

5.1 Übersicht

5.2 Adressen

5.3 Lokale Netze: Bridges und Switches

5.4 Intra-Domain Routing

5.5 Inter-Domain Routing

5.6 Internet Protocol (IP)

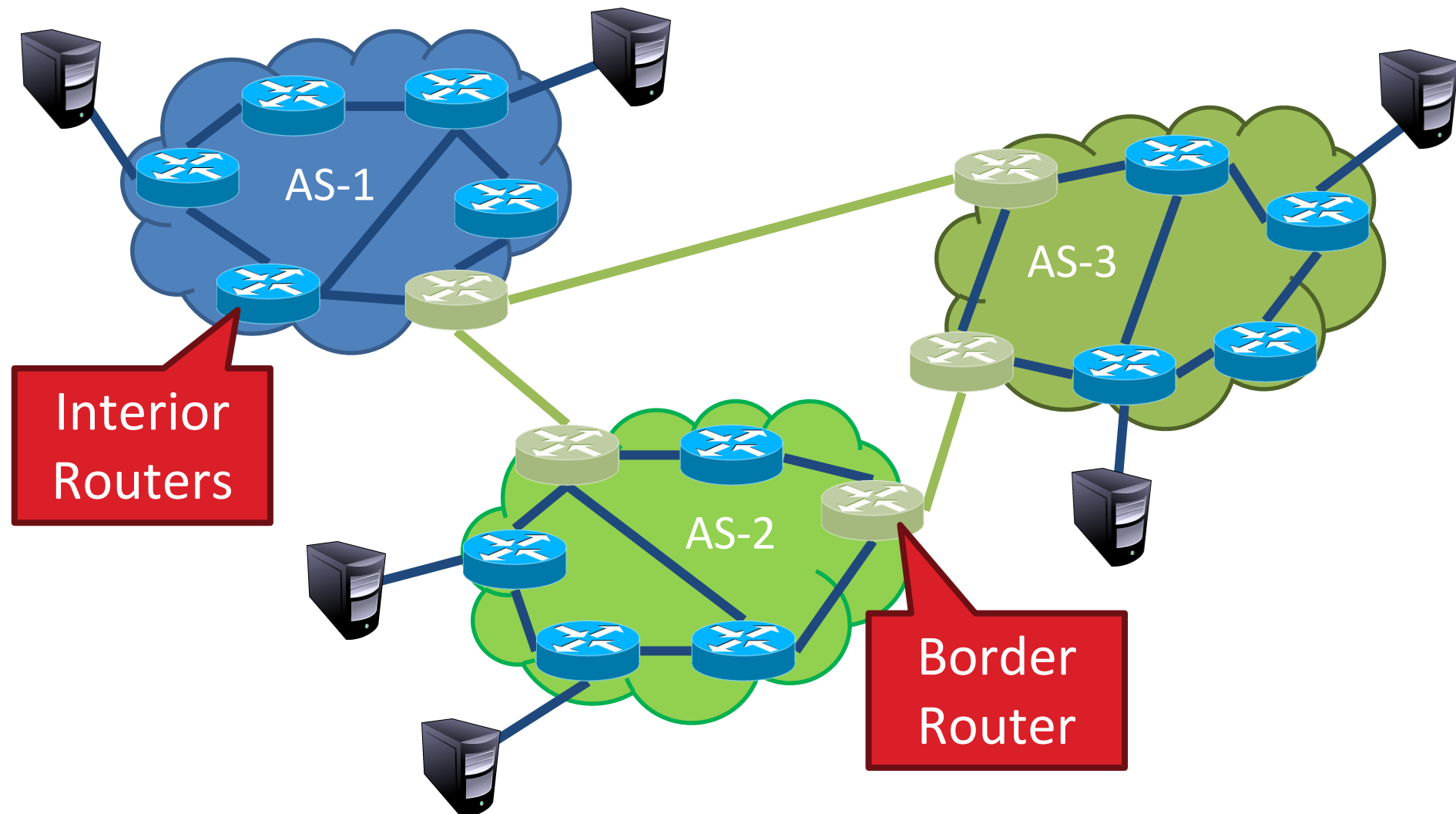
5.7 Network Address Translation (NAT)

5.8 IPv6

5.9 Mobilitätsunterstützung

5.10 Zusammenfassung

Intra- und Inter Domain Routing



Warum Aufteilung in Autonome Systeme?

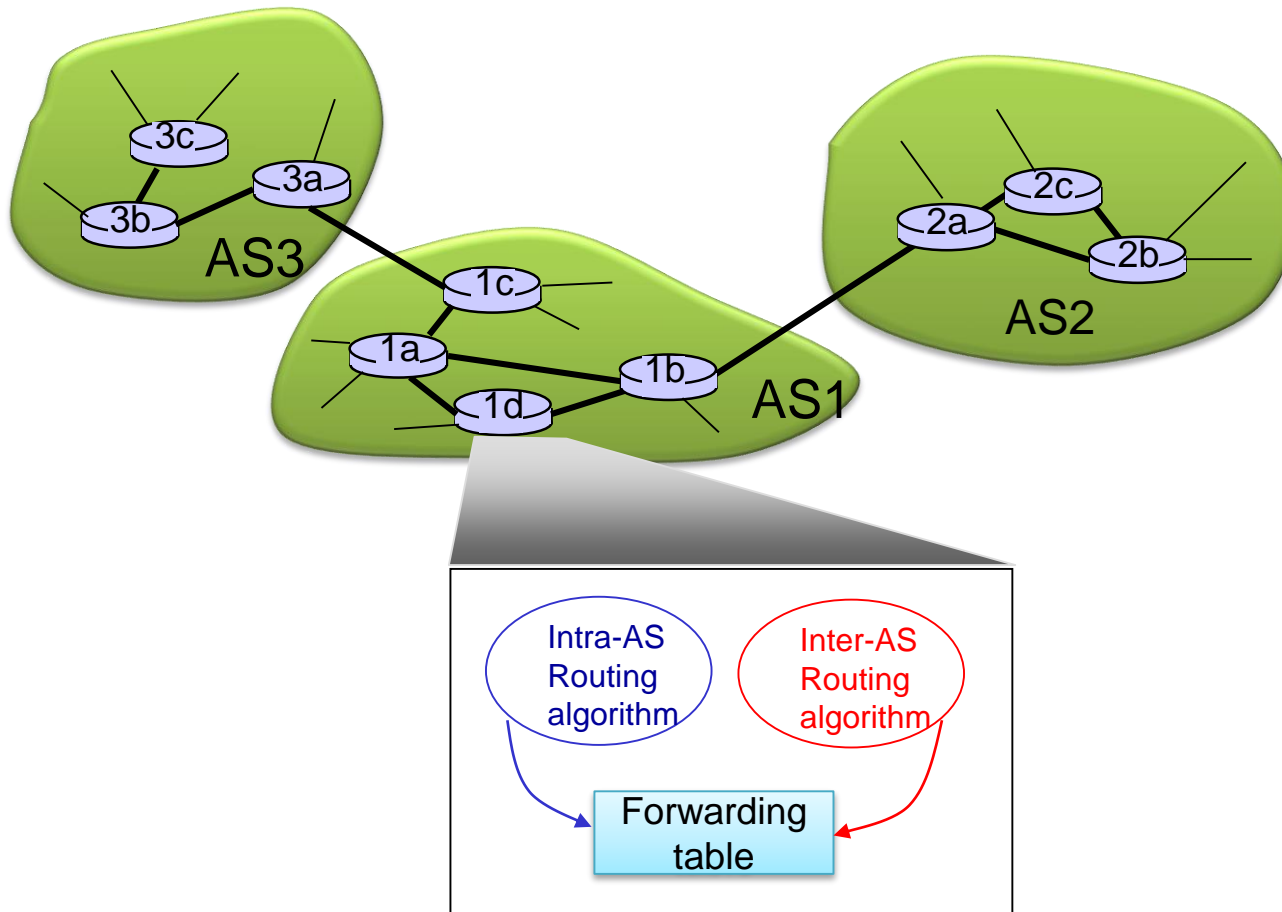
- Geringere Komplexität, geringerer Management-Overhead
- Höhere Flexibilität, größere Individualität
- Strukturierte Zuweisung von Internet-Adressen
 - effizient möglich in einem AS
 - keine hierarchische Struktur im weltweiten Internet
- Routing Algorithmen können nicht effizient für das ganze Internet ausgeführt werden
 - Abstraktionsebenen reduzieren die Komplexität
- Unterschiedliche Routing-Metriken innerhalb eines AS und zwischen AS
 - Intra AS: hohe Bandbreite, niedriger Delay
 - Inter AS: geringe Kosten
- Organisationen verbergen ihre interne Struktur

- Jedes AS wird durch eine AS Nummer (ASN) identifiziert
 - 16 Bit
- Derzeit mehr als 20000 ASNs vergeben
 - AT&T: 5074, 6341, 7018, ...
 - Sprint: 1239, 1240, 6211, 6242, ...
 - Northeastern: 156
 - Belwue: 553

Border Router

Routing-Tabelle des Border-Routers wird durch Intra-AS- und Inter-AS-Routing-Algorithmen gefüllt

- Intra-AS-Einträge für interne Ziele
- Inter-AS- & Intra-AS-Einträge für externe Ziele



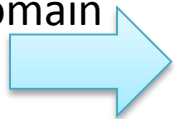
Adress-Typen und resultierende Netze

- Class A: Aufteilung in Subnetze möglich, aber es gibt nur 126 Netze mit mehr als 65534 Adressen
- Class C: einfach zu klein
- Class B: Größe passt, aber es gibt nur etwa 16000
- Routing-Tabellen: zu groß, 2.1 Millionen Einträge pro Router

Class	Prefix Bits	Network Bits	Number of	Hosts per Class
A	1	7	$2^7 - 2 = 126$ (0 and 127 are reserved)	$2^{24} - 2 = 16,777,214$ (All 0 and all 1 are reserved)
B	2	14	$2^{14} = 16,398$	$2^{16} - 2 = 65,534$ (All 0 and all 1 are reserved)
C	3	21	$2^{21} = 2,097,512$	$2^8 - 2 = 254$ (All 0 and all 1 are reserved)
			Total: 2,114,036	

Viel zu groß

Inter-Domain



zu viele Netz-IDs

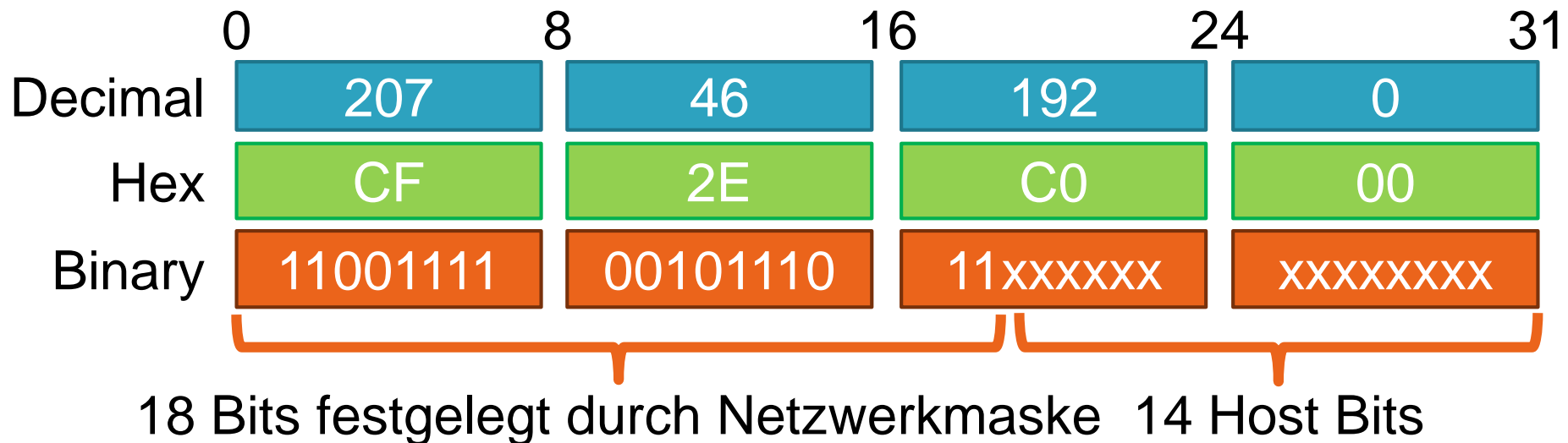
Eigentlich zu klein

CIDR – Classless Inter Domain Routing

- CIDR verallgemeinert das Prinzip der Subnetze auf alle Internetbereiche und führt auch „Supernetting“ ein also das Zusammenlegen mehrere Class A oder Class B-Netze zu einem Präfix
- CIDR, gesprochen ‘cider’
- Grundprinzip:
 - Konzept der “IP Adressklassen” abschaffen
 - Bit-Masken auf allen Routing-Ebenen nutzen
 - insbesondere auch Inter-Domain
- Beliebige Aufteilung zwischen Netzwerk- und Hostadresse
 - festgelegt über Netzwerkmaske bzw. Präfixlänge
 - Beispiel: Northeastern
 - 129.10.0.0 mit **Netzwerkmaske** 255.255.0.0
 - 129.10.0.0 / 16 mit Präfixlänge 16
 - kann jetzt aber auch 3, 7, oder 12 sein

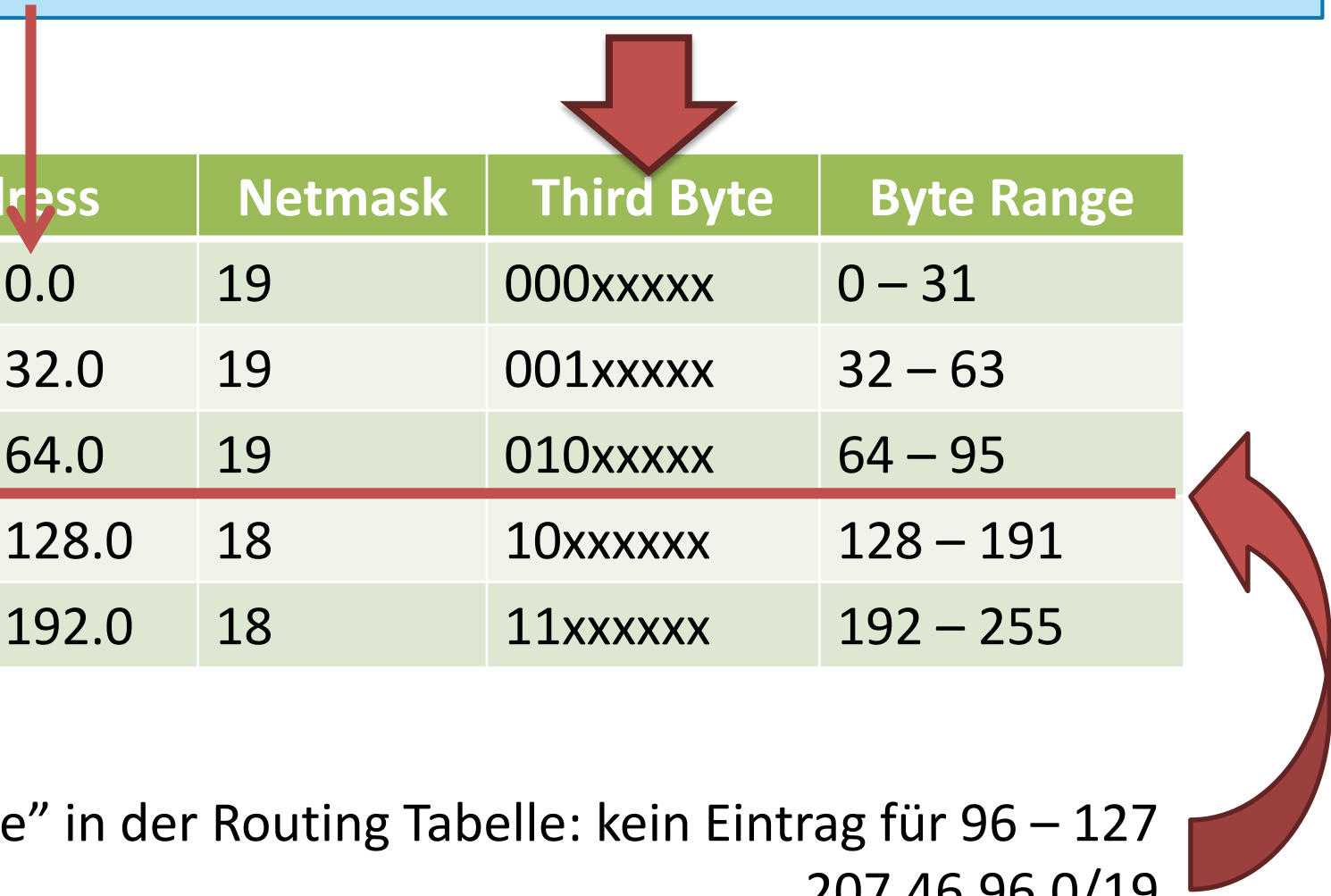
Aggregation mit CIDR

- Ursprünglicher Zweck: Zusammenfassen von Class C Adressbereichen
- Eine Organisation erhält kontinuierlichen Class C Adressbereich
 - Beispiel: Microsoft mit 207.46.192.* – 207.46.255.*
 - entspricht $2^6 = 64$ Class C Adressbereichen
 - oder CIDR Adresse 207.46.192.0/18



Routing Tabelle mit CIDR

- Anzahl der Einträge in Routing-Tabelle wird durch Präfix-Aggregation reduziert
- Lücken in der Routing-Tabelle können entstehen, wenn kein AS eine Route für ein Präfix anbietet

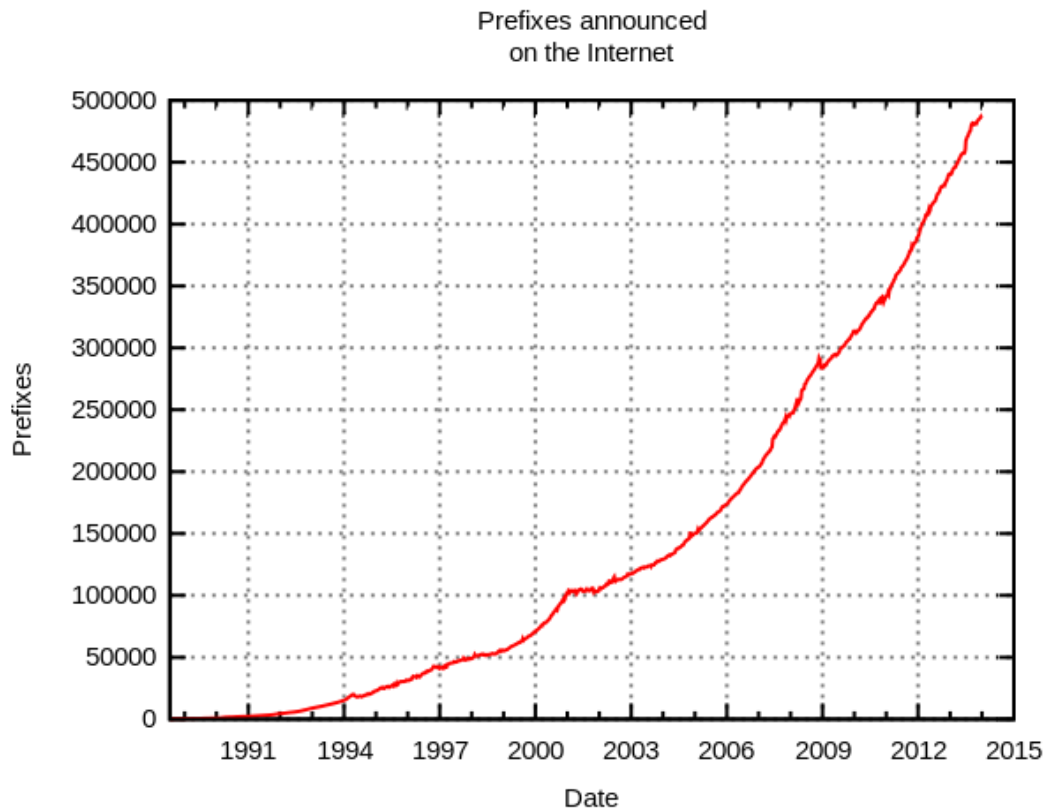


Address	Netmask	Third Byte	Byte Range
207.46.0.0	19	000xxxxx	0 – 31
207.46.32.0	19	001xxxxx	32 – 63
207.46.64.0	19	010xxxxx	64 – 95
207.46.128.0	18	10xxxxxx	128 – 191
207.46.192.0	18	11xxxxxx	192 – 255

“Lücke” in der Routing Tabelle: kein Eintrag für 96 – 127
207.46.96.0/19

Größe von CIDR Routing Tabellen

- CIDR hält die Größe der IP Routing-Tabelle in Grenzen
 - derzeit knapp 500000 Einträge in vollständiger Internet Routing-Tabelle von Backbone Routern
- Trend zu Multi-homing und Mobilität vergrößert die Tabelle
 - im Rahmen von Netz-Generation-Internet-Initiativen (Geni (US), G-Lab (D)) Entwicklung von neuen Internetarchitekturen basierend auf dem “Locator-Identifier-Split)



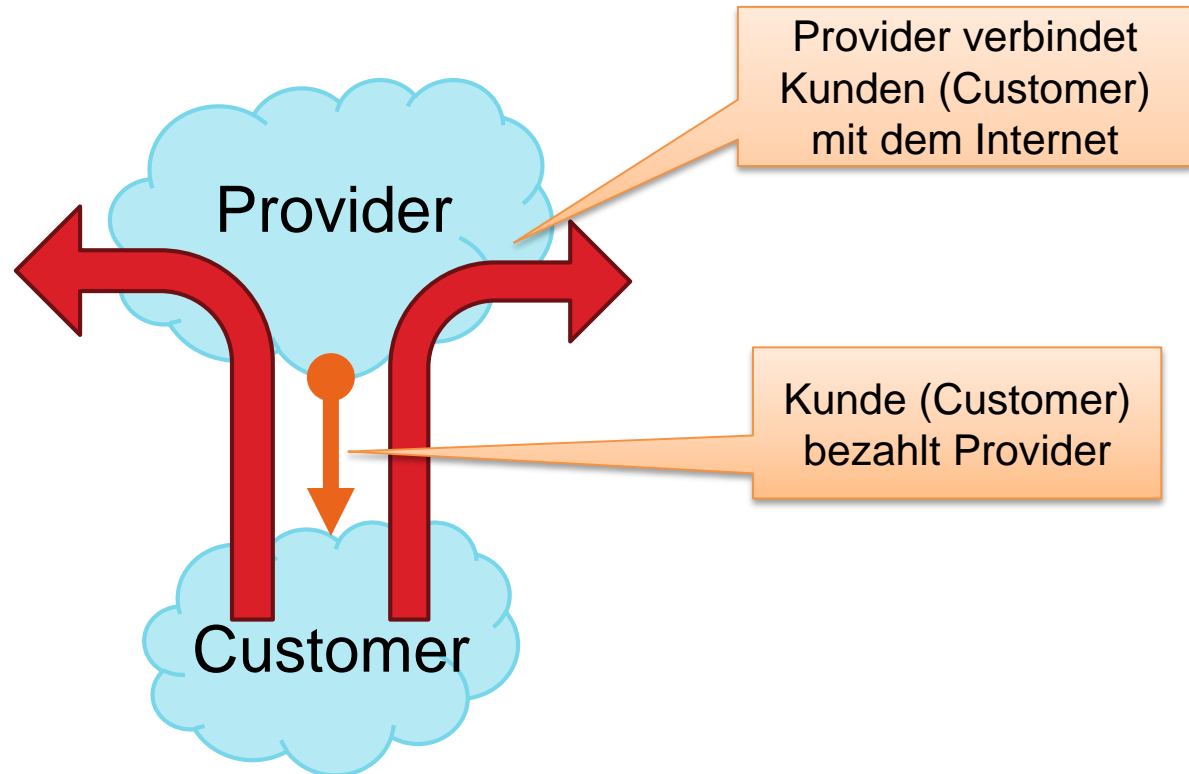
IPv4	Ende 2005	April 2011	April 2012	Mai 2014
Routing-einträge	170.000	360.000	411.000	500.000
Autonome Systeme	26.000	37.000	40.900	47.010

Quelle: wikipedia

- Es geht um globale Konnektivität!!!
 - Alle Border Router in allen Autonomen Systemen **müssen** das gleiche Protokoll benutzen
 - Intra Domain Routing: jedes AS benutzt ein eigenes oder sogar mehrere eigene Routing-Protokolle parallel
- Was sind die Anforderungen?
 - Skalierbarkeit
 - Flexibilität in der Routing-Entscheidung
 - Kosten
 - Fehlerhafte Links und Netze umgehen
- Frage: Link State oder Distance Vector?
 - weder noch: BGP (Border Gateway Protocol) ist ein **Path Vector** Protokoll

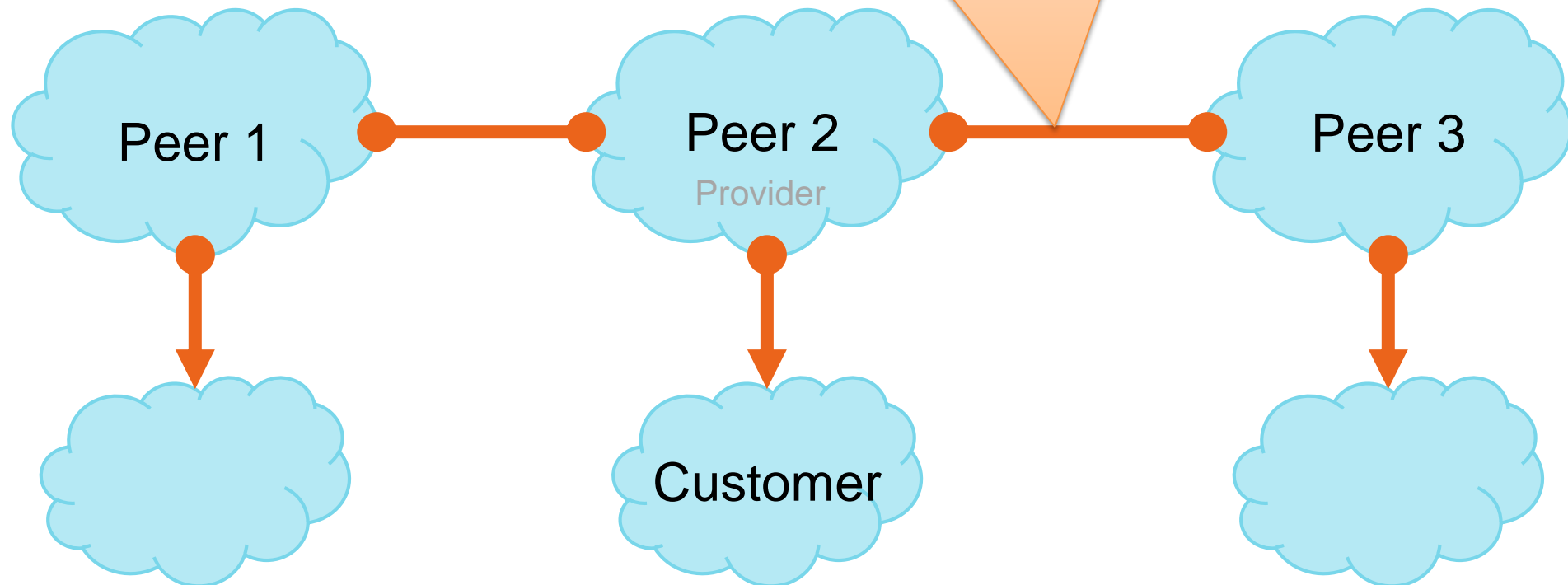
- Border Gateway Protocol
 - de Facto Inter-Domain Protokoll im Internet
 - Routing Protokoll basierend auf “Policies”
 - basiert auf einem Bellman-Ford Pfad-Vektor-Algorithmus
- Recht einfaches Protokoll, aber ...
 - komplex durch Notwendigkeit manueller Konfiguration
 - die ganze Welt sieht die “Advertisements” (lokale Routing-Tabelle)
 - Fehler können globale Auswirkungen haben
 - ... und das globale Internet lahmlegen/beeinträchtigen
 - ... Pakistan hijacked YouTube
 - Routing-Policies sind von ökonomischen Interessen getrieben
 - Wie viel €€€ kosten es, auf diesem oder jenem Pfad zu routen?
 - Performance spielt eine untergeordnete (keine) Rolle
 - kürzester oder bester Pfad nicht das Hauptkriterium

BGP Beziehungen



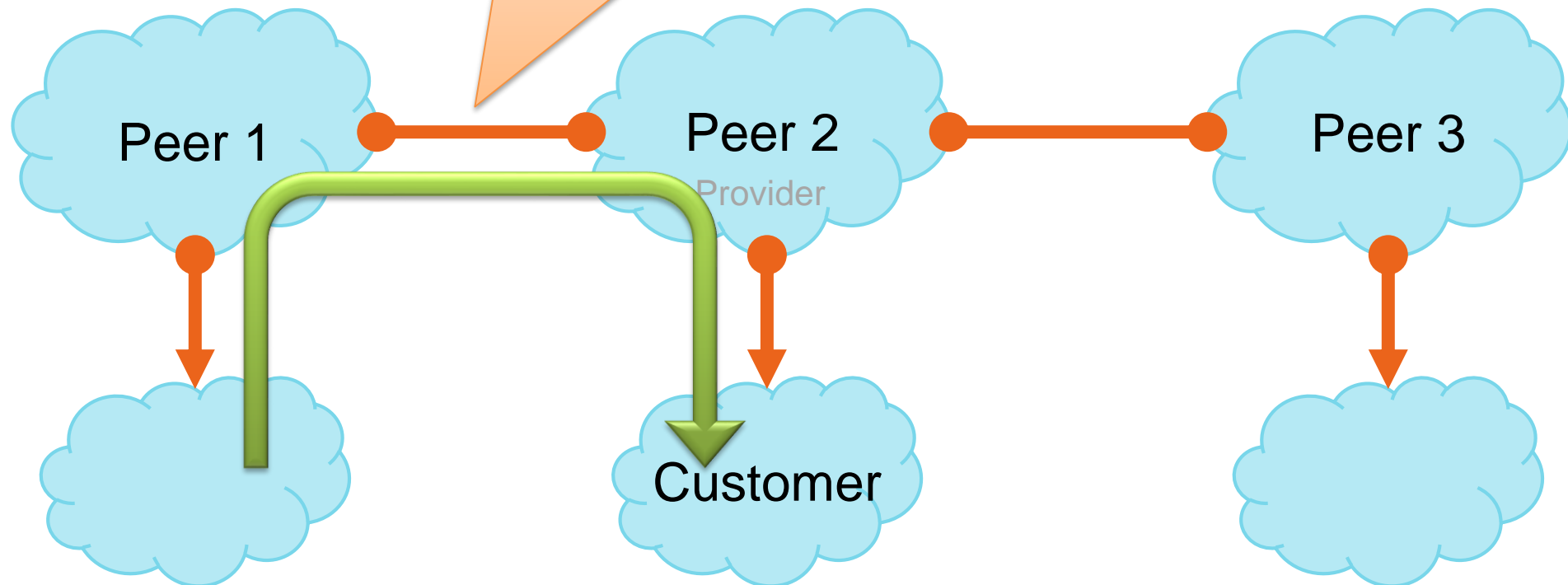
BGP Beziehungen

Peers treffen Abkommen zur kostenlosen Verkehrsweiterleitung. Peers sind ISPs mit ähnlichem Geschäftsmodell z.B. Deutsche Telecom und France Telecom. Grundlage für Peering-Abkommen ist, dass ungefähr die gleiche Verkehrsmenge in beide Richtungen fließt.



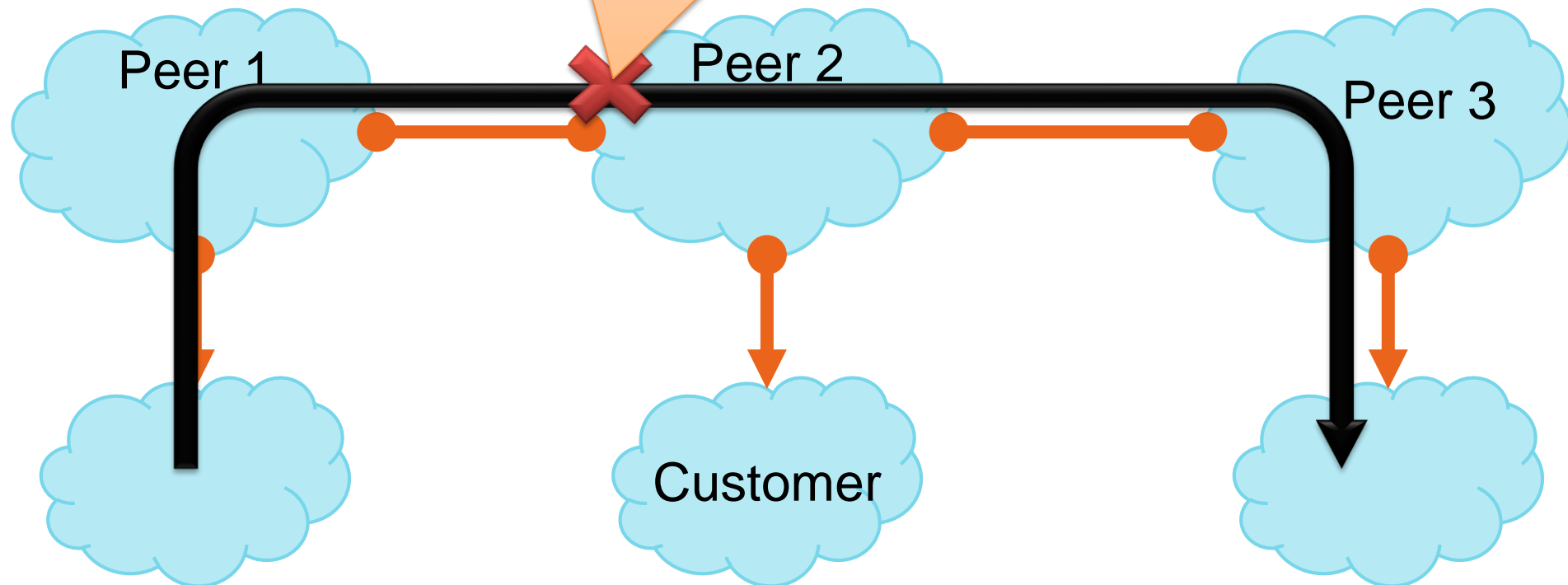
BGP Beziehungen

Peer 2 leitet beispielsweise kostenlos IP Pakete von einem Kunden von Peer 1 an seine eigenen Kunden weiter.

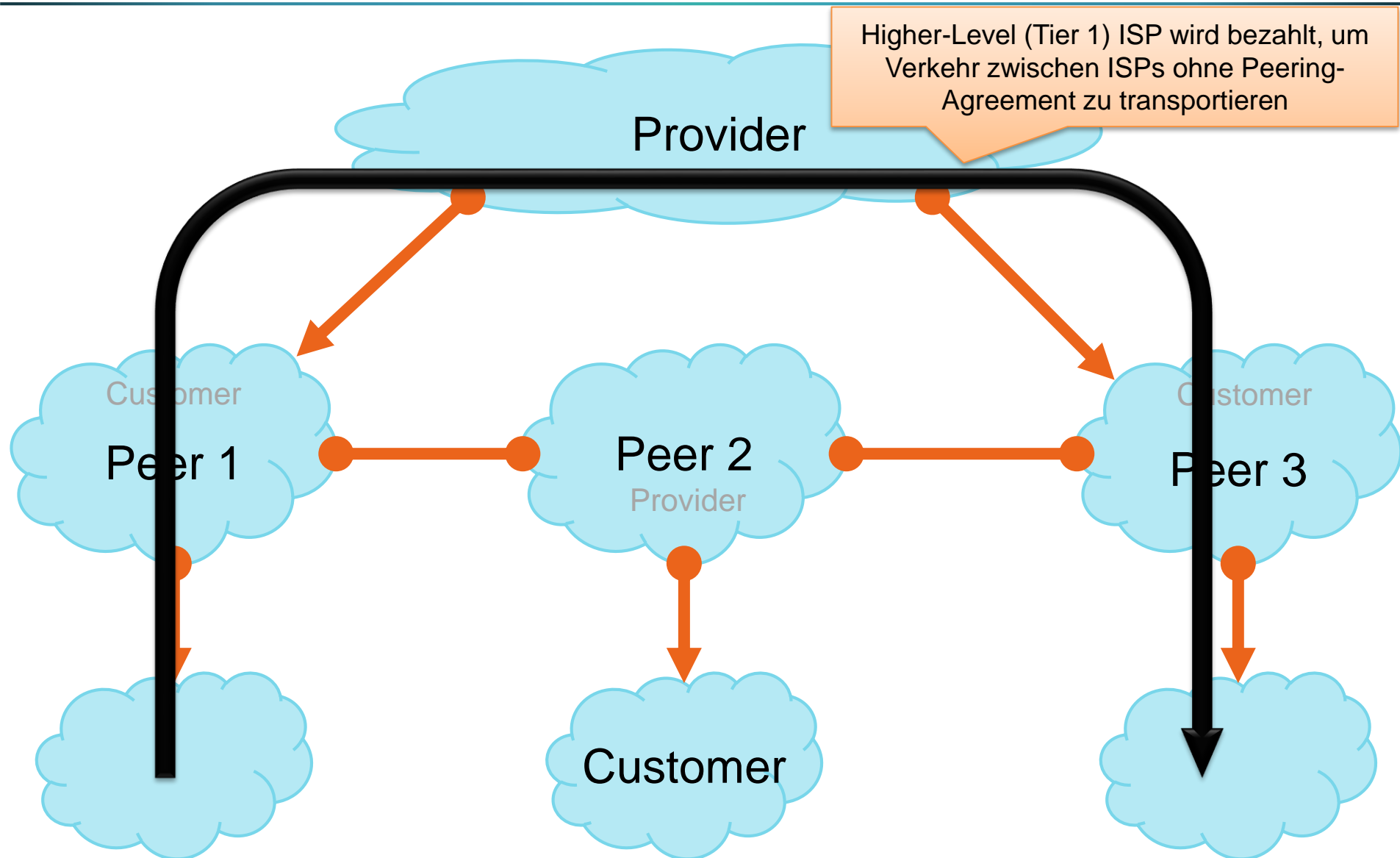


BGP Beziehungen

Peer 2 hat aber keine Veranlassung, Transit-Verkehr von einem Kunden von Peer 1 zu einem Kunden von Peer 3 kostenlos weiterzuleiten. Das ist nicht Teil des kosten-neutralen Peering-Abkommens.

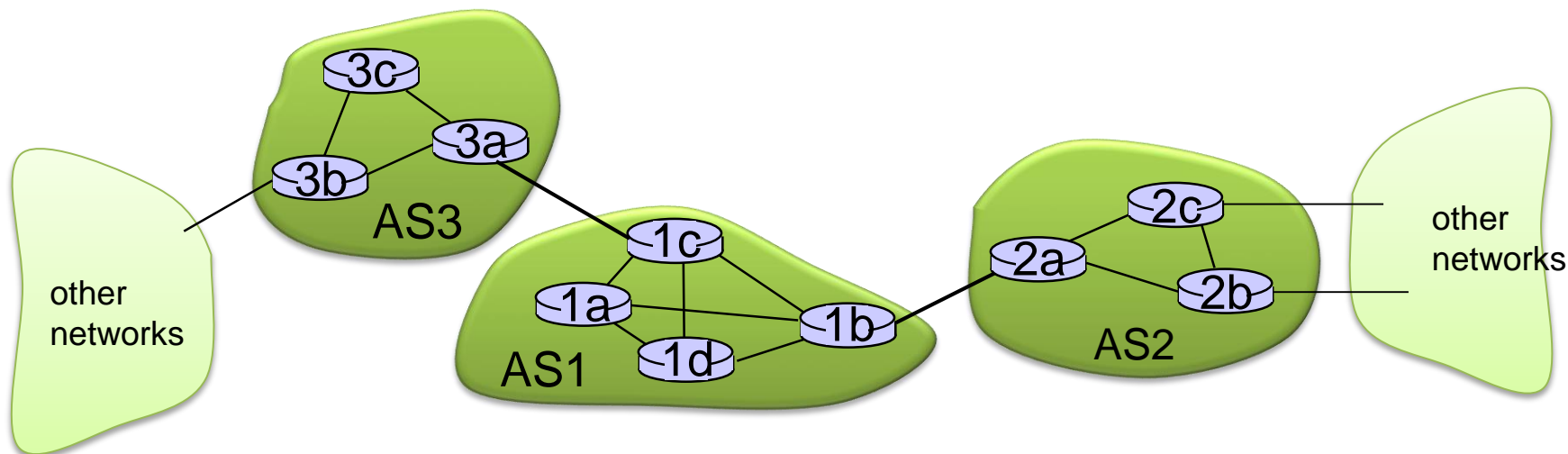


BGP Beziehungen

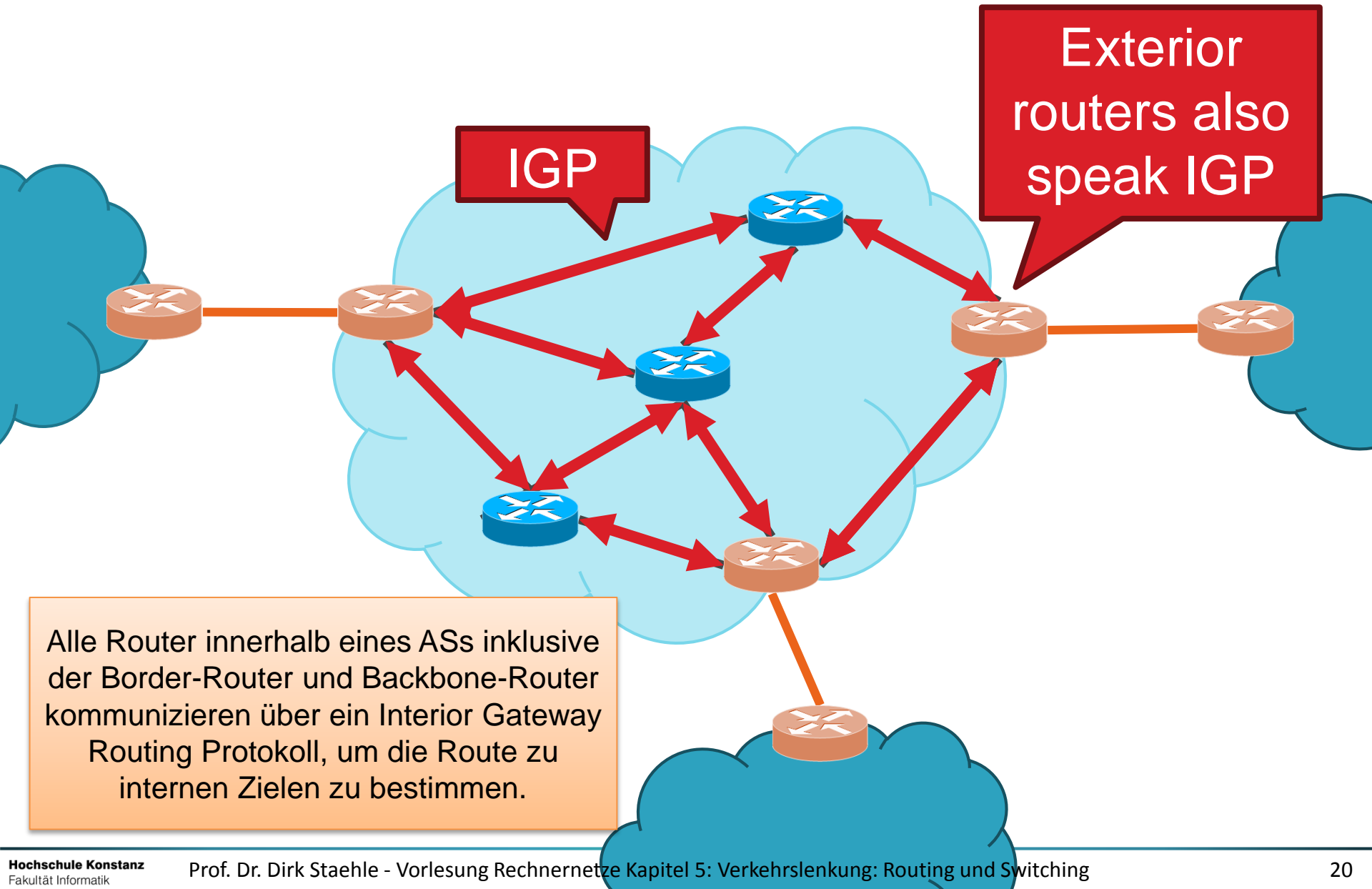


Aufgaben des Inter-AS-Routings

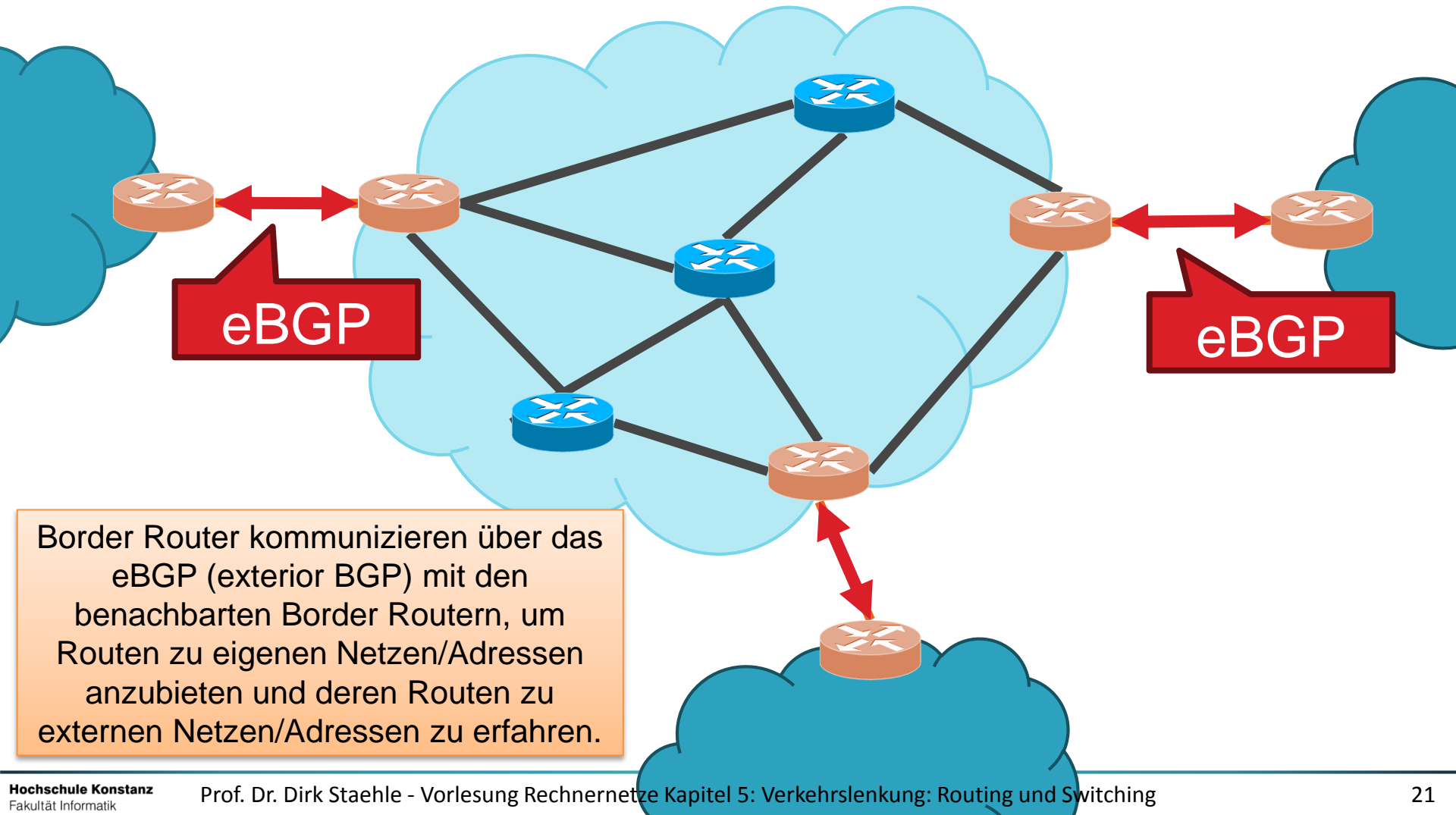
- Wenn ein Router in AS1 ein Paket für ein Ziel außerhalb von AS1 erhält:
 - Router sollte das Paket zu einem der Gateway-Router in AS1 weiterleiten
 - Aber zu welchem?
- Aufgaben des Inter-AS-Routing Protokolls (aus Sicht von AS 1):
 - Herausfinden, welche Ziele über die Autonomen Systeme AS2 und AS3 erreichbar sind
 - Verteilen dieser Informationen an alle Router in AS1



Internes und externes BGP

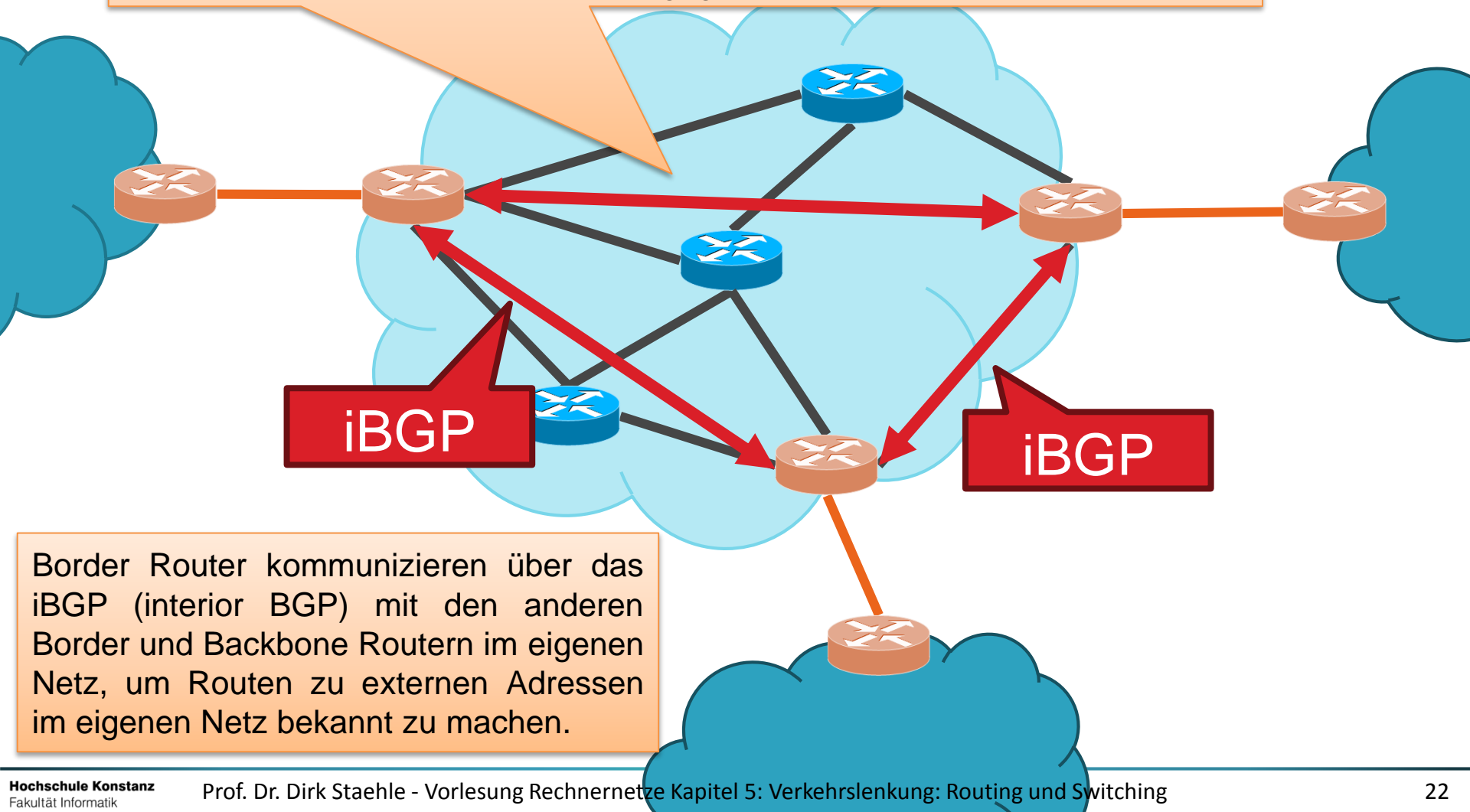


Internes und externes BGP



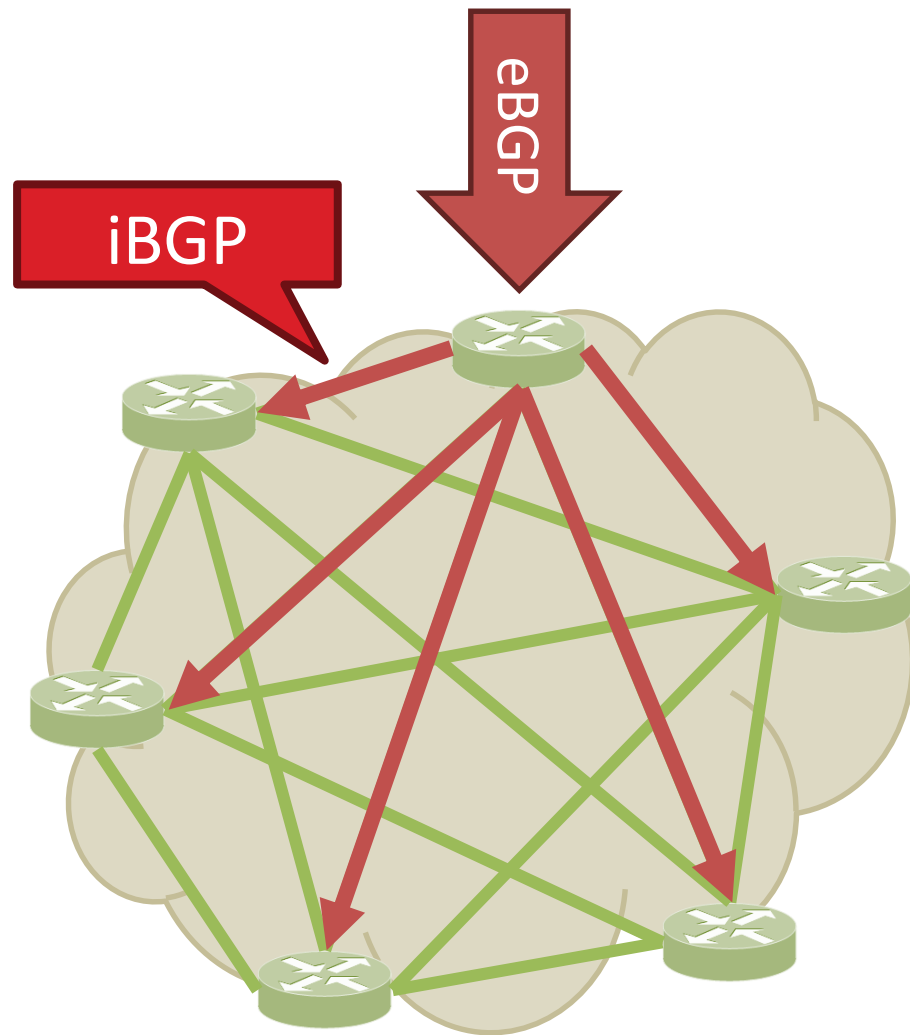
Internes und externes BGP

Border und Backbone Router müssen voll-vermascht sein, d.h. jeder BGP Router kommuniziert direkt mit jedem anderen BGP Router, um Loops zu vermeiden. Routen von iBGP Backbone Routern (nicht eBGP) werden nicht weitergegeben.



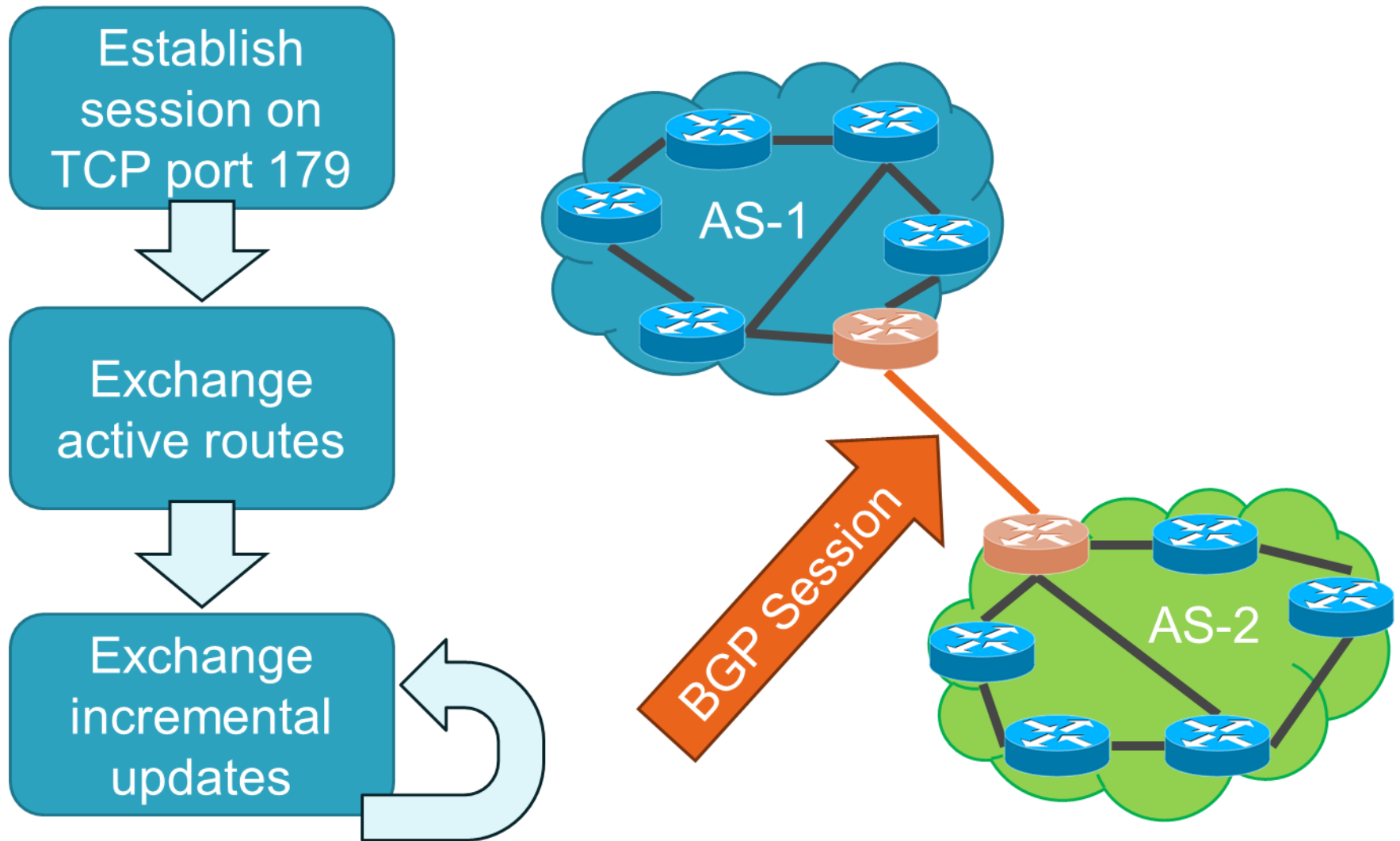
Border Router kommunizieren über das iBGP (interior BGP) mit den anderen Border und Backbone Routern im eigenen Netz, um Routen zu externen Adressen im eigenen Netz bekannt zu machen.

Warum iBGP?



- Prinzip von iBGP
 - eBGP läuft auf Border Routern
 - iBGP läuft auf Backbone Routern
 - Vollvermaschung der Backbone Router
 - Border-Router geben eBGP Routen an Backbone-Router weiter
 - Backbone Router können besten Border Router bestimmen
 - andere Router kennen Default-Route zu einem Backbone-Router und können externe Routen nicht unterscheiden
- Warum wird iBGP benötigt? Es gibt doch Intra-Domain-routing Protokolle z.B. OSPF?
 - OSPF kann keine Informationen über die BGP Policy mitteilen
 - Vermeidung von Intra-AS Routing-Loops

BGP Sessions



- **Open**: Aufbau einer Peering Session
- **Keep Alive**: Handshake in regelmäßigen Abständen
- **Notification**: Abbruch einer Peering Seesion
- **Update**: Ankündigung (Announcement) einer Route/Routenänderung
 - Route Announcement= Versprechen, Pakete mit diesem IP Prefix weiterzuleiten

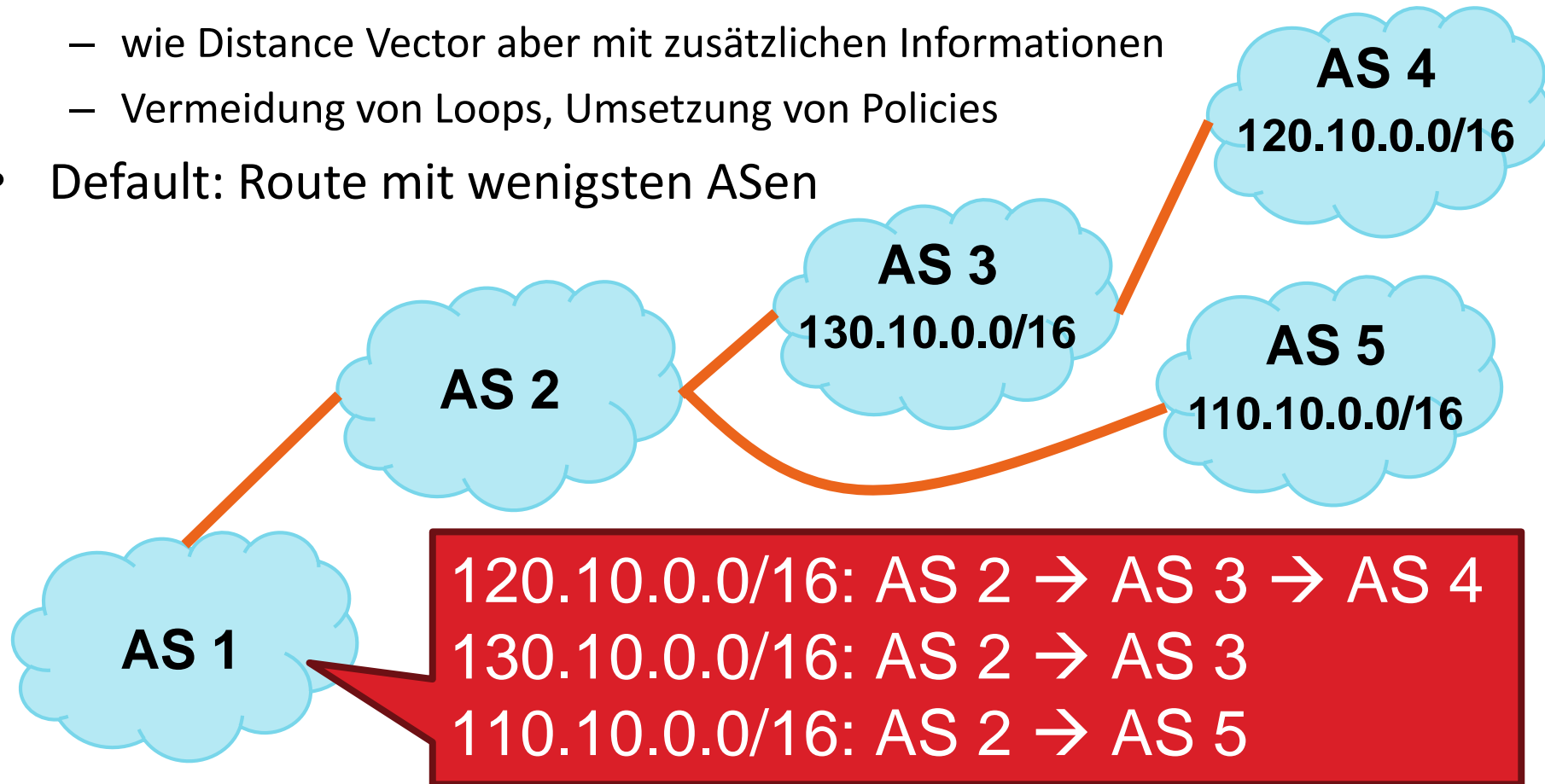
announcement = IP prefix + attribute values

IP Prefix (Destination)	Next Hop	Attribut	Pfad-Vektor
12.6.126.0/24	207.126.96.43	1021	0 6461 7018 6337 11268 i

- Arten der Update-Nachrichten: Neue Route, Änderung der Pfad-Attribute, Zurücknehmen einer Route

Pfad Vektor Protokoll

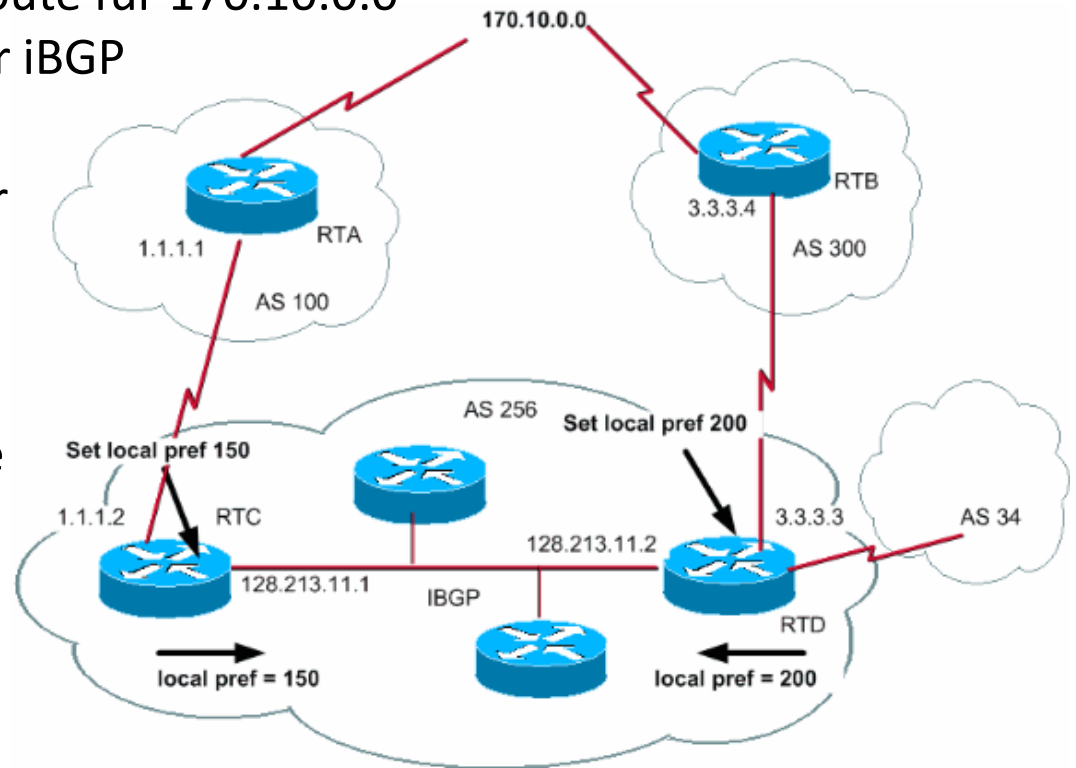
- AS-path: Sequenz von ASen, durch die die Route verläuft
 - wie Distance Vector aber mit zusätzlichen Informationen
 - Vermeidung von Loops, Umsetzung von Policies
- Default: Route mit wenigsten ASen



- Attribute:
 - LocalPREF: Local preference policy
 - Verwendung: interne Pfadauswahl
 - Policy des Netzbetreibers, berücksichtigt Peering- und Transit-Agreements, ökonomische Aspekte, etc.
 - überschreibt Default-Route (wenigste AS Hops)
 - MED: Multi-exit Discriminator
 - Verwendung: mit Peering-Partnern
 - Mitteilung an anderes AS, welcher Border-Router als Eingang bevorzugt wird
- Regeln auf Border Routern:
 - Import Rules: Filter, die spezifizieren, welche Route-Advertisements angenommen werden
 - Beispiel: Vermeidung von Outgoing-Traffic über (teure) Transit-Link
 - Export Rules: Filtern, die spezifizieren, welche Routen an wen weitergeleitet werden
 - Beispiel: Vermeidung von Transit-Traffic

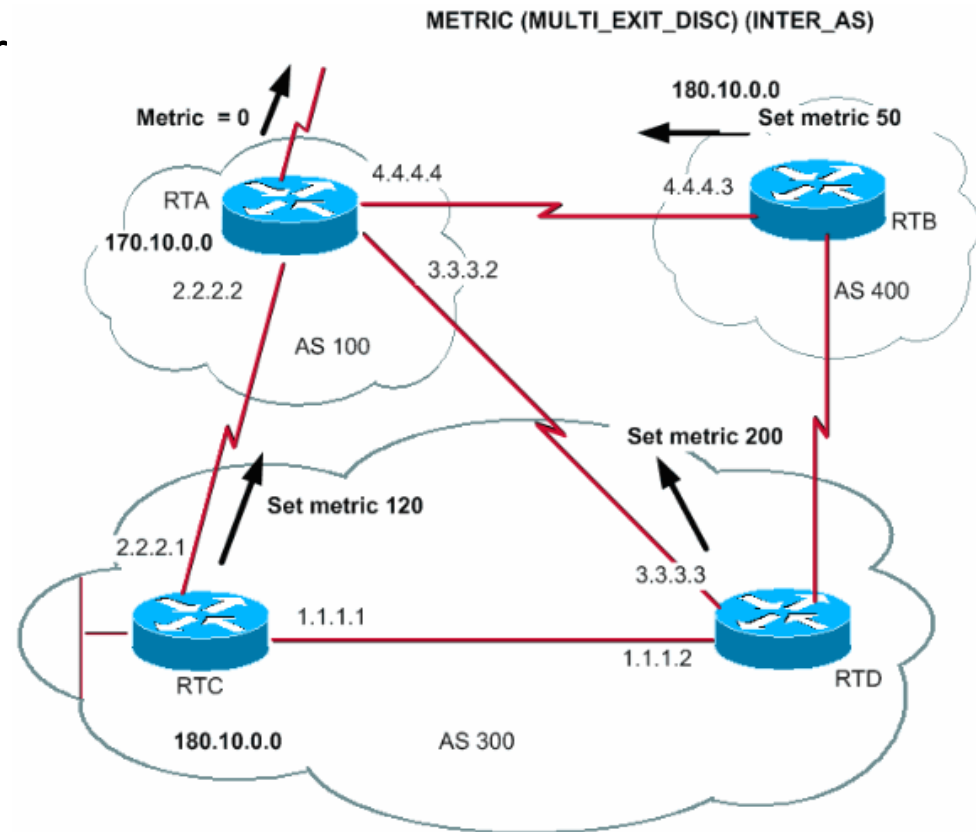
Local Preference

- AS 256 hat zwei Möglichkeiten, den Adressbereich 170.10.0.0 zu erreichen: über AS 100 oder AS 300 und möchte erreichen, dass der ausgehende Verkehr über AS 300 fließt
- Border Router RTC gibt die Route für 170.10.0.0 mit Local Preference 150 über iBGP an das eigene Netz weiter
- Border Router RTD gibt die Route für 170.10.0.0 mit Local Preference 200 über iBGP an das eigene Netz weiter
- die internen Backbone Router leiten Pakete mit Zielen aus 170.10.0.0 an RTD weiter, da dieser die Route mit größerer Local Preference gesendet hat



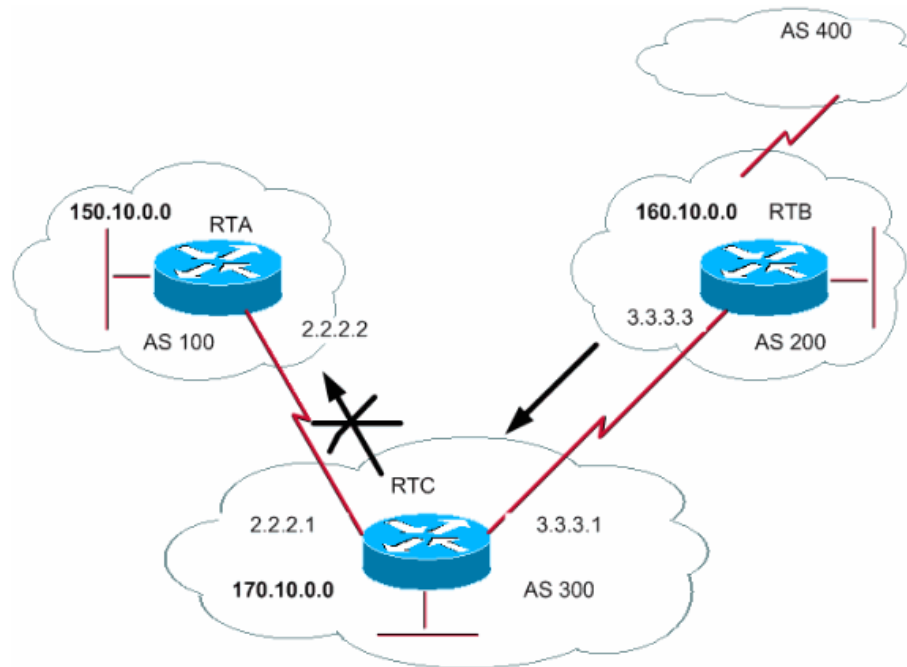
Multi-Exit-Discriminator

- AS 300 ist über zwei BGP Router mit AS 100 verbunden und möchte erreichen, dass AS 100 Pakete über BGP Router RTC und nicht RTD schickt
- Router RTC schickt daher Routen mit MED 120 und RTD mit MED 200 an Border Router von AS 100
- AS 100 schickt den Verkehr über die Route (den Link) zu AS 300 mit niedrigstem MED
- AS 100 gibt die MED Metrik nicht an andere AS weiter sondern setzt den MED auf 0



Beispiel: Export-Filter

- AS 300 gibt keine Route über AS 200 an AS 100 weiter
 - z.B. weil AS 200 und AS 100 beides Peers sind
 - AS 100 soll keine anderen AS über die Route AS 300->AS 200 erreichen
 - alle Routen mit AS 200 als Next-Hop im Pfadvektor werden gefiltert und nicht an AS 100 weitergegeben



Routen Auswahl

- Generell ist die Auswahl der Routen nicht standardisiert sondern hängt von der Konfiguration in einem AS ab



Highest Local Preference

Enforce relationships

Shortest AS Path

Lowest MED

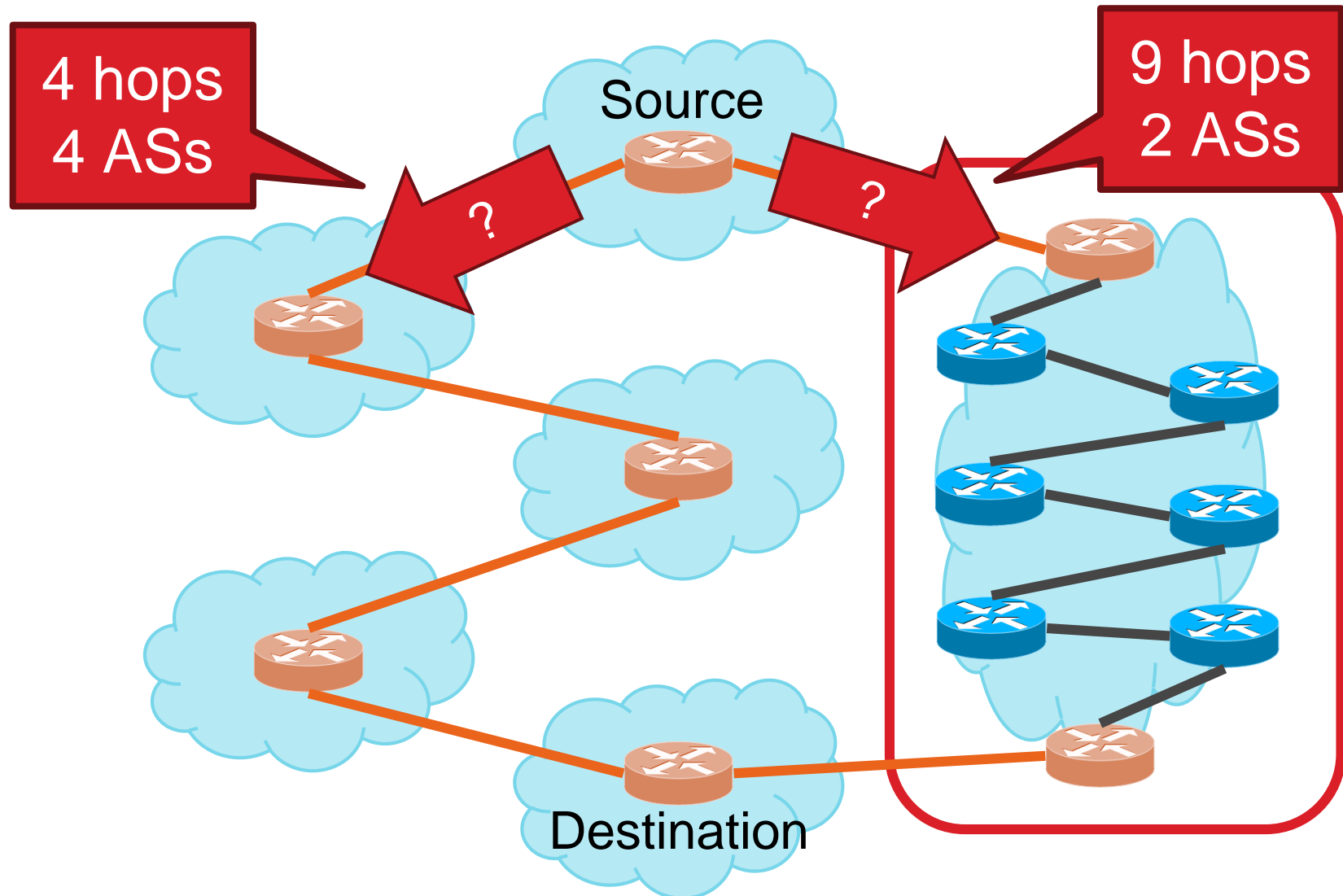
Lowest IGP Cost to BGP Egress

Traffic engineering

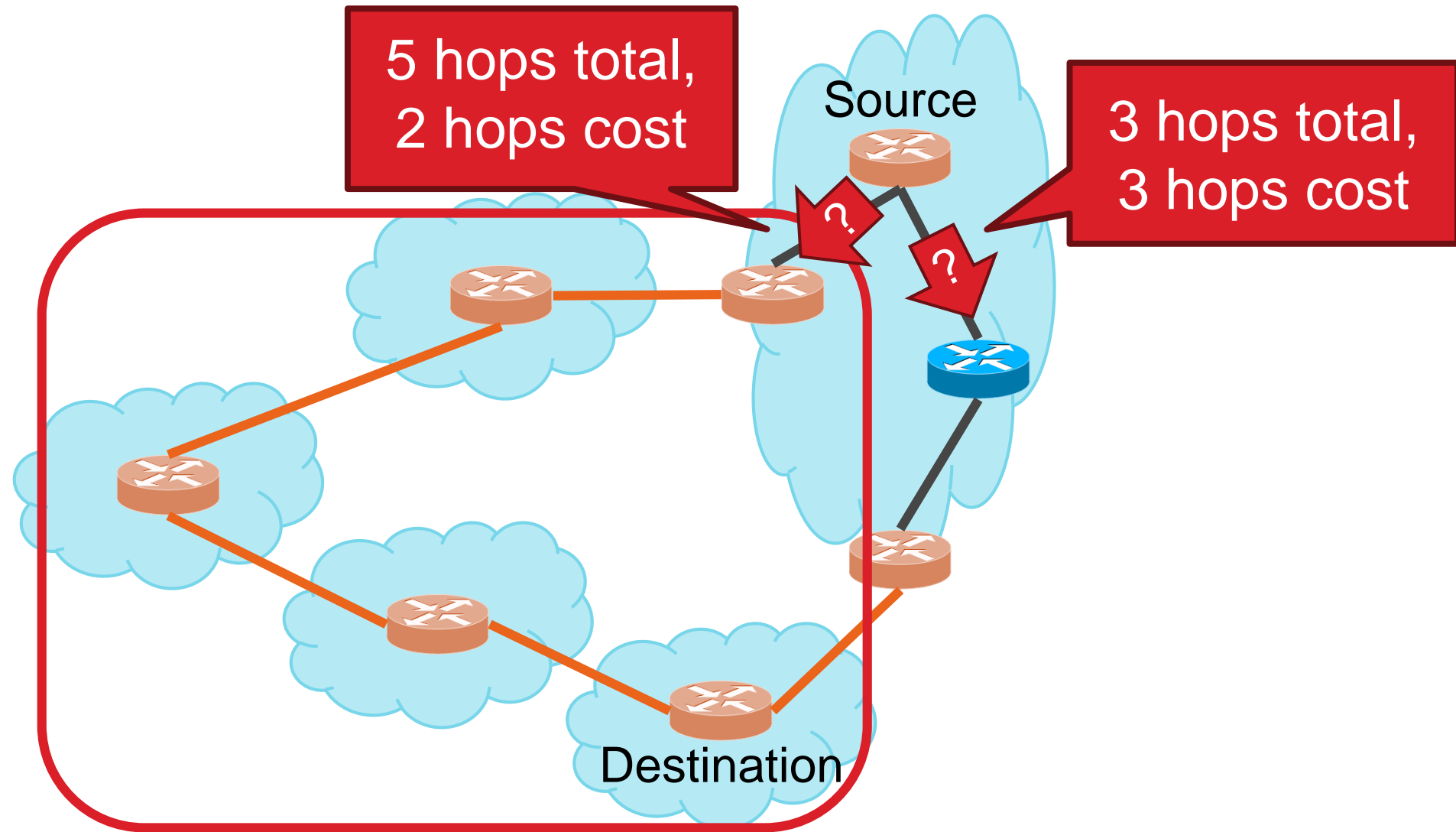
Lowest Router ID

**When all else fails,
break ties**

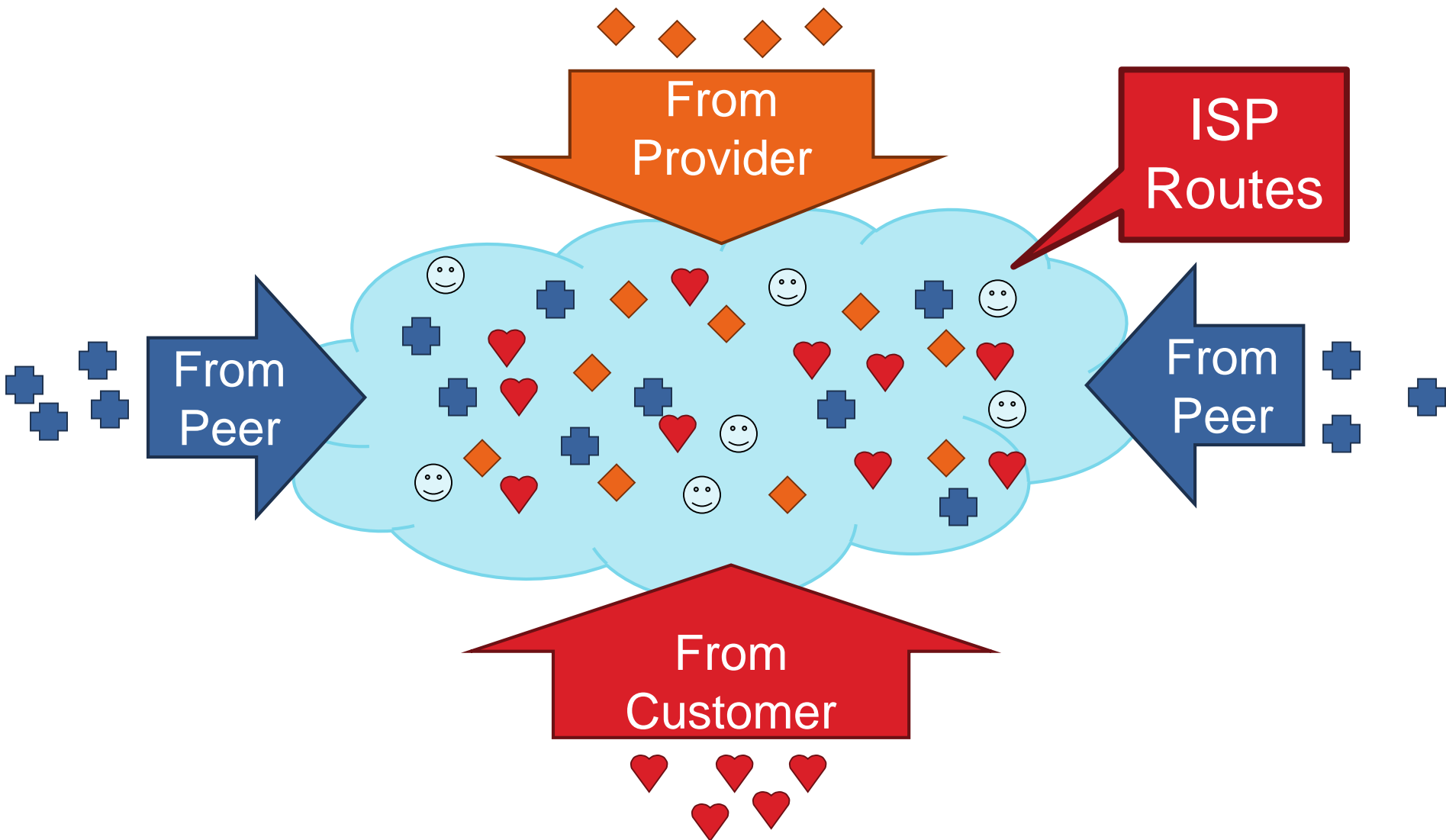
Shortest Path: Hops vs. ASes



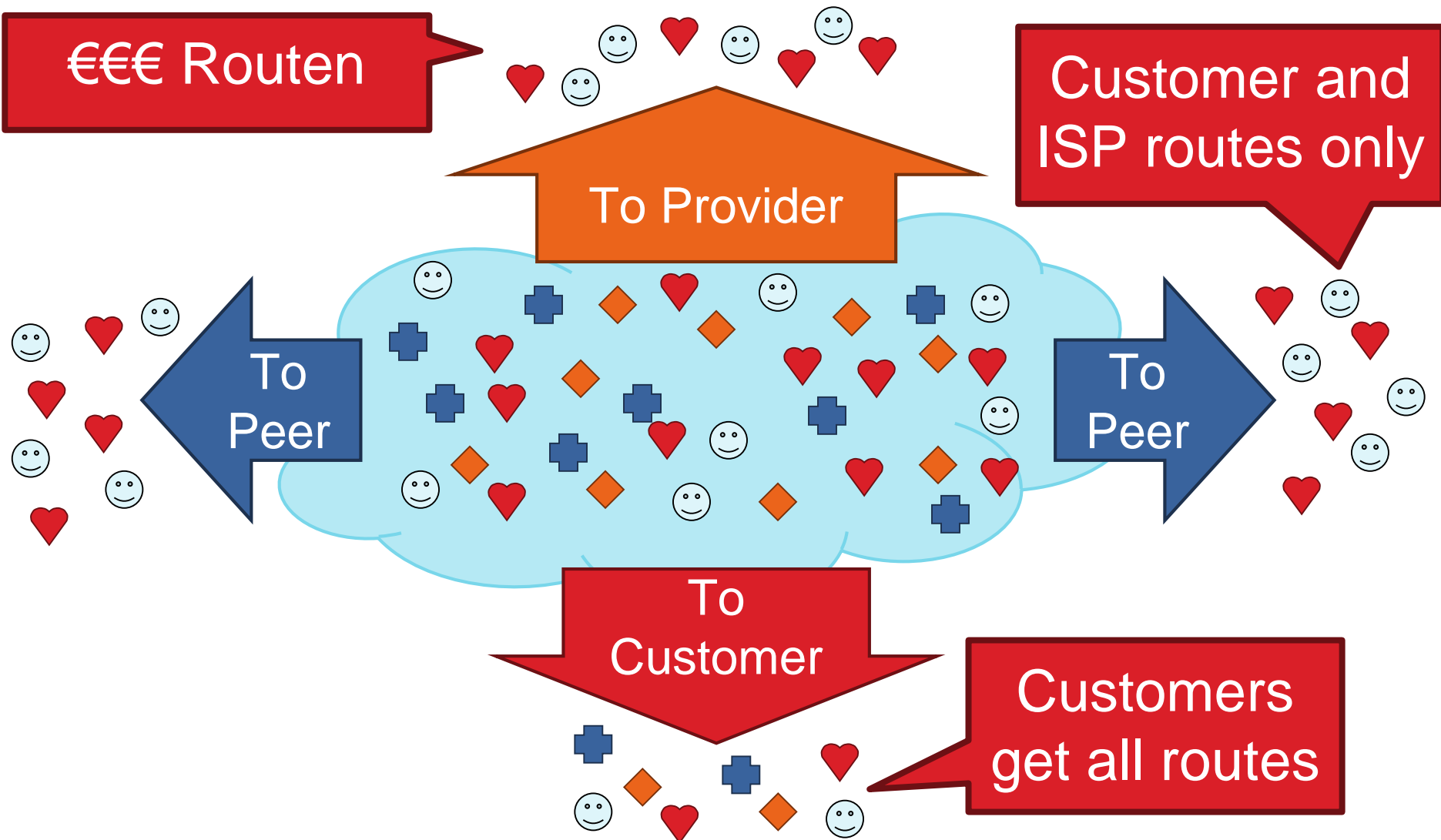
Hot Potatoe Routing



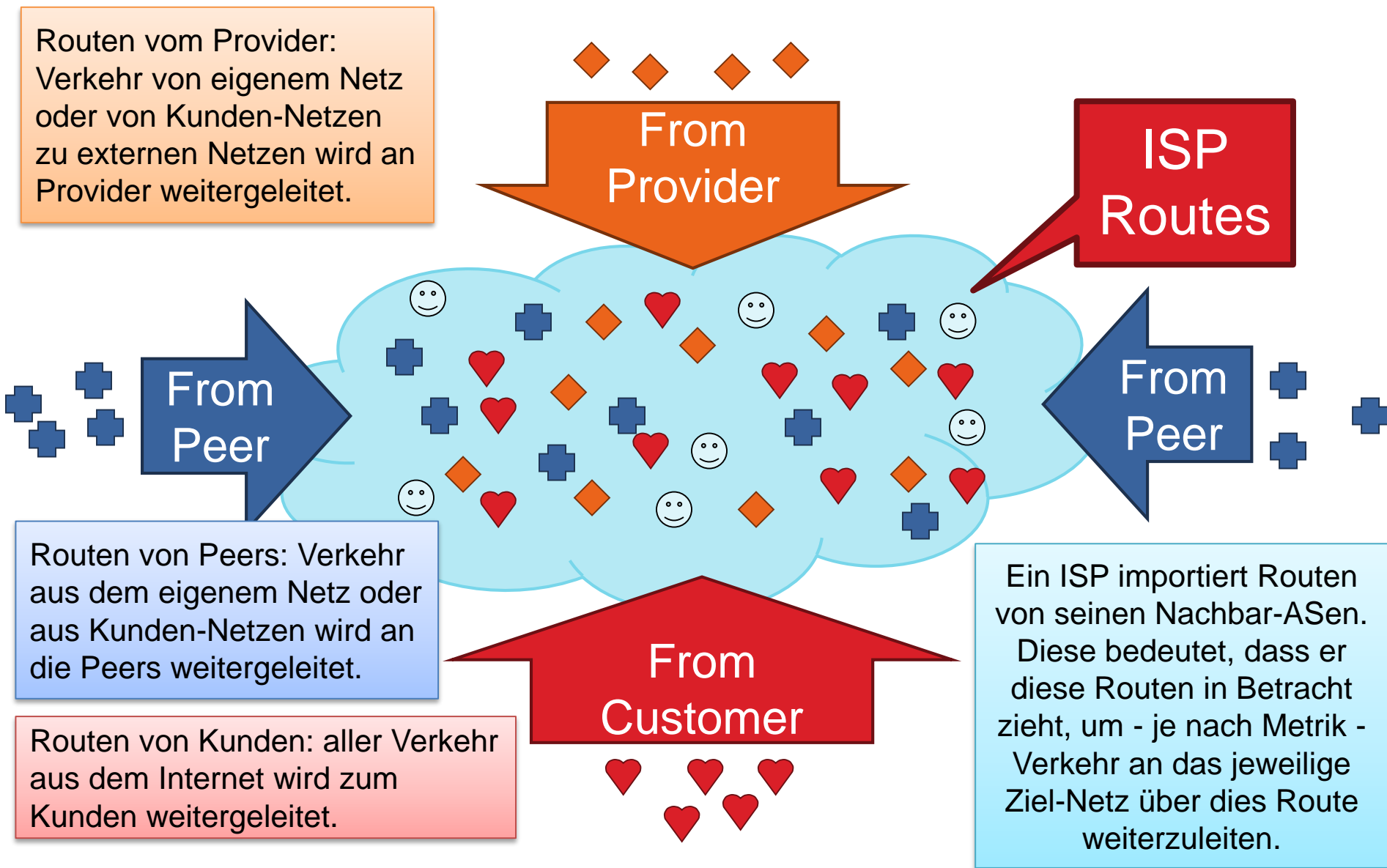
Routen importieren (animiert für Vorlesung)



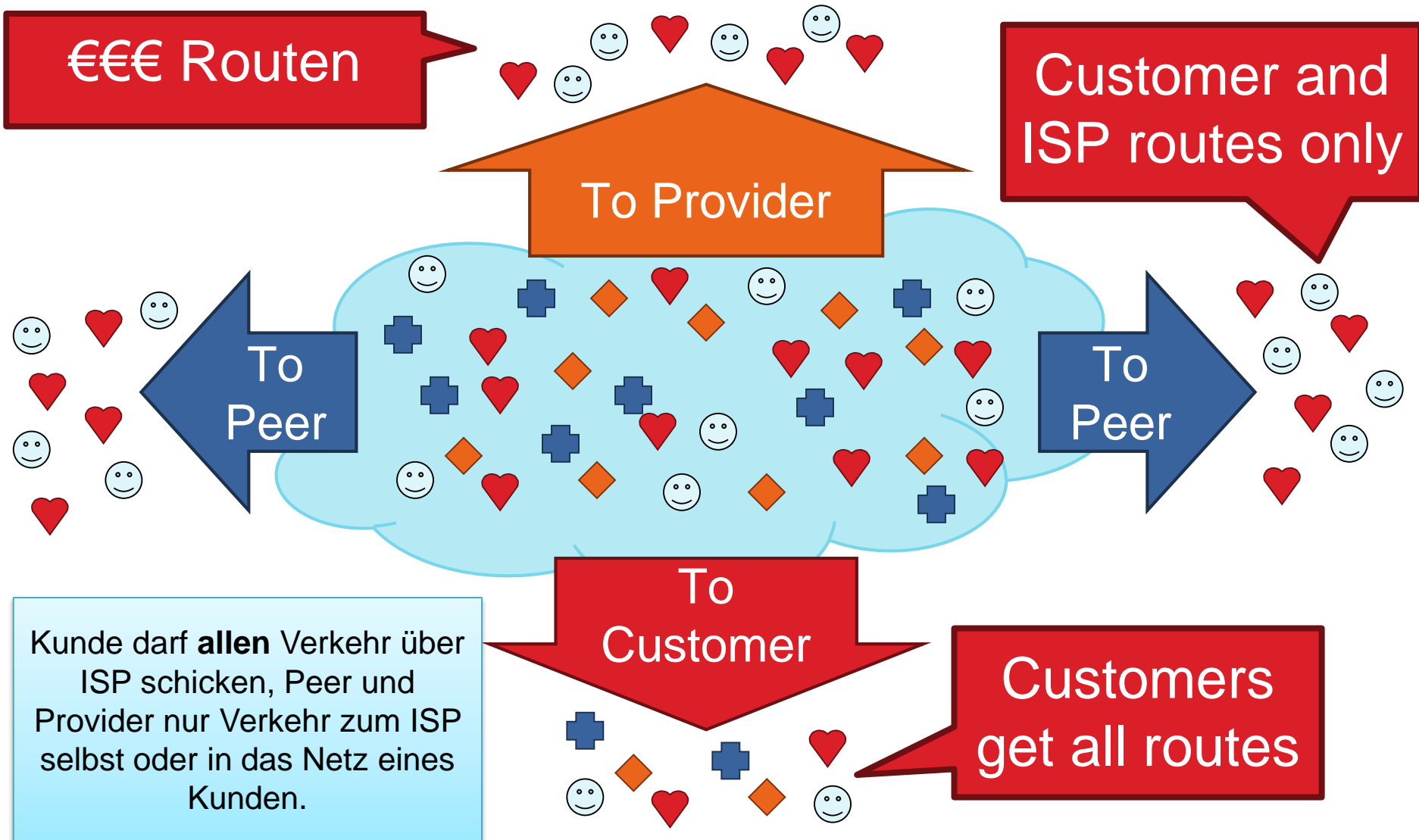
Routen exportieren (animiert für Vorlesung)



Routen importieren (mit Erklärung für Skript)



Routen exportieren (mit Erklärung für Skript)



- de facto: BGP
- Hauptkriterium: €€€
 - Peering- und Transit Agreements dominieren Routing-Entscheidungen
 - Bester Pfad im Sinne von Delay und Bandbreite spielt keine Rolle
- Herausforderung durch lokal unterschiedliche Metriken
 - keine global konsistente Routing-Metrik
 - lange Zeit (Minuten) bis zu Konvergenz
- Zusätzlich:
 - schwerwiegende Folgen bei falschem Route-Announcement
 - Pakistan und YouTube
 - iBGP zur internen Auswahl des richtigen Border Routers
 - Kombination von internen und externen Metriken
 - BGP Routing-Tabellen wachsen
 - Zunehmende Zersplitterung des Adressbereichs durch Multi-Homing und globale Unternehmen
 - Alternativen wie LISP (Locator-Identifier-Split) im Rahmen von Next-Generation-Internet-Projekten entwickelt und standardisiert -> aber noch(?) nicht genutzt