# Towards a global IP Anycast service

# IP Anycast : Affinity and Proximity Measurements

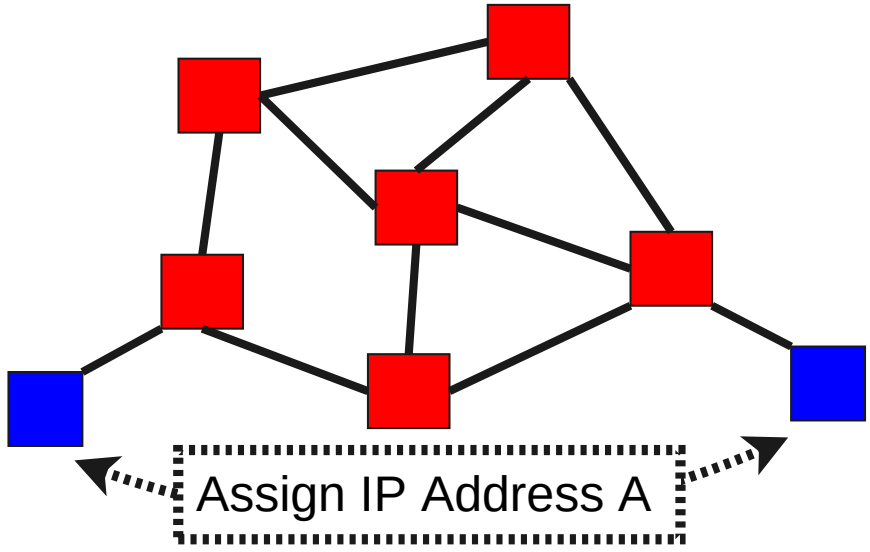
Hitesh Ballani, Paul Francis

Cornell University

DNS-OARC Workshop

# What is IP Anycast?

A paradigm for communicating with any member of a group

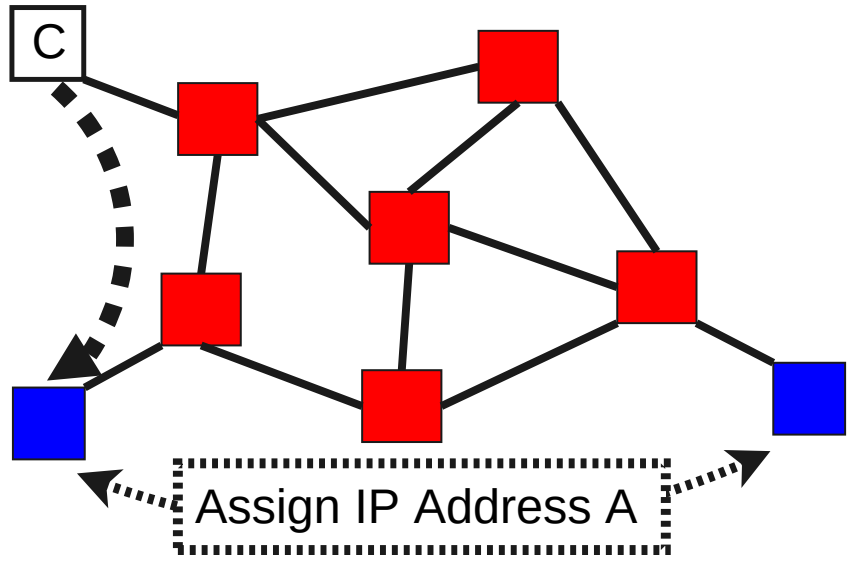



## Robust and efficient service discovery

- Query-Reply Services : DNS Root-Servers etc.
- Routing Services : 6to4

# What is IP Anycast?

A paradigm for communicating with any member of a group

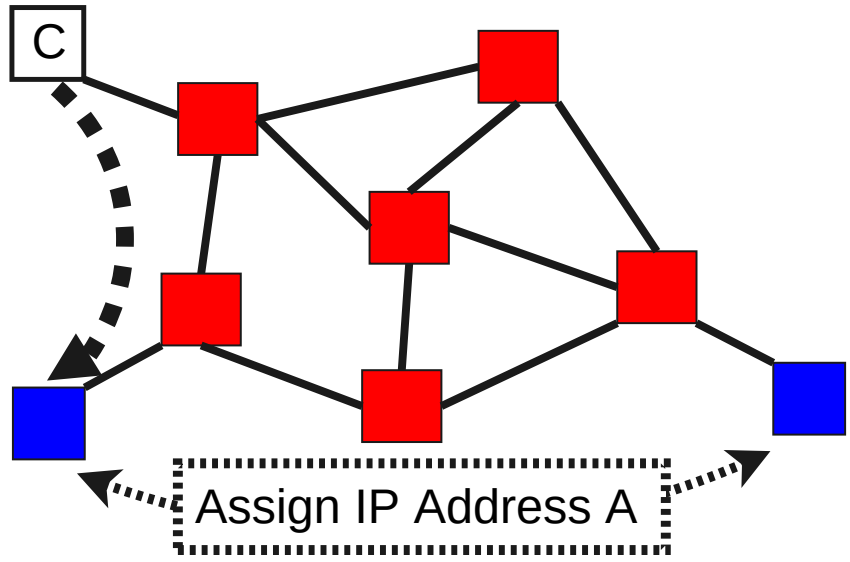



## Robust and efficient service discovery

- Query-Reply Services : DNS Root-Servers etc.
- Routing Services : 6to4

# What is IP Anycast?

A paradigm for communicating with any member of a group




## Robust and efficient service discovery

- Query-Reply Services : DNS Root-Servers etc.
- Routing Services : 6to4

## But its use is limited?

# Limitations of IP Anycast

Incredibly wasteful of address space

Scales poorly by the number of anycast groups

Difficult to deploy

- obtain an address prefix and an AS number
- a certain level of technical expertise

Subject to the limitations of IP routing

# Limitations of IP Anycast

Incredibly wasteful of address space

Scales poorly by the number of anycast groups

Difficult to deploy

- obtain an address prefix and an AS number
- a certain level of technical expertise

Subject to the limitations of IP routing

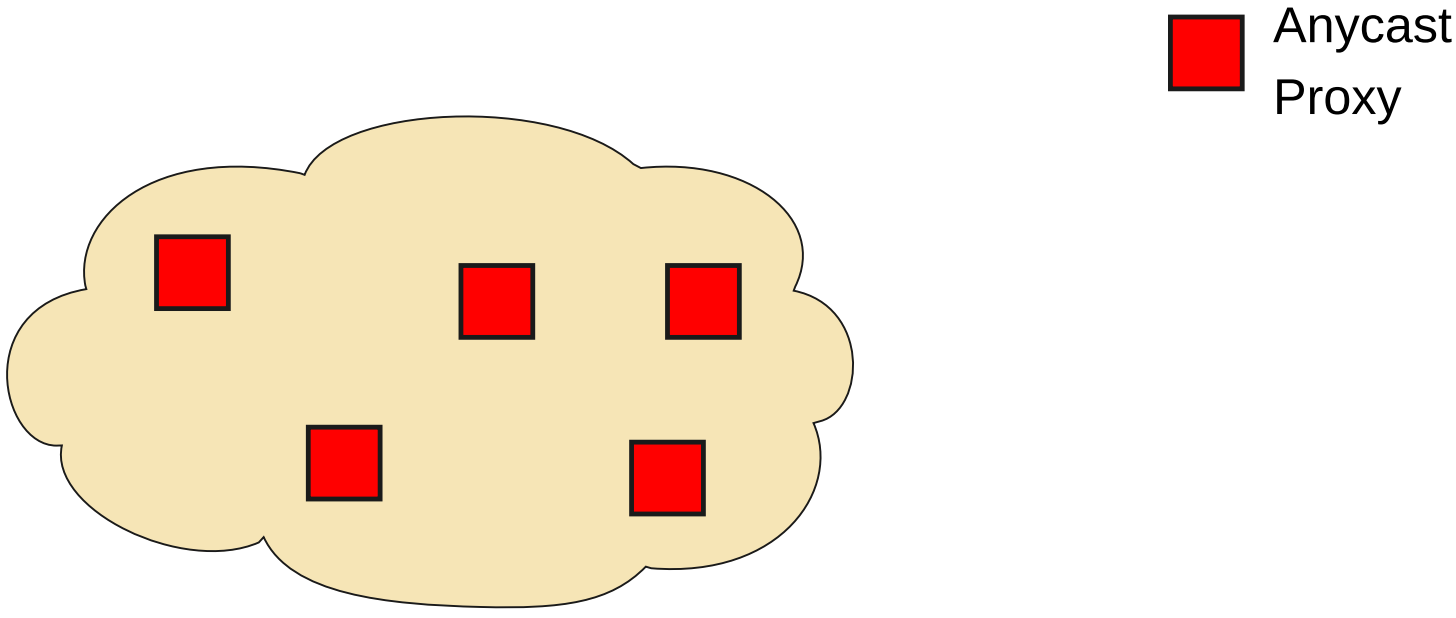
Application-layer anycast

- DNS-based load balancing
- used in current applications of anycast

# Proxy IP Anycast Service (PIAS)

## Deploy Anycast Proxies

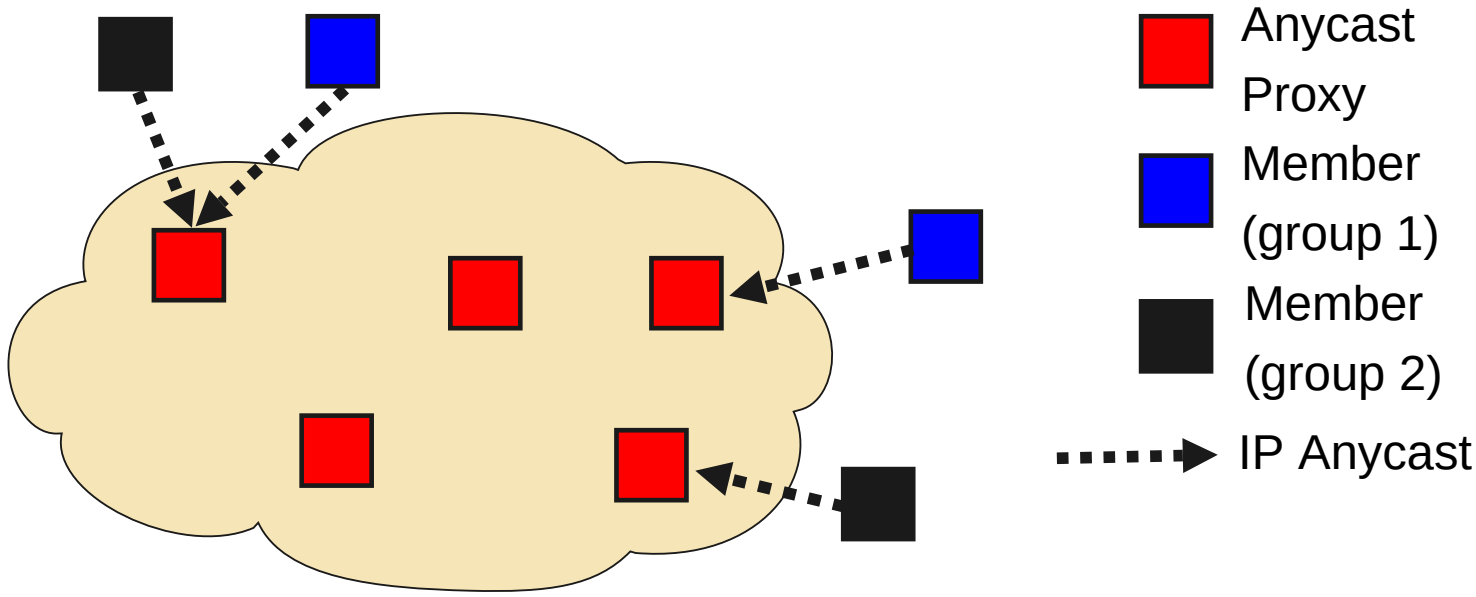
- Group members register with the proxies
- Native IP Anycast delivers packets to proxies
- Proxies forward them to appropriate member



# Proxy IP Anycast Service (PIAS)

## Deploy Anycast Proxies

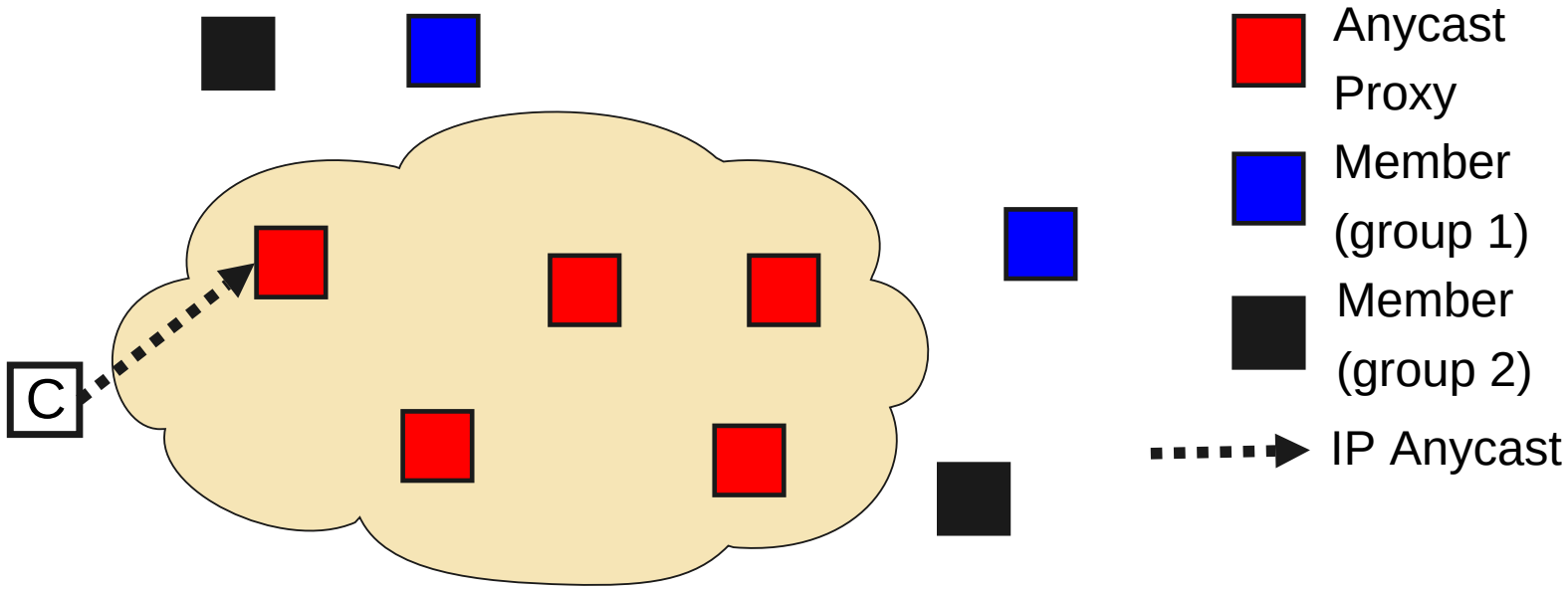
- Group members register with the proxies
- Native IP Anycast delivers packets to proxies
- Proxies forward them to appropriate member



# Proxy IP Anycast Service (PIAS)

## Deploy Anycast Proxies

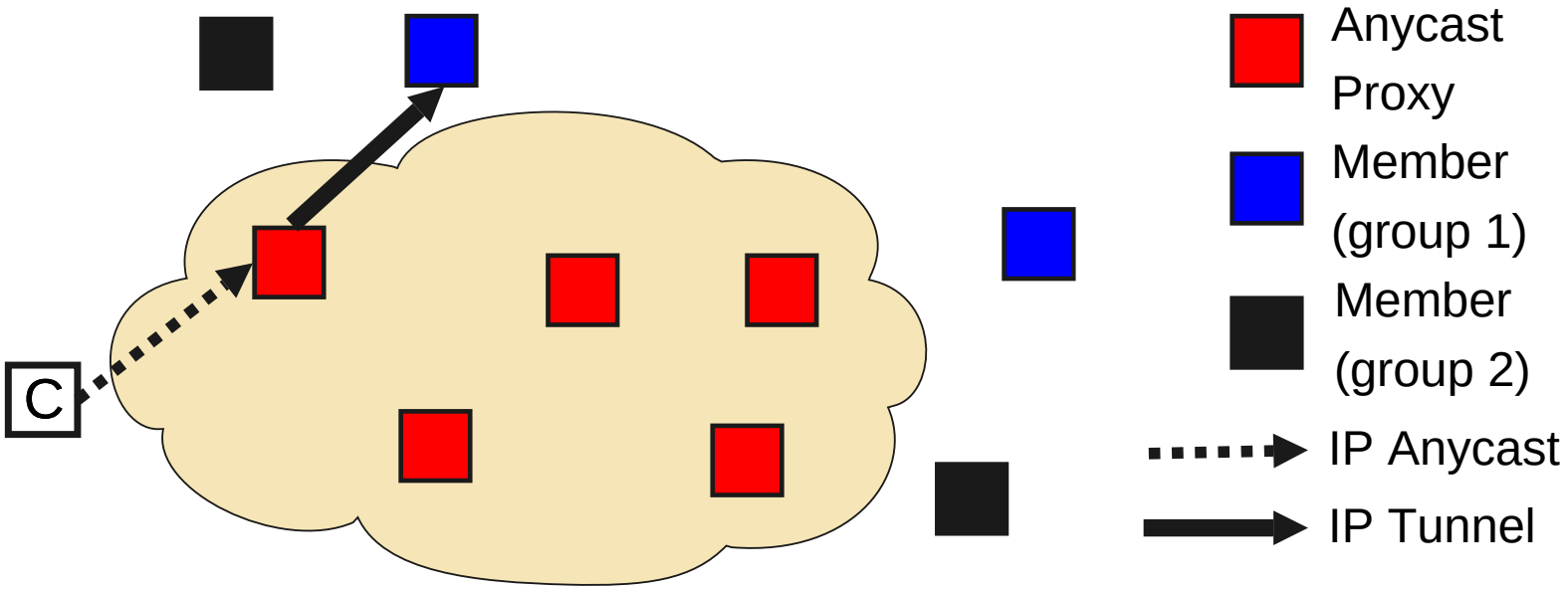
- Group members register with the proxies
- Native IP Anycast delivers packets to proxies
- Proxies forward them to appropriate member



# Proxy IP Anycast Service (PIAS)

## Deploy Anycast Proxies

- Group members register with the proxies
- Native IP Anycast delivers packets to proxies
- Proxies forward them to appropriate member

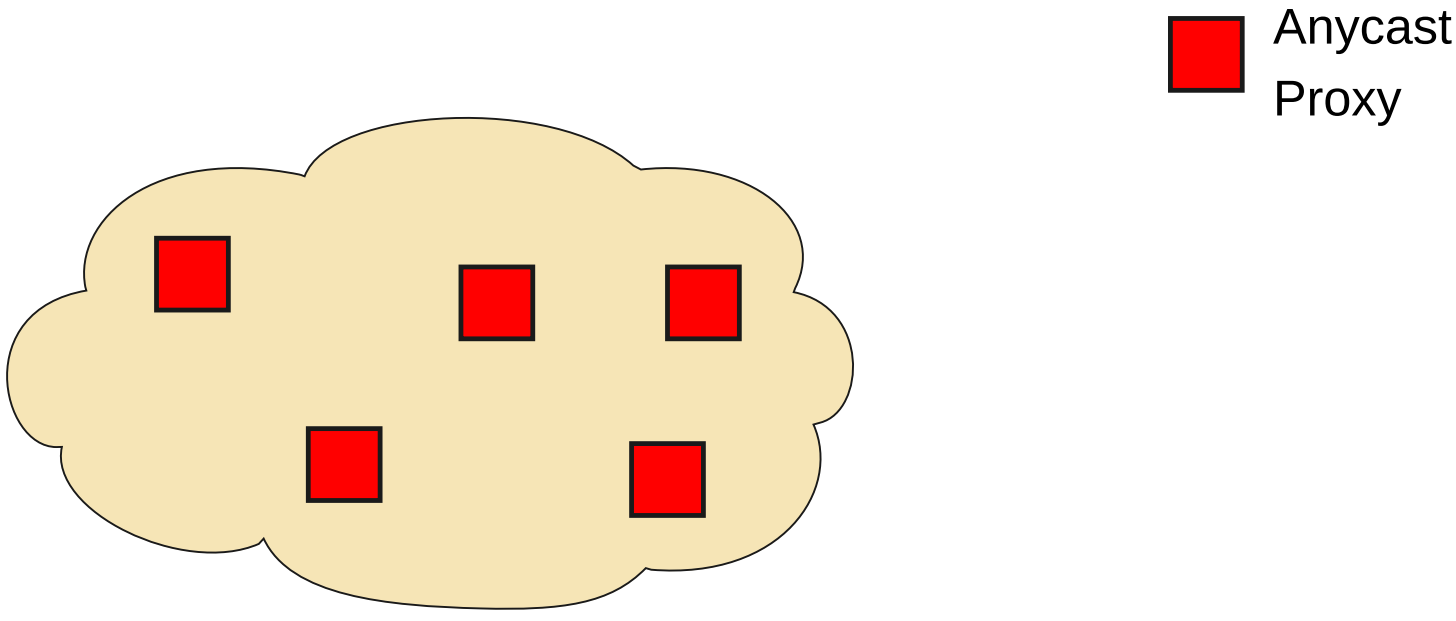


# What does PIAS solve?

- Practical deployment model
- Efficient use of address space
- No changes to clients
- Smart selection by the proxies

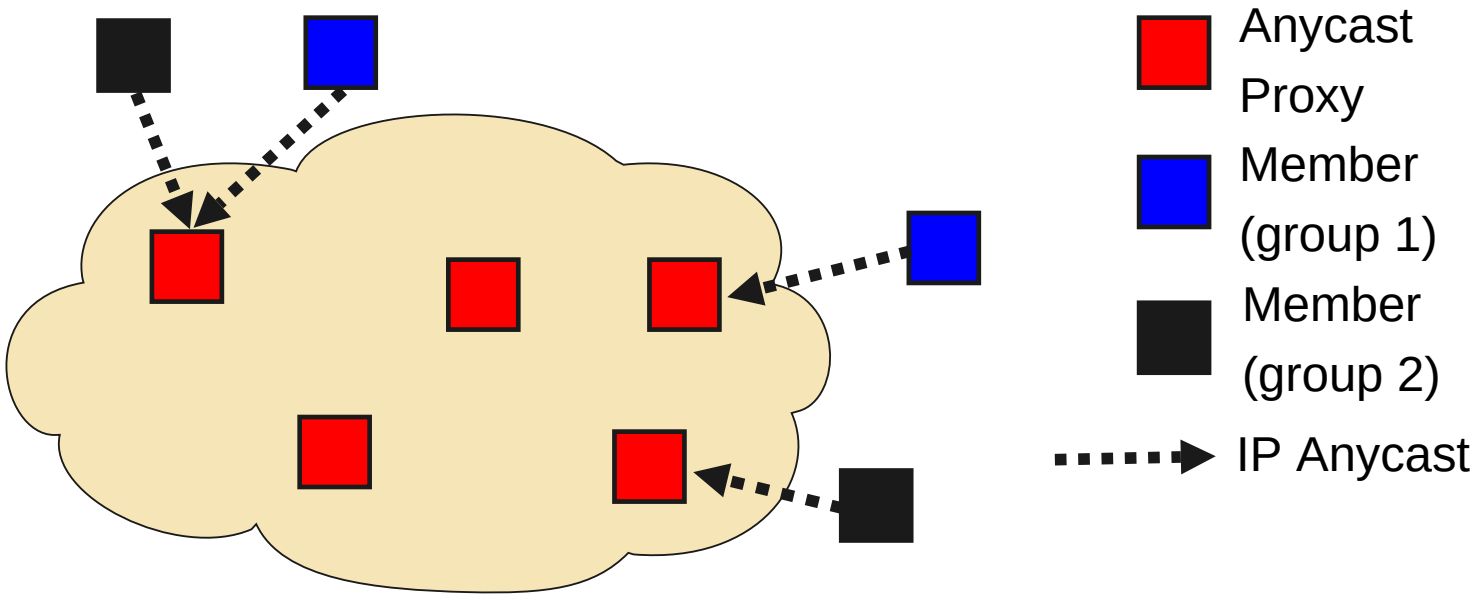
# What does PIAS solve?

- Practical deployment model
- Efficient use of address space
- No changes to clients
- Smart selection by the proxies



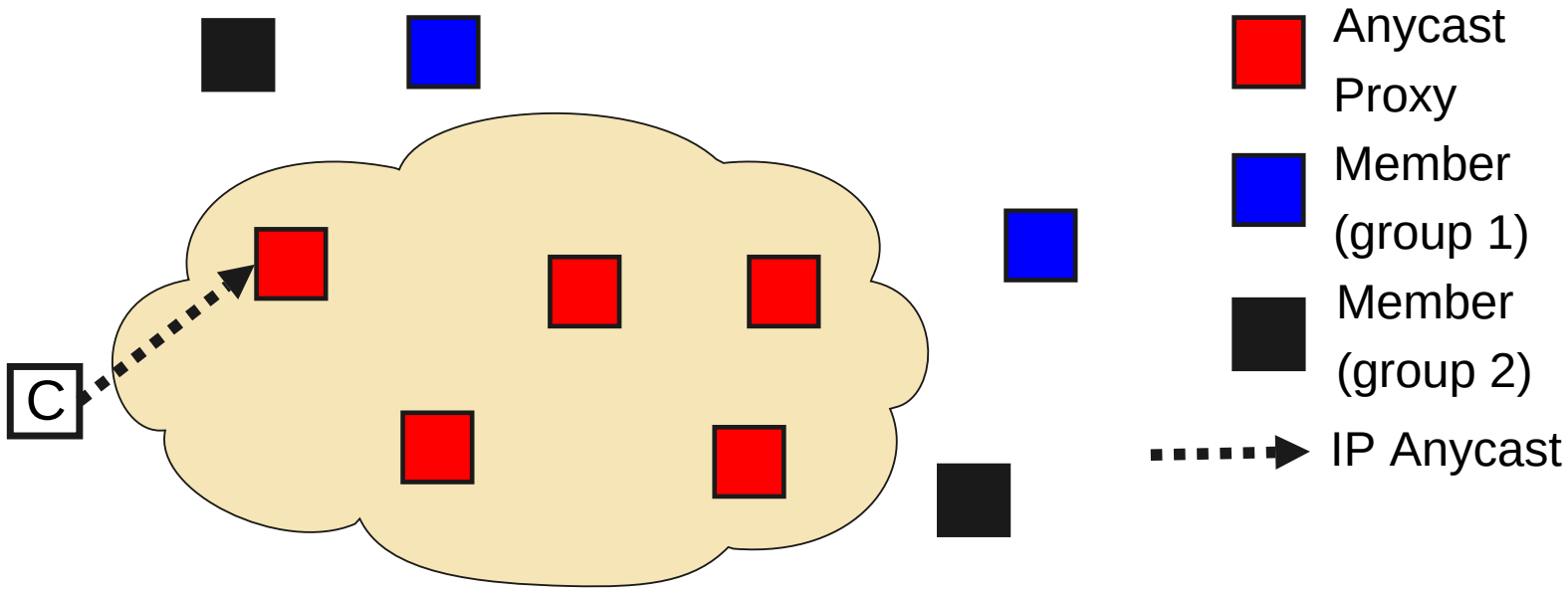
# What does PIAS solve?

- Practical deployment model
- Efficient use of address space
- No changes to clients
- Smart selection by the proxies



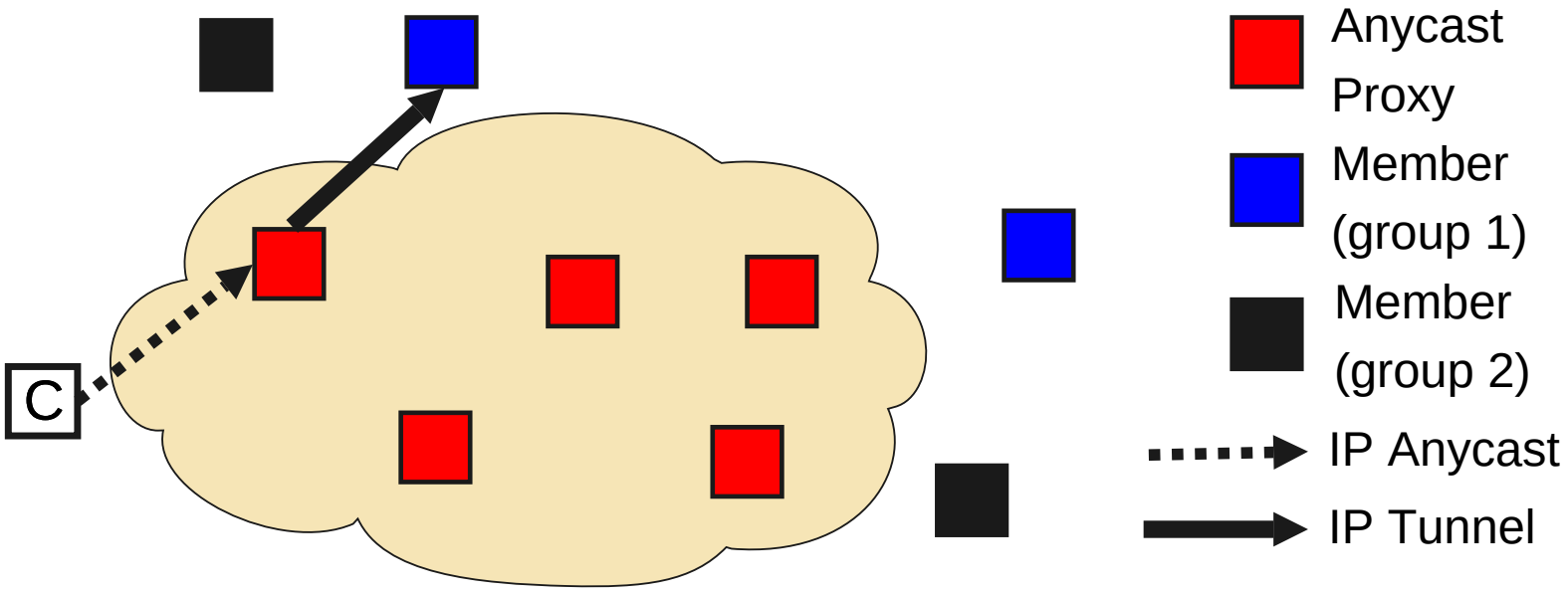
# What does PIAS solve?

- Practical deployment model
- Efficient use of address space
- No changes to clients
- Smart selection by the proxies

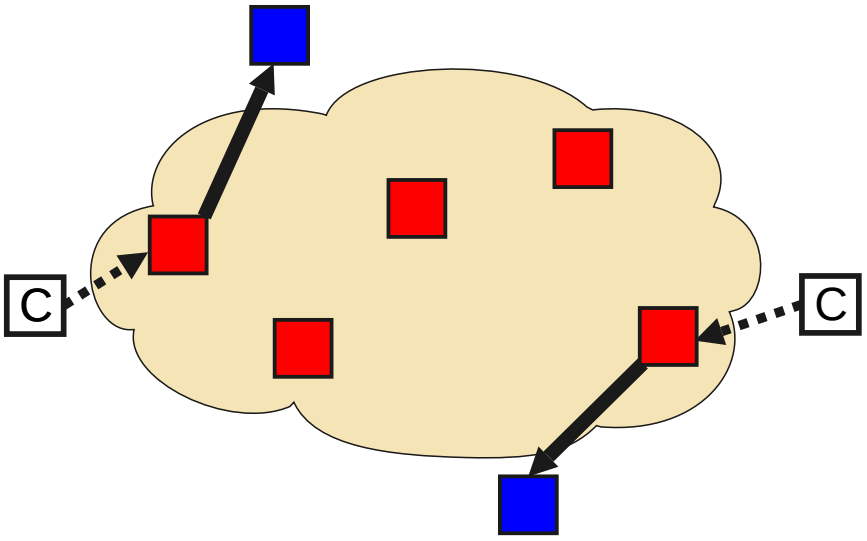


# What does PIAS solve?

- Practical deployment model
- Efficient use of address space
- No changes to clients
- Smart selection by the proxies

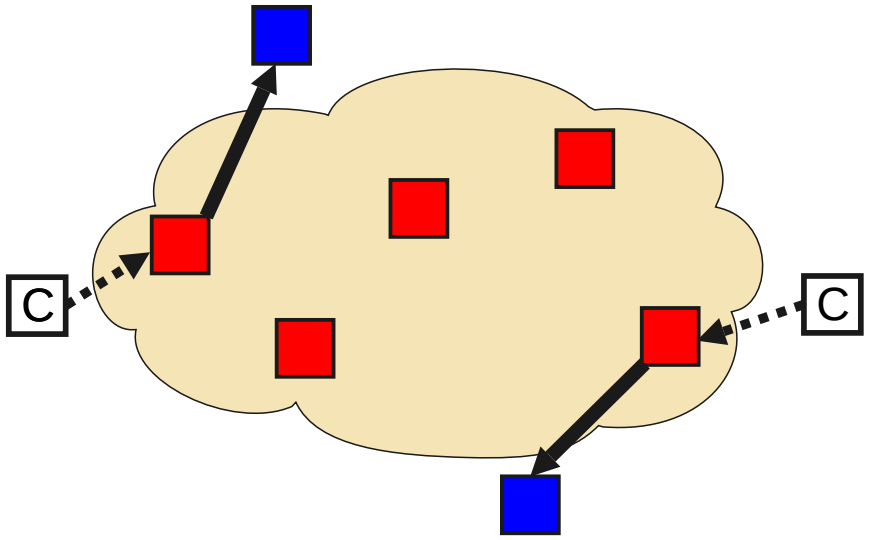


# Wouldn't it be nice if . . .

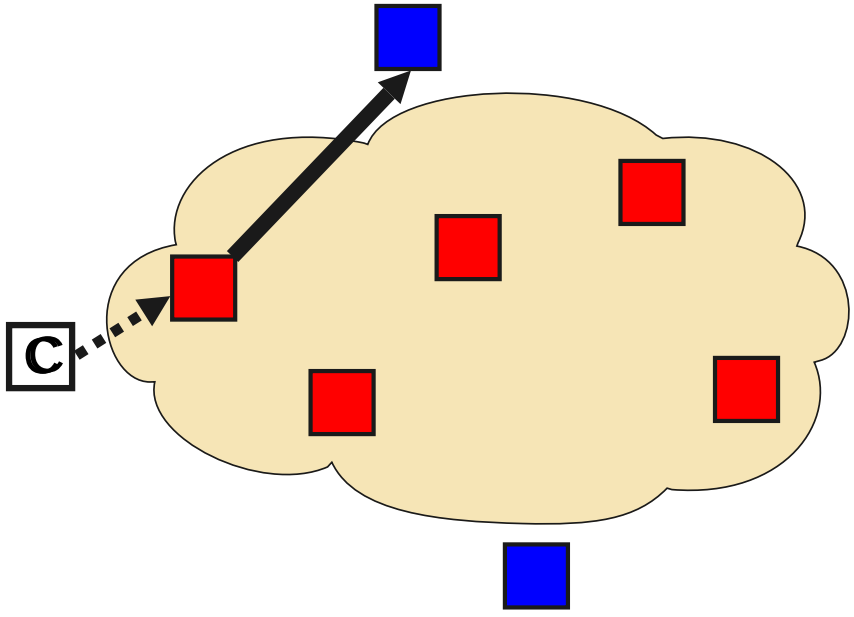



Native IP Anycast offered proximity!!

# Wouldn't it be nice if . . .

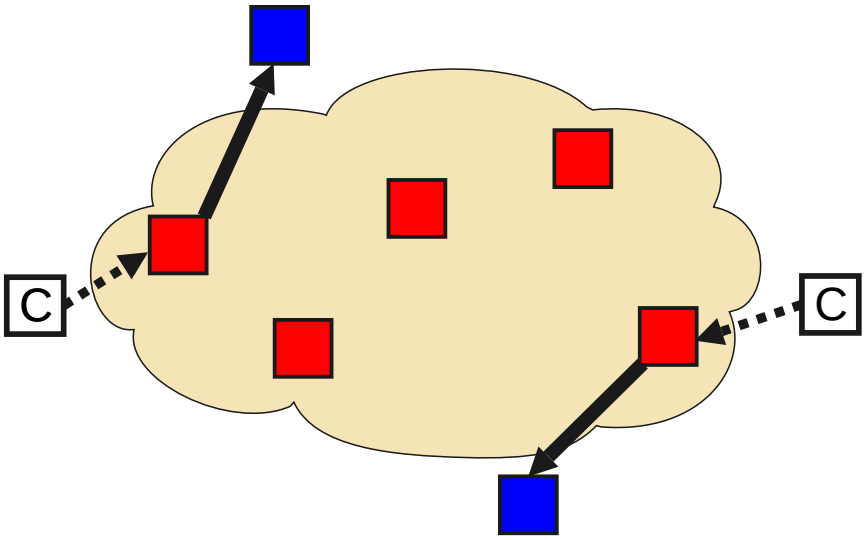



Native IP Anycast offered proximity!!

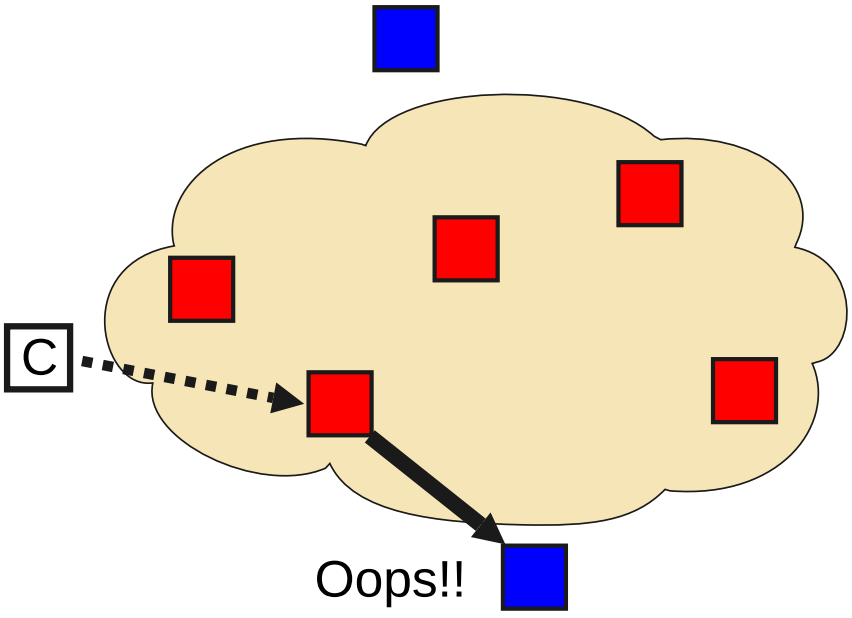



Native IP Anycast offered affinity!!

# Wouldn't it be nice if . . .




Native IP Anycast offered proximity!!




Native IP Anycast offered affinity!!

# Proximity

## IP Anycast as a server-selection primitive

- Robustness
- Efficiency
  - Proximity for free!
  - Anycast packets delivered to nearest* server

# Proximity

## IP Anycast as a server-selection primitive

- Robustness
- Efficiency
  - Proximity for free!
  - Anycast packets delivered to nearest* server

## Nearest

- Topologically (in terms of routing protocol metrics)
- Proximity in terms of other metrics?
  - Latency-based proximity

# Proximity

## IP Anycast as a server-selection primitive

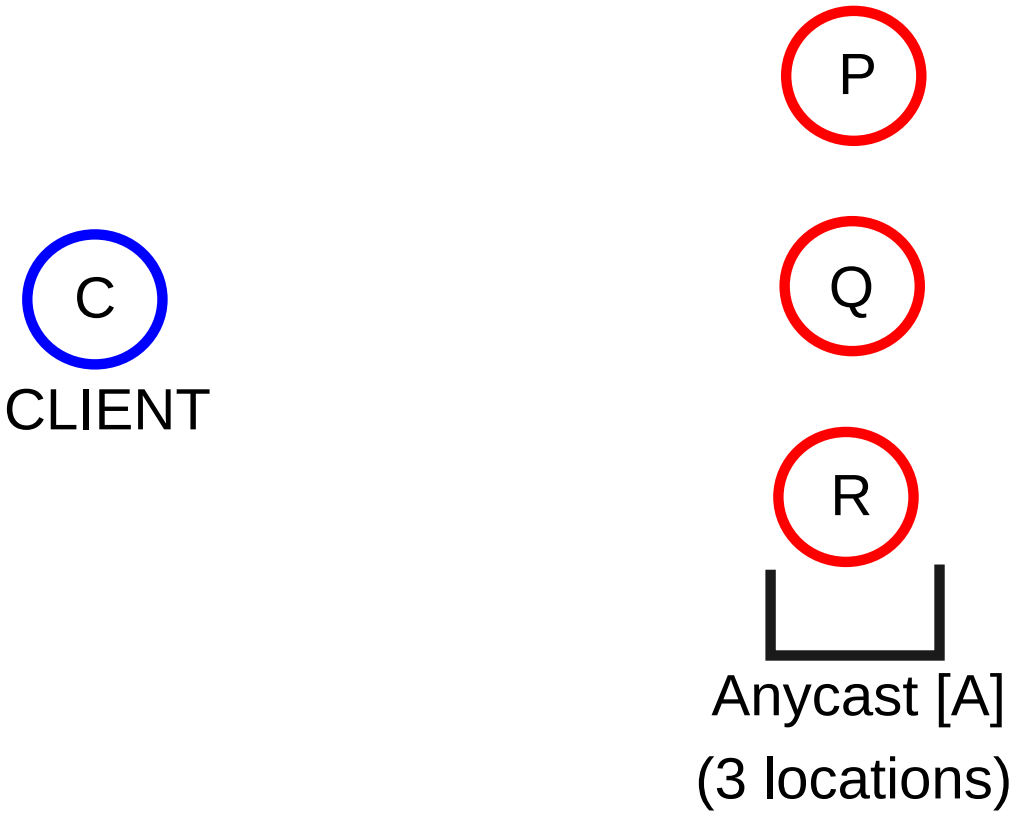
- Robustness
- Efficiency
  - Proximity for free!
  - Anycast packets delivered to nearest* server

## Nearest

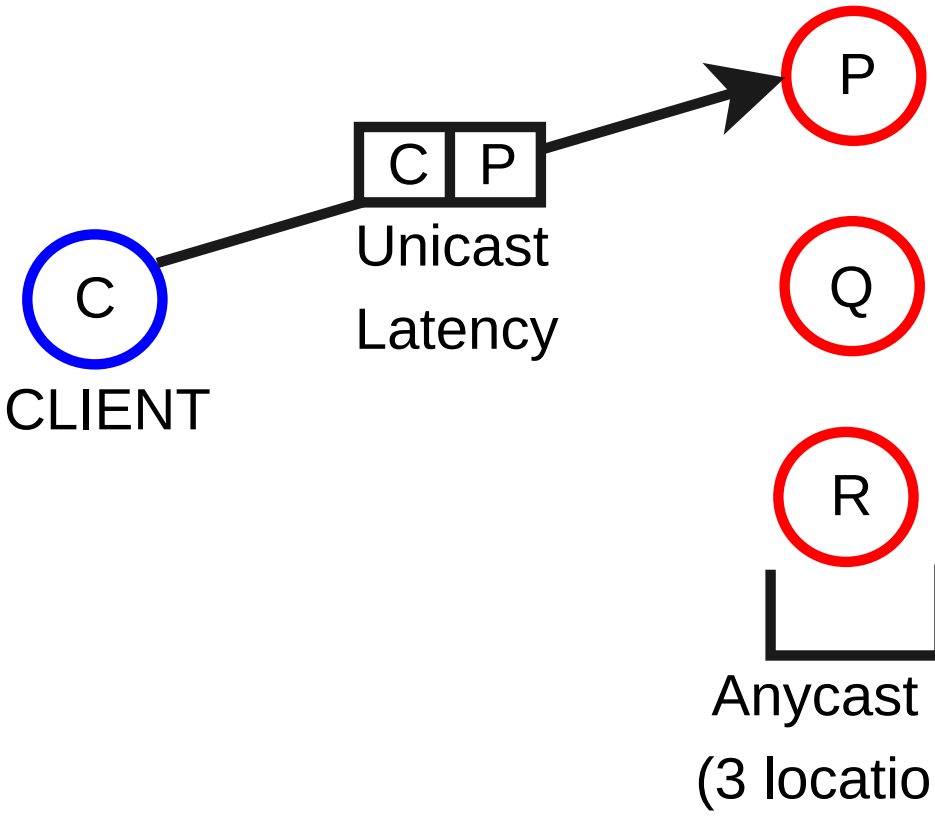
- Topologically (in terms of routing protocol metrics)
- Proximity in terms of other metrics?
  - Latency-based proximity

How good is the latency-based proximity offered by current IP Anycast deployments?

# Measuring Proximity : Methodology



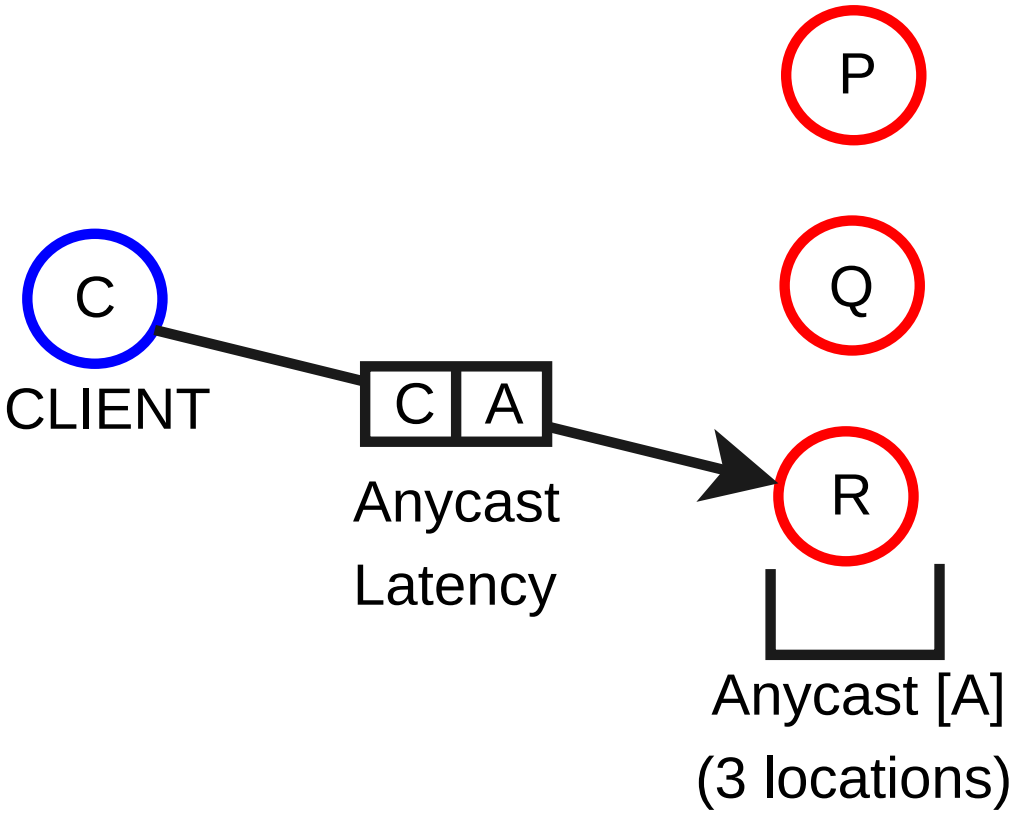
# Measuring Proximity : Methodology




## Unicast Latency

C-P, C-R and C-Q

# Measuring Proximity : Methodology



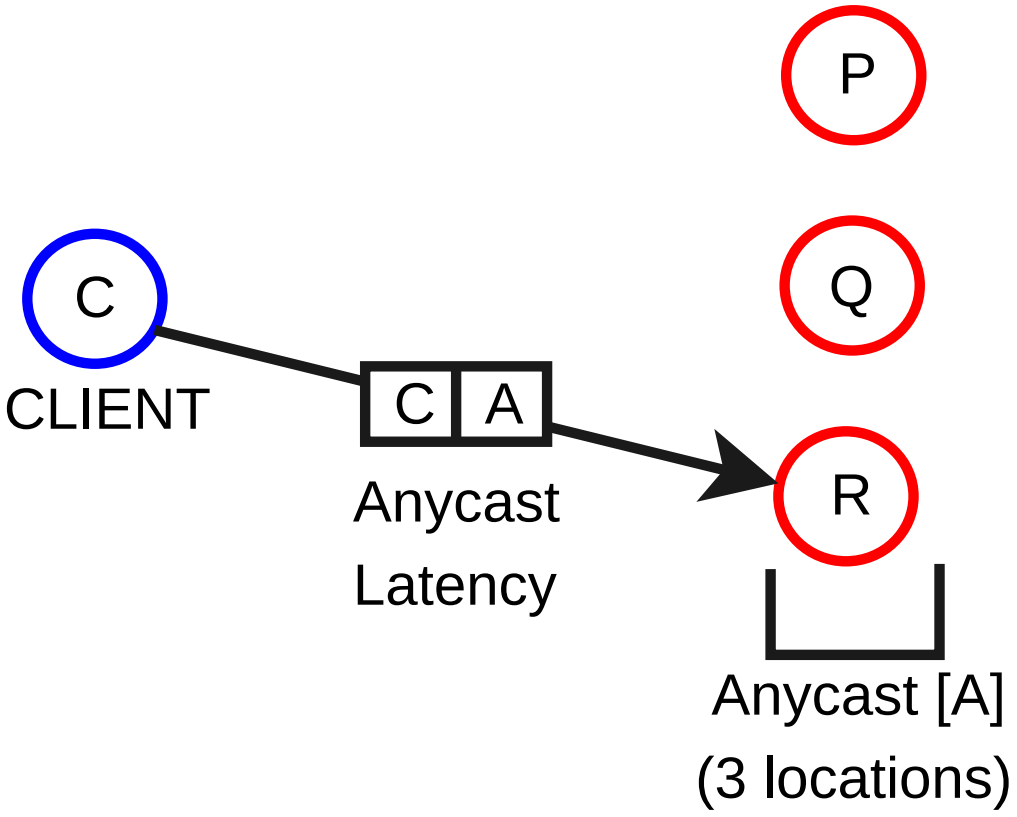

## Unicast Latency

C-P, C-R and C-Q

## Anycast Latency

C-A

# Measuring Proximity : Methodology




## Unicast Latency

C-P, C-R and C-Q

## Anycast Latency

C-A

## Metric for quality of proximity

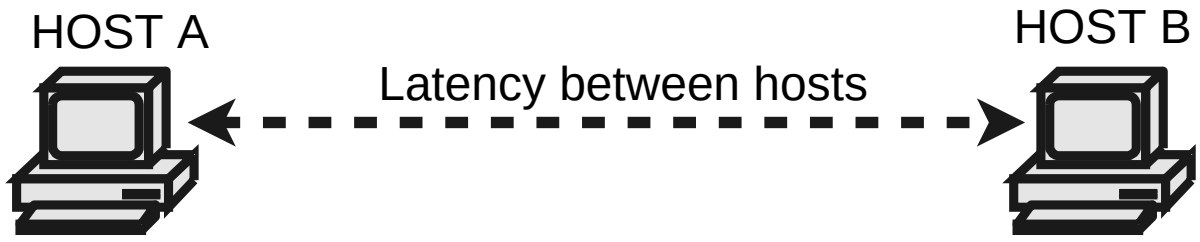
Ratio of anycast to minimum unicast latency

# Measuring Proxmity : Methodology

King : Latency between any two Internet hosts

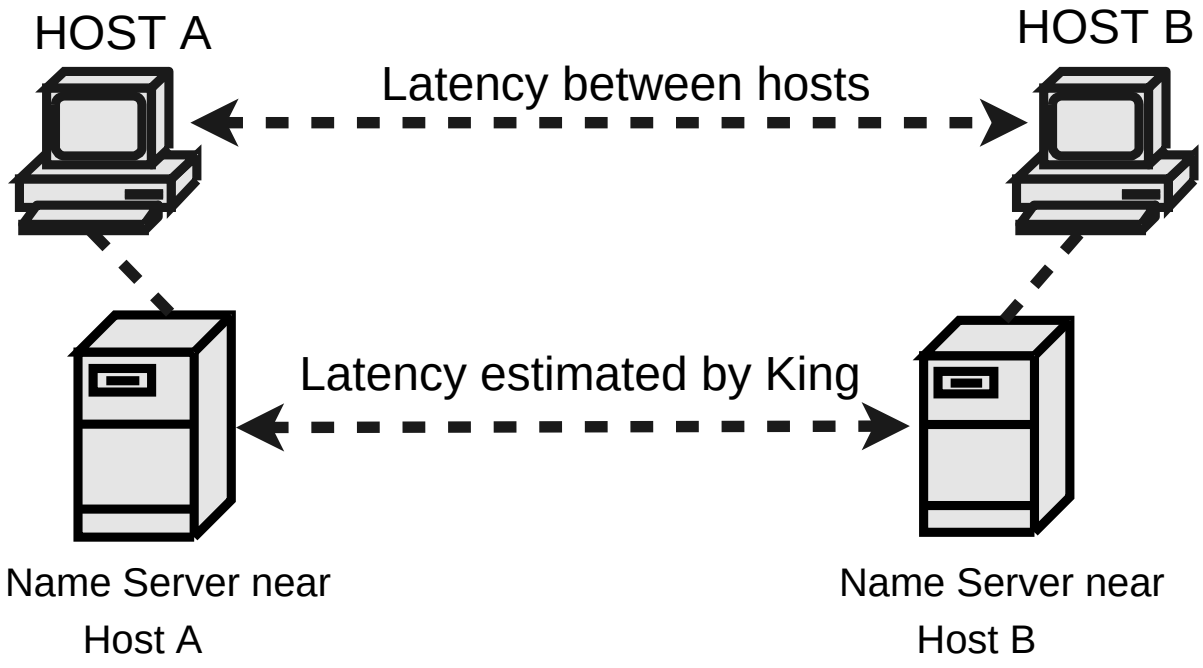
# Measuring Proxmity : Methodology

## King : Latency between any two Internet hosts



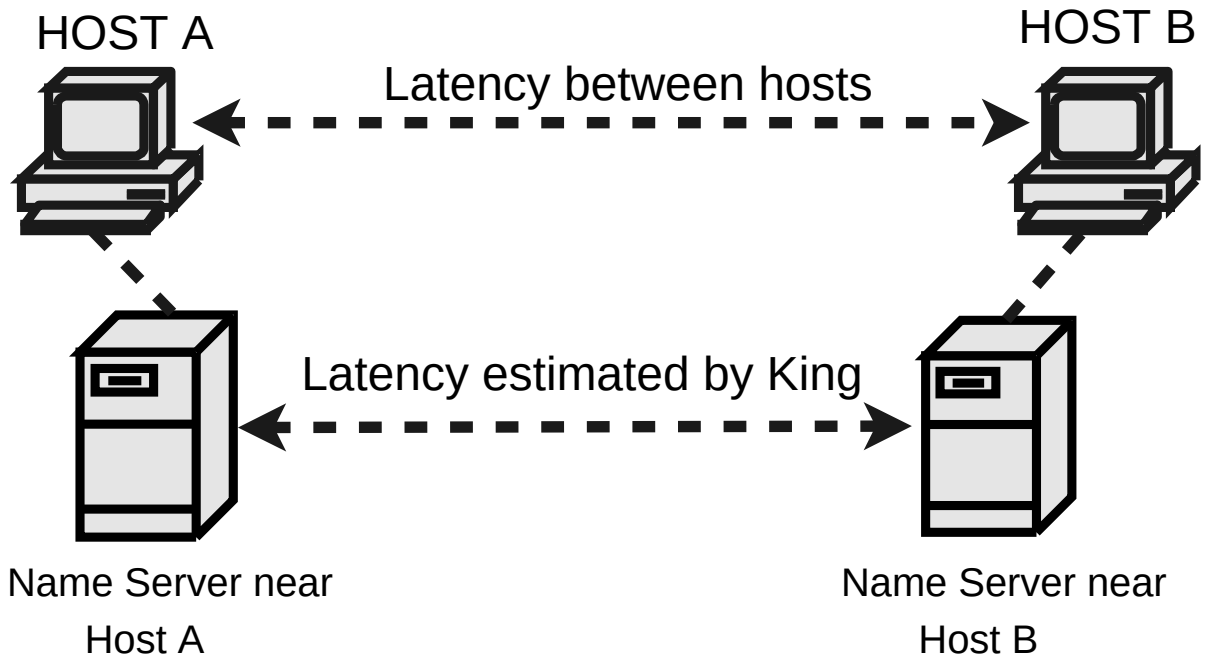
# Measuring Proxmity : Methodology

## King : Latency between any two Internet hosts



# Measuring Proxmity : Methodology

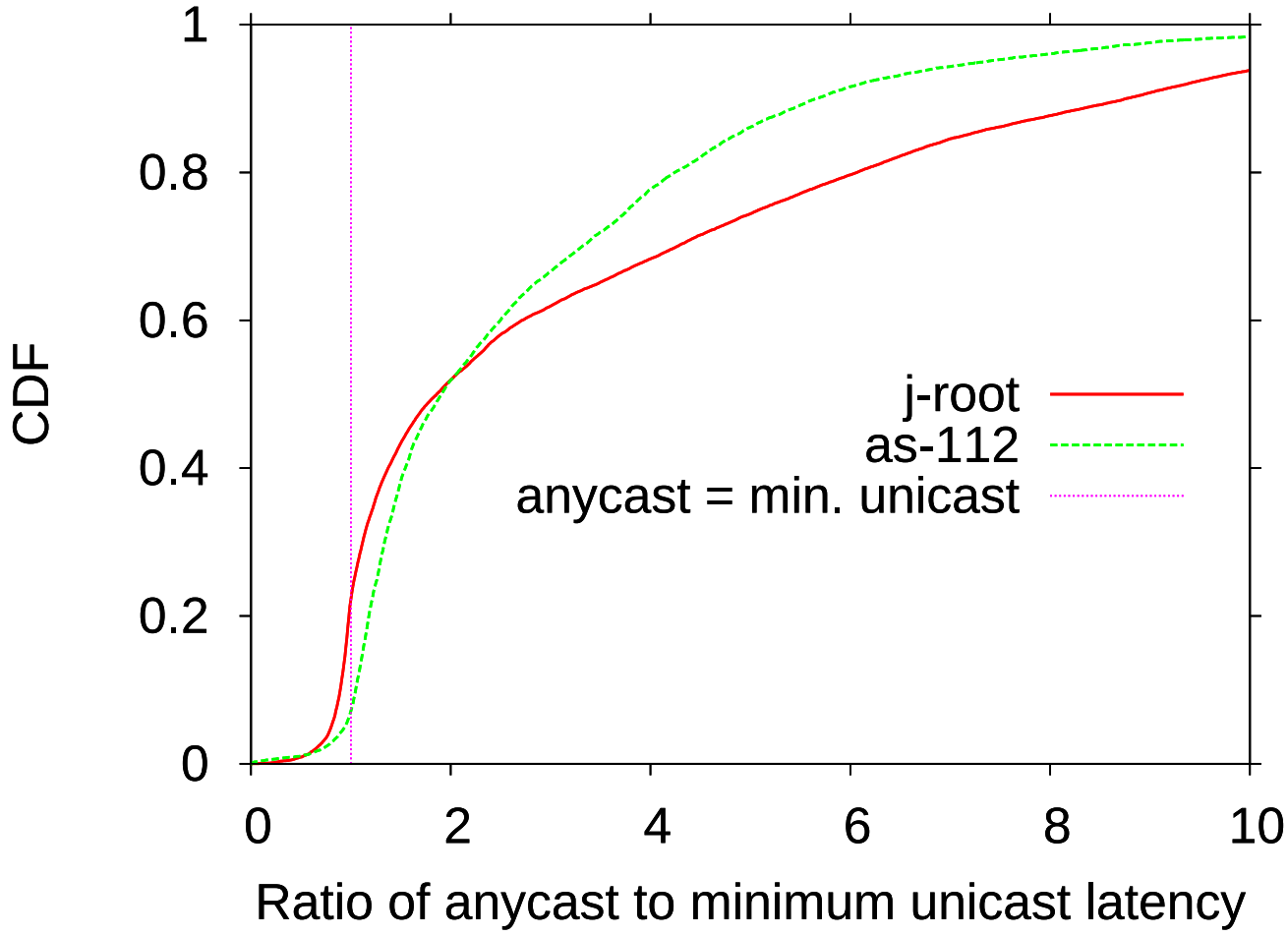
## King : Latency between any two Internet hosts




## Measured anycast deployments : J-Root, AS112

- Measured latencies from 30000 clients

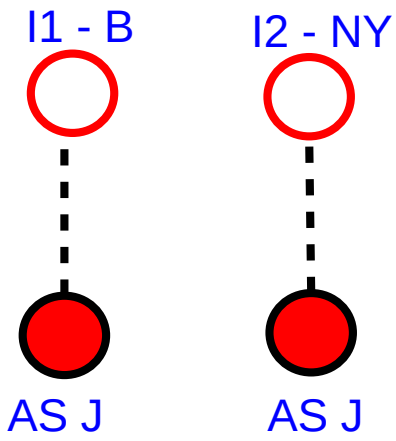
# Measuring Proximity : Results



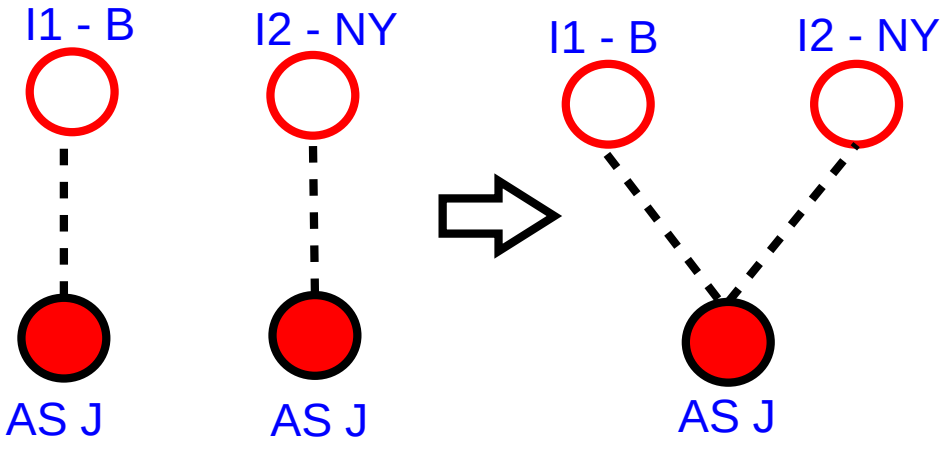

AS112 and J-Root : poor proximity

≈50% clients had a ratio >= 2

# Inter-domain routing and Anycast!!

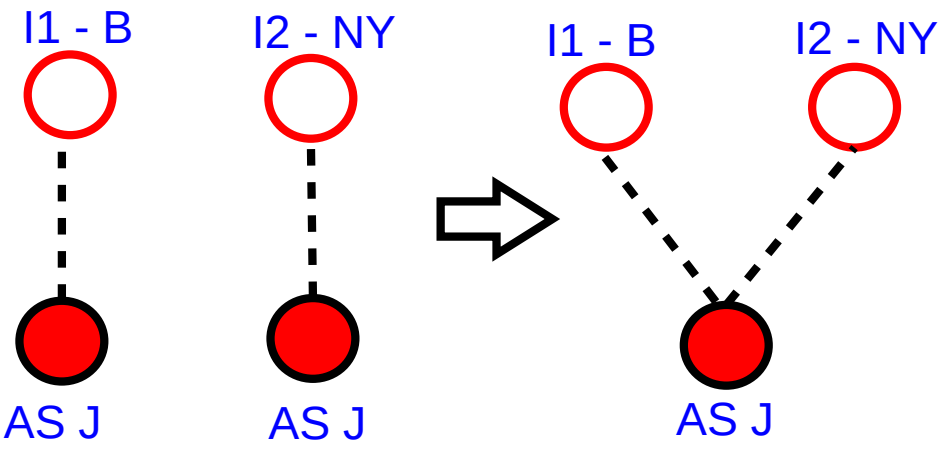


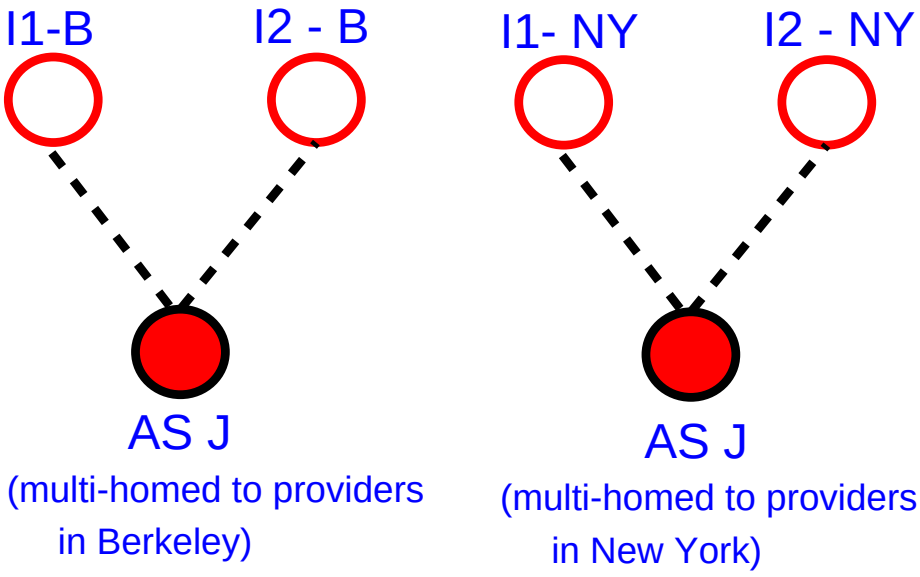
# Inter-domain routing and Anycast!!




Anycast'ed AS appears similar to a multihomed AS

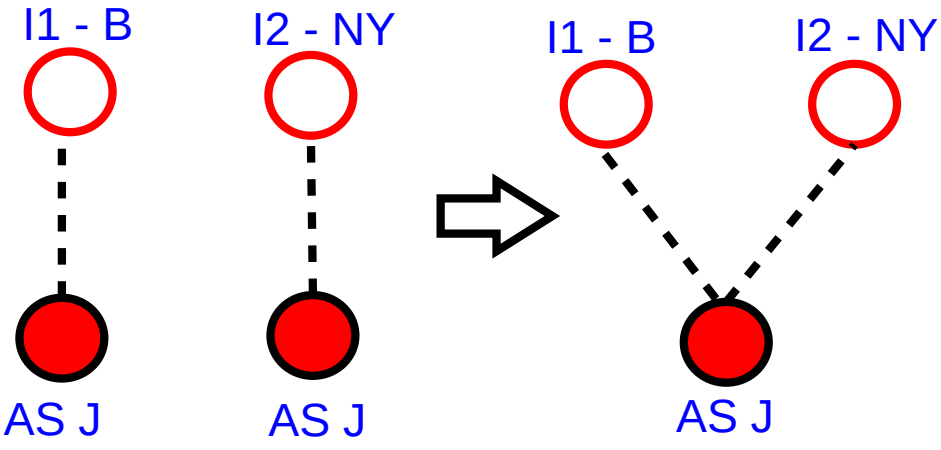
# Inter-domain routing and Anycast!!



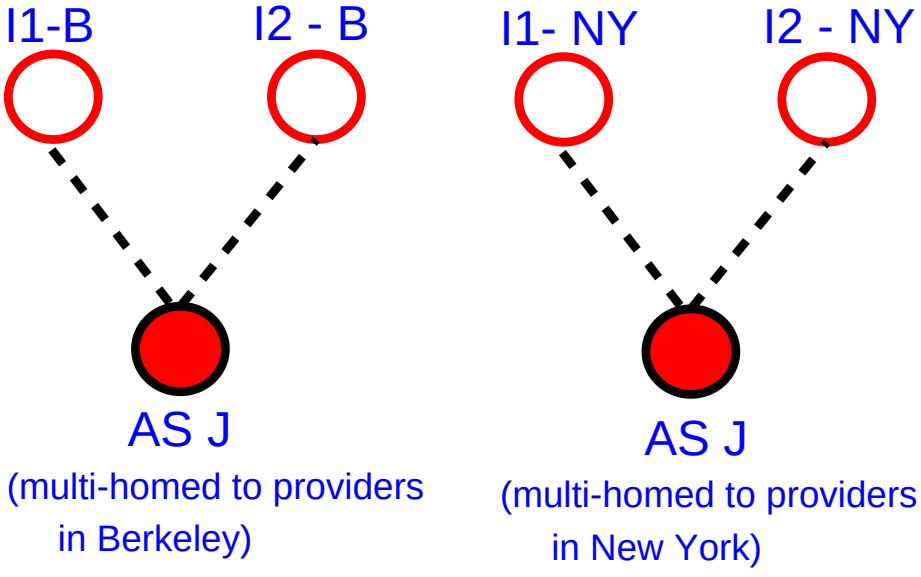




Anycast'ed AS appears similar to a multihomed AS

# Inter-domain routing and Anycast!!




Anycast'ed AS appears similar to a multihomed AS




But is different from typical multihoming!

# Our Conjecture

## Anycasting of a prefix introduces

- atypical connectivity in the AS-level Internet topology

## Current Inter-domain routing

- supports anycast out-of-the-box
- but hurts the quality of anycast

# Our Conjecture

## Anycasting of a prefix introduces

- atypical connectivity in the AS-level Internet topology

## Current Inter-domain routing

- supports anycast out-of-the-box
- but hurts the quality of anycast

## Alleviative

- Planned Deployment with proximity in mind
- Details in the technical report.

# IP Anycast : a network layer service

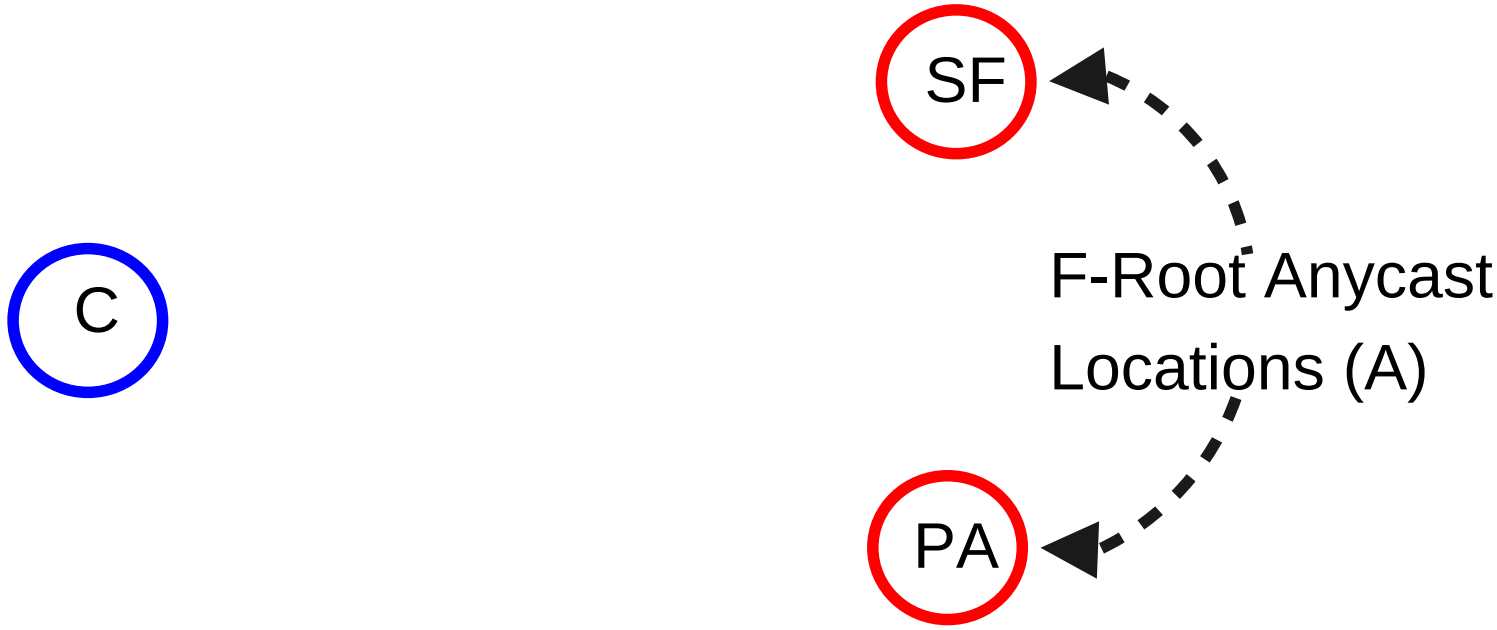
Anycast packets sent by a client can be

- Lost
- Duplicated
- Delivered to different anycast locations

# IP Anycast : a network layer service

Anycast packets sent by a client can be

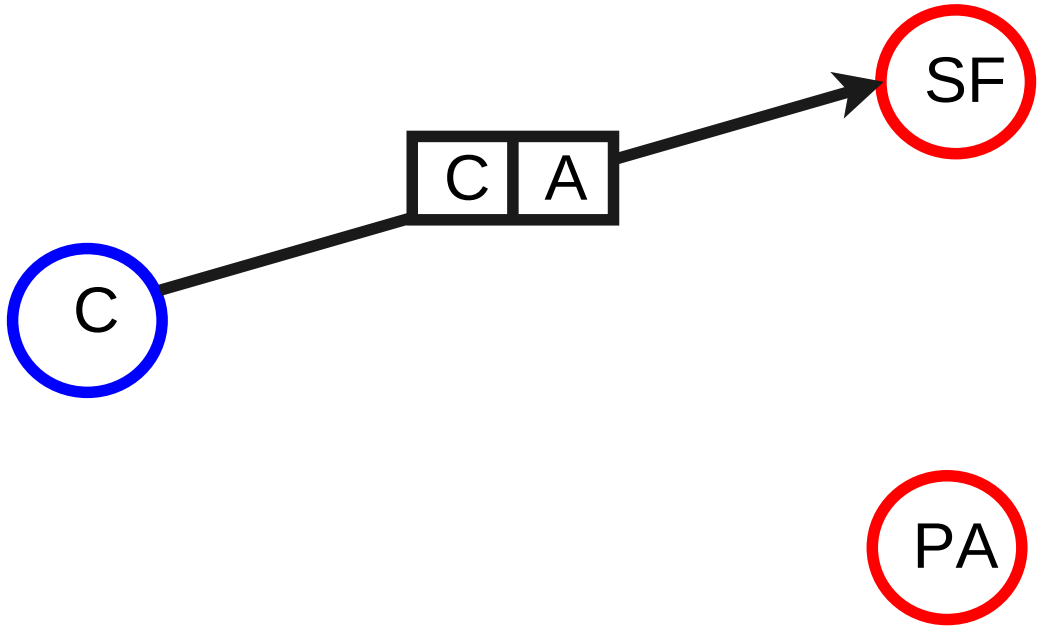
- Lost
- Duplicated
- Delivered to different anycast locations



# IP Anycast : a network layer service

Anycast packets sent by a client can be

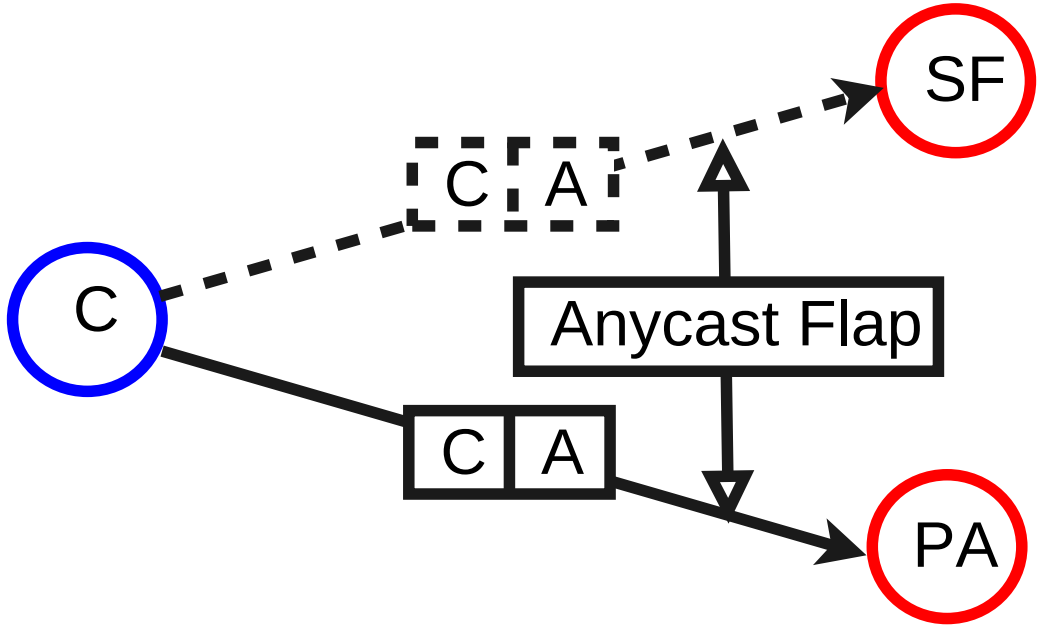
- Lost
- Duplicated
- Delivered to different anycast locations



# IP Anycast : a network layer service

Anycast packets sent by a client can be

- Lost
- Duplicated
- Delivered to different anycast locations



# Anycast Flaps

## Affinity

- Tendency of subsequent packets of a connection to be delivered to the same anycast location
- Anycast Flaps $\Rightarrow$ Lack of Affinity
- What is the affinity offered by native IP Anycast?
  - How often do anycast destinations flap?

## Why bother about affinity?

- IP Anycast affinity $\Rightarrow$ PIAS affinity
- Anycast based connection-oriented services
- Better understand inter-domain routing
  - Does anycasting interact badly with existing Internet elements?

# Anycast Flaps

## Affinity

- Tendency of subsequent packets of a connection to be delivered to the same anycast location
- Anycast Flaps $\Rightarrow$ Lack of Affinity
- What is the affinity offered by native IP Anycast?
  - How often do anycast destinations flap?

## Why bother about affinity?

- IP Anycast affinity $\Rightarrow$ PIAS affinity
- Anycast based connection-oriented services
- Better understand inter-domain routing
  - Does anycasting interact badly with existing Internet elements?

# Measuring Affinity : Methodology

## Affinity for exisiting anycast deployments

- Anycast Root-Servers (C,F,I,J,K,M)
- AS112 Servers (answer PTR queries for private addresses)

## Which anycast location is a client accessing?

- Location querying supported by aforementioned destinations
- TXT-type DNS query
  eg. dig +norec @F.ROOT-SERVERS.NET HOSTNAME.BIND CHAOS TXT

## Active probing to measure affinity

- Location-probe (UDP) every 10 seconds
- Flap: consecutive probes to different locations

# Measuring Affinity : Data Collected

## Planetlab [PL-set]

- 163 Planetlab sites
- Duration : 3 months (Dec'04-Mar'05)

| | |
|---|---|
| Africa | 0 |
| Asia | 22 |
| Australia | 3 |
| S.America | 1 |
| Canada | 12 |
| Europe | 31 |
| US | 94 |
| Total | 163 |

# Measuring Affinity : Data Collected

## Planetlab [PL-set]

- 163 Planetlab sites
- Duration : 3 months (Dec'04-Mar'05)

| | |
|---|---|
| Africa | 0 |
| Asia | 22 |
| Australia | 3 |
| S.America | 1 |
| Canada | 12 |
| Europe | 31 |
| US | 94 |
| Total | 163 |

## Traceroute-Servers [TS-set]

- 244 vantage points
- Traceroute'd to anycast destinations
- Load restrictions
  - Probe every 60 seconds
  - Duration : one week each

| | |
|---|---|
| Africa | 3 |
| Asia | 26 |
| Australia | 12 |
| S.America | 8 |
| Canada | 1 |
| Europe | 152 |
| US | 42 |
| Total | 244 |

# Measuring Affinity : Results

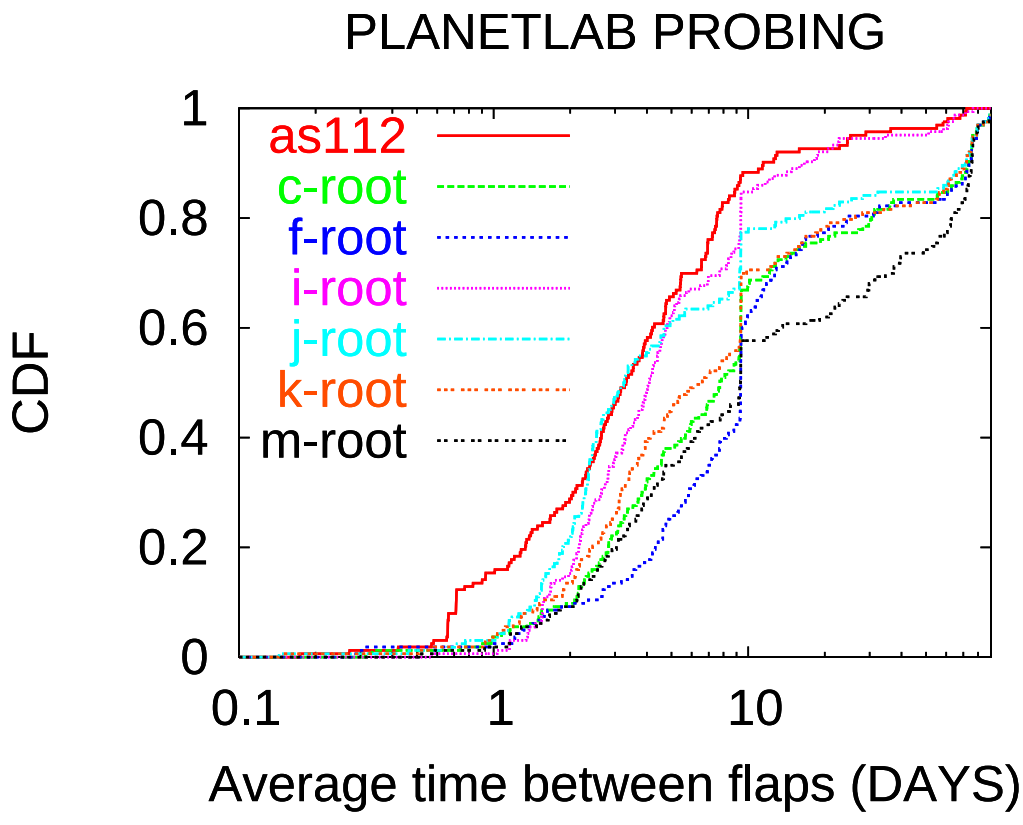
## Metric of Affinity

Average Inter-flap interval

# Measuring Affinity : Results

## Metric of Affinity

Average Inter-flap interval



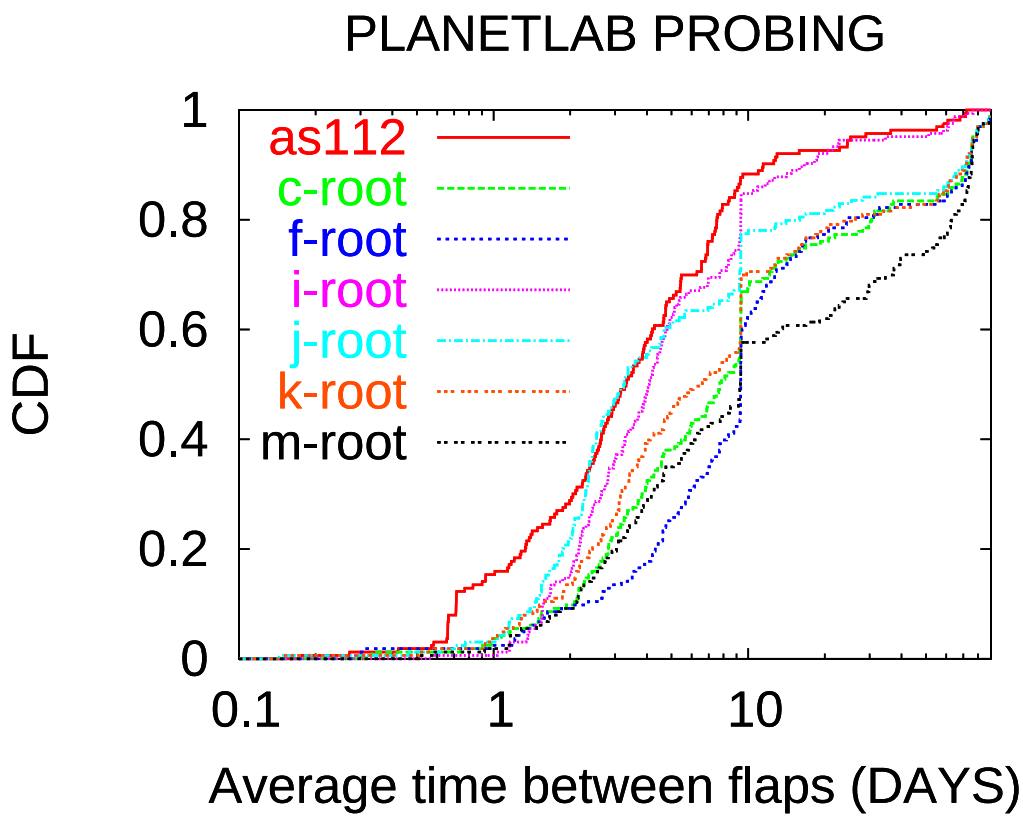

Planetlab

Very Stable

# Measuring Affinity : Results

## Metric of Affinity

Average Inter-flap interval




Less than 1 flap per day for $> 95\%$ of nodes

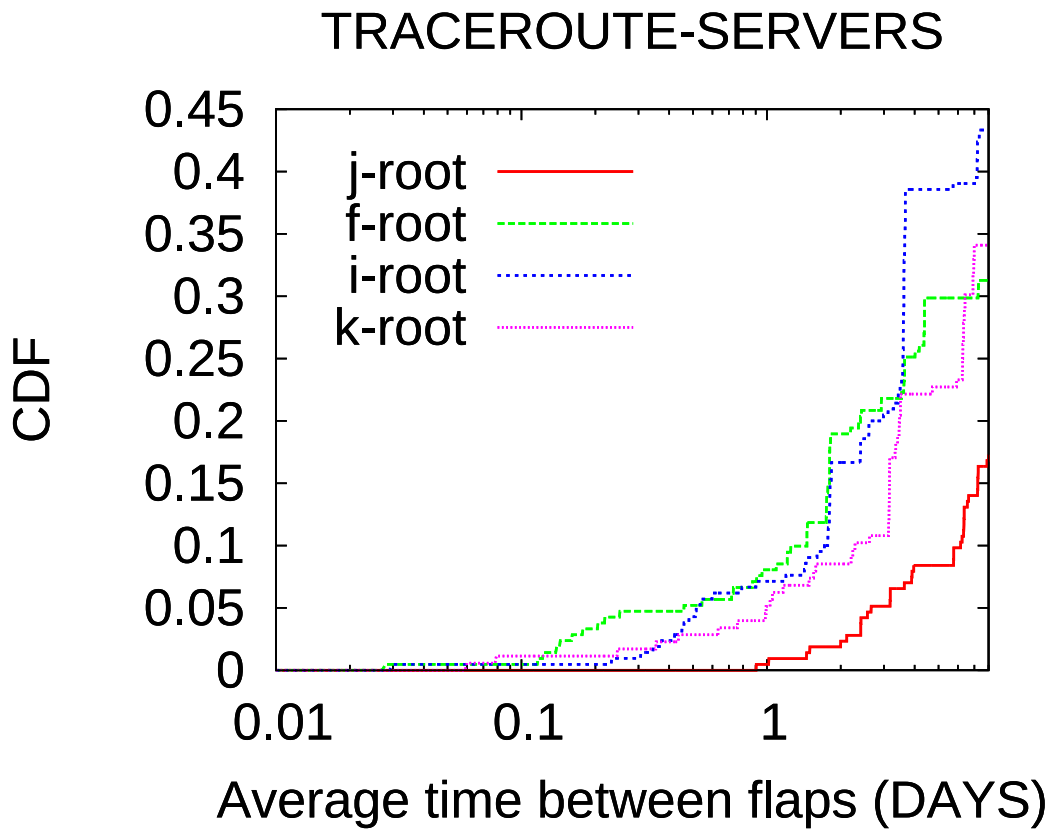
## Planetlab

Very Stable

# Measuring Affinity : Results

## Metric of Affinity

Average Inter-flap interval




## Traceroute-Servers

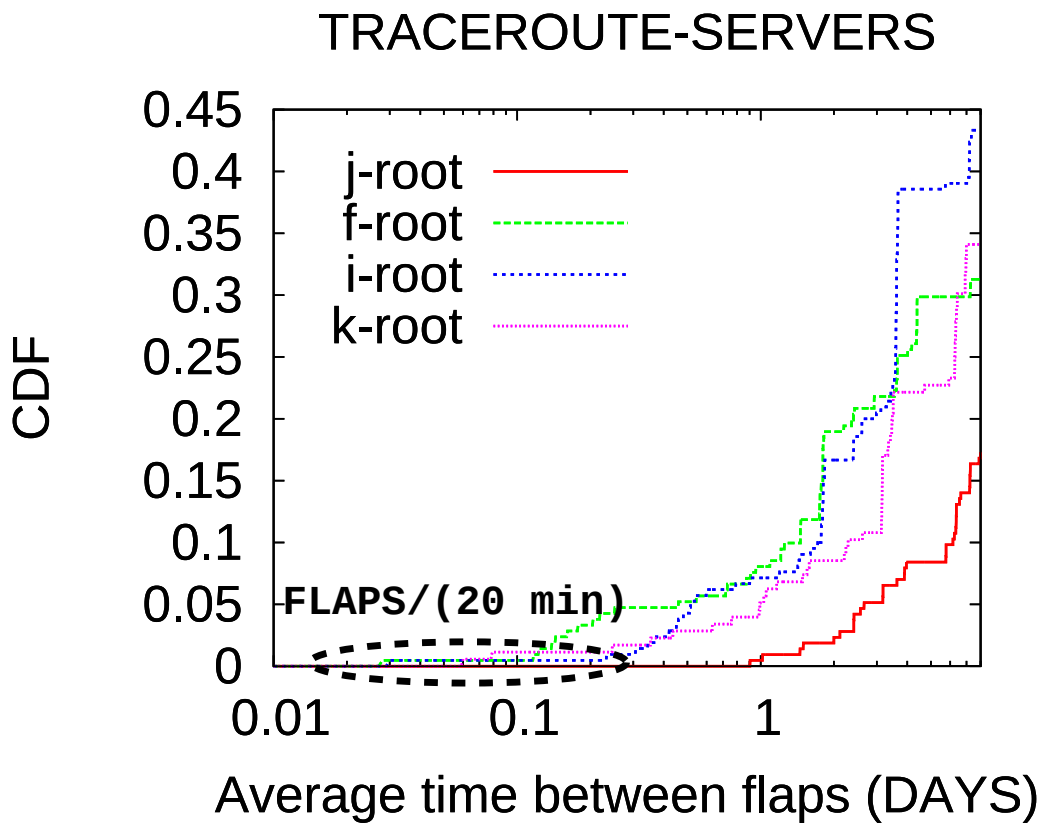
Very Stable

- with exceptions!

# Measuring Affinity : Results

## Metric of Affinity

Average Inter-flap interval




### Traceroute-Servers

Very Stable

- with exceptions!

# Measuring Affinity : Concluding Remarks

## BGP-level analysis

- Data from Route-Views and RIPE RIS
- Low activity for anycast prefixes

# Measuring Affinity : Concluding Remarks

## BGP-level analysis

- Data from Route-Views and RIPE RIS
- Low activity for anycast prefixes

## Our findings . . . summarized

- Measured anycast deployments offer good affinity
- Confirmed by BGP-level stability analysis
- ✗ Fast switching across providers

# Measuring Affinity : Concluding Remarks

## BGP-level analysis

- Data from Route-Views and RIPE RIS
- Low activity for anycast prefixes

## Our findings ... summarized

- Measured anycast deployments offer good affinity
- Confirmed by BGP-level stability analysis
- ✗ Fast switching across providers

## Other affinity studies : Barber et. al., Boothe et. al. and Daniel Karrenberg

- Reported lack of affinity in anycast

# Measuring Affinity : Concluding Remarks

## BGP-level analysis

- Data from Route-Views and RIPE RIS
- Low activity for anycast prefixes

## Our findings . . . summarized

- Measured anycast deployments offer good affinity
- Confirmed by BGP-level stability analysis
- ✗ Fast switching across providers

## Other affinity studies : Barber et. al., Boothe et. al. and Daniel Karrenberg

- Reported lack of affinity in anycast
  - Not sure why :(
  - Bias due to the vantage points chosen
  - Data may be the same . . . interpretations differ

# Thanks!

PIAS: pias.gforge.cis.cornell.edu

e-mail : hitesh@cs.cornell.edu