Network Security and Measurement

- BGP (Border Gateway Protocol) -

Prof. Dr. Thomas Schmidt

http://inet.haw-hamburg.de | t.schmidt@haw-hamburg.de

Agenda

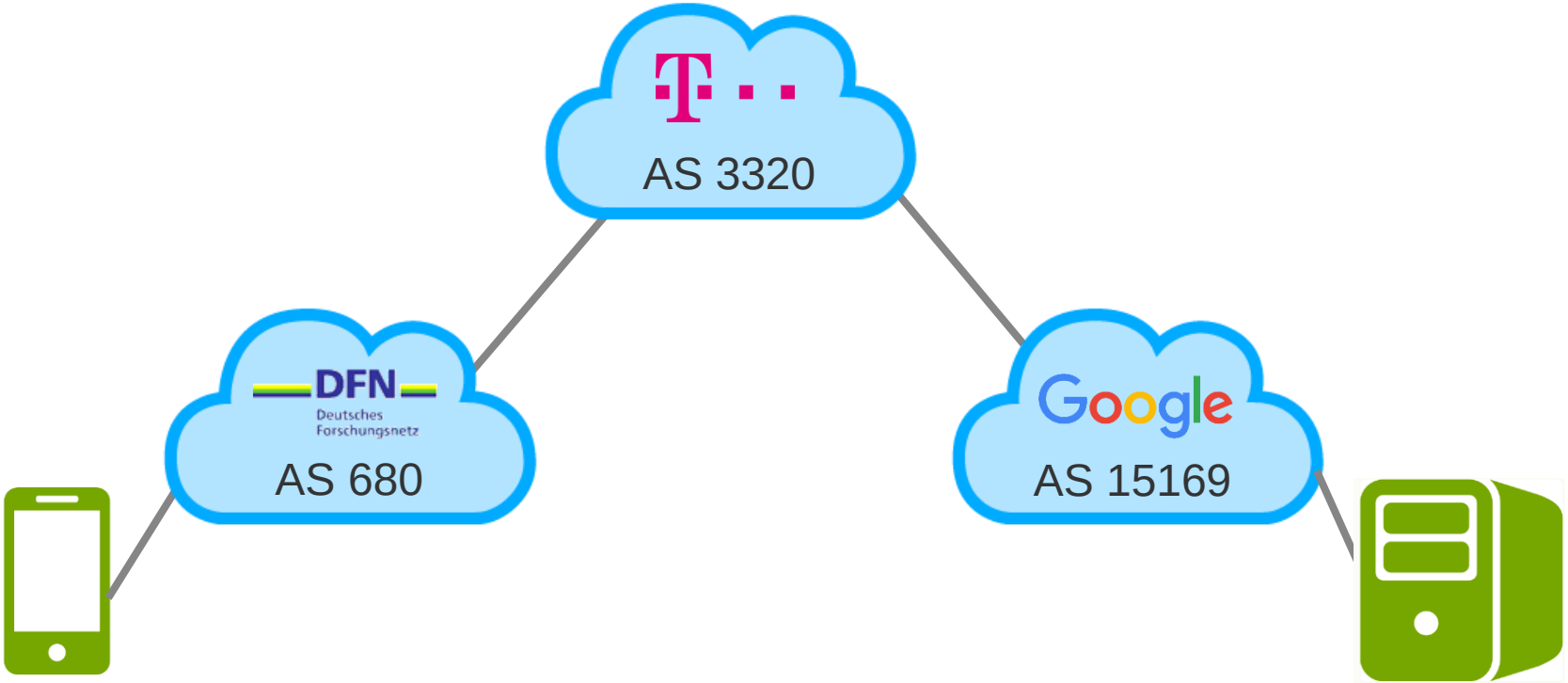
AS Level Topology

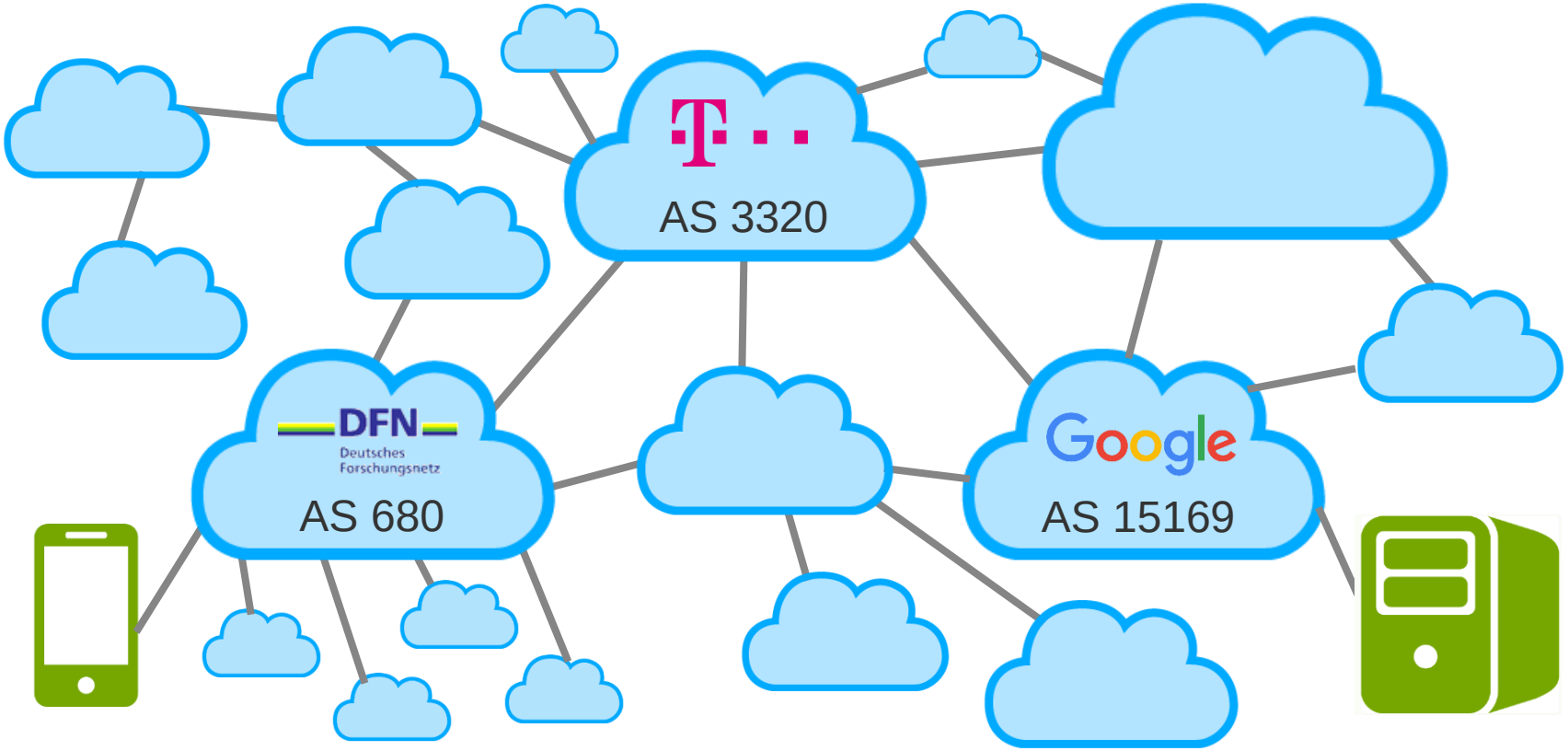
BGP Routing

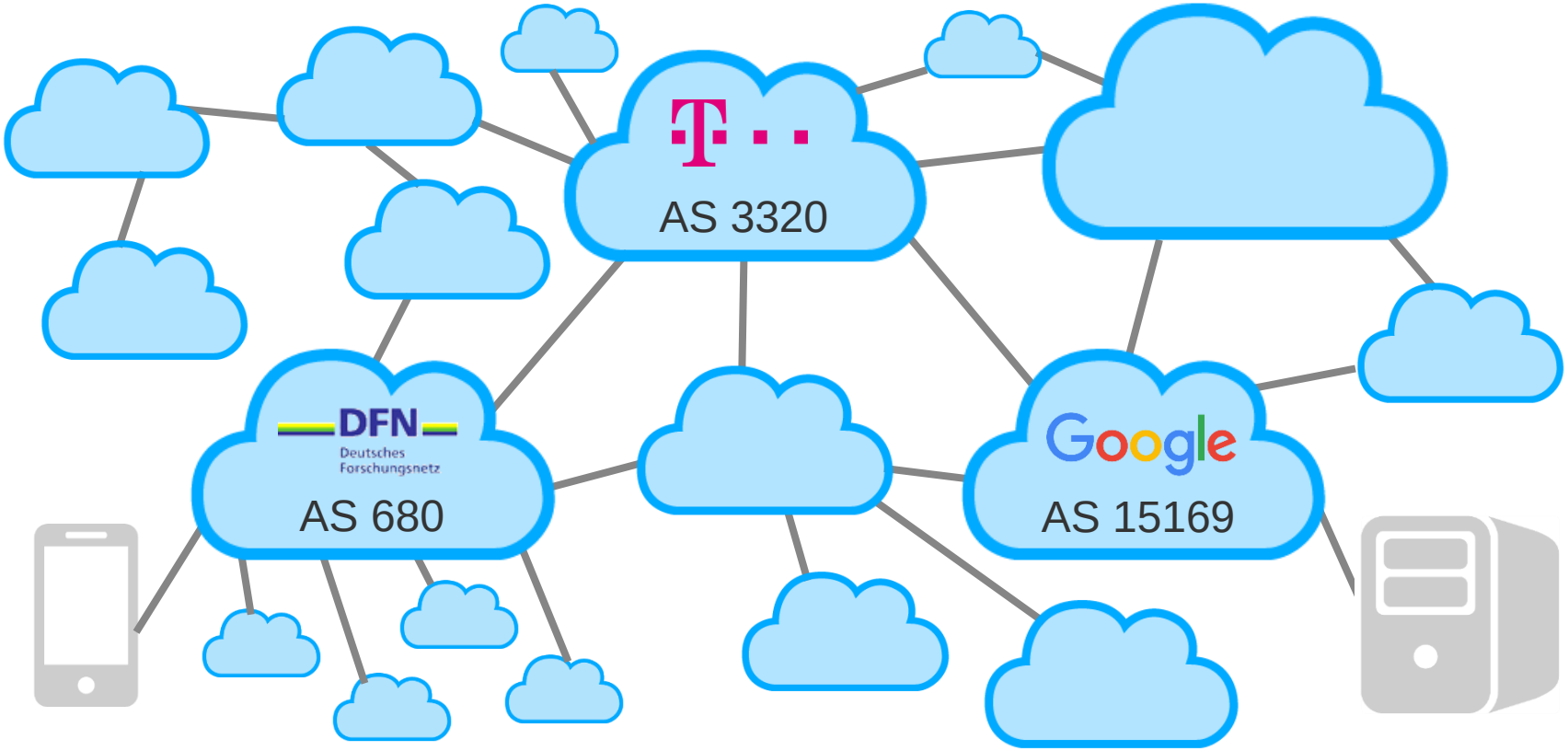
BGP Monitoring

Building a Network of Networks

AS LEVEL INTERCONNECTS


AS 3320
DFN
Deutsches Forschungsnetz
AS 680
Google
AS 15169


AS 3320
DFN
Deutsches Forschungsnetz
AS 680
Google
AS 15169


AS 3320
DFN
Deutsches Forschungsnetz
AS 680
Google
AS 15169

Basically, we are interested in the AS Topology

Characterizing autonomous systems (ASes)

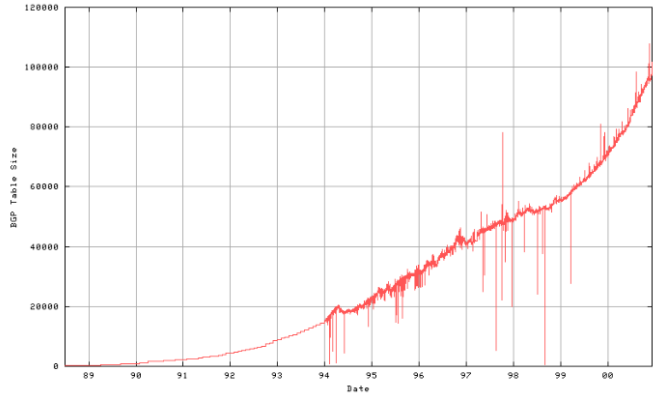
Characterizing AS connectivity

Motivating example 1: Router performance

Scaling the Internet – The Routing View

Geoff Huston
February 2001

Last month, in looking at the structure of routing within the Internet I ended with the Big Question: how will routing deal with the demands of tomorrow's Internet? Lets take a quick look.



There's quite a story behind this chart, and it can tell us a lot about what is likely to happen in the future. The chart appears to have four distinct phases: exponential growth between 1988 and 1994, a correction through 1994, linear growth from 1995 to 1998 and a resumption of exponential growth in the past two years.

heise online heise+

23.12.2006 11:08 Uhr

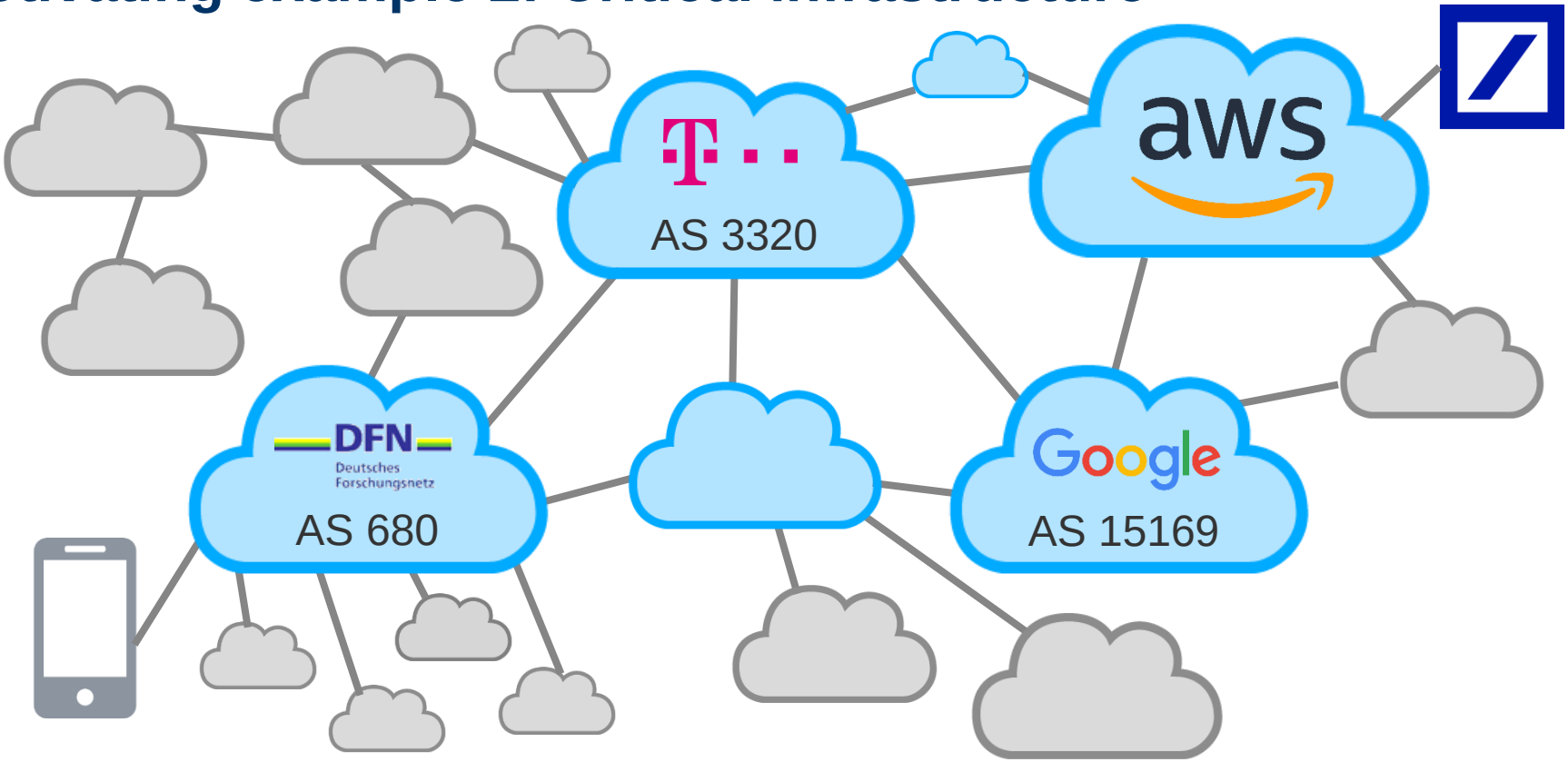
Routing-Tabellen im Internet werden zu groß

Von Monika Ermert

191

Das Internet Architecture Board (IAB) und die Internet Engineering Task Force (IETF) suchen nach einer Lösung für das Problem überbordender Routingtabellen. Einzelne Experten warnen davor, dass das enorme Wachstum der Routing-Informationen von vielen Routern nicht mehr bewältigt werden könne und man daher rasant auf gravierende Adressierungsprobleme im Netz zusteuere. IAB-Chefin Leslie Daigle informierte die Mitglieder der IETF diese Woche darüber, dass man beim IAB zu dem Schluss gekommen sei, es handle sich um ein echtes Problem. Allerdings sei man der Ansicht, dass genug Zeit bleibe, um über eine koordinierte Lösung zu diskutieren. Das IAB will einen Kreis von Experten (Directorate) berufen; eine oder mehrere IETF-Arbeitsgruppen sollen bei dem für Ende März angesetzten sechsten IETF-Treffen in Prag an den Start gehen.

Motivating example 2: Critical infrastructure



It’s challenging

PAST ESTIMATES OF LINKS IN THE AS-GRAPH.

Paper	Date	Measured	Estimated
Zhang *et al.* [66]	2004-10-24	45,058	55,388
He *et al.* [37]	2005-05-12	47,199	59,500
Mühlbauer *et al.* [45]	2005-11-13	49,241	58,903
Roughan *et al.* [67]	2004-01	38,397	42,818
	2005-01	45,814	54,582
	2006-01	50,129	59,319
	2007-01	57,038	68,856
	2008-01	63,536	76,944
Dhamdhere *et al.* [68]	end of 2007	70,000	-

[M. Roughan et al., “10 Lessons from 10 Years of Measuring andModeling the Internet’s Autonomous Systems,” IEEE JSAC 2011.]

It’s challenging

Year/Methodology	Est. # of customer-provider links in the Internet	Est. number of peering links in the Internet
2008 (BGP)*	~60,000	~15,000
2010 (BGP + traceroute)**	~90,000	~30,000
2012 (ground truth from a large IXP)***	~90,000	**>200,000**

[Slide from Philipp Richter, 2018]


MOMMY, WHO MAKES THE INTERNET
THE ELVES!
WOW REALLY? EVEN THE ROUTING?
SURE!
OMG WTF IS THIS SHIT.
BGP ROUTING FOR DUMMIES

The Backbone Control Protocol

BGP ROUTING

Border Gateway Protocol (BGP) in a nutshell

BGP is a **path vector protocol** (details see RFC 4271)

BGP routers exchange path vectors with BGP neighbors (**peers**)

- Typically, neighbors are connected directly or via a switch
- Multihop BGP peering: Neighbors need not be topologically adjacent

BGP peers accept or discard paths based on **policies**

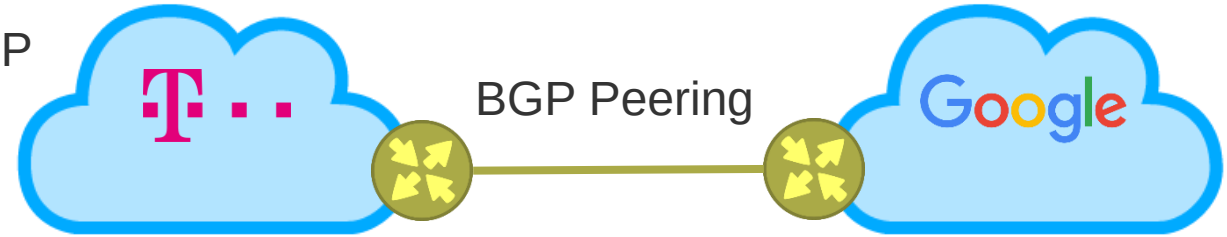
- e.g., preferred neighbors, hot potato, shortest paths

BGP router decides on outgoing advertisements based on policies

BGP is the “glue” of the Internet

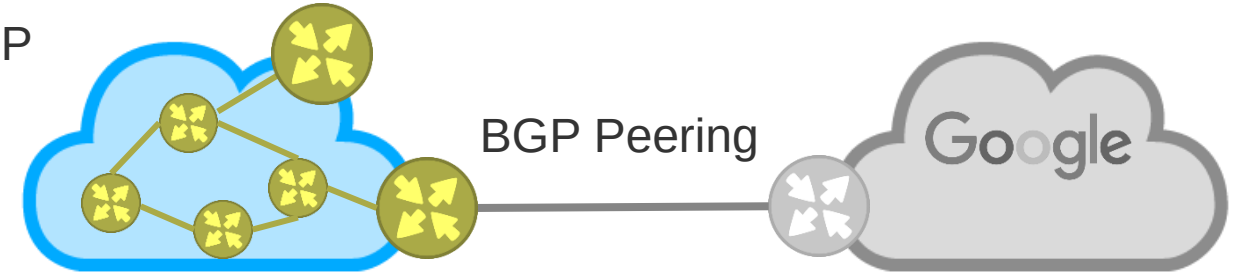
Two faces of BGP

Exterior BGP
(eBGP)

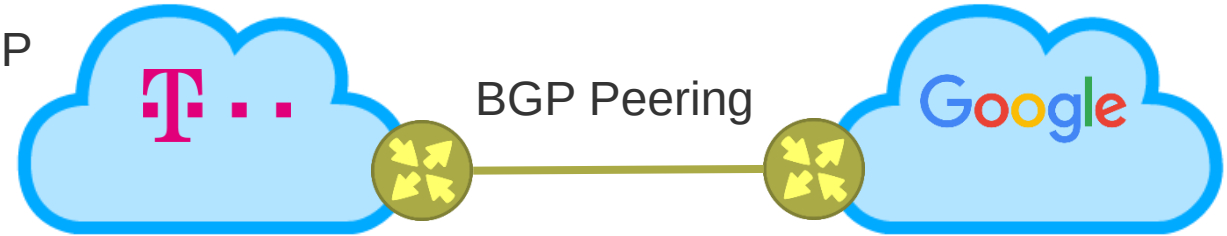


Tow faces of BGP

Interior BGP
(iBGP)



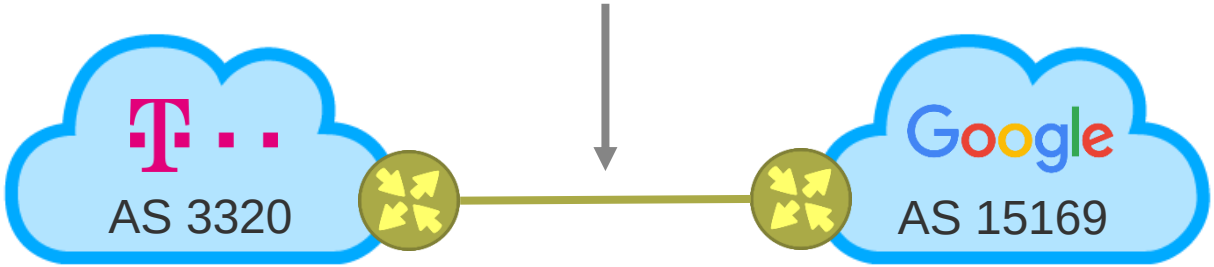
Exterior BGP
(eBGP)



BGP implements three basic three steps

Learning
prefixes

First, establish TCP connection (port 179) in which routers exchange BGP messages



Four BGP message types

Open

Establishes a BGP session

Update

Announce or withdraw prefix(es)

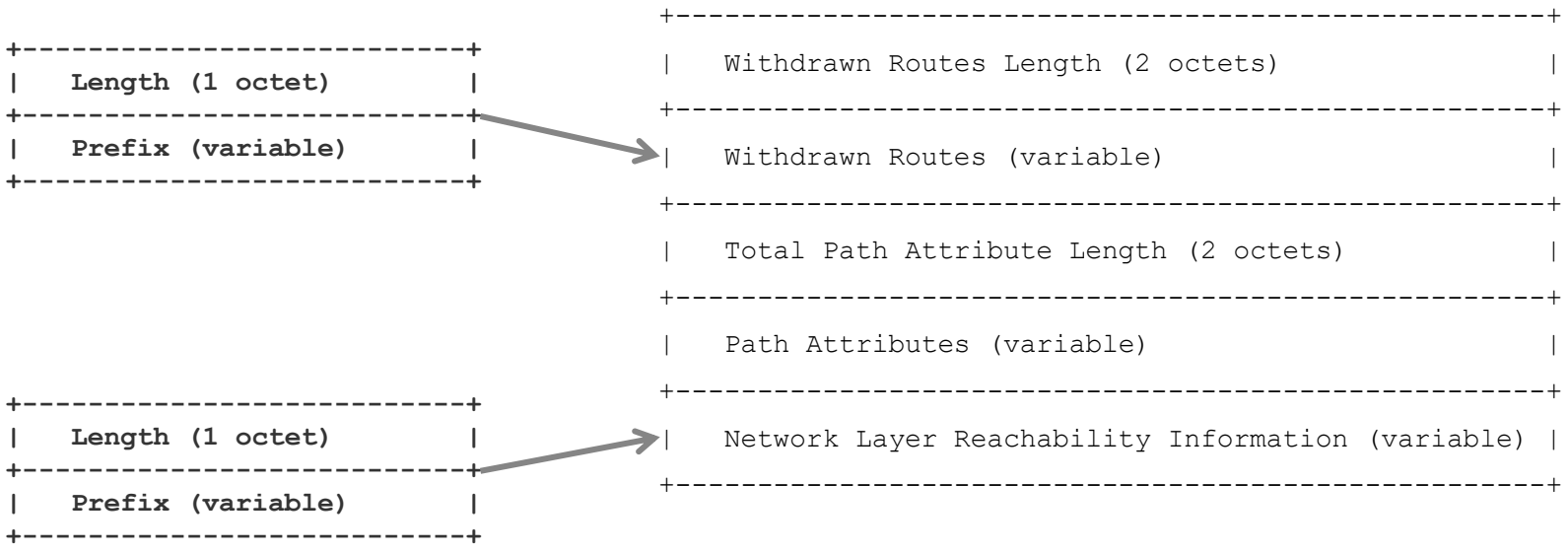
Keepalive

Check if adjacent peers are still available

Notification

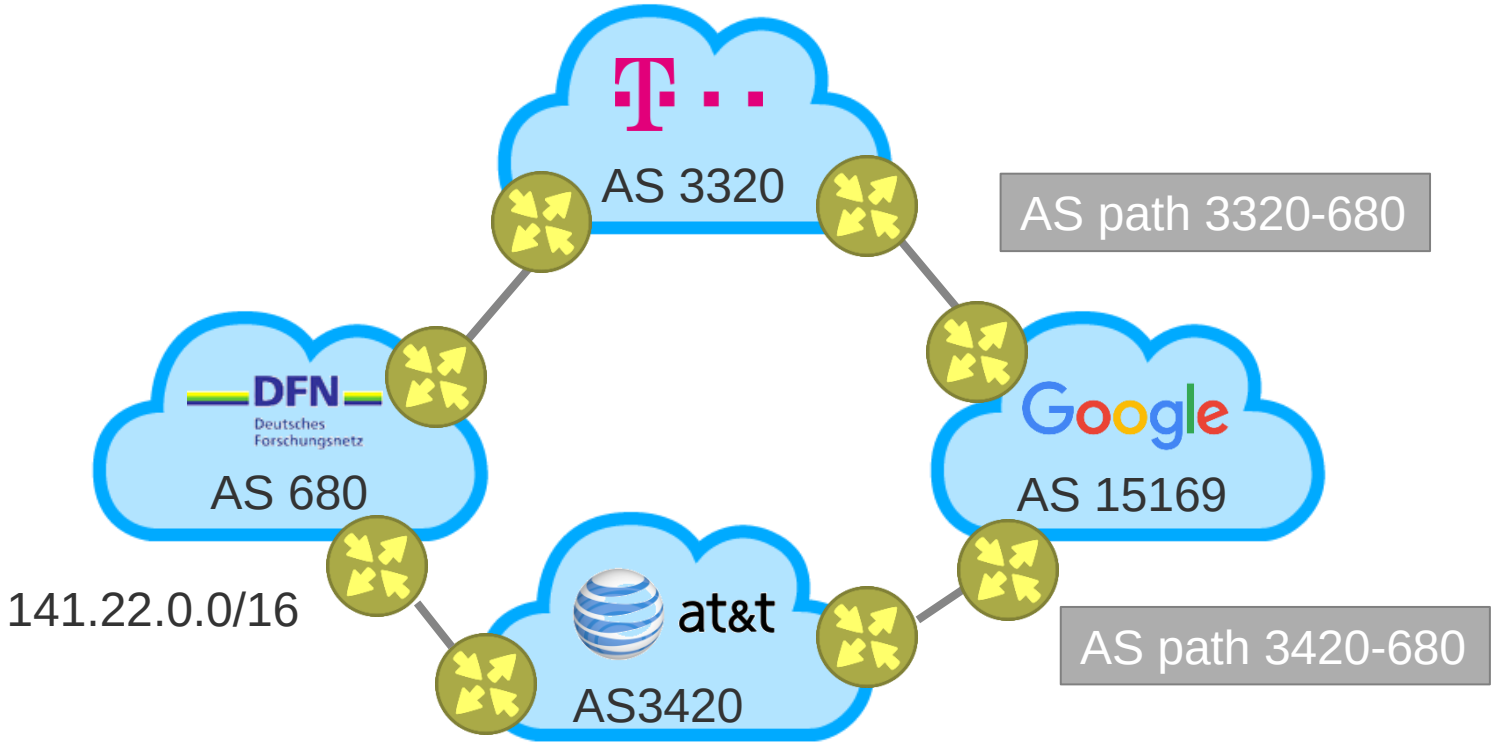
Handle errors

A closer look into a BGP update message



Path attributes describe the way to the prefix(es)

mandatory	Origin of the path information IGP, EGP, Incomplete
mandatory	AS_PATH, sequence of ASNs the update traversed e.g., 3320, 780
mandatory	NEXT_HOP, (unicast) IP address of the router that should be used as the next hop to the destinations
optional, non-transitive	MULTI_EXIT_DISC (MED), discriminate among multiple entry points to a neighboring autonomous system
optional, non-transitive	LOCAL_PREF, inform other internal peers of the advertising speaker's degree of preference for an advertised route
optional, transitive	Communities, aid in policy administration and reduce the management complexity


AS 3320
AS path 3320-680
DFN
Deutsches Forschungsnetz
AS 680
Google
AS 15169
141.22.0.0/16
at&t
AS3420
AS path 3420-680

Example: BGP RIB Entry

TIME: 2008-7-1 02:36:49

TYPE: MSG_TABLE_DUMP/AFI_IP6

VIEW: 0 SEQUENCE: 2702

PREFIX: 1.2.0.0/19

ORIGINATED: Mon Jun 30 10:29:18 2008

FROM: 2001:0418:0000:1000:0000:0000:0000:f000 **AS15169**

AS_PATH: 15169 20 10

MULTI_EXIT_DISC: 1

COMMUNITIES: 15196:420 15169:2000 15169:3000

BGP implements three basic three steps

Learning prefixes

Selecting best path

Exporting prefixes

Which path to use?

Controls outbound traffic

Phase 1: Calculation of Degree of Preference

Based on local policies and attributes, a preference will be assigned to all RIB entries

Phase 2: Route Selection

For each IP prefix, select one best route

Phase 3: Route Dissemination

Before RIB entries will be announced to neighbors, they will be filtered again based on **local policies**

Route selection in detail

BGP speaker identifies the route that has:

a) the highest degree of preference of any route to the same set of destinations, or

b) is the only route to that destination, or

c) is selected as a result of the tie breaking rules

Speaker shall install that route in the local RIB

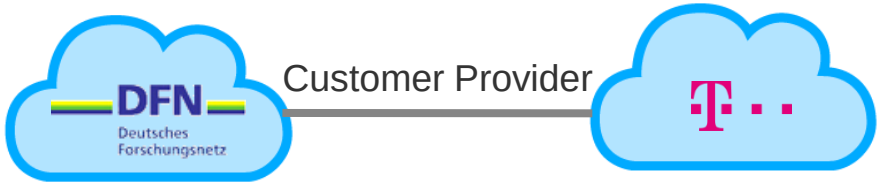
Tie breaking rules

a) Shortest AS path wins
b) Prefer IGP over EGP over incomplete
c) Lowest MED wins
d) Prefer eBGP over iBGP
e) Lowest IGP metric to BGP NEXT_HOP wins
f) Oldest path wins*
g) Lowest speaker ID wins
h) Lowest peer IP address wins

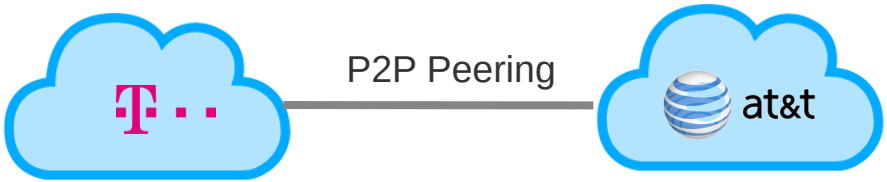
* Not defined in RFC 4271 but implementation practice.

How do you assign local preferences?

Types of peering relations: Two main business relations

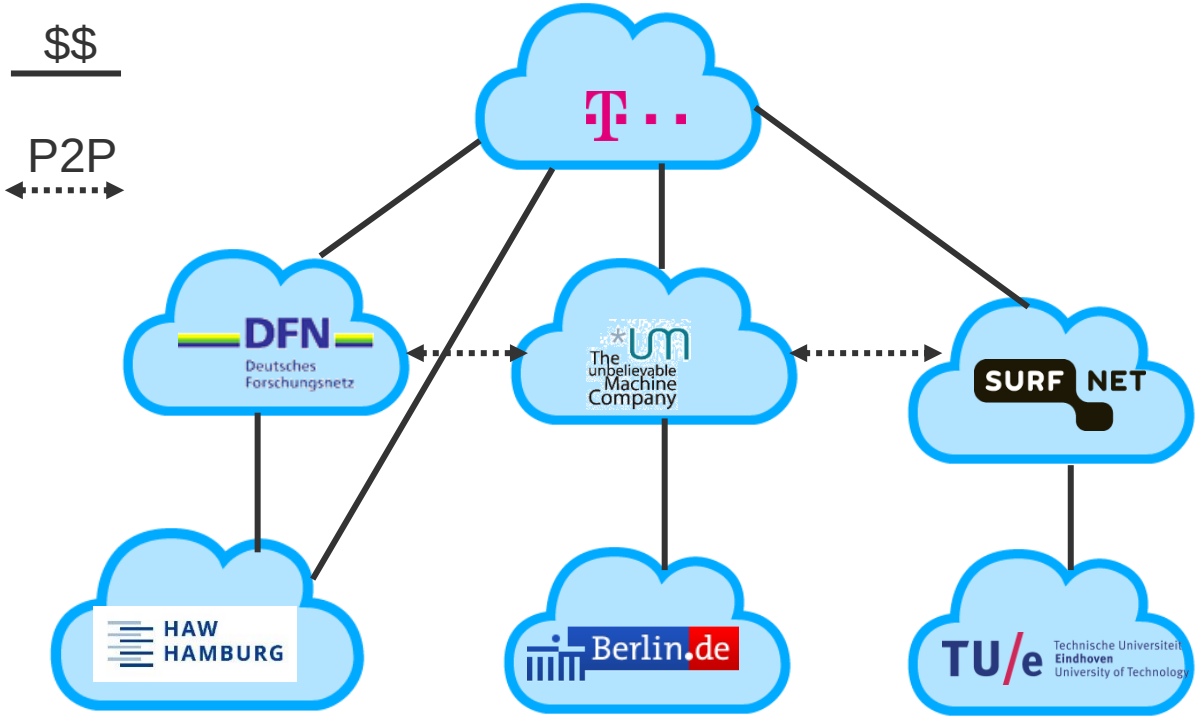


Customer pays $$ provider to get Internet connectivity



Peers don’t pay each other, exchange own and prefixes of customers

Which routes do you prefer?



Rule of thumb, which works most of times

For a destination prefix, prefer routes from

customers over

peers over

providers

BGP implements three basic three steps

Learning prefixes

Selecting best path

Which path to use?

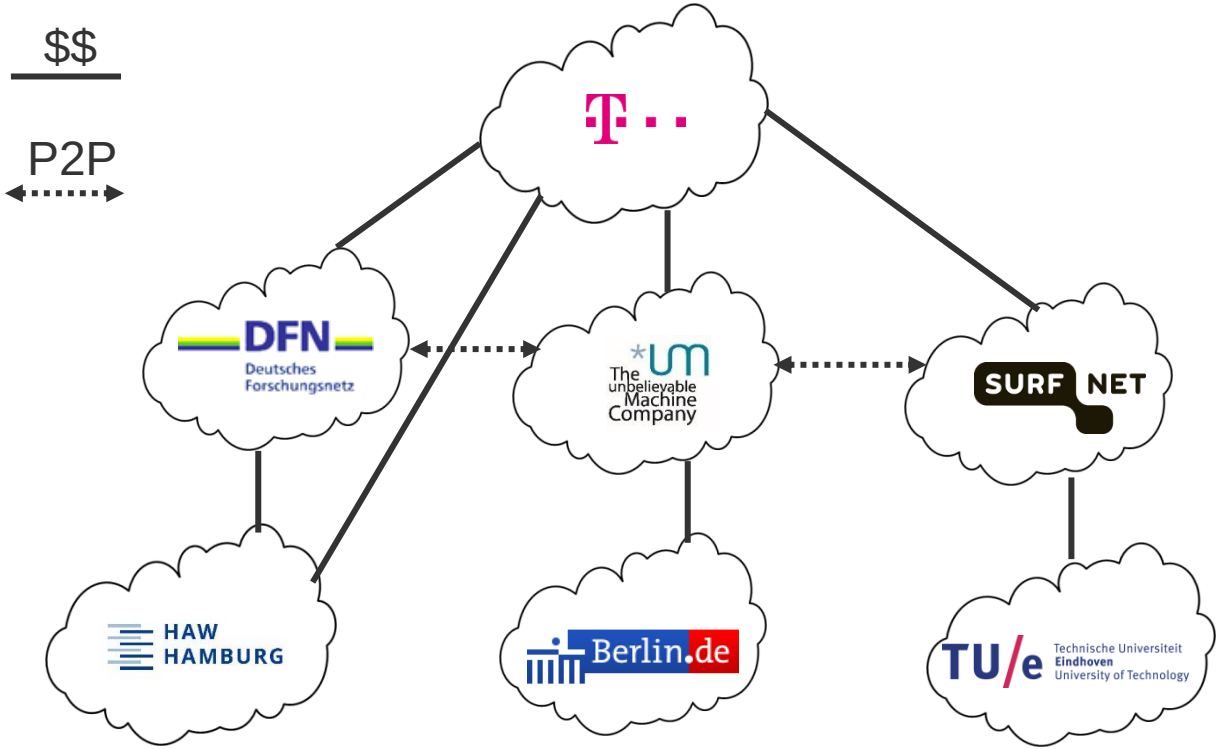
Controls outbound traffic

Exporting prefixes

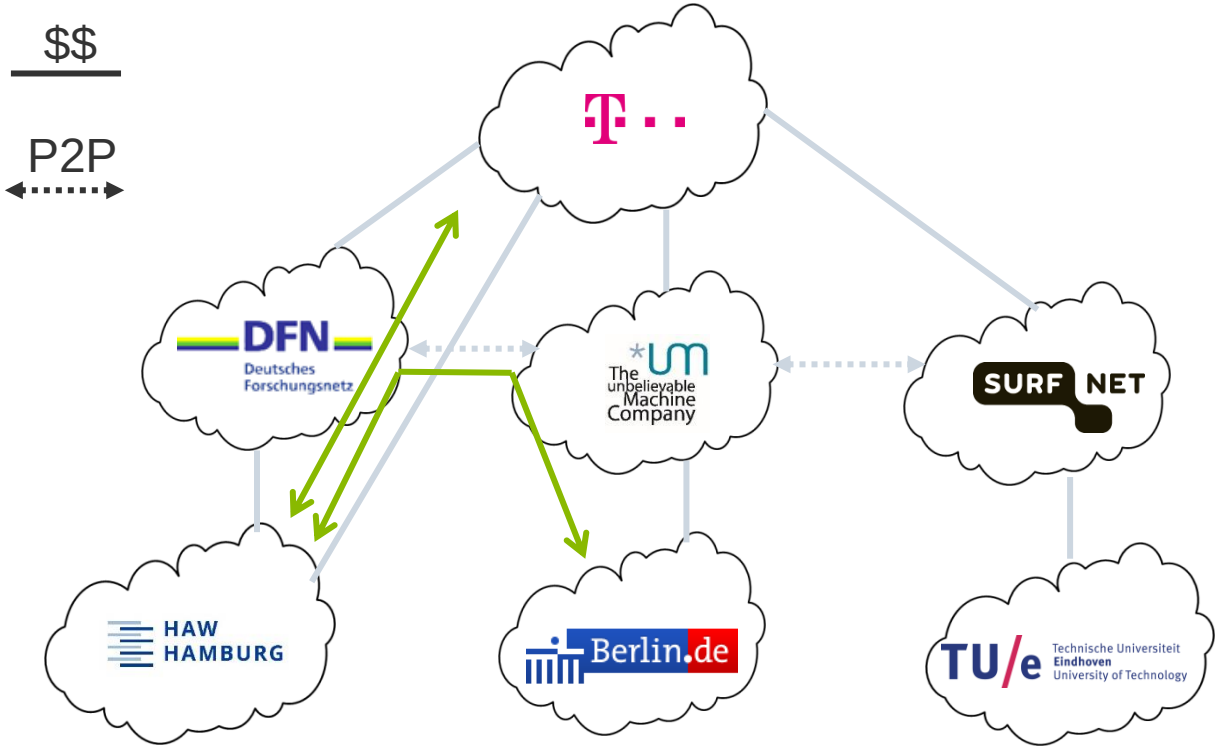
Which path to advertise?

Controls inbound traffic

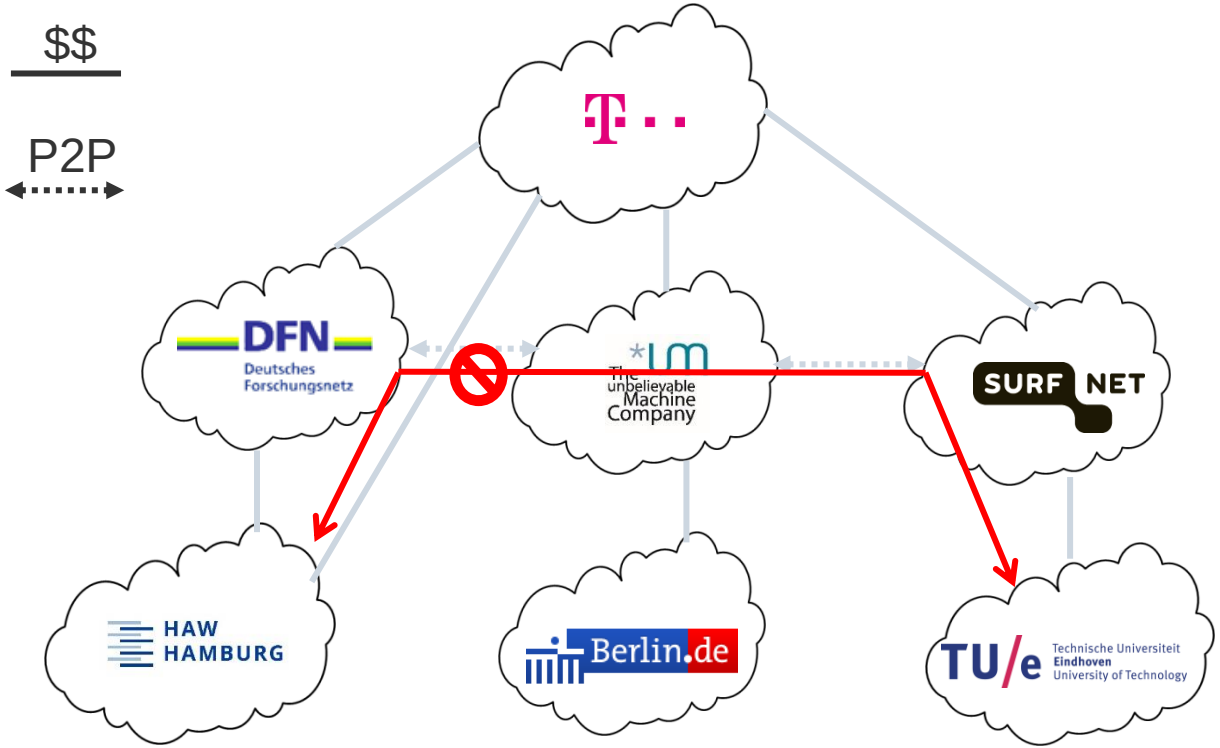
Which routes do you export to whom?



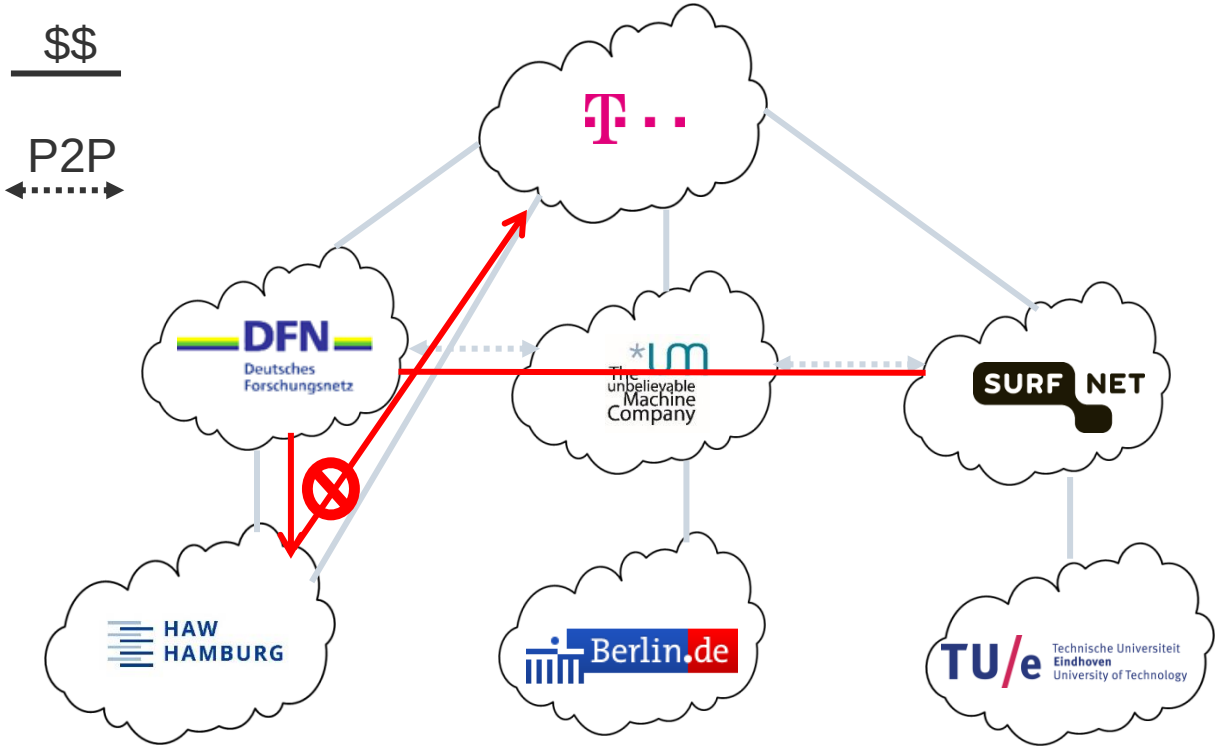
Providers transit traffic for their customers



Peers don’t transit between each other



Customers don’t transit between providers ($$)



Common exporting rules

Announce to

Prefixes from	Customer	Peer	Provider
Customer	Yes	Yes	Yes
Peer	Yes	-	-
Provider	Yes	-	-

Rule of thumb: Follow the money.

Attention

Common import and export do not always hold

Why do researchers still mostly use them?

Simplicity.

What did we learn from the protocol design?

BGP is a highly scalable and expressive information hiding protocol by design.

To keep BGP scalable, only best paths are distributed.

To hide internal network structure, BGP allows operators to exchange routing data w/o revealing information about their own network.

To allow full configuration flexibility w/o conflicting with business, BGP hides policies.

Observing the Wild

BGP MONITORING

BGP data sources

BGP Looking glasses

Routing information services

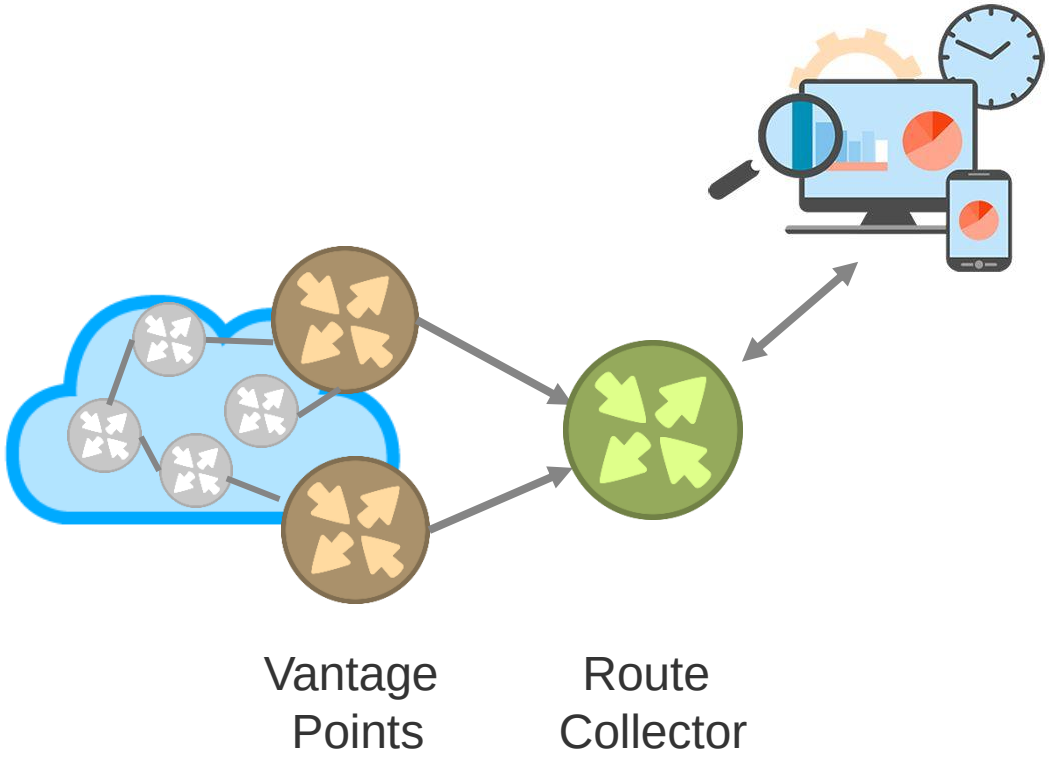
Route Views

RIPE RIS (Routing Information Service)

How does a route collector work?

Route collector peers with other ASs: Vantage Points

- Takes role of a customer
- Should receive full BGP tables
- Route collector receives the view of how the Vantage Point sees the Internet
- Route collector grants access to its table



Data from BGP collectors has limitations

Collectors sees what the vantage point sent.
We do not see the Internet as seen by the connected router.

Type of collected information is not always the same.
Most feeds are “full-feed”, some are “partial feeds”.

Some ASes span multiple continents, you usually see data from one peer.
Most operators keep traffic local, which may result in different views.

Most collectors are biased towards core networks and IXPs.

Connections between VPs and monitors are not 100% reliable.

Single triggering event may cause multiple updates to be observed at the collector.

Various artifacts appear in the data, which may conflict with the actual inter-domain topology.