# The Role of IP Address in the Internet Architecture

Lixia Zhang
UCLA

**Asia Future Internet Summer School**

**August 2008**

# Disclaimer

- Personal observations and understanding
- Presented for discussion

# Why Talk about IP Address

- A fundamental building block in the original Internet architecture
- In articulating a future Internet architecture:
  - Would it still have IP address as a fundamental building block?
  - If so, what should be the new role of IP address?
  - If not, what is the replacement?

→ How do we answer these questions?

# In the Original Internet Design

An IP address

- identifies an attachment point to Internet
- has the following basic properties:
  - Globally unique
  - Globally routed
  - Globally visible

  → a foundation for end-to-end model
- used in the following functions:
  - E2E datagram delivery to specified destinations
  - borrowed by TCP as part of connection identifier

# Function 1: Datagram Delivery

From "The Design Philosophy of the DARPA Internet Protocols" SIGCOMM'88

- Primary goal: developing an ***effective*** technique for multiplexed utilization of all existing networks
- Second goal: Continued operation despite partial (physical component) failures

# The Internet Started Simple

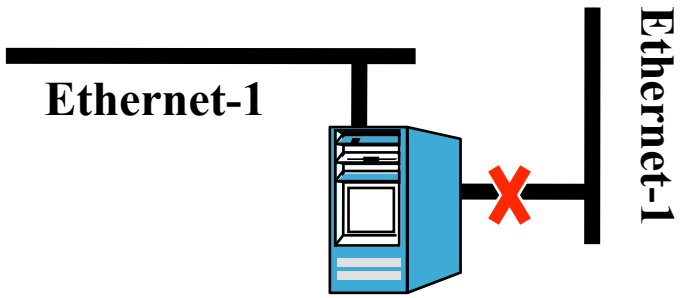
- Different networks connected through gateways
- *All* gateways trying to find best paths to forward *all* packets
- Datagram routing: which way to forward each packet towards its *destination address*?
  - Routing entry granularity: ***network***
- There was no "Internet service provider" at the time

→All networks were equal

→All addresses *provider-independent*

# IP address *is* topological dependent

- From RFC791 IP Specification: “provision must be made for a host to have several physical interfaces to the network with each having several logical internet addresses.”
- A 2-way multihomed host may have one interface failed but still reachable through another
- but a TCP connection using the IP address of the failed interface will fail!



# Why TCP borrowed IP address as part of its connection identifier

- Each TCP connection wanted a globally unique connection identifier
  - To assure each packet being delivered to the right connection
- IP address is globally unique
  - any identifier derived from it is also globally unique
- *It's an engineering design decision*

# Consider the alternative:

- Had the TCP design required a host ID from a separate identifier space, this topology-independent host ID would allow a TCP connection to persist over IP address changes
- But the benefit would show up only if
  - the host is multihomed
  - A failure occurred during a TCP connections life time, or
  - the host changes IP address during a TCP connection's life time

# Is the benefit worth the cost?

- Need to answer this question in the context of 30 years ago
- Unclear benefit?
  - At the time: single-homed hosts dominate
  - No host mobility?
  - Perhaps connections were short-lived?
- Clear costs:
  - Managing another identifier space
  - Requiring a mapping system to match a host ID to the corresponding IP address

# Weighing the benefit, saving, simplicity

- Using IP address as connection identifier is the *simplest* design to reach the entity the identifier identifies
  - With little loss of benefit (at the time)
  - A SSN is a unique identifier, but does not say anything about where to find the person
- In addition: making it difficult to hijack a TCP connection
  - An IP address cannot be easily hijacked as long as the routing system is not compromised
    - This fact has been used for security enhancement, e.g. TCP SYN cookie

# Engineering design versus "correctness"

- Protocol design is engineering
- When a host is connected through a single interface, IP address semantic overload worked out quite well
- This semantic overloading represents a good engineering design tradeoff under the given condition
  - If/when the conditions change, the conclusion is likely to change as well

# What have changed since 1981?

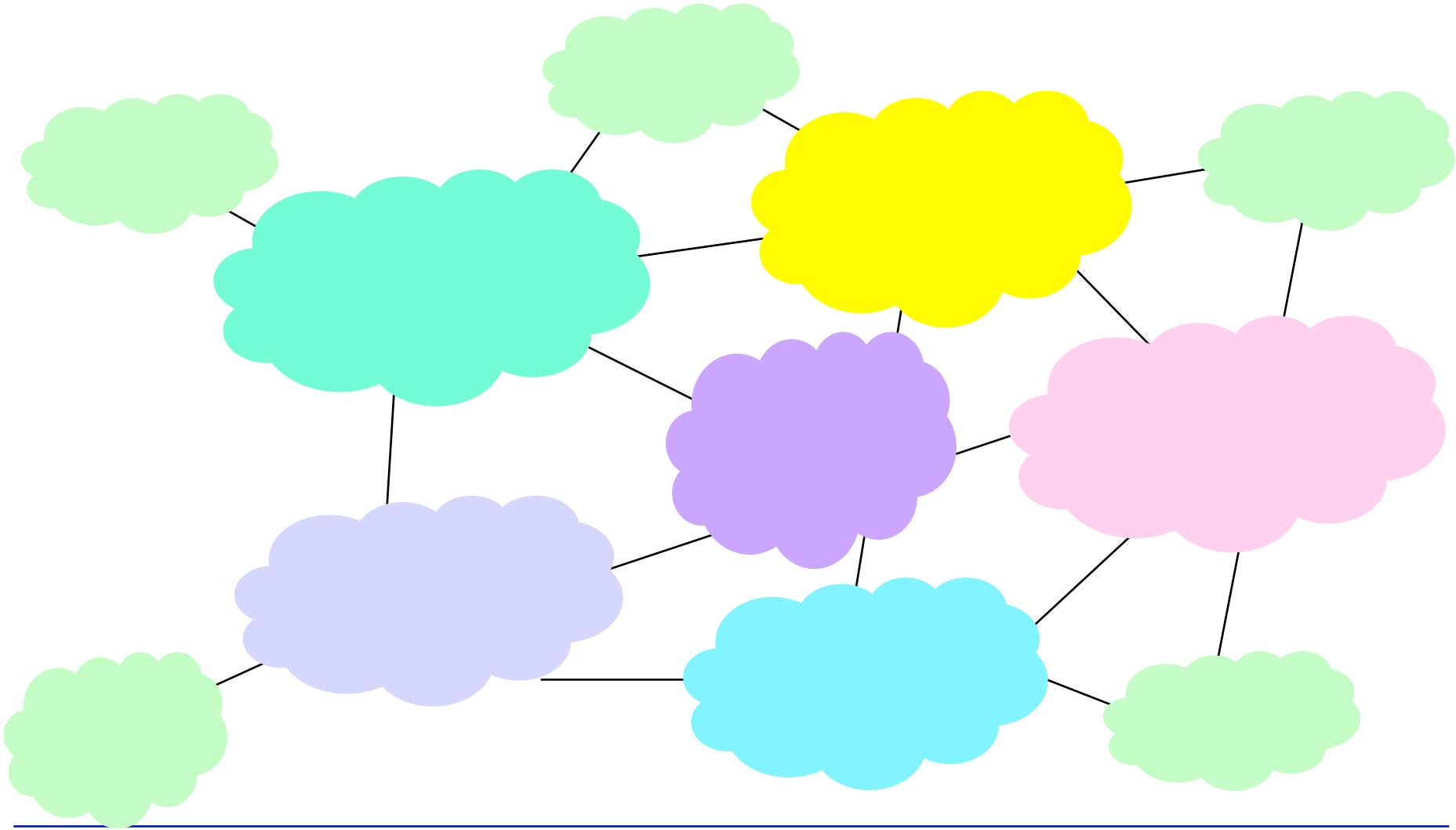
- First and foremost: Internet has grown by orders of magnitude!
  - Beyond the wildest dreams of the original designers
  - NAT deployment became pervasive

    "*A Retrospective View of Network Address Translation*"

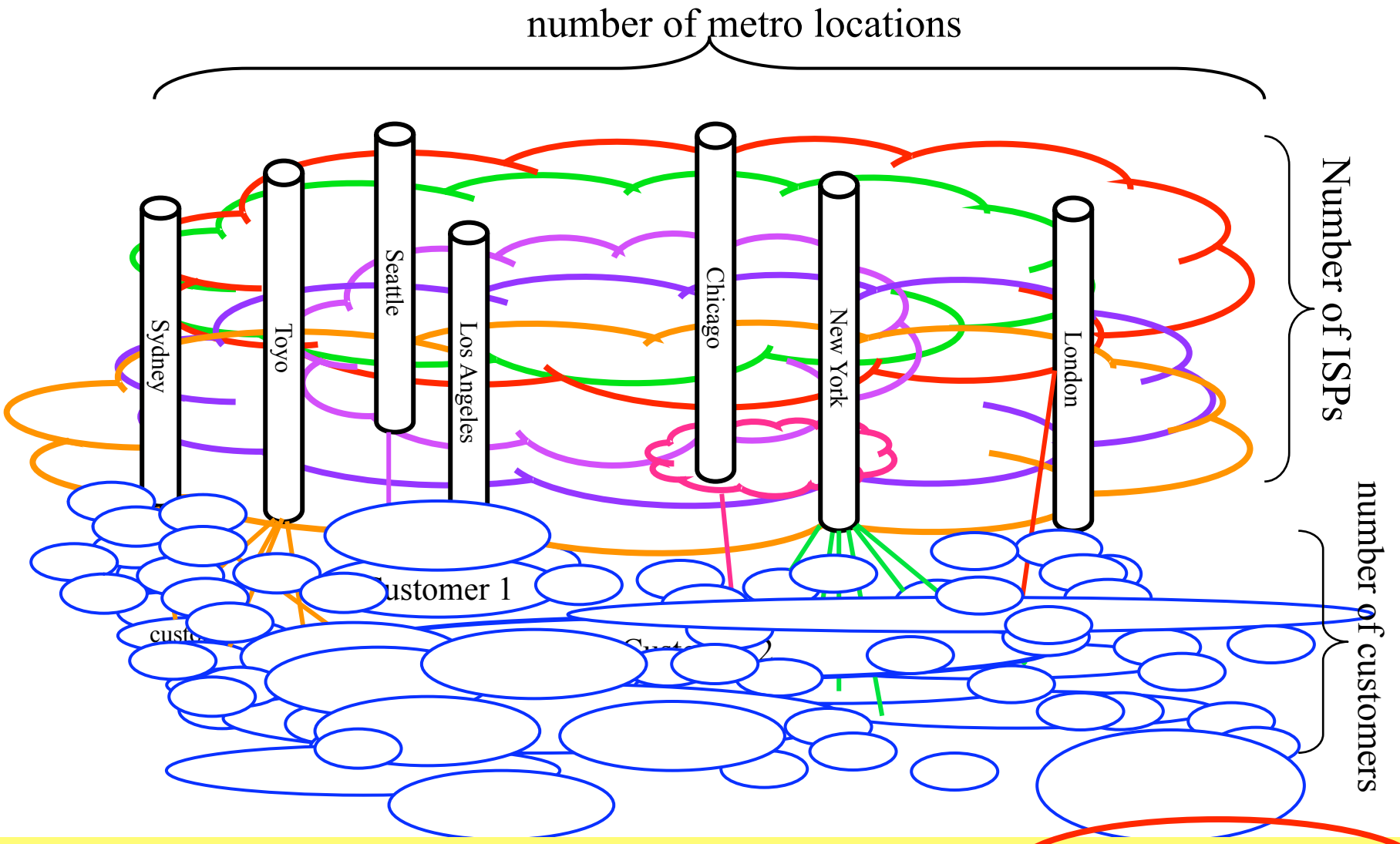
    IEEE Network September 2008
- Site multihoming
- Host multihoming
- Mobility
- ***Ever increasing security threats***

# How Networks Look like Today

When we draw network graphs, it tends to look like this



# But in reality, it is more like this



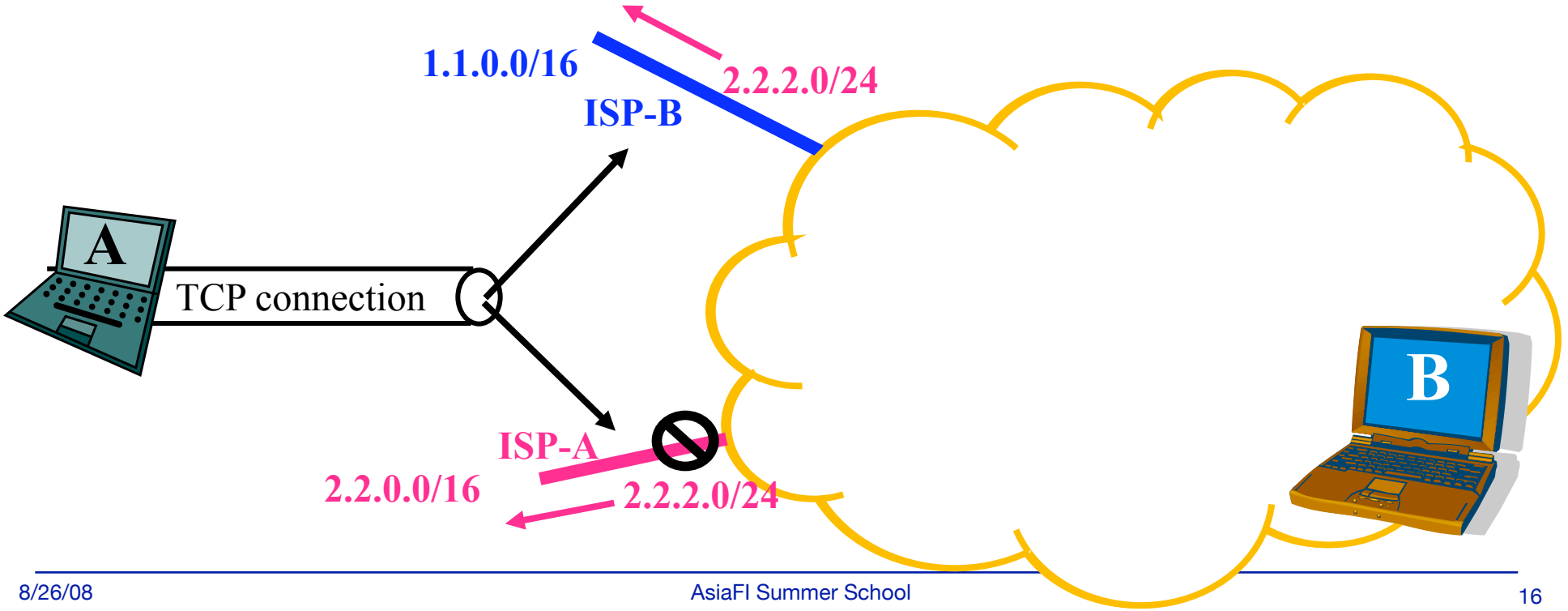

**DFZ Routing table size = Function(# of ISPs X # of PoPs X # of user sites X TE)**

# We now have ISPs→2 new things happened

- **Provider-Assigned address (PA)**
- User site multihoming

*ISPs are not happy!*

*Users are not happy!*

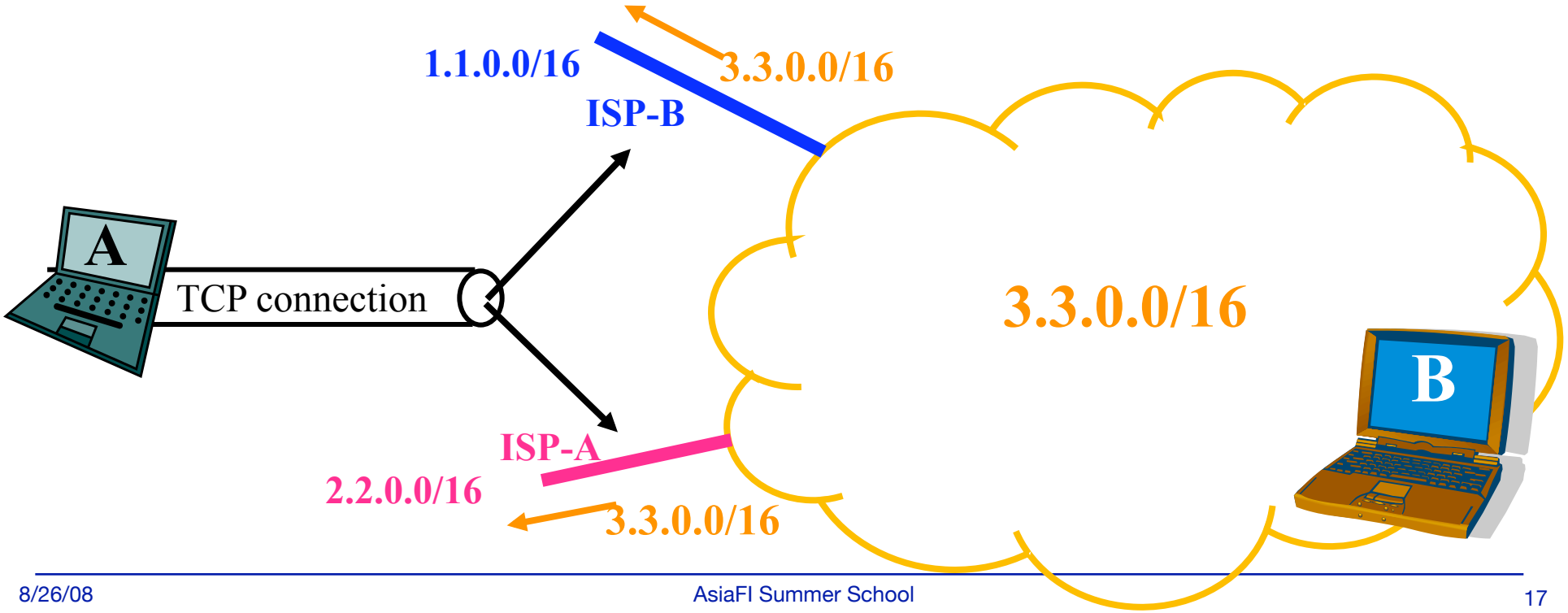


# Provider-Independent Addressing

- User site multihoming

*Users would be happy!*

*ISPs would not be happy!*



# Tensions between user sites and providers

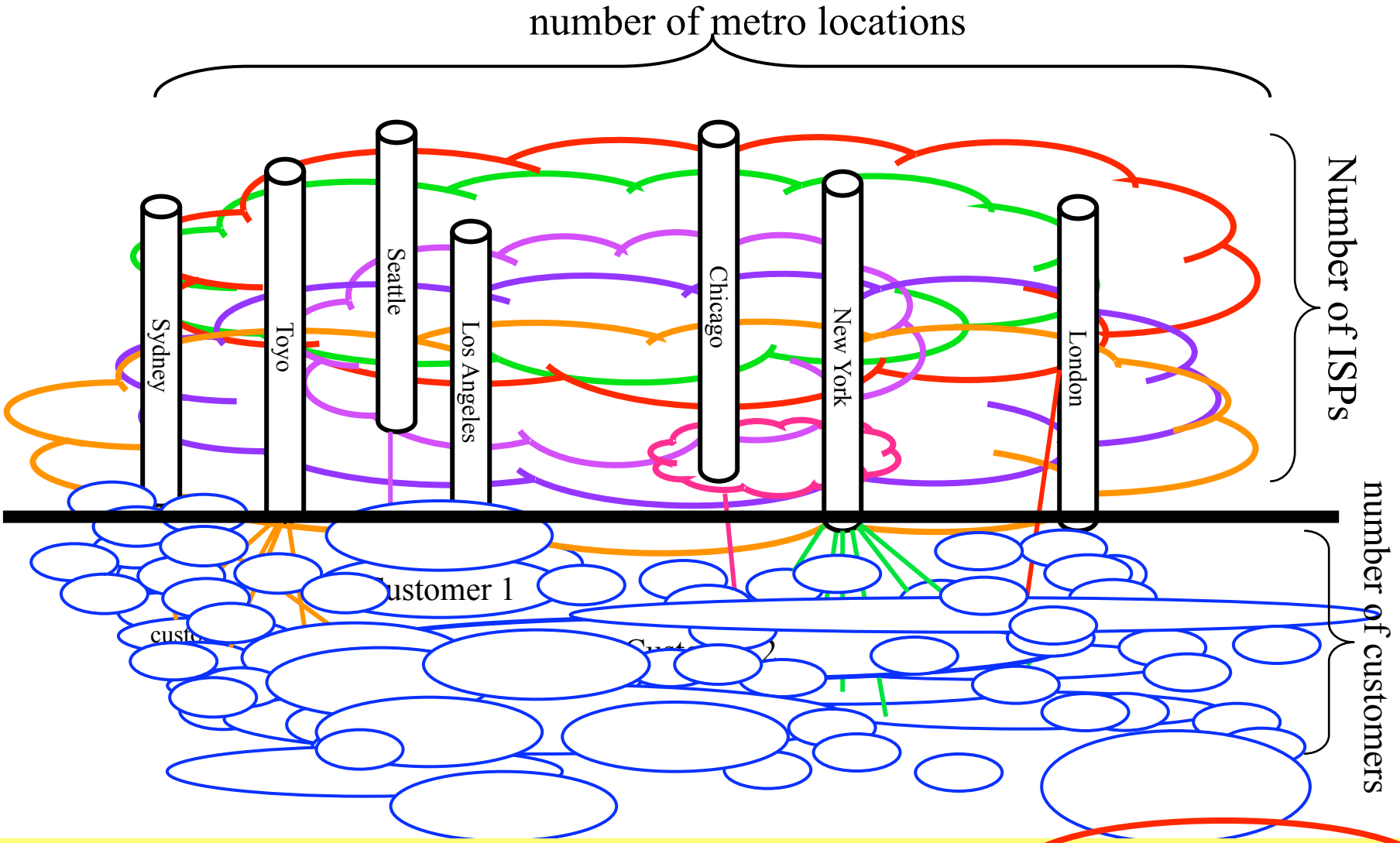
- Providers want provider-based addressing, which can be aggregated to scale the routing system
- User sites want Provider Independent (PI) address
  - Most user sites are multihomed today
  - no one desires renumbering

⇒ Head-on conflict

⇒ Whoever paying wins

The result: ever increasing global routing table size

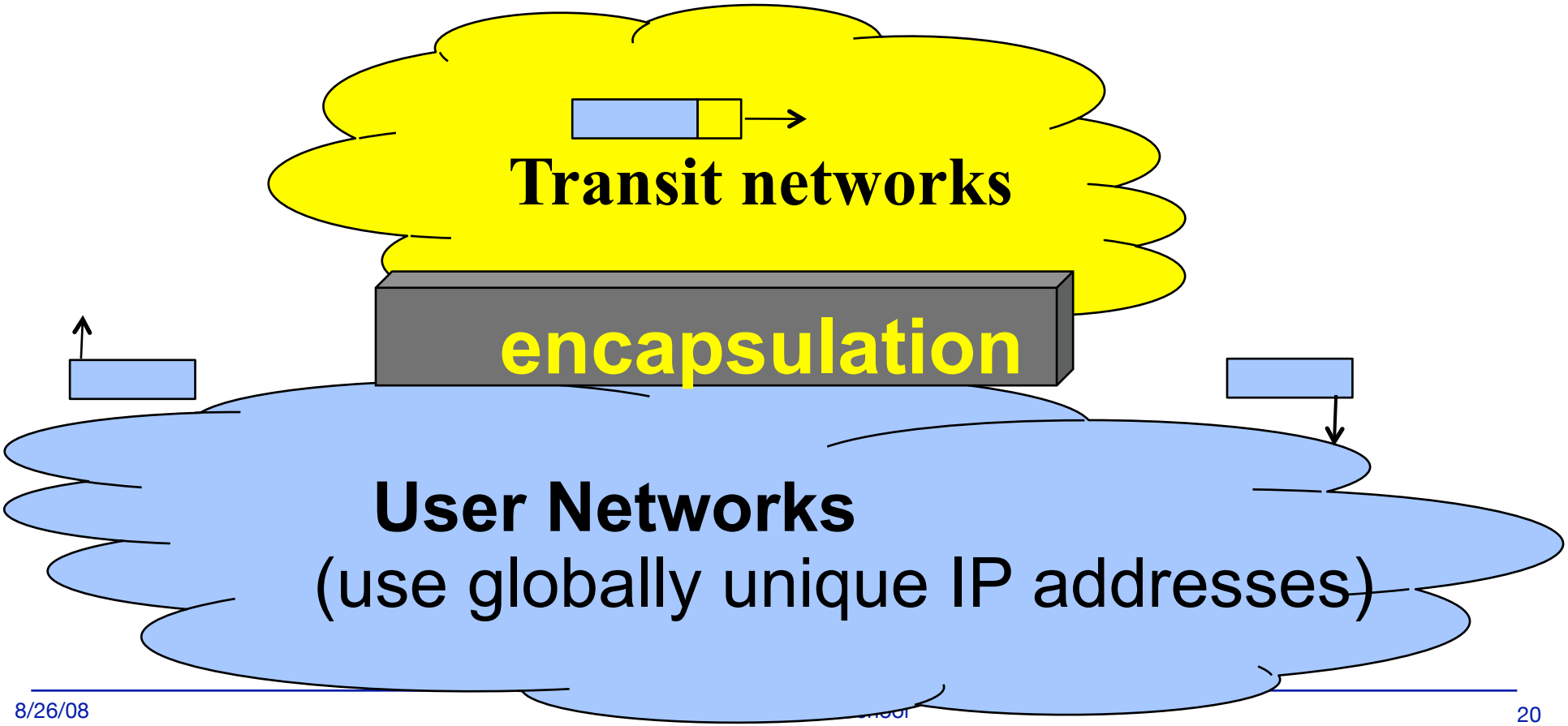
# Proposed solution:
# Removing PI prefixes from global routing system




**DFZ Routing table size = Function(# of ISPs X # of PoPs X # of user sites X TE)**

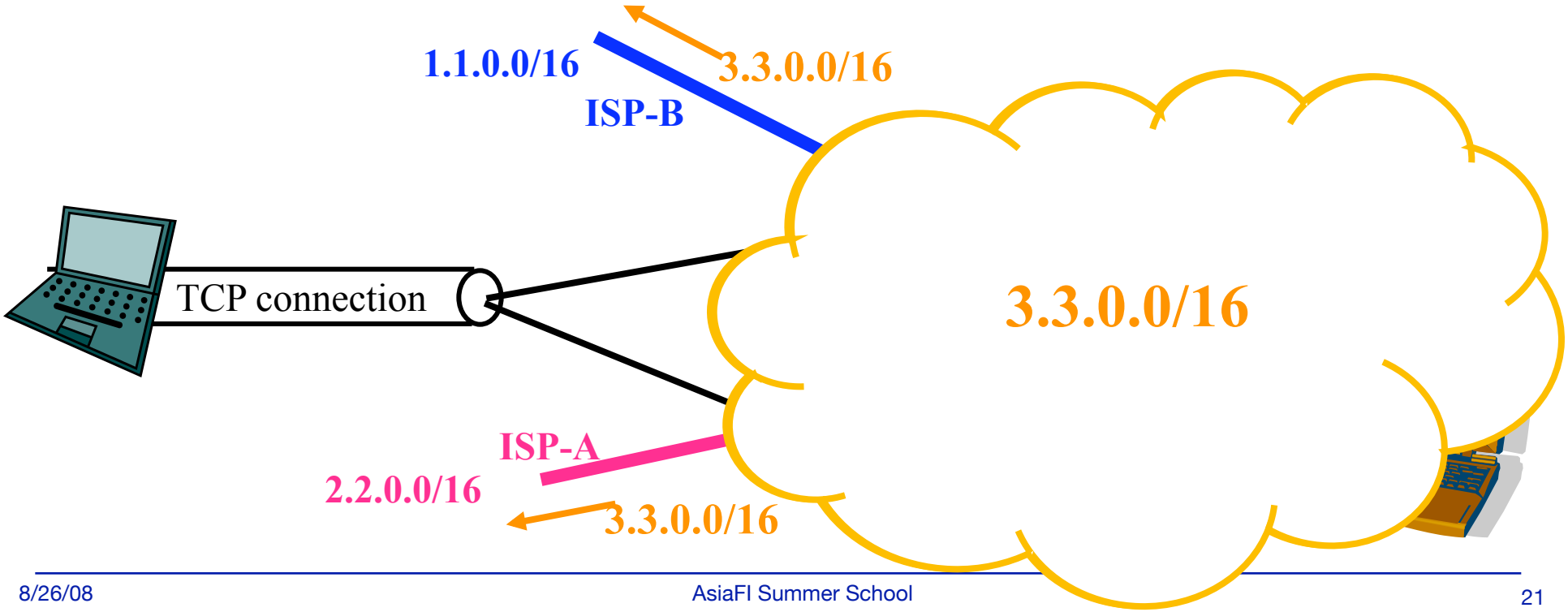
# One class of solution: Map-n-Encap

- First proposed in RFC1955
- Changing the *scope of IP address routability*
  - See more details in tomorrow's talk



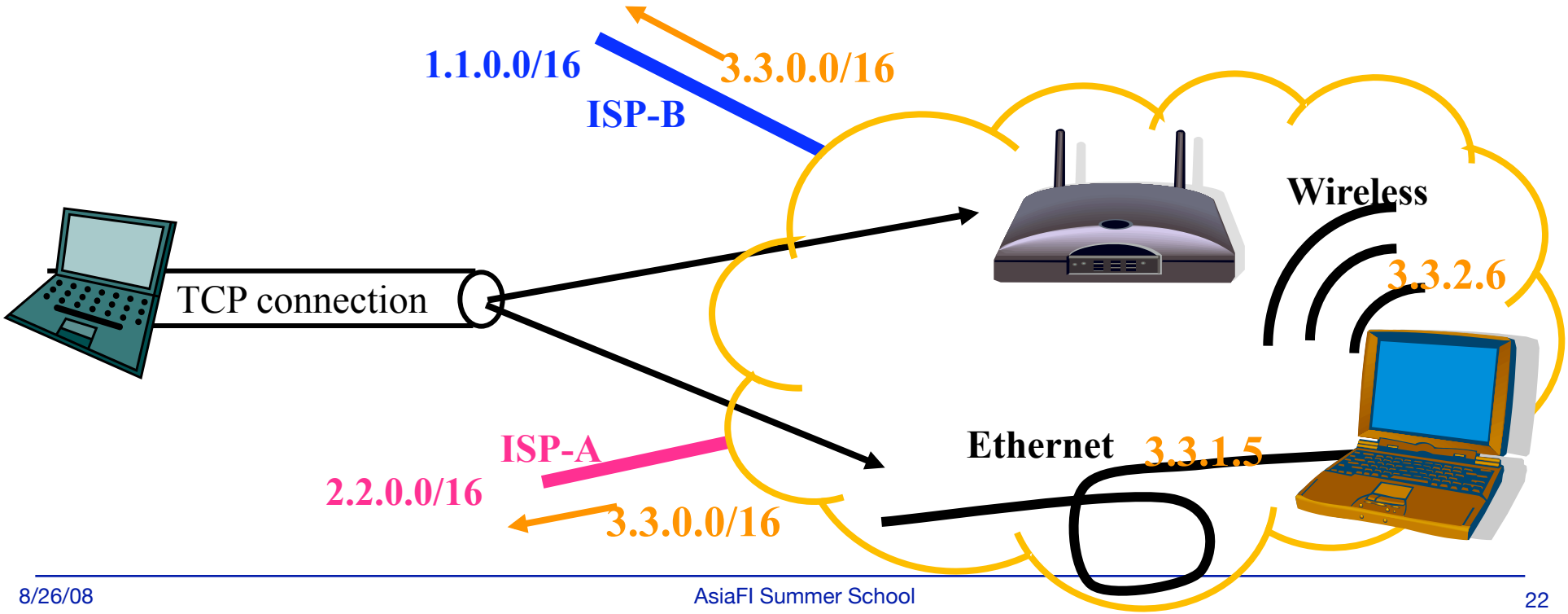
# Host Multihoming

- TCP uses IP address as part of connection identifier
- IP address identifies one attachment point!



# Host Multihoming

- TCP uses IP address as part of conn. identifier
- IP address identifies attachment point!



# Do we need a host identifier now?

- Technology advances → multihomed hosts dominant
  - Desktop, laptop, pulm top
- The condition 30 years ago (single-homed hosts) changed forever
- If one wants to identify a host independent from its connectivity → need a host identifier

***Why now? Why not from day one?***

# Addresses, Identifiers, locators: Exactly what are we separating from what?

- Providers: want topologically aggregatable address prefixes
- Sites: want provider-independent address blocks
- TCP (high level protocols in general): want IP address-independent end-point identifiers

To scale DFZ routing: separate these two

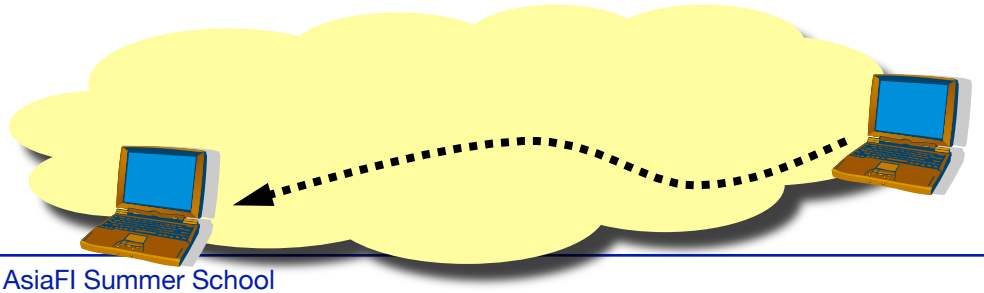
To make TCP conn. survive change of delivery path: ***separate*** IP-addr and end identifiers

# Technology advances → Arising of Mobility

- Mobility of individual hosts
- Mobility of all the nodes in a network (Ad hoc networks)

# Supporting host mobility

- Goal: delivering packets to the right IP *interface* in the *global* Internet
- IP address: defines attachment point
- Moving from one place to another ⇒ change of IP addresses
- The *fundamental* design question: *who*/*where* to keep the state (=new address) of a moving host?



# Mobile IP design

- Who: individual mobile hosts to choose
- Where:
  - Within IP layer
  - Outside network routing infrastructure
- How: let the moving host report back to its chosen home agent
- *Simplest* fix to support host mobility
  - In general "*simplest*" is unlikely to yield "*optimal*" performance

# Is mobile IP design a patch-on?

- Yes it *was* added on later
- If we were to start from scratch, would it have been done differently?

# Many alternative designs possible

- The network routing infrastructure could take over the responsibility of keeping tracking mobile hosts
- The address change could be directly reported to a name lookup service
  - Keeping state outside (above) IP layer
- And a number of others

Q: how do we judge which one is better?

# What should be our yardstick for measure?

- Scalability as #1 objective
  - We'll see increasing number of mobile devices
- Delegation of responsibility
- Keep it simple; Must be prepared for things to go wrong
- Performance is important, but below any of the above
  - Performance is always second to reachability

# Does Mobile IP Design Measure UP?

- Keeping mobile state at “home agent”

→ No impact on routing scalability

→ Keep the matter in your own hand

  - One implements/chooses his own home agent right!
  - X's mistake has no impact on Y
  - Pre-settlement for relation/accounting/security

Admittedly,

- Not giving *highest possible* performance
  - Especially in case of a single home agent
- Not very efficient
  - Especially when facing rapid host movement
  - Additional engineering improvement possible

# Is Ad hoc networking a different beast?

**topology**

| | |
|---|---|
| Fixed | Topology *does* change<br>• semi-static structure<br>• link/node failures<br>• routing: Baran's hot-potato flooding ⇒ separate routing protocols for scalability |
| changing | Structure-free ⇒ host routing<br>Resource constrained ⇒ On-demand routing<br>To handle high dynamics ⇒ flooding<br>To scale better ⇒ Cluster/landmark routing<br>Moving towards structured routing |

# IP Address Today

- IP address is still used for data delivery
  - topological dependency unchanged
- Pervasive IPv4 NAT deployment led to a large number of hosts using addresses that are
  - → No longer globally unique (locally unique), nor globally routed (locally routed)
- Plethora solutions to mobility support
- What have changed/may change:
  - The scope of uniqueness
  - The scope of routability
  - The need for indirection

# The changing nature of IP address

- Wide existence of private IP addresses (RFC1918)
  - with *scoped uniqueness*
  - Private: non-visible outside the local scope
- The usefulness (or lack of it) of IP addresses with *scoped visibility*
  - Do addresses of scoped visibility have value (for some purpose)?
  - If so, should they be globally unique?
- IP addresses with *scoped routability*

In addition: the need for connectivity-independent node identifier

- and how many different name spaces may be necessary?

# IP Address and Internet Security

- On day one: it's given that each packet carried correct source IP address
- Today: source address spoofing as one of the malicious attacking weapons
- One needs explicit effort to *enforce* correct source address
- It is important to do so
  - Measuring network traffic: monitoring
  - Identifying problems: diagnosis
  - Identifying attackers: mitigation

# Summarizing

- IP address remains a fundamental building block in the architecture
- IP address is used for packet delivery, as such they are topology-dependent to make routing scale
  - Mobility being handled outside the routing system
- Multihoming occurring with multiple granularity, leading to necessary changes to the original use of IP address
- Understand scoped uniqueness, visibility, routability: their roles and implications on the overall Internet architecture

# Look into future

- The fundamental value of IPv6: restore IP address' *global uniqueness*
- Global visibility: Different views on whether allowing private IPv6 address
  - If allow: should it have guaranteed globally uniqueness?
- Global routability: May not stay, to make routing scalable
  - Separating uniqueness from routability

**Thank you!**

# Questions?

*lixia@cs.ucla.edu*