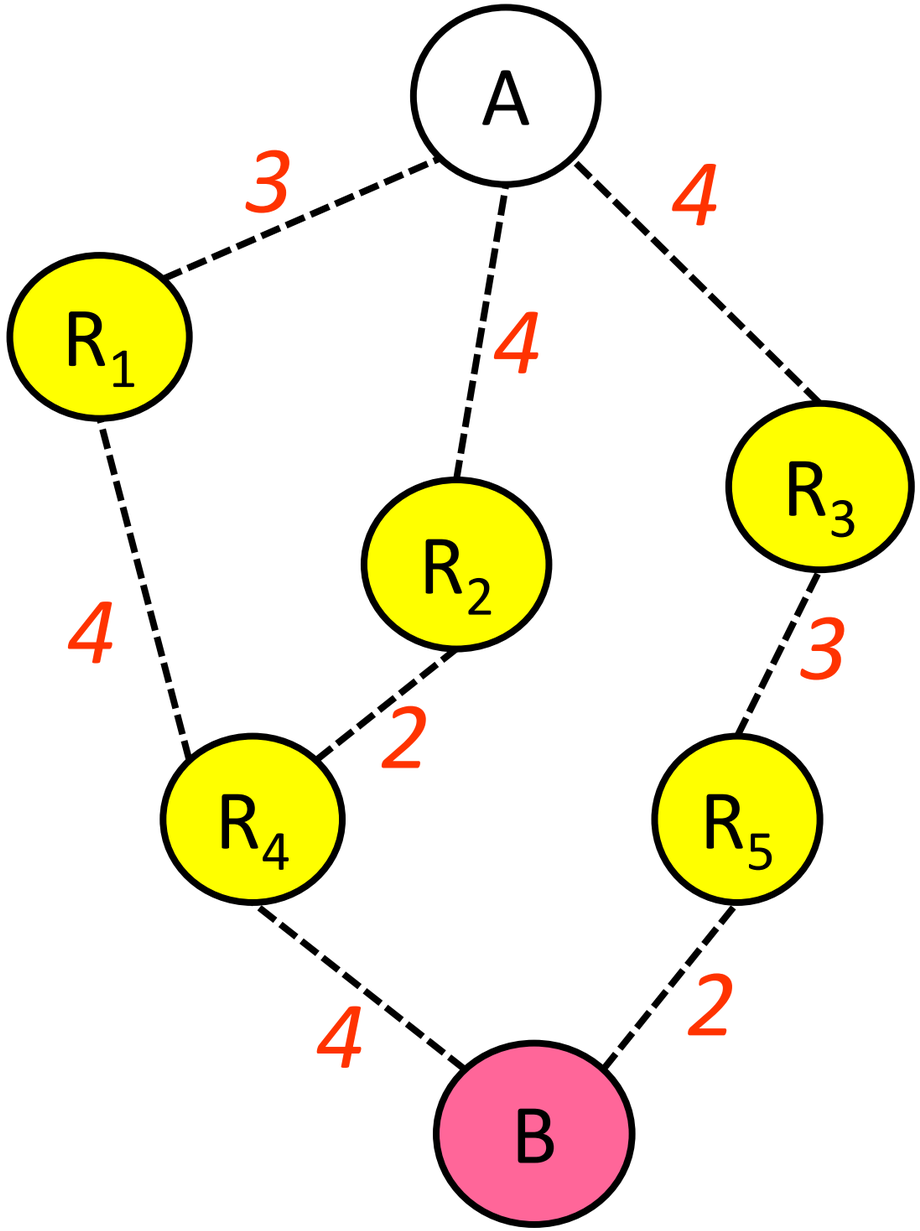
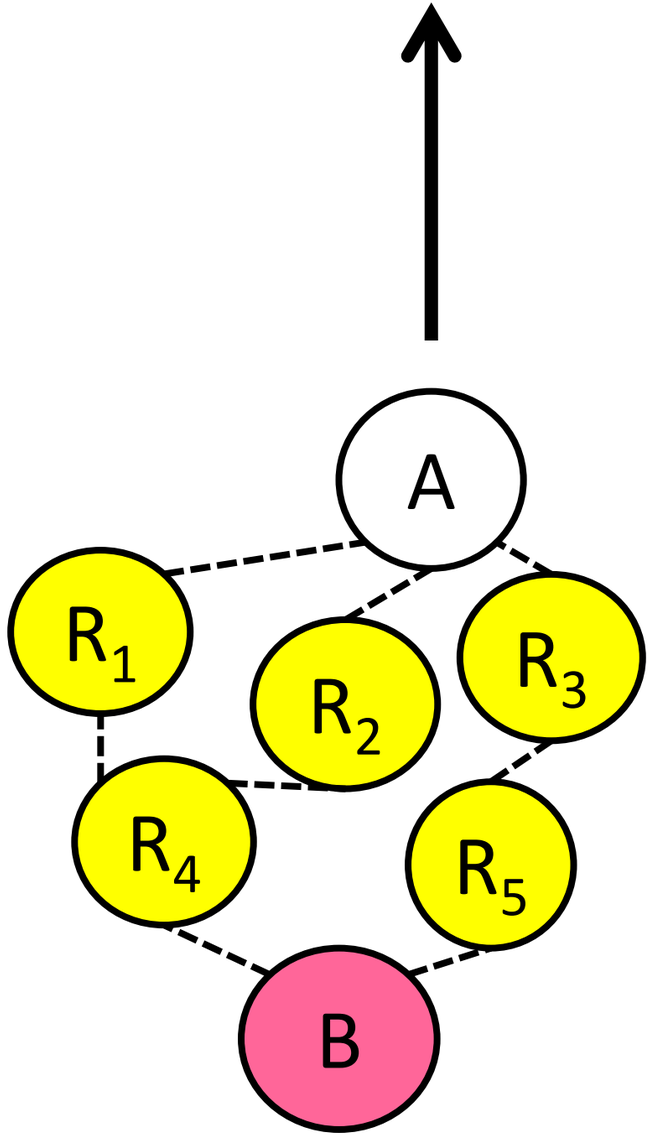
CS144
An Introduction to Computer Networks

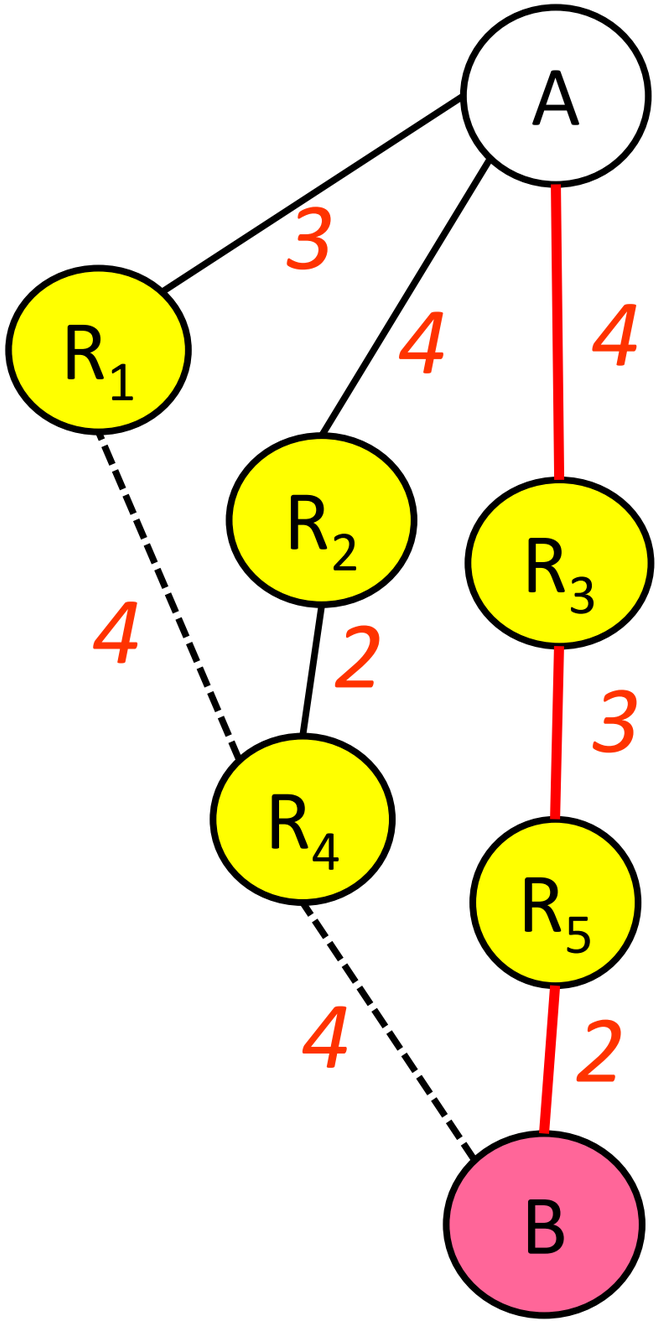
Routing – Lecture 2

Nick McKeown
Professor of Electrical Engineering
and Computer Science, Stanford University

Another view of Dijkstra…

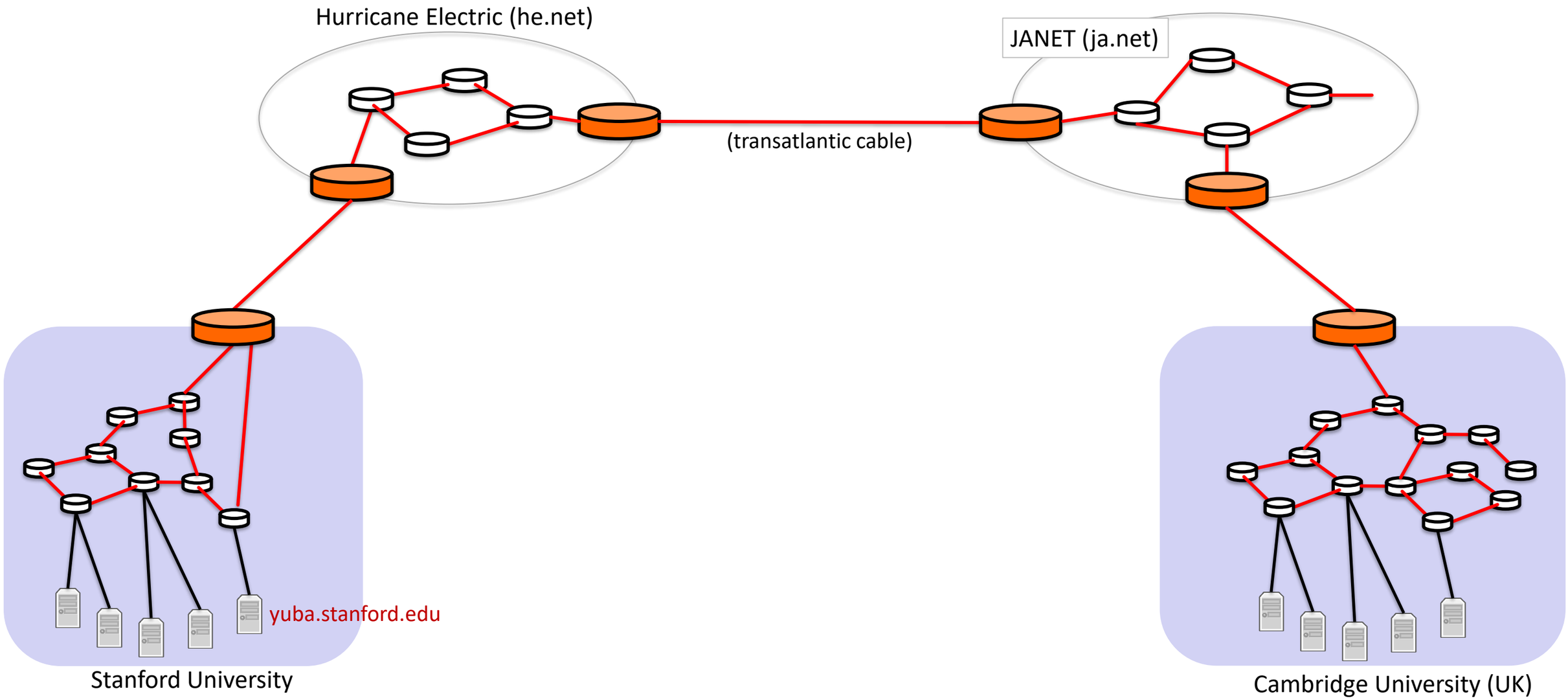

A
3
4
R1
4
R3
R2
4
3
2
R4
R5
4
2
B


A
R_1
R_2
R_3
R_4
R_5
B

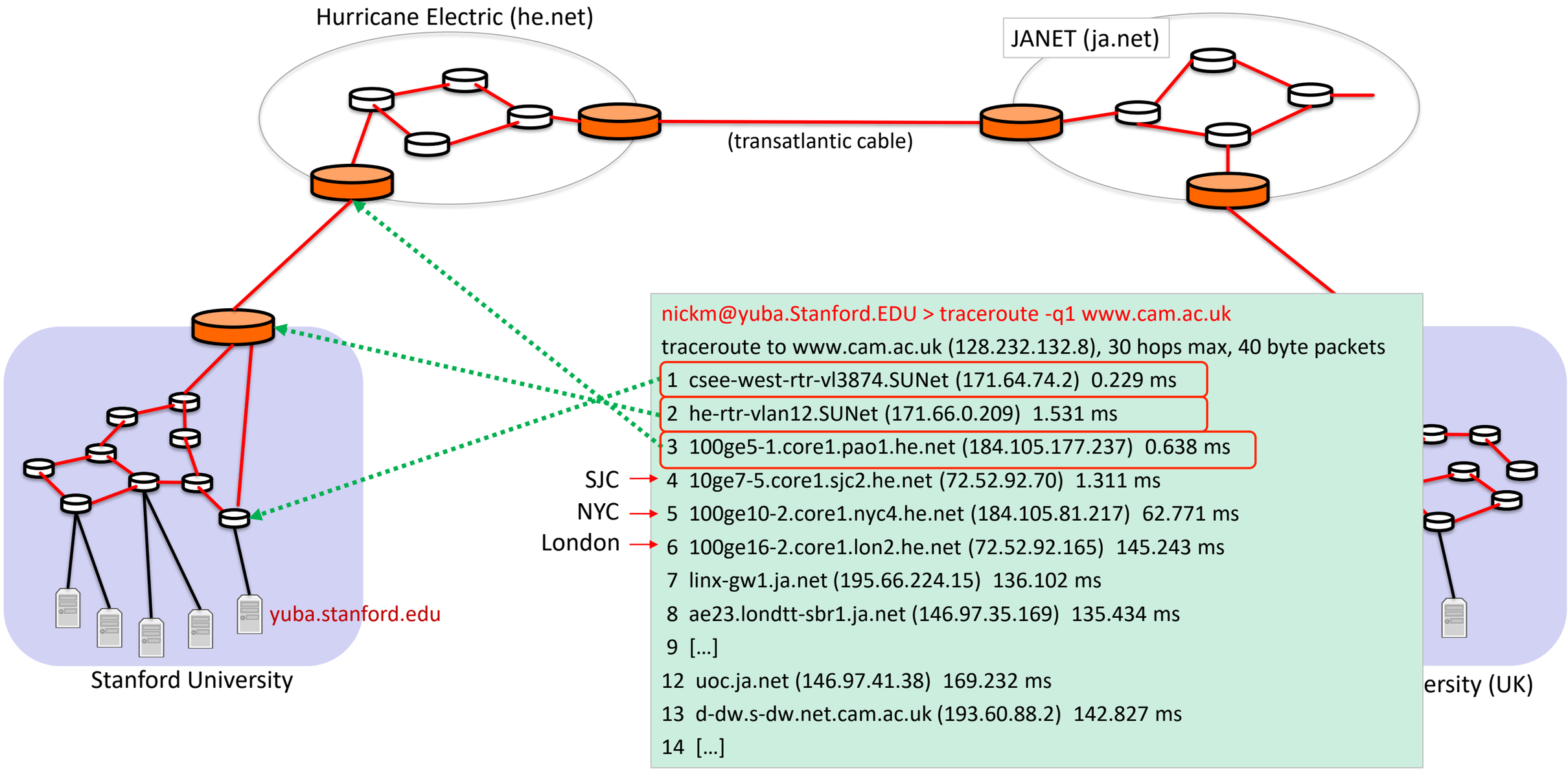

A
3
R1
4
4
R2
R3
4
2
3
R4
R5
4
2
B

Internet routing is hierarchical

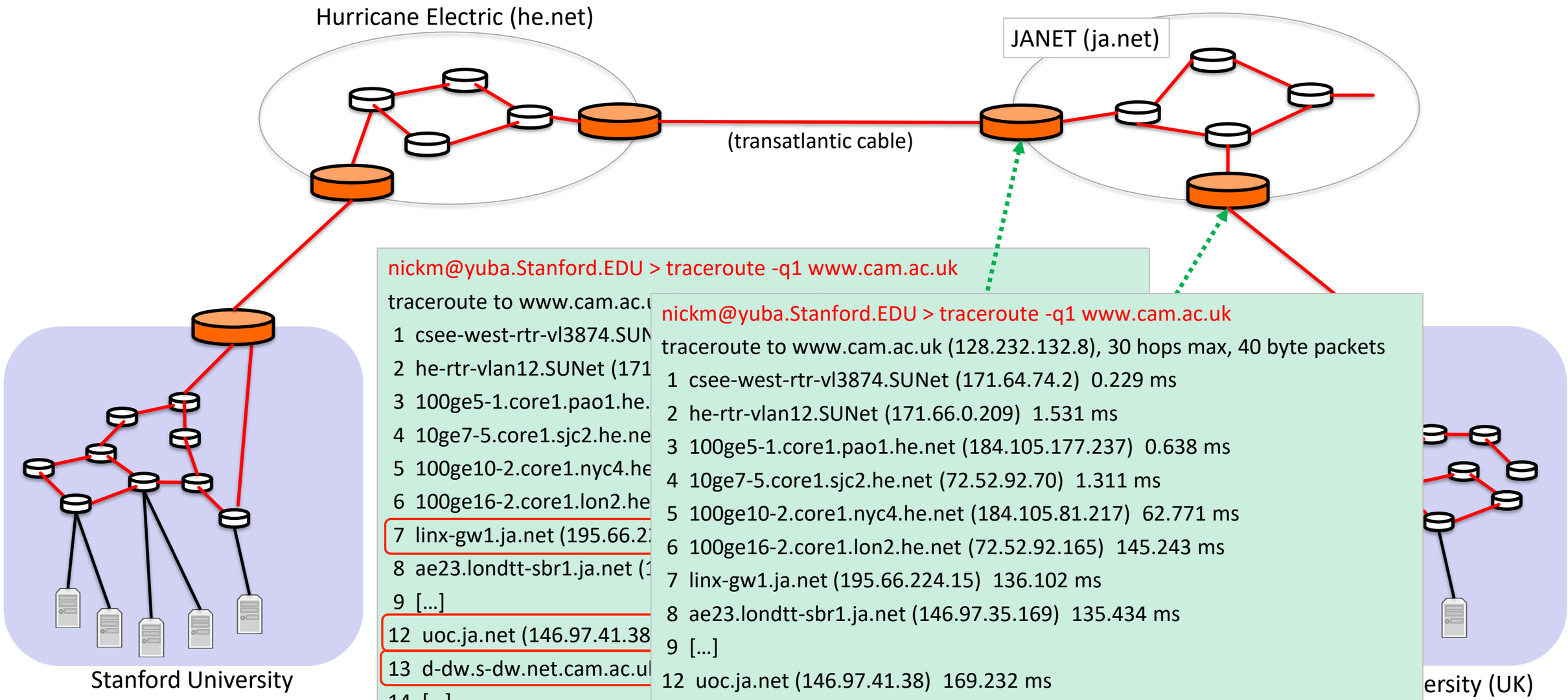
In the Internet, Autonomous Systems (AS's) have Border Routers (orange).
The border routers route packets to each other using the Border Gateway Protocol.



In the Internet, Autonomous Systems (AS's) have Border Routers (orange).
The border routers route packets to each other using the Border Gateway Protocol.



In the Internet, Autonomous Systems (AS's) have Border Routers (orange).
The border routers route packets to each other using the Border Gateway Protocol.




```
nickm@yuba.Stanford.EDU > traceroute -q1 www.cam.ac.uk
traceroute to www.cam.ac.u
 1  csee-west-rtr-vl3874.SUN
 2  he-rtr-vlan12.SUNet (171
 3  100ge5-1.core1.pao1.he.
 4  10ge7-5.core1.sjc2.he.ne
 5  100ge10-2.core1.nyc4.he
 6  100ge16-2.core1.lon2.he
 7  linx-gw1.ja.net (195.66.2
 8  ae23.londtt-sbr1.ja.net (1
 9  […]
12  uoc.ja.net (146.97.41.38
13  d-dw.s-dw.net.cam.ac.ul
14  […]
```

```
nickm@yuba.Stanford.EDU > traceroute -q1 www.cam.ac.uk
traceroute to www.cam.ac.uk (128.232.132.8), 30 hops max, 40 byte packets
 1  csee-west-rtr-vl3874.SUNet (171.64.74.2)  0.229 ms
 2  he-rtr-vlan12.SUNet (171.66.0.209)  1.531 ms
 3  100ge5-1.core1.pao1.he.net (184.105.177.237)  0.638 ms
 4  10ge7-5.core1.sjc2.he.net (72.52.92.70)  1.311 ms
 5  100ge10-2.core1.nyc4.he.net (184.105.81.217)  62.771 ms
 6  100ge16-2.core1.lon2.he.net (72.52.92.165)  145.243 ms
 7  linx-gw1.ja.net (195.66.224.15)  136.102 ms
 8  ae23.londtt-sbr1.ja.net (146.97.35.169)  135.434 ms
 9  […]
12  uoc.ja.net (146.97.41.38)  169.232 ms
13  d-dw.s-dw.net.cam.ac.uk (193.60.88.2)  142.827 ms
14  […]
```

AS (Autonomous System) numbers

```
nickm@yuba.Stanford.EDU > traceroute -q1 www.cam.ac.uk
traceroute to www.cam.ac.uk (128.232.132.8), 30 hops max, 40 byte packets
 1  csee-west-rtr-vl3874.SUNet (171.64.74.2)  0.229 ms
 2  he-rtr-vlan12.SUNet (171.66.0.209)  1.531 ms
 3  100ge5-1.core1.pao1.he.net (184.105.177.237)  0.638 ms
 4  10ge7-5.core1.sjc2.he.net (72.52.92.70)  1.311 ms
 5  100ge10-2.core1.nyc4.he.net (184.105.81.217)  62.771 ms
 6  100ge16-2.core1.lon2.he.net (72.52.92.165)  145.243 ms
 7  linx-gw1.ja.net (195.66.224.15)  136.102 ms
 8  ae23.londtt-sbr1.ja.net (146.97.35.169)  135.434 ms
 9  […]
12  uoc.ja.net (146.97.41.38)  169.232 ms
13  d-dw.s-dw.net.cam.ac.uk (193.60.88.2)  142.827 ms
14  […]
```

```
nickm> whois -h whois.cymru.com 146.97.35.169

AS      | IP               | AS Name
786     | 146.97.35.169    | JANET Jisc Services
                             Limited, GB
```

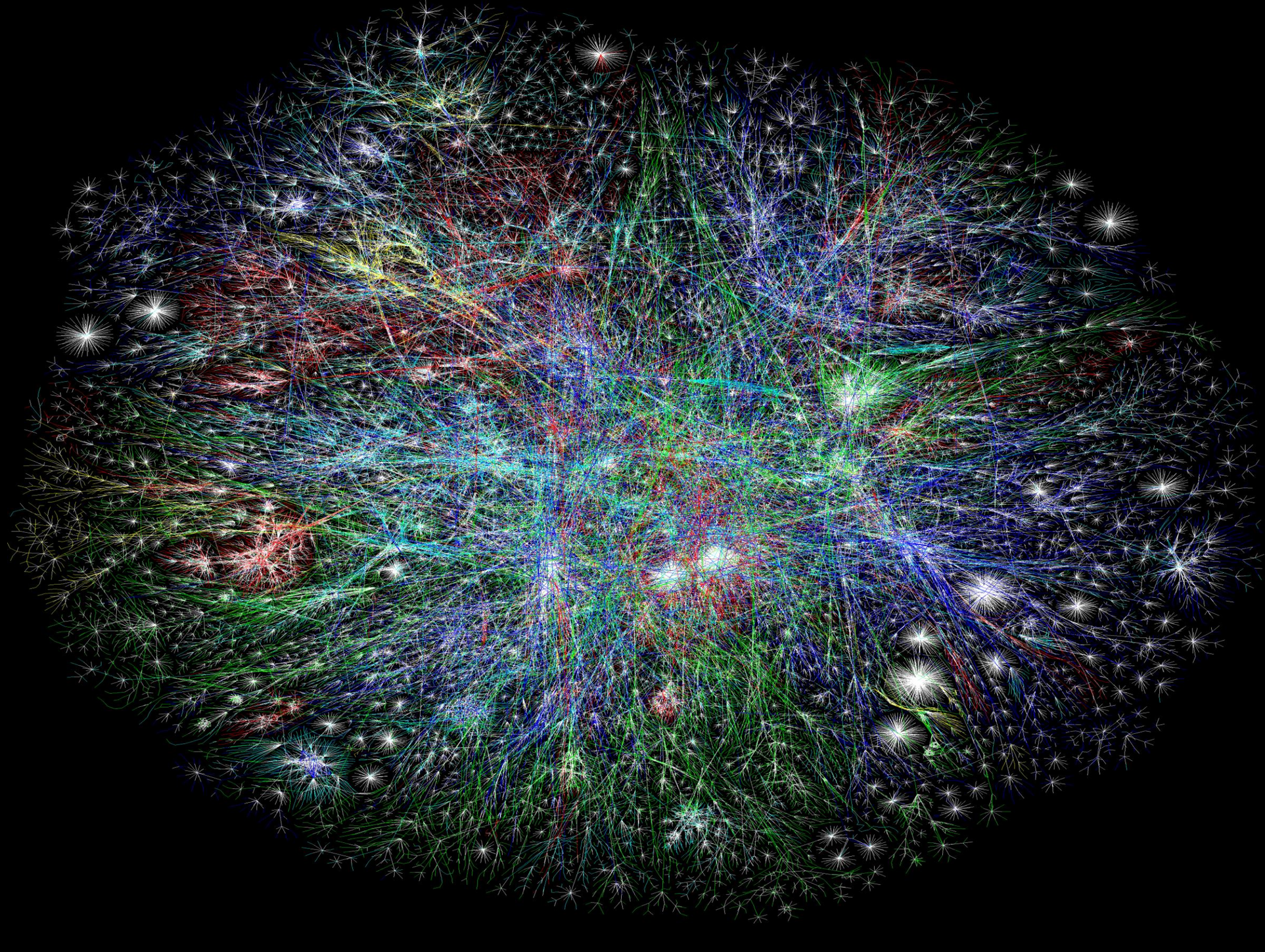
JANET is AS 786

e.g. yuba.Stanford.edu

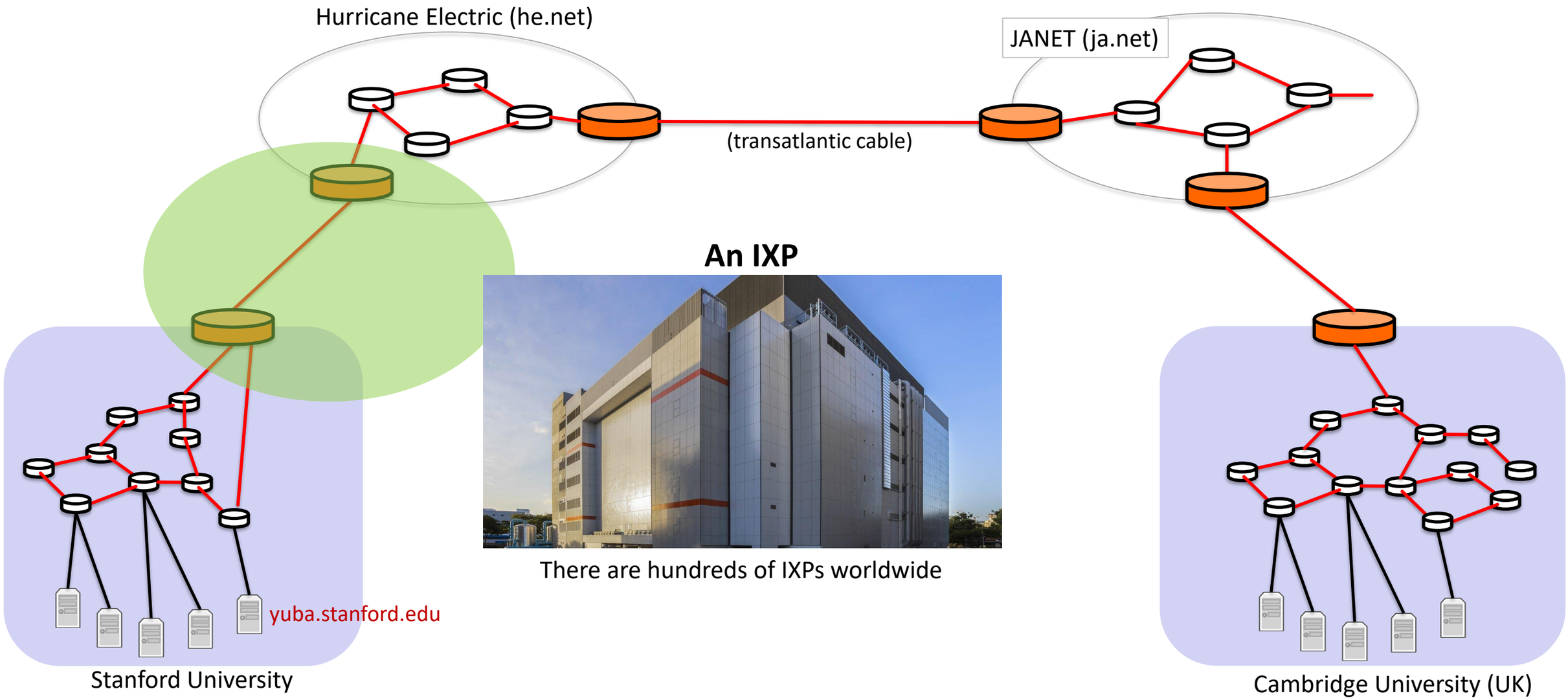
```
nickm> whois -h whois.cymru.com 171.64.74.155

AS      | IP               | AS Name
32      | 171.64.74.155    | STANFORD, US
```

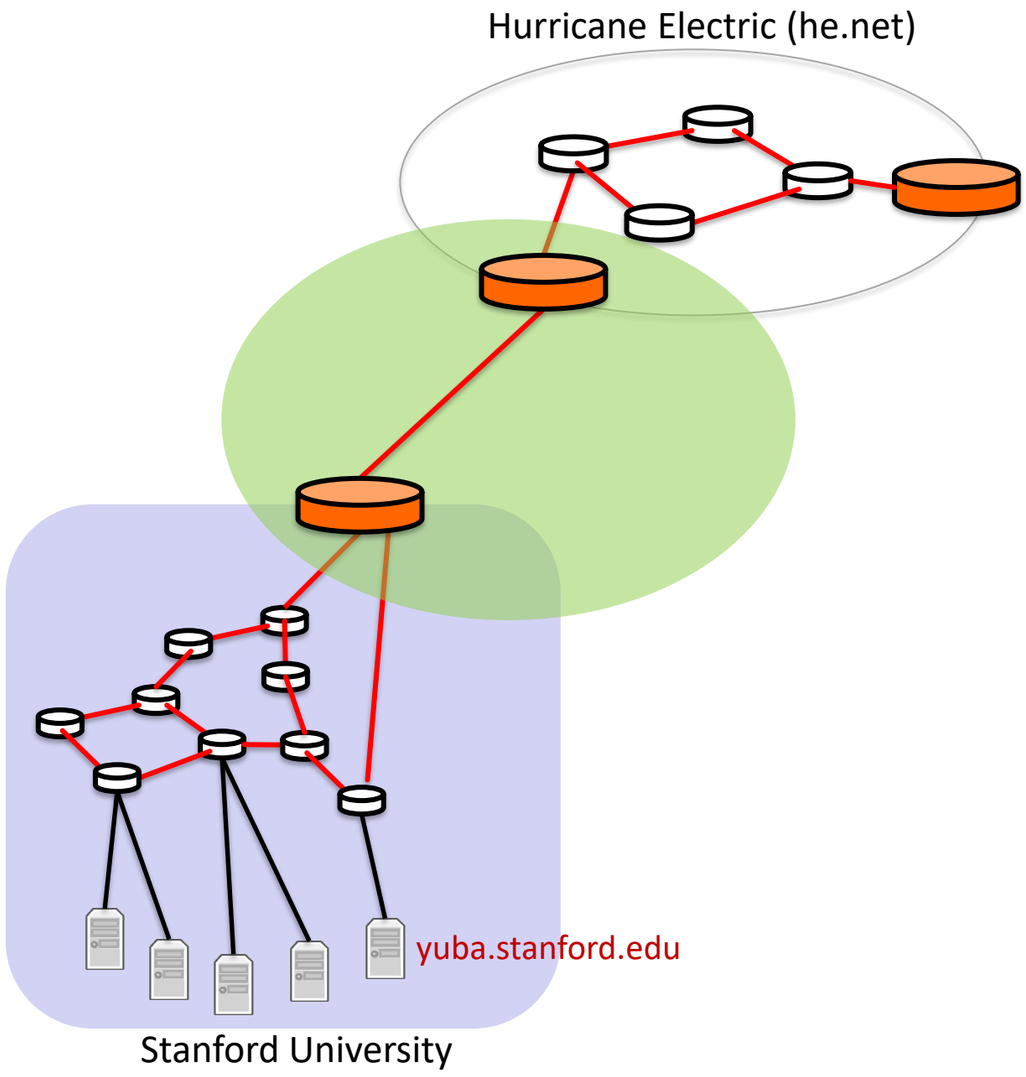
Stanford is AS 32

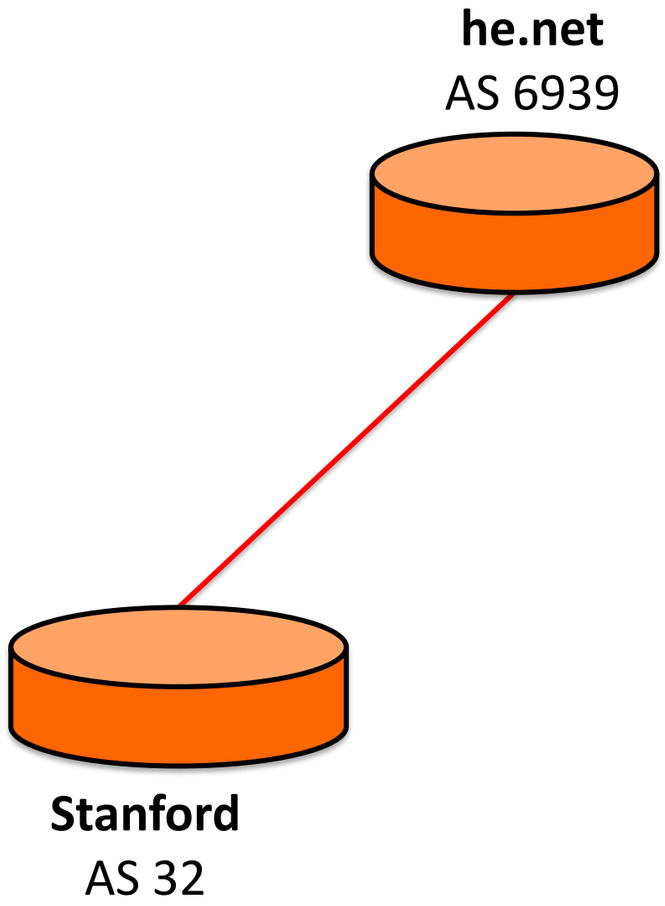


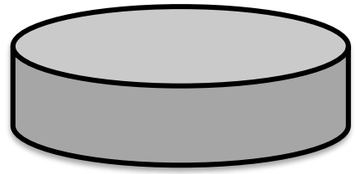
Autonomous Systems (AS's) usually connect to each other in an Internet eXchange Point (IXP)

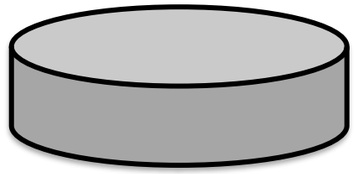


Autonomous Systems (AS's) usually connect to each other in an Internet eXchange Point (IXP)



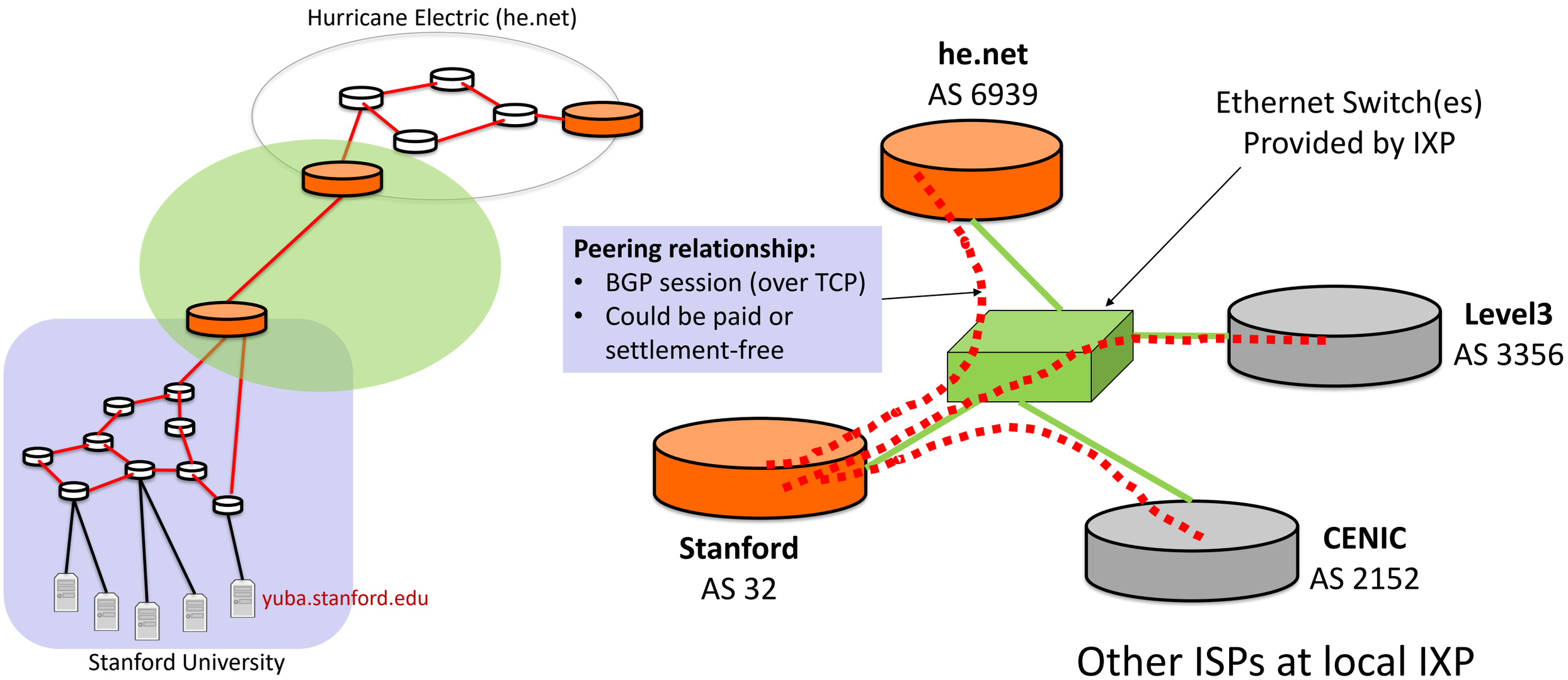






Other ISPs at local IXP

Autonomous Systems (AS's) usually connect to each other in an Internet eXchange Point (IXP)



Border Gateway Protocol (BGP)

BGP routers advertise routes to their neighbors, containing:

- A prefix
- The list of AS's indicating the path the packet will take to reach the prefix

Example of path advertisement:

"The network 171.64/16 can be reached via the path {AS1, AS5, AS13}"

Q: Why advertise a path of AS's for each prefix, rather than

a. The next hop for each prefix

b. The path of IP addresses

Border Gateway Protocol (BGP)

"The network 171.64/16 can be reached via the path {AS1, AS5, AS13}"

Paths with loops are detected locally and ignored.

A BGP router may connect to several peers and receive multiple different advertised paths for the same prefix.

Local policies chosen by the AS administrator pick the preferred path.

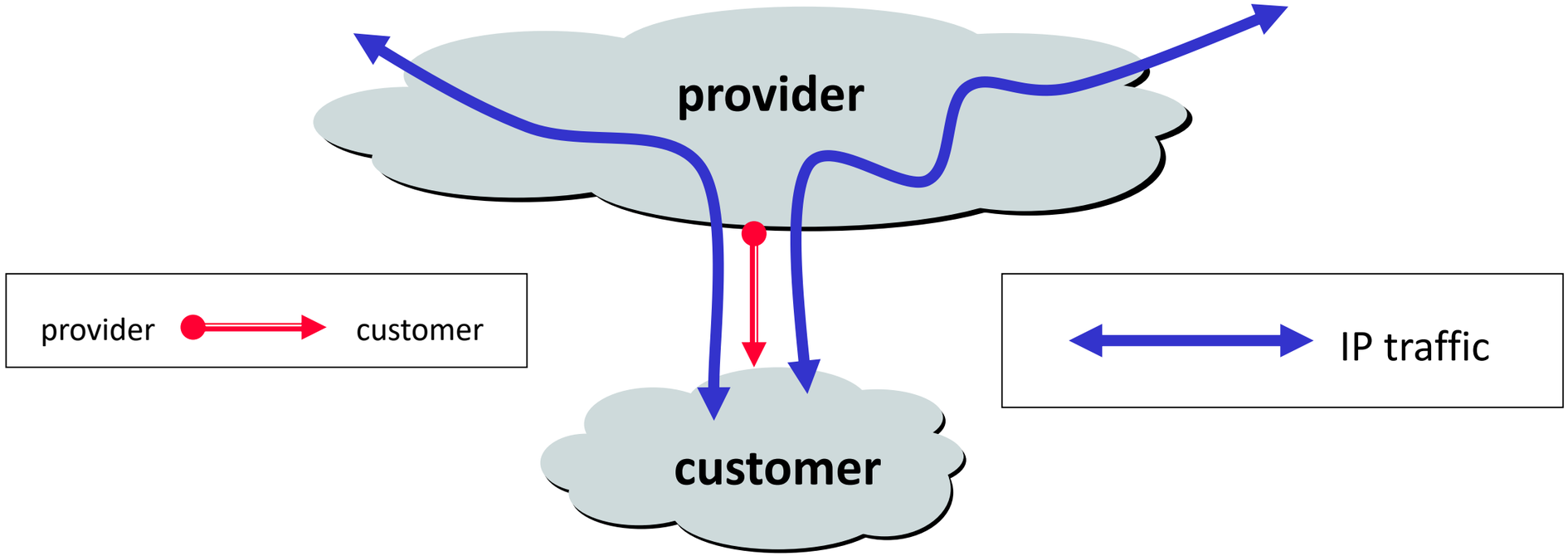
Border Gateway Protocol (BGP): Details

- BGP neighbors (“peers”) establish a TCP connection.
- The TCP connection is manually configured at both ends.
- Neighbors send “keepalive” messages every 60 seconds.

- BGP is sometimes called a “Path vector” algorithm.
- It is not a link-state or a distance-vector routing protocol.

- When an advertised path changes, the path vector is first “withdrawn”, then the new one is advertised.

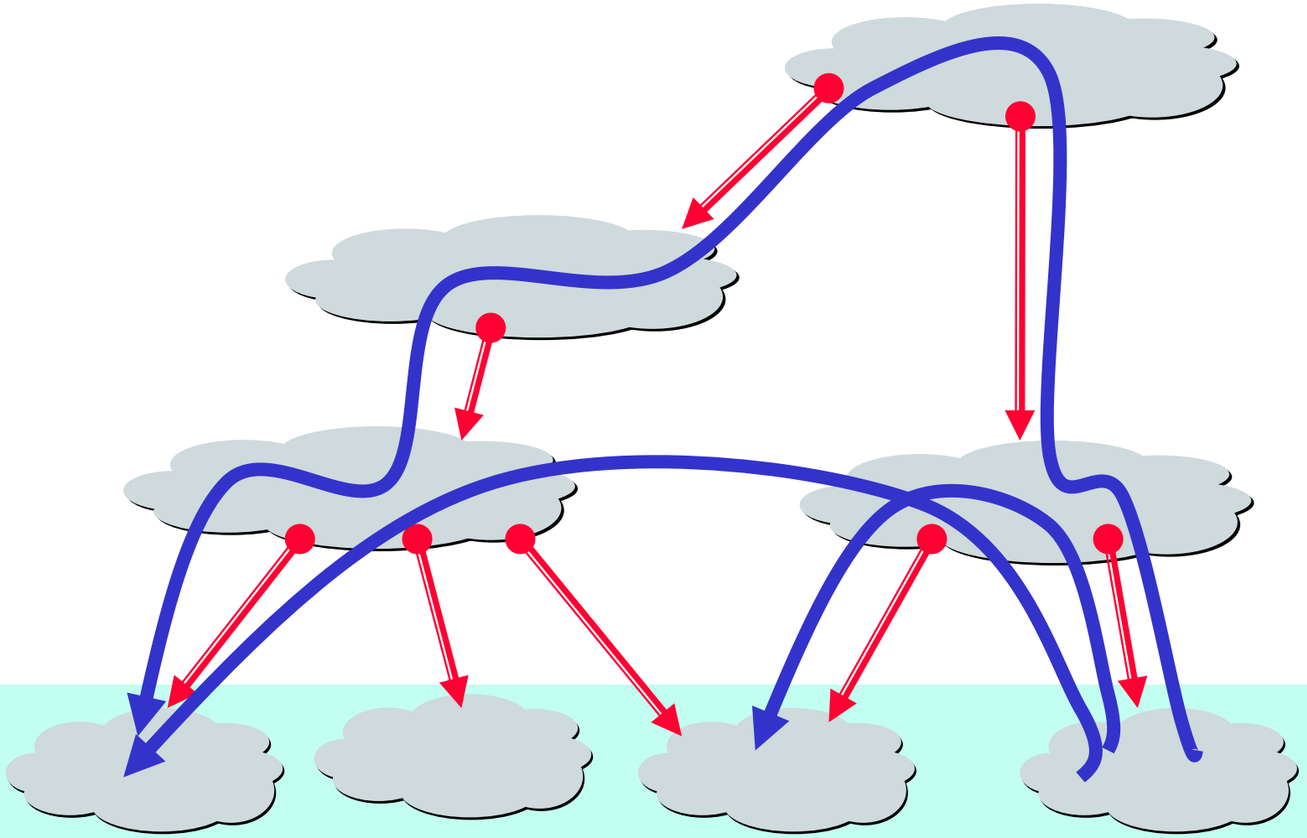
Customers and Providers



Customers pay providers to carry their packets.

Customer-Provider Hierarchy

Transit AS's



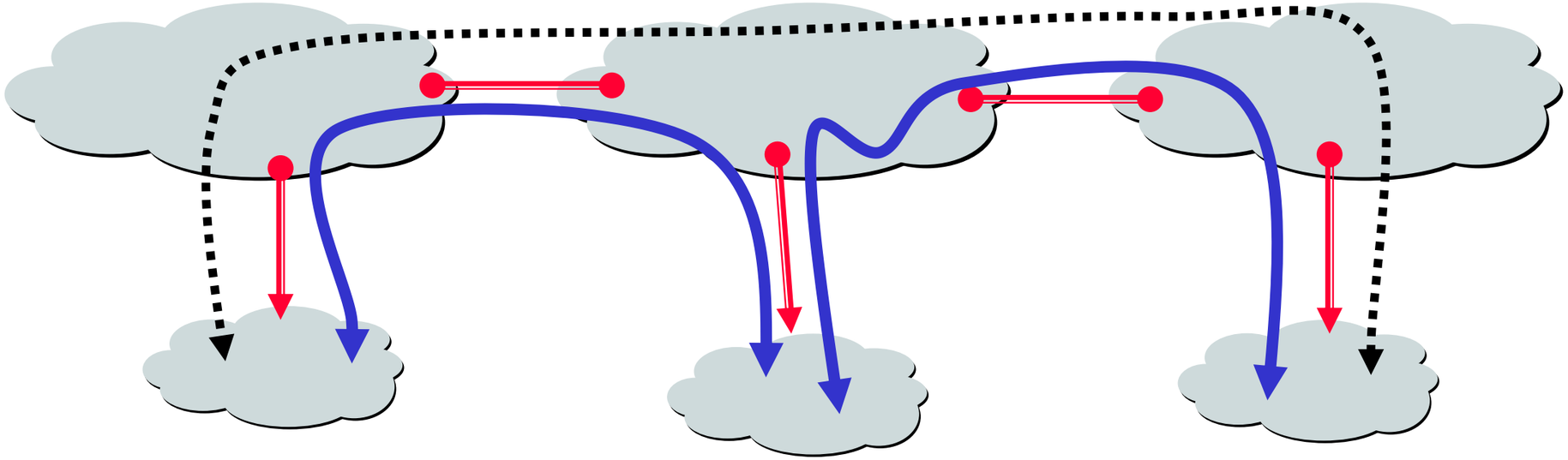
Stub AS's

Routers inside Stub AS's can use a "default route" for unknown IP addresses. The default route is the Border router.

provider → customer

← IP traffic

The Peering Relationship



peer — peer

provider → customer

traffic allowed

Transit traffic NOT allowed

Peers provide transit between their respective customers

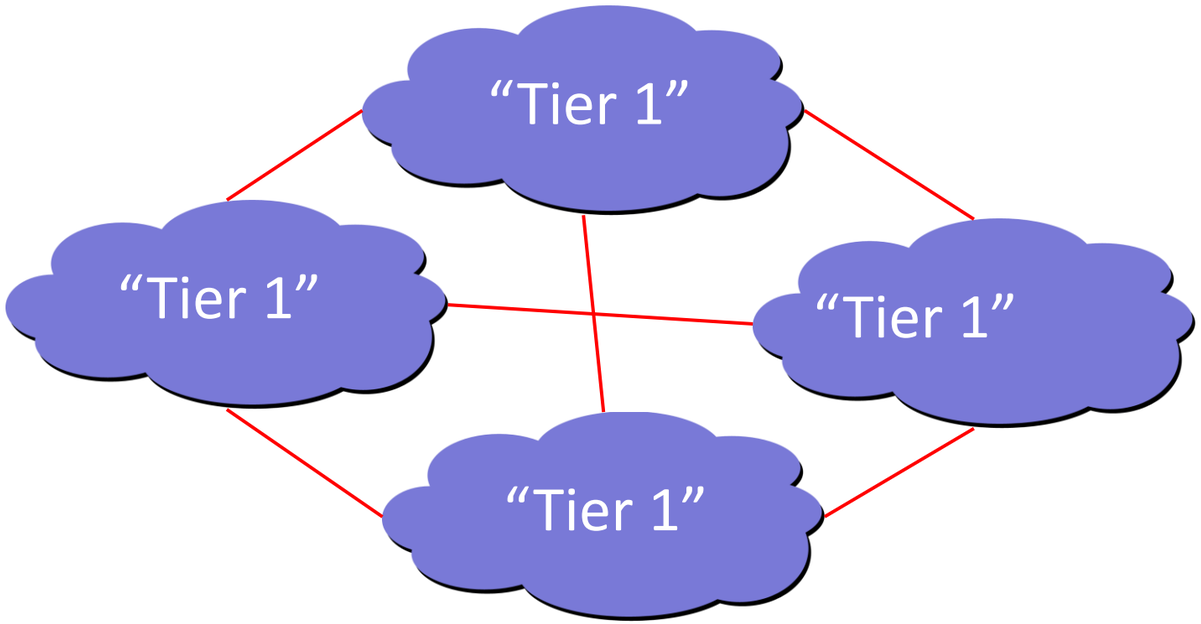
Peers do not provide transit between peers

Peers (typically) do not exchange $$$

So how does traffic from the left side reach the right side?

“Tier 1” Providers

A **Tier 1** network is a transit-free network that peers with every other tier 1 network



Tier 1 ISPs

Definition: A **Tier 1 ISP** has access to the entire ***Internet Region*** solely via its free and reciprocal peering agreements.

Definition: An **Internet Region** is a portion of the Internet network typically bounded by a country's geographical boundaries.

Each Internet Region has its own set of "**Tier 1 ISPs**."

The litmus test:

"*Does an ISP pay anyone to reach any destination in the Internet Region*?"
If the answer is "No" then it is a **Tier 1 ISP**, and
If the answer is "Yes" then it is a **Tier 2 ISP**.

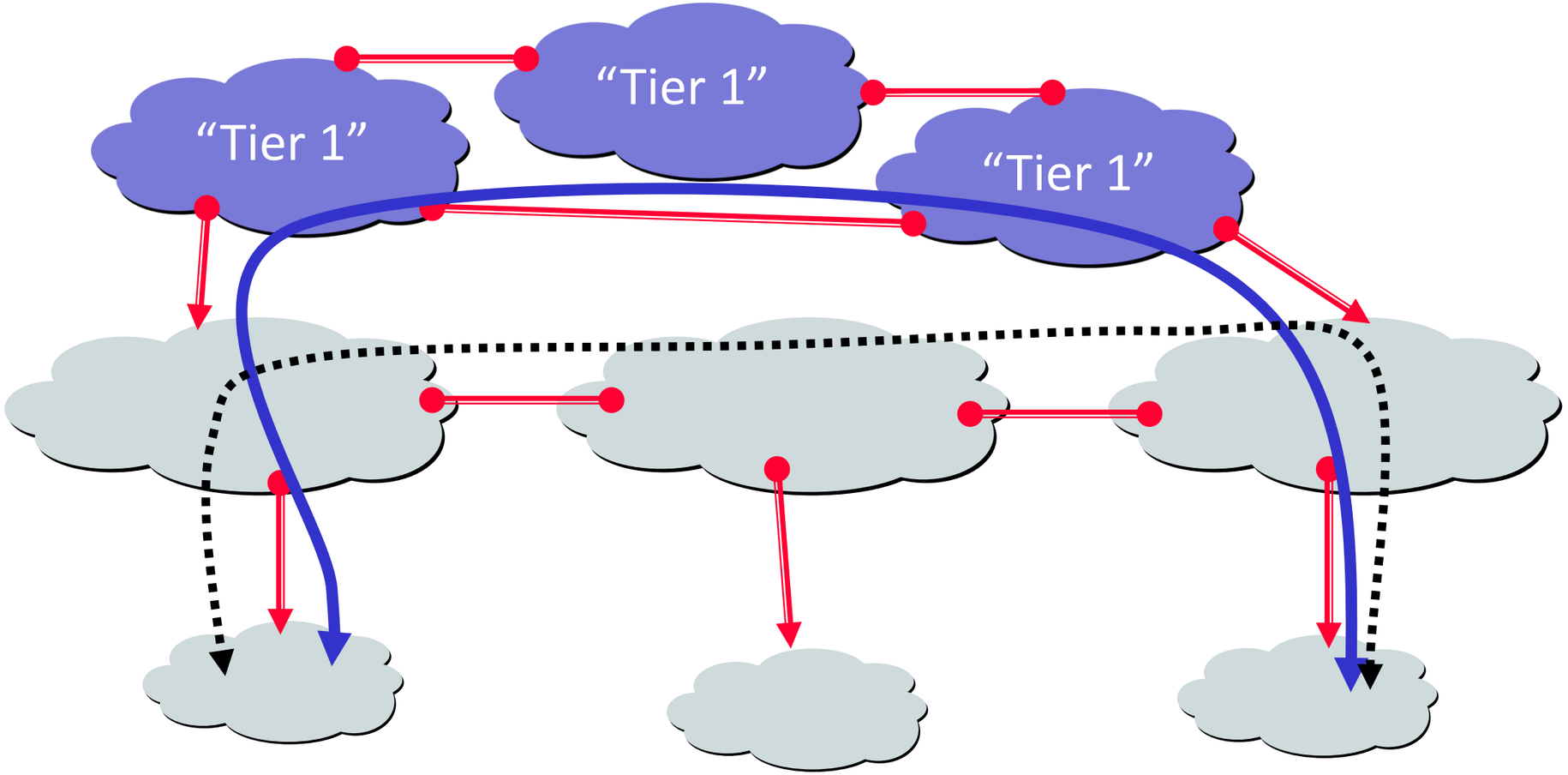
Tier 1 ISPs by country

The U.S. Internet Region Tier 1 ISPs

1. AT&T
2. Verizon
3. Sprint (Softbank Broadband)
4. Century Link (Qwest)
5. Level 3 (with Global Crossing now)
6. NTT/Verio
7. Cogent

The Japan Internet Region Tier 1 ISPs

1. NTT
2. Japan Telecom (Softbank)
3. KDDI
4. IIJ
5. Powered.com



peer — peer

provider — customer

traffic allowed

Transit traffic NOT allowed

Peers provide transit between their respective customers

Peers do not provide transit between peers

Peers (typically) do not exchange $$$