Future Internet and Data Service

2010. 10.28 목 KOREN Workshop

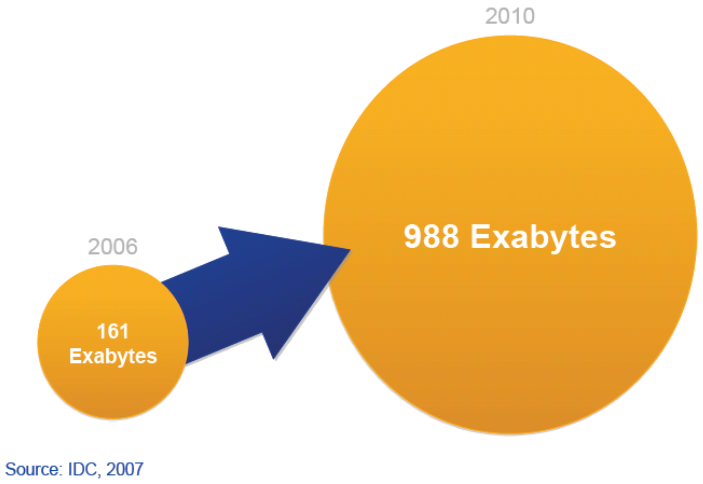
삼성종합기술원 최진혁
jinchoe@samsung.com

contents

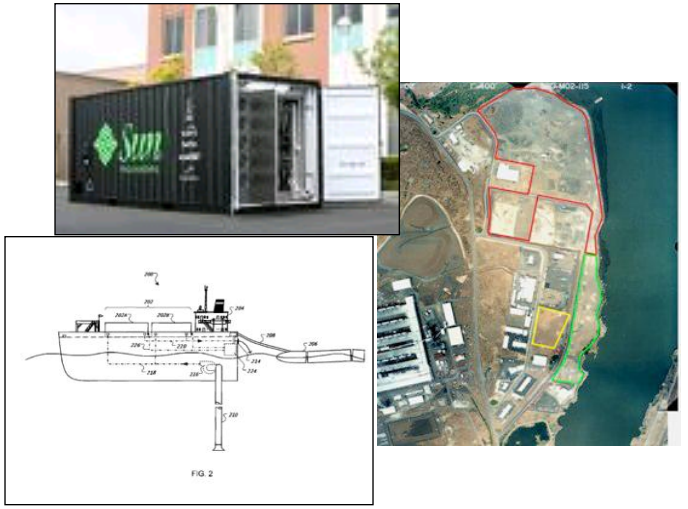
- current status
- needs & limits
- existing schemes
- generic data service model
- new data service architectures
  - DONA, CCN, PSIRP, NetInf
- summary

rise of data

- Data volume increase
  - Multi-media, UGC, streaming service in Internet
  - UC Davis expects 10 peta byte in 2010 (51 tera byte in 2005)
  - PPLive: 110m users, 600+ channels, 20%-30% outside of China
- Data center expand
  - Arms race in the cloud: many millions of servers.
  - Power consumption (1.5% of US electricity production 2006)
- Data dominance in Internet
  - The overwhelming use (>99% by most measurements) of today's networks is to acquire named chunks of data.
  - Making data is easy, Moving not.

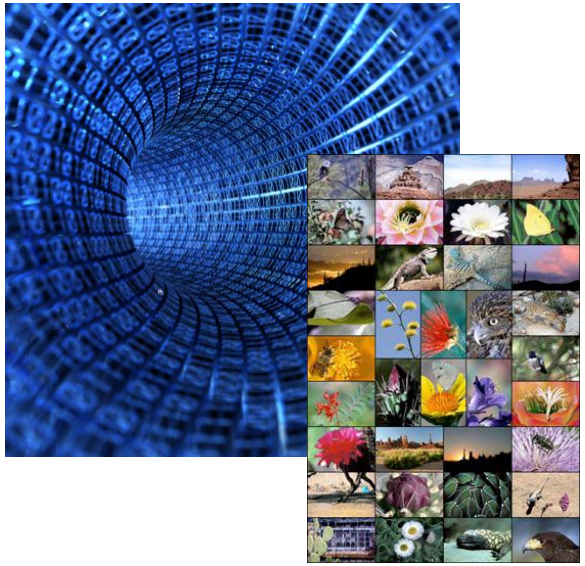


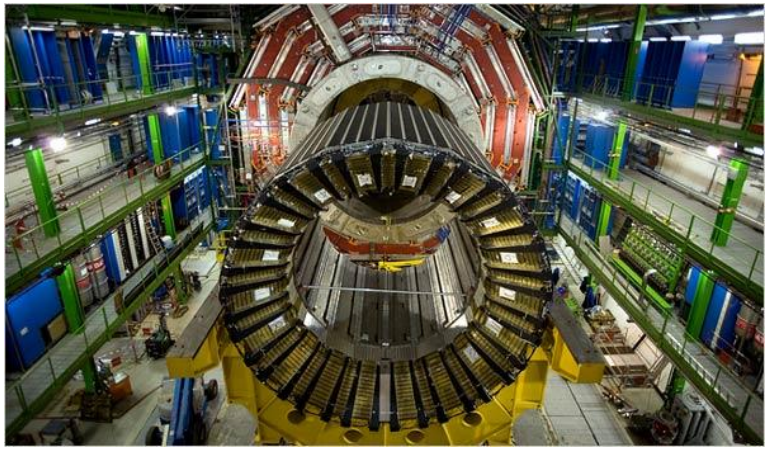
Source: IDC, 2007



current situation

- Is the Internet the preferred medium for distributing bulk (delay tolerant) digital content?
  - Not beyond a certain size. Either private network or postal
  - Ex) movies, data backup, data replication
- NETFLIX
  - 10M + users & 1.5 M DVDs per day → 2.5 PB/ day
    - All US P2P traffic 14 PB/ day (Cisco)
  - Postal still carries vast amount of multimedia traffic
- CERN
  - 20 TB/day scientific data, Private network
  - Large Hadron Collider (LHC) Computing Grid
- Data Center Replication
  - CDNs ships TB/ day logs via postal mail





needs

- Requirements change.
  - Internet was designed for host-to-host communication, i.e. conversation model. Acquiring named chunks of data is not a conversation, it's a dissemination.
- Persistence of names
  - Independent of location or host. Follow data migration
  - Today: HTTP redirects, email forwarding
- Availability of data: (both latency and reliability)
  - Take advantage of replicated data
  - Today: Akamai/BitTorrent
- Authenticity of data:
  - Know that data came from intended source
  - Today: securing the channel (IPsec, TLS), or PKI

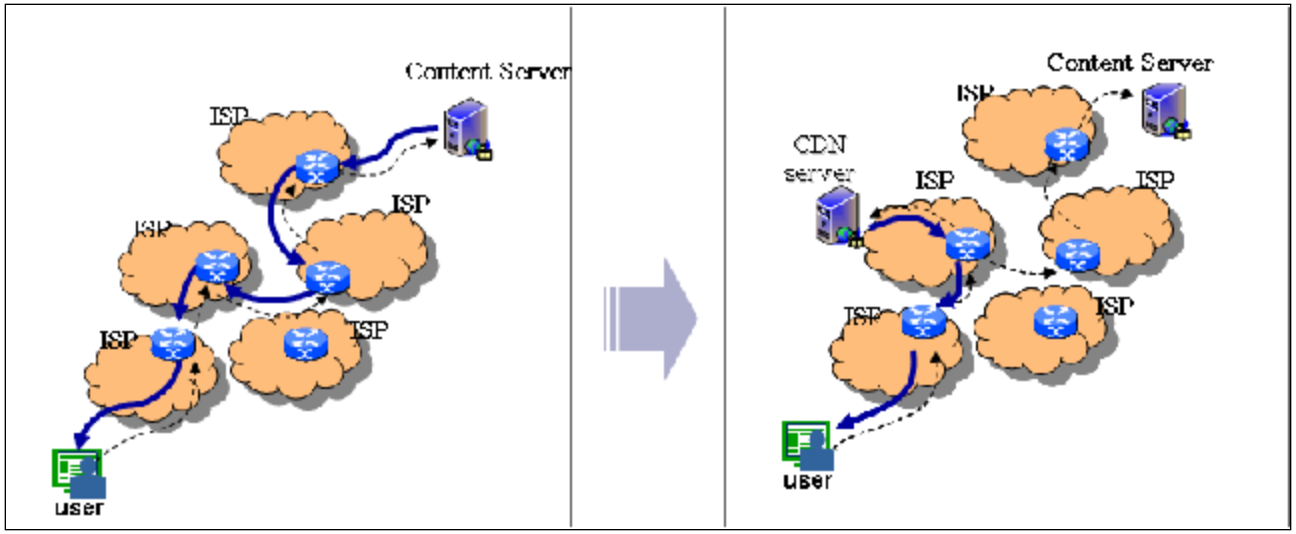
problems

- Internet was designed for conversation model. Mismatch between usage and design
  - data migration and replication unnecessarily hard
  - well-suited for short bursty traffic, not for long duration flows

- Limits
  - Rigid and Weak Naming: hostname/path
    - Ties data to host, making migration/replication hard
  - Protocol Mess: e.g., DNS, TCP
    - TCP session is tied to IP addresses, not DNS names, because DNS was developed later.
  - Lack of authentication
    - Channel secured, not data
  - Difficulty with bulk content beyond a certain size
    - Moving 1.2 TB of bulk data via Internet was a nightmare.
    - Postal networks are more efficient & cheaper for bulk data

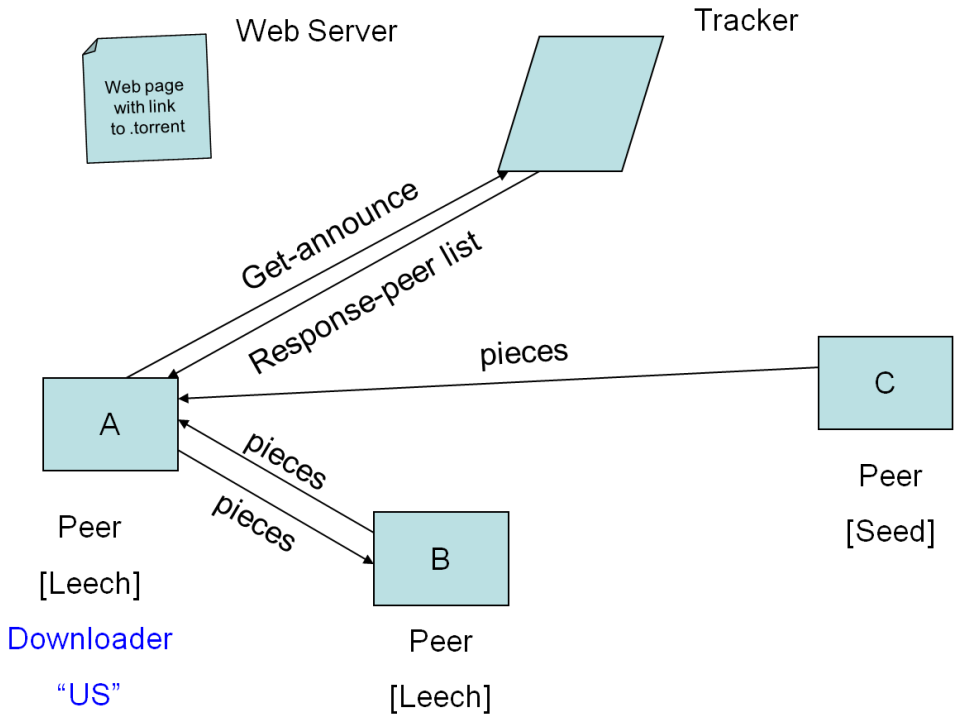
Content distribution networks (CDN)

- Content distribution networks are coordinated caching systems.
- Replicating content over a large number of distributed servers without relying on centralized servers
- CDNs are a multi-million-dollar business already, Akamai



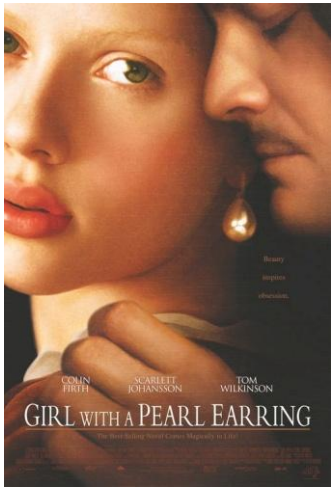
BitTorrent

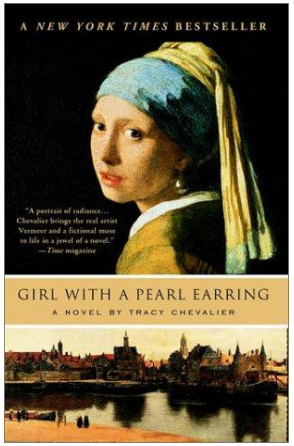
- Peer-to-peer file sharing protocol by Bram Cohen for large amount of data distribution (approximately 35% of all Internet traffic in 2002)
- Tracker, Swarm, Pieces of files
- Incentive based on Tit-for-tat, Uploading while downloading



generic data service model

- Data





- Name (tied to a location/ host)
  - www.vermeer-foundation.org/Girl-with-a-Pearl-Earring-c.-1665.html
- Location
  - 143.21.23.43

generic data service model

- Intent Resolution Service (IRS)
  - Translating user's intent to meta data.
  - Intent specifies what an user is looking for and could be expressed by a set of keywords.
  - Meta data uniquely specifies data and provides the information for data retrieval, such as name (at least), publisher or signature.
  - Meta data may also facilitates data organization, information network, film/ picture/ book, private/ public.
  - IRS could be a search engine.
- Meta-data Resolution Service (MRS)
  - Resolves the location of a entity or entities which can serve the data.
  - Tracker in bitTorrent, DNS
- Data retrieval
  - Contact the serving entities (server or client) to retrieve the data.

additional mechanism

- Persistence of names
  - Location independent name
- Availability of data: (both latency and reliability)
  - Replication
  - Swarm
  - storage and forward
- Authenticity of data:
  - self-certifying with meta-data

research areas

- Naming
  - Location independent
  - Self-certifying (with meta-data)
- Intent Resolution
  - Resolve meta-data from user's intent
  - Information model, meta-data attachment
- Meta-data (Name) resolution
  - Resolve data location from meta-data (Name)
  - Name based routing, DHT, DNS type
- Data retrieval
  - Replication
  - Swarm
  - Store & forward

new data service architectures

- Data Oriented Network Architecture (DONA)
  - By Teemu Koponen, Scott Shenker and Ion Stoica (UC Berkeley)
  - A FIND project: A New Approach to Internet Naming and Name Resolution
- Content Centric Network (CCN)
  - By Van Jacobson (PARC)
  - http://www.ccnx.org/
- Publish-subscribe Internet Routing Paradigm (PSIRP)
  - A FP7 project
  - http://psirp.org/
- Network of Information (NetInf)
  - A sub-project of 4ward, a FP7 project
  - http://www.4ward-project.eu/

PSIRP
PUBLISH-SUBSCRIBE
INTERNET ROUTING
PARADIGM

new data service architectures

- Content-centric / data-oriented paradigm
  - Information is indexed by keys & retrieved by subscription.
- Publish/Subscribe communication model
  - When a node publishes data, no data transfer actually takes place (the rendezvous system is informed of its existence).
  - Only when a node subscribes to a named piece of data, the network finds the publication and creates a delivery path from the publisher to the subscriber(s).
- Protocols are declarative
  - Say what you want, not where/who to get it from
  - Get/ Fetch/ Retrieve ()
- Data is self-certified
  - Self-validating data (hash, signature, PKI)
  - Secure the data, not the channel
- Content-routing
  - Routing data messages based on their content (semantic & syntactic) rather than network host addresses.
  - “name-based routing” where name ≠ host/interface name.
- Network nodes more than simple routers
  - Are caches of content, indexes, and buffers.
  - Forward information while caching,
  - in the style of DTNs and P2P.

new data service architectures

Original Internet	Content Centric/ Data oriented
Sender	Content Provider (Publisher)
Receiver	Content Consumer (Subscriber)
Sender based control	Receiver based control
Host-to-Host communication	Service/ Information access, Data retrieval
One-to-one conversation	Pub/ Sub Uncoupled Publisher & Subscriber
Channel/ Container security via authentication	Self-certifying Data/ Content
Unicast	Unicast, Multicast, Anycast
Host naming Look-up oriented naming	Data naming Data oriented search naming
Destination host routing	Content based routing
Content agnostic network	Content aware network (router)

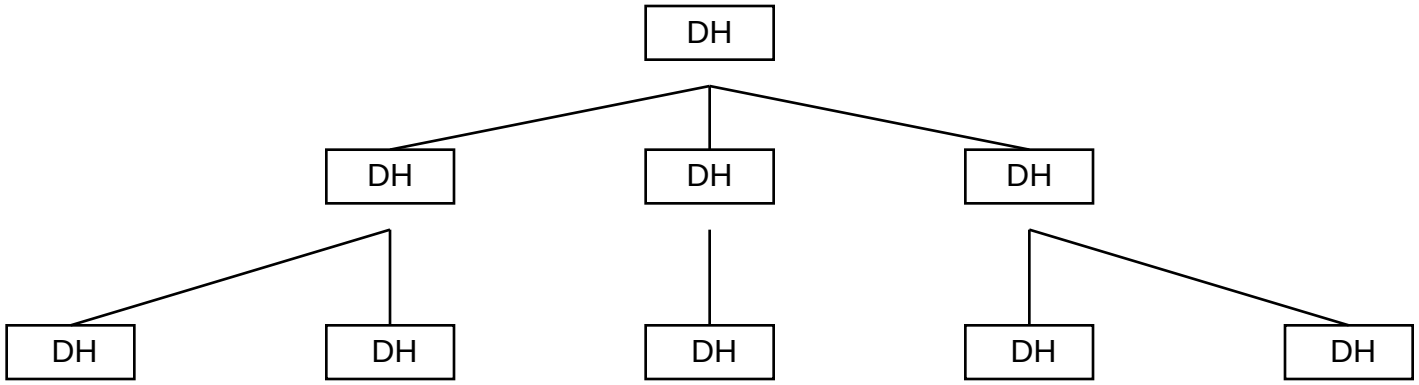
DONA

- Flat & self-certifying naming
  - Data associated with principal P
    - with public key K_p & identifier $P^{\#} = Hash\ (K_p)$
    - ex) principal: CNN, public key K_{CNN}, $CNN^{\#} = Hash\ (K_{CNN})$
  - Names are of the form: $P^{\#}$: label
    - $P^{\#}$ represents a data owner & (Unique) label a specific data.
    - Name not tied to location, so naming isn't rigid
    - Users acquire $P^{\#}$: label via outside mechanism (google)
  - Self certifying $< K_p$, Data, Signature>
    - Upon requesting a data with the name $P^{\#}$: label
    - returns a data with meta-data $< K_p$, Data, Signature>
      - Signature = $K_{private}$ (Data)
    - Can verify the data by checking
      - Hash (K_p) = $P^{\#}$ & K_p(Signature) = Data
  - Names take care of persistence, authenticity

DONA

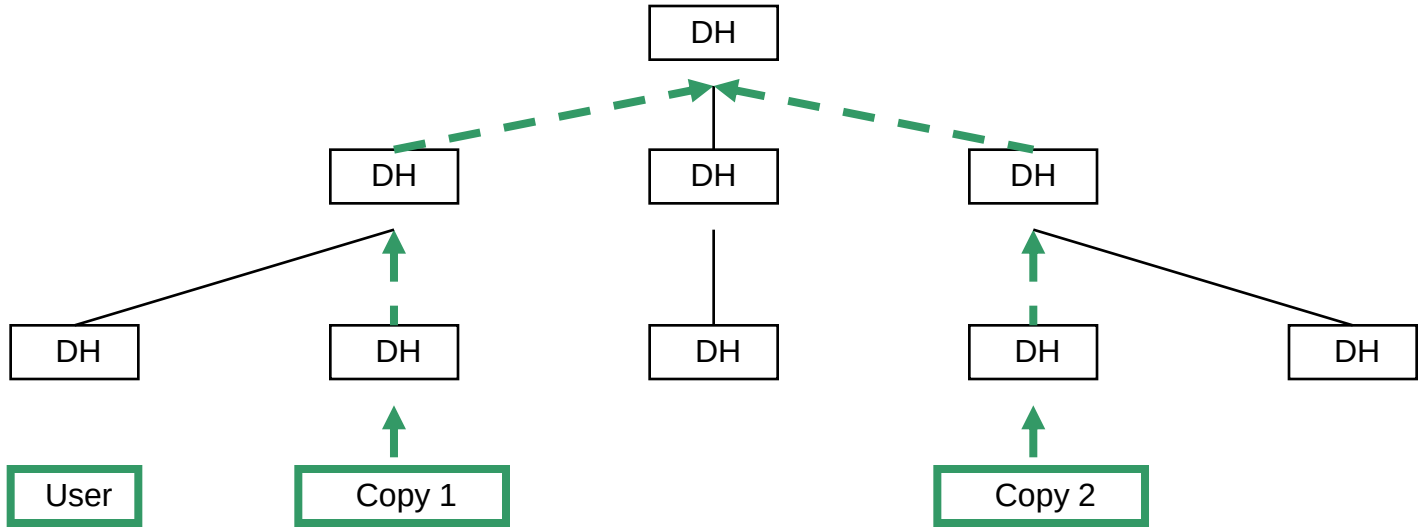
- Name resolution
  - Routing by name. Anycast Name Resolution (ANR)
- Data Handlers (DHs)
  - DHs do name-based routing and caching
  - a logical DH per administrative unit
    - DH hierarchy according to AS hierarchy (and finer grain below)
- New network primitives
  - Fetch(name): data request
  - Register(name): offer to serve data (authenticated)

DONA



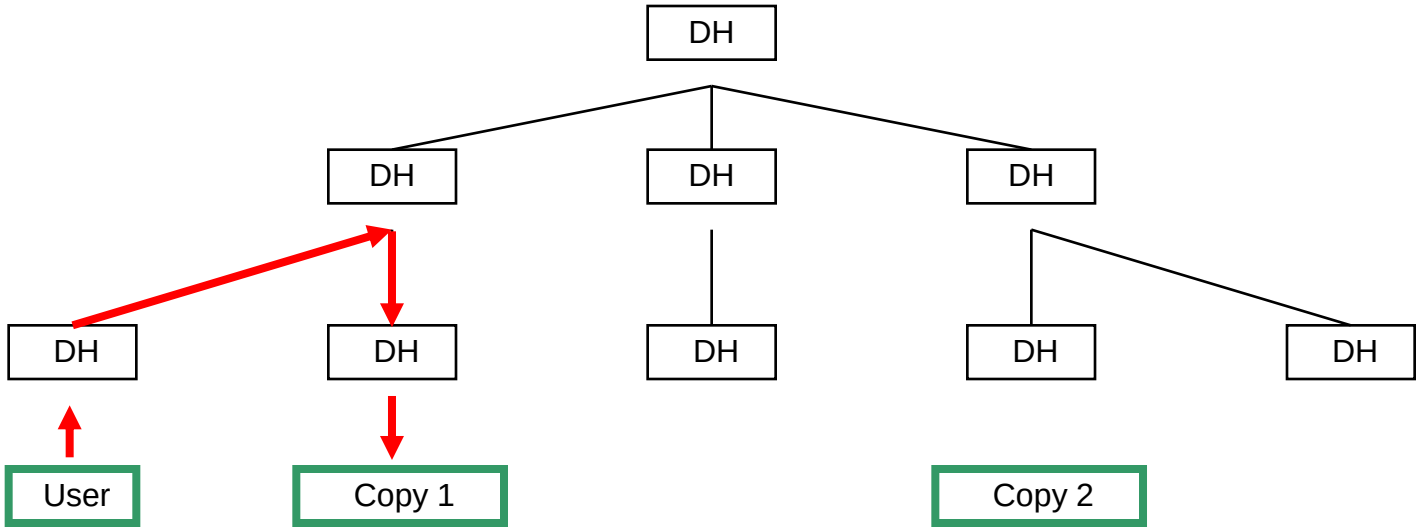
- ANR (Anycast Name Resolution)
  - Need to implement anycast routing at DONA-layer
  - Use DH hierarchy to guide routing

DONA



- Data Register
  - DHs forward register commands to parents and peers
- Scaling: DHs only hold state for items below
  - core: few TBs
  - edge: typically far less than a GB

DONA



- Resolution procedures
  - Clients configured with local DH to send their fetch requests
  - DHs respond to fetch if data is in cache
  - Otherwise, DH routes fetch towards nearest copy of data by sending to a next-hop DH
  - If name isn't in routing table, fetch routed upward to the core.

CCN

- Data dissemination with Request & response.
  - acquiring named chunk of data, is not a conversation. Unwise to tie data to a fixed location or host.
  - Consumer ‘broadcasts’ an ‘interest’ over any and all available communication media:

    want ‘/parc.com/people/van/presentation.ppt’

  - Interests identifies a collection of data – all data items whose name has the interests as a prefix
  - Anything that hears the interest request & has a valid copy of the data can respond.

    hereis ‘/parc.com/van/presentation.ppt’ <data >

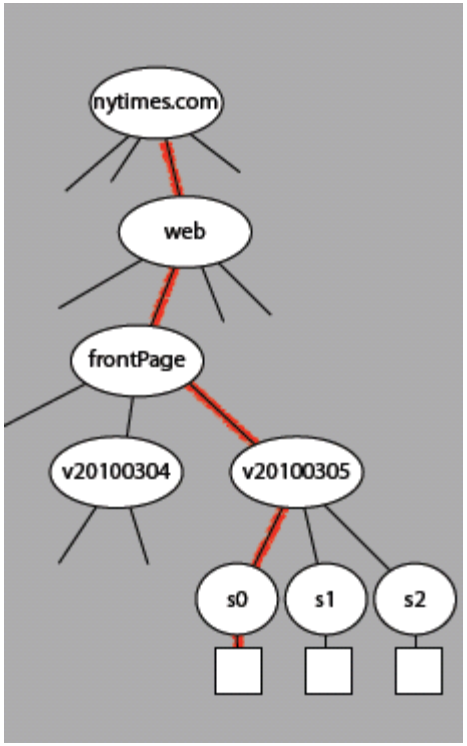
  - The returned data is signed to validate its integrity & association with the name.

CCN

- Name

  /parc.com/van/cal/v3/s0/xode3fd…

- Name tree solves ‘discovery problem’



- Messages

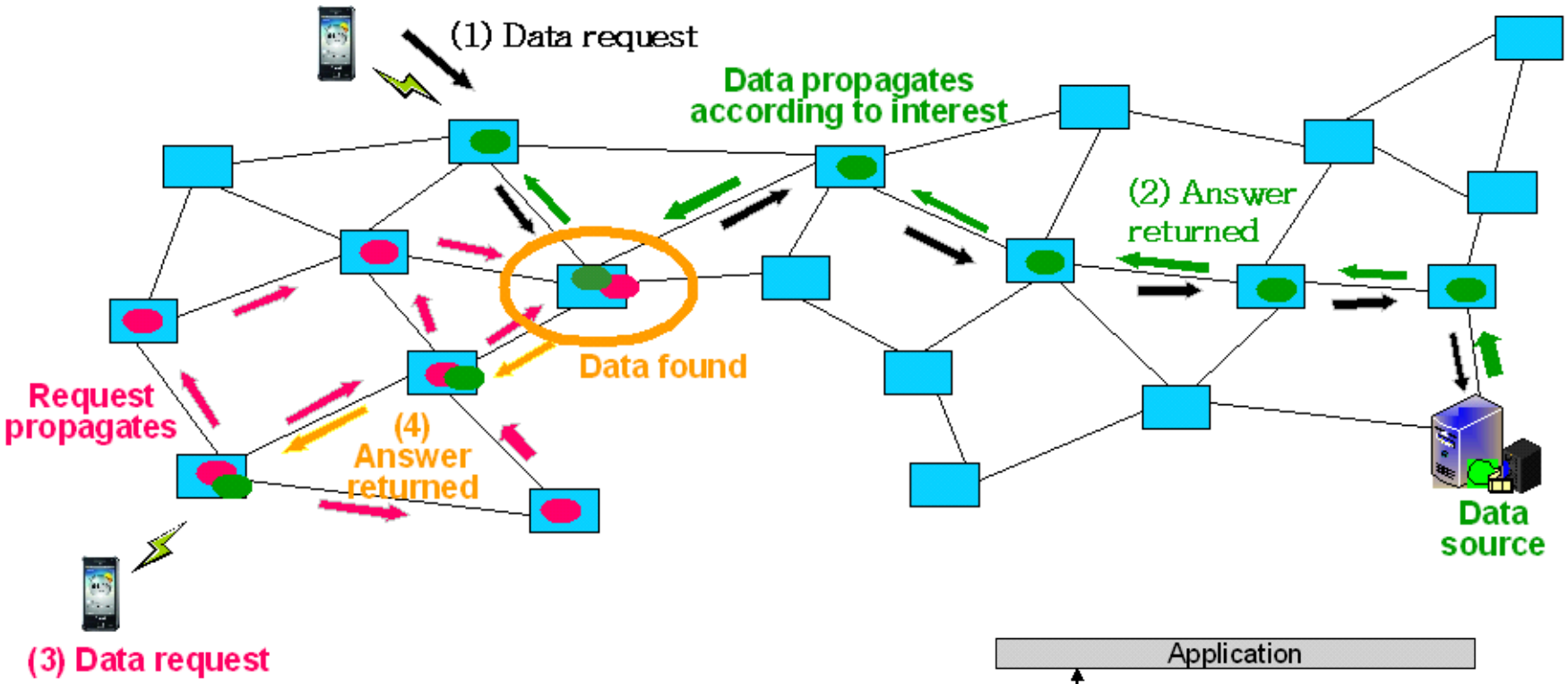
“interests”

Content Name
Cryptographic Signature

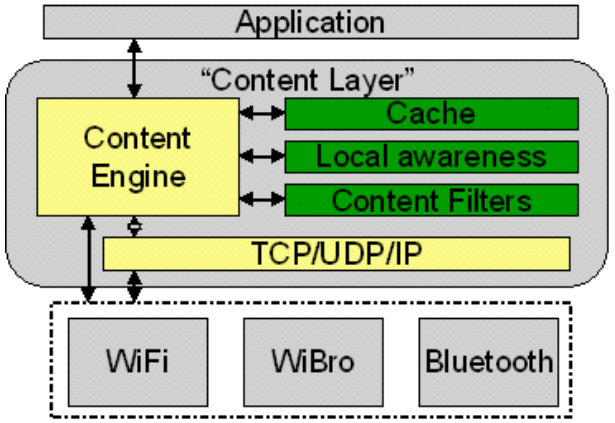
“data”

Content Name
Cryptographic Signature
Content Data

CCN

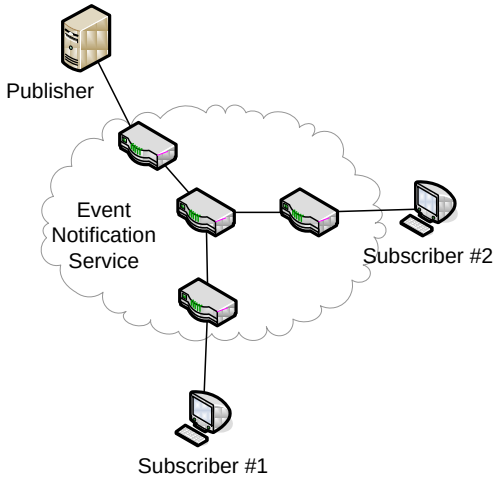


- Initial Content Request
- Content transmission & Caching from Content source.
- Subsequent Content Requests
- Cached Contents Returned



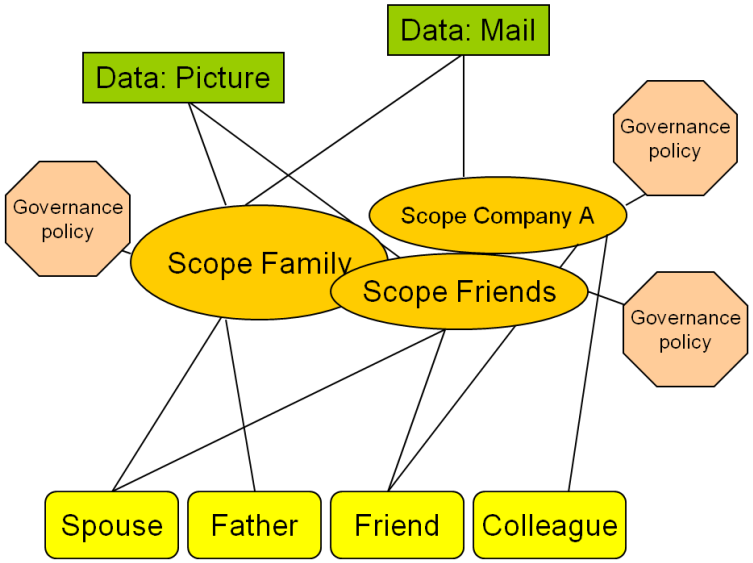
PSIRP

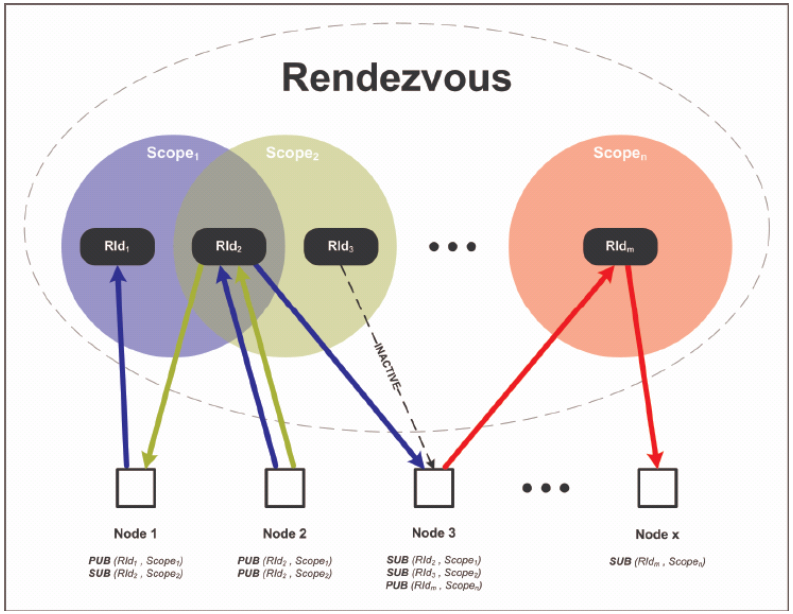
- Information becomes available through publications
- Endpoints:
  - Publishers: data owners
    - Provide pieces of information in the form of publications
  - Subscribers (data consumers)
    - Express interest in pieces of information via subscriptions
- Network:
  - Event notification service (broker substrate): matching publications and subscriptions
- End-to-end decoupling
  - Publishers/Subscribers need not be aware of corresponding Subscribers/Publishers
  - Asynchronous communication
- Multicast
  - Multiple subscriptions can be grouped
  - brokers merge data streams
  - Norm in pub/sub
- Caching
  - Pub/sub state and multicast suitable for in-network caching



PSIRP

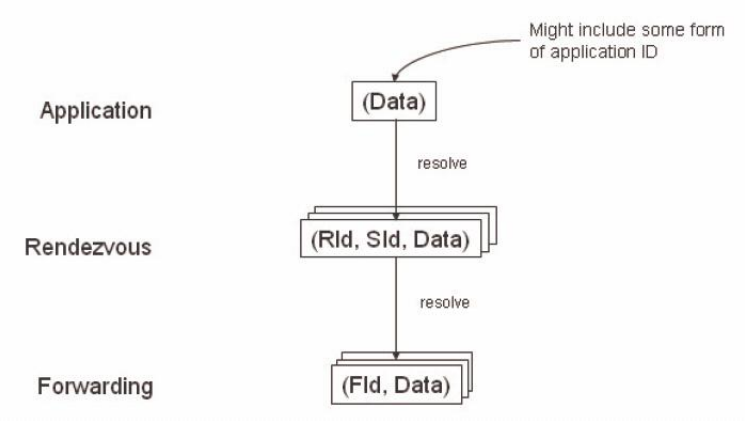
- Each publication is identified by a unique identifier (rendezvous identifier – RId)
- Information is organized in networks called scopes, each one identified by a scope identifier (SId)
  - Physical networks or Social one, e.g. university campus or social network
  - Used for: locating information (context), access control
  - Hierarchically organized (algorithmic identifiers, AIds)
  - Scope builds information network network
- Rendezvous point (RP) is responsible for the specific scope

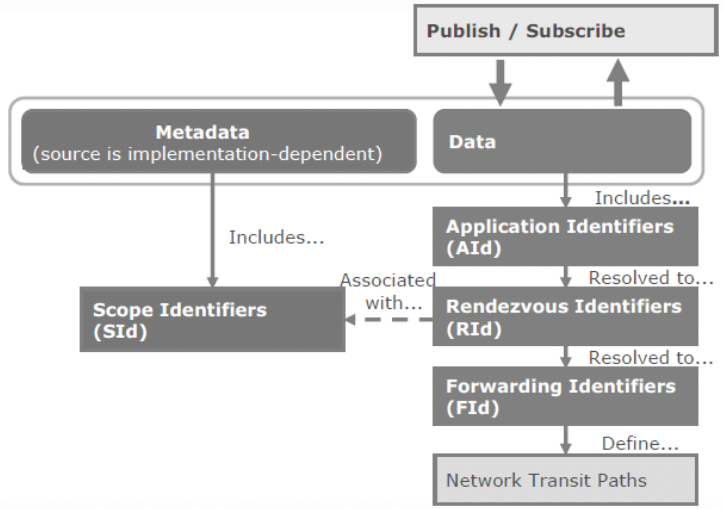




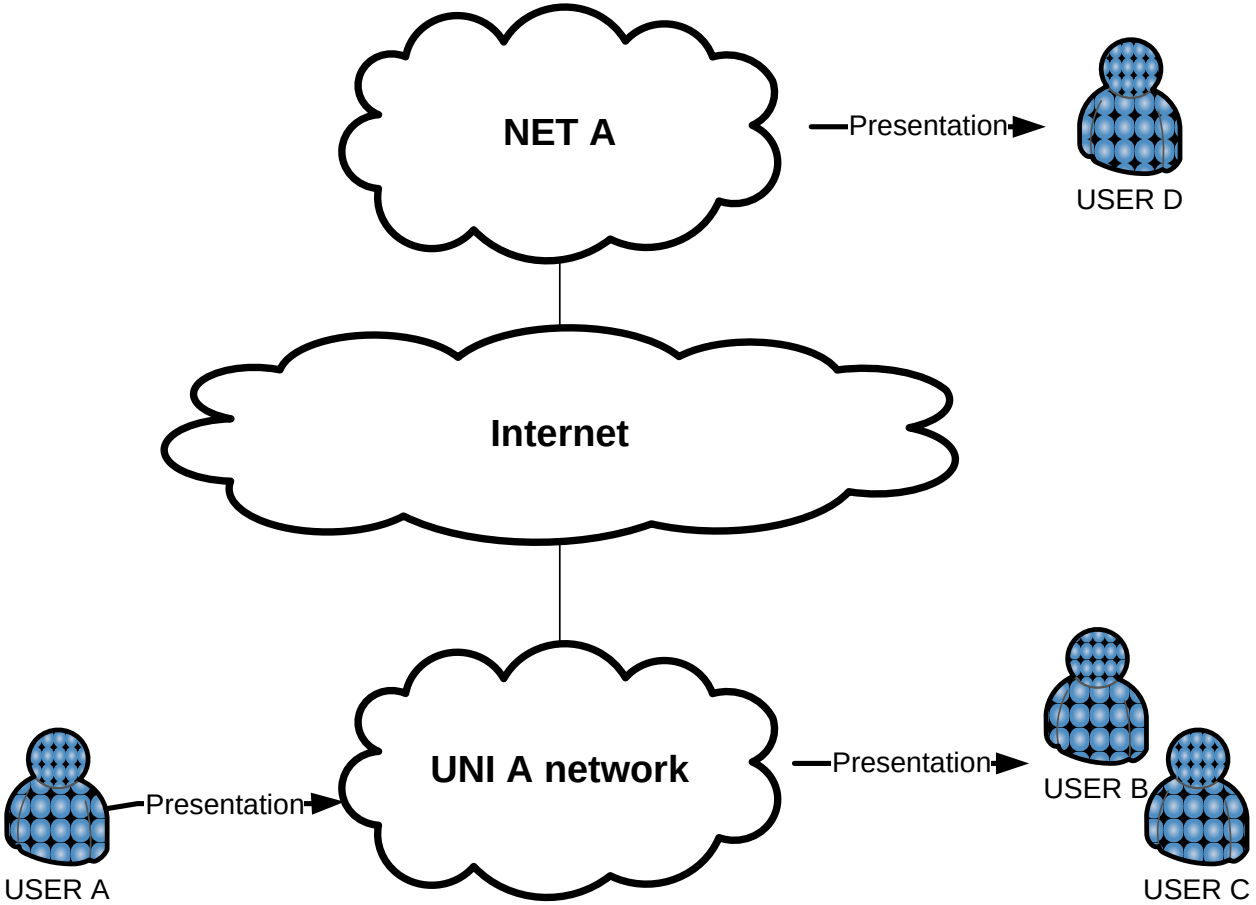
PSIRP

- Publishers initially publish metadata to the rendezvous point (RP) of the information
  - RP responsible for the specific Sid
- Information is accessed through subscriptions issued to the rendezvous point (RP) of the information
- RP is responsible for *matching* publications with subscriptions
  i.e. matching RIds within a certain scope (SId)
- Information dissemination is achieved using a stack of forwarding identifiers (FIds) similar to MPLS
- Data do not necessarily pass through RP
- All identifiers are flat and location independent
- SIds and RIds can be of local or global significance

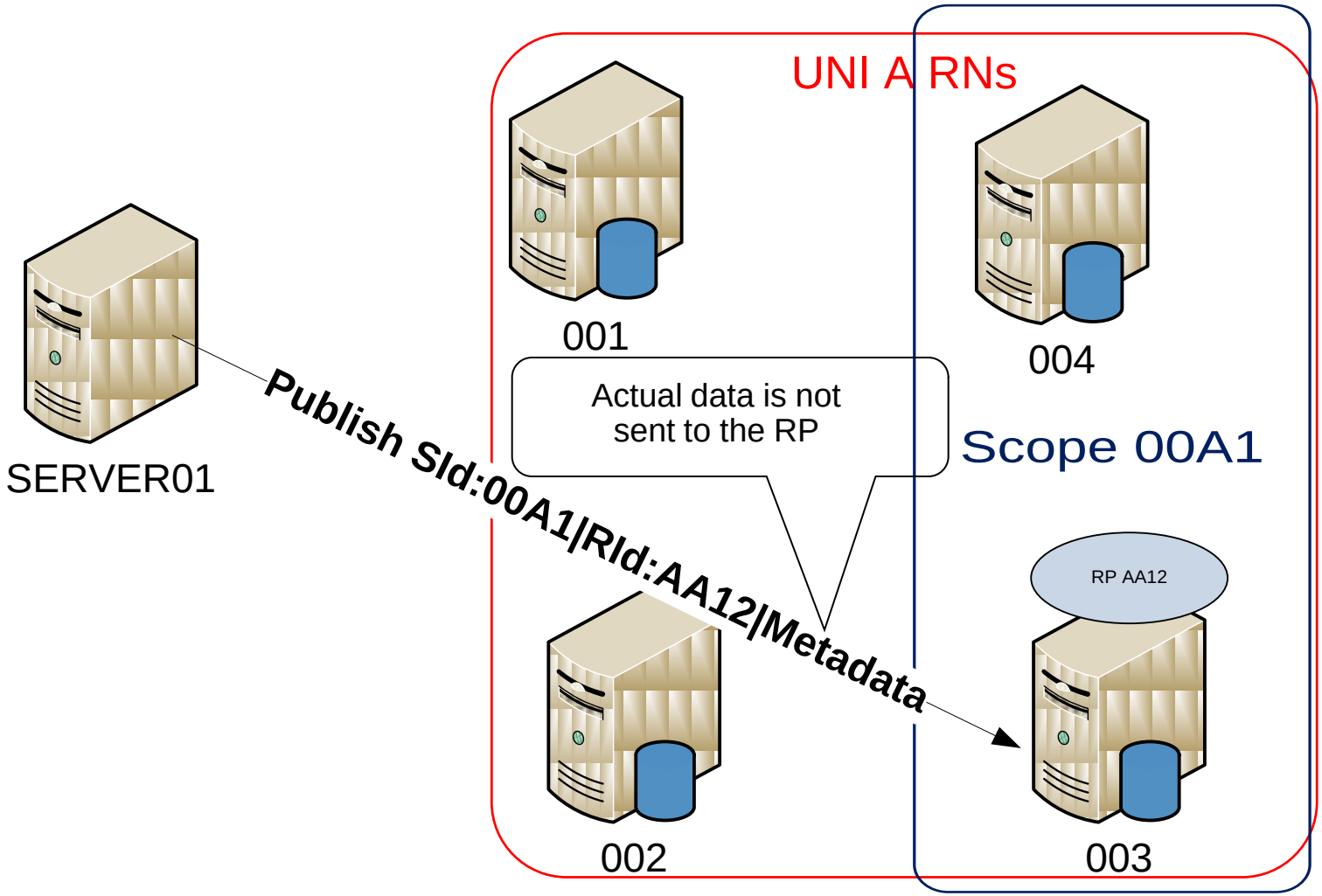




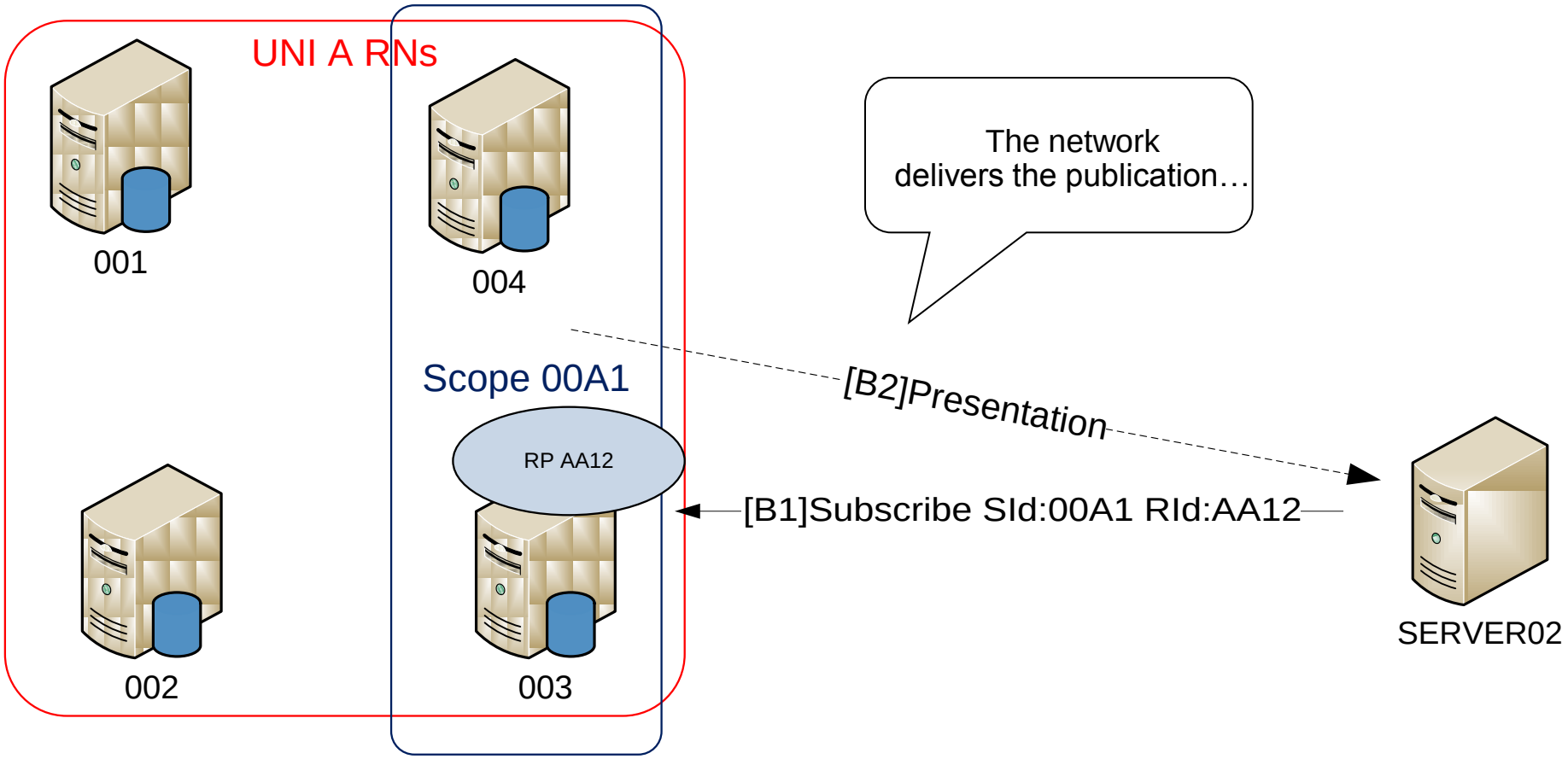
PSIRP Usage Scenario Overview



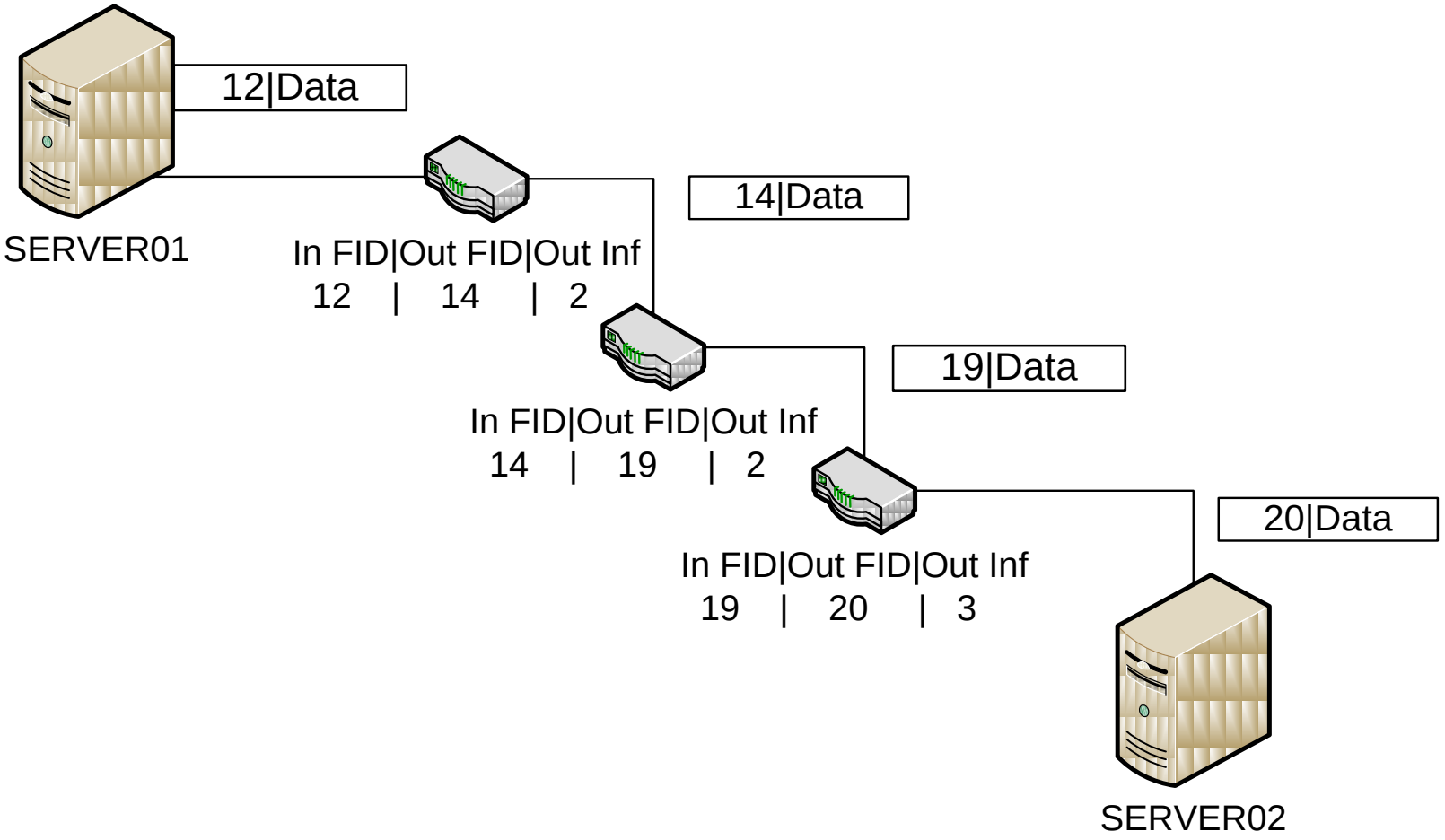
PSIRP publish



PSIRP subscribe

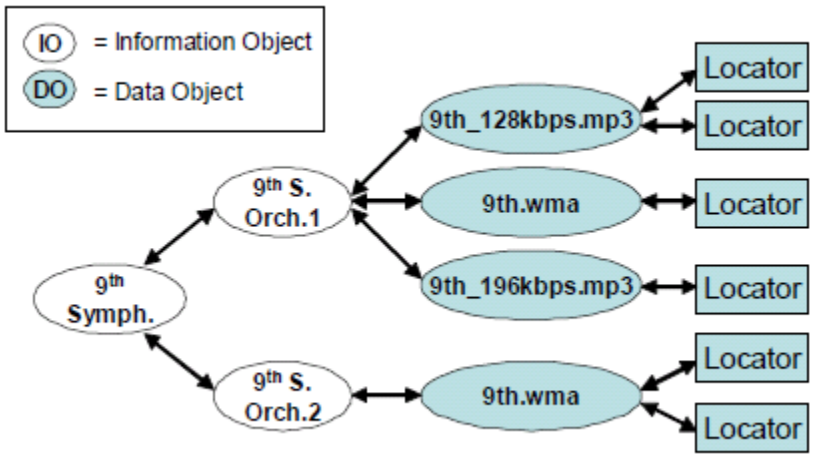


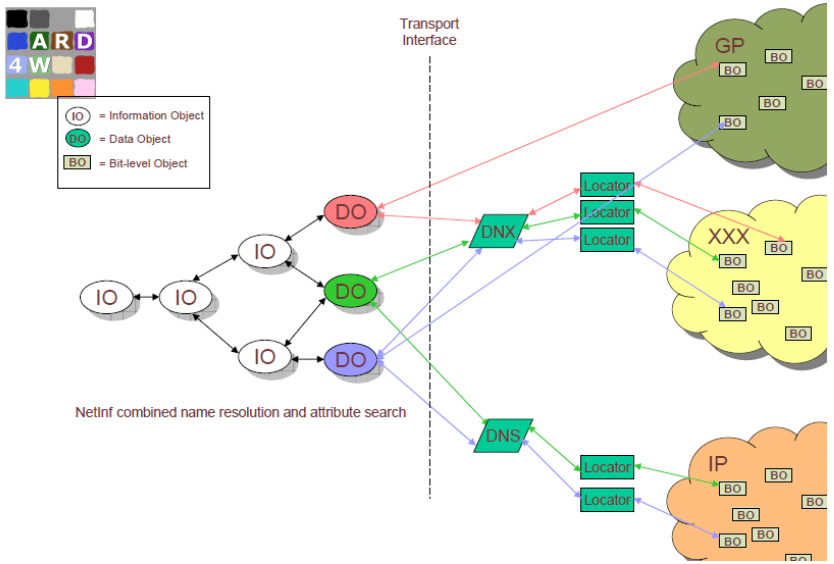
PSIRP forwarding



NetInf

- Organize Information – IO, DO and BO
  - Enables efficient information dissemination
- Information Object (IO)
  - a set of attributes defining the semantics of a data object.
  - IO may refer to a piece of music, a film or a webpage.
  - Can be static, dynamic or realworld objects, including streams and services
- Data Object (DO)
  - Sub-class of IO holding attributes for bit-level objects and pointer(s) to the actual data.
- Bit-level Object (BO)
  - Sub-class of IO holding attributes for bit-level objects and pointer(s) to the actual data.



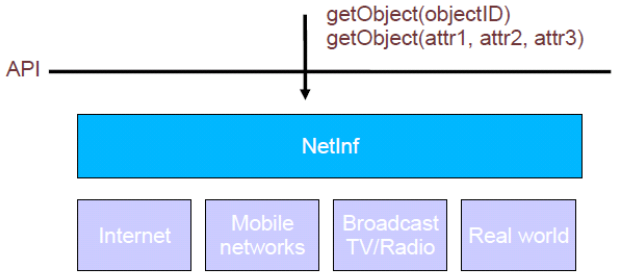


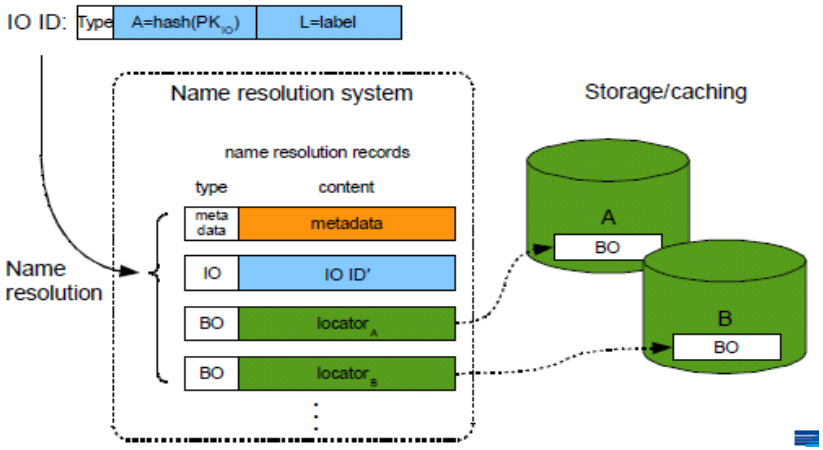
NetInf

- NetInf Naming scheme

Type	A = Hash (P_{IO})	L = {identifier attri}

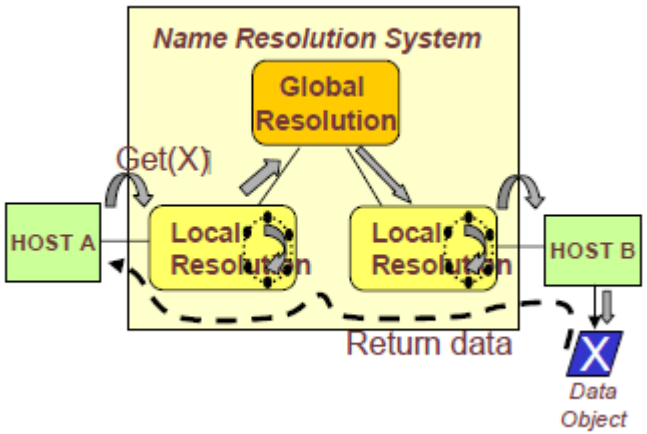
  - Type: Defines the format
    - e.g. Hash algorithm used (SHA1, MD5, …)
  - Authenticator (A): Binds the ID of the IO to a public key P_{IO}
    - Hash function used to compress length of P_{IO}
  - Label (L): Identifying individual object published by Authenticator
    - contains a number of identifier attributes associated with an IO
    - (A, L) combination needs to be globally unique

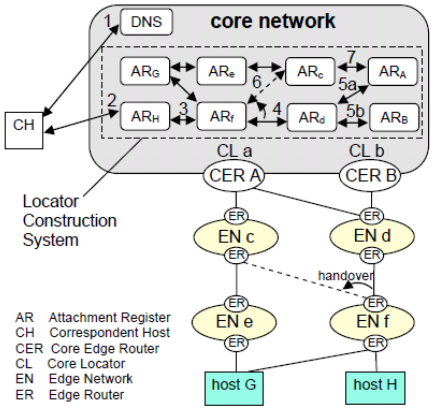




NetInf

- Two key processes:
  - Name resolution: provides network locator(s) for the object ID
  - Routing: forwards request towards storage location and returns data
- A Name-based routing approach is being pursued by NetInf
  - Combines name resolution and routing into single process
  - Routing of requests for Data Objects (e.g. put/get) based on their IDs instead of locations, directly mapping DO identifiers to a route.
  - The routing mechanism has to ensure fast data forwarding, while keeping table size manageable in spite of the huge entry size.
- Multiple DHT & LLC (Late Locator Construction)
  - Under consideration for name-based routing schemes for NetInf
  - Multiple DHT
    - Hierarchical DHC perform locality-aware name-based routing in hierarchically structured domains
  - LLC
    - Attachment registered to keep track of immediately attached neighbors
    - Hierarchical locator construction on demand at the time of session initiation.





summary

	DONA	CCN	PSIRP	NetInf
Name	$P^{\#}$: Label $P^{\#} = Hash(K_p)$ with public key K_p	/parc.com/van/cal/v3/s0/xode3fd	SId: RId Scope ID Rendezvous ID Forward ID	$P^{\#}$: Label $P^{\#} = Hash(K_{IO})$ with public key K_{IO}
Information model	Flat	Name tree	Organized by Scope	Information Object
Name Resolution	Anycast Name Resolution	Content based routing	Rendezvous Point	Name based routing (DHT, LLC)
Routing	Separate from name resolution	Interests towards publisher, data reverse path	Separate from name resolution	Separate from name resolution
Replication		Along the route		
Layer	Above IP	Above IP, for the time being	Replacing IP	

Work items

- Name
  - Location independent, Machine or Human friendly
- IRS
  - Google? Mr. Know All?
  - Content organization? Information network? Ontology?
- MRS
  - DNS type? Name routing?
  - DHT, LLC?
- Data retrieval
  - Client & server, P2P
  - Replication with synchronization, Swarm
- Bulk data transfer
  - Better than DHL
  - High-rate transport, Congestion avoiding transport for P2P
  - In-network storage, Internet postal service, Nano-data center

감사 합니다.