Hierarchical Indirection Multicast

2008. 8. 28

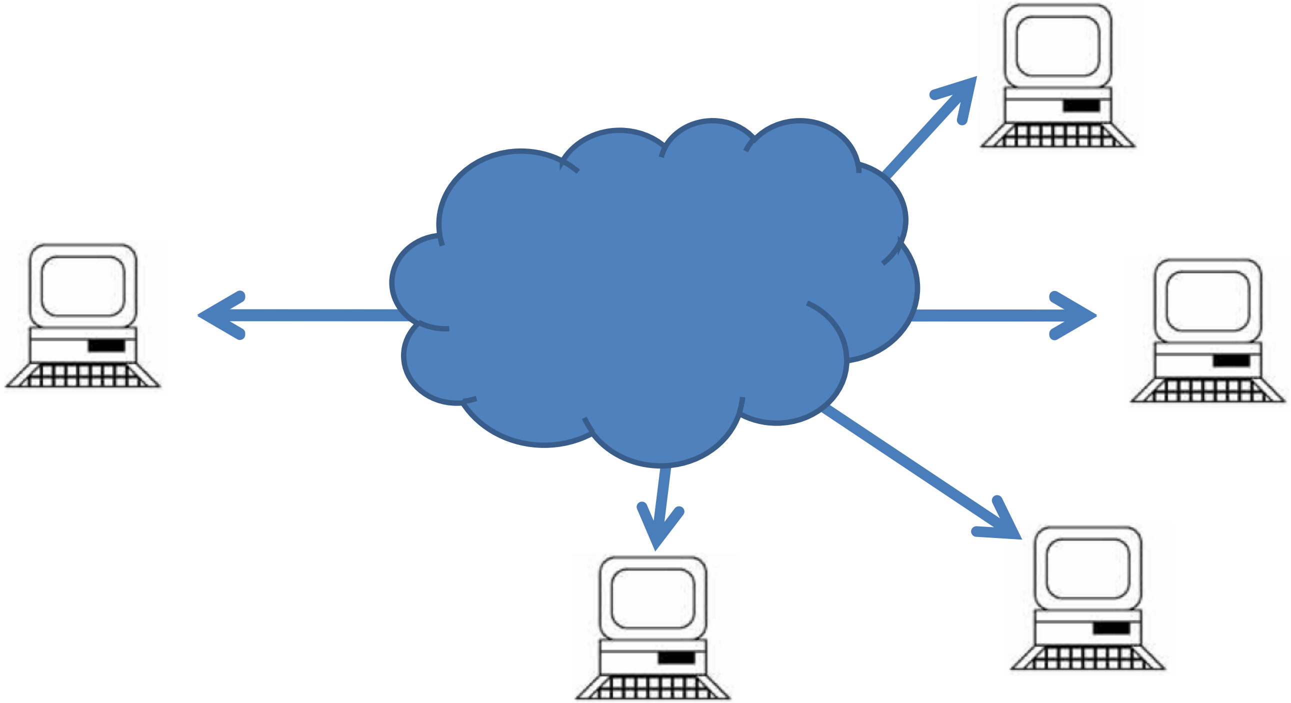
Park, Soyoung
sypark@mmlab.snu.ac.kr

Contents

- Internet Indirection Infrastructure
- End-host controlled multicast routing
- Hierarchical Indirection Multicast

Internet Indirection Infrastructure

Motivation



- Internet : designed for **unicast** between **fixed** locations
  - Today, we want **multicast**, **anycast**, **mobility**

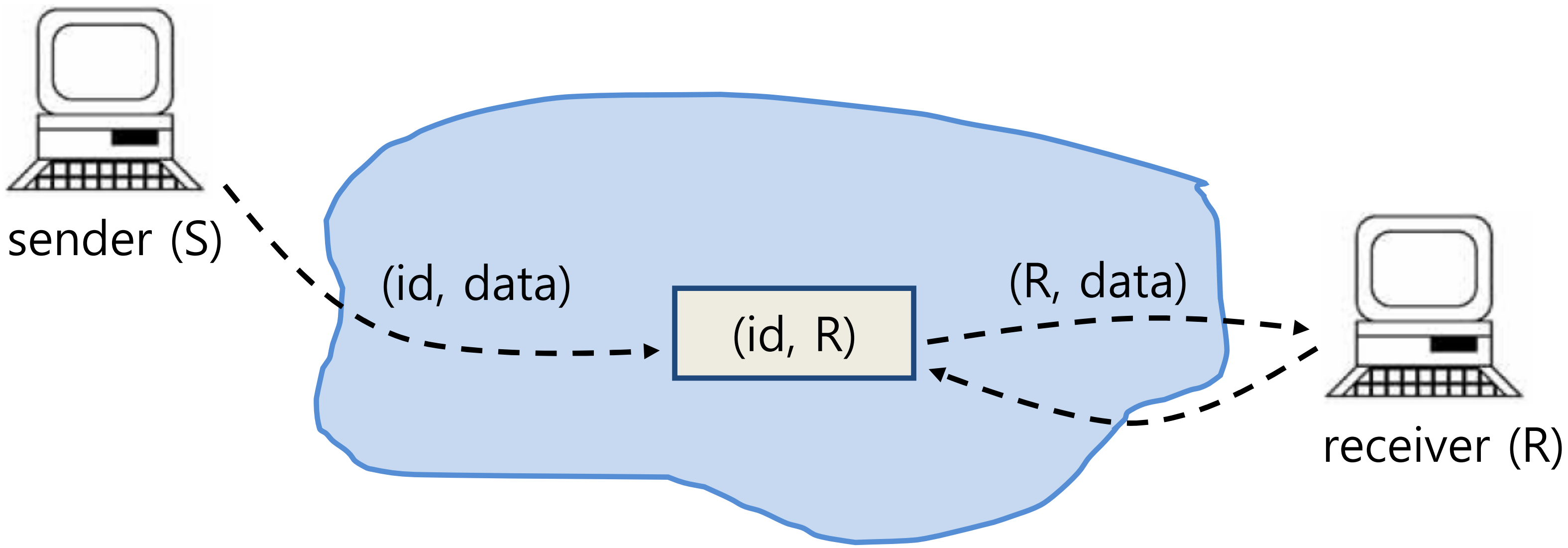
Propose

- **a single new overlay network**
  - Provides **a general communication abstraction**
    - unicast, multicast, anycast, mobility
  - Uses application layer solutions

How

- Using “**Indirection**”
  - the ability to reference something using a name, reference, or container *instead of the value itself* [wikipedia]

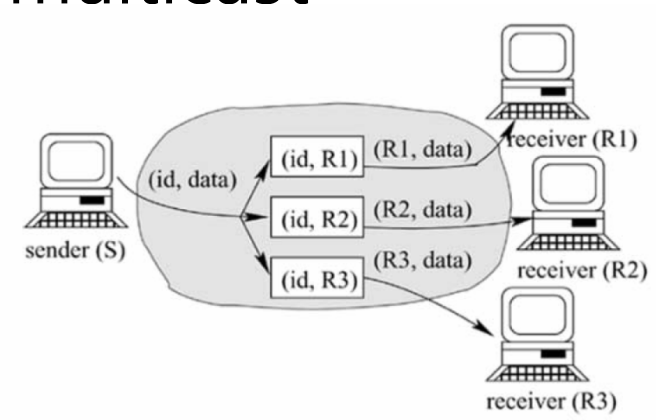
Internet Indirection Infrastructure



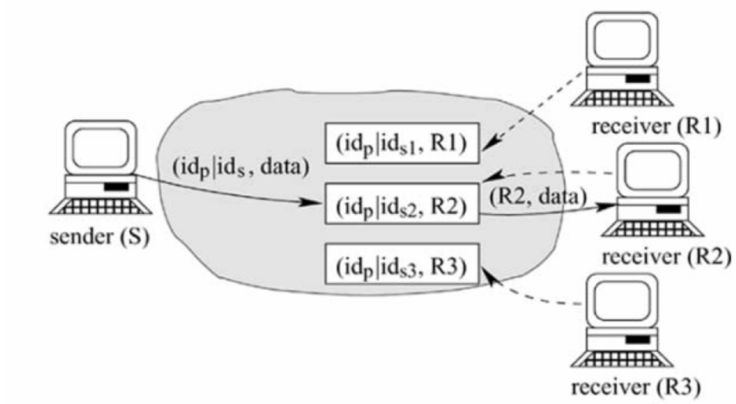
– Each packet is associated with an id
– Receiver : inserts trigger (id, addr) into the network
– Sender : sends packet (id, data)

Internet Indirection Infrastructure

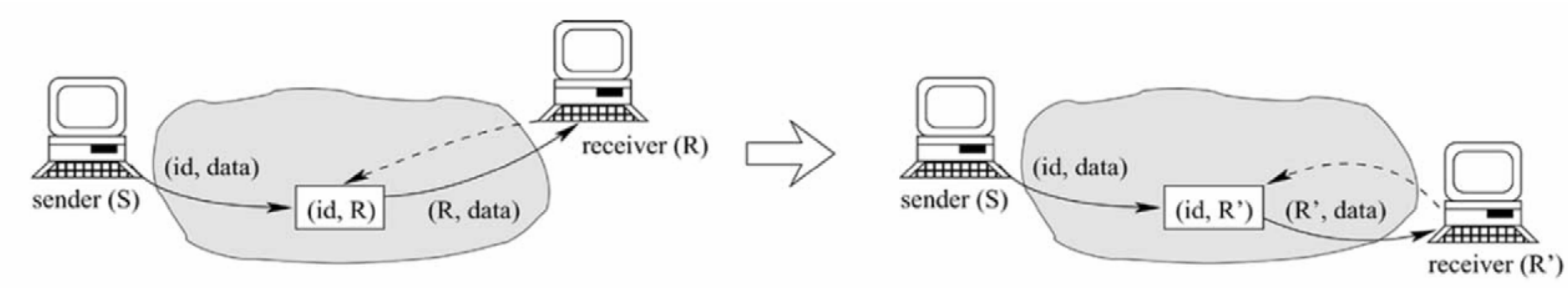
- multicast



- anycast



- mobility



Identifier

- i3 over Chord
  - Node identifier : i3 server identifier
    - i3 servers construct the Chord ring
    - generated the same as the node identifier in Chord
  - Key identifier : id
    - Generated using the random sampling
  - i3 servers store (id, the address of and end-host) pairs

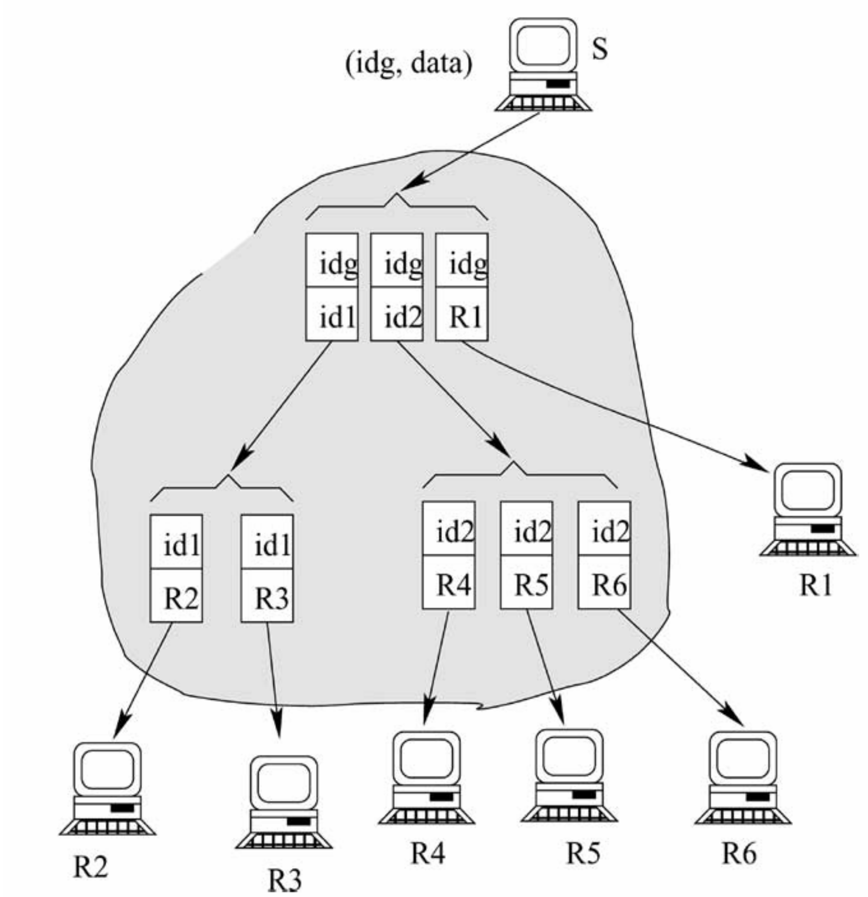
Chord	equivalents in i3
node identifer	i3 server identifier
key identifer	id
value	address of an end host

Random sampling

- In i3, to avoid triangular routing, end-hosts should **select an id which a nearby i3 server is responsible** for
  - id sampling
    - An end-host selects an id randomly, inserts a trigger and sends a packet with the id → measure RTT to the i3 server
  - Offline, chose id set with low RTT latency

Large Scale Multicast

- Problem
  - Only one indirection point to every member
    - not scalable
- Solution
  - A hierarchy of triggers (trigger chain)
    - $(id_g, R_i) \rightarrow (id_g, x_1), (x_1, x_2), \ldots, (x_i, R_i)$



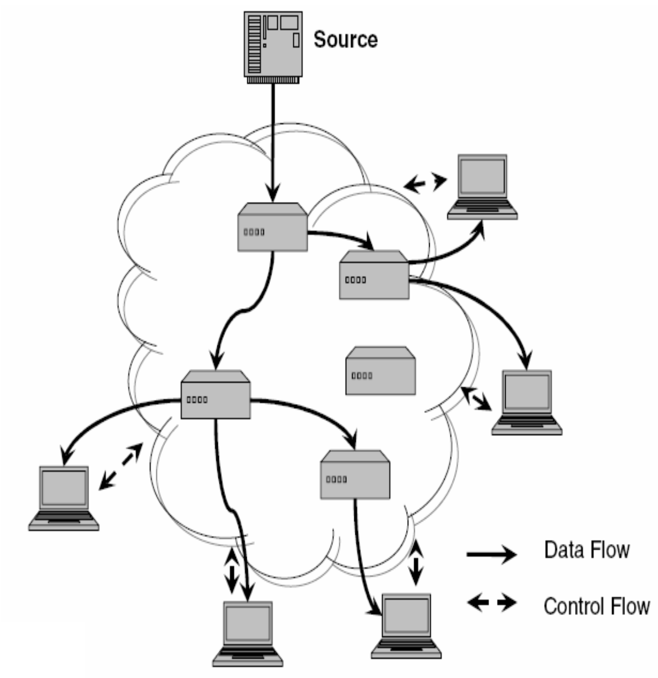
End-host controlled multicast routing

Motivation

- Existing multicast design
  - Infrastructure-based approach
    - More efficient and scalable multicast tree
  - Host-based approach
    - Easy to deploy and modify
    - More flexibility
      - ability to change the tree construction algorithms as per user needs
- Goal: realize the best of both approaches

How

- On top of i3, build multicast tree by using trigger chain
  - Infrastructure-based
    : export simple primitives
  - Host-based
    : use only the primitives to construct multicast trees in a distributed fashion



Join algorithm

- Selects a server to join (join_server) using triggers
  - Top-down method
- Makes a tree branch by inserting trigger
  - (join_server, close_server), (close_server, my_address)

Limitations of ECM

- The optimal multicast tree is not guaranteed
  - Greedy algorithm : The multicast tree is build differently according to the order of host join
- Physical topology is not considered at all
  - We propose a way to aggregate branches using physical topology

Hierarchical Indirection Multicast

Identifier: revisited

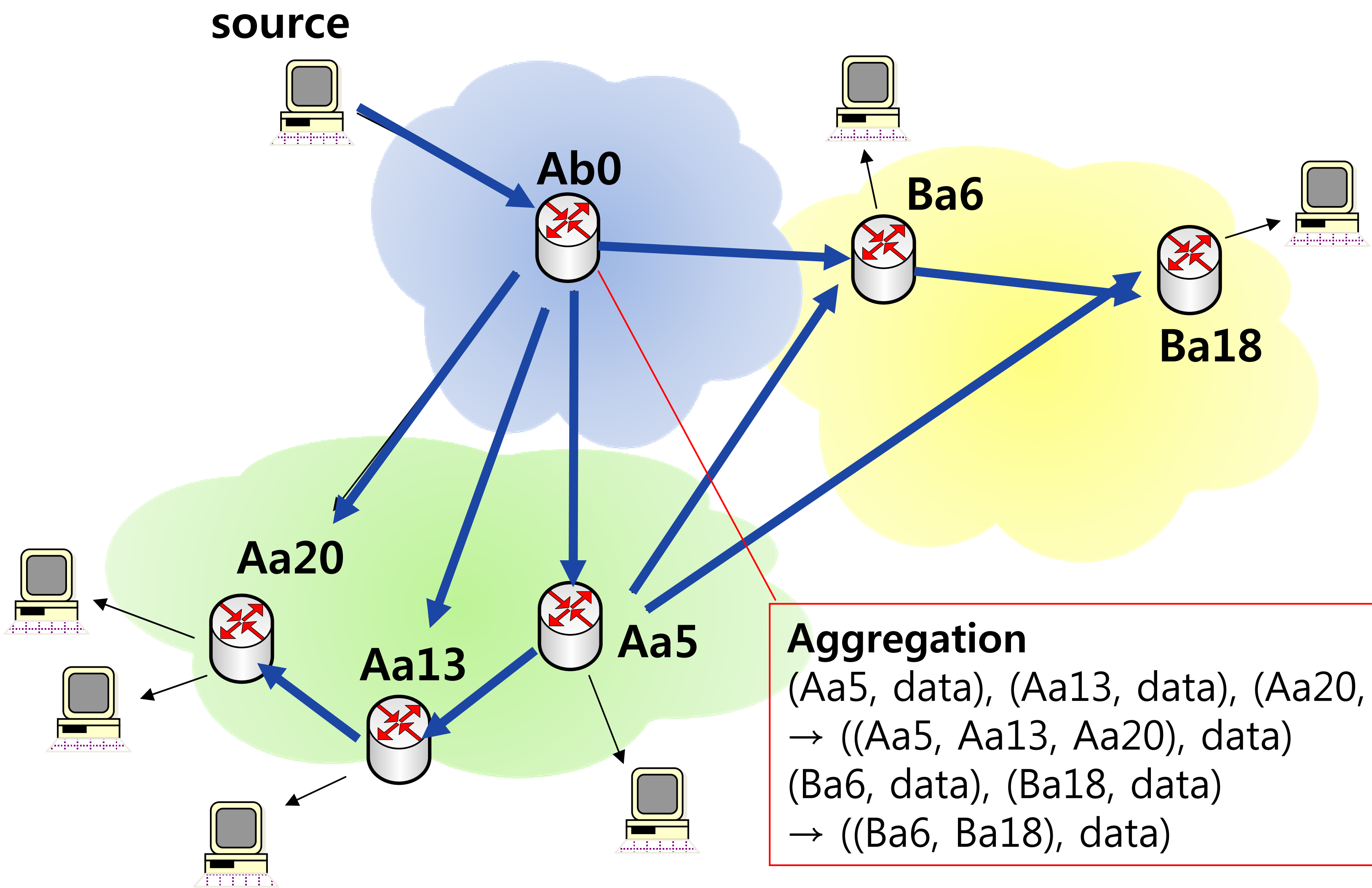
- i3 over Chord
  - Node identifier : i3 server identifier
  - Key identifier : id

- Identifier in Chord : flat label
  - Node identifier : SHA-1(IP address of node)
  - Key identifier : SHA-1(key)
  - It doesn't reflect any physical information. Just logical identifier.

Hierarchical Indirection Multicast

- Identifier = (physical topology) hierarchy part + random part
- Aggregation
  - Instead of individual forwarding to nearby i3 servers, forward to one of them aggregately and the i3 server does to the others
  - "nearby" can be detected hierarchy part of identifier

Example



Aggregation

(Aa5, data), (Aa13, data), (Aa20, data)

→ ((Aa5, Aa13, Aa20), data)

(Ba6, data), (Ba18, data)

→ ((Ba6, Ba18), data)

Identifier in HIM (1/2)

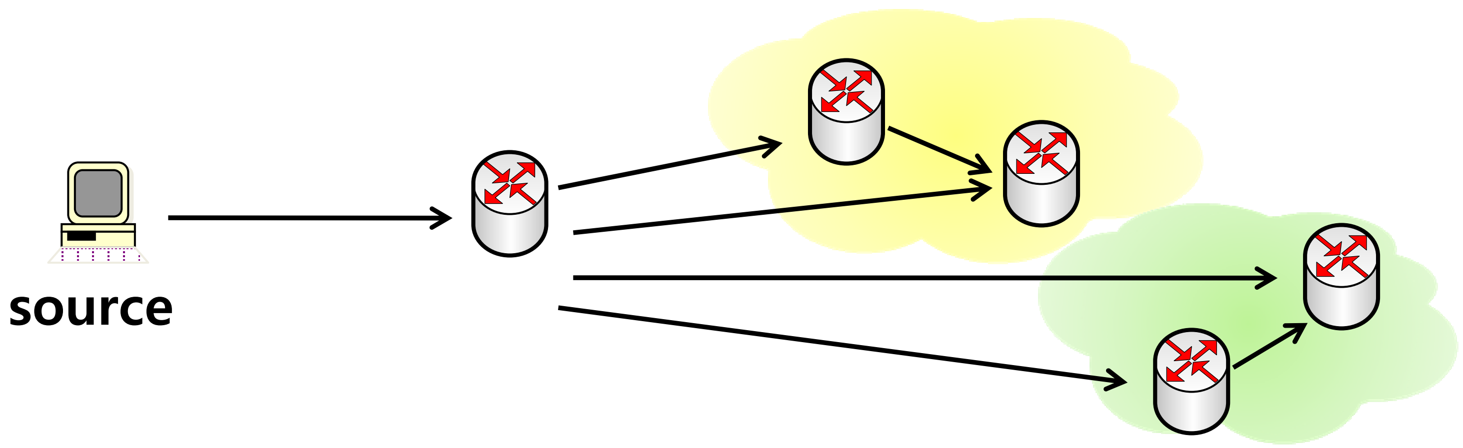
- (Physical topology) Hierarchy part + random part
- i3 servers identifier
  - Random part: generated the same as the original
  - Hierarchy part: assigned by i3 servers physical topology
  - i3 servers in the same domain have same hierarchy part → proximity, aggregation

Identifier in HIM (2/2)

- id (identifier for indirection, trigger)
  - Instead of random sampling
  - Assume the end-host knows a nearby i3 server
  - End-host request id space to the nearby i3 server
    - i3 servers know their responsible id space

Multicast forwarding in HIM

- Instead of individual forwarding to i3 servers on the multicast tree,
- HIM uses aggregation
  - From hierarchical part of i3 identifier, we know whether separate logical link end points (i3 server) are located closely or not



Q & A

Thank you.

sypark@mmlab.snu.ac.kr