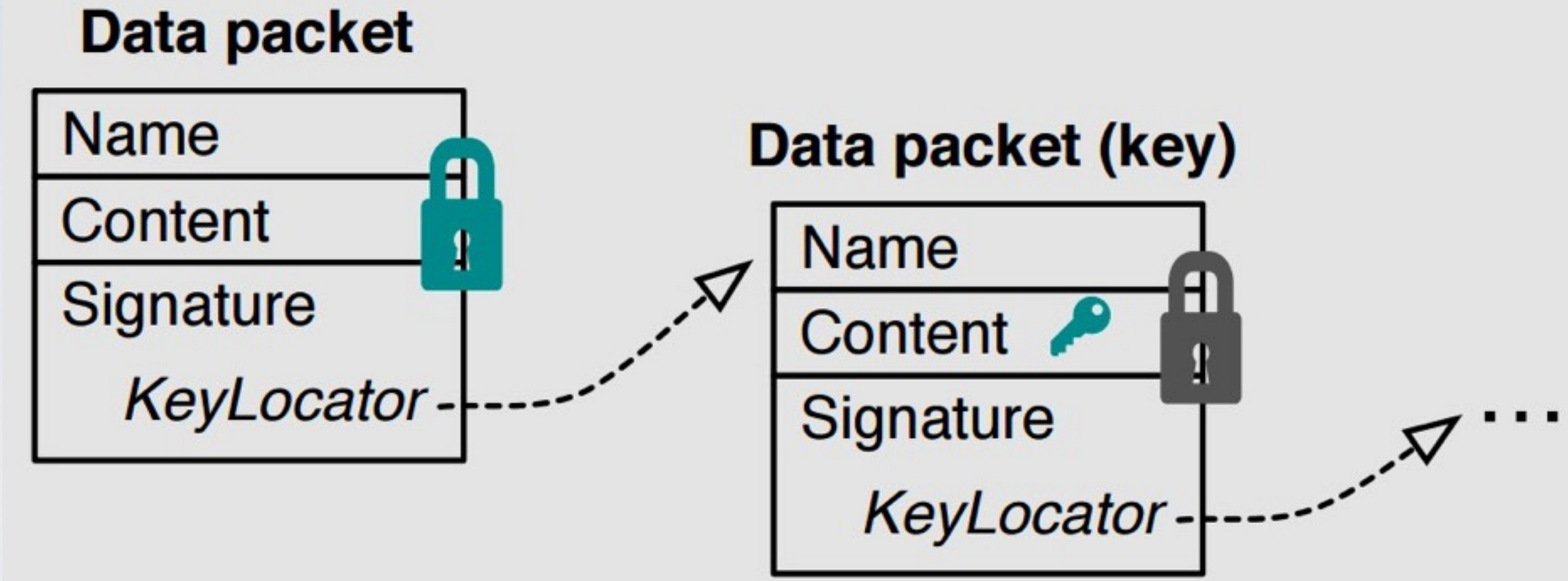
Validation policy defined by applications.

+ Flexible foundation for many security properties:

  Integrity, authentication, access control, provenance



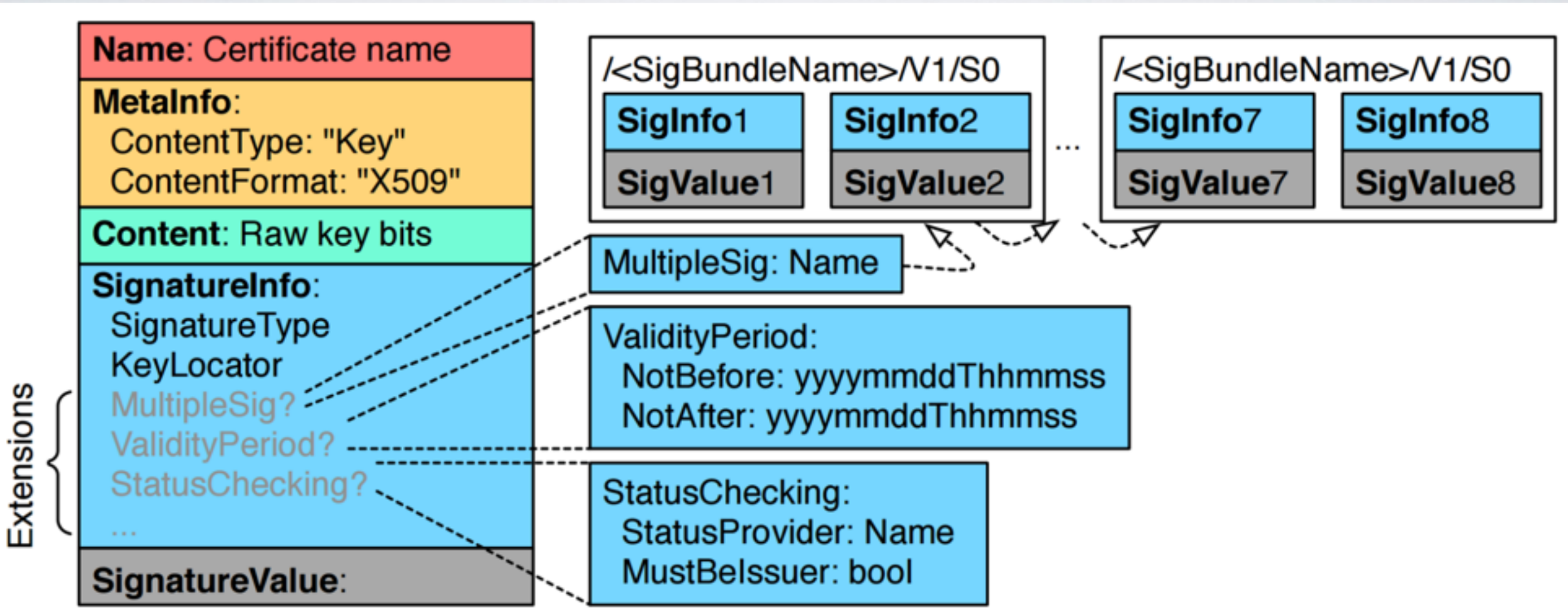
SIGNATURES IN NDN



Big idea: Certificates are just named, signed data.
Get them “for free” in the data-centric security approach.

SIGNATURE FORMAT DETAILS

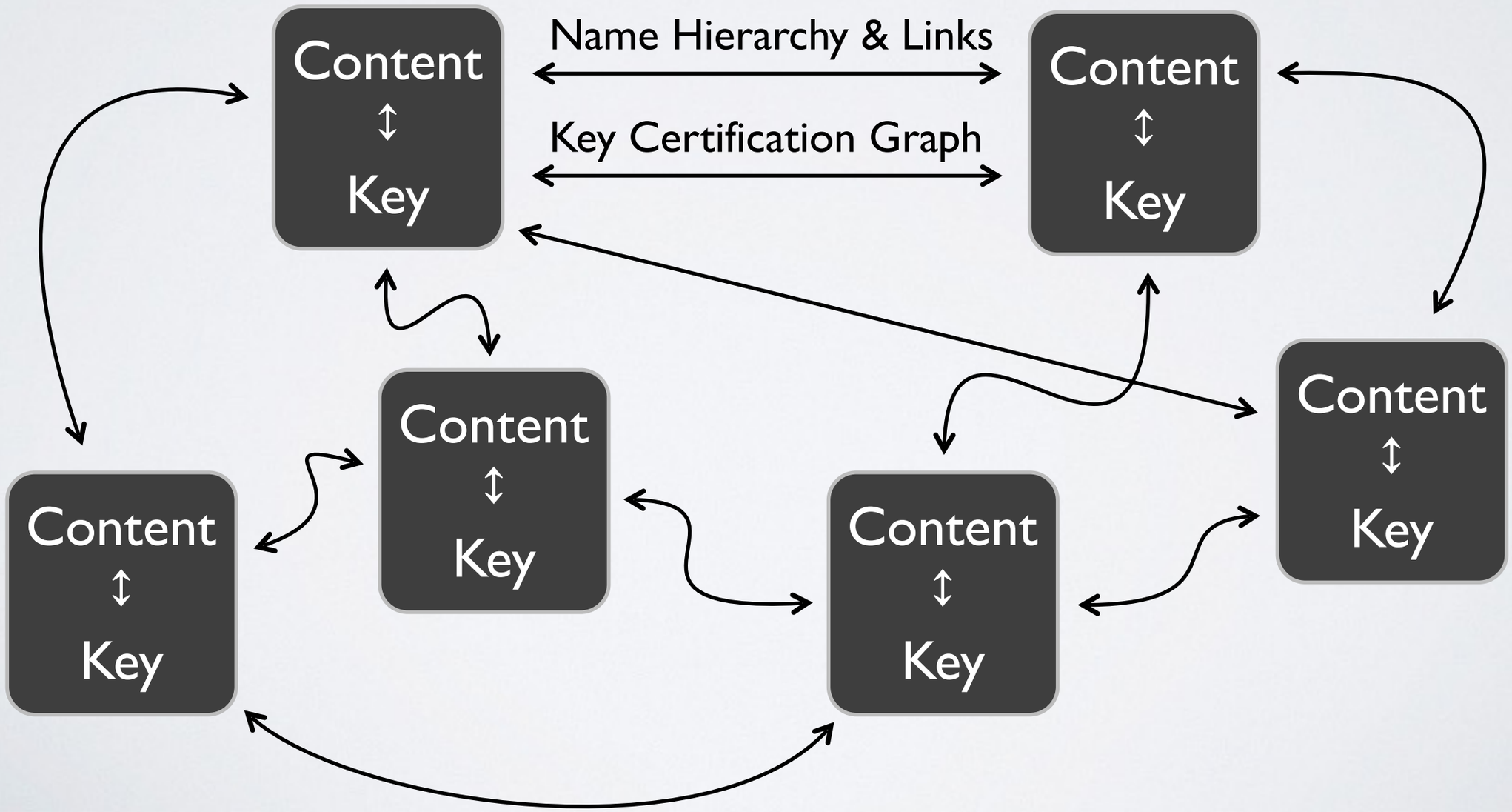
Ensure flexibility, trust agility,
robustness for long-lived signatures.



Big idea: With appropriate mechanisms, signatures can outlive the keys that signed them, even if compromised.

EVIDENTIARY TRUST

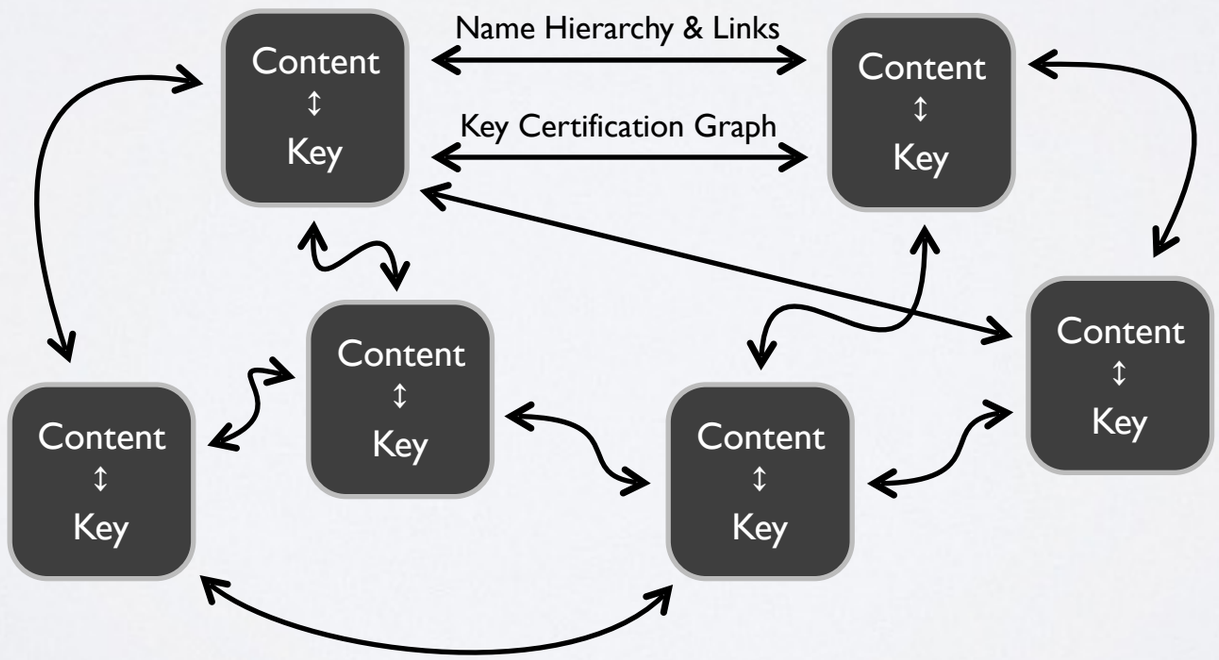
A rich web of trustworthy information arises from named, signed data:



EVIDENTIARY TRUST

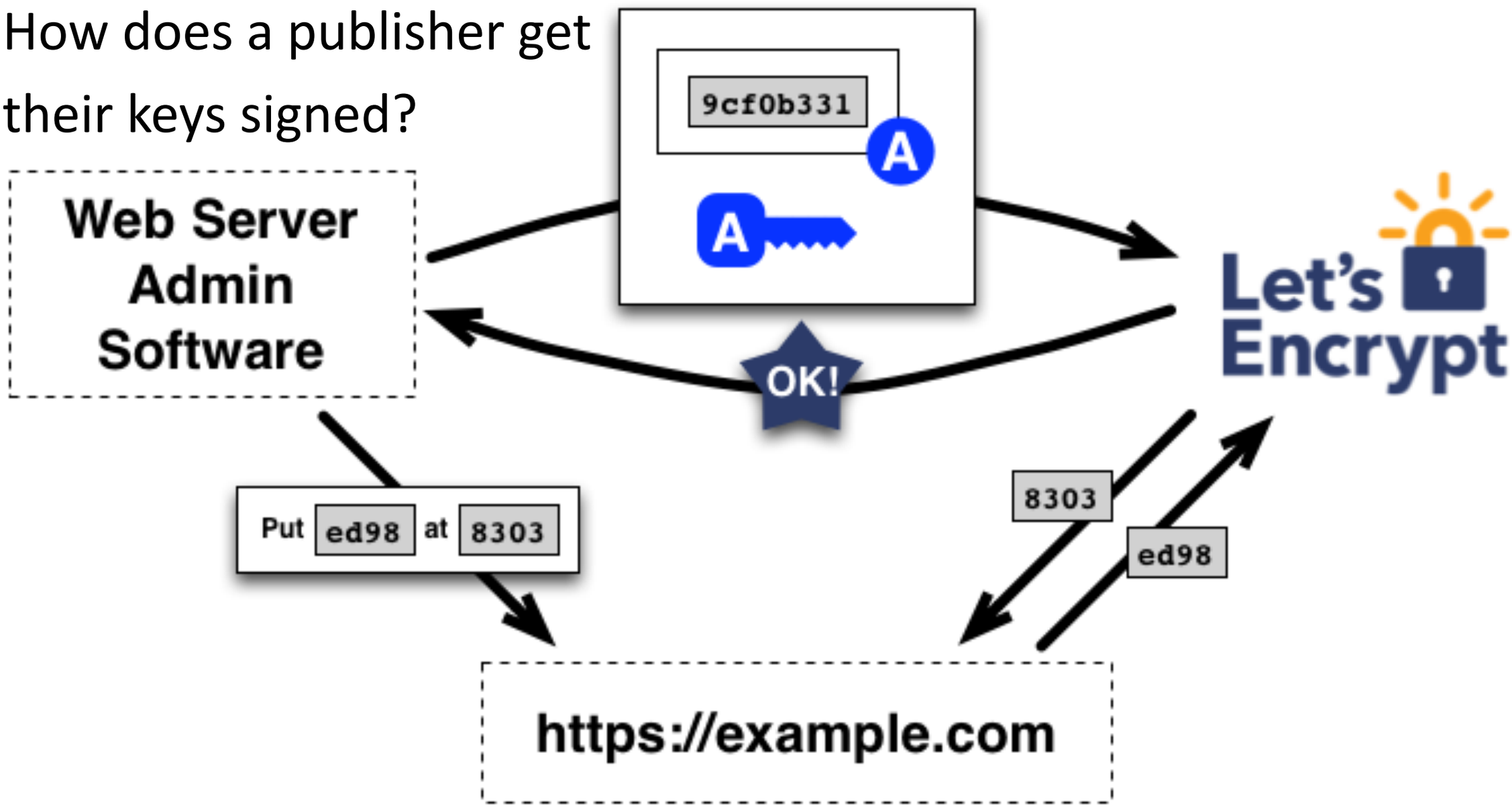
- Attacker's job gets exponentially harder as you accumulate information.

➡Security is emergent property of the system.



AUTOMATICALLY PROVISIONING TRUST

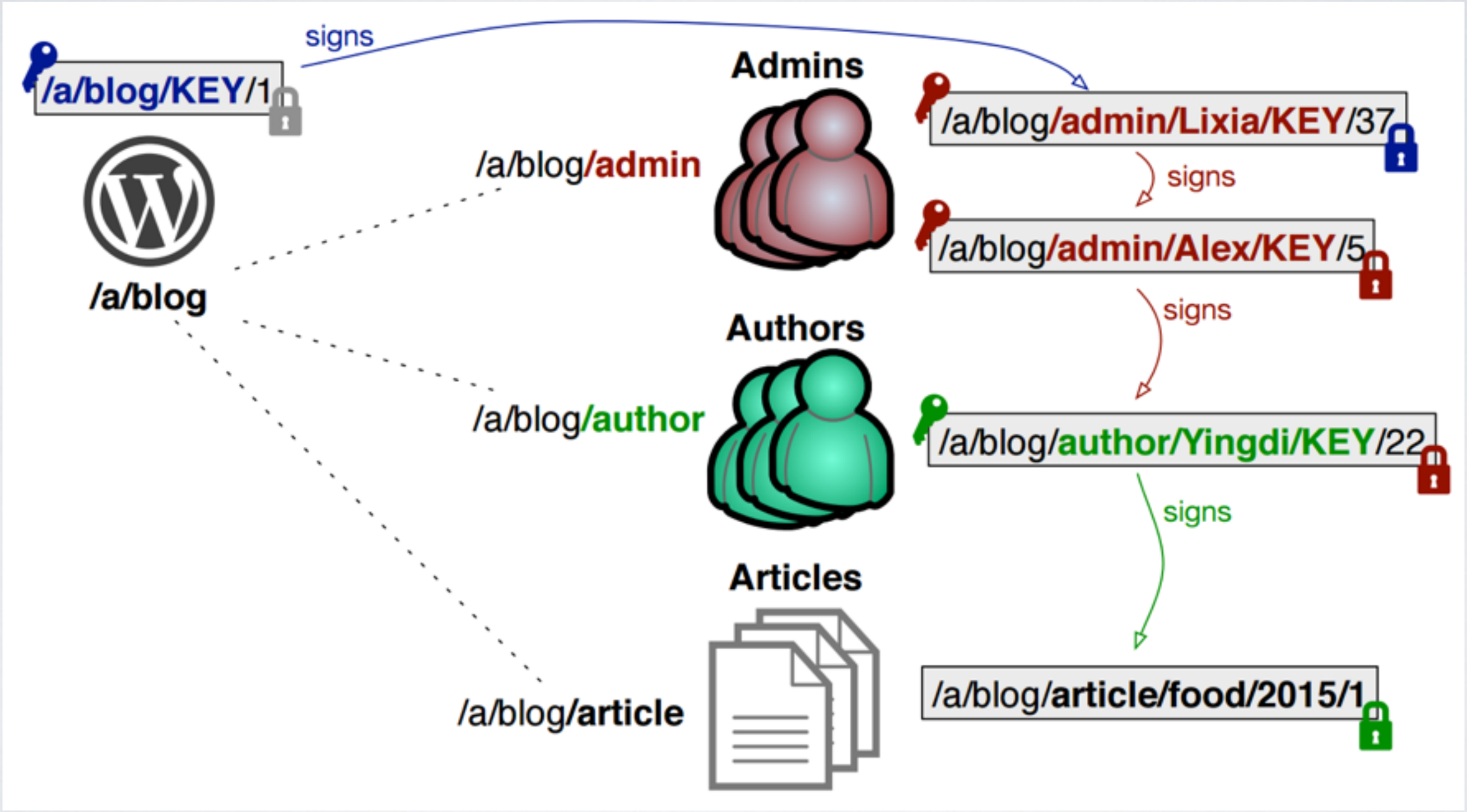
How does a publisher get their keys signed?



Big idea: Abstract identity verification and automate issuance.

TRUST SCHEMAS

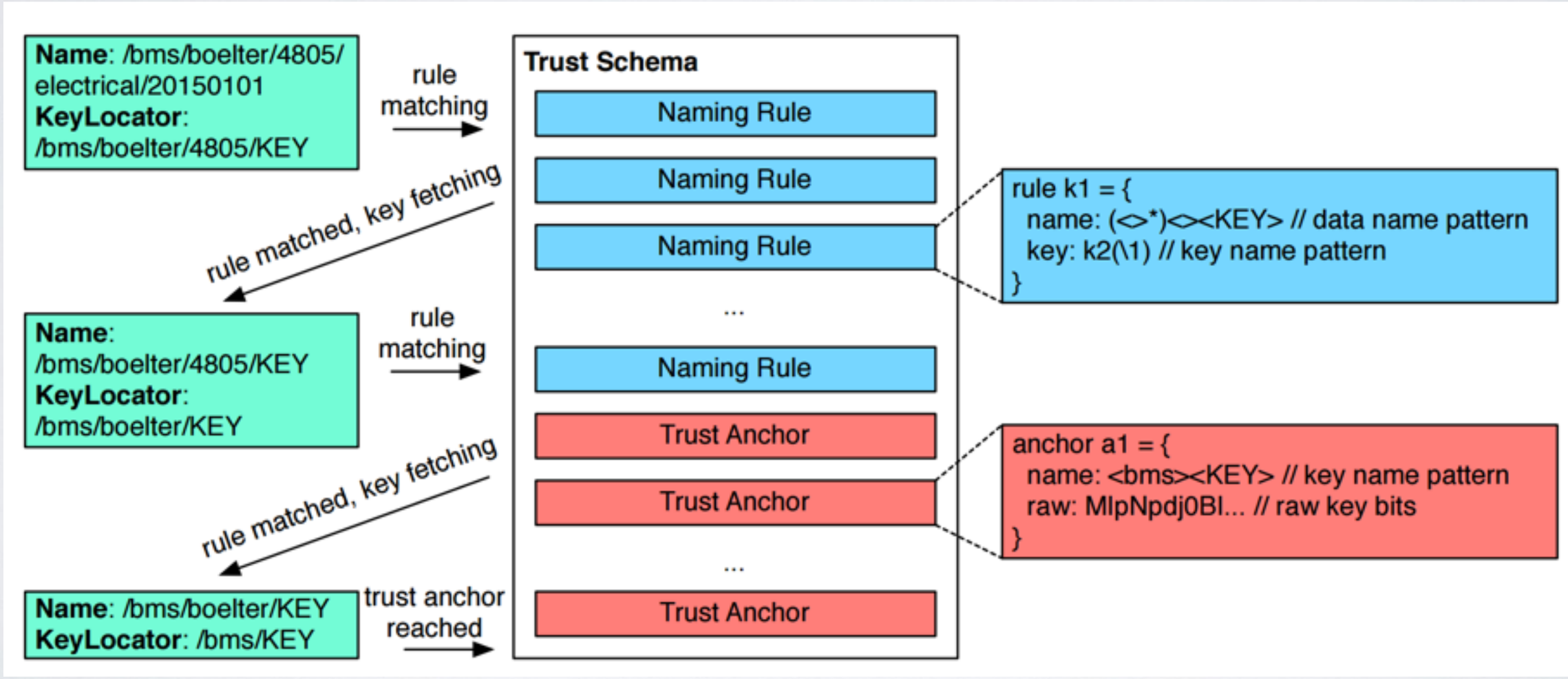
Who is allowed to sign what?



Big idea: Namespace design can convey capabilities, structure trust.

TRUST SCHEMAS

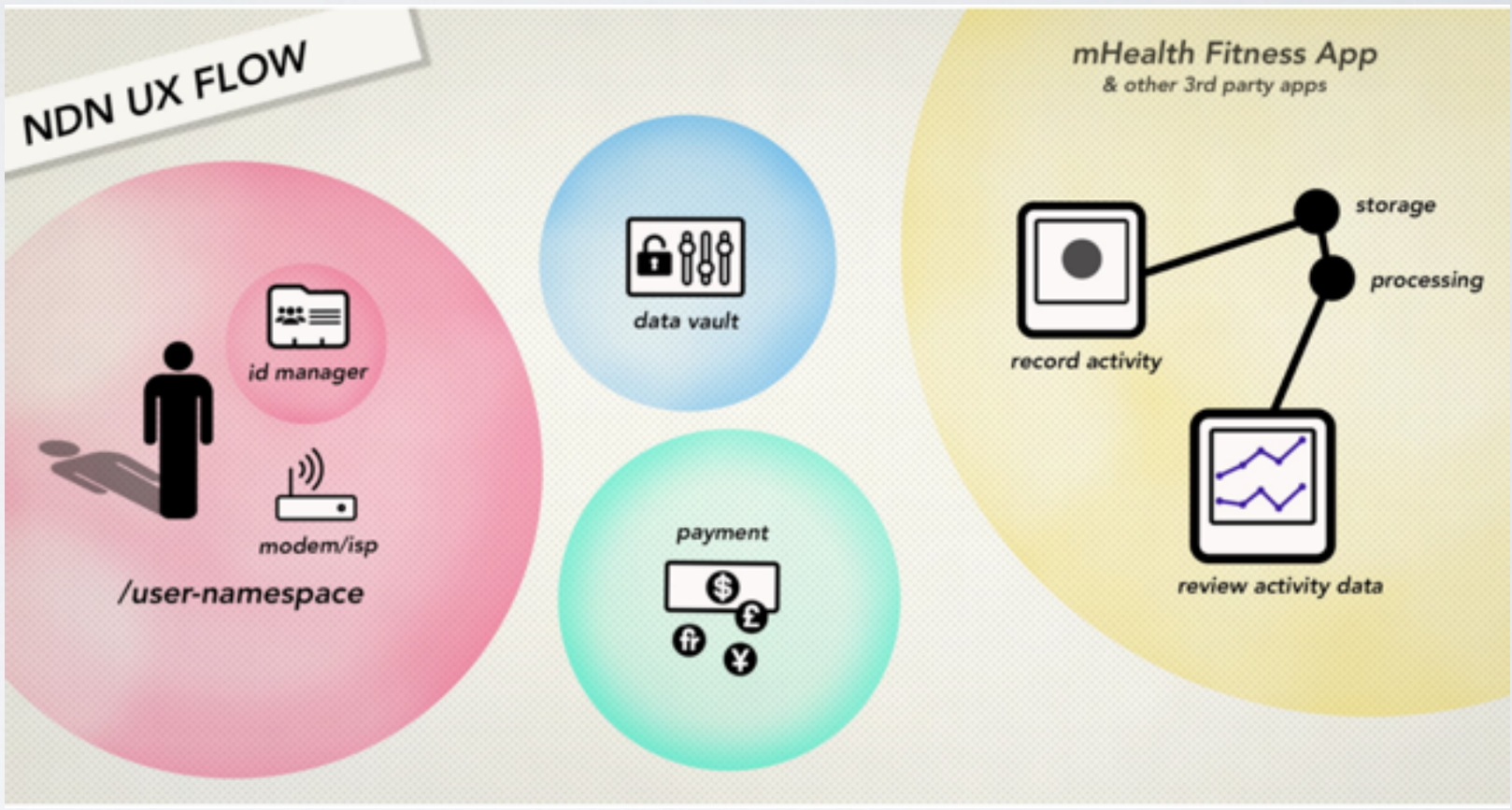
Big idea: Abstract validation based on structure of namespace, allow applications to define rules for trust or adopt pre-defined templates designed by experts.



Achieves vastly greater flexibility and security than existing TLS PKI.

LEARNING FROM APPLICATIONS: OPEN MHEALTH

Granular, user-centric data access control in an ecosystem of composable services



- An old idea: Encryption-based access control
- New opportunities: Use namespace hierarchy to express fine-grained access policies

SECURITY LESSONS

Data-centric security philosophy allows us to convert hard security problems (e.g., host security) into ones that are relatively easier (crypto, key management).

Security priorities will continue to evolve, and no network architecture will solve them all for all time—but architecture can give us a more solid foundation.

NDN has yielded insights on problems and solutions in the IP/TLS architecture.

WHO IS USING NDN NOW?

not your father's Internet (yet)

- leading edge users in a lot of pain.
  - big data applications, e.g., high energy physics
- emerging commercial interest in narrow slice of it, e.g., video content distribution
- one instance of secure data storage services (Telehoc)

[see NDNCOMM 2015 report, named-data.net]

RESEARCH AGENDA

App Design
- Namespace
- Trust models
- In-network storage
- Synchronization
- Rendevous, discovery, boot-strapping

Security
- Fast signing
- Usable Trust
- Privacy
- Attack resistance

Routing
- Fast Forwarding
- New models

Fundamental Theory
- Any-to-Any communication
- Bandwidth / Memory /
- Distance tradeoffs

WHO IS MAKING NDN NOW?

- Highly collaborative effort, 10 different campuses
- Software is open source and freely available.
- Tutorials, tech reports, videos of tutorials and meetings

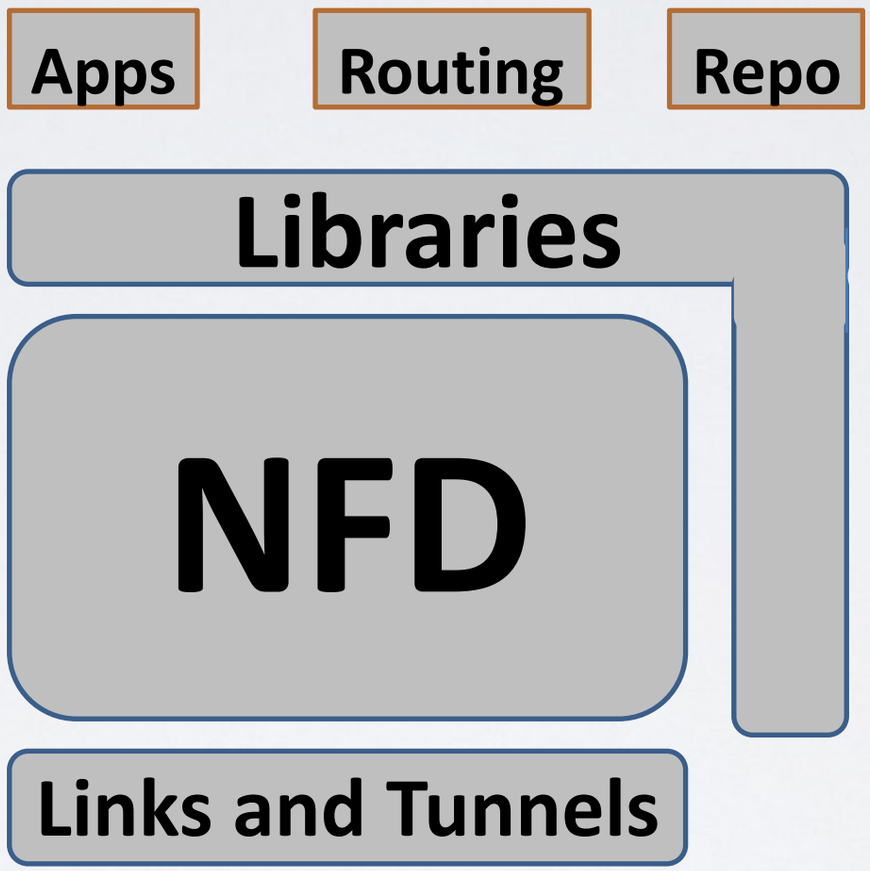
named-data.net

WHY SHOULD YOU CARE?

operators appreciate new ways of looking at problems that remove unnecessary detail

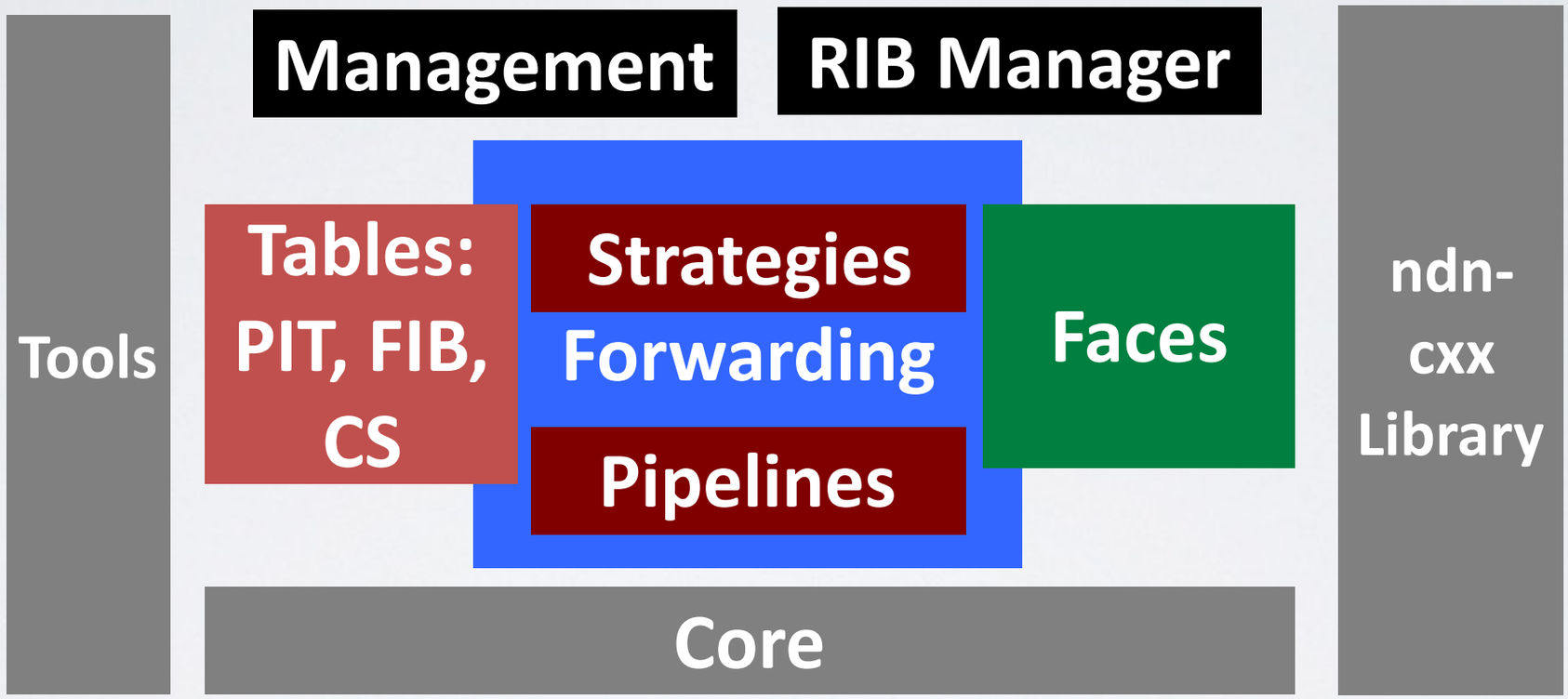
- like plumbing issues (IP address management)

NDN NUTS AND BOLTS



See:
https://github.com/named-data

NFD'S MAJOR PIECES



See:
https://github.com/named-data

NDN PLATFORM

Core: NFD, the NDN Forwarding Daemon

Libraries: full featured implementations in a variety of languages

Applications: rich and growing software ecosystem

NLSR
repo-ng
ndn-tlv-ping
ndn-traffic-generator
ndndump
Federated Wiki
ndn-bms

ndn-lighting
ndn-protocol
ndnfs
ChronoShare
NDNoT
ndnrjs
ndnrtc

Chronochat-js
Matryoshka
ndnstatus
NDNVideo
NDNFit
OpenPTrack-NDN
ndn-dissect

See:
https://github.com/named-data

ICN TUTORIAL ONLINE

Goal: Help guide NDN research & application development

Use chat application to illustrate intermediate concepts:

Synchronization: Abstractions beyond Interest/Data exchange

Storage Options: Alternatives to in-network Content Stores

Trust & Verification: Specifying what content to trust

http://named-data.net/icn2015-tutorial

VISION FOR FUTURE INTERNET

Secured, immutable data with hierarchical names

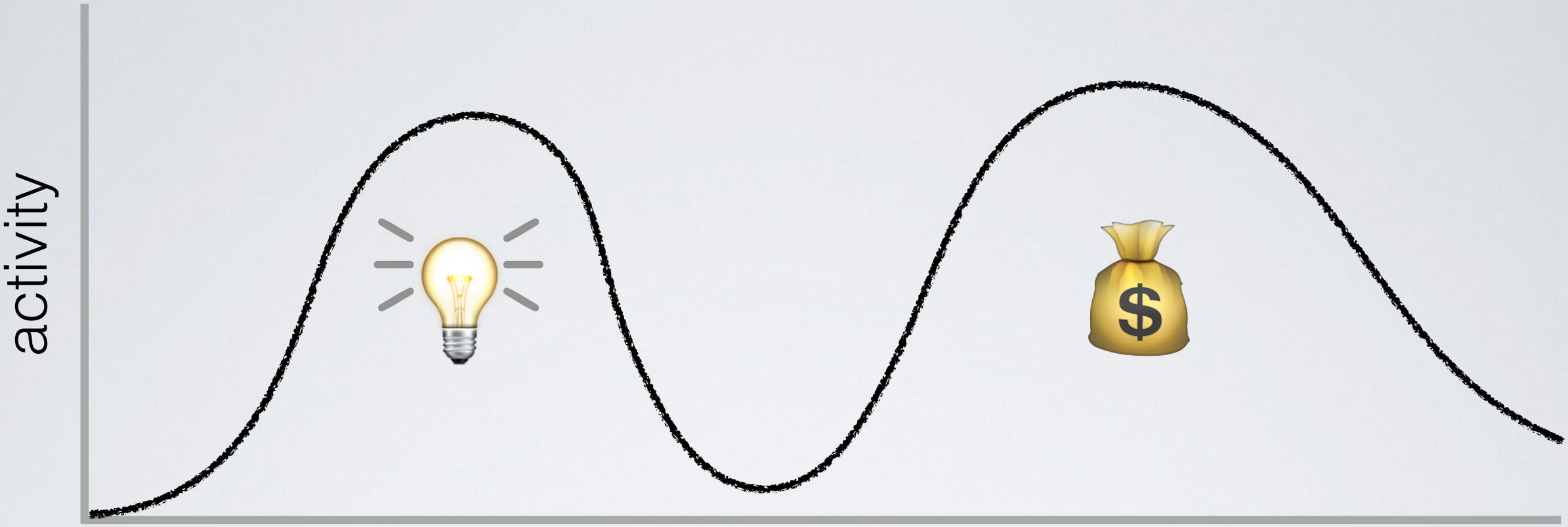
- Big science, small IoT, mobility, intermittent connectivity
- Promotes data management and efficient sharing
- Naming data directly simplifies protocol stack

Applications focus on their data and trust management.

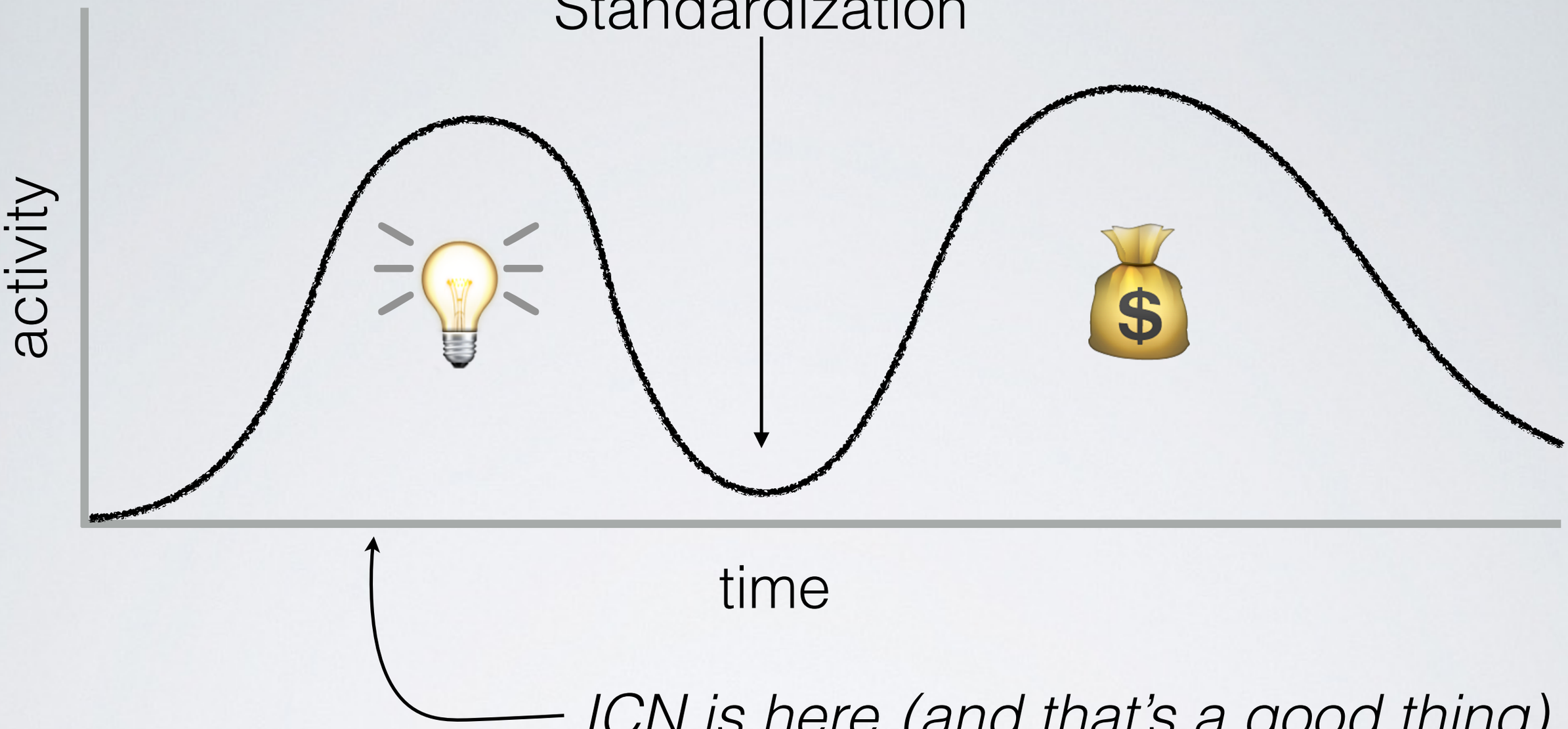
- Networking simply happens, at all scale

In-network storage, multicast to any available interfaces

- Mitigate traffic growth
- Eliminate heavy reliance on cloud
- Enable "O3B" to leverage ad hoc, DTN, P2P, intermittency
  Miinimize energy consumption, delay, facilitates privacy



(Dave Clark ~1985)



(Dave Clark ~1985)

groups.csail.mit.edu/ana/People/DDC/Apocalypse.html