



BGP, Traffic Engineering, & Akamai

Niels Bakker – NLnog-dag 2014

The Akamai Intelligent Platform

The world's largest on-demand, distributed computing platform delivers all forms of web content and applications

The Akamai Intelligent Platform:

154,000+
Servers

2,500+
Locations

1,250+
Networks

700+
Cities

92
Countries

Hoe werkt een CDN?

Eindgebruikers die content opvragen worden doorgestuurd naar de beste server:

- Dit gebeurt vooral op basis van DNS, vooral voor CDN's zonder eigen netwerk zoals Akamai.
- Kan ook door middel van HTTP redirects of anycast.

Verschillende gebruikers van op DNS gebaseerde CDN's kunnen andere antwoorden (A of AAAA) krijgen voor dezelfde hostname.

Dit heet “mapping”. Hoe beter de mapping, hoe beter de prestaties van een CDN.

Voorbeeld van Akamai mapping in actie

: [Amsterdam] \$; dig www.symantec.com

www.symantec.com.	7112	IN	CNAME	www.symantec.com.akadns.net.
www.symantec.com.akadns.net.	24	IN	CNAME	migration-www.symantec.com.akadns.net.
migration-www.symantec.com.akadns.net.	24	IN	CNAME	www.symantec.com.edgekey.net.
www.symantec.com.edgekey.net.	6	IN	CNAME	e5211.b.akamaiedge.net.
e5211.b.akamaiedge.net.	20	IN	A	2.16.107.29

: [Munich] \$; dig www.symantec.com

www.symantec.com.	7044	IN	CNAME	www.symantec.com.akadns.net.
www.symantec.com.akadns.net.	144	IN	CNAME	migration-www.symantec.com.akadns.net.
migration-www.symantec.com.akadns.net.	144	IN	CNAME	www.symantec.com.edgekey.net.
www.symantec.com.edgekey.net.	16899	IN	CNAME	e5211.b.akamaiedge.net.
e5211.b.akamaiedge.net.	20	IN	A	23.57.67.29

: [Boston] \$; dig www.symantec.com

www.symantec.com.	5944	IN	CNAME	www.symantec.com.akadns.net.
www.symantec.com.akadns.net.	66	IN	CNAME	migration-www.symantec.com.akadns.net.
migration-www.symantec.com.akadns.net.	66	IN	CNAME	www.symantec.com.edgekey.net.
www.symantec.com.edgekey.net.	12198	IN	CNAME	e5211.b.akamaiedge.net.
e5211.b.akamaiedge.net.	20	IN	A	23.34.7.14

Hoe werkt Akamai mapping?

- Op basis van DNS query source IP address
- Opgevraagde hostname bevat een hint naar het type content
- Lage TTL, dus lage reactietijden op veranderende serverbelasting

```
: [Amsterdam] $; host whoami.akamai.com  
whoami.akamai.com is an alias for whoami.akamai.net.  
whoami.akamai.net has address 62.216.31.206
```

CDN's en Traffic Delivery

Hoe krijgen CDN's hun traffic richting jullie klanten?

- Over transit
- PNI (private peering)
- Peering via exchange points (AMS-IX, NL-IX etc.)
- Direct bij ISP's geplaatste servers (niet alle CDN's kunnen dit)

Sommige CDN's peeren met meer enthousiasme dan bepaalde andere CDN's, ondanks de duidelijke voordelen (performance, redundancy, kostenreductie, extra capaciteit voor tijdens evenementen bovenop in netwerk geplaatste servers).

Akamai traffic delivery

Akamai is speciaal:

- Akamai heeft geen eigen backbone netwerk.
- Niet elk cluster heeft dezelfde hoeveelheid servers, of servers van hetzelfde type, of zijn geconfigureerd om dezelfde content te leveren.
- Clusters zijn niet onderling verbonden.
- Dit betekent dat elk cluster eigen, aparte IP space heeft.
- Dit betekent ook dat de standaard BGP traffic engineering trukendoos niet gaat werken, of nieuwe bijwerkingen heeft.

BGP traffic engineering

Trucs die anders werken met Akamai dan met network-based CDN's:

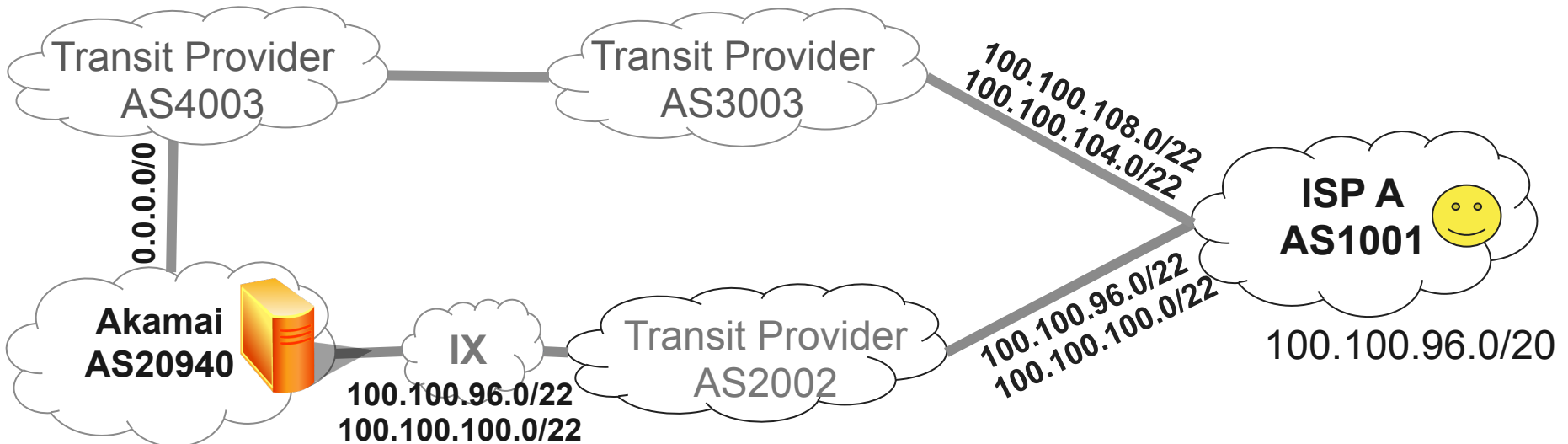
- AS Path prepending
- MED
 - Clusters zijn niet onderling verbonden, er is geen andere BGP route die korter wordt
- More specifics announcen over een andere transit

Het Akamai mapping system kijkt naar latency, throughput, packet loss van zowel test pings naar nameservers als daadwerkelijke client requests.

Zorg ervoor dat je nameservers pingbaar zijn!

Praktijkvoorbeeld: incomplete route announcement

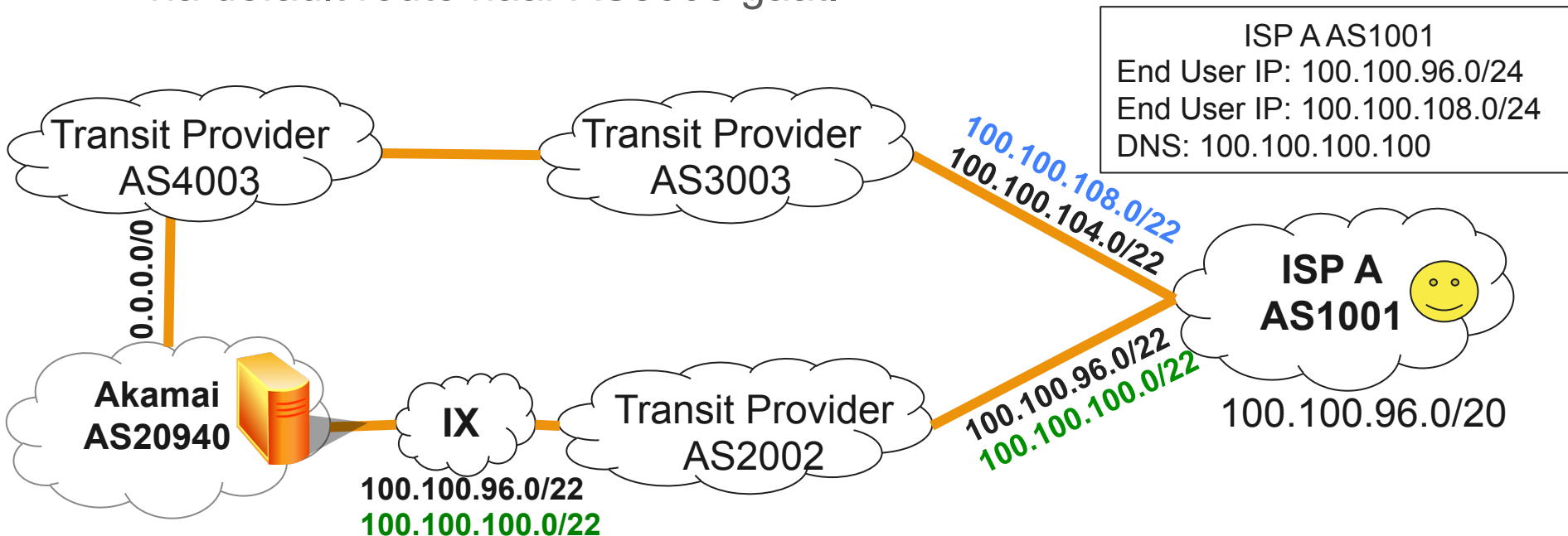
- ISP A is multihomed naar transit providers AS2002 en AS3003
- Transit provider AS2002 peert met Akamai
- Transit provider AS3003 peert niet met Akamai
- ISP A announcet verschillende prefixes naar verschillende ISP's
- ISP A kan het volledige internet bereiken.



100.100.96.0/22	AS2002 AS1001
100.100.100.0/22	AS2002 AS1001
0.0.0.0/0	AS4003

Hoe gaat het verkeer naar de eindgebruikers?

- ISP eindgebruikers in 100.100.96.0/22, 100.100.100.0/22, 100.100.104.0/22, 100.100.108.0/22
- Eindgebruikers gebruiken ISP A nameserver 100.100.100.100
- Akamai ontvangt prefix 100.100.100.0/22 van AS2002, dus verkeer van ISP A wordt naar dit cluster gemapt
- 100.100.96.0/22 en 100.100.100.0/22 verkeer wordt naar AS2002 gerouteerd, terwijl verkeer naar 100.100.104.0/22, 100.100.108.0/22 via default route naar AS3003 gaat.

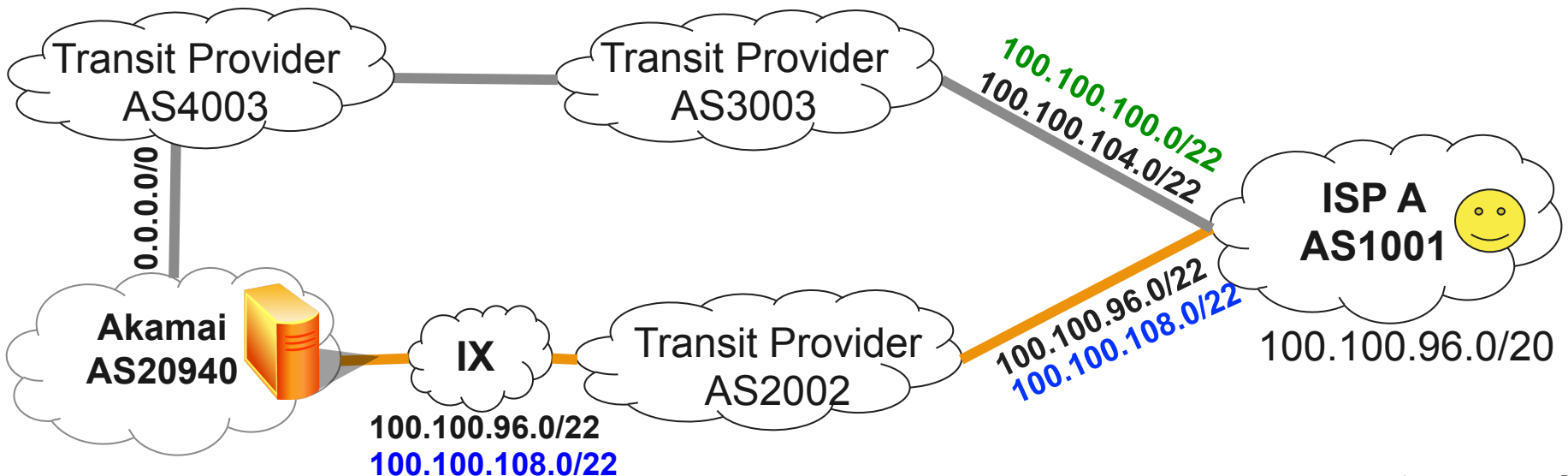


Is dit een probleem?

- Sommige gebruikers in [100.100.108.0/22](#) hebben mogelijk andere performance.
- Wat gebeurt er als de ISP hun announcements wijzigt?
- Dit (*inconsistent announcements*) komt in de praktijk voor.

ISP A wijzigt hun announcements

- ISP A merkt dat AS3003 performance lager is dan AS2002
- ISP A past route announcements aan
- Zowel 100.100.96.0/22 and 100.100.108.0/22 gaan nu over AS2002 en eindgebruikers in deze twee prefixes hebben identieke performance
- Voorheen klagende eindgebruikers van ISP A zijn nu blij; ISP sluit ticket.



Een dag later...

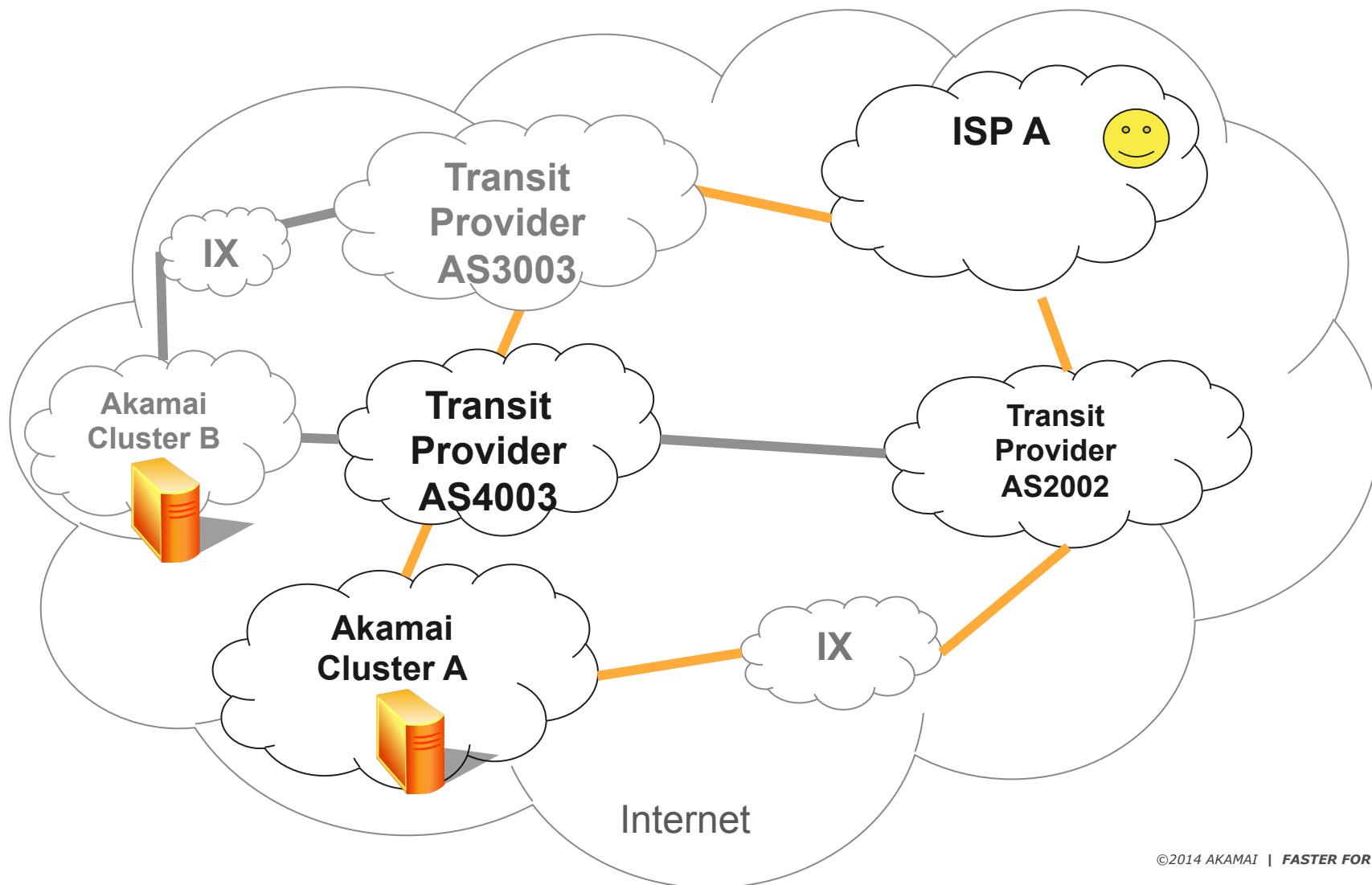
- Het Akamai mapping system merkt dat de prefix echt verdwenen is en stopt geheel met eindgebruikers van ISP A naar dit cluster te mappen.

Veranderde performance voor ISP A eindgebruikers!

- Akamai ontvangt DNS prefix **100.100.100.0/22** niet langer van AS2002
- Akamai mapt het verkeer van ISP A naar andere clusters dan Cluster A
- ISP A ontvangt nog steeds verkeer van beide upstreams
- Eindgebruikers van ISP A klagen weer 😞

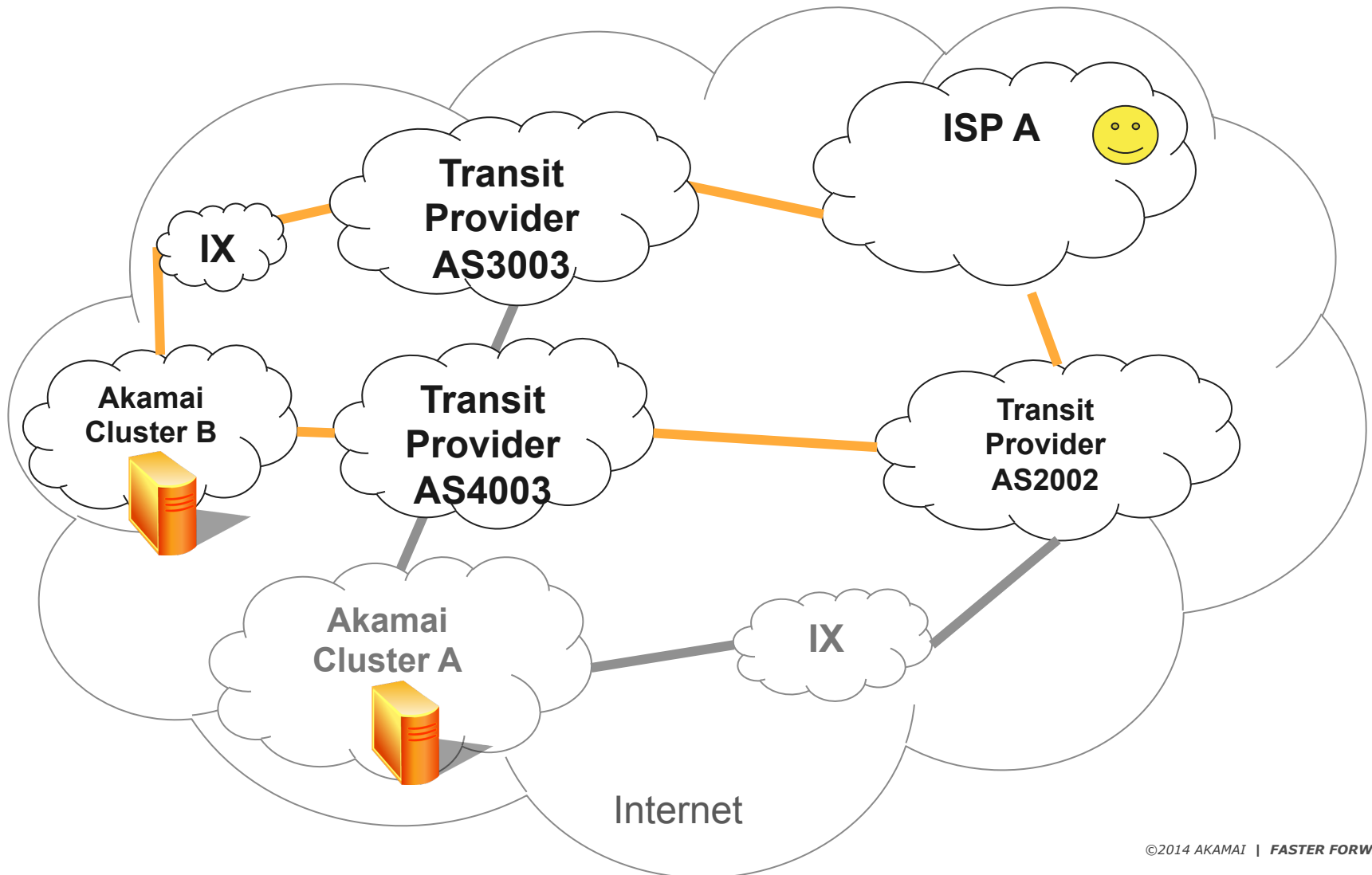
Situatie voor de mapping system refresh

- Akamai mapt het verkeer naar Cluster A



Situatie na de mapping system refresh

- Akamai mapt het verkeer naar Cluster B

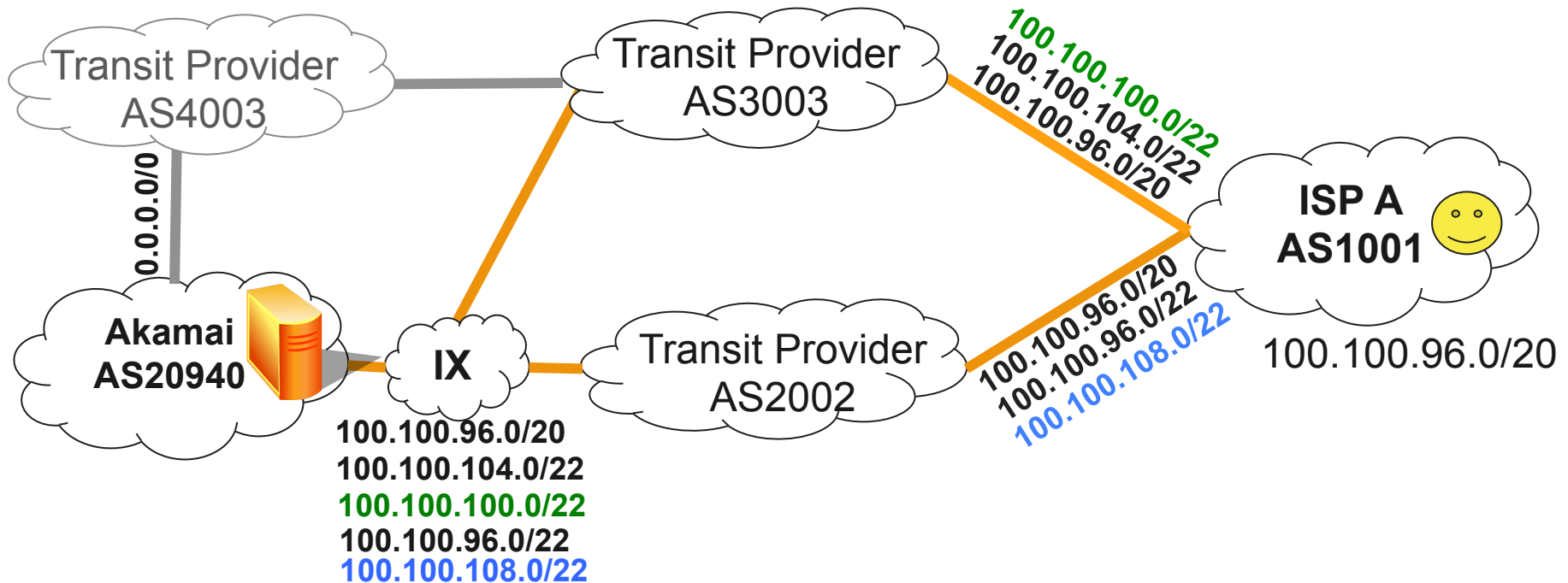


Conclusie

- Houd complete en consistente prefix announcements
- Neem contact met ons op als er problemen zijn met performance of hoeveelheid verkeer over een bepaald pad
- Samen kunnen we tot een oplossing komen.

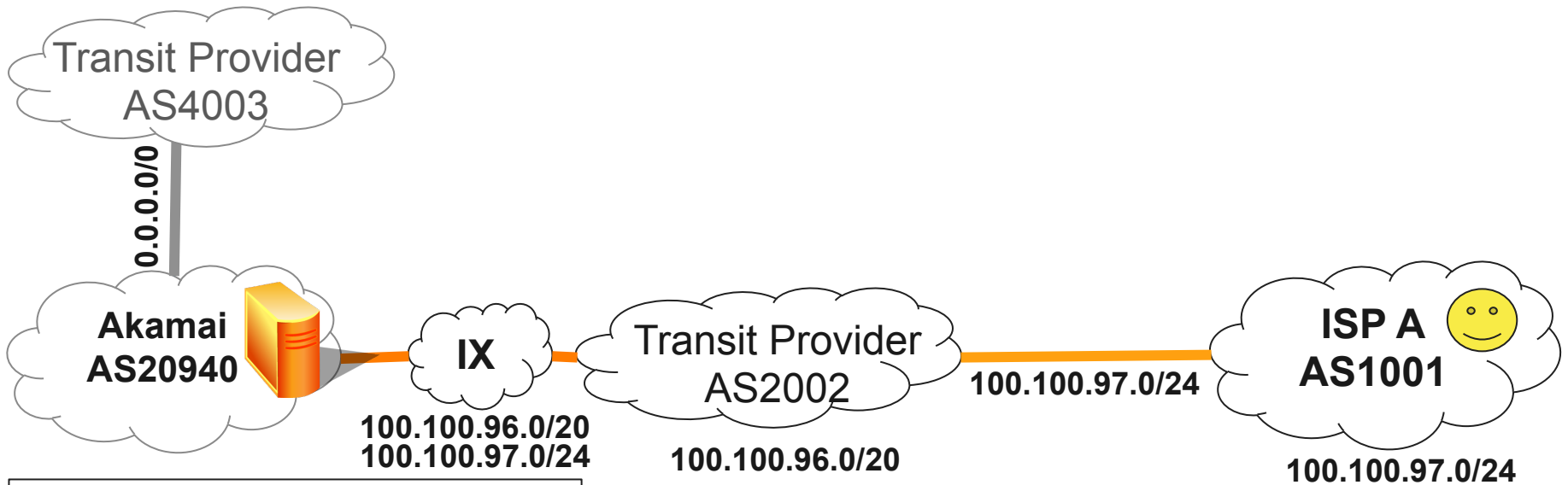
De ideale situatie

- ISP A heeft volledige announcements naar beide upstreams
- ISP A werkt samen met Akamai en met de upstream provider
- Transit provider AS3003 peert met Akamai.



Praktijkvoorbeeld 2: de verhuisde klant

- ISP A is singlehomed naar transit provider AS2002
- ISP A heeft een /24 van transit provider AS2002
- Akamai stuurt verkeer naar ISP A via transit provider AS2002.



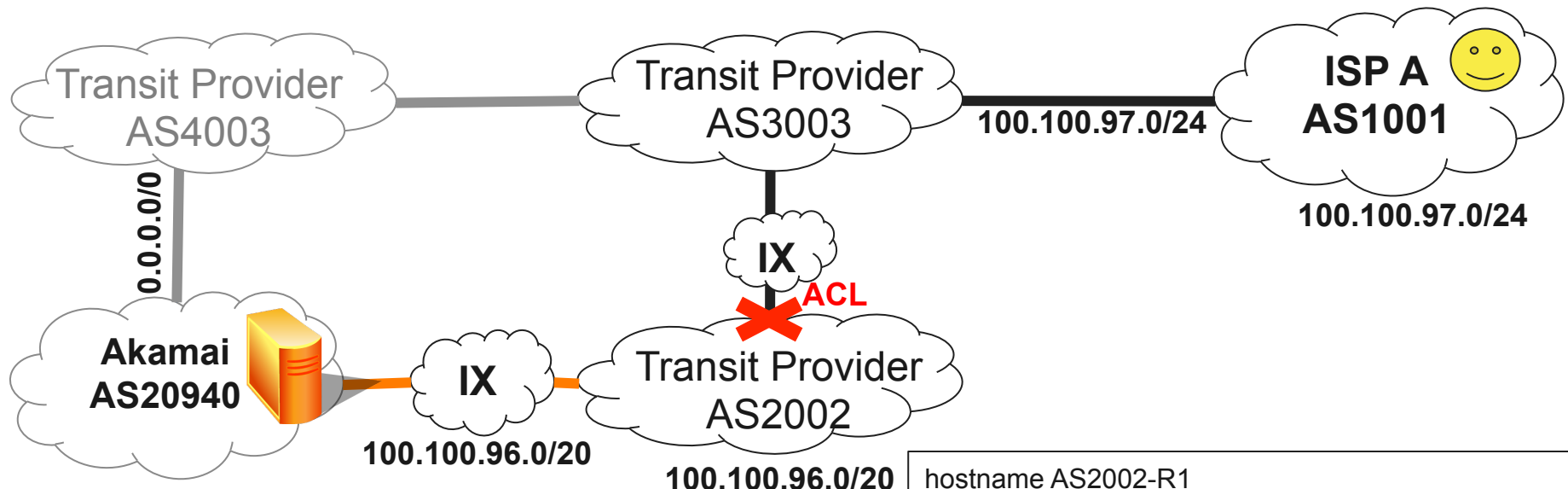
100.100.96.0/20	AS2002
100.100.97.0/24	AS2002 AS1001
0.0.0.0/0	AS4003

Wat gaat er hier fout?

- Verloren omzet voor transit provider AS2002: ook al wordt hun backbone gebruikt, de klant is toch weg, dus geen inkomsten
- Wat gebeurt er als AS2002 dit peering verkeer niet leuk vindt en tegenmaatregelen neemt?

Transit provider AS2002 filtert peering verkeer

- Transit provider AS2002 plaatst een traffic filter op hun internet exchange poort om van dit verkeer af te komen
- ISP A kan nu niet bij bepaalde CDN content vanwege dit blackhole.



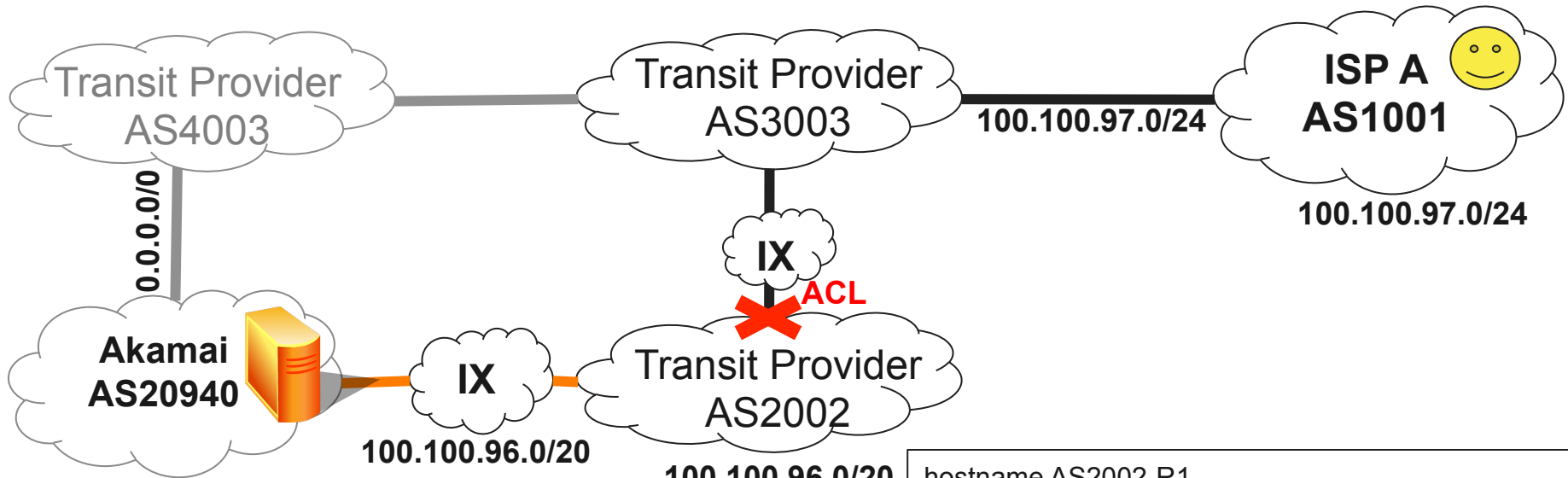
Akamai AS20940 Routing Table

100.100.96.0/20	AS2002
0.0.0.0/0	AS4003

```
hostname AS2002-R1
!  
interface TenGigabitEthernet1/1  
ip access-group 101 out  
!  
access-list 101 deny ip any 100.100.97.0 0.0.0.255  
access-list 101 permit ip any any
```

Akamai workaround om het filter heen

- Akamai merkt dat ISP A eindgebruikers onbereikbaar zijn
- Akamai blokkeert alle prefixen ontvangen van transit provider AS2002, dus traffic verplaatst zich naar transit provider AS4003
- ISP A kan alle CDN content weer bereiken
- Transit provider AS2002 ziet internet exchange verkeer wegvallen.



100.100.96.0/20	AS2002
0.0.0.0/0	AS4003

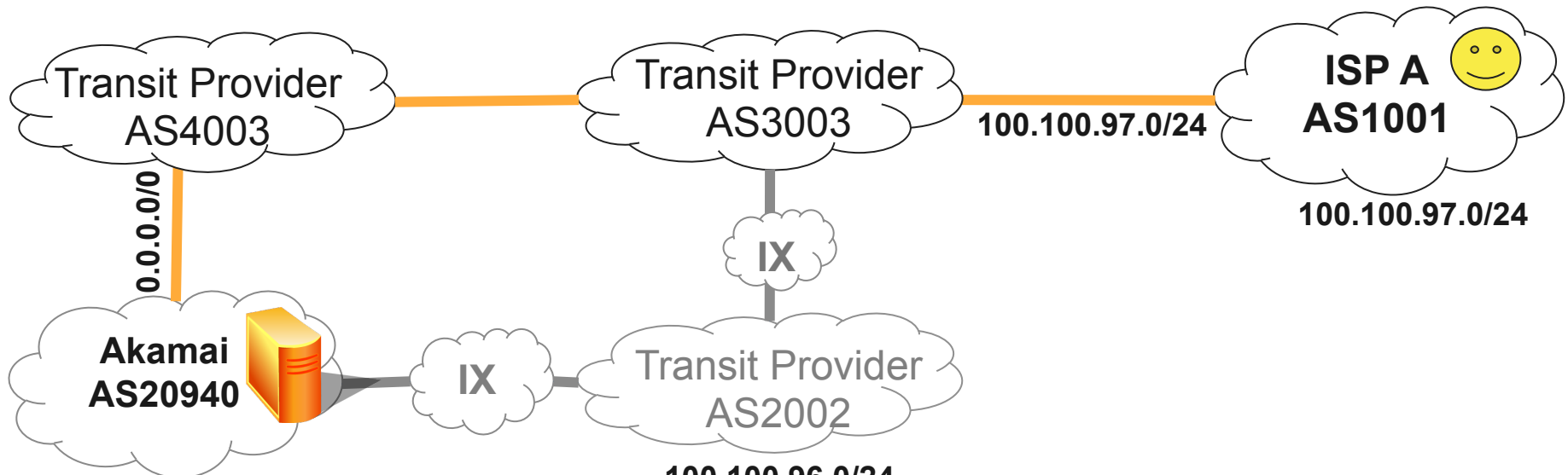
```
hostname AS2002-R1
!
interface TenGigabitEthernet1/1
ip access-group 101 out
!
access-list 101 deny ip any 100.100.97.0 0.0.0.255
access-list 101 permit ip any any
```

Is dergelijke filtering een goed idee?

- We zien in de praktijk dat sommige transit providers dit soort verkeer filteren op hun peering of transit poorten.
- Als je aangeeft verkeer voor een netblock (bijv. /20) af te willen handelen horen daar geen gaten in te zitten en hoort alle verkeer afgeleverd te worden.
- Als een ISP een netblock (zeg /24) permanent aan een klant delegeert (het oude Assigned Portable), dan moet het supernet (bijv. /20) niet meer geannounced worden.
- Connectivity van eindgebruikers is in het geding met dit soort filters!

De 'ideale' (...) situatie

- AS2002 kan het supernet (/20) opbreken in sub-blocks
- AS2002 kan hierna stoppen met annoucen van de prefix van ISP A



Akamai AS20940 Routing Table	
100.100.96.0/24	AS2002
100.100.98.0/23	AS2002
100.100.100.0/22	AS2002
100.100.104.0/21	AS2002
0.0.0.0/0	AS4003

100.100.96.0/24
100.100.98.0/23
100.100.100.0/22
100.100.104.0/21

Conclusies

- CDN's in het algemeen en Akamai in het bijzonder werken anders dan gewone BGP routing.
- Ga peeren voor meer capaciteit, lagere kosten en betere performance!

Questions?



Akamai Careers

@akamaijobs

We're Hiring!! Senior Network Support
Consultant Job (Amsterdam, NL)
bit.ly/1mRijLX #akamai #job



RETWEETS

2



12:12 PM - 15 Aug 2014

Met dank aan Christian Kaufmann, Matt Jansen, Kams Yeung, Torsten Blum