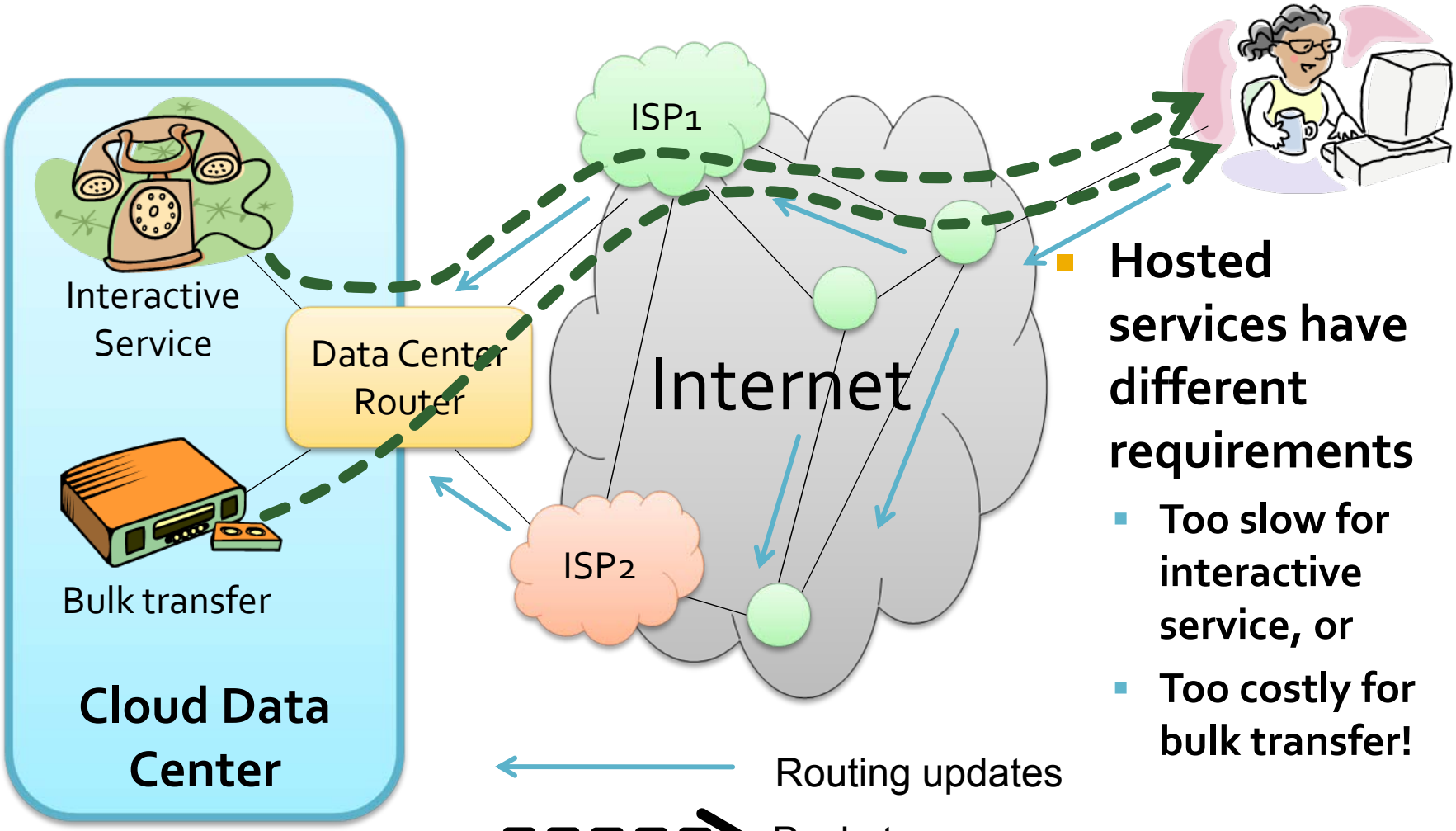
Vytautas Valancius, Nick Feamster, Akihiro Nakao, and Jennifer Rexford

Wide-Area Route Control for Distributed Services

Cloud Computing

- Cloud computing is on the rise
- Provides computing resources and storage in cloud data centers
- Hosting on the steroids for Internet services

Accessing Services in the Cloud



- **Hosted services have different requirements**
  - **Too slow for interactive service, or**
  - **Too costly for bulk transfer!**

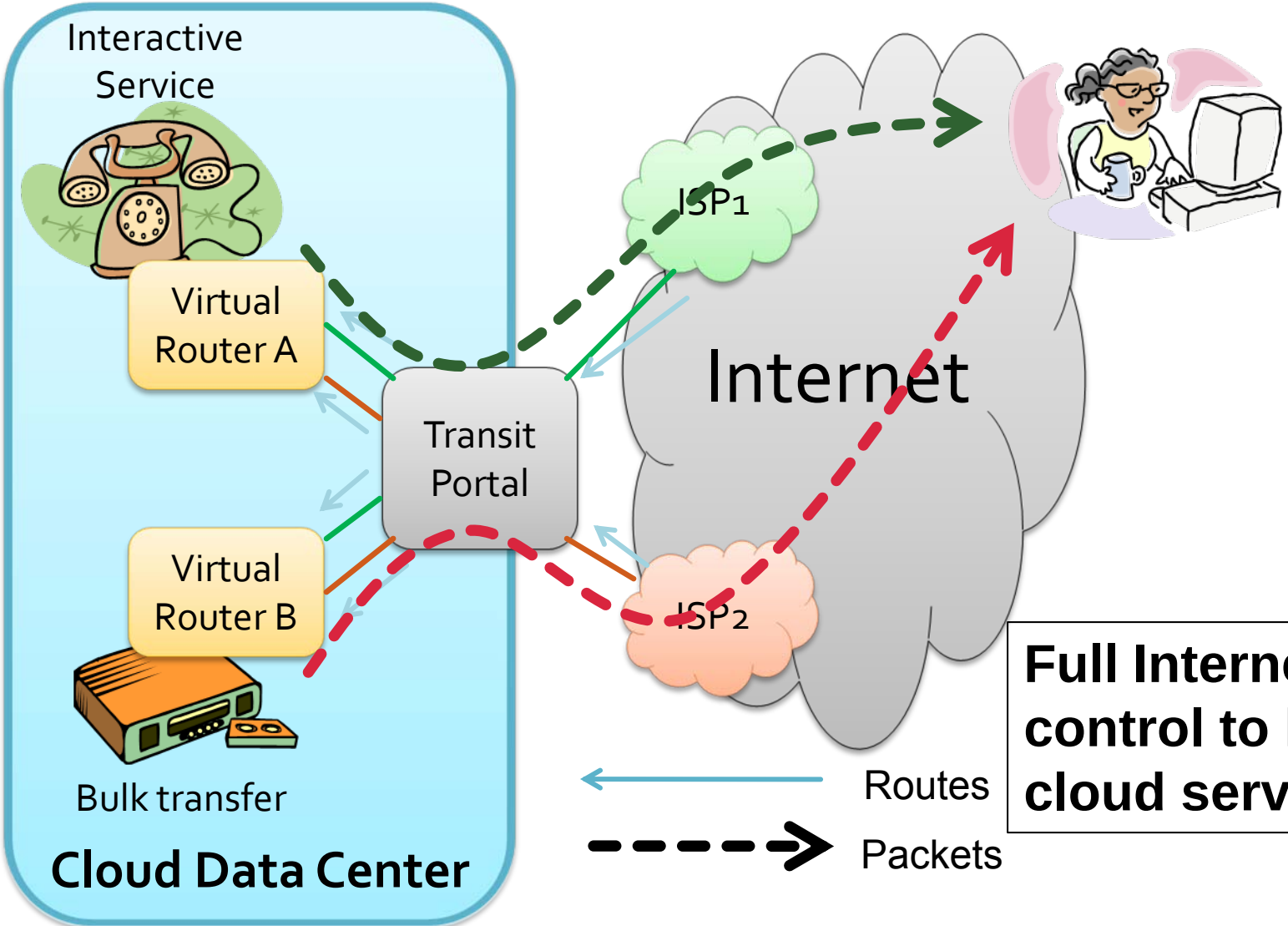
Cloud Routing Today

- Multiple upstream ISPs
  - Amazon EC2 has at least 58 routing peers in Virginia data center
- Data center router picks ***one*** route to a destination for all hosted services
  - Packets from all hosted applications use the same path

Route Control: "Cloudless" Solution

- Obtain connectivity to upstream ISPs
  - Physical connectivity
  - Contracts and routing sessions
- Obtain the Internet numbered resources from authorities
- **Expensive and time-consuming!**

Routing with Transit Portal (TP)



Full Internet route control to hosted cloud services!

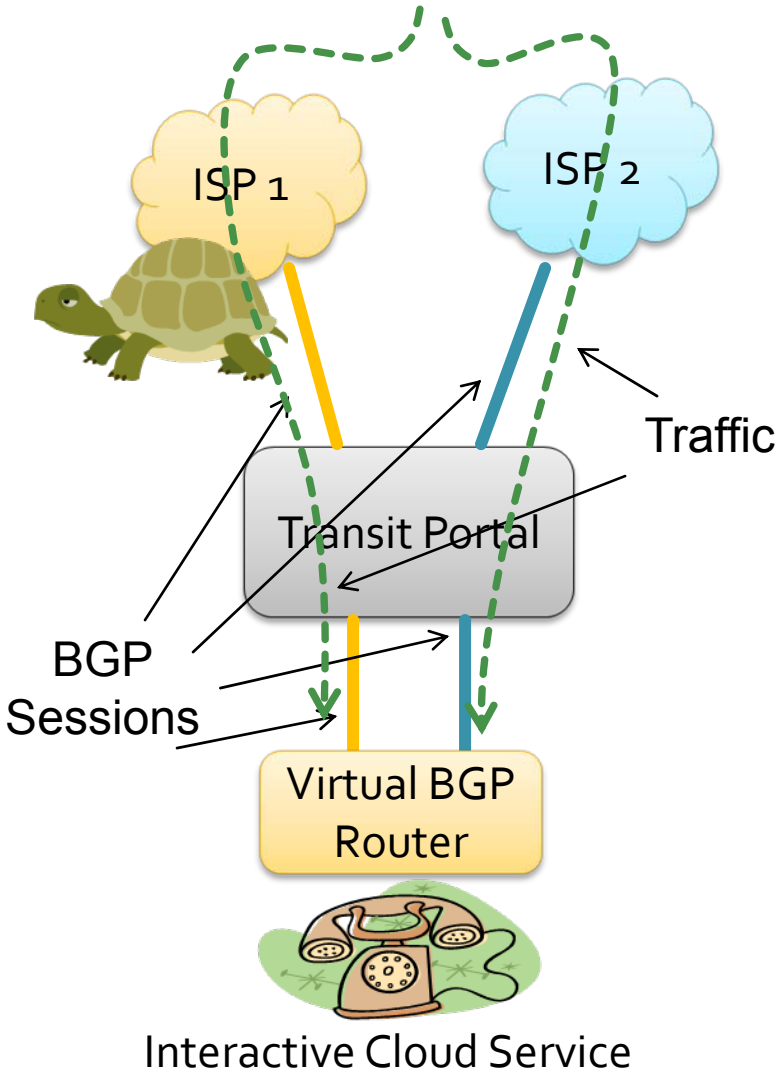
Outline

- Motivation and Overview
- Connecting to the Transit Portal
- Advanced Transit Portal Applications
- Scaling the Transit Portal
- Future Work & Summary

Connecting to the TP

- Separate Internet router for each service
  - Virtual or physical routers
- Links between service router and TP
  - Each link emulates connection to upstream ISP
- Routing sessions to upstream ISPs
  - TP exposes standard BGP route control interface

Basic Internet Routing with TP



- Cloud client with two upstream ISPs
  - ISP 1 is preferred
- ISP 1 exhibits excessive jitter
- Cloud client reroutes through ISP 2

Current TP Deployment

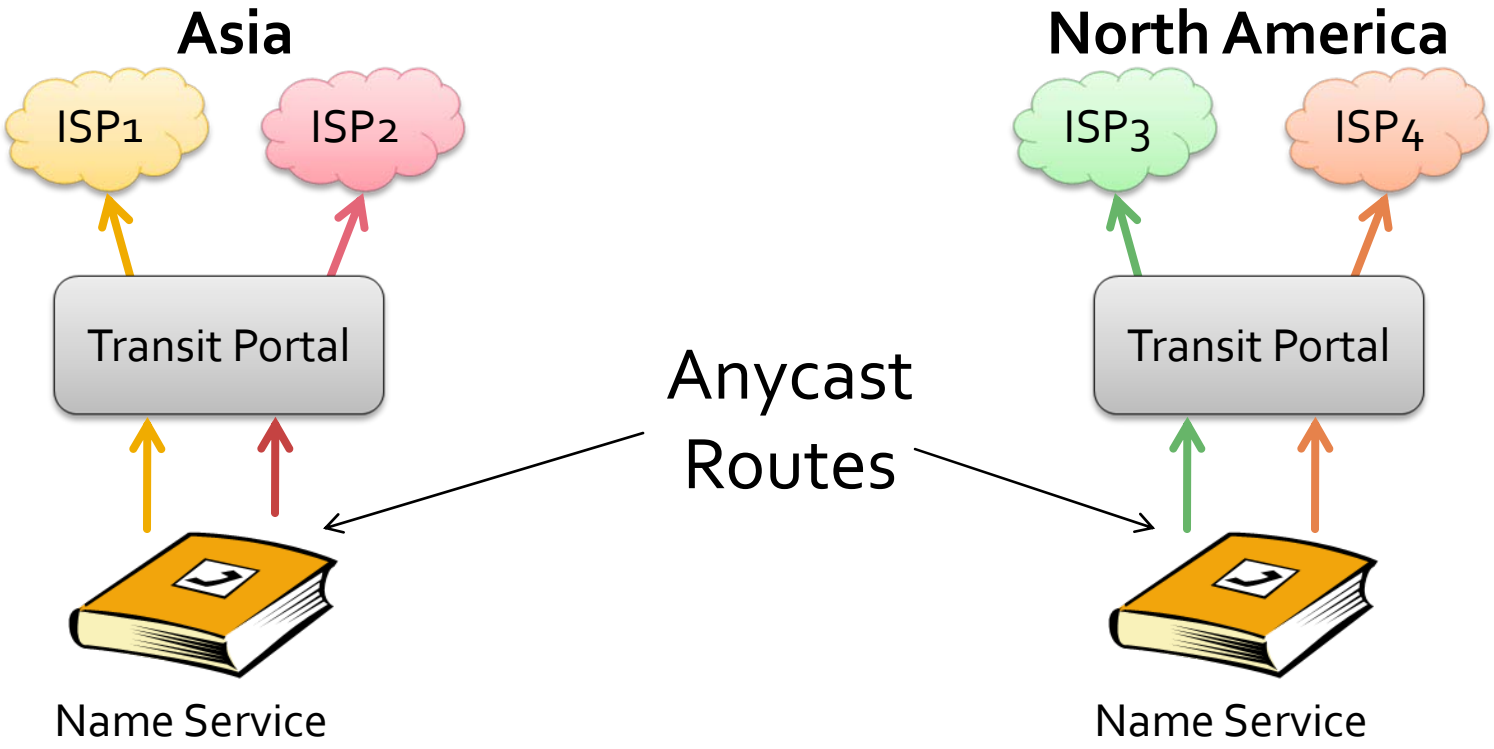
- Server with custom routing software
  - 4GB RAM, 2x2.66GHz Xeon cores
- Three active sites with upstream ISPs
  - Atlanta, Madison, and Princeton
- A number of active experiments
  - BGP poisoning (University of Washington)
  - IP Anycast (Princeton University)
  - Advanced Networking class (Georgia Tech)

TP Applications: Fast DNS

- Internet services require fast name resolution
- IP anycast for name resolution
  - DNS servers with the same IP address
  - IP address announced to ISPs in multiple locations
  - Internet routing converges to the closest server
- Available only to large organizations

TP Applications: Fast DNS

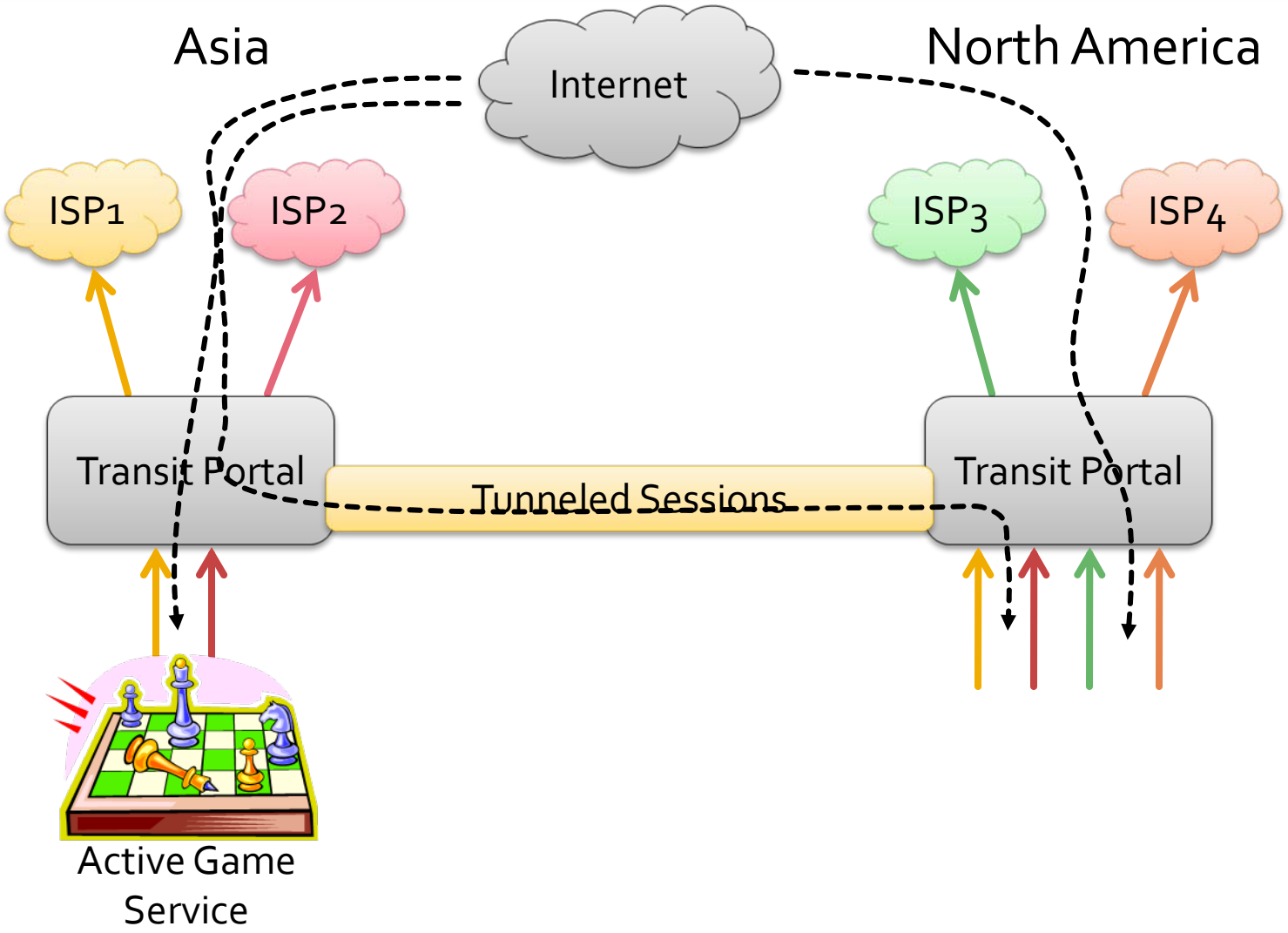
- TP allows hosted applications use IP anycast



TP Applications: Service Migration

- Internet services in geographically diverse data centers
- Operators migrate Internet user's connections
- Two conventional methods:
  - DNS name re-mapping
    - Slow
  - Virtual machine migration with local re-routing
    - Requires globally routed network

TP Applications: Service Migration

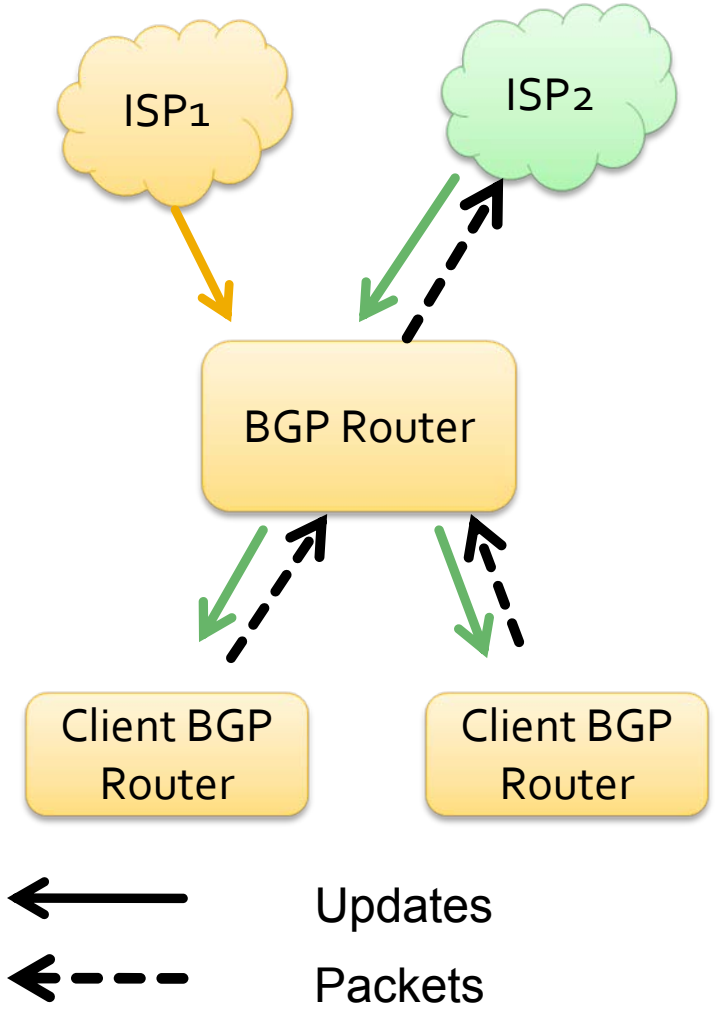


Scaling the Transit Portal

- Scale to dozens of sessions to ISPs and hundreds of sessions to hosted services
- At the same time:
  - Present each client with sessions that have an appearance of direct connectivity to an ISP
  - Prevented clients from abusing Internet routing protocols

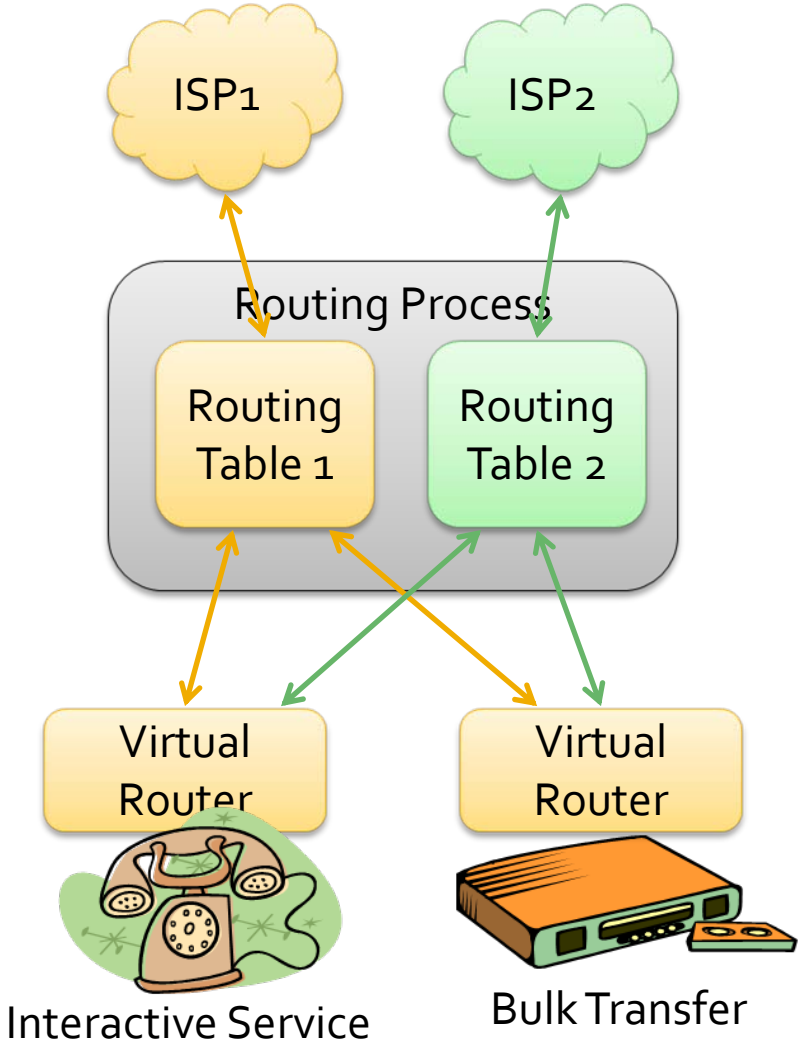
Conventional BGP Routing

- Conventional BGP router:
  - Receives routing updates from peers
  - Propagates routing update about one path only
  - Selects one path to forward packets
- **Scalable but not transparent or flexible**



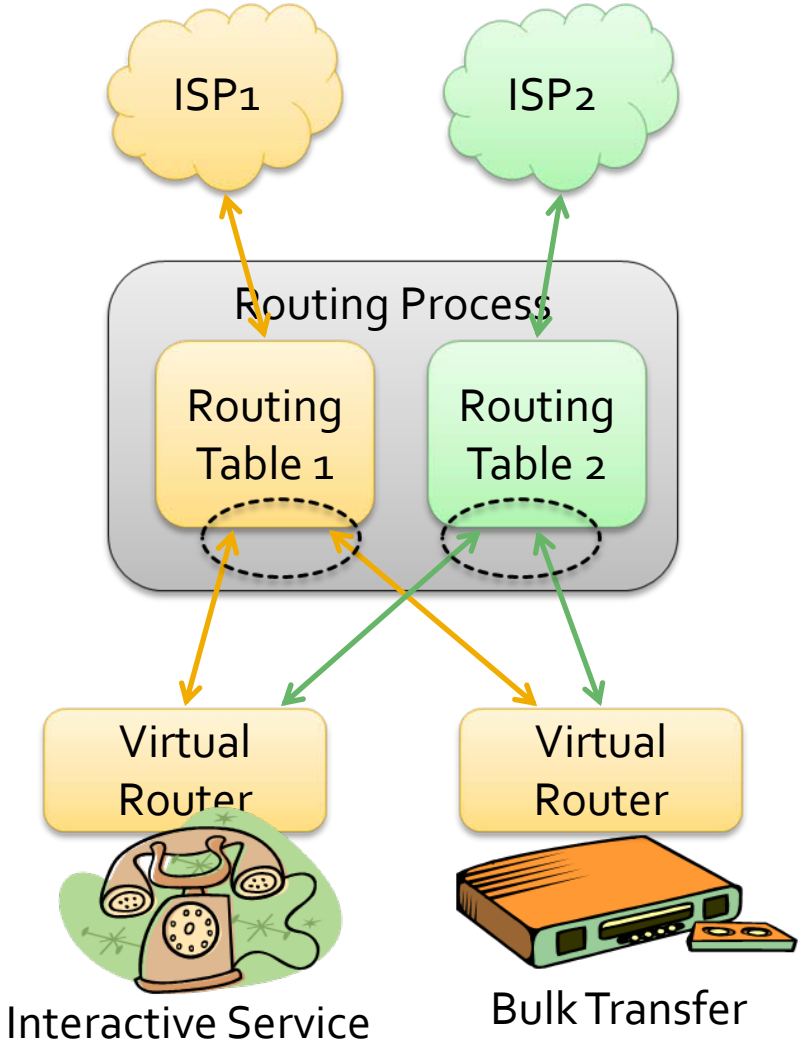
Scaling BGP Memory Use

- Store and propagate all BGP routes from ISPs
  - Separate routing tables
- Reduce memory consumption
  - Single routing process - shared data structures
  - Reduce memory use from 90MB/ISP to 60MB/ISP



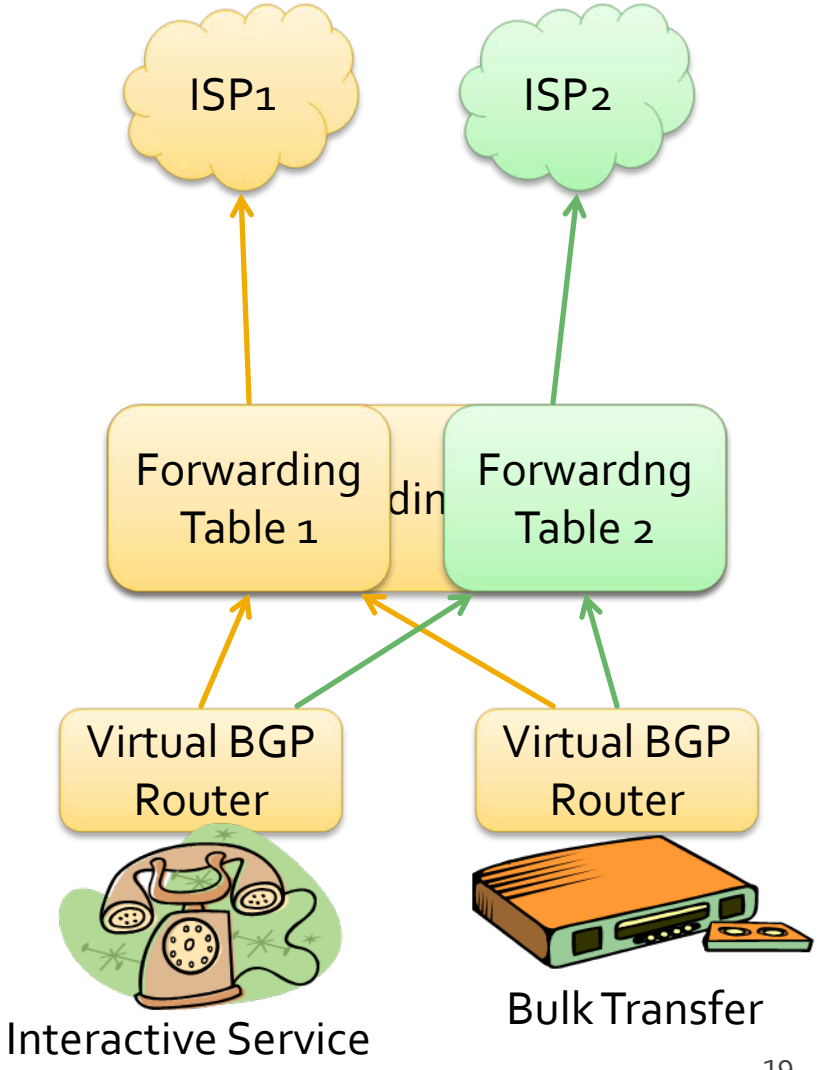
Scaling BGP CPU Use

- Hundreds of routing sessions to clients
  - High CPU load
- Schedule and send routing updates in bundles
  - Reduces CPU from 18% to 6% for 500 client sessions



Scaling Forwarding Memory for TP

- Connecting clients
  - Tunneling and VLANs
- Curbing memory usage
  - Separate virtual routing tables with default to upstream
  - 50MB/ISP -> ~0.1MB/ISP memory use in forwarding table



Future Work

- Future work:
  - More deployment sites
  - Making TP accessible for network research test-beds (e.g., GENI, CoreLab)
  - Faster forwarding (NetFPGA, OpenFlow)
  - Lightweight interface to route control

Conclusion

- Limited routing control for hosted services
- Transit Portal gives wide-area route control
  - Advanced applications with many TPs
- Open source implementation
  - Scales to hundreds of client sessions
- The deployment is real
  - Can be used today for research and education
  - More information http://valas.gtnoise.net/tp

Questions?