Inferring AS Relationships from BGP Attributes

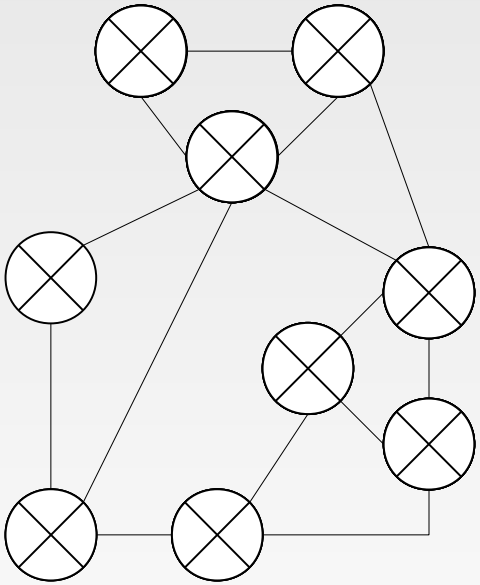
Vasileios Giotsas, Shi Zhou

Department of Computer Science, UCL

MoN10: Tenth Mathematics of Networks meeting, September 2011

Introduction

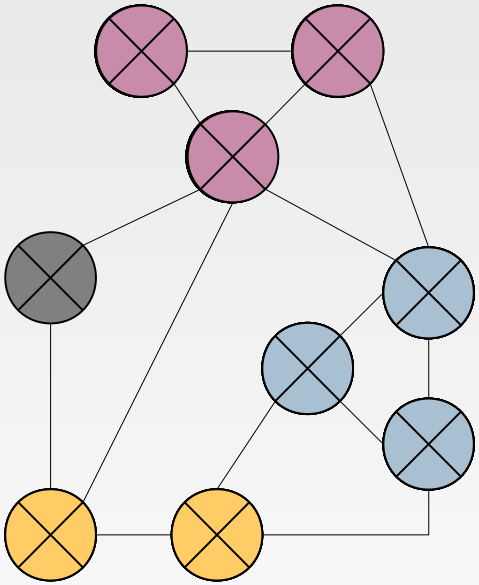
- The Internet is a *Network of Networks*



Routers Topology

Introduction

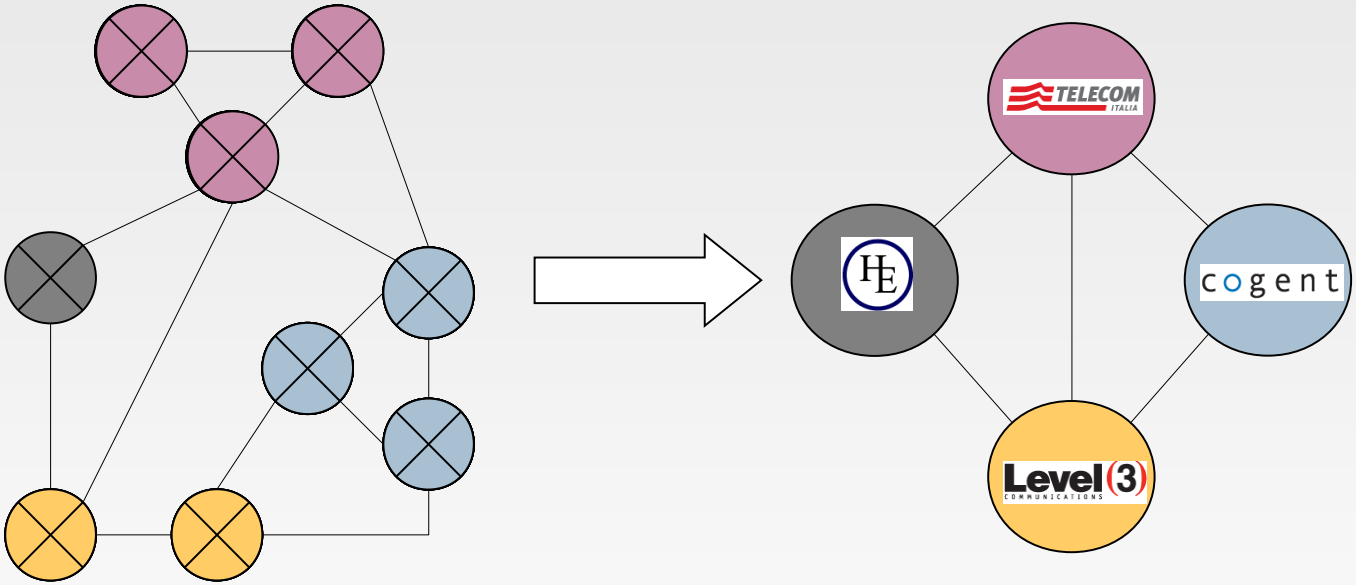
- The Internet is a *Network of Networks*



Routers Topology

Introduction

- The Internet is a *Network of Networks*



Routers Topology

Autonomous Systems Topology

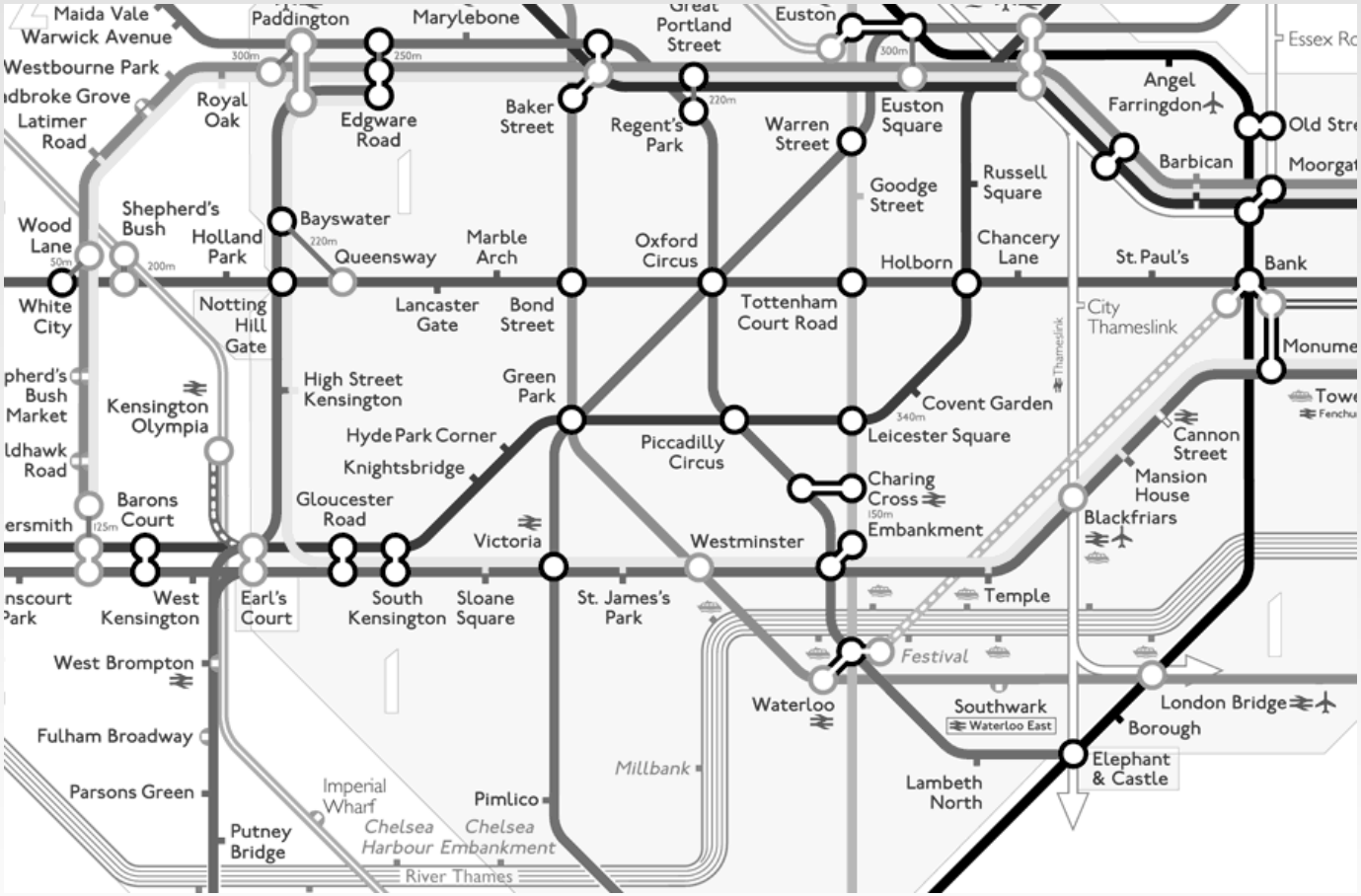
Why AS Topology?

- Two levels of routing
  - *Intra-domain routing*
  - *Inter-domain routing* – *Border Gateway Protocol (BGP)*
- Performance
- Traffic Engineering
- Security
- Business policies/economics

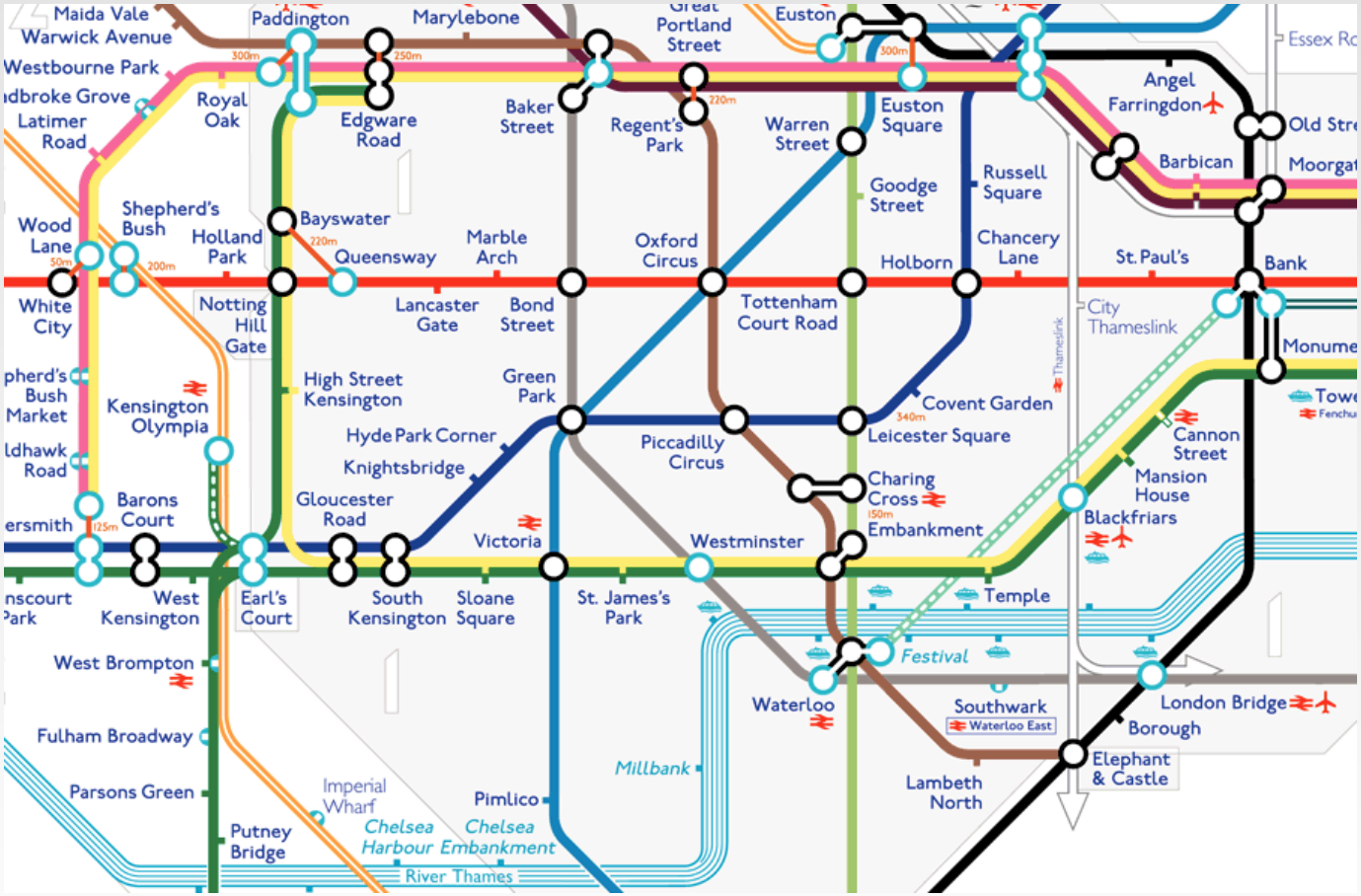
Autonomous Systems Business Relationships

- Customer-to-Provider (c2p)
  - Paid transit
- Peer-to-Peer (p2p)
  - Free bilateral transit, routing restrictions
- Sibling-to-Sibling (s2s)
  - Free bilateral transit, no restrictions

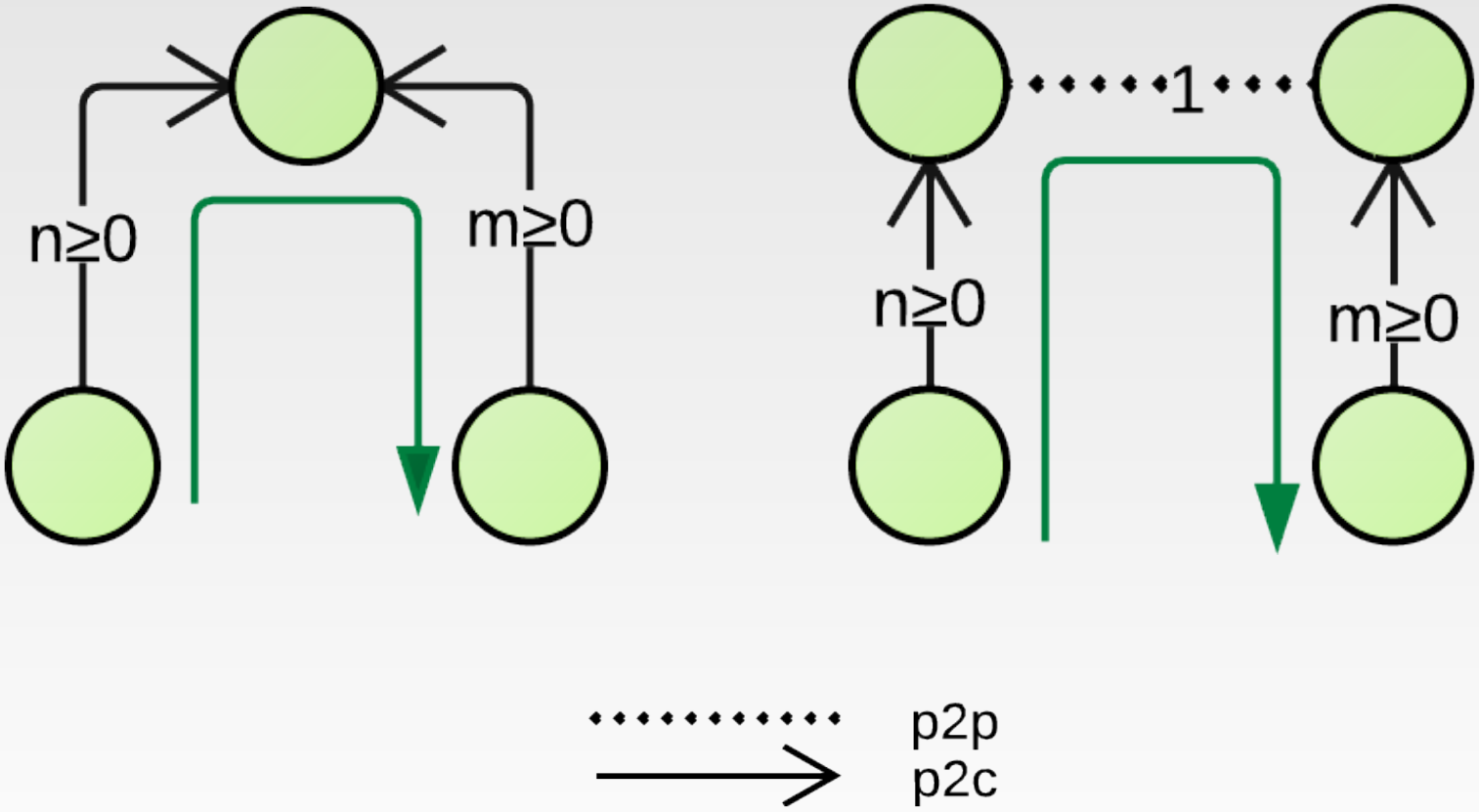
Why AS Relationship?



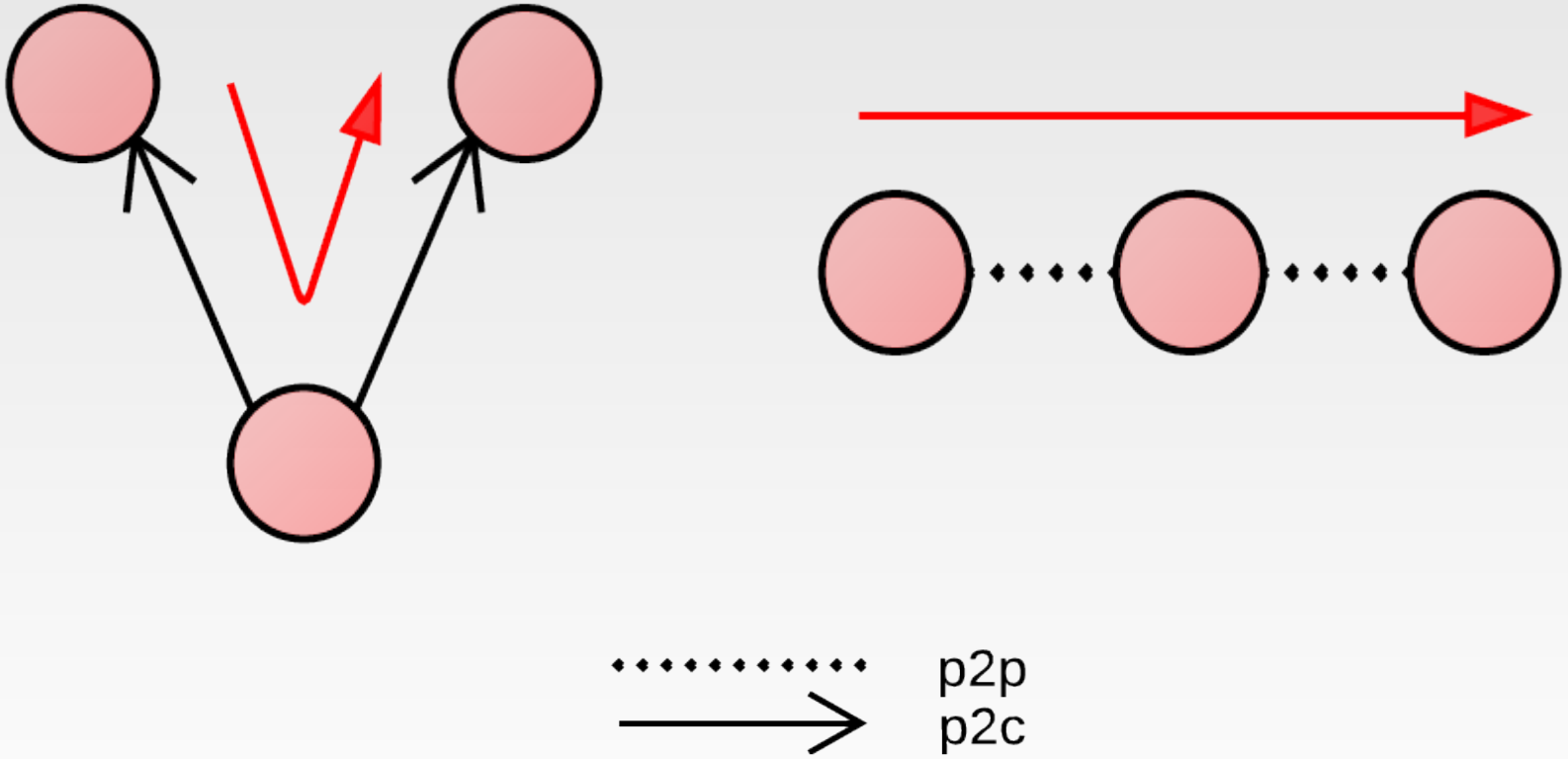
Why AS Relationship?



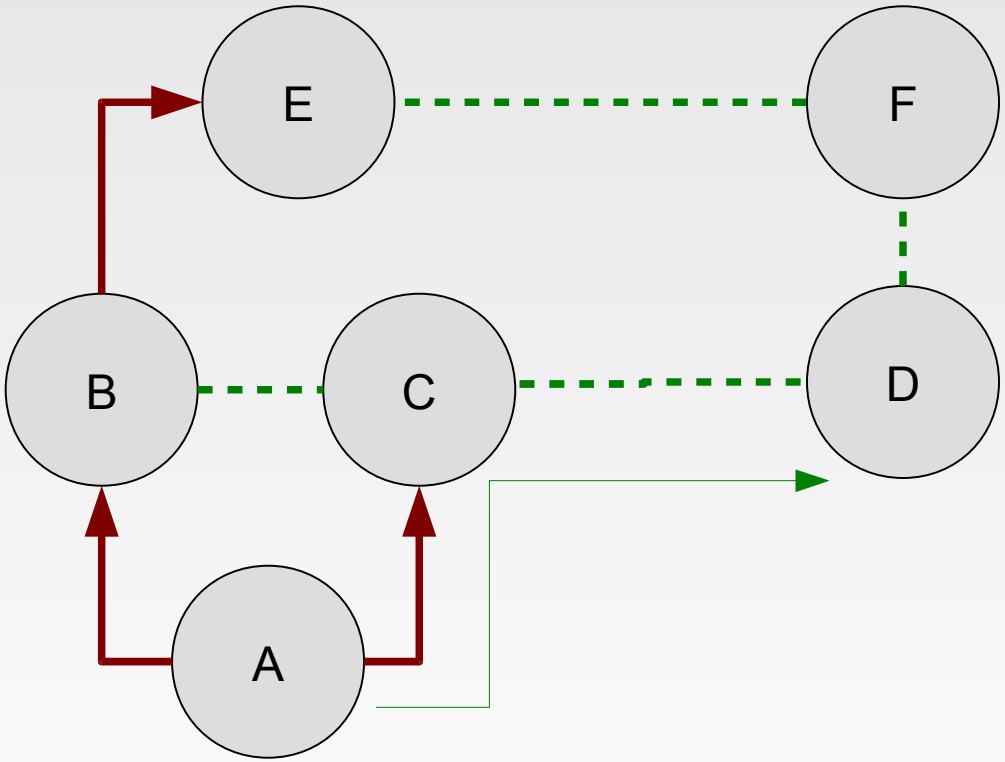
Valley-Free Paths



Non Valley-Free Paths



Valley-Free Routing

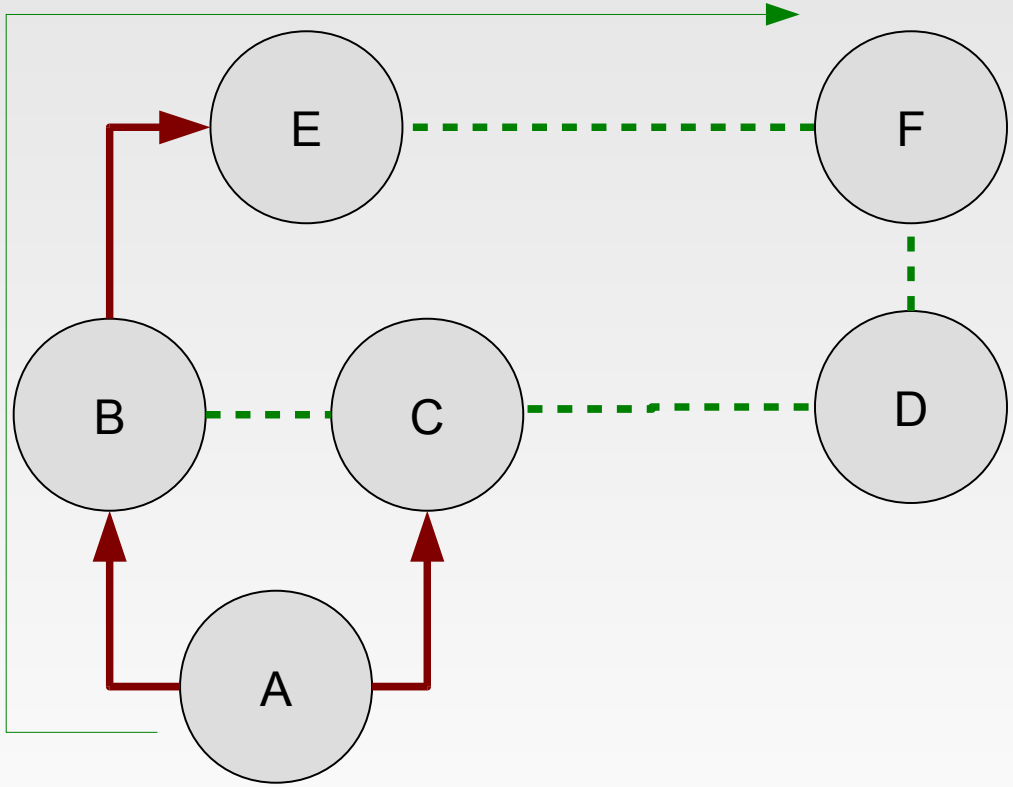


A C D : Valley-Free

c2p

p2p

Valley-Free Routing



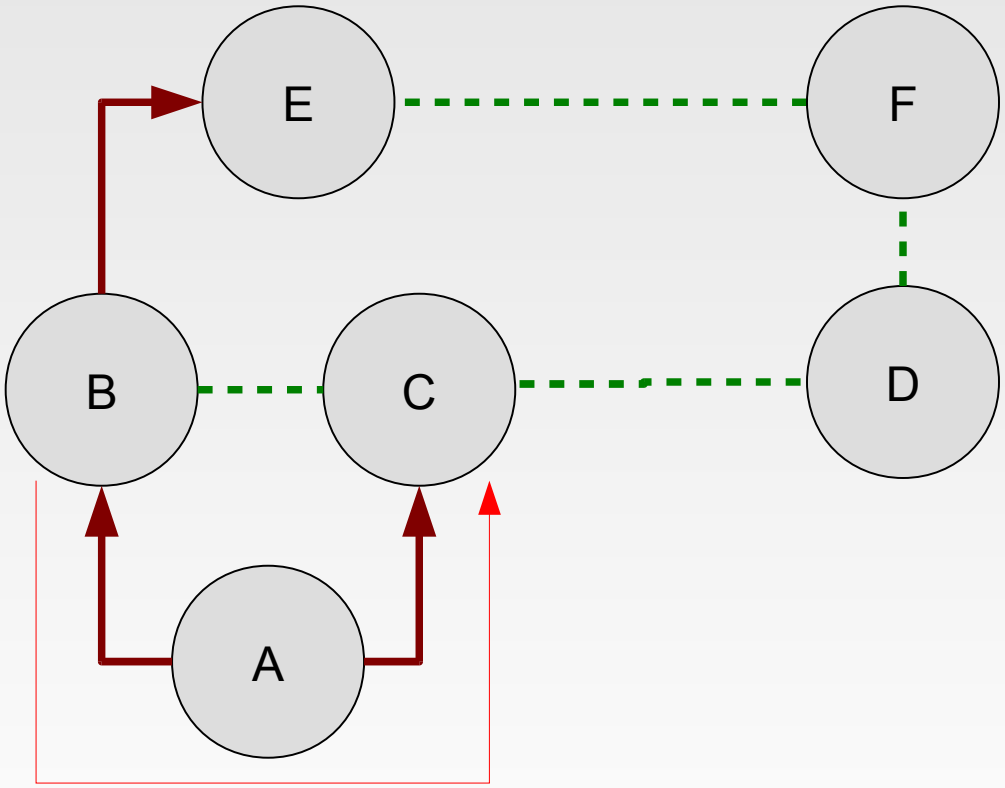
A C D : Valley-Free

A B E F : Valley-Free

c2p

p2p

Valley-Free Routing



A C D : Valley-Free

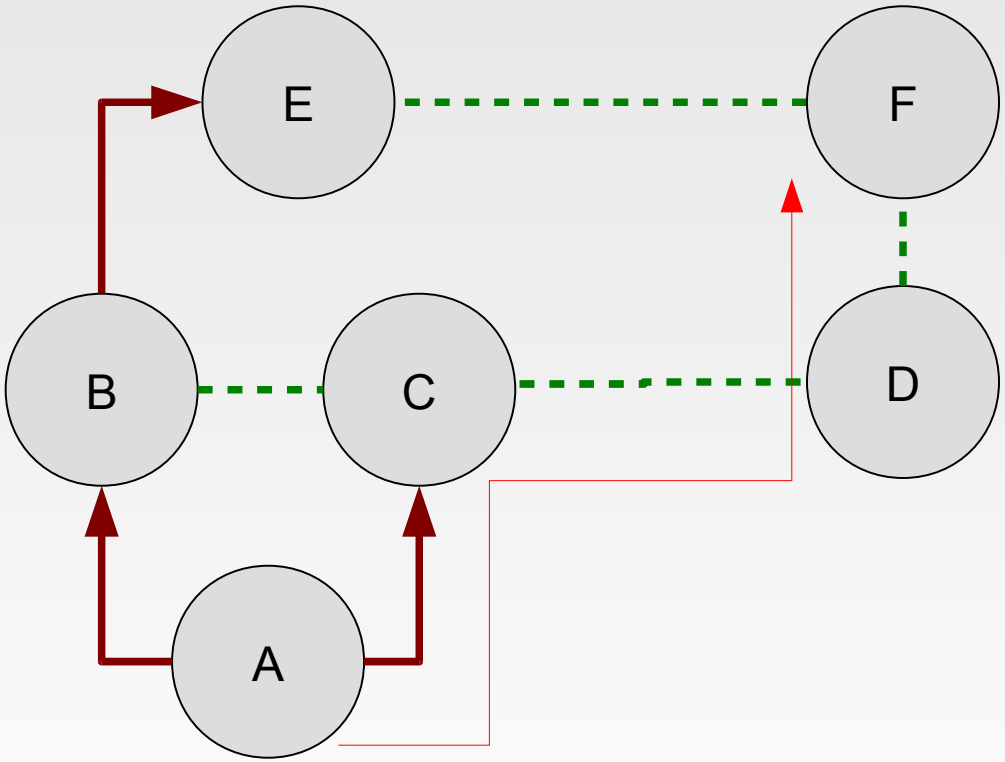
A B E F : Valley-Free

B A C : Non Valley-Free

c2p

p2p

Valley-Free Routing



A C D : Valley-Free

A B E F : Valley-Free

B A C : Non Valley-Free

A C D F: Non Valley-Free

c2p

p2p

AS Relationship Inference Problem

- AS relationships are not publicly disclosed
- **How to assign AS relationships to AS edges given the publicly available BGP/traceroute data?**

AS Relationship Inference: Existing Approaches

- *AS Topology* + *Heuristics*

- Maximize the number of valley-free paths
- p2p relationships are agreed between ASes of comparable degree
- All p2c AS edges will cross the Tier-1
- All long-lived paths (> 2 days) are valley-free

AS Relationship Inference: Existing Approaches

- *AS Topology* + *Heuristics*

Different Algorithms result in significantly conflicting results!

comparable degree

- All p2c AS edges will cross the Tier-1
- All long-lived paths (> 2 days) are valley-free

BGP Communities

- Optional BGP attribute that encodes meta-data on an AS Path
  - AS Relationships, Routing policies, Geographical information
- Non-standardized values, each AS defines its own

  32-bit values *xxxx:yyyy*
  - *xxxx*: Autonomous System Number
  - *yyyy*: Community value

BGP Communities

```
TYPE: TABLE_DUMP_V2/IPV4_UNICAST
PREFIX: 1.22.73.0/24
FROM: 206.223.115.10 AS4589
ORIGIN: IGP
ASPATH: 4589 15412 18101 45528
NEXT_HOP: 206.223.115.10
COMMUNITY: 4589:2 4589:410 4589:612
4589:14413 15412:604 15412:614 15412:621
15412:705 15412:1431 18101:1344
18101:50120 18101:50420
```

Sample of BGP entry

BGP Communities

```
TYPE: TABLE_DUMP_V2/IPV4_UNICAST
PREFIX: 1.22.73.0/24
FROM: 206.223.115.10 AS4589
ORIGIN: IGP
ASPATH: 4589 15412 18101 45528
NEXT_HOP: 206.223.115.10
COMMUNITY: 4589:2 4589:410 4589:612
4589:14413 15412:604 15412:614 15412:621
15412:705 15412:1431 18101:1344
18101:50120 18101:50420
```

Sample of BGP entry

BGP Communities

```
TYPE: TABLE_DUMP_V2/IPV4_UNICAST
PREFIX: 1.22.73.0/24
FROM: 206.223.115.10 AS4589
ORIGIN: IGP
ASPATH: 4589 15412 18101 45528
NEXT_HOP: 206.223.115.10
COMMUNITY: 4589:2 4589:410 4589:612
4589:14413 15412:604 15412:614 15412:621
15412:705 15412:1431 18101:1344
18101:50120 18101:50420
```

Sample of BGP entry

Interpretation of BGP Communities

```
Two-digit communities
---------------------

Customers can set two-digit communities to control which local preference
prefixes receive.

        Community       Local Preference
      ------------------------------------------------------------

        4589:10          50     (equiv. to last resort transit)
        4589:20         100     (equiv. to peering and transit)
        4589:25         130     (depreferred customer route)
        4589:30         150     (default for customers)
        4589:35         170     (preferred customer route)


Three-digit communities
-----------------------

Prefixes coming from peers and transit will be tagged with three-digit
community values, e.g. a prefix received at DECIX will be tagged with
4589:641. Only the most specific community is added, e.g. a route from
DECIX will not have 4589:640 set.
Additionally prefixes from peers will be tagged with a 4xx community
based on speed of the interconnection.

        Community        Entry point
    ------------------------------------------------------------------------
        4589:4xx        Special Markings
        4589:410        From a high capacity IXP or Private Peer
        4589:420        From a low capacity IXP or Private Peer
```

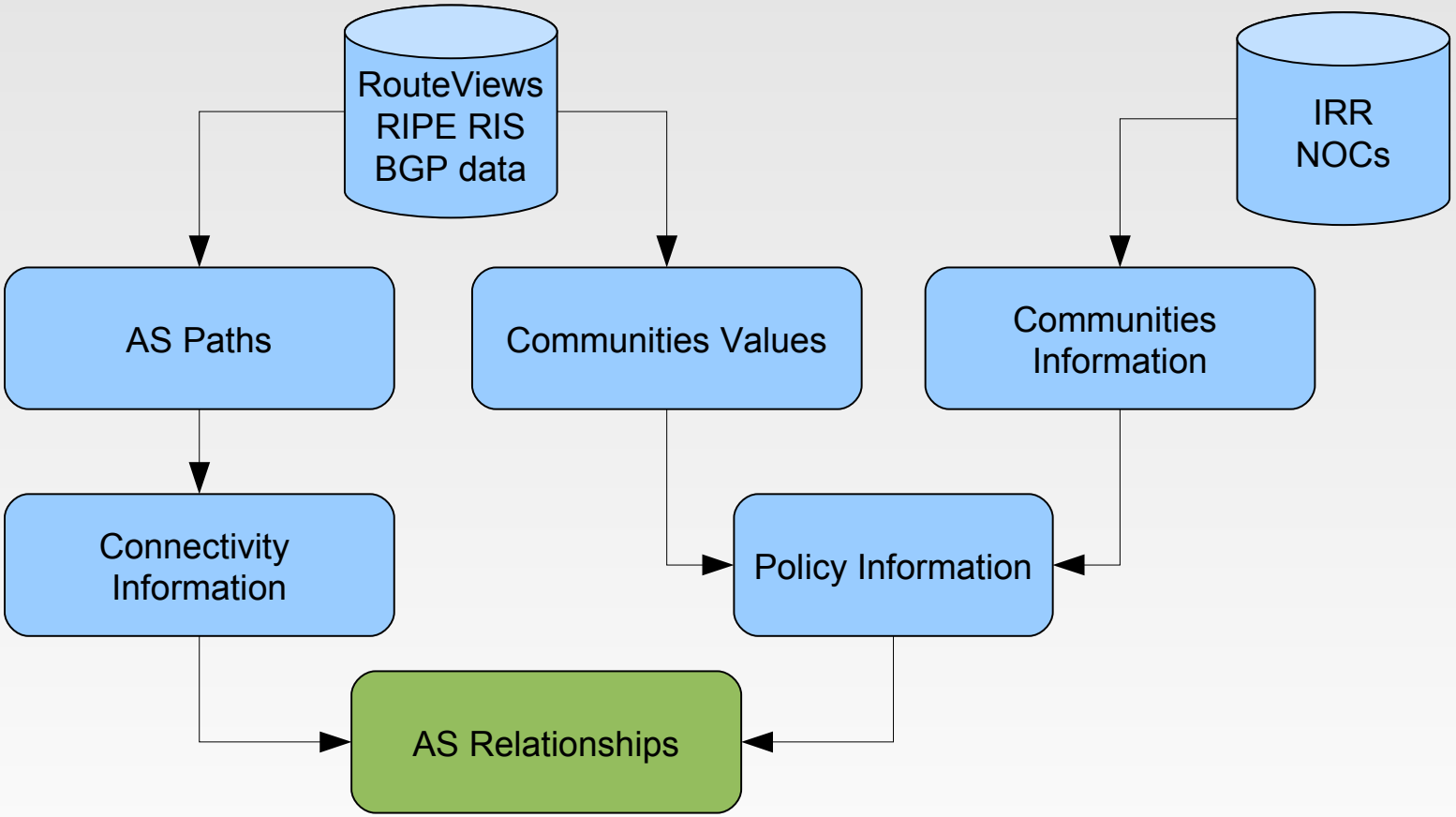
Network Operation Centers (NOCs) (e.g. lg.easynet.com/bgppolicy.php)

Interpretation of BGP Communities

```
remarks:        15412:1514 Amsterdam
remarks:        ------------------------------------------------------
remarks:        15412:7xx Customer
remarks:        15412:701 Aggregate
remarks:        15412:702 Statically Routed
remarks:        15412:703 BGP Routed
remarks:        15412:705 BGP Routed (Suppress MED to upstreams)
remarks:        ------------------------------------------------------
remarks:        15412:8xx Peer
remarks:        15412:800 PRIVATE PEER
remarks:        15412:801 PAIX
remarks:        15412:802 NYIIX
remarks:        15412:803 JPIX
remarks:        15412:804 KINX
remarks:        15412:805 HKIX
remarks:        15412:806 LINX
remarks:        15412:807 SFINX
remarks:        15412:808 LAIX
remarks:        15412:809 AMSIX
remarks:        15412:810 DECIX
remarks:        15412:813 JPNAP
remarks:        15412:814 EQUINIX ASHBURN VA
remarks:        15412:815 EQUINIX SINGAPORE
remarks:        15412:816 EQUINIX TOKYO
remarks:        15412:817 ANY2
remarks:        15412:820 EQUINIX PARIS
remarks:        15412:821 EQUINIX HONG KONG
remarks:        ------------------------------------------------------
remarks:        15412:9xx Upstream
remarks:        15412:902 LEVEL3 AS3356
remarks:        15412:903 NTT/VERIO AS2914
remarks:        ------------------------------------------------------
remarks:        BGP Communities available to customers for traffic engineering
remarks:        ------------------------------------------------------
remarks:        Modify LocalPref
remarks:
remarks:        15412:80  = 80
remarks:        15412:200 = 200 (e.g. backup link)
remarks:        15412:300 = 300
remarks:        Default (Customer/Transit/Peer) = 250/100/100
remarks:        ------------------------------------------------------
remarks:        Suppression/Prepend
remarks:        ------------------------------------------------------
remarks:        15412:4100 Do not announce to any upstream
remarks:        ------------------------------------------------------
remarks:        15412:4120 Do not announce to LEVEL3 AS3356
```

Internet Routing Registries (e.g. whois -h whois.radb.net AS15412)

Data Collection Architecture



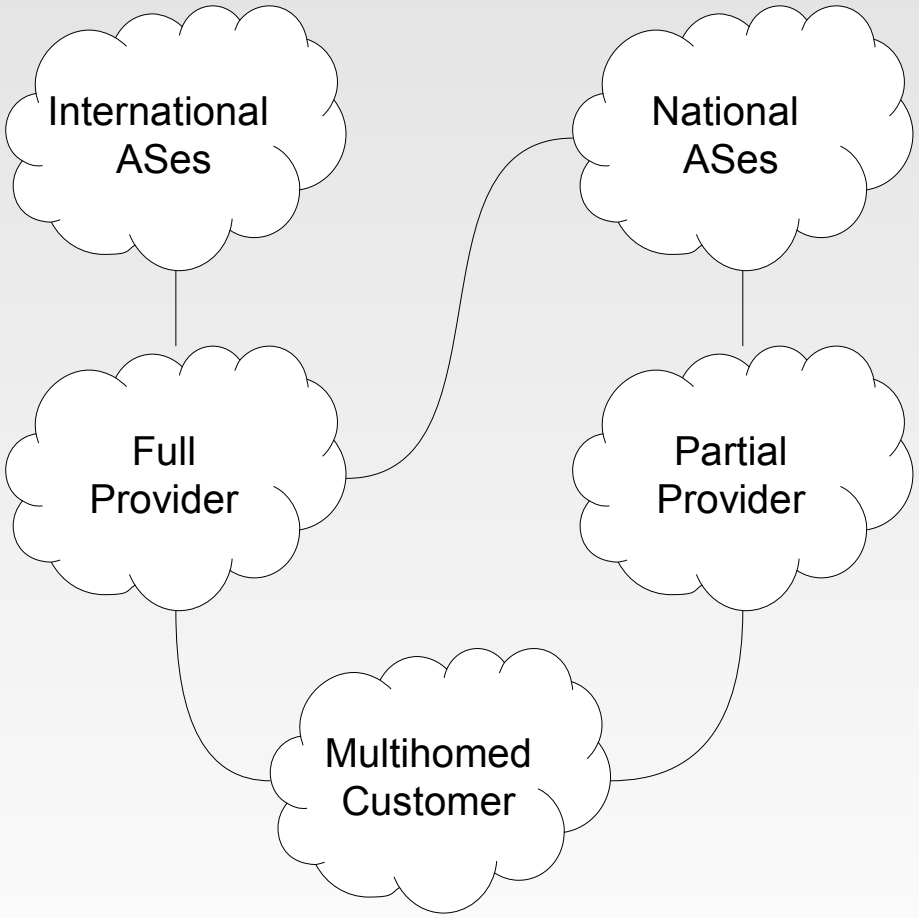
Results (February 2011)

Total number of observed links	109,807
Number of inferred relationships	38,704 (35%)
c2p links	23,012
p2p links	15,375
s2s links	174

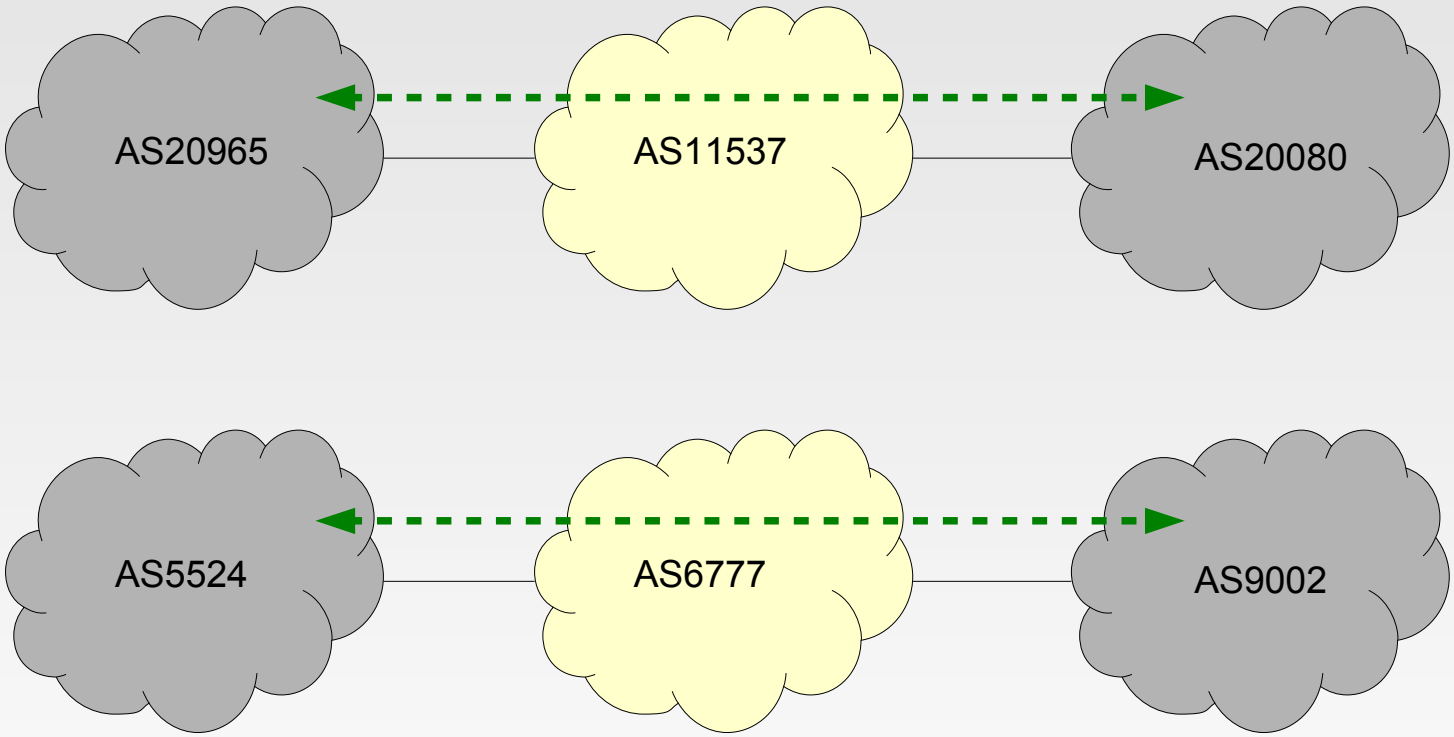
“Special” Relationship Types

- Relationships not described by the c2p, p2p, s2s model
- Little attention, difficult to detect
  - Partial transit: 1,828
  - Indirect peering: 811
  - Hybrid relationships: 1,034

Partial Transit



Indirect Peering



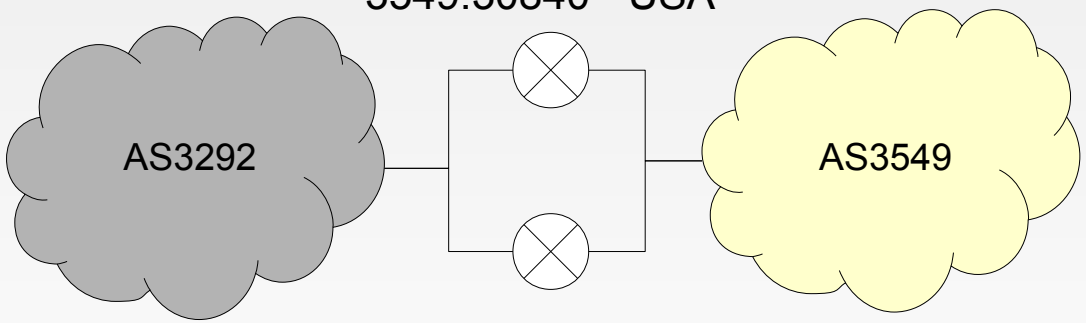
Educational/
Research
Networks
(e.g. Internet2)

Public peering
At IXPs
(e.g. AMS-IX)

Hybrid Links

IP-version depended

3548:4354 – customer
3549:30840 - USA



Location depended

3548:2771 – peer
3549:31208 - Denmark

IPv6 Relationships

- 7,618 **AS links** carry both IPv4 and IPv6 traffic
  - 13% of these have different relationship between IPv4 and IPv6
- 47% of the IPv6 **AS paths** contain at least one hybrid AS link
- 10% of the IPv6 **AS paths** are non valley-free
  - Same during IPv6 day

Conclusions

- Unexploited wealth of BGP attribute data
- Complex relationship types widely disregarded become increasingly popular
- IPv6 relationships should be studied separately

Conclusions & Future Work

- Extend the interpretation of Communities values
- Extend to more AS links
- Use traceroute data to verify/evaluate inferences
- Performance impact on IPv6

THANK YOU!