Can We Improve Internet Performance? An Expedited Internet Bypass Protocol

Dr. –Ing. Nirmala Shenoy

Professor, ISchool, School of Information

Director, Lab for Networking and Security

Golisano College of Computing and Information Sciences

Rochester Institute of Technology, Rochester, New York 14623

nxsvks@rit.edu

Agenda

- Growing Internet Complexity
- Escalating Proprietary Solutions & Infrastructure Costs
- Can we improve Internet performance?
  - A Cost Effective – Low Complexity Solution
  - The Expedited Internet Bypass Protocol (EIBP)
- Performance tested an EIBP prototype on the GENI Tested
  - Compared with IP &BGP, IP&OSPF
- Future work
- Discussions / Questions

Growing Networks and Needs

- Number of Internet Users and Networks continue to grow
- Current Layer 3 Protocols (IP, BGP, OSPF)
  - IP to forward Internet packets, BGP and OSPF are routing protocols
  - Are they addressing the growing needs?
  - Challenges
    - Developed decades ago – **Severe Limitations**
    - **Sluggish and unstable**
- The **Needs** – Next Slide

The Demand Scenario

USERS

- Federal, Defense and Emergency networks..
- Need secure, reliable and fast delivery of data

SERVICES

- Content delivery
  - Growing CDN providers and networks
  - High infrastructure investment
- Proprietary solutions
  - GAFAM

  (Google, Amazon, Facebook, Apple, Microsoft)
- Private CDNs

Internet Today

- Internet Infrastructure – widely deployed
  - Challenges
    - Heavy traffic
    - Security
    - Reliability
  - BGP Scalability
  - Complex interworking OSPF, iBGP, eBGP (for inter-AS and intra-AS)

Internet Today (contd)

- DATA travels across several networks, several tens of routers
  - ***Routing Path through the networks defined by Routing tables***
  - ***Routing Table Size > 800,0000***
  - ***Severe Security Concerns at Layer 3***
- ***Consequences***
  - Non-deterministic Delays
  - Unpredictable Loss of Data
  - Vulnerable to security attacks
  - Privacy Compromised

Solution?

- Improve the Internet? – We are trying
- Replace the Internet? …….

- Bypass the Internet – possible
  - Turn on bypass services for specific IP users when needed
  - The Expedited Internet Bypass Protocol (EIBP)

The Expedited Internet Bypass Protocol

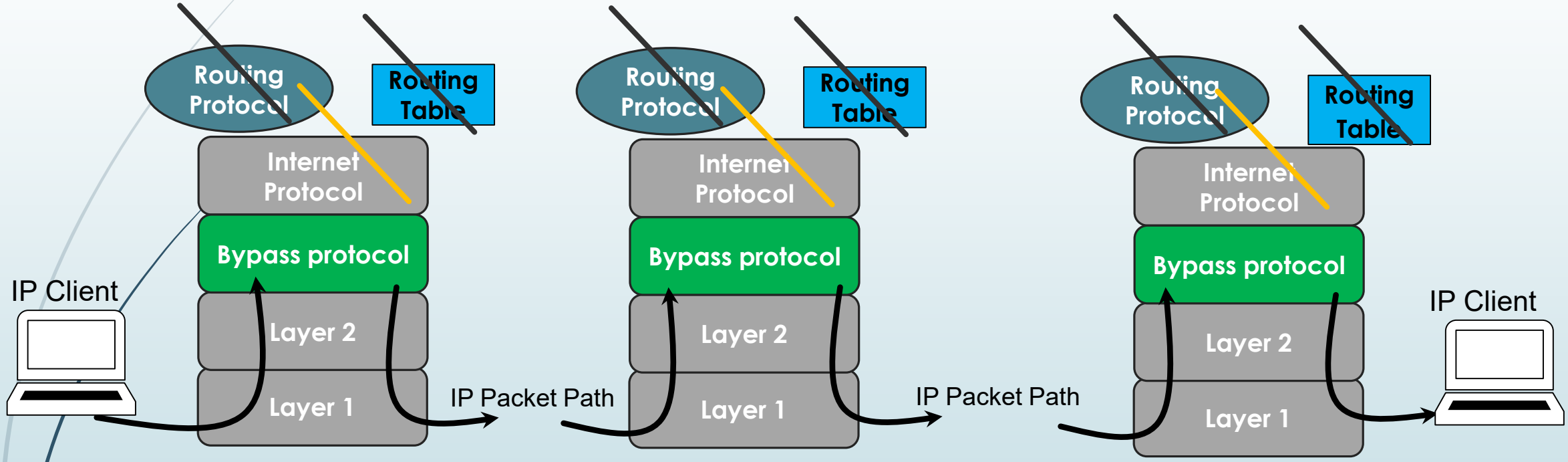
- EIBP for end to end IP packet delivery (IP Network or user)
- Uses no routing protocols
  - No global dissemination of routes
  - No routing tables
- ***Auto-configured addresses at routers provide routing information***
  - Multiple routing Paths
  - Topology changes have localized impact
- Extremely Fast Recovery on component Failures
- **A Single Protocol** to route and forward
  - Integrates control and data planes
  - Simple and robust

The Expedited Internet Bypass Protocol

- Expedites selected traffic –
  - EIBP traffic flows below IP, hence IP traffic is avoided
  - EIBP traffic bypasses layer 3 security threats
- EIBP has no dependency on any Layer 3 protocol
- Traffic flow at Layer 3 is not impacted
  - EIBP operations are transparent to operations at Layer 3

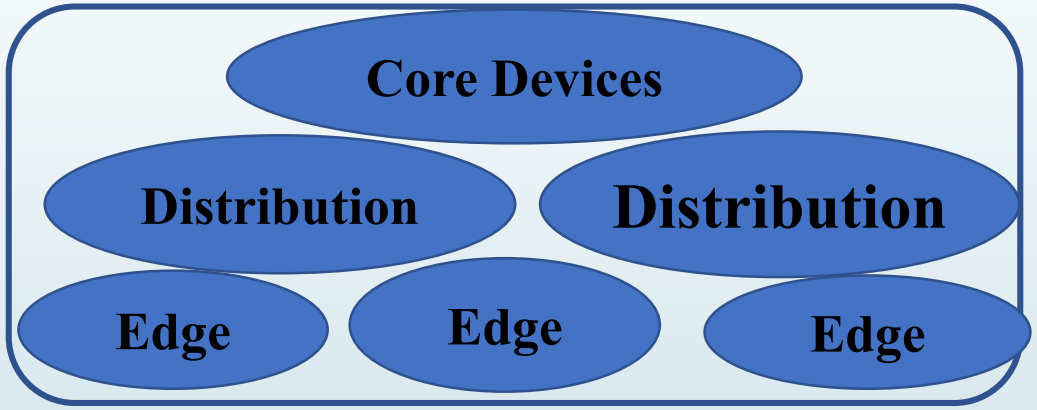
- EIBP has been coded and prototype tested (GENI testbed)
- Performance compared to IP &OSPF, IP&BGP

The Expedited Internet Bypass Protocol



Routing with EIBP

- EIBP routes using structures
  - Physical or Virtual Structures
  - Scalable and Modular
  - Avoids loops

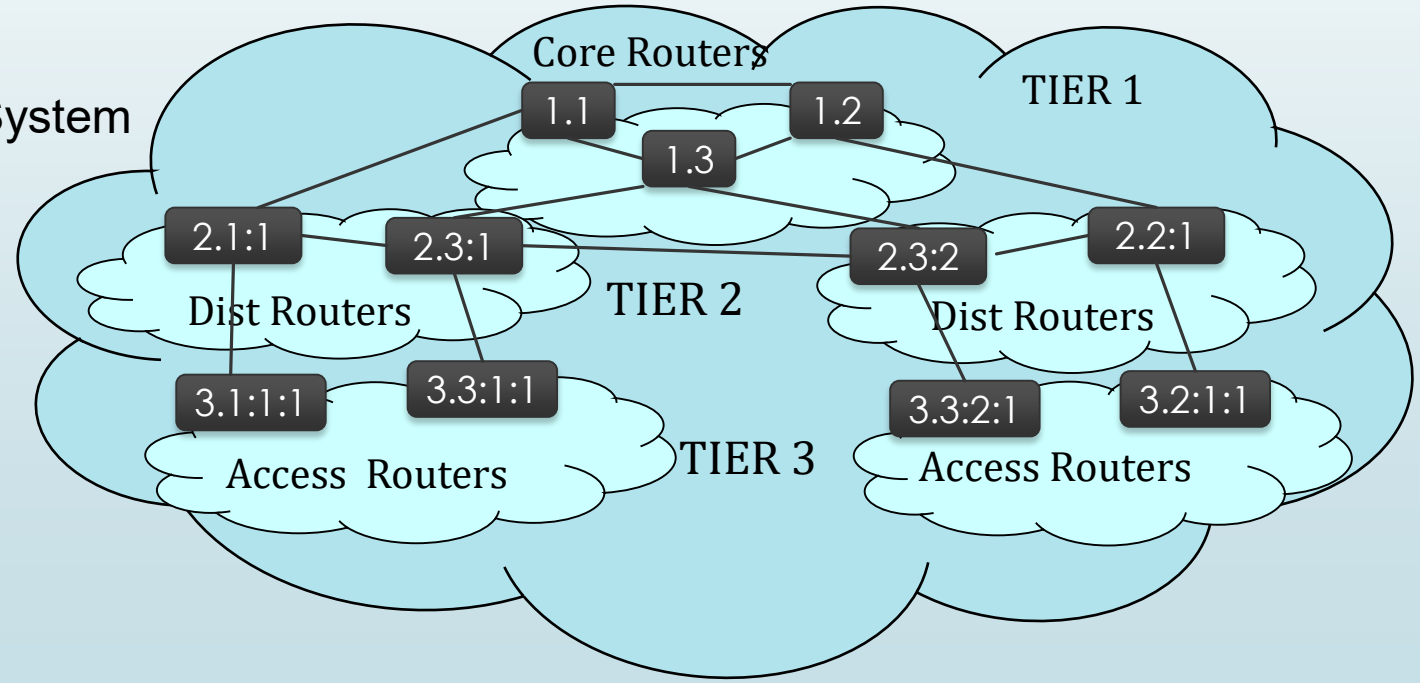


- Example – Three Tier Structure in networks

Structed Addresses

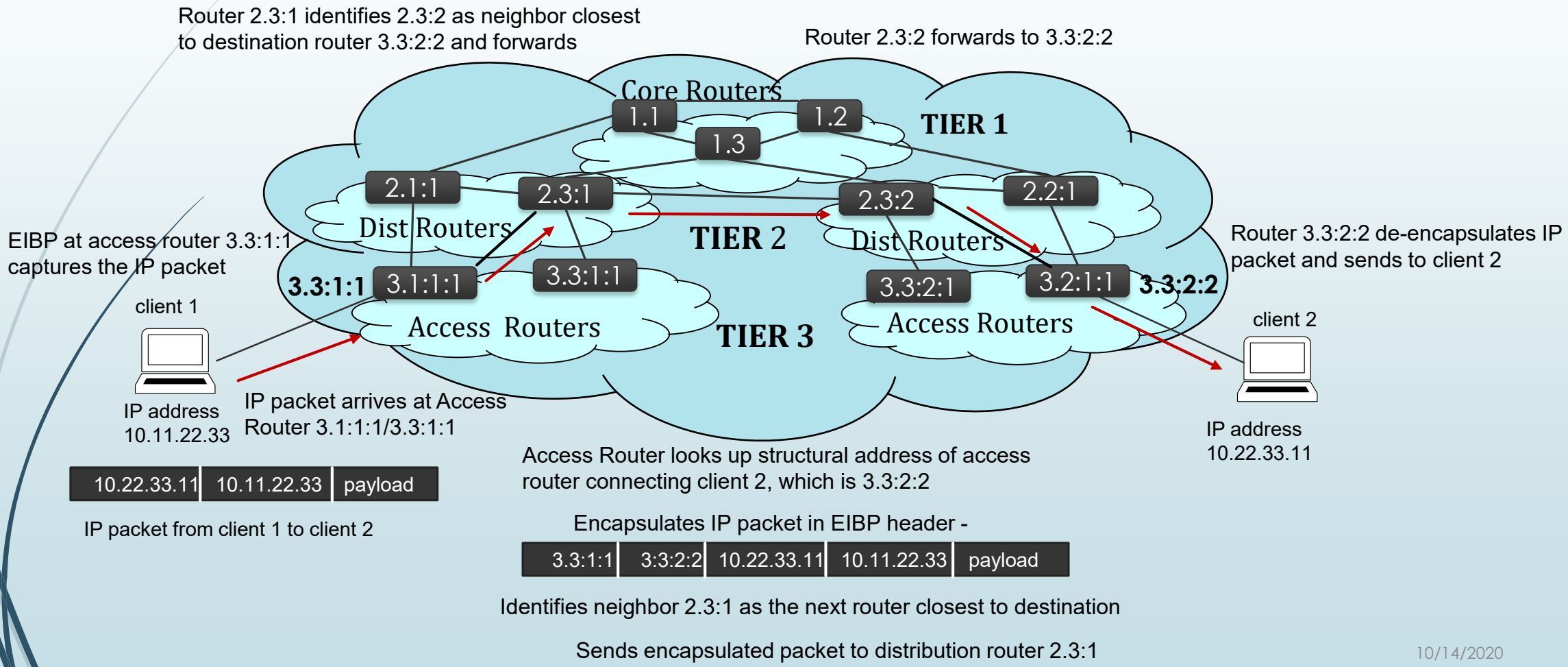
- Addresses carry routing Information
- Simple address assignment – auto-configuration except in Tier 1,
- Addresses updated on topology changes
  - Changes are localized
- Self-configuring, self-healing

Example - Autonomous System



Routing with Structured Addresses (ANIMATED SLIDE)

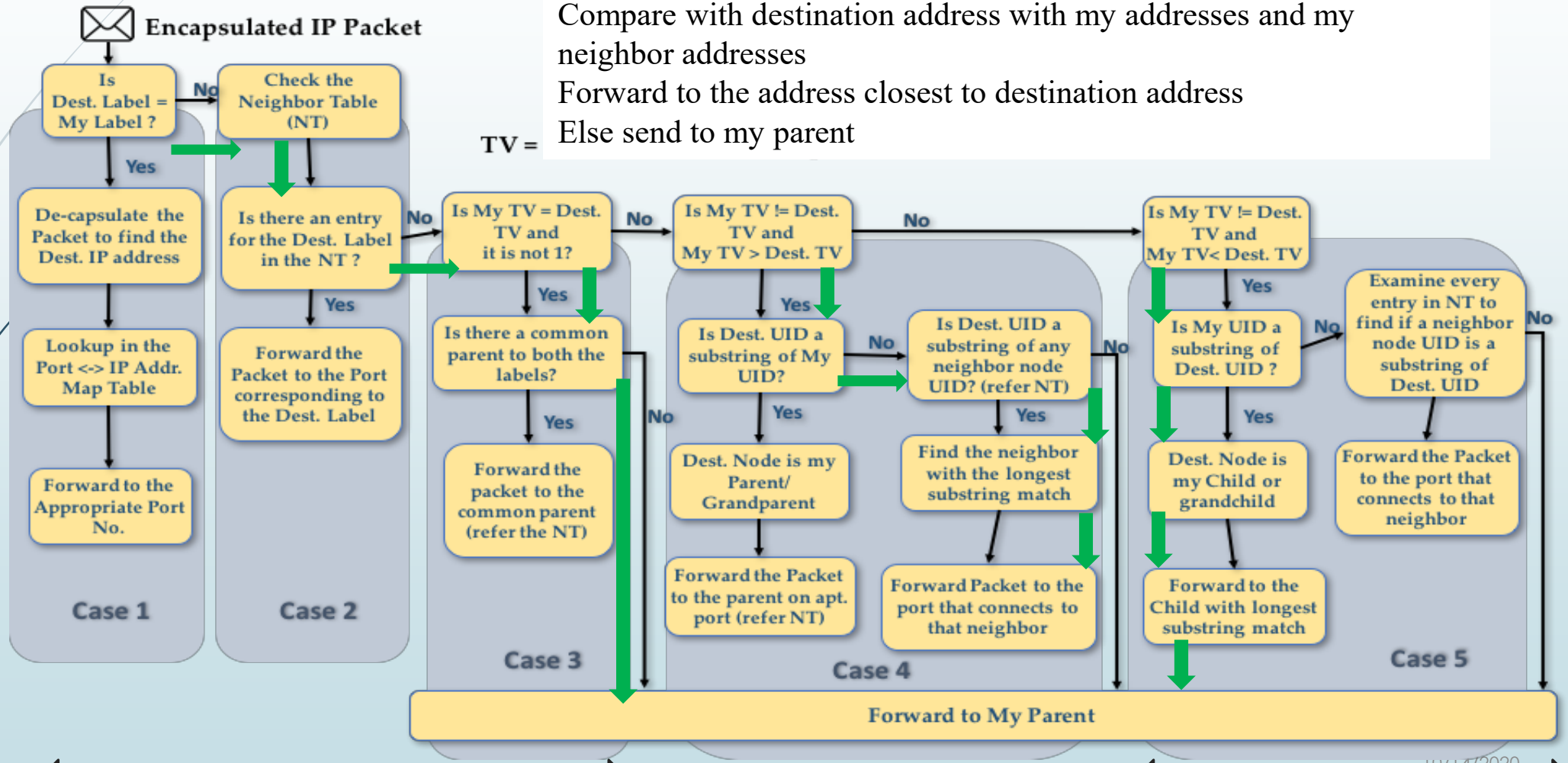
Knowledge of edge router labels and networks they connect



Flow Chart to Route with EIBP

→ Decision path followed in previous example

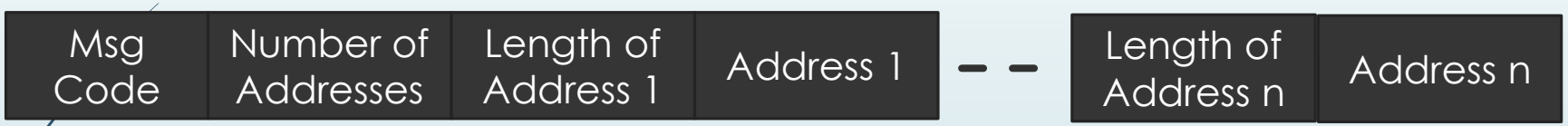
Compare with destination address with my addresses and my neighbor addresses
Forward to the address closest to destination address
Else send to my parent



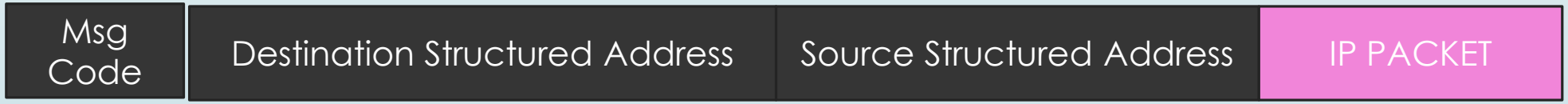
EIBP Implementation

Knowledge of edge router labels and networks they connect

- EIBP messages carried in Ethernet frames - uses an unused type value in the protocol type field
  - Captured on arrival at the sockets by EIBP
- Hello Message – variable addresses- only if addresses change



- Encapsulation of IP Packet



- Join Request Message – lower tiers send to upper tiers

Bypass Protocol Implementation

- *Implemented as a software that operates below the Internet Protocol*
  - *Prototype Tested for intra-AS*

- The EIBP code was written in C language and ported into Linux Systems (Ubuntu 16.04) in the GENI testbeds
- Code Available on gitlab

http://www.rit.edu/news/story.php?id=61939

EIBP Implementation Flexibility

- Code ported into routers – runs below IP without disrupting normal IP operation
- All routers in a network must run a copy of EIBP

- Turn on EIBP– WHEN NEEDED
  - For specific end IP networks/hosts

Prototype Tests on GENI Testbed

Performance Compared with IP&OSPF and IP&BGP

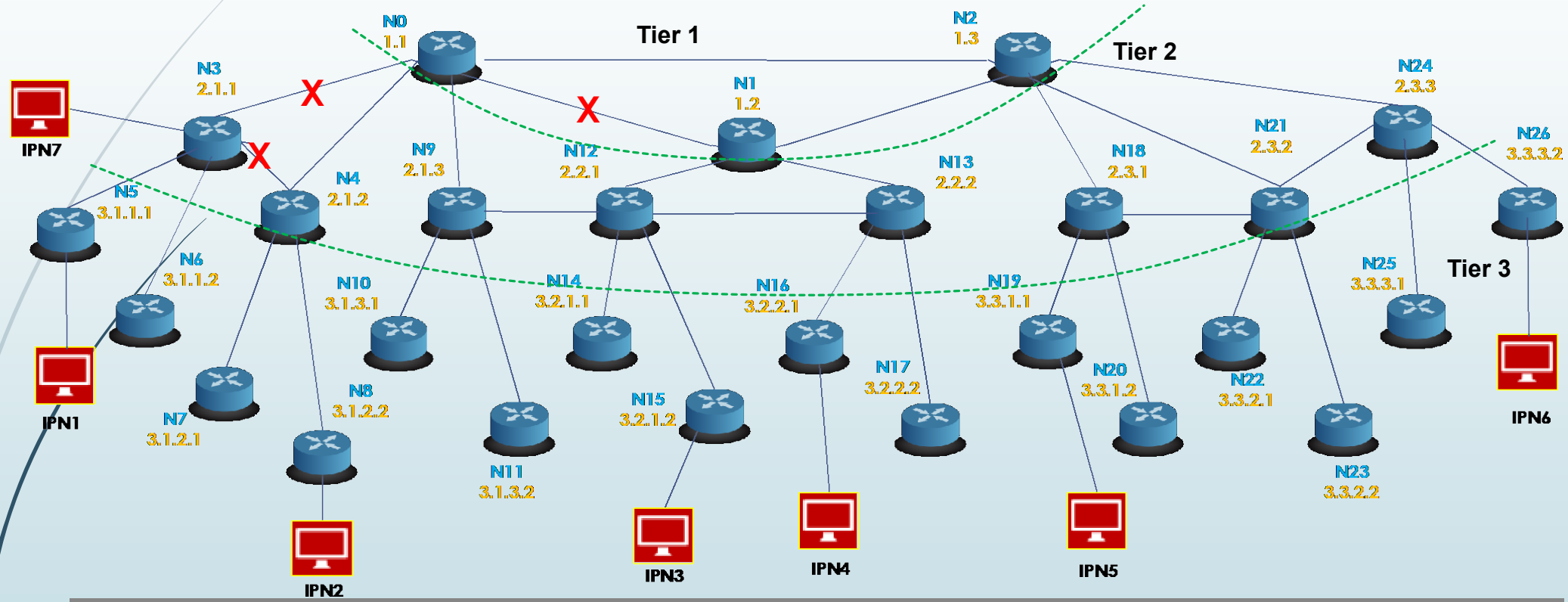
What is the GENI testbed?

GENI (Global Environment for Network Innovations) provides a virtual laboratory for networking and distributed systems research and education. It is well suited for exploring networks at scale, thereby promoting innovations in network science, security, services and applications. GENI allows experimenters to:

- **Obtain compute resources from locations around the United States;**
- **Connect compute resources using Layer 2 networks in topologies best suited to their experiments;**
- **Install custom software or even custom operating systems on these compute resources;**
- **Control how network switches in their experiment handle traffic flows;**
- **Run their own Layer 3 and above protocols by installing protocol software in their compute resources and by providing flow controllers for their switches.**

→ https://www.geni.net/about-geni/what-is-geni/

Prototype Evaluation on GENI Test Bed
17 Routers with IP Clients



17 NODE TEST TOPOLOGY ON GENI TESTBED

X – Failure Points (only one address shown)

This is one of many tests conducted. Please check Nirmala Shenoy, Shashank Rudroju and Jennifer Schneider, " An Emergency Internet Bypass Lane Protocol", High Performance Computing and Communications (HPCC-2018) Exeter, England, UK, 28-30 June 2018

Convergence Process on Failures

- Convergence time = Failure detection time + Protocol recovery time
- Failure Detection Time
  - The node with the failed interface knows first. Node across from the failure has to miss hello messages to detect failure and take action
  - Bidirectional Forwarding Detection can speed up failure detection
- Protocol Recovery Time – is a true measure of a protocol's recovery process and its robustness to failures

Convergence Delays Recorded

- In the tests, protocol timers were used for failure detection
- BGP failure detection averages to 180 seconds (60 second hello timer and 3 missing hellos) – default values
  - Future tests will optimize these values
- OSFP failure detection averages to 30 seconds (10 second hello timer and 3 missing hellos)
  - Recent tests with BFD enabled
- OSPF convergence was calculated by recording the time when updates messages stopped and Link State database stabilized after a failure
- EIBP does not flood network with route changes –
  - Hello timer was set to 1 sec, and in the event of failure, the next path was used
  - To avoid flapping interfaces – hysteresis in reinstating an address was adopted
- Convergence delays record = Failure detection + Protocol Recovery time

Failure Recovery and Convergence

FAILURE BETWEEN N3 AND N4		
Protocol	Convergence (seconds)	Impact Ratio
BGP	FD+100 (PR)	26/27
OSPF	FD+30 (PR)	8/27
EIBP	1	2/27
FAILURE BETWEEN N0 AND N1		
BGP	FD+100 (PR)	19/27
OSPF	FD+30 (PR)	27/27
EIBP	1	2/27
FAILURE BETWEEN N0 AND N3		
BGP	FD+100 (PR)	27/27
OSPF	FD+30 (PR)	25/27
EIBP	3	5/27
FD – Failure Detection, PR – Protocol Recovery		

- CONVERGENCE TIME in secs is the recovery time after a link failure.
- Deducting the failure detection time BGP records > 80 seconds for its tables to stabilize
- OSPF records > 30 seconds for its tables to stabilize
- EIBP records 1 second.

- IMPACT RATIO is the number of routers that update their routing tables on a link failure.
- With BGP most routers update.
- With OSPF in certain cases impact ratio is slightly below 1/3.
- With EIBP less than 1/5 and in many cases less than 1/13.

Routing Table Sizes

Protocol	Routing Table Size	
BGP	93	multiple backup
OSPF	83	at least 1 backup
EIBL	5	Neighbor table Size

RESULTS INTERPRETATION

ROUTING TABLE SIZE provides a measure of scalability of the protocol. For a 27 router partial meshed topology,

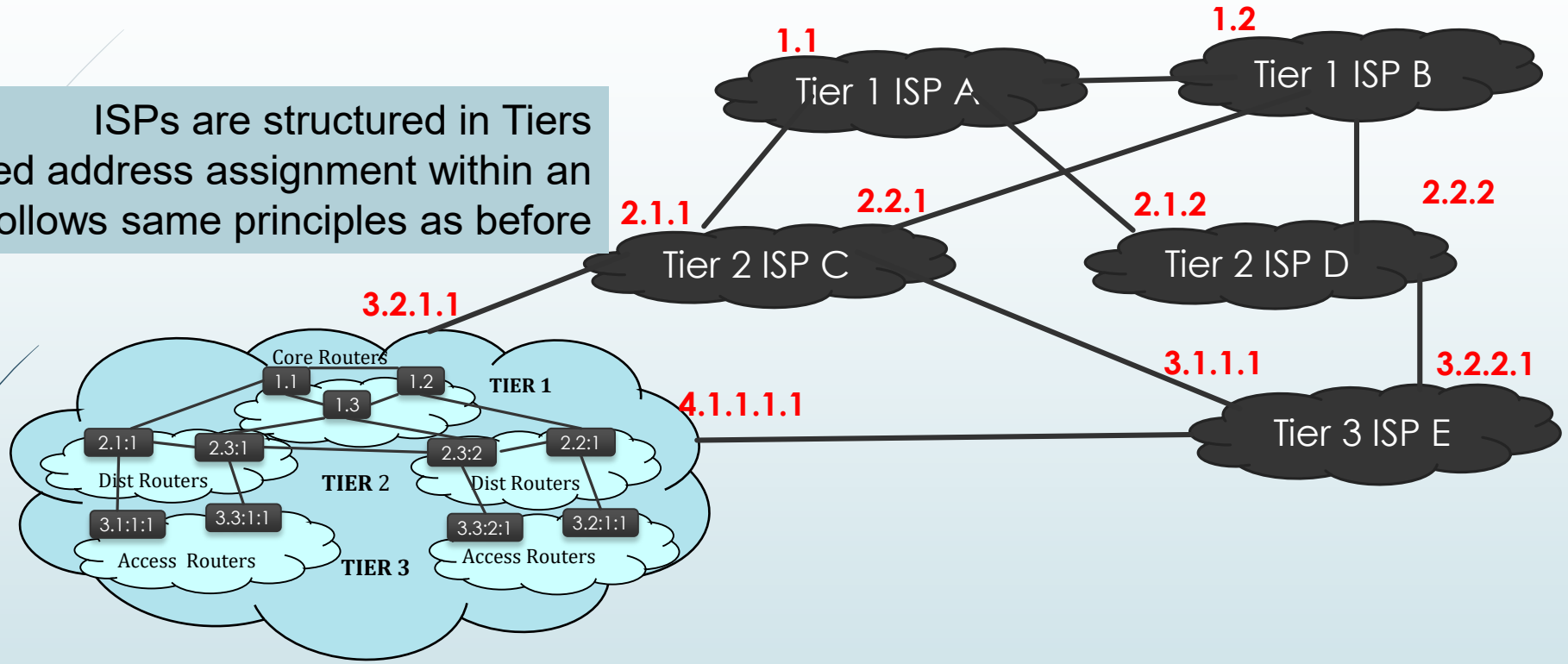
- BGP records 93 entries,
- OSPF records 83 entries and the EIBP recorded a max of 5 entries

Benefits

- Several magnitudes in recovery time on failures
- Routing simplified
  - A single protocol updates routing information + forwards packets
  - Integrated control and data operations
- Improved Security and Privacy for data transfers
- Improved Fault Tolerance
- Seamless interworking of intra-AS and Inter-AS operations
- Easy deployment / migration
- RIT news item
  - http://www.rit.edu/news/story.php?id=61939

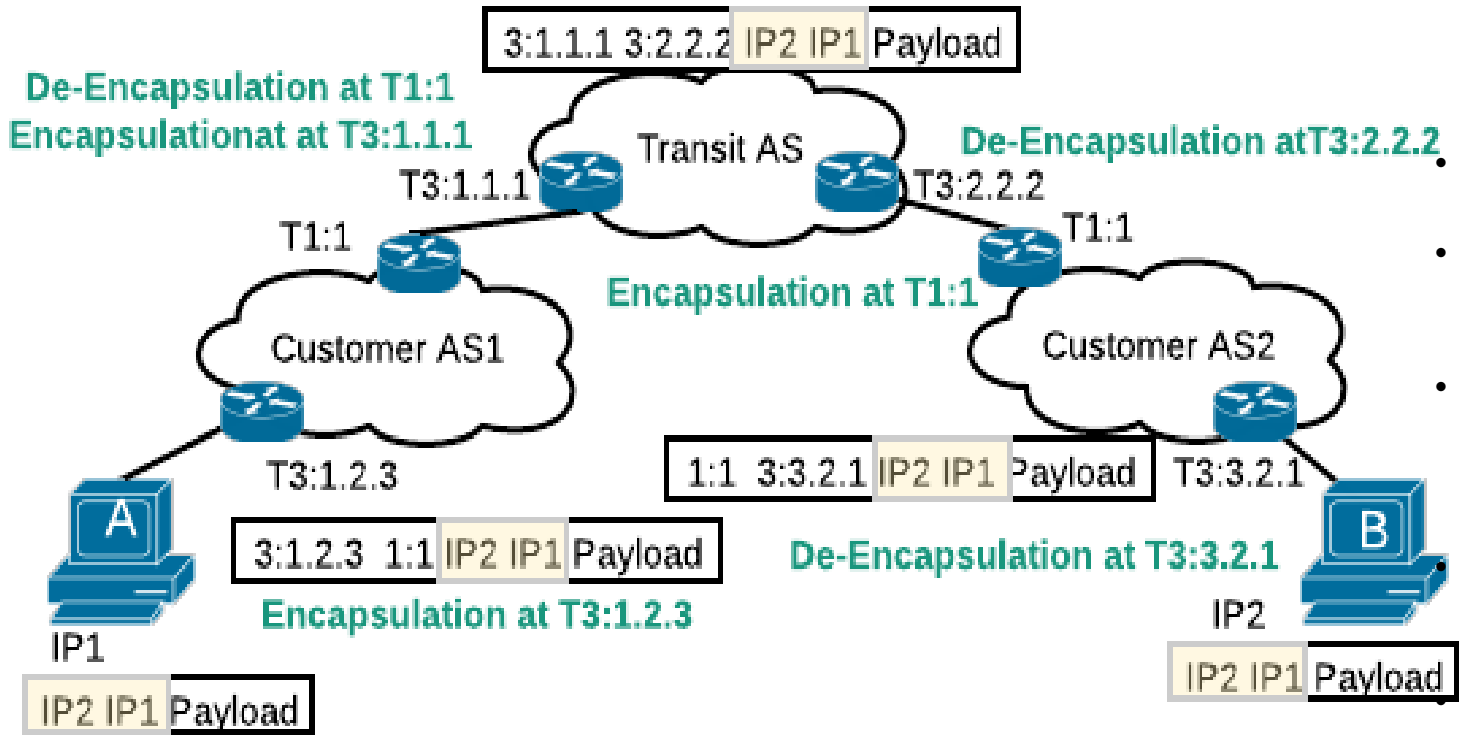
Inter-AS with EIBP

ISPs are structured in Tiers
Auto structured address assignment within an ISP follows same principles as before



Inter-AS forwarding next slide

Extending to inter-AS



A single protocol for intra-AS and inter-AS
No iBGP

- System A sends an IP packet to System B.
- At access router T3:1.2.3 IP packet is encapsulated and sent to the core /border router as destination B' address IP2 is not in AS1.
- Packet reaches T1:1, it forwards the packet to T3:1.1.1 at the transit AS.
- At the transit AS, the access routers have the AS IP addresses that the transit AS is connected to.
- The access routers also have a map of the AS IP addresses (that the transit AS connects) mapped to the structured address of the access routers.
- Router T3:1.1.1 will encapsulate the IP packet with new header and the packet will be delivered to router T3:2.2.2, T3:2.2.2.2 will de-encapsulate and send to T1:1 at Customer AS2.
- The packet is re-encapsulated at T1:1 at Customer AS2, and delivered to access router T3:3.2.1
- Access router de-encapsulates and deliver to System B

Future Features

- Fast failure detection and recovery without the use of BFD
  - Failover with single missing hello (partly implemented)
  - Send 1 byte hello messages at a high frequency
  - Hysteresis based recovery when failed link/device comes up
- Link Failure
  - Router recognizes first
  - Identifies failed addresses and disseminates
- Concepts to be tested

Future Benefits

- Efficient use of Internet infrastructure
  - Leverage the current infrastructure to offer superior services
  - Reduce deployment of proprietary, costly and resource intensive infrastructure
  - Offer expedited services on need.

THANKS

QUESTIONS