2015 Spring：Future Internet Class
2015.6.9

ICN for Future Internet

Jun Bi
Tsinghua University

Outline

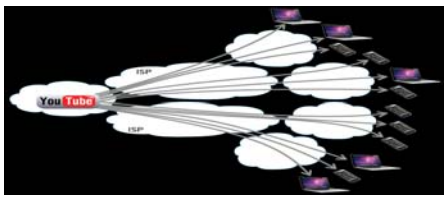
- Why ICN, even IP is such a great success?
  - It is the data that users concern about essentially
  - End-to-end restrictions to data distribution
- How is ICN designed?
  - Naming the data: unique and independent from location/path
  - Retrieving the data: discover the data, and then transmit it
  - Securing the data
- Debate: Only a few consensuses
  - Self-certifying v.s. Hierarchical name
  - Edge v.s. in-network caching
- Conclusions

Why ICN?
IP is already a great success

The Internet is turning to data distribution infrastructure, while IP is designed for connecting two ends

Data Distribution Context

- End-to-end TCP/IP semantic restricts data distribution from utilizing data replicas or redundant paths
  - Patches do not get widely deployed: IP multicast, multipath TCP, SCTP, DCCP, Tng.
  - Overlay suffers from trust, heterogeneity, path stretch, link stress, etc: CDN, p2p, ALM

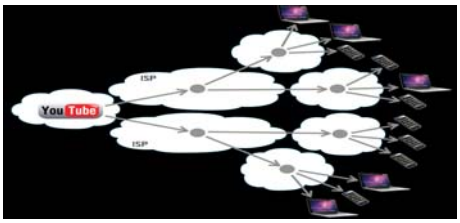


Data Distribution Context

- Basic Observation in current era:
  - End users essentially concerns about the data as long as it is genuine; instead of where the data residents, how the data is reached, or from which path the data is transferred

Information-centric Networking

- ICN, which names the data directly, breaks through the restrictions of end-to-end IP semantics
  - Multiple copies: authoritative sources or delegation
  - Multiple path & multicast & broadcast on broadcast channel
  - Identical request aggregation



How to Design an ICN Architecture ?

Retrieve and secure the data via name-based primitives

Design: Three implications of ICN

- **Naming the data independent from its container**
- **Another two implications of ICN from the perspective of Internet architecture**
  - Retrieving the data
    - Data Discovery: deliver the requests to the target data producer
    - Data Delivery: transmitting the data to requesters
  - Securing the data
    - Validity: the data is a complete, uncorrupted copy
    - Provenance: the data is produced by a trusted party
    - Relevance: the received data is the desired one
- **Primitives of retrieving and securing data are based on given name**

Naming the Data

- **Hierarchical/Human-readable (HR) name**
  - Introduce the binding between the desired data (entity in human mind) and its ICN name (entity in cyberspace)
  - e.g., www.google.com/news/xxx
- **Self-certifying (SC) name**
  - Hashing is the simplest form, and general form contain public key digest of producer
  - Introduce the bing between the name and the data (both are cyberspace entities)
  - e.g., 23azdad:alda23ad

Retrieving the Data

- **Two Steps, both are about routing/forwarding:**
  - Data Discovery: Deliver the request to target replica(s)
  - Data Delivery: Deliver the data to requester
- **Name of data as routing identifier (RID)**
  - Name-based routing: just like IP routing, but with another name namespace
- **Name basked Routing + IP (locator) based routing**
  - Name based Routing to find the content
  - IP based routing to retrieve the data
- **Soft State (no RID):**
  - Routers maintain the state that needed to forward packets from source to the target

Securing the Data

- **Goals: Validity, Provenance, Relevance**
- **Three Steps to verify Validity, Provenance and Relevance**
  1. Verifying content-name mapping is signed by a particular key
  2. Determining something about who that key belongs to whom, in our term, the producer
  3. Deciding whether or not that is an acceptable producer for this particular data
- **Availability: defend DoS led by caching poisoning:**
  - Caching Poisoning: data is faked and distributed among the network

Summary of ICN Proposals

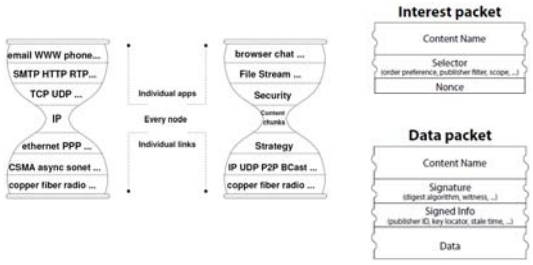
Feature	DONA	NDN	PURSUIT	SAIL
NDO Naming	self-certifying	human-readable	self-certifying	human-readable or self-certifying
Routing Identifier	attachment point	name	path label	attachment point
Data Discovery	name-based routing by infrastructure	name-based routing	name-based routing by infrastructure	name-based routing / (name resolution & locator-based routing)
Data Delivery	locator-based routing / hop-by-hop state	hop-by-hop state	path addressing	locator-based routing / hop-by-hop state

Different proposals advocate different design,
Only a few consensuses

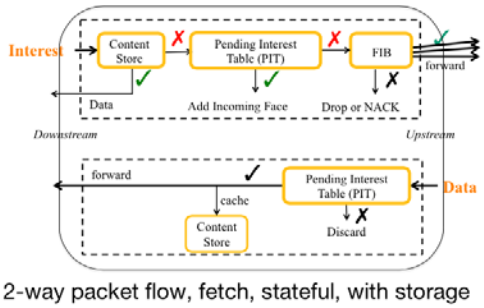
ICN Examples

NDN

- Where -> What
- Internetworking -> Named Data Networking (NDN)



NDN Router

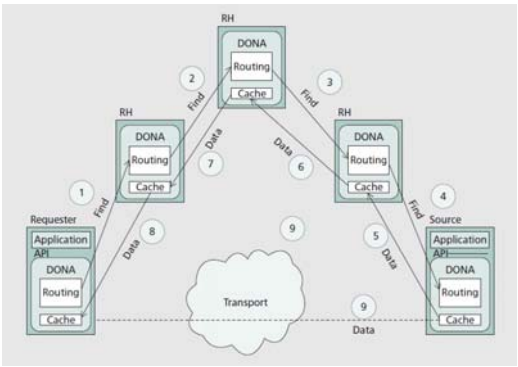


2-way packet flow, fetch, stateful, with storage

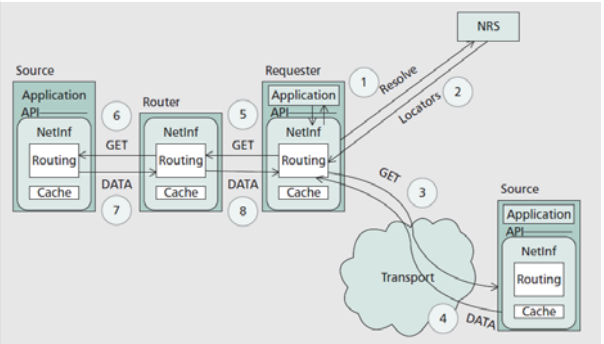
15

DONA

(Route by name + IP based data forwarding)



NetInf (mixed)



Very Different Design Which is the best?

It is all about assumptions and trade-offs

Debate: SC v.s. HR Naming

- **SC name**
  - relies on prior provenance and relevance in the first place
  - Hard to be aggregated – scalability problem
- **HR name provide weak intrinsic relevance**
  - By adding self-certifying component, public key or its pointer which sign the NDO, to provide availability
  - Easy to be aggrated

Debate: Caching

- **S. Shenker et al [sigcomm'13] argue that edge-based caching is enough for ICN based on a dataset from Akamai, where requests follow zipf distribution**
- **C. Imbrenda et al [ICN'14] conclude very different conclusion based on the dataset from access and back-haul Orange S.A., wherein requests follows combination of Weibull (head&tail) and zipf (middle)**

Edge or In-network ?

- **Thus, request distribution is the key factor for caching storage placement – We don't know the real ICN traffic distribution yet!**

Conclusions

- **ICN is about content delivery**
- **ICN Design**
  - Naming the data
  - Retrieving the Data
  - Securing the Data
- **Different ICN designs**
- **ICN is still on-going research.**
  - There are not too many consensuses on the designing of ICN, even for those fundamental design
- **Reference**
  - Xiaoke Jiang, Jun Bi, et. al., A Survey on Information-Centric Networking: Concept, Design and Debate, China Communications (IEEE), Vol.12, No.7, 2015