(The Futility of) Data Privacy in Content-Centric Networking

Cesar Ghali, Gene Tsudik, Christopher A. Wood
University of California Irvine
{cghali,gene.tsudik,woodc1}@uci.edu

ACM WPES 2016

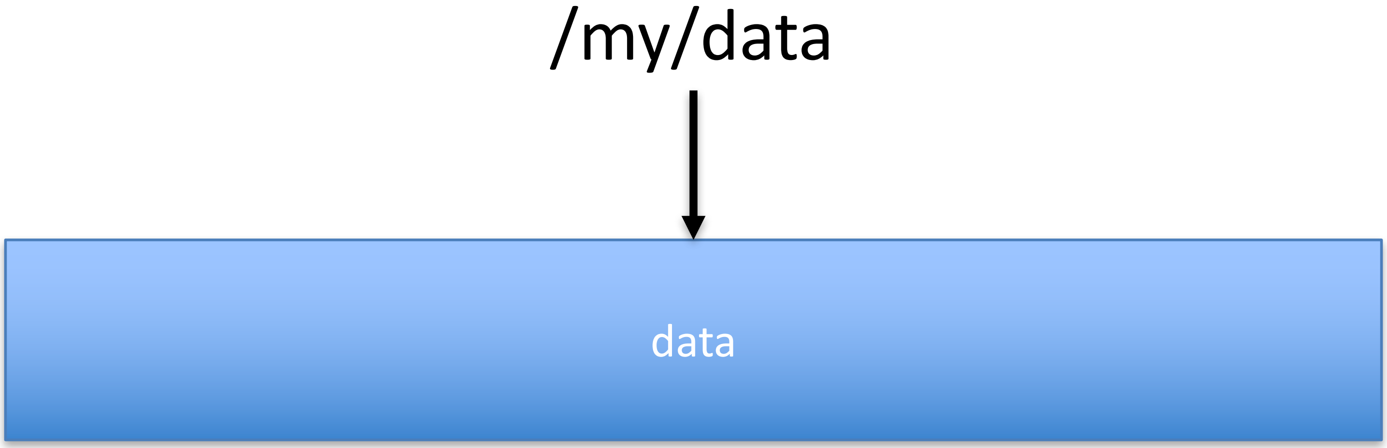
Outline

- CCN overview
- Privacy in IP vs CCN
- Privacy attacks
- Privacy requirements
- Looking ahead

CCN Overview: Named Data

data

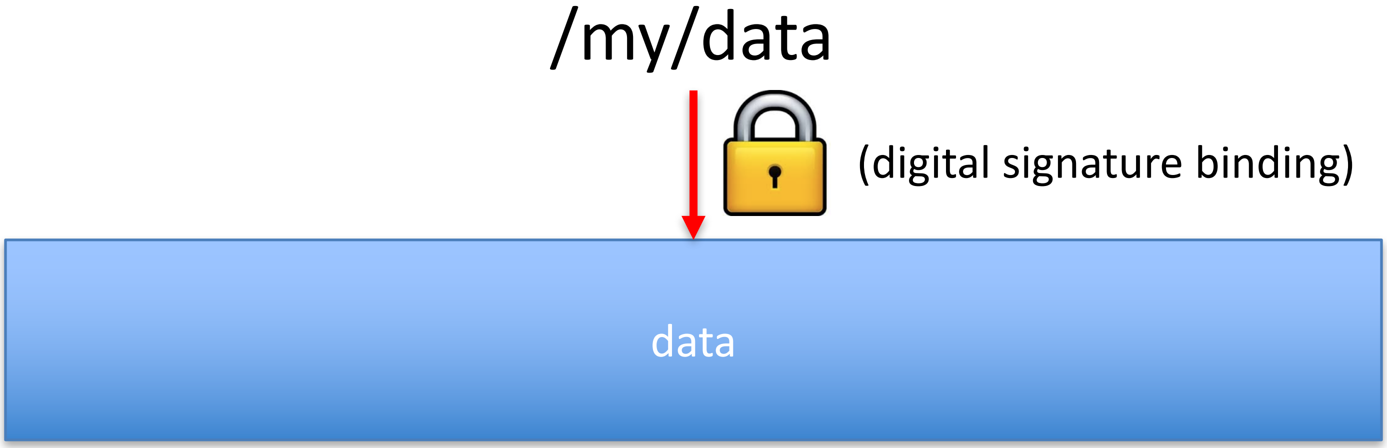
CCN Overview: Named Data



/my/data

data

CCN Overview: Named Data



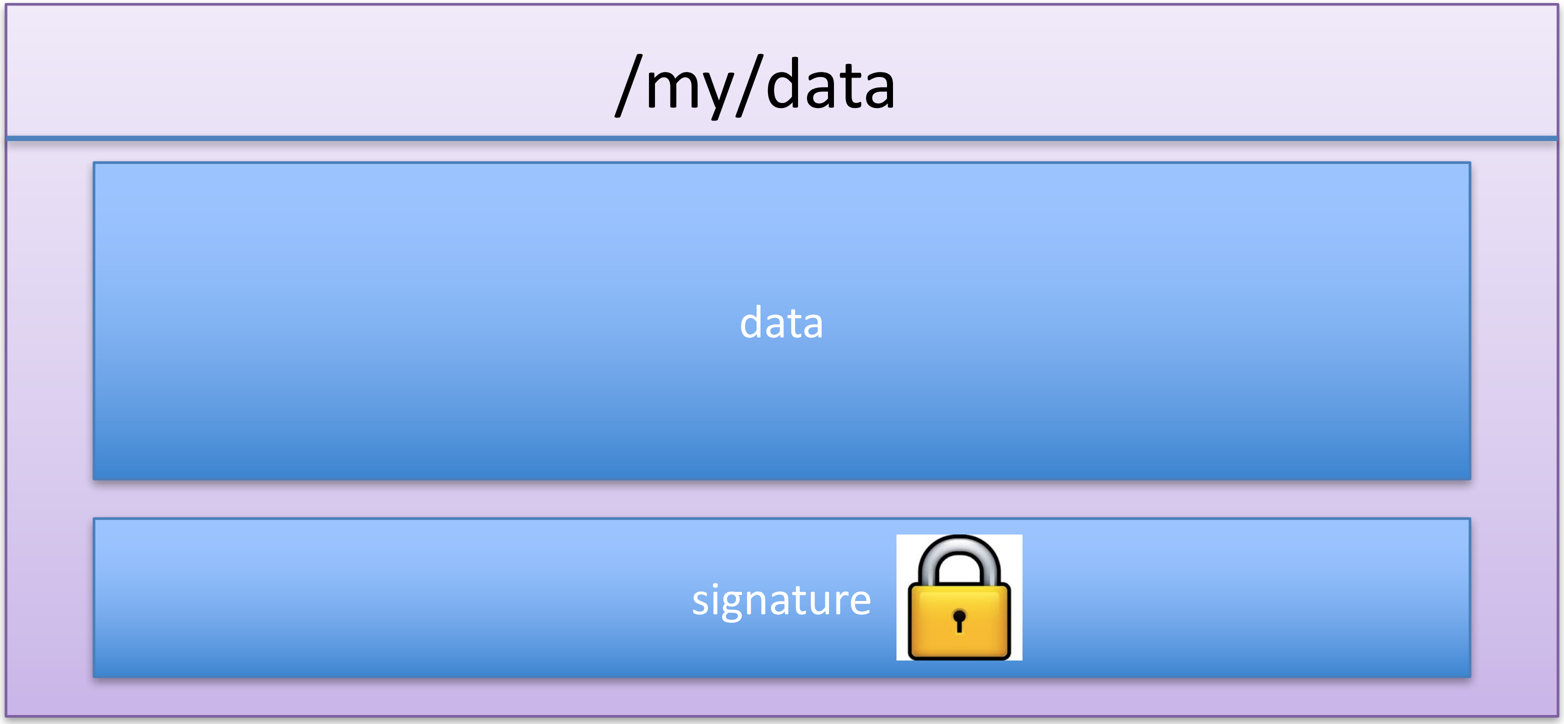
/my/data

(digital signature binding)

data

CCN Overview: Named Data

Named Data (Content) Packet

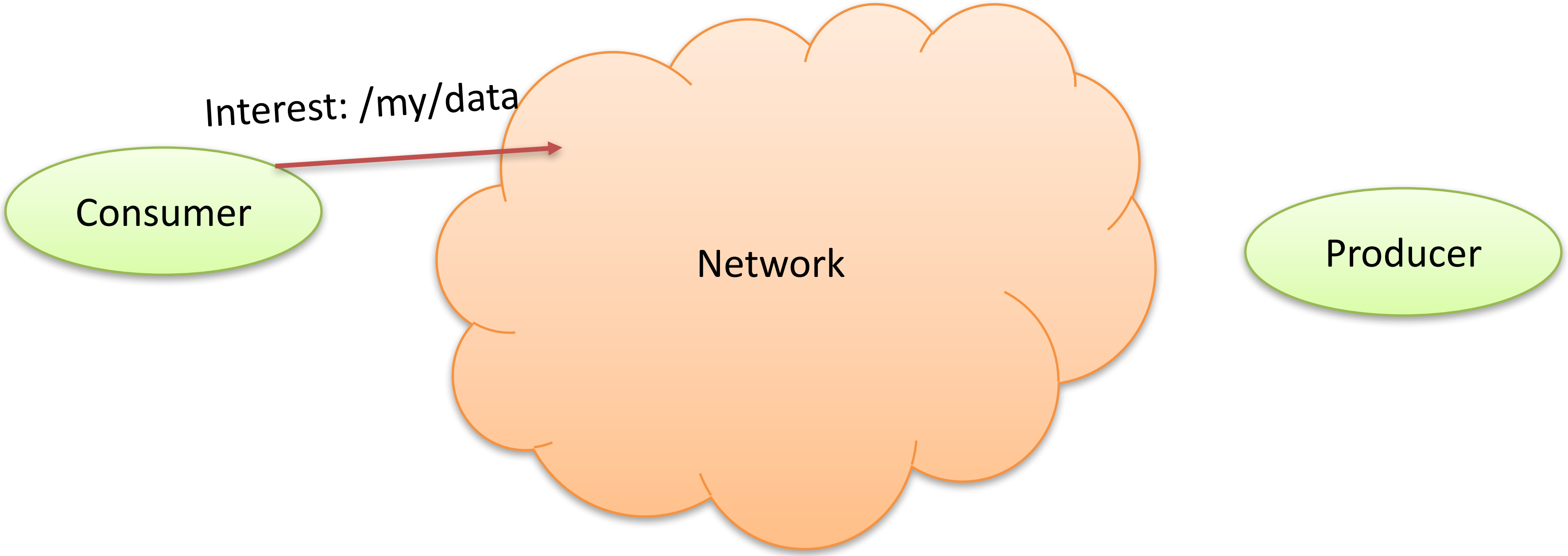


/my/data

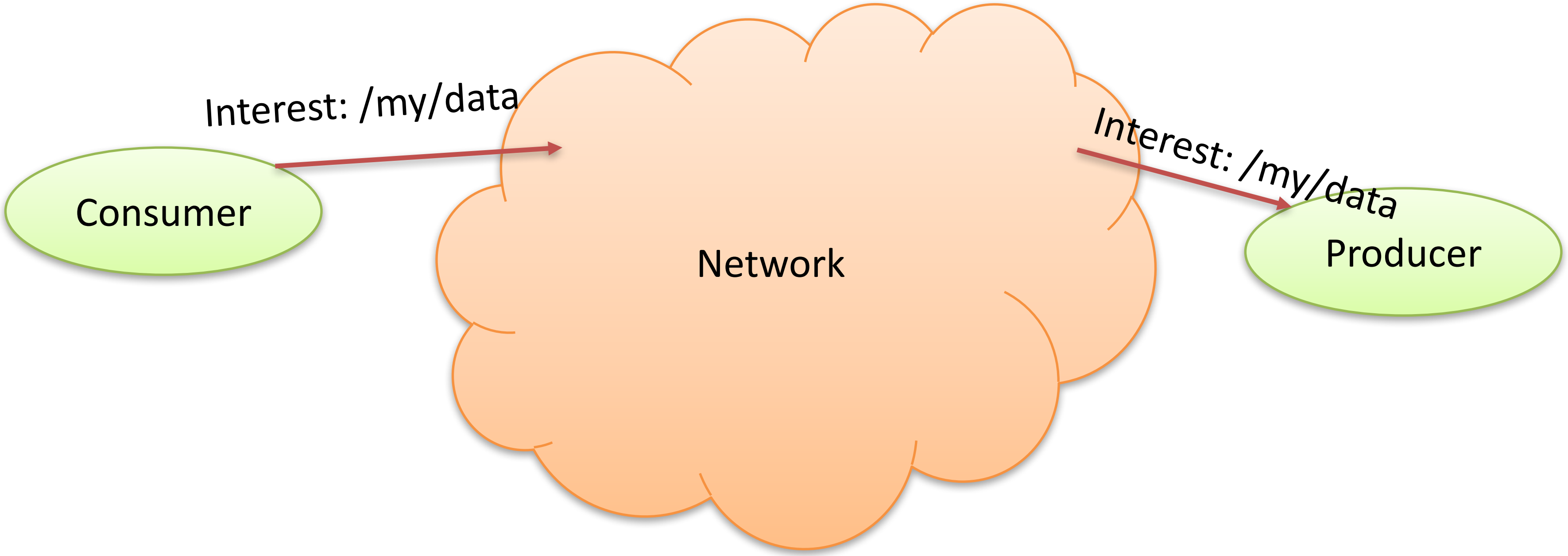
data

signature 🔒

CCN Overview



CCN Overview



Interest: /my/data

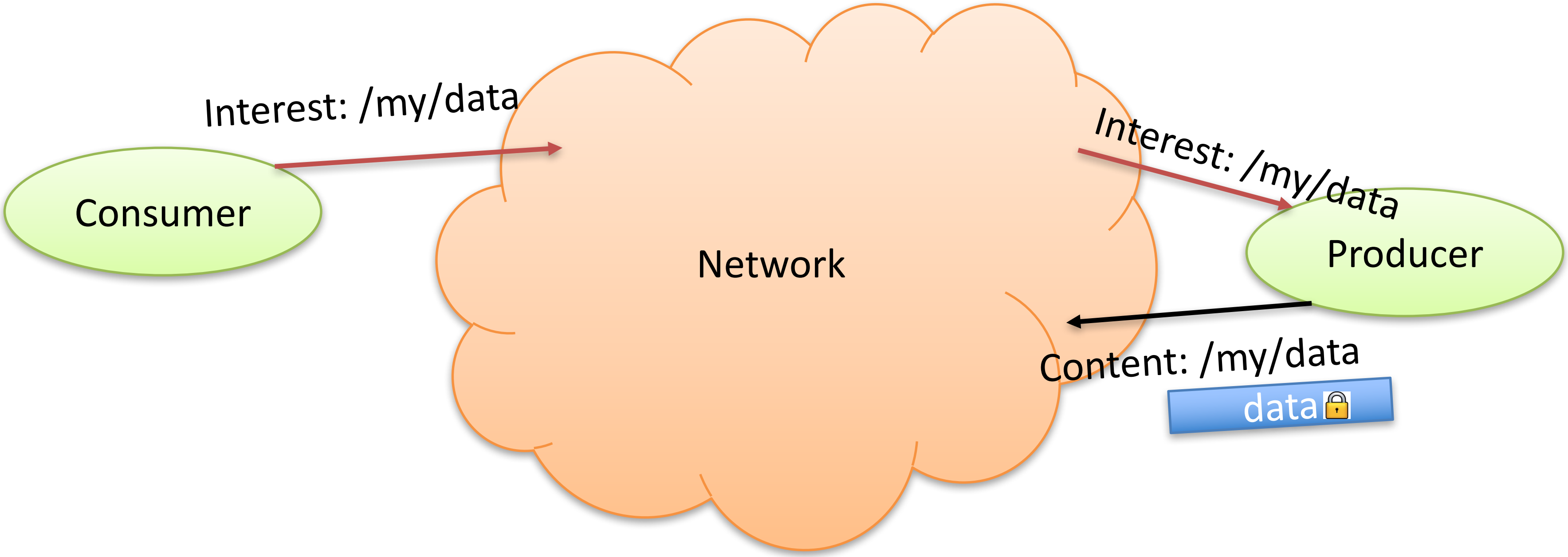
Consumer

Network

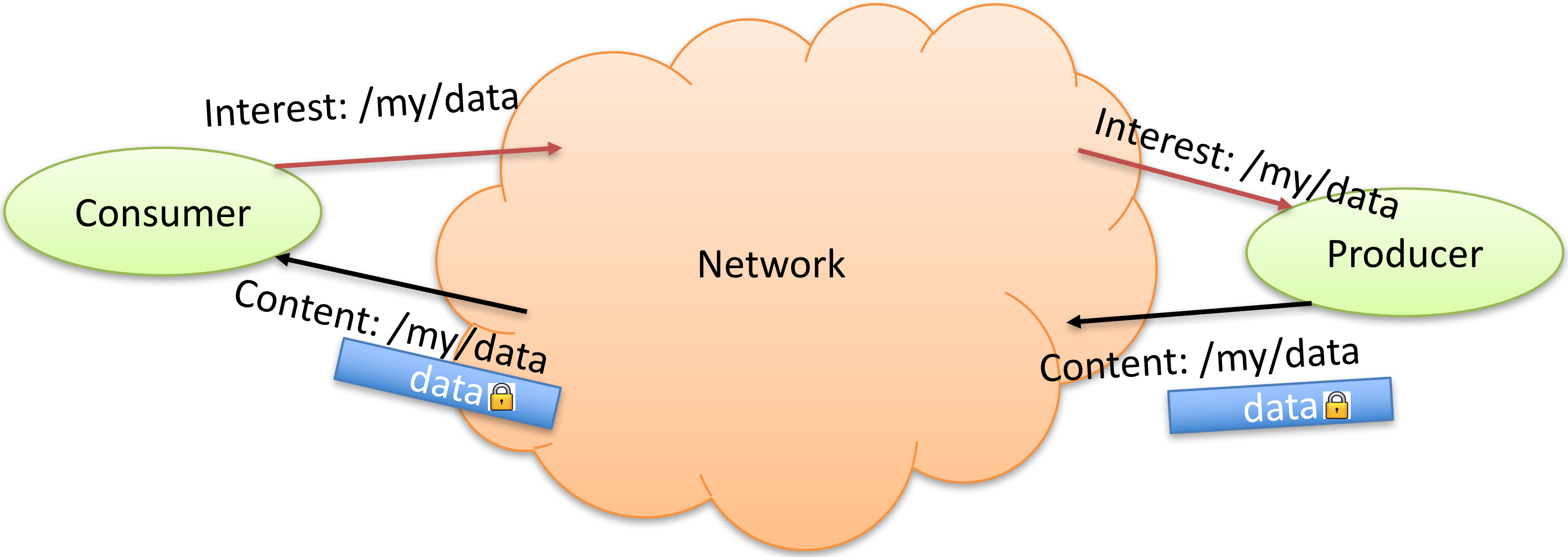
Interest: /my/data

Producer

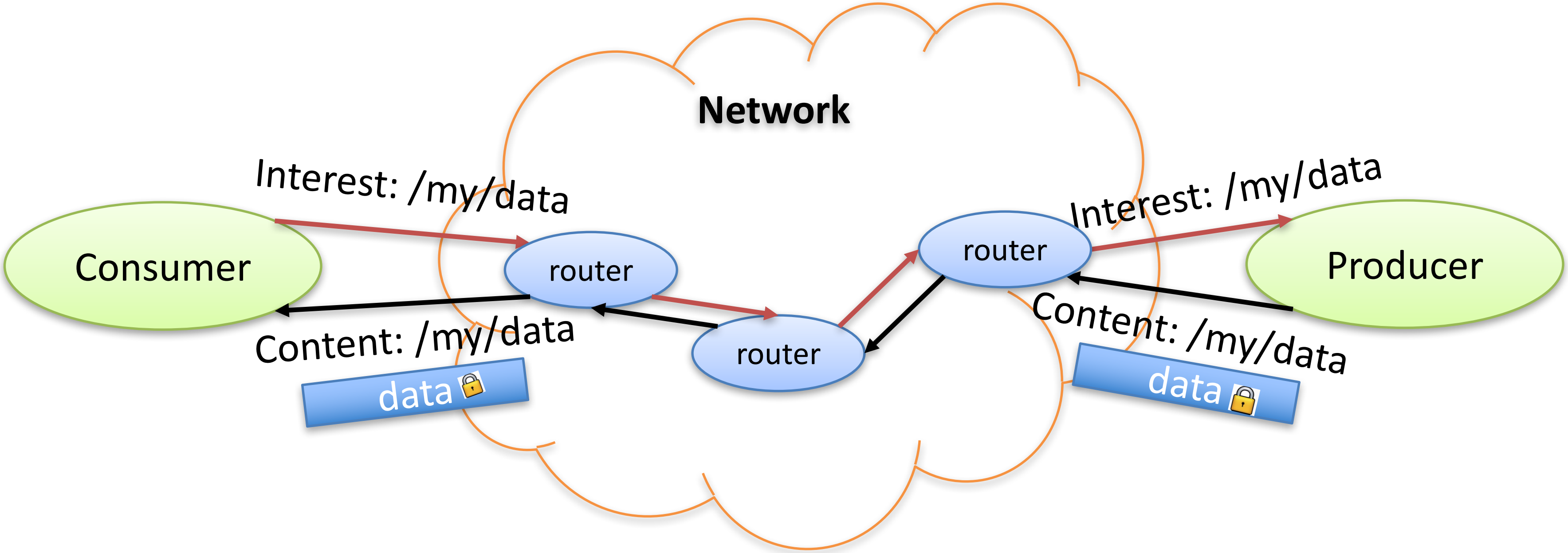
CCN Overview



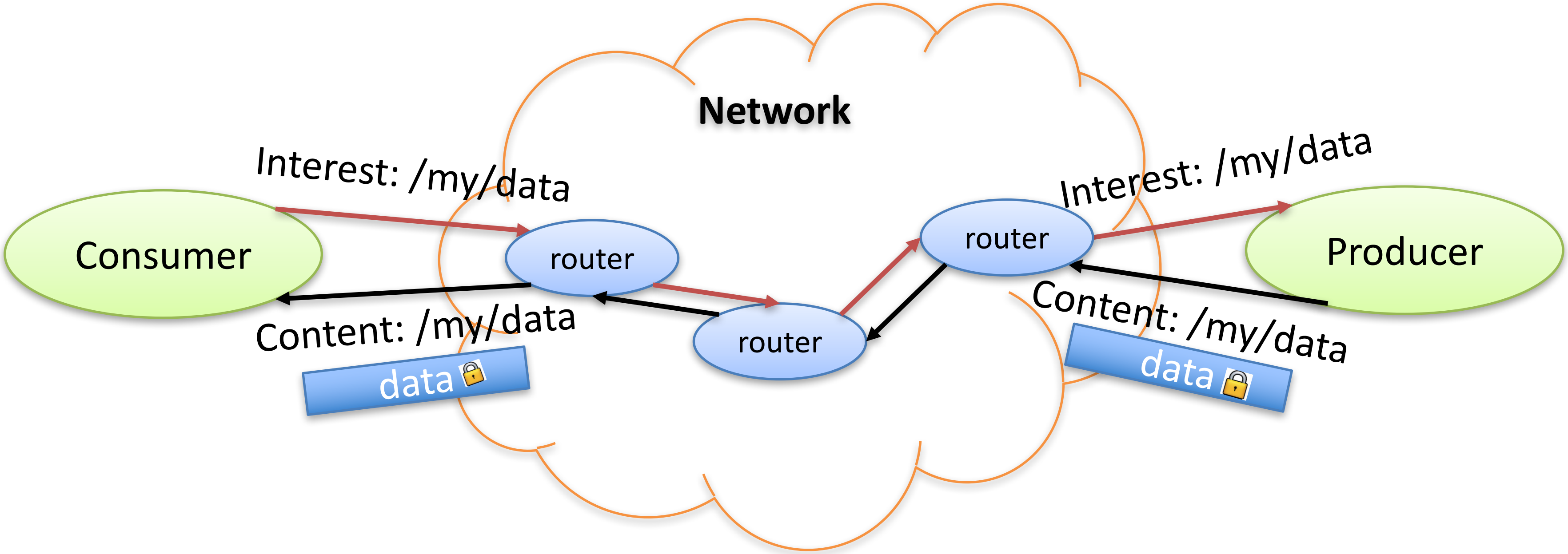
CCN Overview



CCN Overview

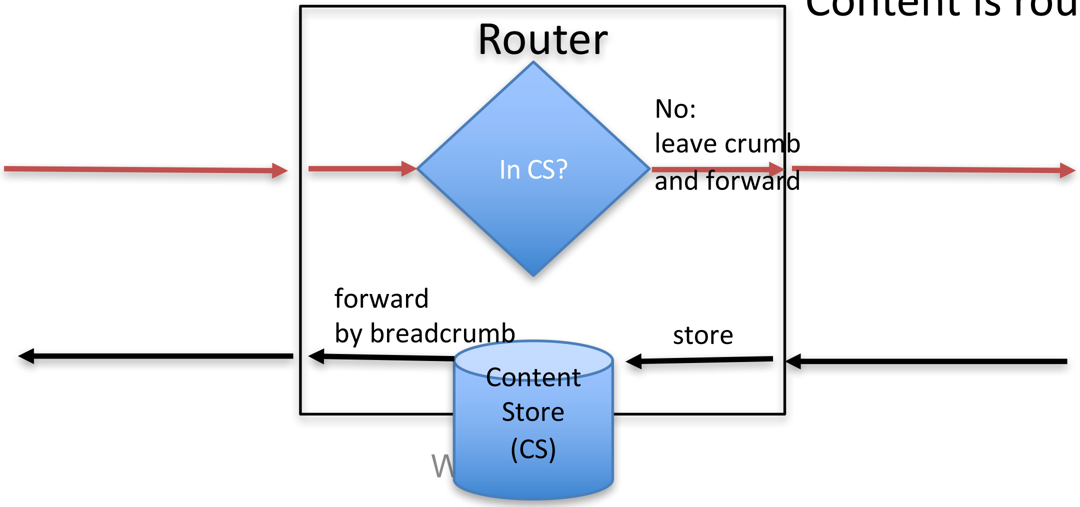


CCN Overview

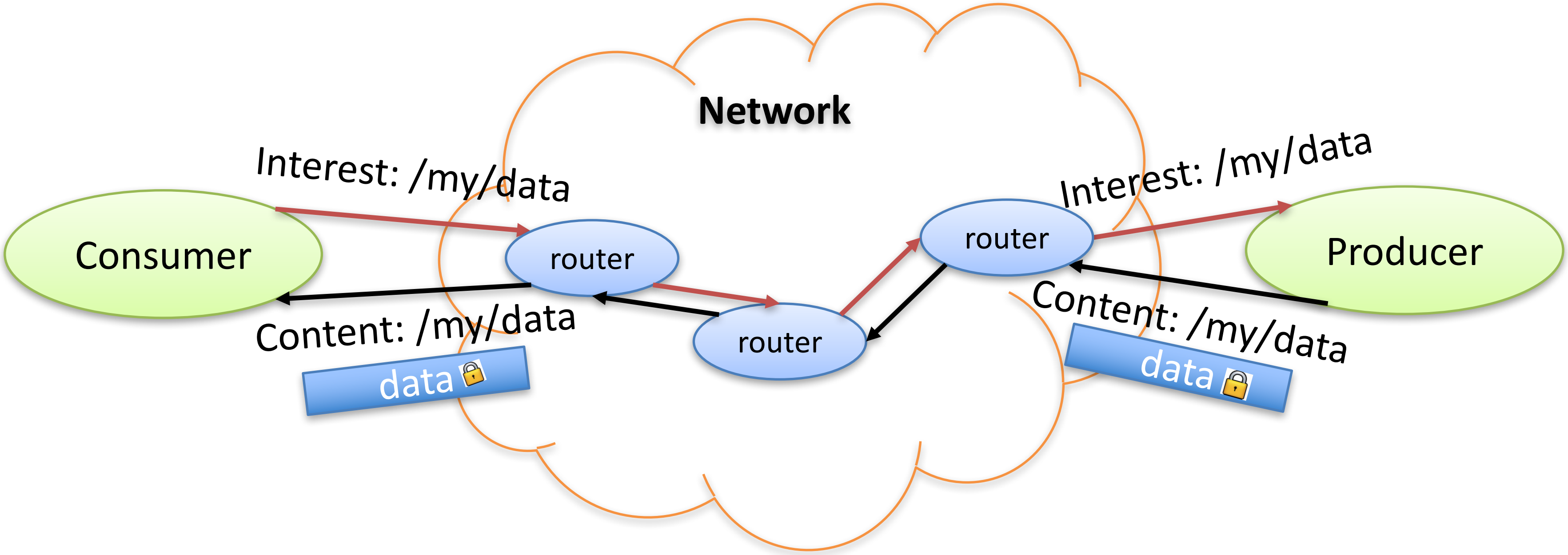


Interests are routed by **name**

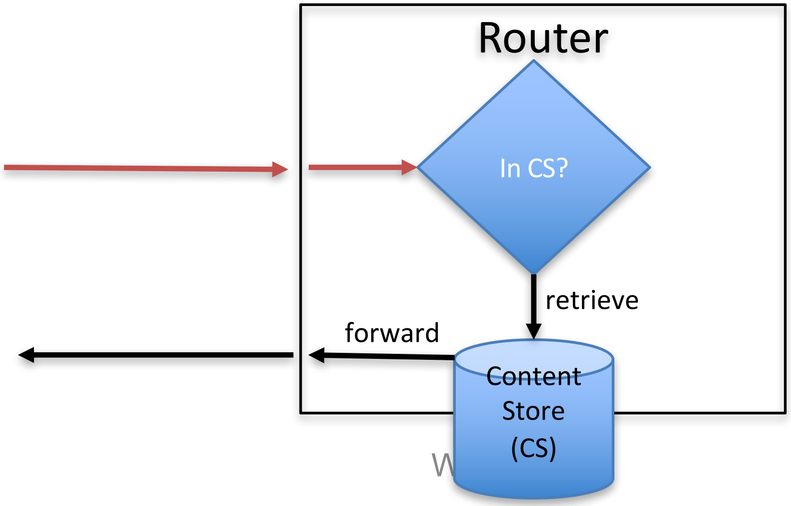
Content is routed by breadcrumbs



CCN Overview



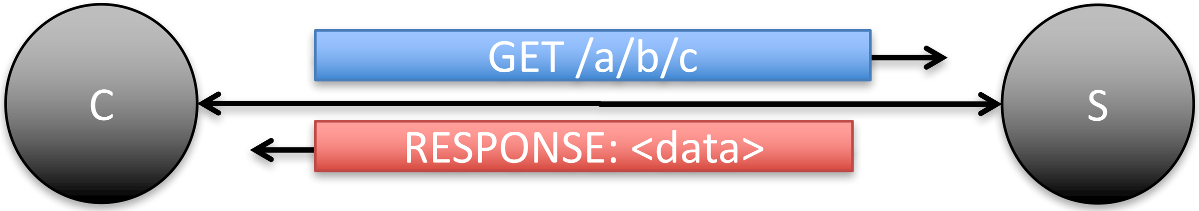
Interests are routed by **name**
Content is routed by breadcrumbs



Onto Privacy

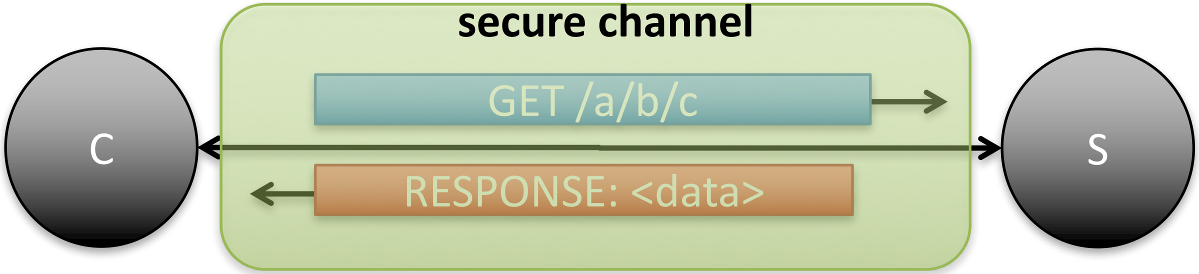
IP Privacy

Turns this…



IP Privacy

Into this… (with IPSec or TLS)

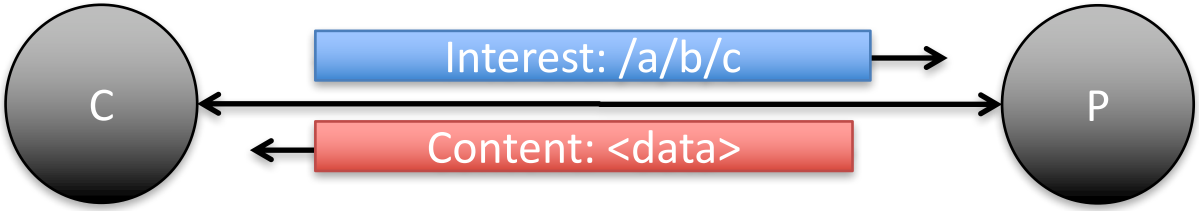


What's revealed?

- Source and destination addresses and port #
- Timing
- Packet sizes

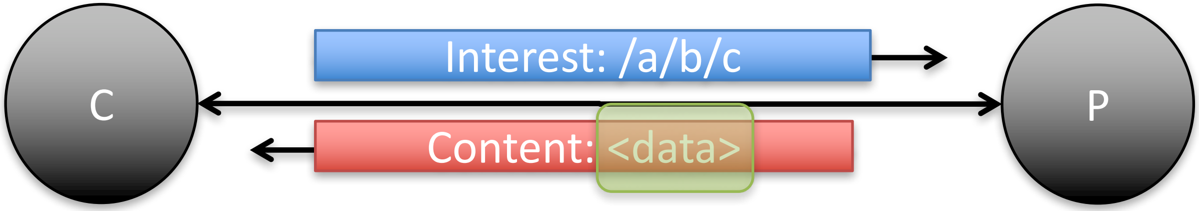
CCN Privacy

Turns this…



CCN Privacy

Into this…



encrypted content?

What's revealed?

- Consumer and producer locations
- Timing
- Packet sizes
- Interest name
- Producer identity
- …

→ Properties of the (application) **data**

Privacy Parity

CCN privacy **<** IP privacy

- What's the "delta"?
  - – Interests for same content can be correlated
  - – Interest names reveal information about content
  - – Content carries explicit names
  - + Location of content not (always) apparent

BTW:

- Anonymity <> privacy
- … *anonymity is out of scope*

Privacy Attacks

- **Correlation**: learn when two requests correspond to same content
- **Identification**: learn when specific content was requested
- **Leakage**: learn **anything** from a request or response

Adversaries

- **Eavesdropper**: a passive interceptor
- **On-path HbC**: router that forwards interest and content packets
- **Distributed:** at least two on-path: one near producer, one near consumer

- **Active & Scary:** as above, also generates its own probes

Main Questions

- What properties must responses have to prevent privacy attacks?
- What about requests?
- What about both?

Weak & Strong Privacy

- **Weak**: Adv can not learn anything from a request or response, but can correlate packets
- **Strong**: Adv can not learn, identify, or correlate

Weak Privacy Requirements

- Responses must be protected via IND-KPA secure encryption

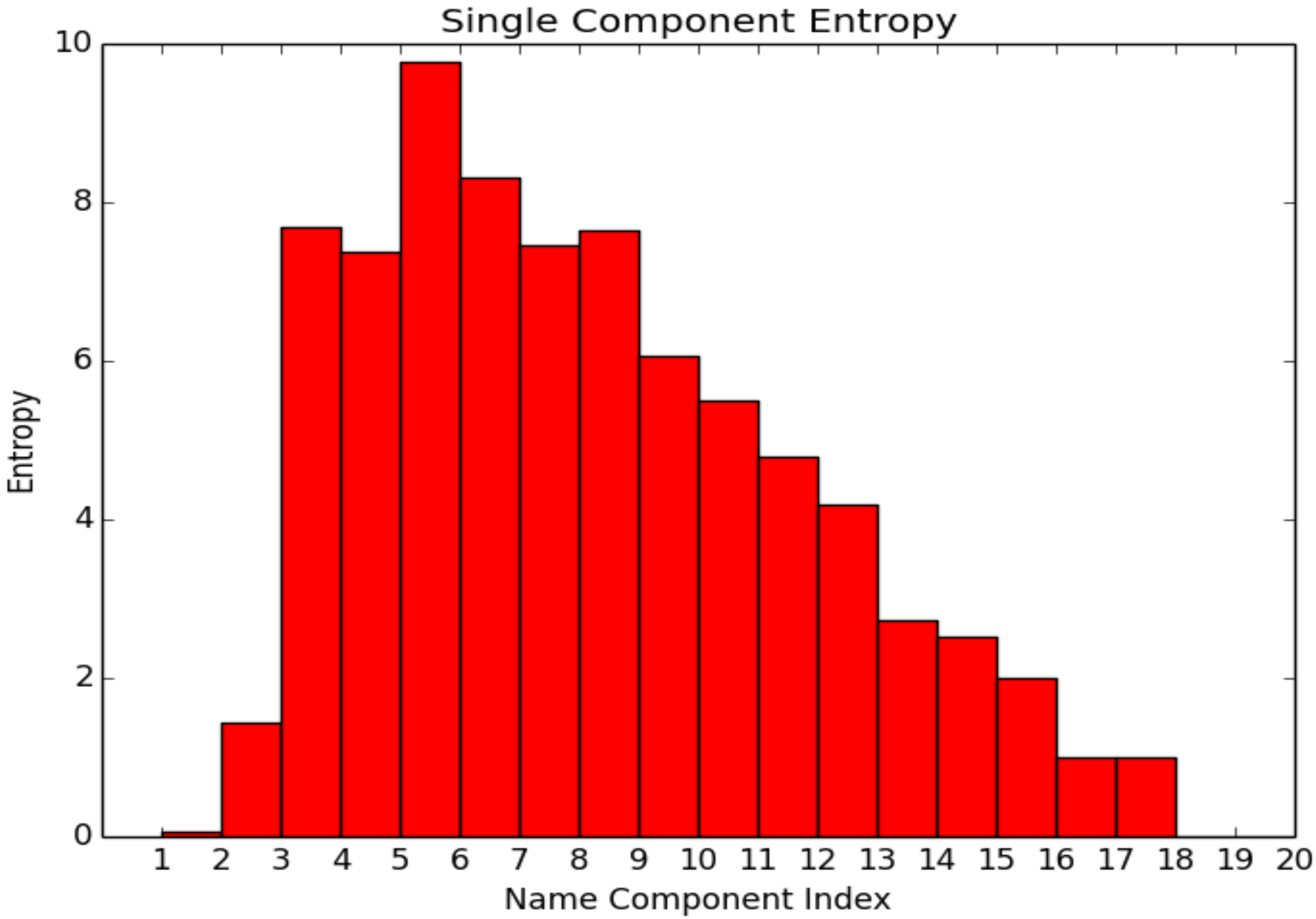
Why? To prevent trivial information leakage

- Requests must be transformed by a **deterministic cryptographic PRF** that is **not length preserving**

Why? Length can be used to distinguish requests from one another and the "network" representation must appear random to Adv

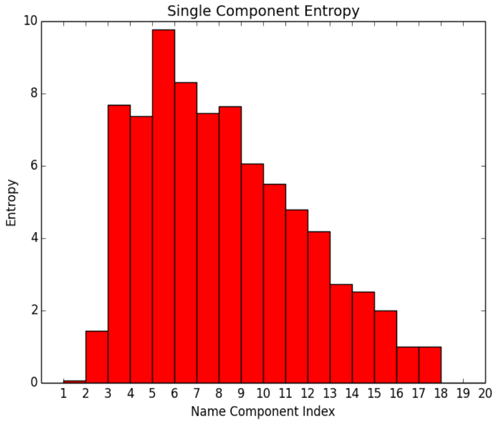
Why deterministic? How to route otherwise? Also: how to preserve the interest collapsing feature?

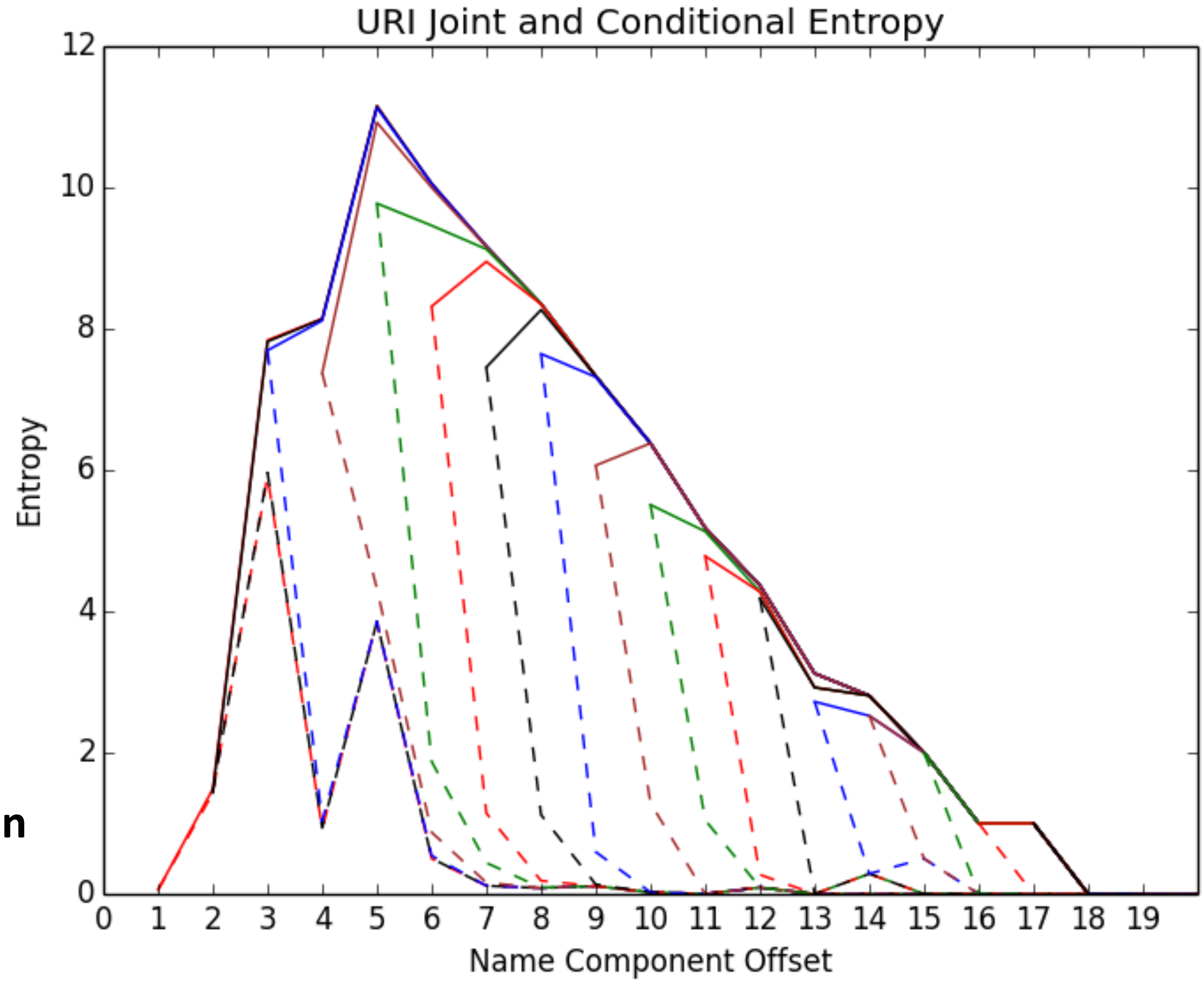
Hash Functions Are Not Enough



Source: Cisco URI dataset

Hash Functions Are Not Enough





Prefix leaks information about the suffix!

Design Patterns

So… many consumers share the same secret?

Shared Secret?	Strategy
Yes	• Consumer and producer: derive ephemeral shared key from secret, use it to encrypt request and response
No	• Consumer: generate random key, encrypt request with producer's public key • Producer: decrypt random key, use it to encrypt response

But what about caches?

Strong Privacy Requirement

Requests and responses must be protected with IND-CCA encryption

Why? To prevent correlation attacks

Design Pattern

Create a secure session (as in TLS) and use it as a pipe to transfer requests and responses

Outcomes?

- Any realistic form of data privacy complicates CCN request-response m.o.
  - It's no longer a simple request-response!
- In most circumstances, privacy inhibits caching
  - How important is caching in CCN?
- To have **strong** (proper) data privacy, CCN is not very different from IP+TLS
  - So what are we doing here?

Open Question

Will

CCN privacy

remain forever elusive

or at least

inferior to IPsec/TLS?

Questions?

Thanks!