Overcoming BGP’s Limitations
SIGCOMM 2017 Preview Session

Ethan Katz-Bassett
Columbia University
Thursday 2-3:15 Session 10 - Peering

- **Paper 1: Engineering Egress with Edge Fabric**
  Brandon Schlinker, Hyojeong Kim, Timothy Cui, Ethan Katz-Bassett, Harsha V. Madhyastha, Italo Cunha, James Quinn, Saif Hasan, Petr Lapukhov, Hongyi Zeng (Facebook, USC, Columbia, Michigan, UFMG)

- **Paper 2: Taking the Edge off with Espresso**
  Kok-Kiong Yap, Murtaza Motiwala, Jeremy Rahe, Steve Padgett, Matthew Holliman, Gary Baldus, Marcus Hines, Taeeun Kim, Ashok Narayanan, Ankur Jain, Victor Lin, Colin Rice, Brian Rogan, Arjun Singh, Bert Tanaka, Manish Verma, Puneet Sood, Mukarram Tariq, Matt Tierney, Dzevad Trumic, Vytautas Valancius, Calvin Ying, Mahesh Kallahalla, Bikash Koley, Amin Vahdat (Google)

Thursday 3:45-5 Session 11 - Routing

- **Paper 2: Bootstrapping evolvability with D-BGP**
  Raja R. Sambasivan, David Tran-Lam, Aditya Akella, Peter Steenkiste (BU, Wisconsin, CMU)
3 papers this year on overcoming BGP limitations

To preview the topic, we’ll look at:

- what is BGP?
- what are its limitations?
- why is it challenging to overcome them?
- what is the basic question asked in each paper?
The Internet

Some Web Server

Client's Computer

The Internet
How to route traffic to destination?
The Internet

Some Web Server

The Internet

Client's Computer
A federation of autonomous networks

(ISP=Internet Service Provider for our purposes equivalent to AS=Autonomous System)
A federation of autonomous networks

(ISP=Internet Service Provider for our purposes equivalent to AS=Autonomous System)
Establishing inter-network routes

BGP (Border Gateway Protocol)
- Selects path based on *opaque policy* of individual *autonomous system*
Establishing inter-network routes

BGP (Border Gateway Protocol)
- Selects path based on **opaque policy** of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)
- Selects path based on **opaque policy** of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)

- Selects path based on **opaque policy** of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)
- Selects path based on **opaque policy** of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)
- Selects path based on **opaque policy** of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)
- Selects path based on **opaque policy**
- of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)
- Selects path based on **opaque policy** of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)
• Selects path based on **opaque policy** of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)

- Selects path based on **opaque policy** of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)
- Selects path based on **opaque policy** of individual **autonomous system**
Establishing inter-network routes

BGP (Border Gateway Protocol)

- Selects path based on **opaque policy** of individual **autonomous system**
- Paths are **per destination prefix**

**Unselected paths** are **not announced** to neighbors
  - So web server’s ISP does not know Sprint-AT&T link
BGP is a great success

Current version has supported Internet since 1994!

BGP

- Selects path based on *opaque policy*

- Paths are *per destination prefix*

- *Unselected paths* are *not announced* to neighbors
BGP is a great success…with big limitations

BGP has limitations

- Selects path based on *opaque policy*
- but *limited information to base decision on*
  - No capacity or performance information
  - No static policy will be optimal as conditions change

Current version has supported Internet since 1994!
BGP is a great success…with big limitations

Current version has supported Internet since 1994!

BGP has limitations

- Selects path based on **opaque policy**
- **limited information to base decision on**
  - No capacity or performance information
  - No static policy will be optimal as conditions change
BGP is a great success... with big limitations

Current version has supported Internet since 1994!

BGP has limitations

- Selects path based on **opaque policy**
  but **limited information to base decision on**
- Paths are **per destination prefix**
BGP is a great success…with big limitations

Current version has supported Internet since 1994!

BGP has limitations

- Selects path based on **opaque policy** but **limited information to base decision on**
- Paths are **per destination prefix** so **limited flexibility**, e.g., for high priority real-time vs bulk backup
BGP is a great success…with big limitations

Current version has supported Internet since 1994!

BGP has limitations

- Selects path based on *opaque policy* but *limited information to base decision on*
- Paths are *per destination prefix* so *limited flexibility*
- *Unselected paths* are *not announced* to neighbors
BGP is a great success...with big limitations

Current version has supported Internet since 1994!

BGP has limitations

- Selects path based on *opaque policy* but *limited information to base decision on*
- Paths are *per destination prefix* so *limited flexibility*
- *Unselected paths* are *not announced* to neighbors so *limited path diversity*
BGP is a great success…with big limitations

Current version has supported Internet since 1994!

BGP has limitations

• Selects path based on **opaque policy**
  but **limited information to base decision on**

• Paths are **per destination prefix**
  so **limited flexibility**

• **Unselected paths** are **not announced** to neighbors
  so **limited path diversity**

• **Vulnerable to attacks**
BGP has limitations

- Selects path based on *opaque policy* but *limited information to base decision on*
- Paths are *per destination prefix* so *limited flexibility*
- *Unselected paths* are *not announced* to neighbors so *limited path diversity*
- *Vulnerable to attacks*

Current version has supported Internet since 1994!

BGP is a great success... with big limitations
BGP is a great success...with big limitations

Current version has supported Internet since 1994!

BGP has limitations

- Selects path based on **opaque policy**
  but **limited information to base decision on**
- Paths are **per destination prefix**
  so **limited flexibility**
- **Unselected paths** are **not announced** to neighbors
  so **limited path diversity**
- **Vulnerable to attacks**
BGP is a great success…with big limitations

Current version has supported Internet since 1994!

BGP has limitations

- Selects path based on **opaque policy** but **limited information to base decision on**
- Paths are **per destination prefix** so **limited flexibility**
- **Unselected paths** are **not announced** to neighbors so **limited path diversity**
- **Vulnerable to attacks**

* Extensions exist but see little use
Many proposals to improve BGP

Current version has supported Internet since 1994!

BGP has limitations

- Selects path based on opaque policy but limited information
- Paths are per destination prefix so limited flexibility
- Unselected paths are not announced so limited path diversity
- Vulnerable to attacks

Many proposed improvements

- EQ-BGP, Wiser, …
- MIRO, Pathlets, …
- RPKI, BGPSec, …

But they have seen little to no adoption, limitations persist, and BGP is essentially unchanged
Barriers exist to adopting improvements

- Tens of thousands of autonomous systems use BGP to talk to each other
  - Need to keep Internet up, so can’t do greenfield design
  - Multiple slow steps: understand problem, design solution, standardize, implement, deploy
  - Incentives can be tricky

- Functionality baked into deployed routers
- Fixed message format with limited information
- Other Internet protocols also slow to change
  - HTTP 1.1: 1997-2015
  - IPv6: proposals in 1992, formalized in 1998, regions started exhausting IPv4 in 2011, only 0.64% of Internet traffic in 2013
3 papers this year on overcoming BGP limitations

- Given we are stuck with BGP, what can we do unilaterally? (Edge Fabric [Facebook] & Espresso [Google])
  - Need to use BGP to talk to ASes, but BGP routing is not flexible enough

- How should we have designed BGP to support evolution? (D-BGP)
Given we are stuck with BGP, what can we do unilaterally? (Edge Fabric [Facebook] & Espresso [Google])

- Need to use BGP to talk to ASes, but BGP routing is not flexible enough
- Achieve **path diversity** by connecting directly to 1000s of ASes
- How to use **performance, capacity, and other info** in decisions?
- Achieve **flexibility** by:
  - Connecting directly to client networks, so only one BGP decision on path
  - Moving control from peering routers to centralized controllers
- Different priorities and settings lead to different designs
  - Espresso replaces BGP for routing, Edge Fabric controls BGP decisions

How should we have designed BGP to support evolution? (D-BGP)
Given we are stuck with BGP, what can we do unilaterally? (Edge Fabric [Facebook] & Espresso [Google])

- Need to use BGP to talk to ASes, but BGP routing is not flexible enough
- Achieve **path diversity** by connecting directly to 1000s of ASes
- How to **use performance, capacity, and other info** in decisions?
- Achieve **flexibility** by:
  - Connecting directly to client networks, so only one BGP decision on path
  - Moving control from peering routers to centralized controllers
- Different priorities and settings lead to different designs
  - Espresso replaces BGP for routing, Edge Fabric controls BGP decisions

How should we have designed BGP to support evolution? (D-BGP)

- Suppose different ASes want to adopt new protocols
- What features would a base protocol need to support this?
Questions?
Further reading

- Limitations of BGP path selection can lead to suboptimal outcomes
  - Examples in Brandon Schlinker’s SIGCOMM 2017 Edge Fabric slides (link not yet live, making educated guess at URL)

- Security limitations of BGP and why improvements see slow adoption.

- Lots more interesting work on BGP limitations and proposed improvements
  - See D-BGP bibliography for citations of some of it