

# Advanced Topics in IP Networks

## Lecture 06

Lecture by Dr. Anat Bremler-Barr

Typeset by Steven Karas

2018-11-22

Last edited 20:44:46 2018-11-22

**Disclaimer** These notes are based on the lectures for the course Advanced Topics in IP Networks, taught by Dr. Anat Bremler-Barr at IDC Herzliyah in the fall semester of 2018/2019. Sections may be based on the lecture slides prepared by Dr. Anat Bremler-Barr.

## 1 Agenda

- Finish up BGP
- Cover a routing protocol, and see why they look the way they do
- SDN
- Programmable switch

A reminder that this is a

## 2 BGP

Border gateway protocol is used to share routes between autonomous systems at either exchange points or peering directly at the same location.

BGP is designed as an incremental protocol. A node learns multiple paths to a destination, but remembers all of them. It then applies policy to select a single active route, and may advertise that route to its neighbors.

An announcement update will add its node id to the path and optionally advertise to each neighbor. A withdrawl update notifies neighbors that a route is no longer available.

### 2.1 BGP Dynamics

Convergence in BGP is nontrivial to define. We have several empirical tests that show that it takes several minutes until a new route is taken up by the entire internet.

Announcements are limited to 30 seconds, but withdrawals may be immediate.

#### 2.1.1 Route flapping

Route flapping is when a route flaps back and forth between two different paths. IGP can expose and export internal instability (for example a link that changes weight/fails repeatedly).

Last week we covered in depth the impact of a single route flapping.

**Route Flap Dampening** A single AS failing can generate a disproportionate amount of announcements. If we prevent route flapping by rate limiting announcements from a single neighbor, we can prevent lots of messages, but this effectively causes incorrect routing, so this approach was reverted shortly after it was published/added.[3]

#### 2.1.2 Convergence

We also covered last week the impact of how long it takes for a route announcement to propagate throughout the internet.

For example, a route disappearing from a network converges in  $30 \cdot n$ , where  $n$  is the number of ASs (because we can't have cycles). Note that as an effect of ghost information, an AS disappearing takes  $n$  steps, instead of  $d$ . An example of this using a clique network is in the slides.

**MinRouteAdver** Only announce routes every 30 seconds. Without this, it would take  $O(n!)$  messages to converge.

### 2.1.3 Message Volume

Cisco had a bug in their routing code, and because of this, there were upwards of 4 million withdrawl messages worldwide daily. After the fix, this dropped off significantly.

## 3 Intradomain Routing

Interdomain routing is between ASs, where there is little to no cooperation or symmetric trust. Intradomain routing in contrast is within the same AS, and typically relies on shortest path routing selected based on link metrics. In many cases, internal routing are based on configuration, not on current link status. Historically, there were attempts to base routing on link status/load, but they were not successful.

OSPF and IS-IS are link-state routing protocols, and RIP and EIGRP are distance-vector routing protocols.

Gateways will combine both internal and external routing tables and be aware of the next hop for a given destination. How to get to a gateway is a complex topic but outside our current scope.

### 3.1 Link-State Routing

Effectively runs Dijkstra's Algorithm, where each router is aware of the entire network. An example can be found on slides 6-8.

As a practical effect, this converges fast, but at the cost of a message per link in the network.

### 3.2 Distance-Vector Routing

Effectively runs Bellman-Ford, where each router is aware for each destination the next hop and the distance. In contrast to path-vector routing, this can suffer from the "count to infinity" problem, where a single failed link can cause nonconvergence.

### 3.3 History and Future

As a practical effect, a relatively low cost link can be oversubscribed. Load-sensitive routing needs to change the forwarding table based upon link metrics. The first ARPANET algorithm from 1968 tried to include this, but congestion advertisements were causing oscillations. Long routes were also bad because they utilize more network resources for longer, and attempts to limit long routes were put into place. The second ARPANET algorithm from 1979 attempted to fix this by averaging load over time. Out of date information caused congestion in alternative links.

As of today, most networks use link-state routing with static weights, which converges quickly in response to topology changes. Some enterprise networks will use distance-vector, but this is largely in decline. There is growing use of MPLS, but also of SD-WAN and other similar technologies.

## 4 Software Defined Networks

We watched [this video](#) as an introduction to SDN. The gist is that it creates an abstraction for basic packet forwarding, and then instructs individual devices to forward packets based on a central authority's software.

Generally speaking, a OpenFlow switch will have a Flow Entry Table. A controller instructs switches how to populate their tables. An entry in a flow table has 3 sections: the rule with many different fields, an action, and a stats section to track metrics such as packet and byte counters.

The standard for openflow is still under development, and new fields/actions/extensions are being added all the time.

[Mininet](#) is a simulator that we will use in the coming weeks to explore SDN.

## References

- [1] Mark Crovella and Balachander Krishnamurthy. *Internet Measurement: Infrastructure, Traffic and Applications*. John Wiley & Sons, Inc., New York, NY, USA, 2006.
- [2] James F. Kurose and Keith Ross. *Computer Networking: A Top-Down Approach Featuring the Internet*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 2002.
- [3] Zhuoqing Morley Mao, Ramesh Govindan, George Varghese, and Randy H. Katz. Route flap damping exacerbates internet routing convergence. *SIGCOMM Comput. Commun. Rev.*, 32(4):221–233, August 2002.
- [4] George Varghese. *Network Algorithmics, : An Interdisciplinary Approach to Designing Fast Networked Devices (The Morgan Kaufmann Series in Networking)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.