Zorgeloos online

Informatiesessie *DDoS Resiliency*

Hoe BGP Anycast het .nl domein meer *resilient* maakt

Marco Davids / Marc Groeneweg

7 juni 2023, 15:00 – 15:45

Hi! 👋

- Over SIDN
- BGP Anycast

~66%

- Stappen naar BGP Anycast
- Toekomstplannen

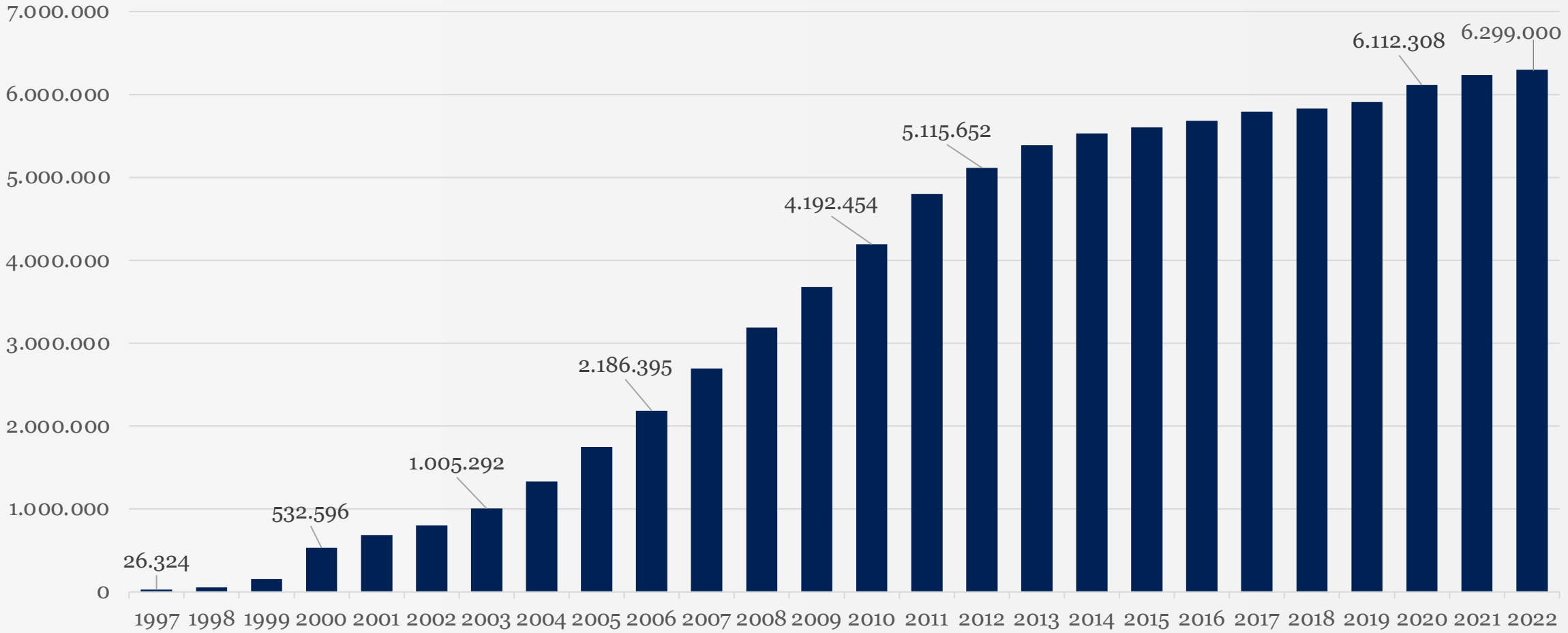
~33%

Over SIDN

- *Registry* voor het .nl *country code top-level* domein
  - Tegenwoordig ook .amsterdam, .politie en .aw (technisch beheer)
  - 6,3 miljoen .nl domeinnamen - 58% DNSSEC
  - Merkbewaking, .nl-Control, portfoliochecker, abuse204.nl, etc.
- Yivi.app
- SIDN Fonds
- SIDN Labs

https://www.sidn.nl/over-sidn/wat-wij-doen

Aantal .nl-domeinnamen: op 4 na grootste ccTLD



Over SIDN Labs

Toegepast technisch onderzoek
naar de veiligheid van internetinfrastructuur

- Drie thema's:
  - Domeinnaambeveiliging
  - Infrastructuurbeveiliging
  - *Emerging Internettechnologies*

https://www.sidnlabs.nl/over-sidnlabs

Over SIDN Labs - voorbeelden

- ENTRADA
- DMAP
- LogoMotive
- RegCheck
- SPIN
- DDoS DB
- TimeNL
- IETF / RIPE / DNS-OARC / CENTR
- ICANN-onderzoeken

https://www.sidnlabs.nl/over-sidnlabs

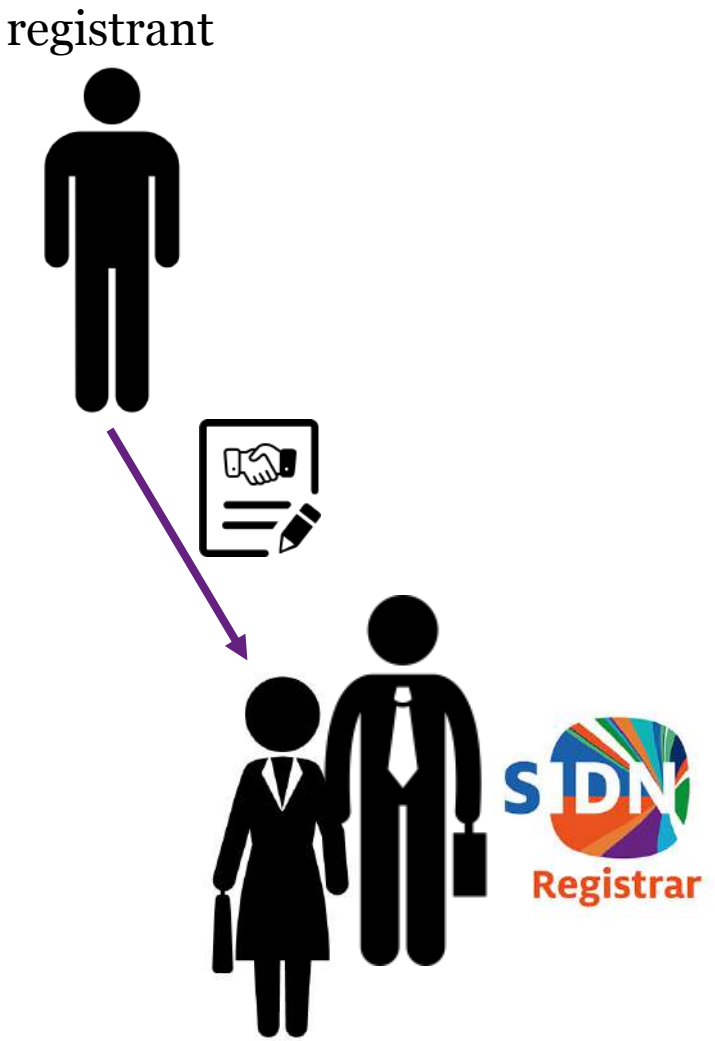
Terminologie

- *Registrant* wil domeinnaam

registrant

Terminologie

- *Registrant* wil domeinnaam
- Gaat naar *Registrar*

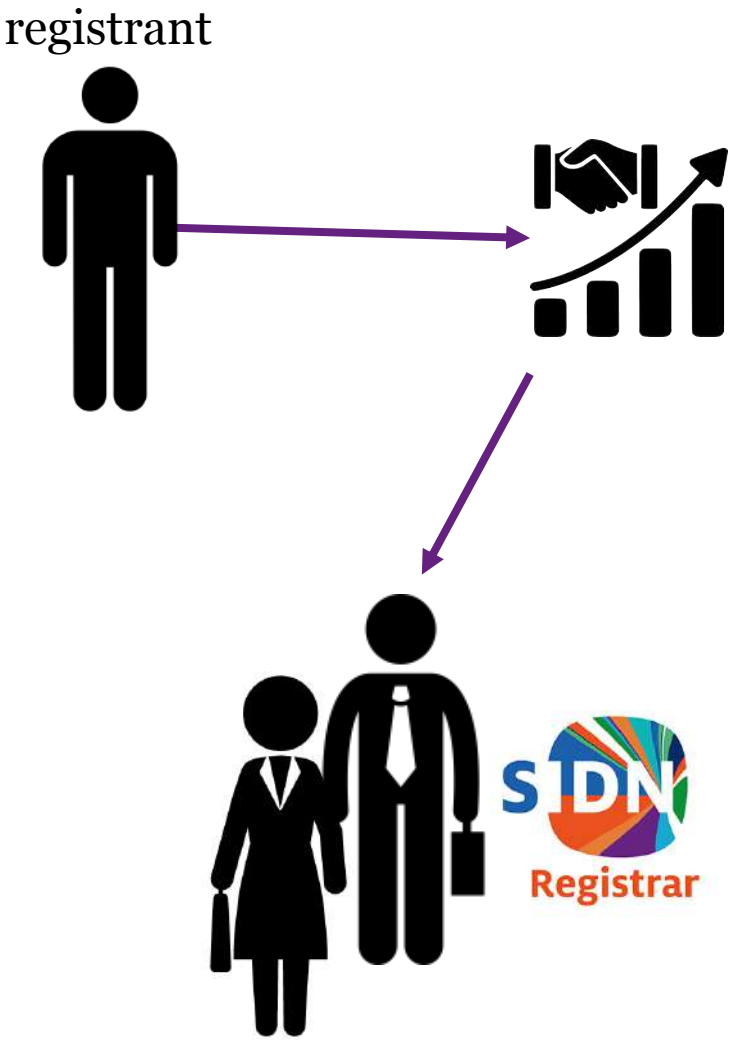


https://www.sidn.nl/nl-domeinnaam/registrar-zoeken

SIDN LABS

Terminologie

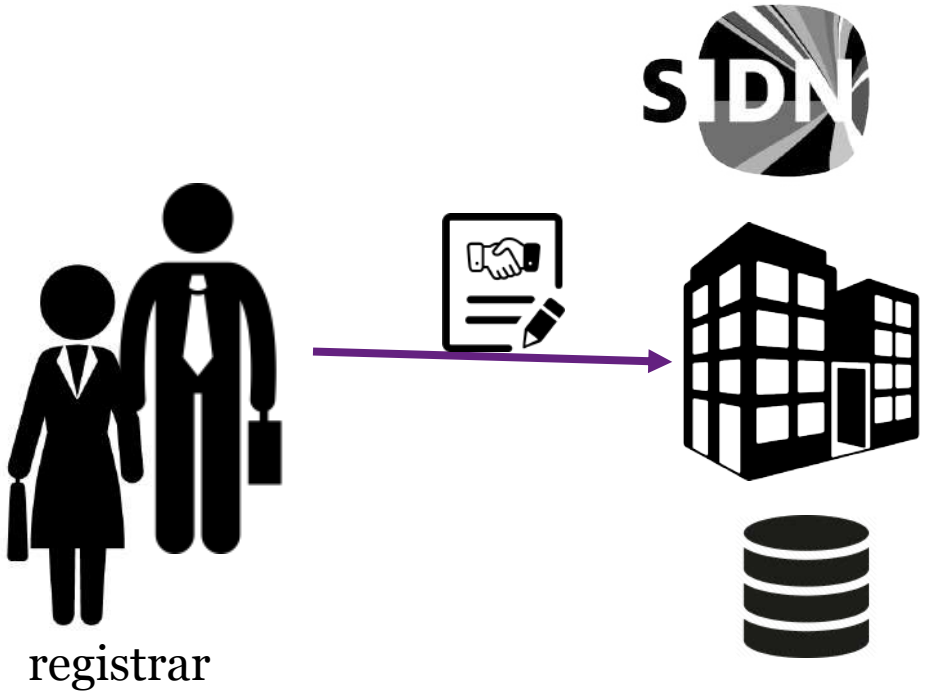
- *Registrant* wil domeinnaam
- Gaat naar *Registrar*
  - (eventueel via *Reseller*)



https://www.verenigingvanregistrars.nl/

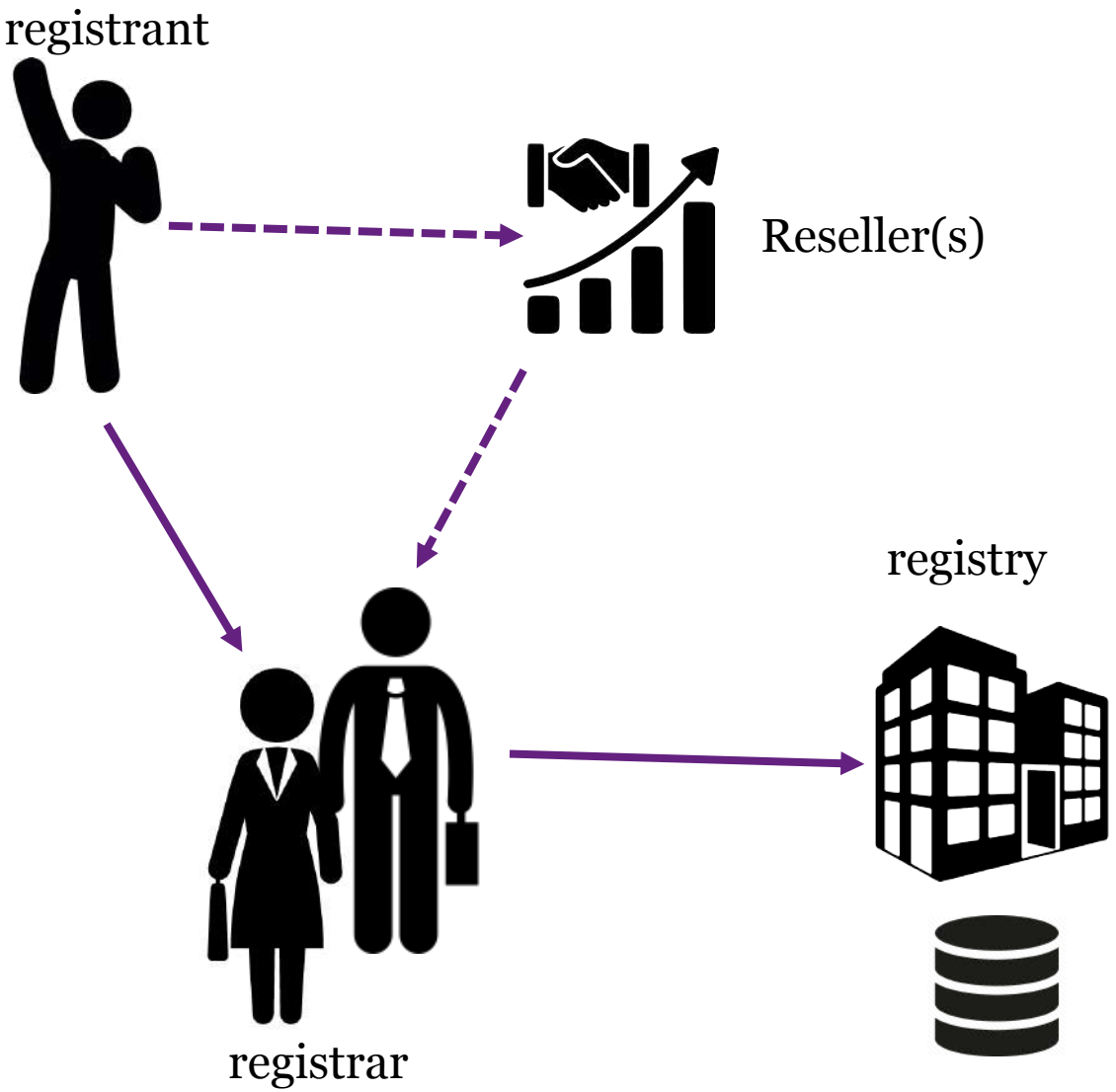
Terminologie

- *Registrar* is aangesloten bij *Registry*
- Dat zijn wij ☺



Terminologie

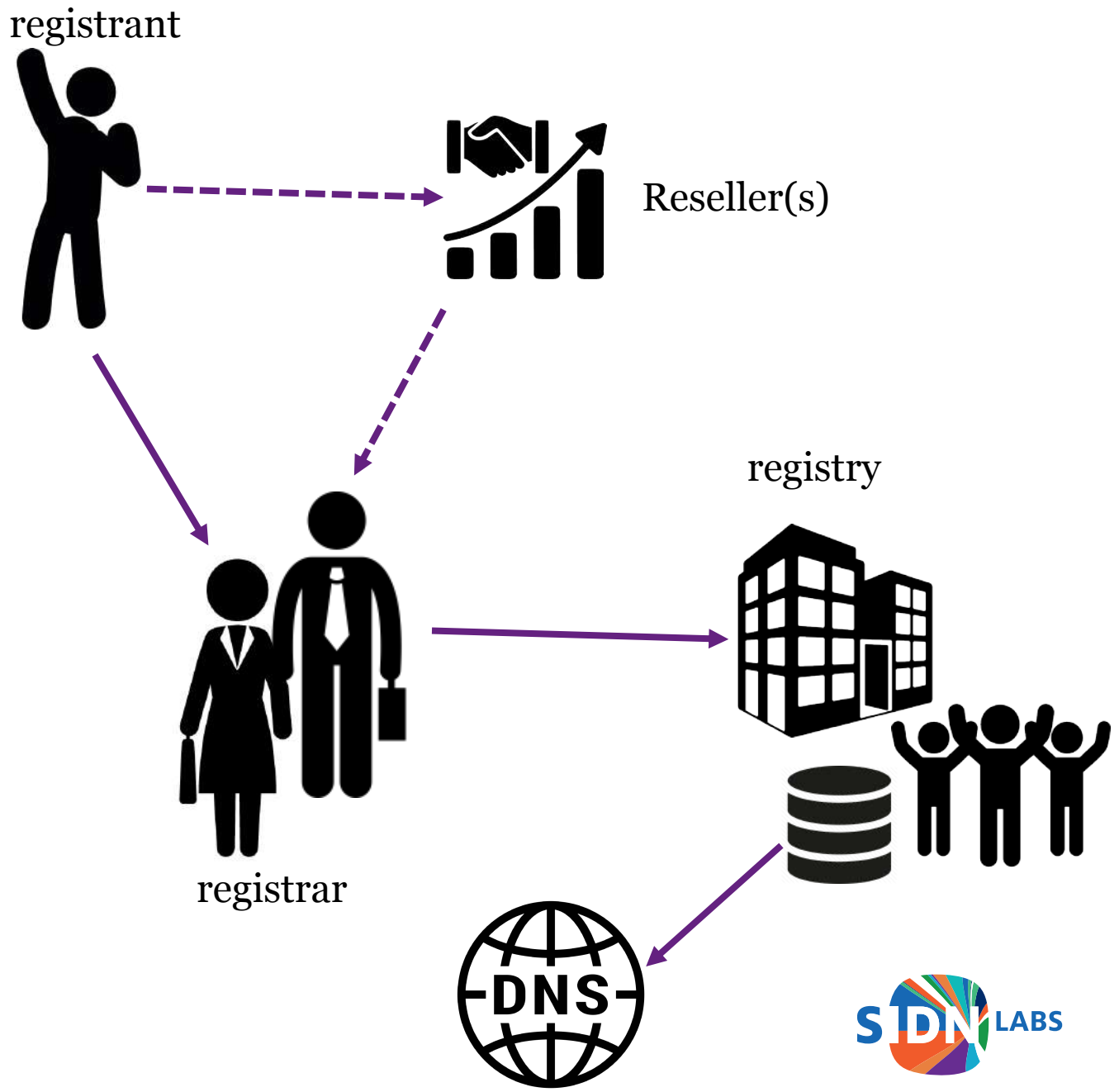
- De domeinnaam is geregistreerd!



https://whois.nl/

Terminologie

- De domeinnaam is geregistreerd
- De domeinnaam is gepubliceerd!
  - (pas dan werkt hij op het internet)



Belangrijke taak!

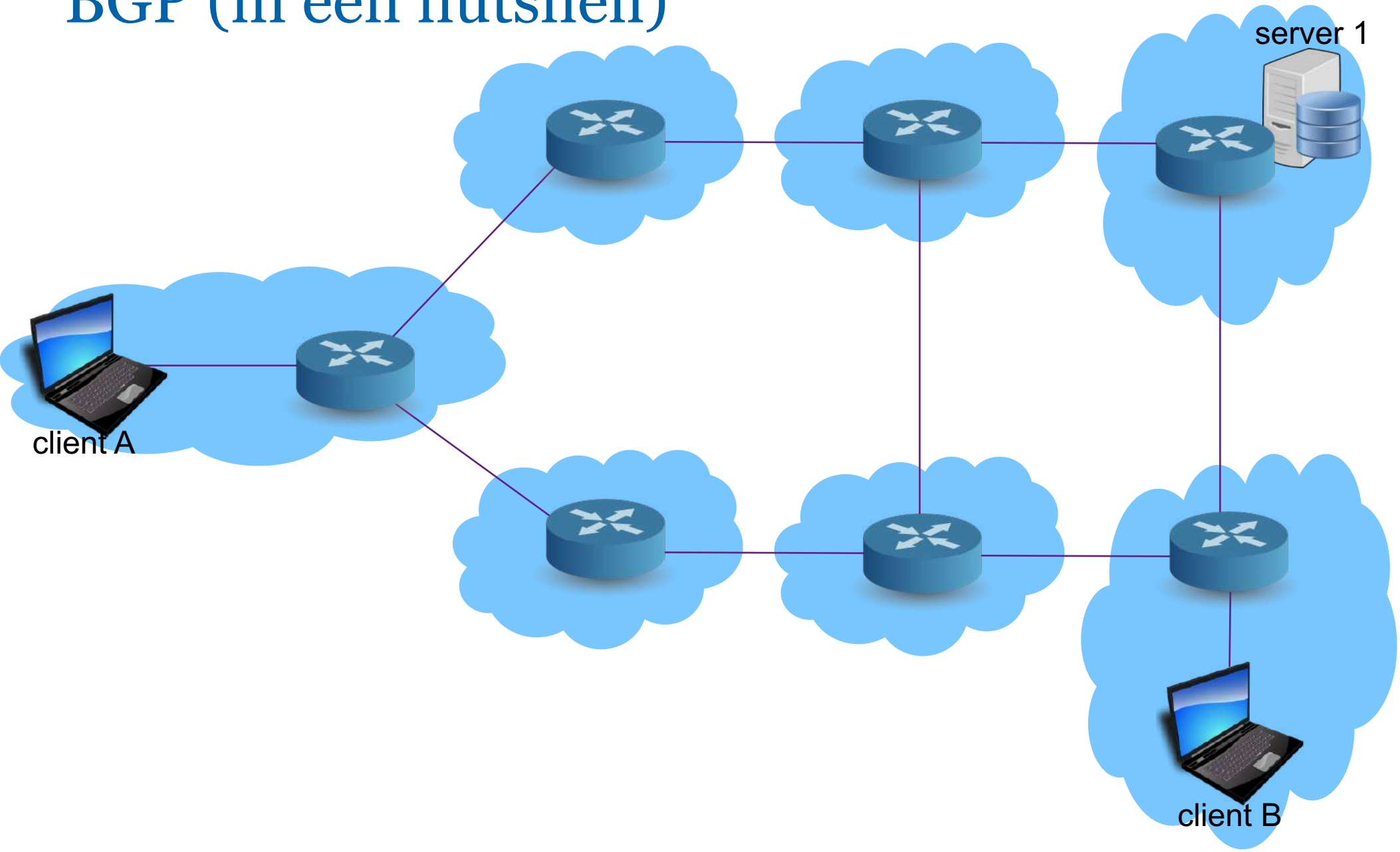
DNS

Immers:

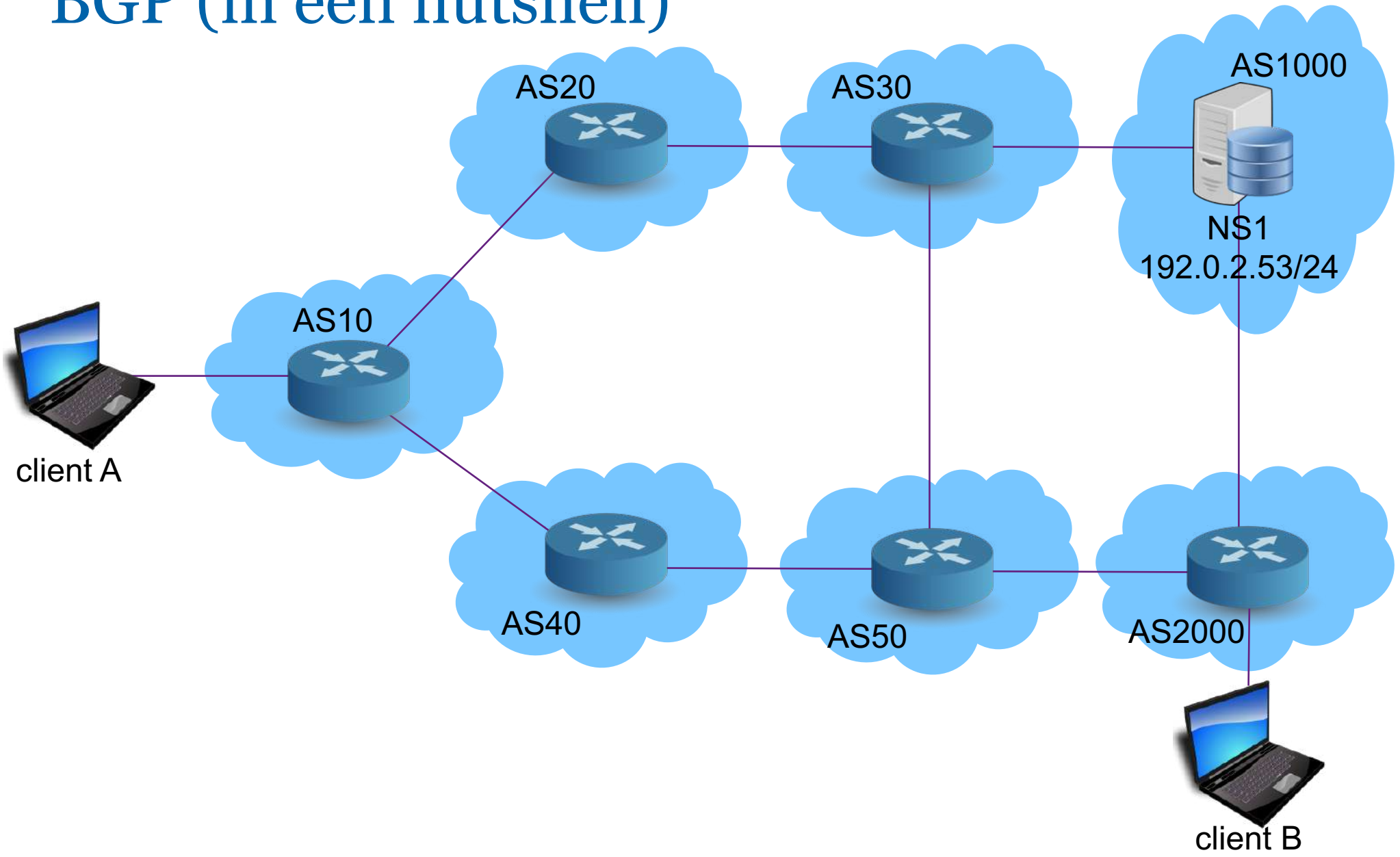
- Als je DNS resolver stuk is, is zo ongeveer effectief het internet stuk
- Als onze .nl *authoritative* DNS stuk is, zijn 6,3 miljoen .nl domeinnamen stuk
- Als jouw *authoritative* DNS stuk is, …. ?
- Kortom: DNS is belangrijk (maar dikwijls onderschat)

Onze oplossing: BGP Anycast

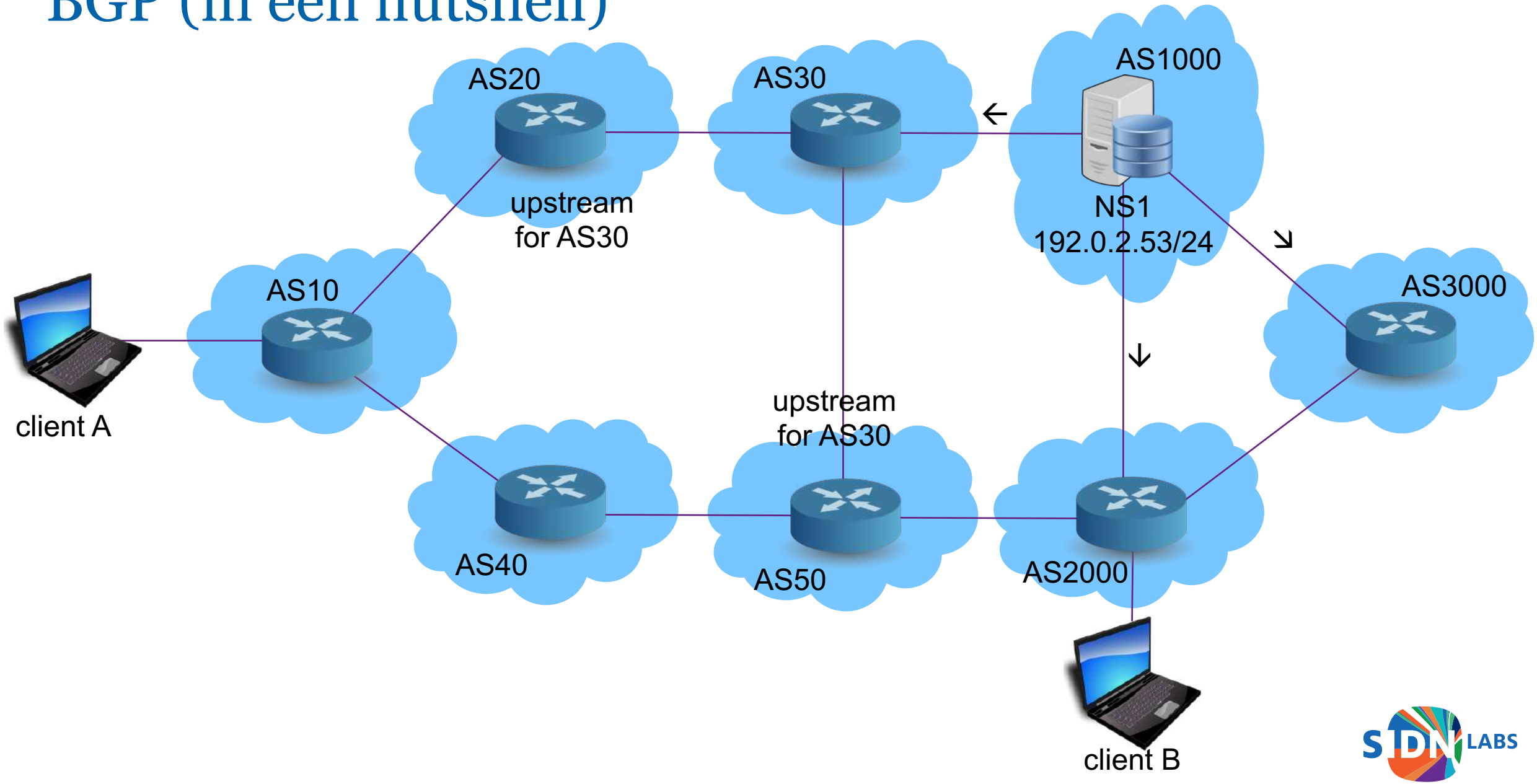
BGP (in een nutshell)



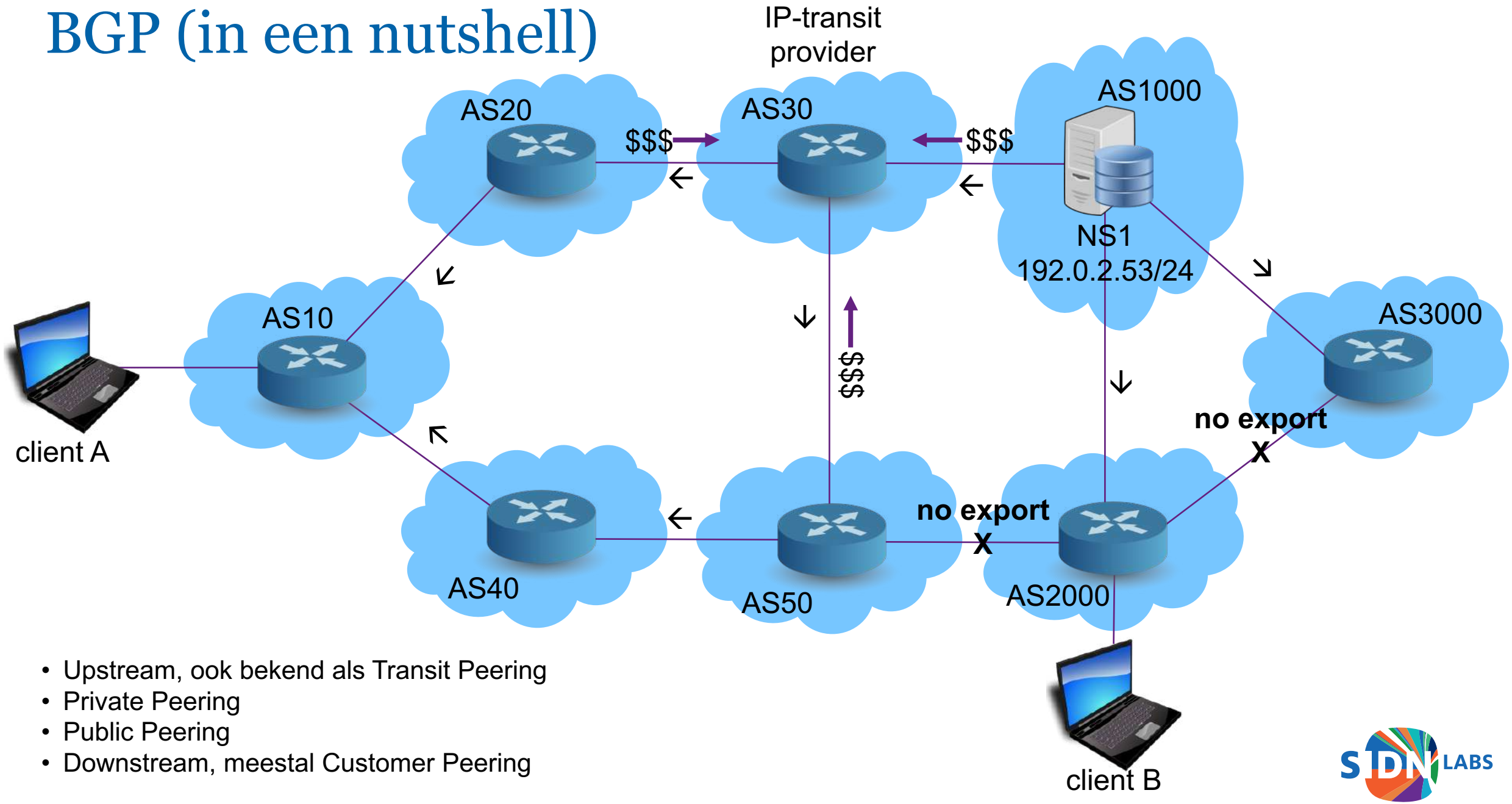
BGP (in een nutshell)



BGP (in een nutshell)

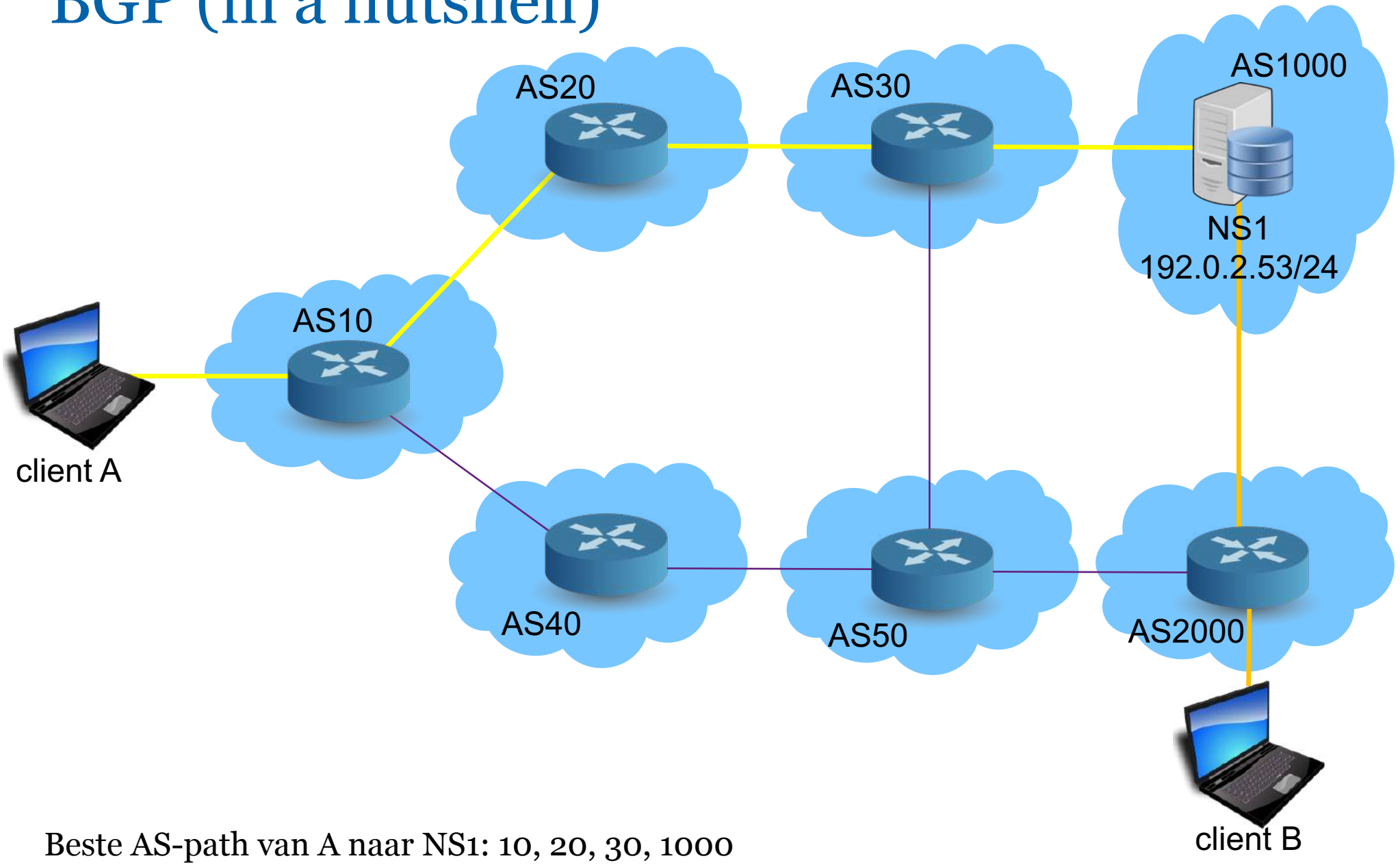


BGP (in een nutshell)



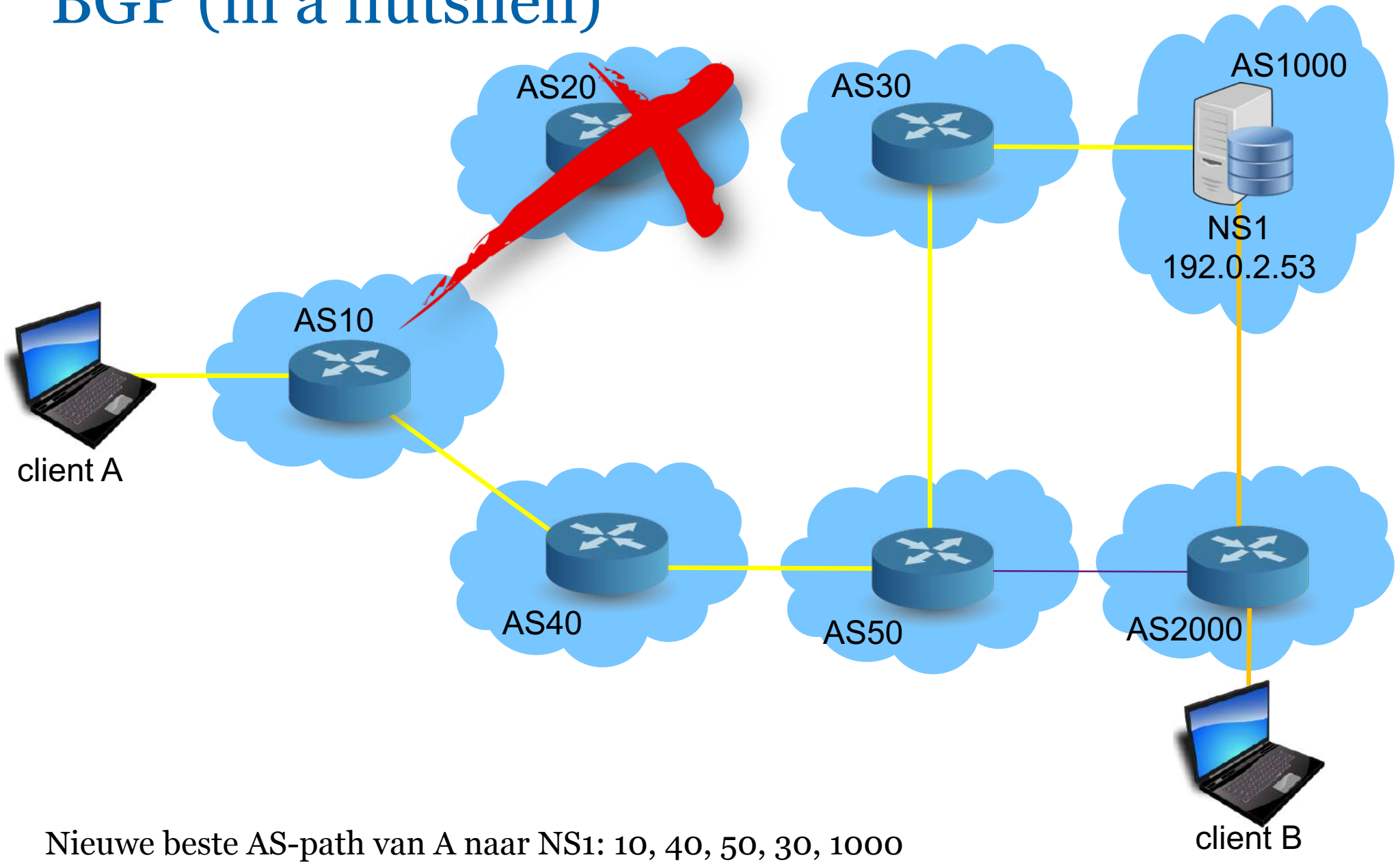
- Upstream, ook bekend als Transit Peering
- Private Peering
- Public Peering
- Downstream, meestal Customer Peering

BGP (in a nutshell)



Beste AS-path van A naar NS1: 10, 20, 30, 1000

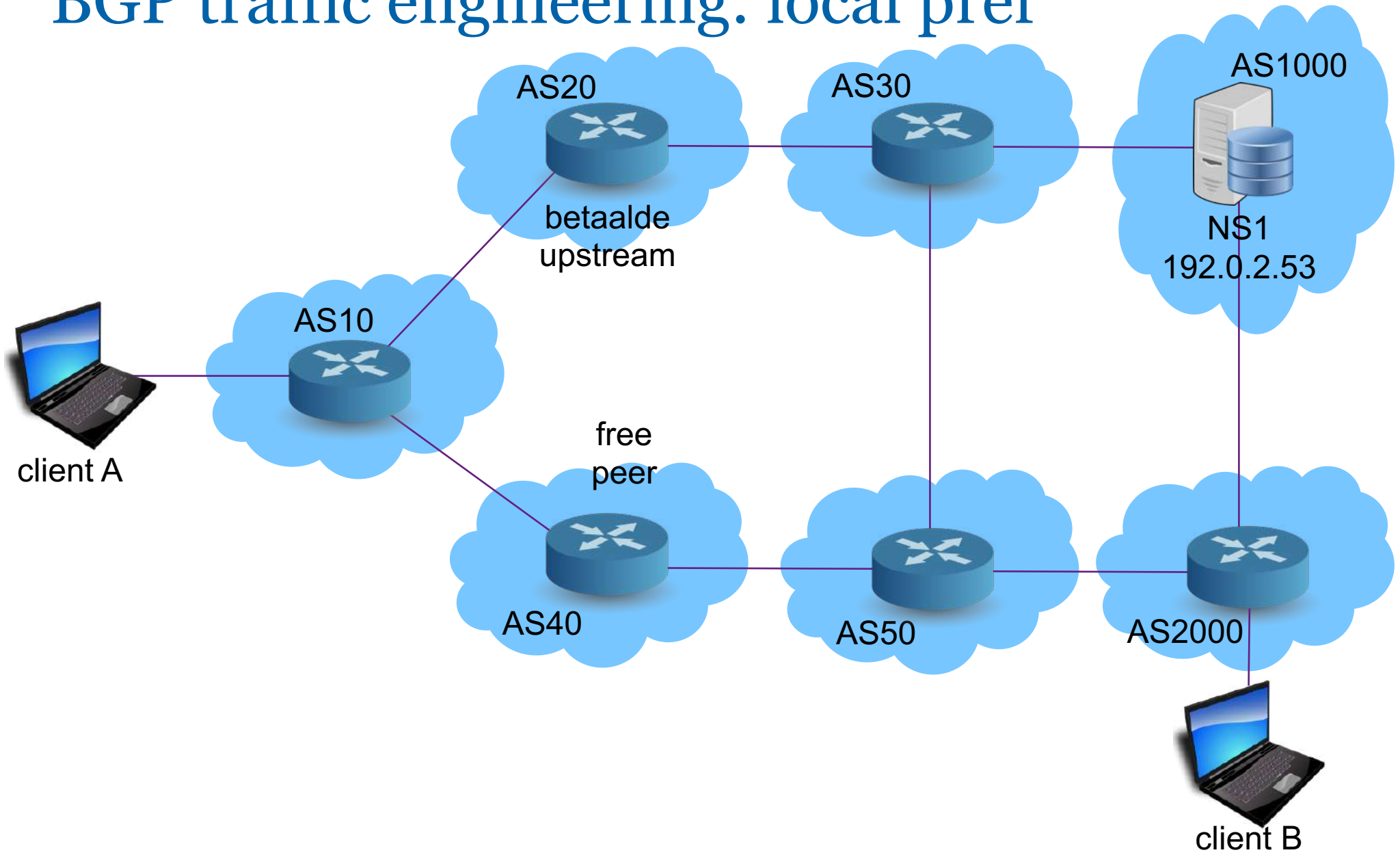
BGP (in a nutshell)



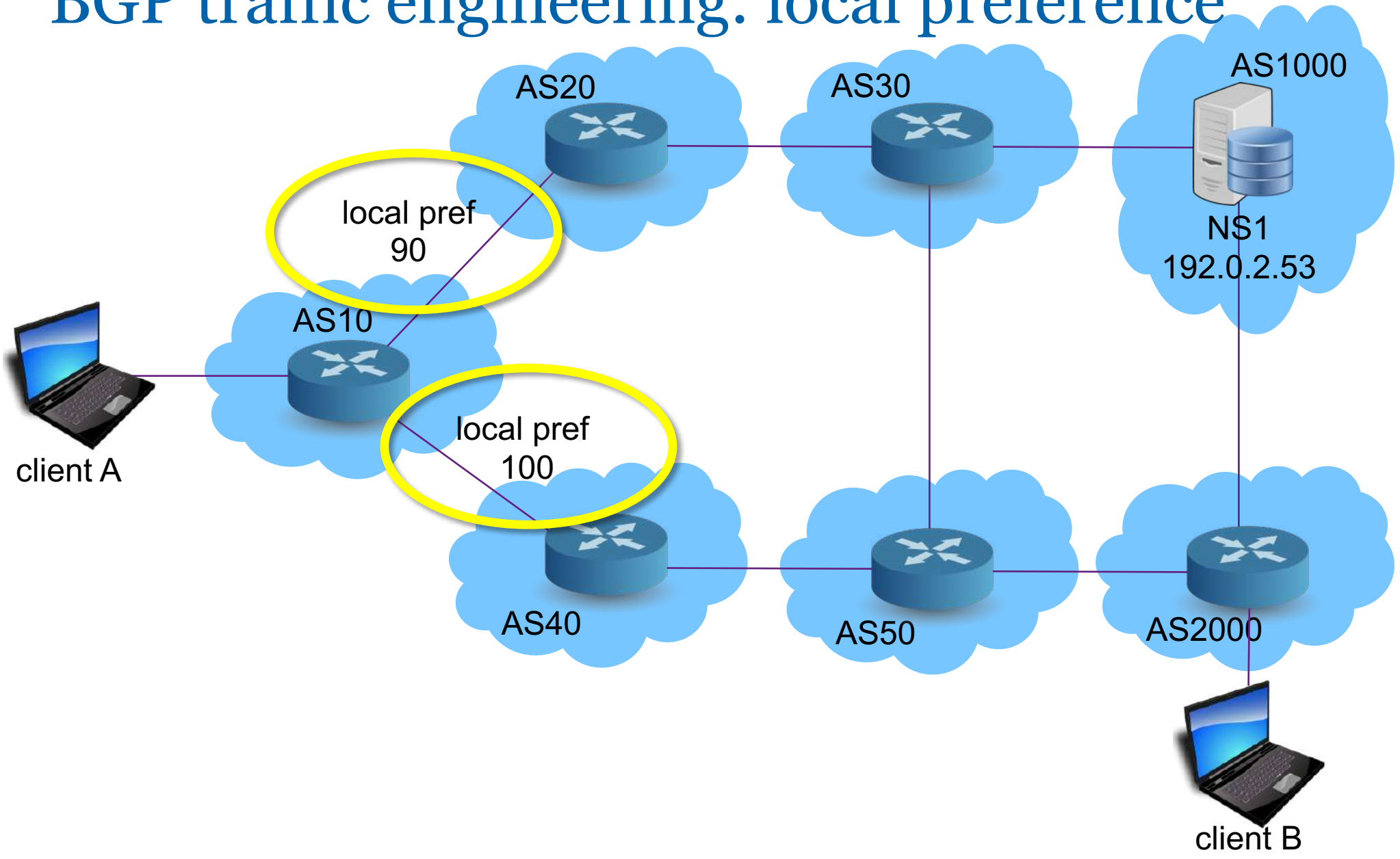
Nieuwe beste AS-path van A naar NS1: 10, 40, 50, 30, 1000

Traffic engineering

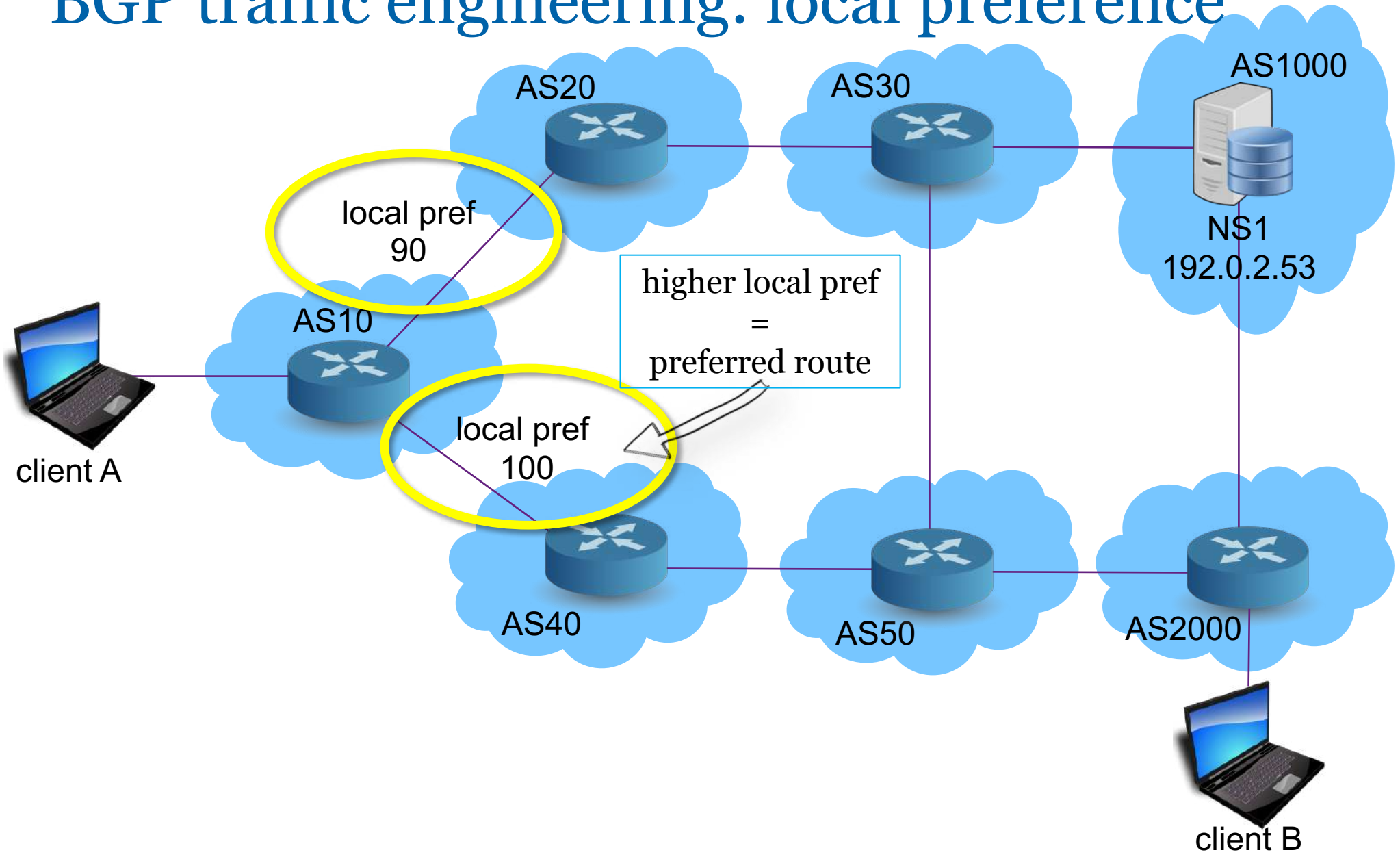
BGP traffic engineering: local pref



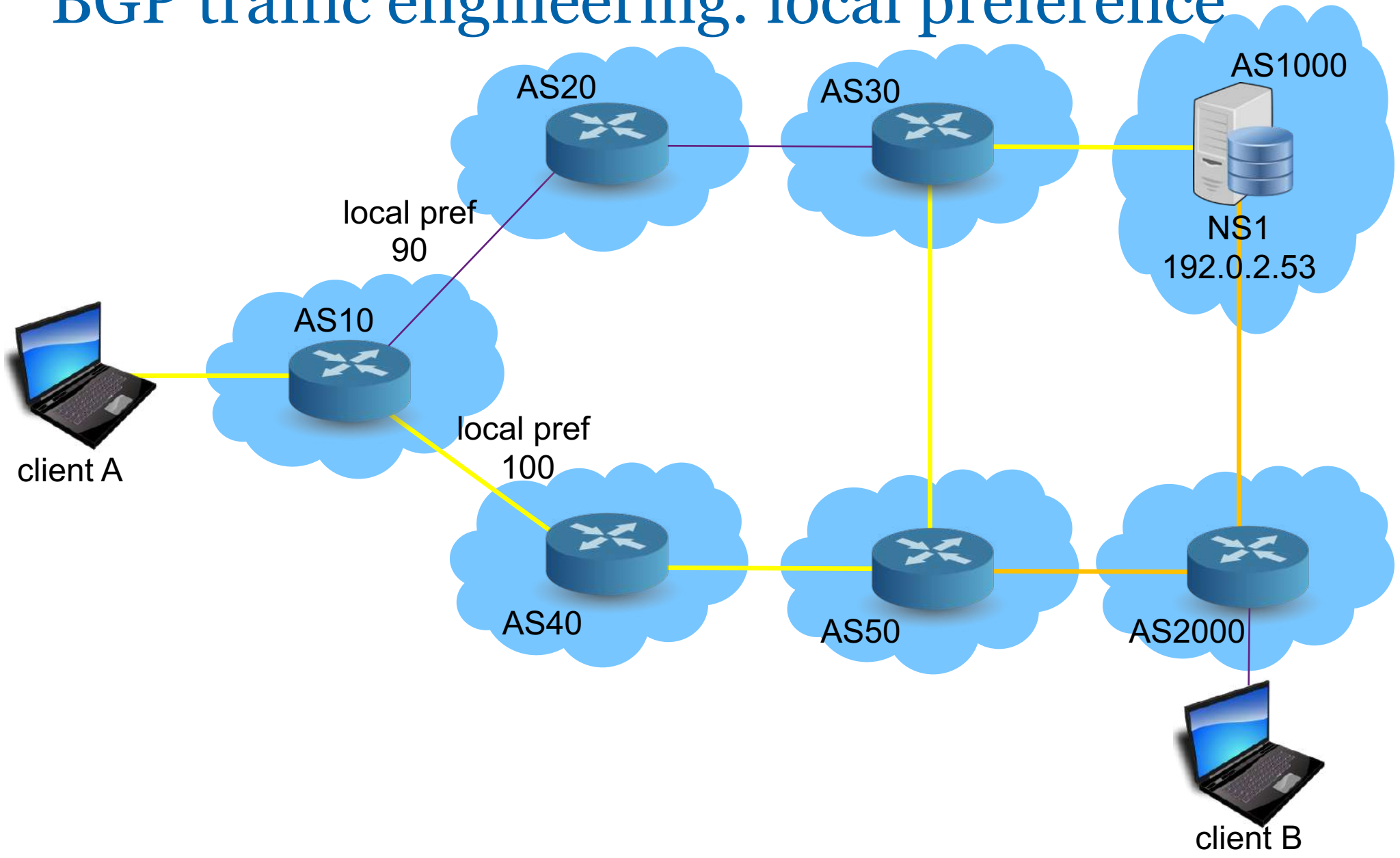
BGP traffic engineering: local preference



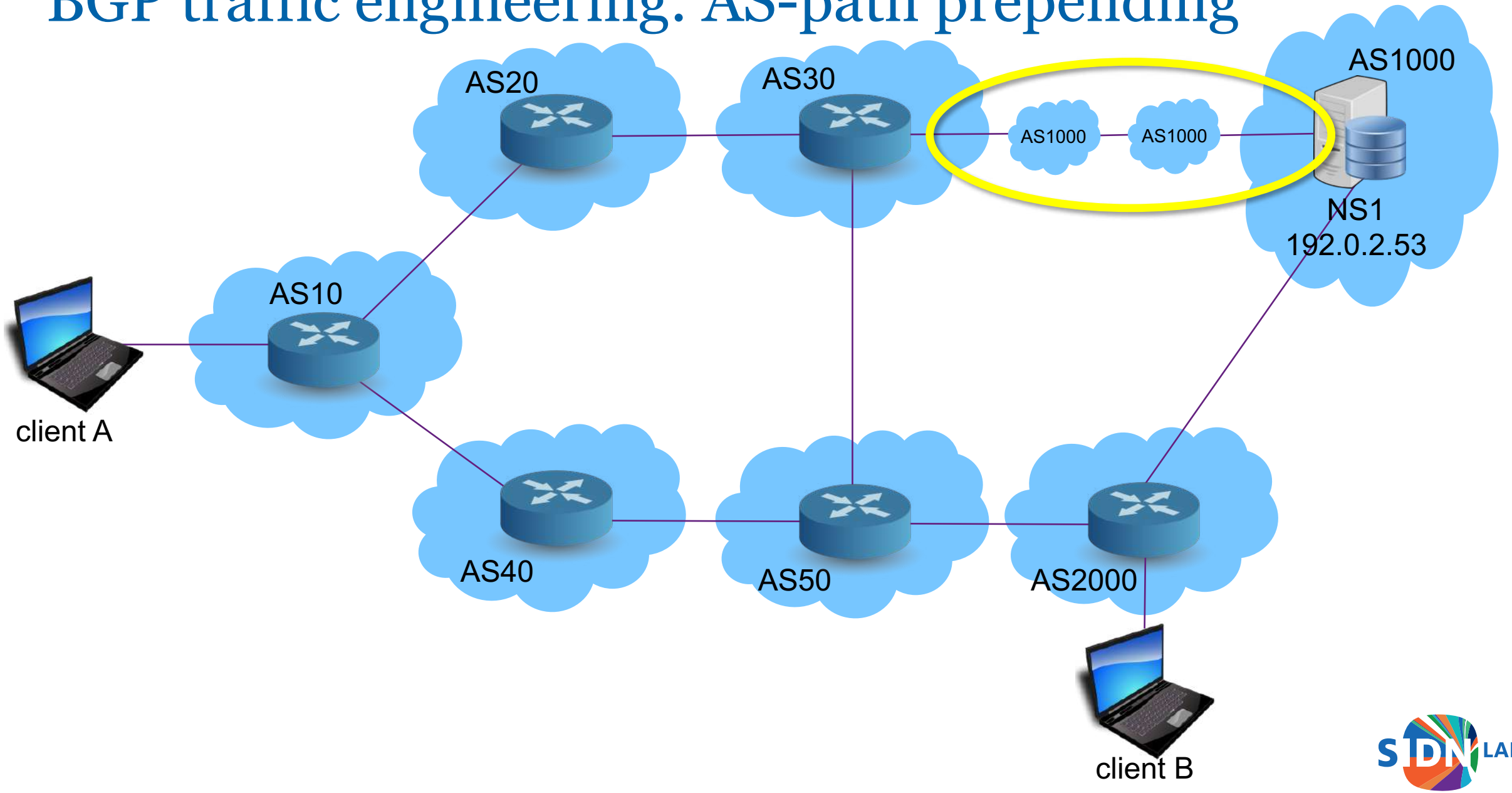
BGP traffic engineering: local preference



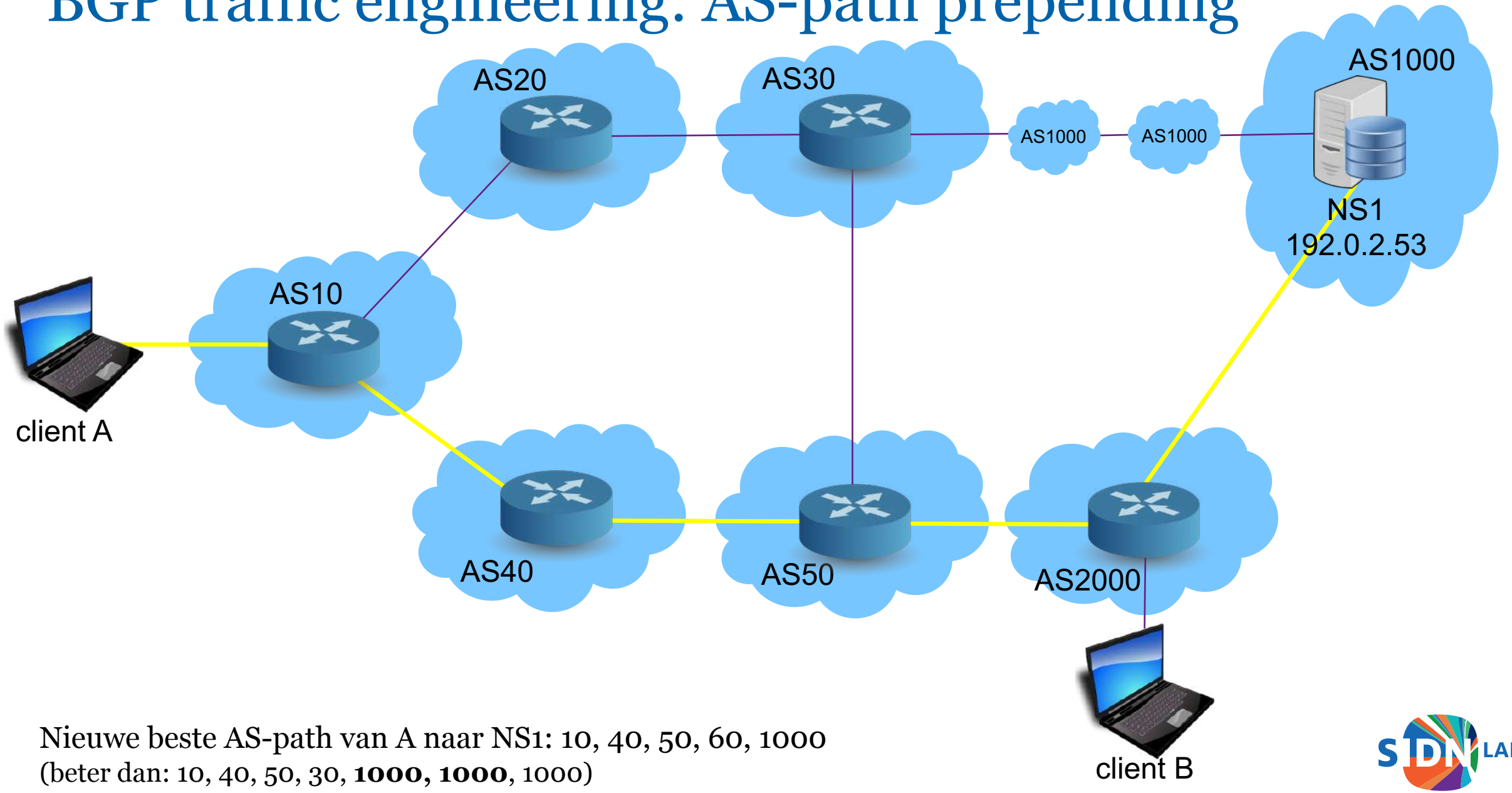
BGP traffic engineering: local preference



BGP traffic engineering: AS-path prepending



BGP traffic engineering: AS-path prepending



Nieuwe beste AS-path van A naar NS1: 10, 40, 50, 60, 1000
(beter dan: 10, 40, 50, 30, **1000, 1000**, 1000)

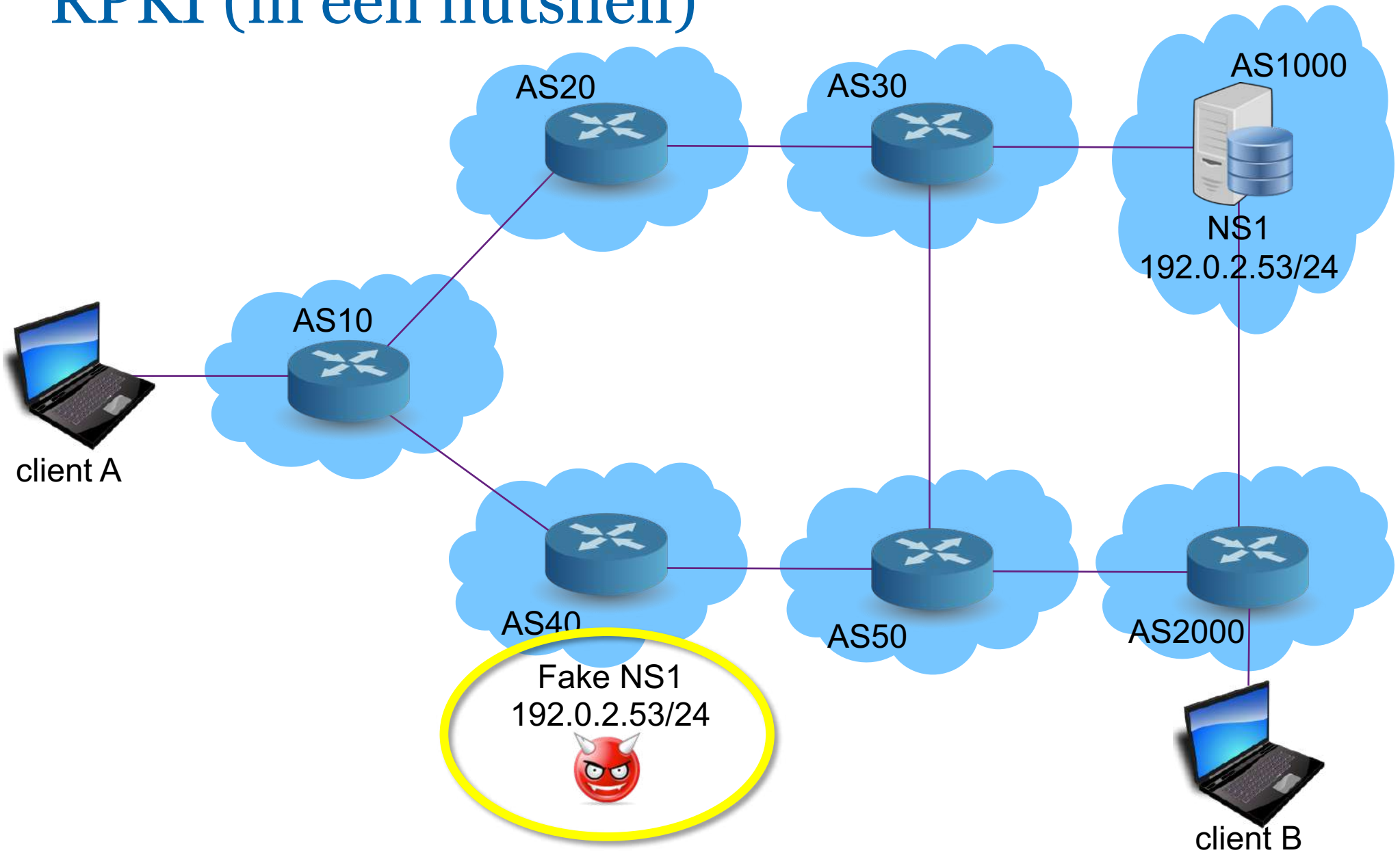
RPKI: Resource Public Key Infrastructure

- Een public key infrastructuur om BGP te beveiligen
- Resource certification van IP-prefixes / ASN combinatie
- Voorkomt (tot op zekere hoogte) route hijacking
- Er zijn twee kanten: het publiceren van ROA's en het valideren ervan.
- Origin validation, **geen** path validation (dat is BGPSEC, nog steeds in de maak)

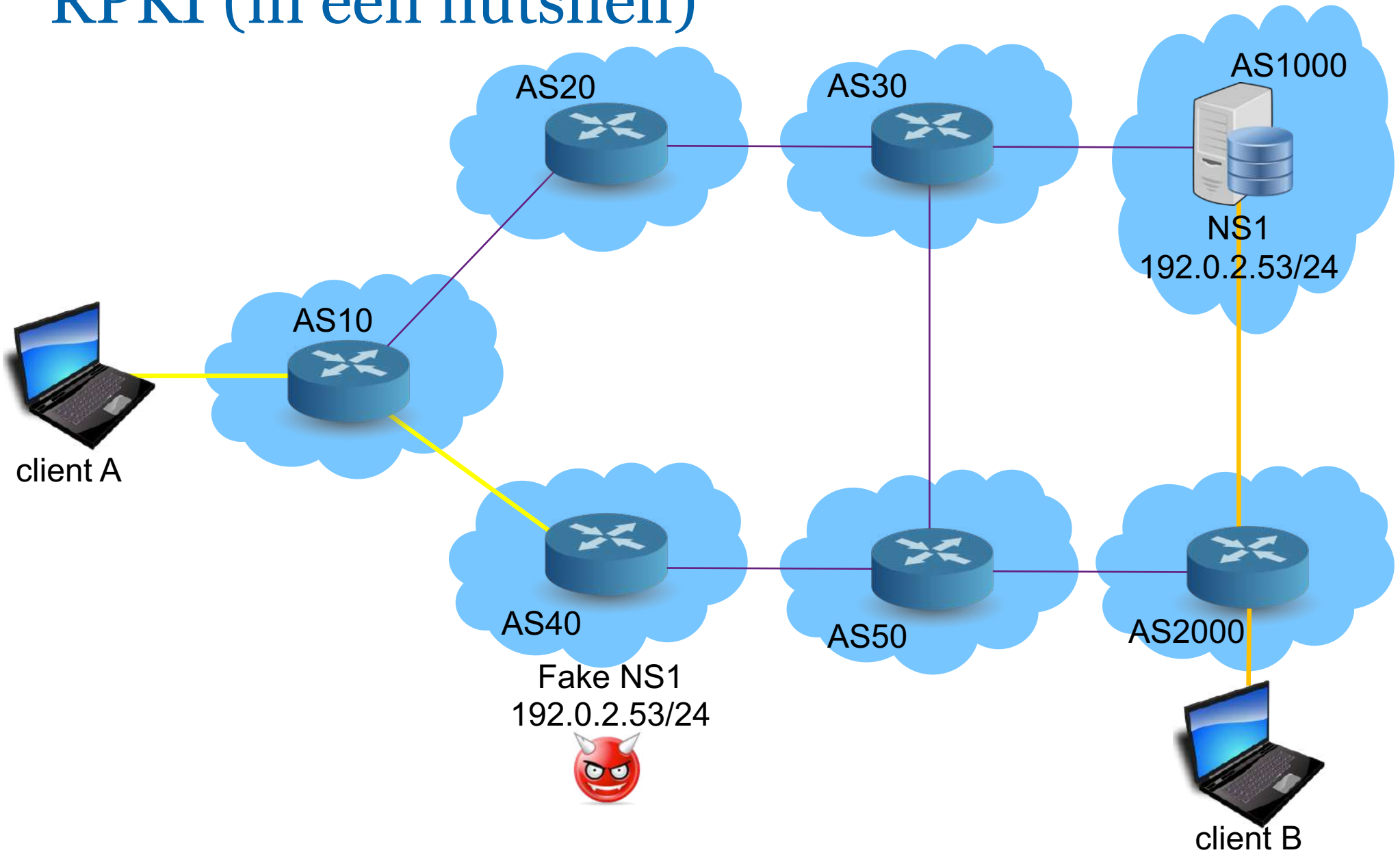
Probeer je eigen ISP: https://isbgpsafeyet.com/ of https://rpkitest.nlnetlabs.net/

SIDN LABS

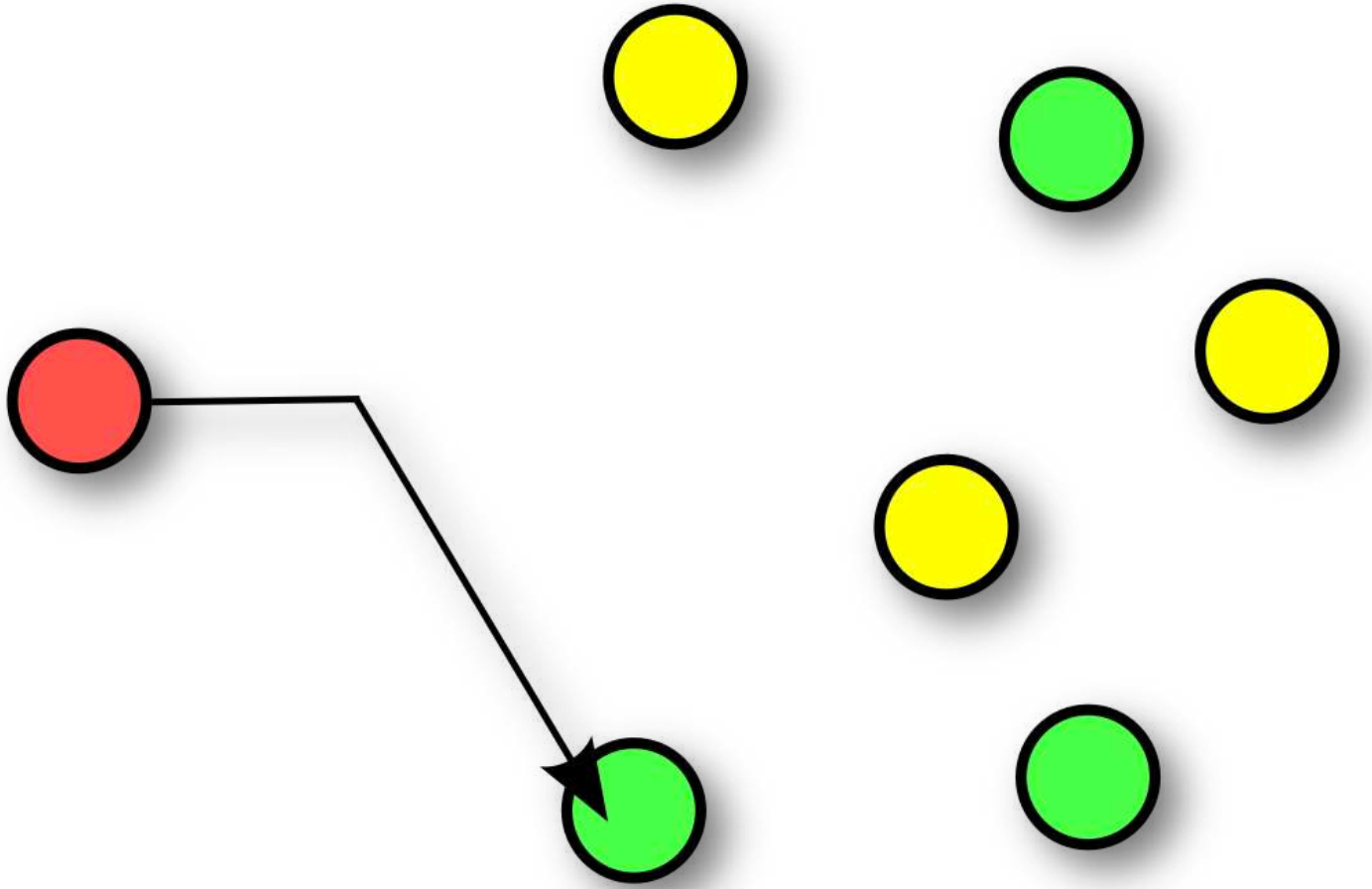
RPKI (in een nutshell)



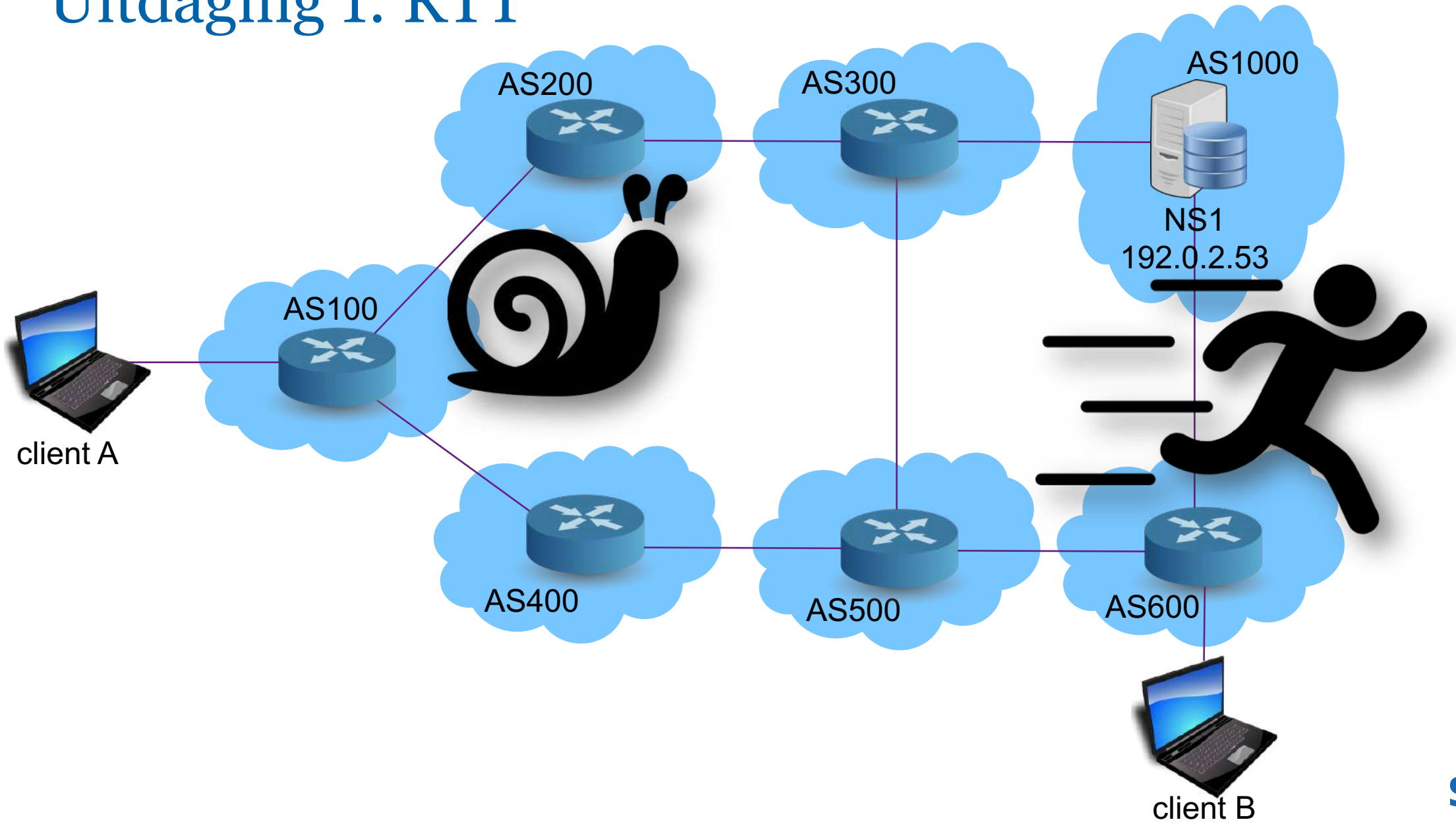
RPKI (in een nutshell)



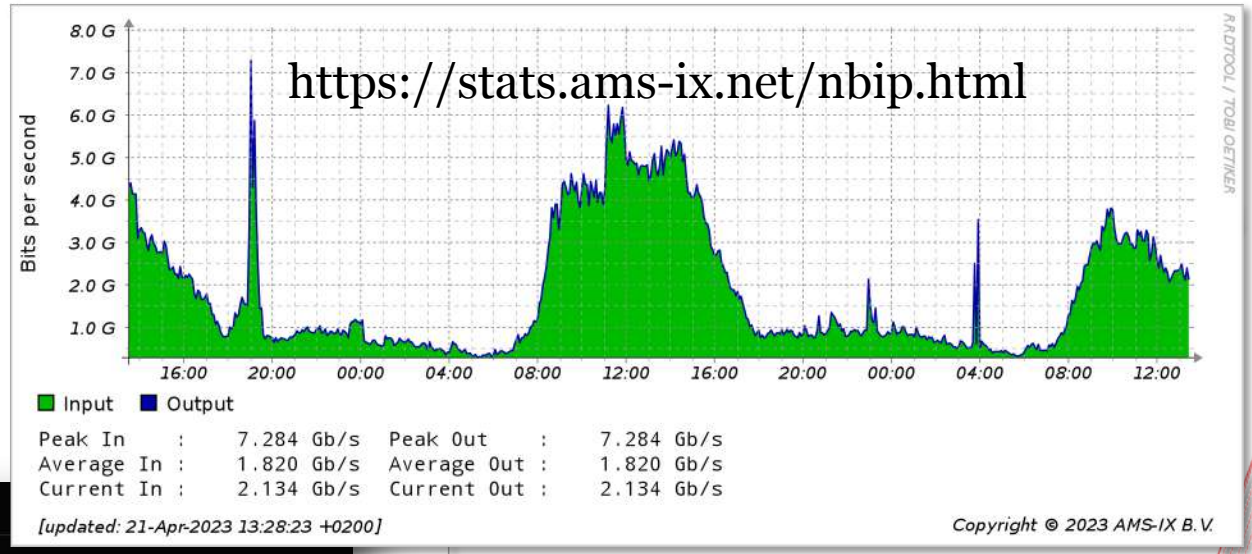
Anycast en waarom het een goed idee is

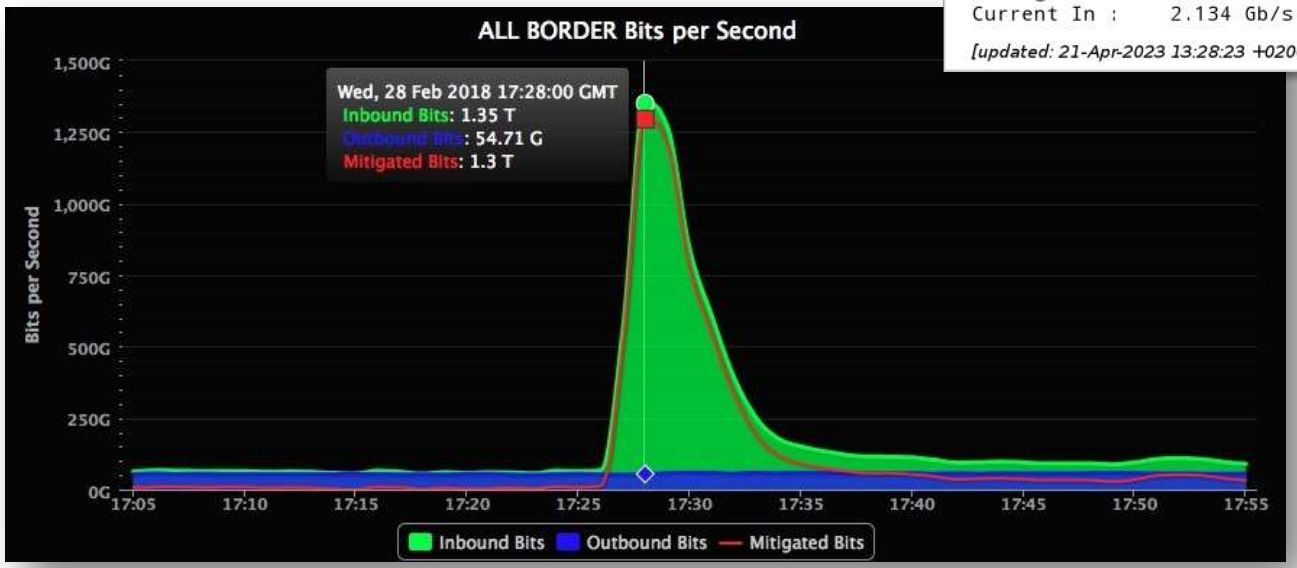


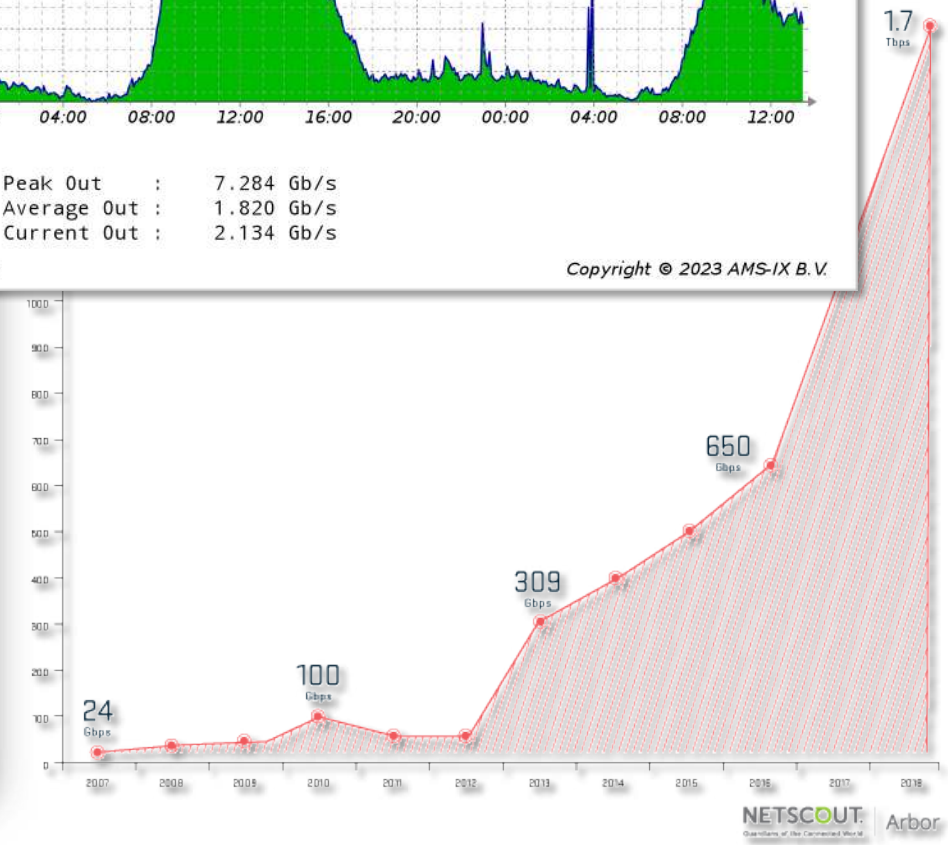
Uitdaging 1: RTT



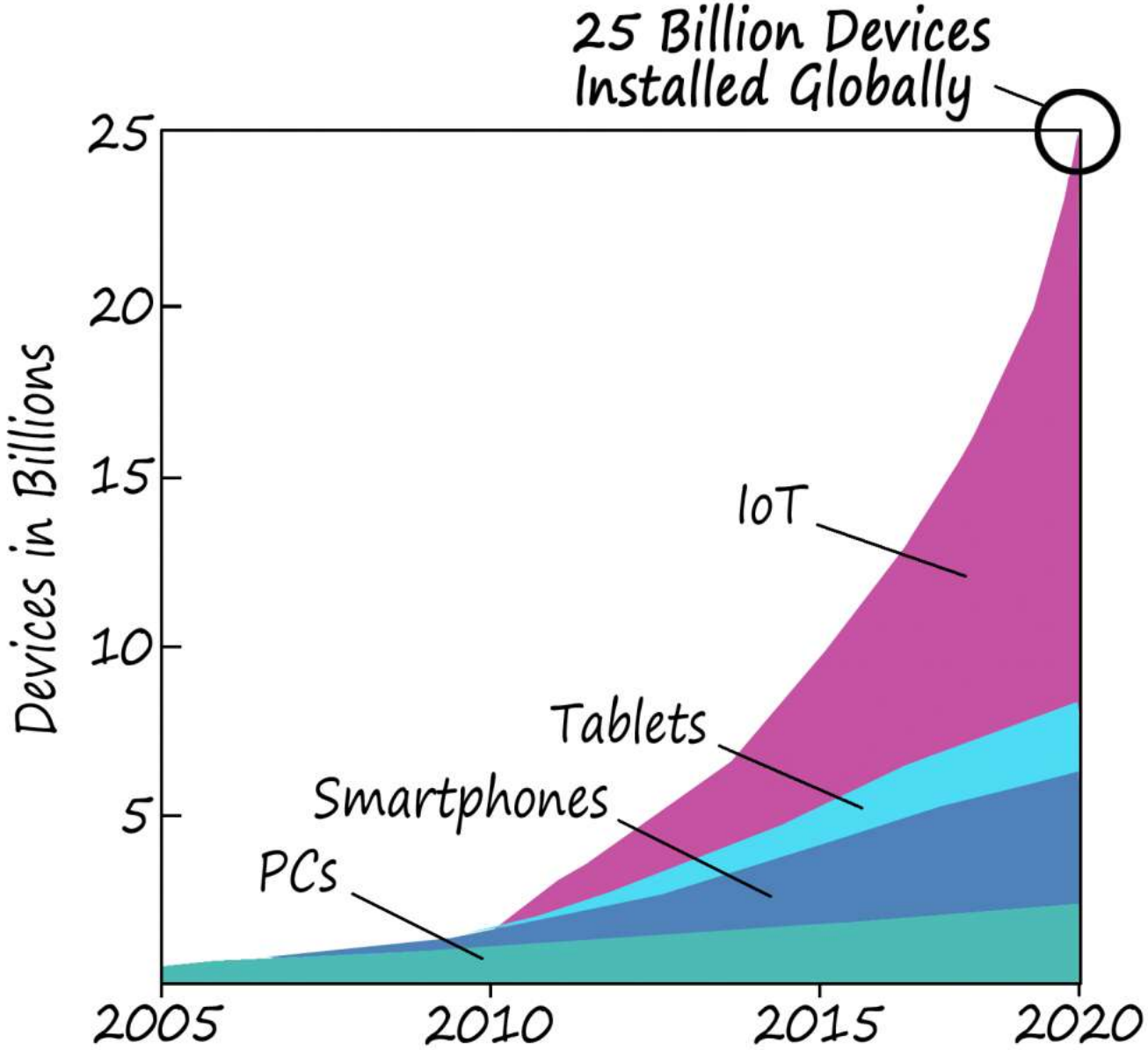
Uitdaging 2: DDoS



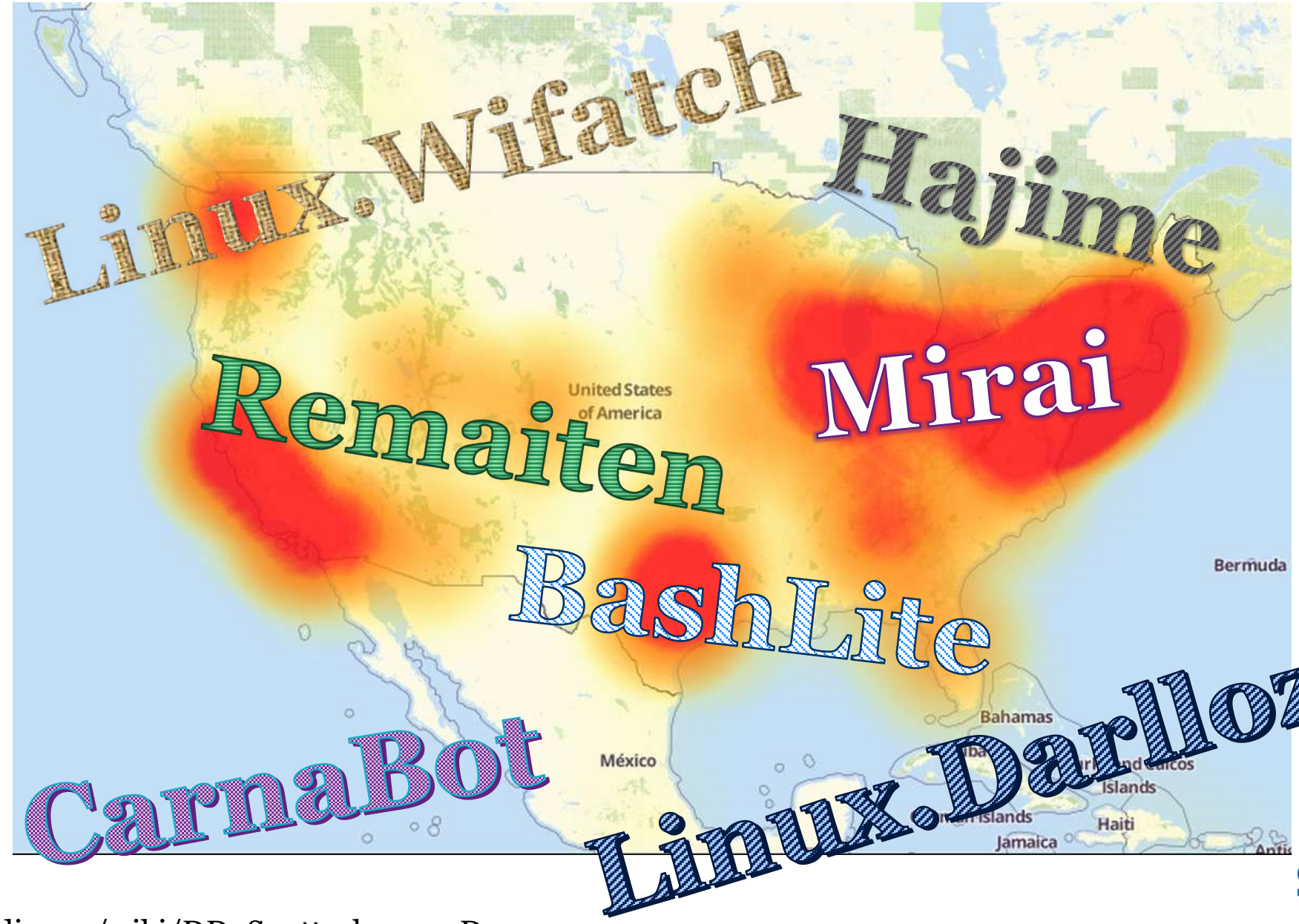




Belangrijkste reden: IoT devices

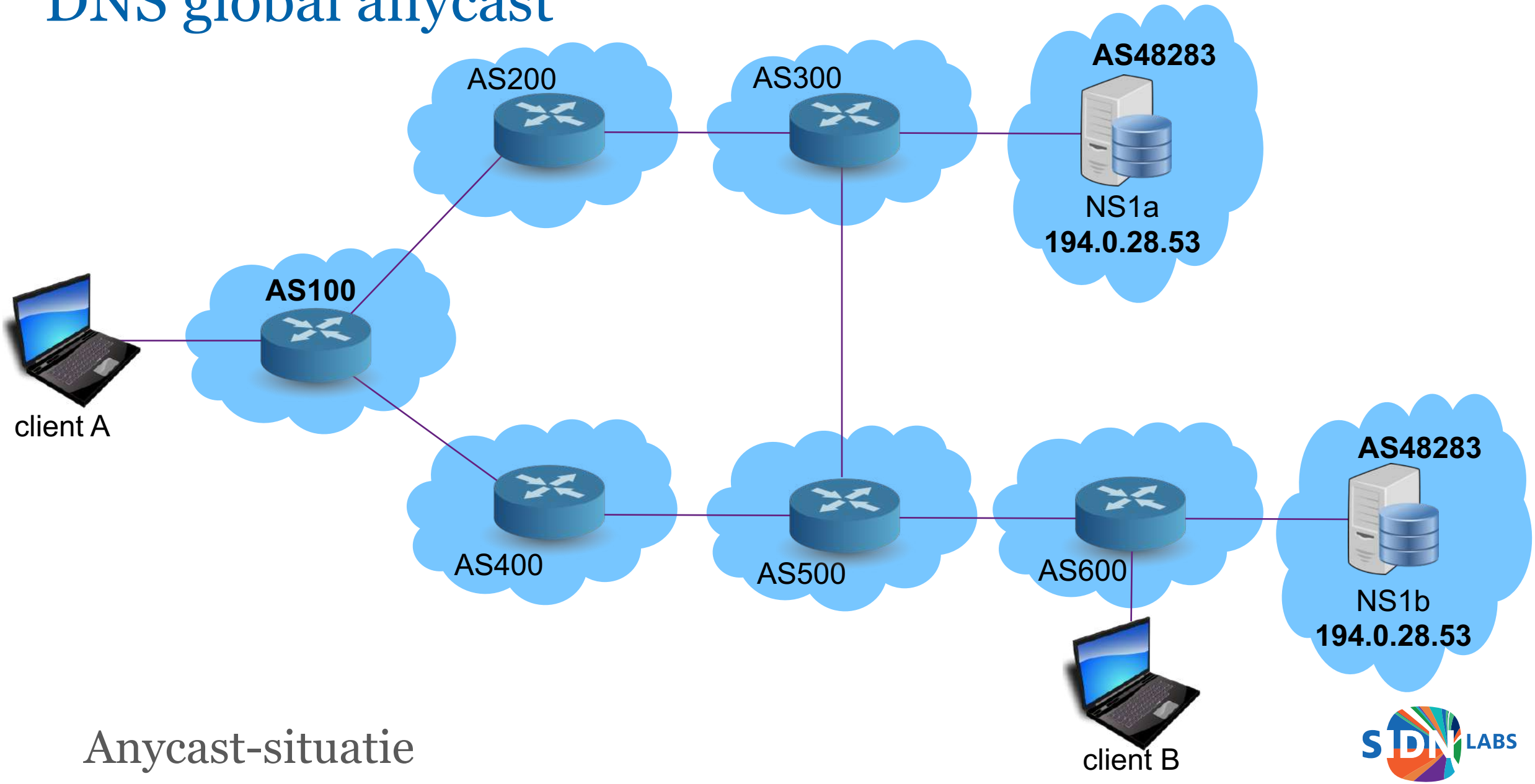


IoT botnets



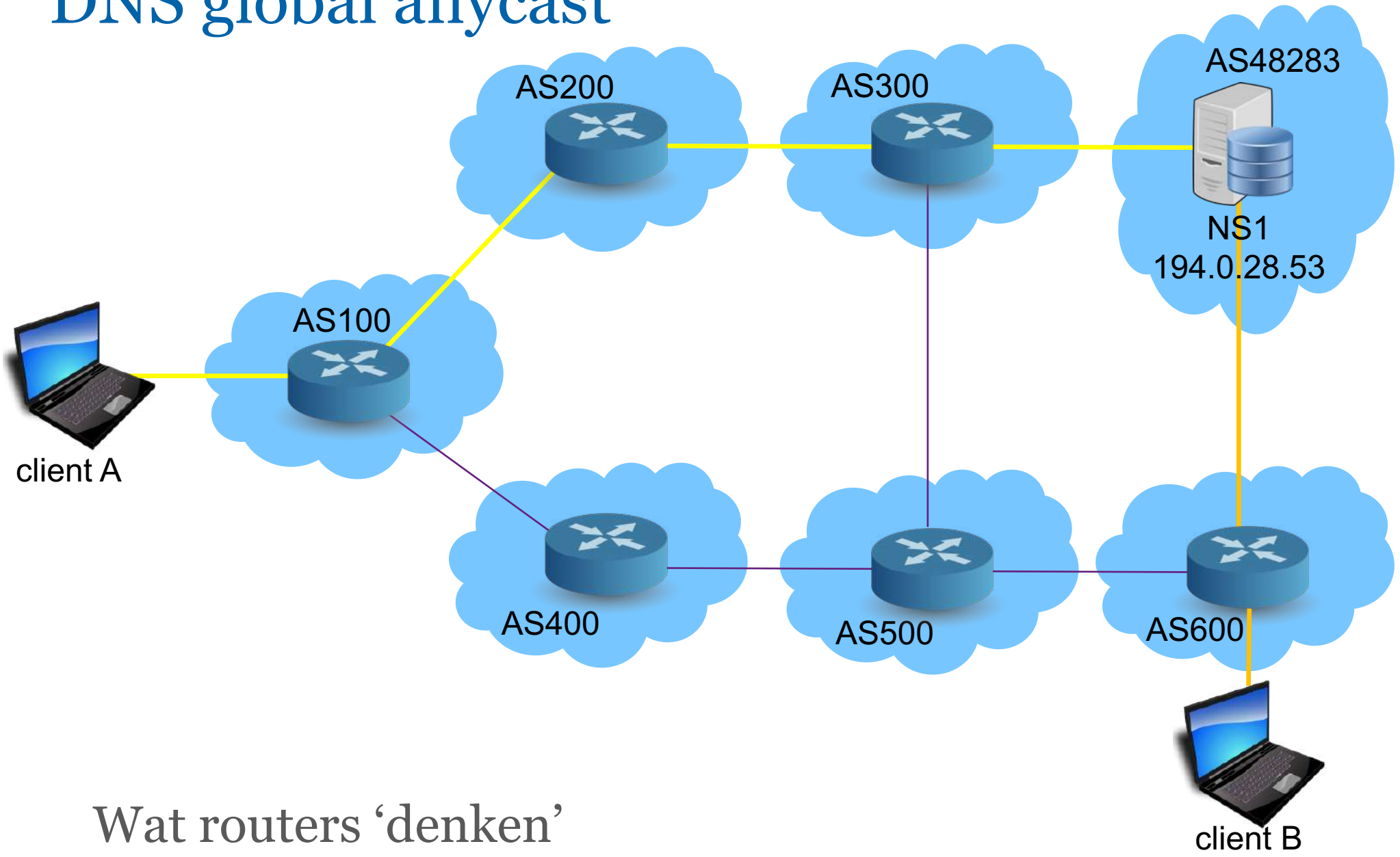
Bron: https://en.wikipedia.org/wiki/DDoS_attacks_on_Dyn

DNS global anycast



Anycast-situatie

DNS global anycast

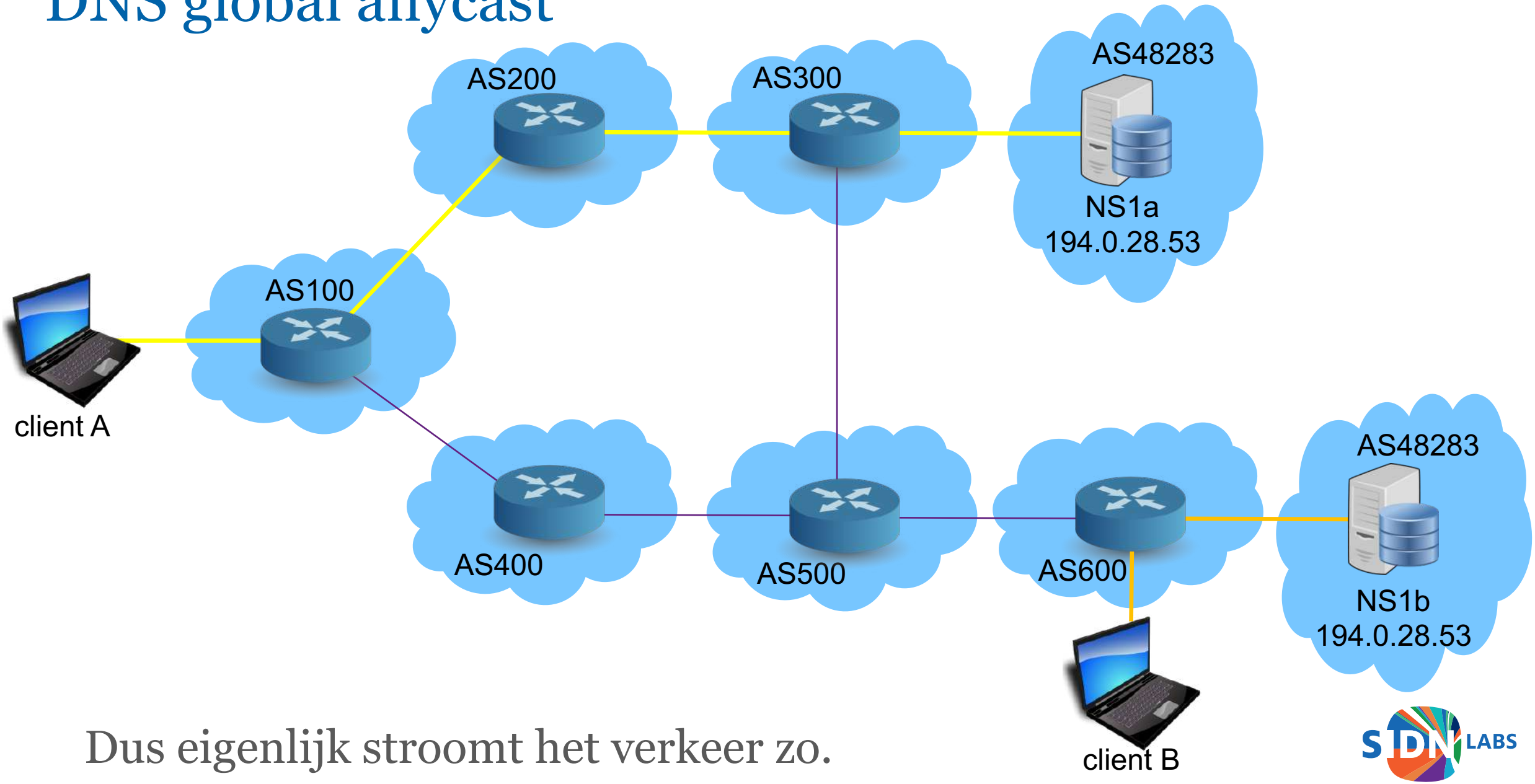


Wat routers ‘denken’

DNS global anycast

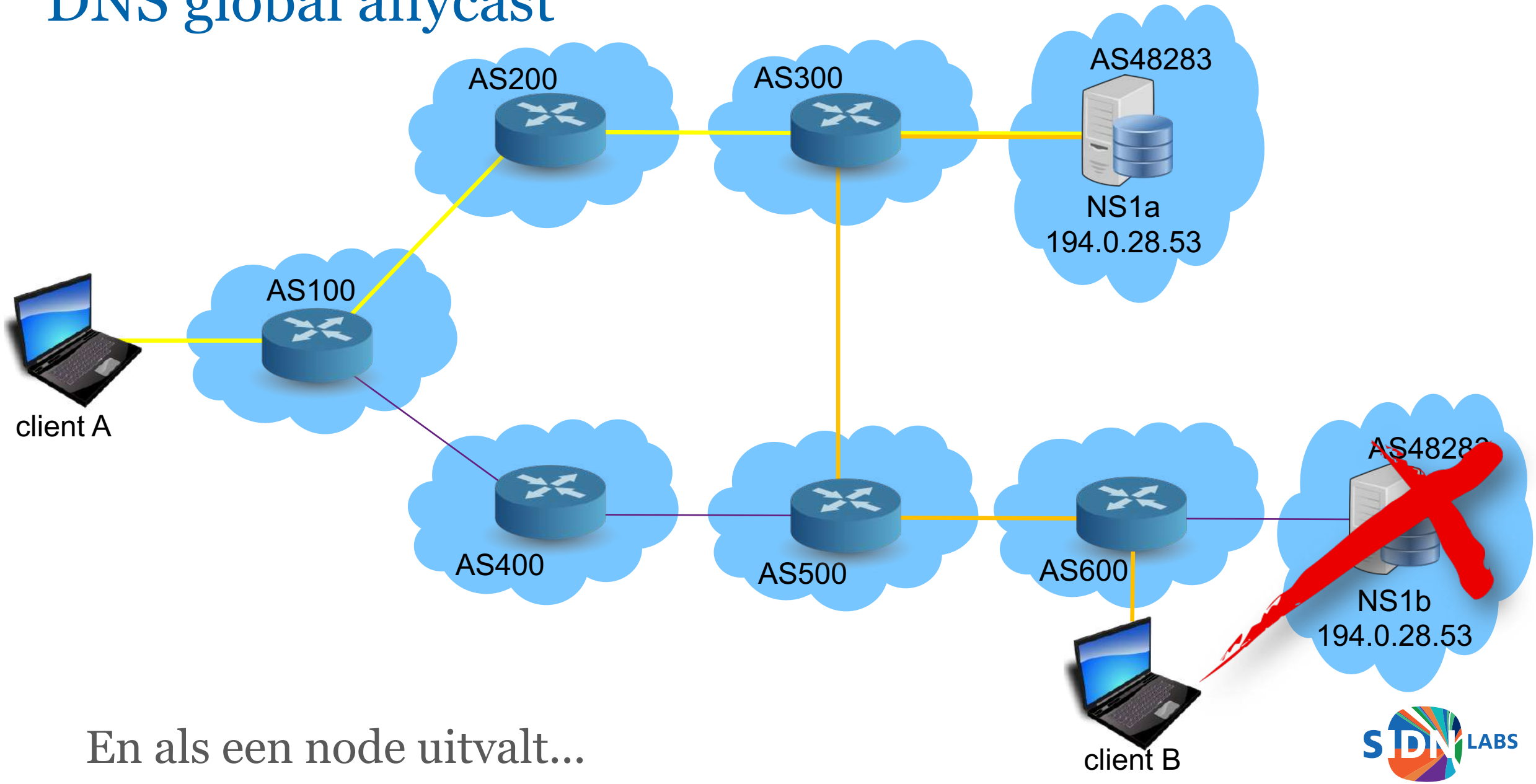
Maar onthoud, dit is de werkelijke situatie!

DNS global anycast



Dus eigenlijk stroomt het verkeer zo.

DNS global anycast



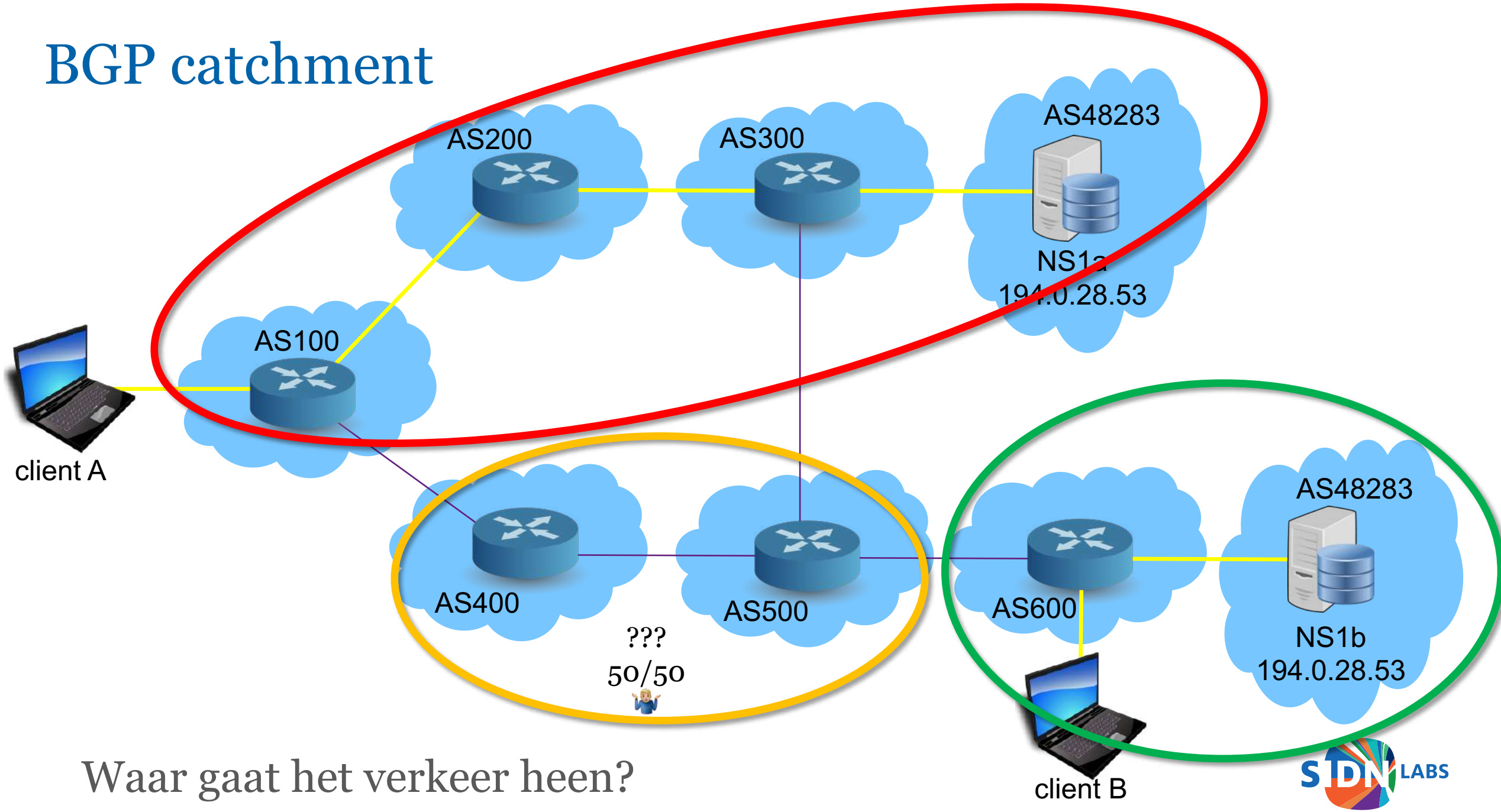
En als een node uitvalt…

DNS **global** anycast

- Gewoon een slimme 'network hack' voor extra *resilience*.
- En betere prestaties (kortere RTT's)
- Werkt met BGP
- Goed begrepen oplossing, reeds op veel plaatsen ingezet
- De DNS root servers (al vele jaren)
- 1.1.1.1, 8.8.8.8, 9.9.9.9, 64.6.64.6, OpenDNS en meer
- Oorspronkelijk alleen in UDP-omgevingen
- Maar ook bewezen in TCP-omgevingen (zoals Cloudflare)

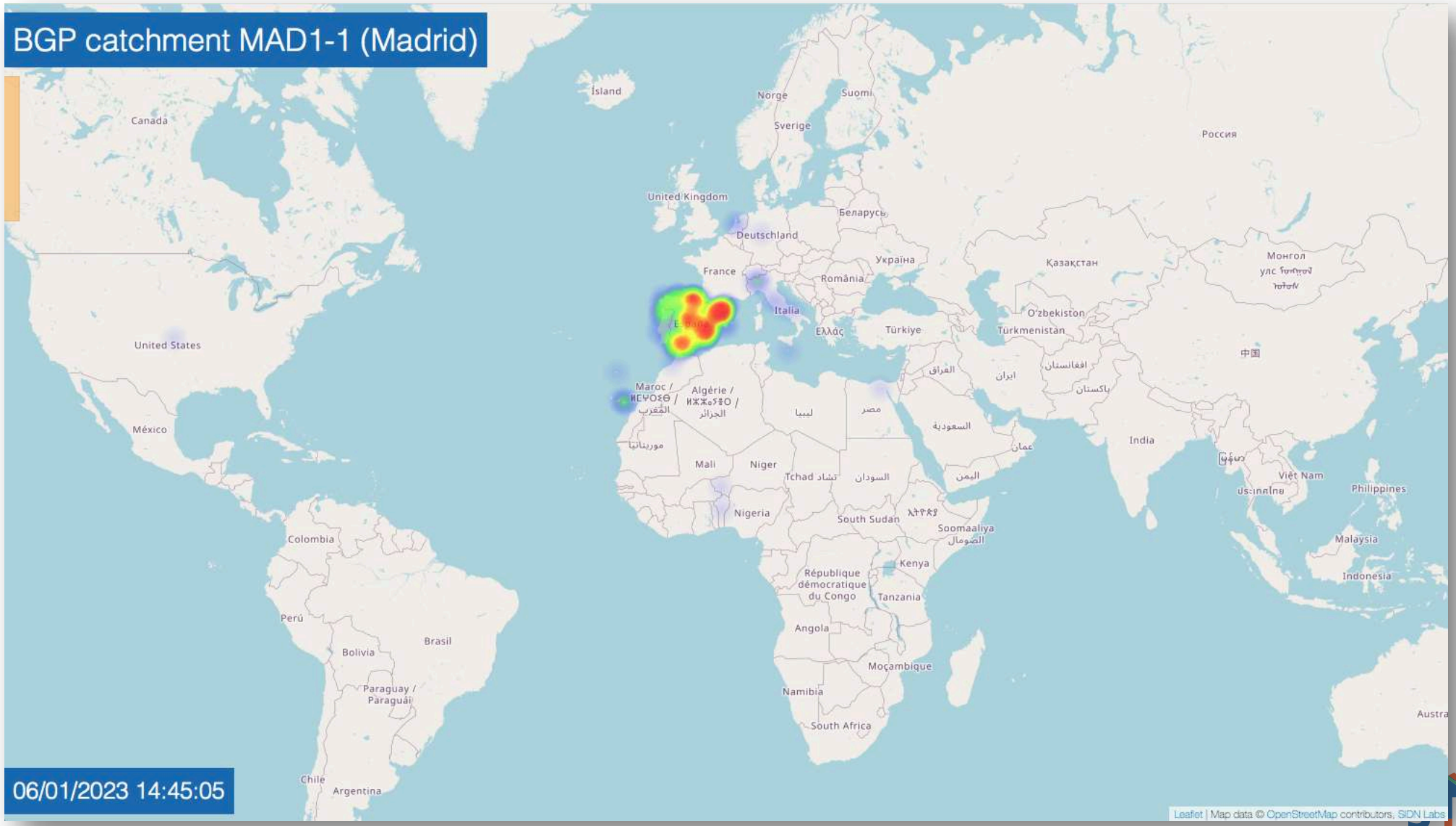
SIDN LABS

BGP catchment



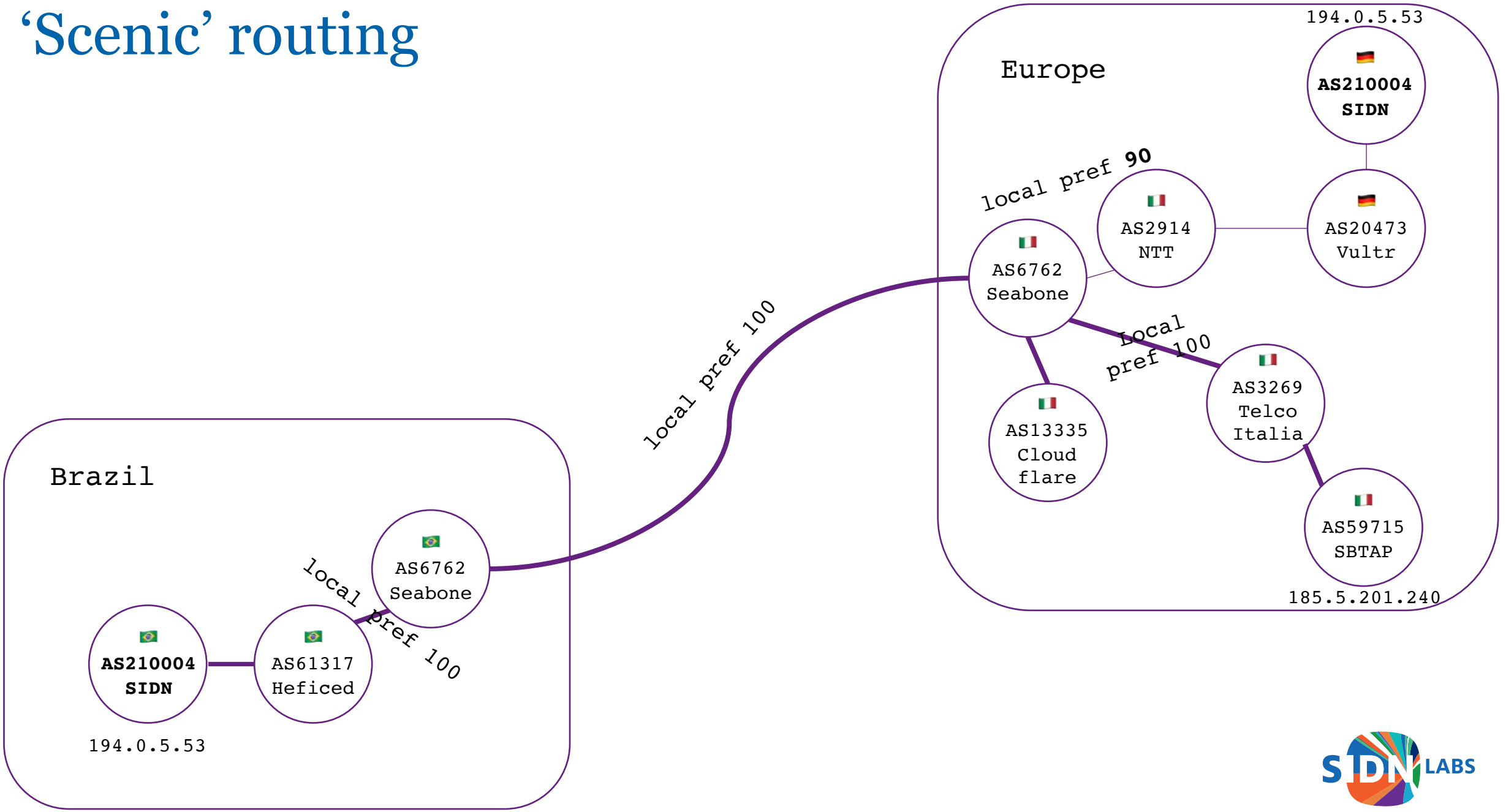
Waar gaat het verkeer heen?

BGP *catchment* Madrid

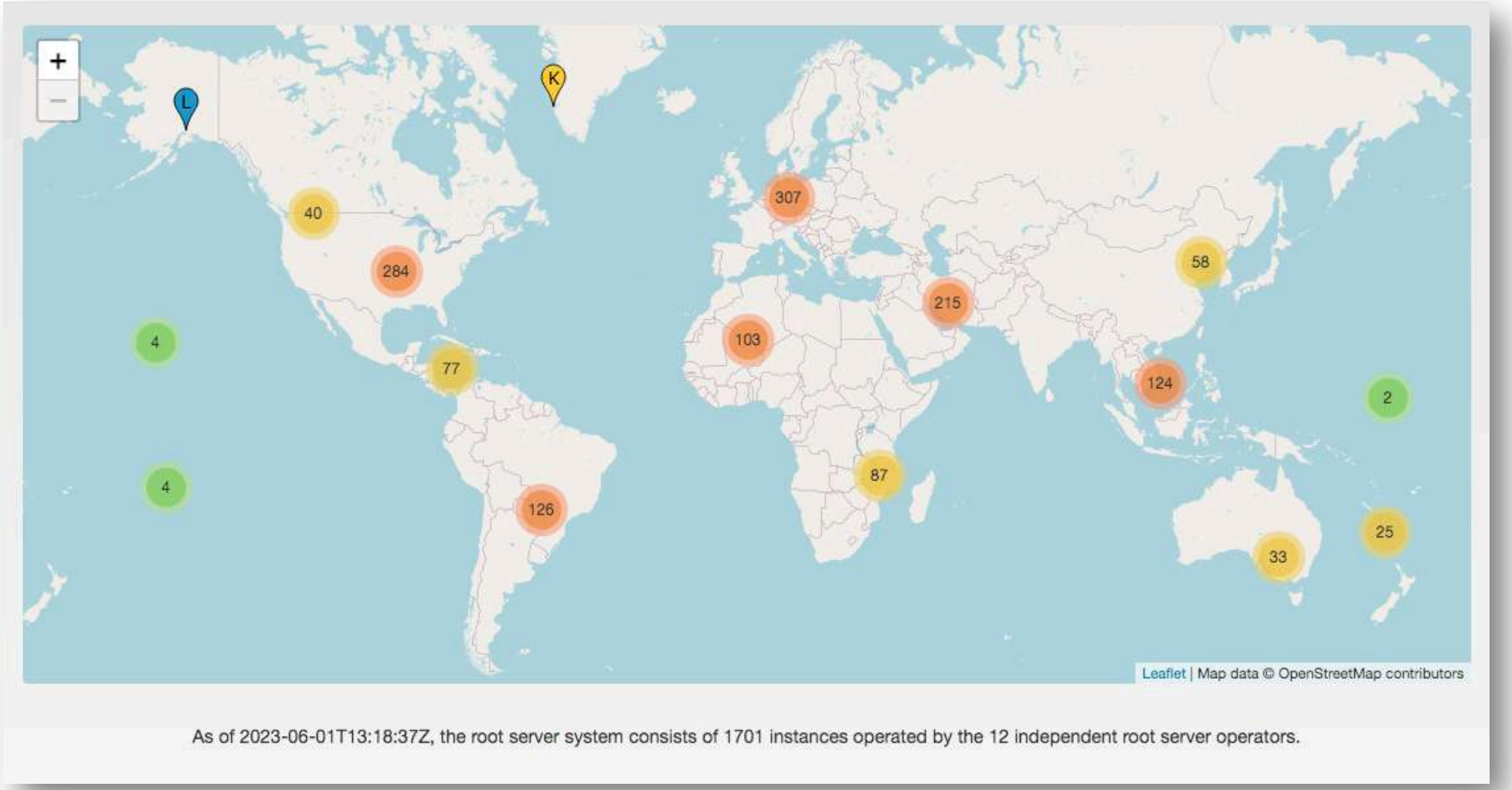


http://dnstest.nl/anycast2020/heatmaps/mad1-1

‘Scenic’ routing



DNS globale anycast (voor de ‘root’)



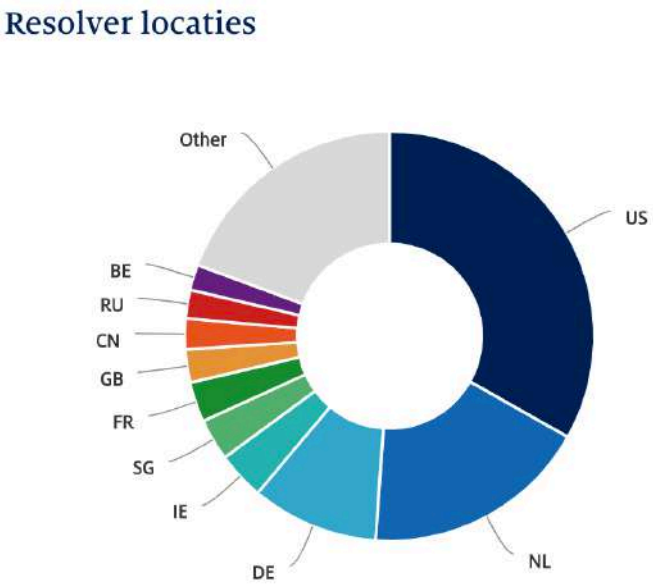
1701 servers!

http://www.root-servers.org/

[this page was intentionally left blank]

Intro en achtergrond

- 5 jaar geleden: unicast DNS met focus op NL
- DDoS aanvallen
- Samenvoegingen van providers
- 60% queries van Big Tech
- Gebruik van anycast-diensten van anderen



Bron:https://stats.sidnlabs.nl

Doelen: terug aan het roer

- ns1.dns.nl door SIDN
- Bouw van Team DNS
- Globale dekking (anycast)
- Snel uit te breiden, zowel virtueel als hardware
- DevOps mindset
- Geen eigen hardware
- Gestuurd door data

Bron: ThousandEyes (Cisco)

Uitdagingen

Zowel technisch als organisatorisch:

- CI/CD was grotendeels nieuw
- Bouw een DevOps team voor DNS
- Scrum was nu ineens logisch

GitOps: Beheers alle bewegende delen

Code in git zou een echte afspiegeling moeten zijn van de productie servers

- Bouw, test en uitrollen van een **gouden image** met een CI/CD pijplijn
- Alle wijzigingen hebben een versie/*commit*
- Mogelijk om voor- en achteruit te gaan met releases
- **Dwing eenzelfde configuratie af** door regelmatig opnieuw uitrollen

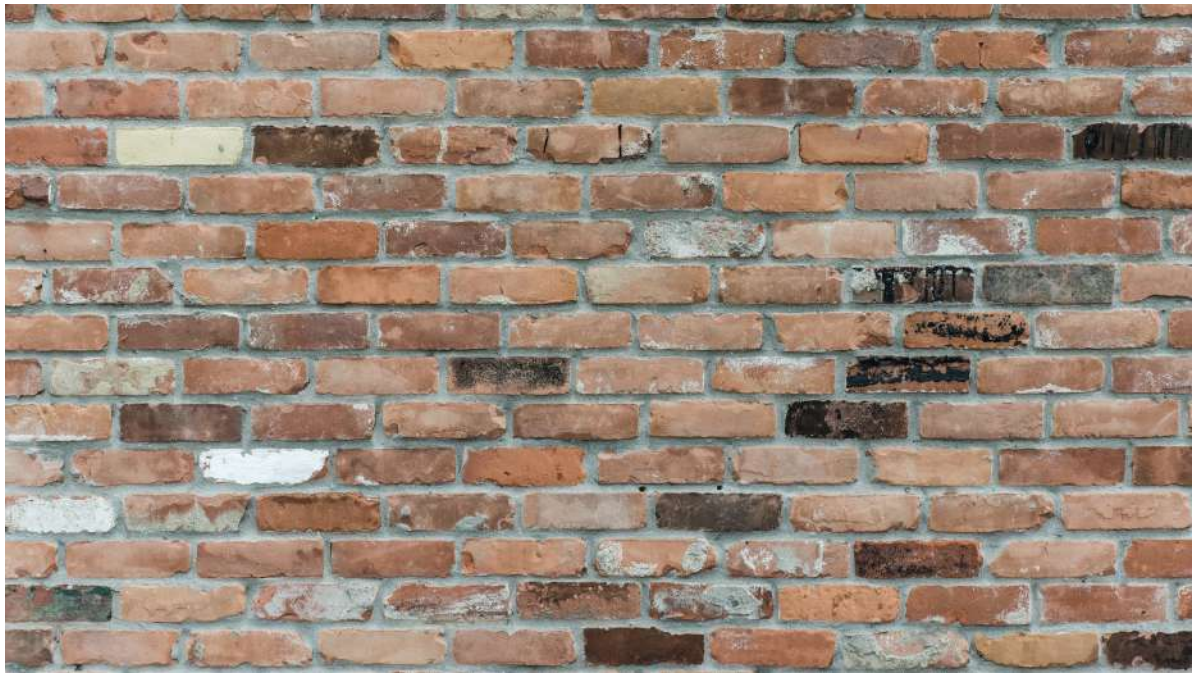


Foto door Math op Unsplash

Een artefact uitrol

Past bij perfect bij DNS

- Anycast zorgt voor:
  - *loadbalancing*
  - *fault tolerance*
- DNS is (bijna) zonder state
- Heel weinig configuratiemanagement nodig na de uitrol
- Zone-data kan automatisch worden opgehaald
- PCAPs en andere statistieken moeten snel overgedragen worden

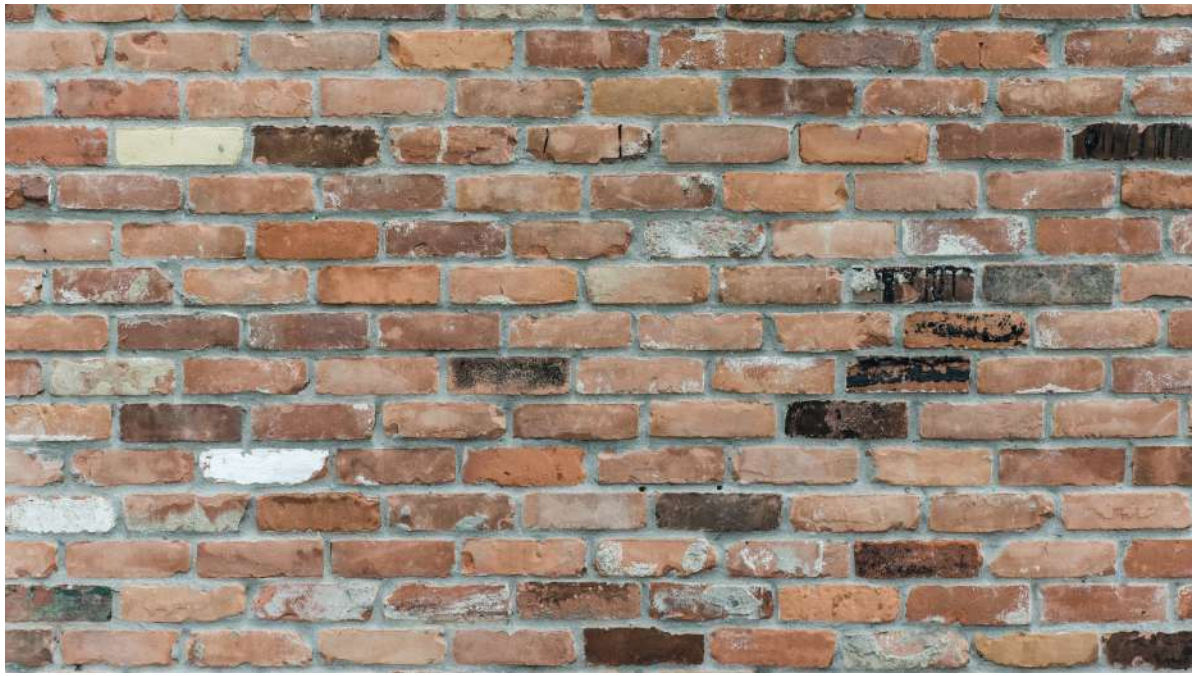
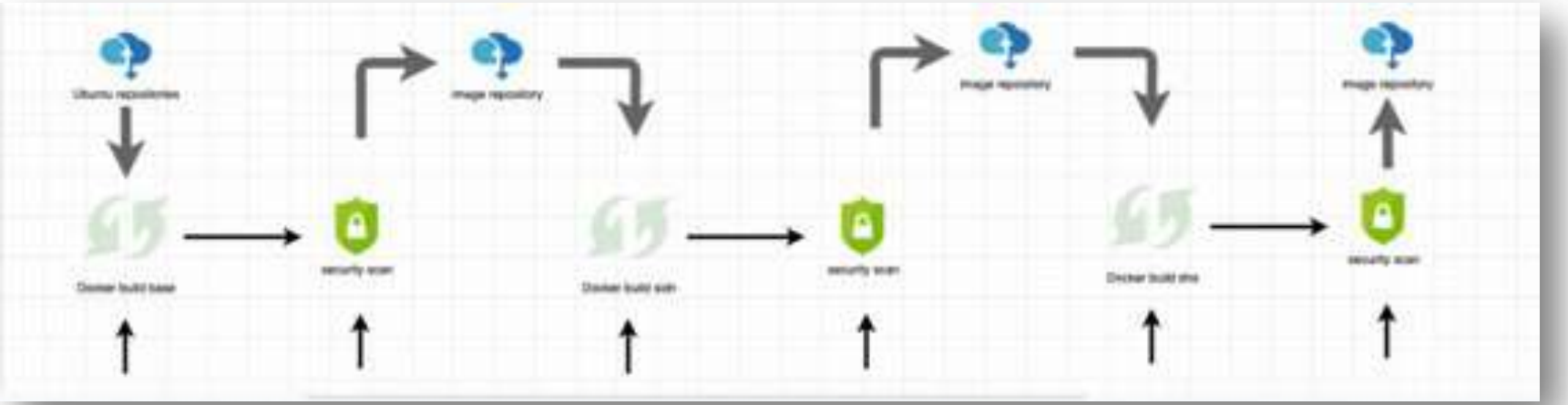


Foto door Math op Unsplash

Basis image bouw pijplijn

- Bouw een minimaal OS-image in 3 stappen
- Beveiliging scannen
- Platslaan en upload naar cloudopslag

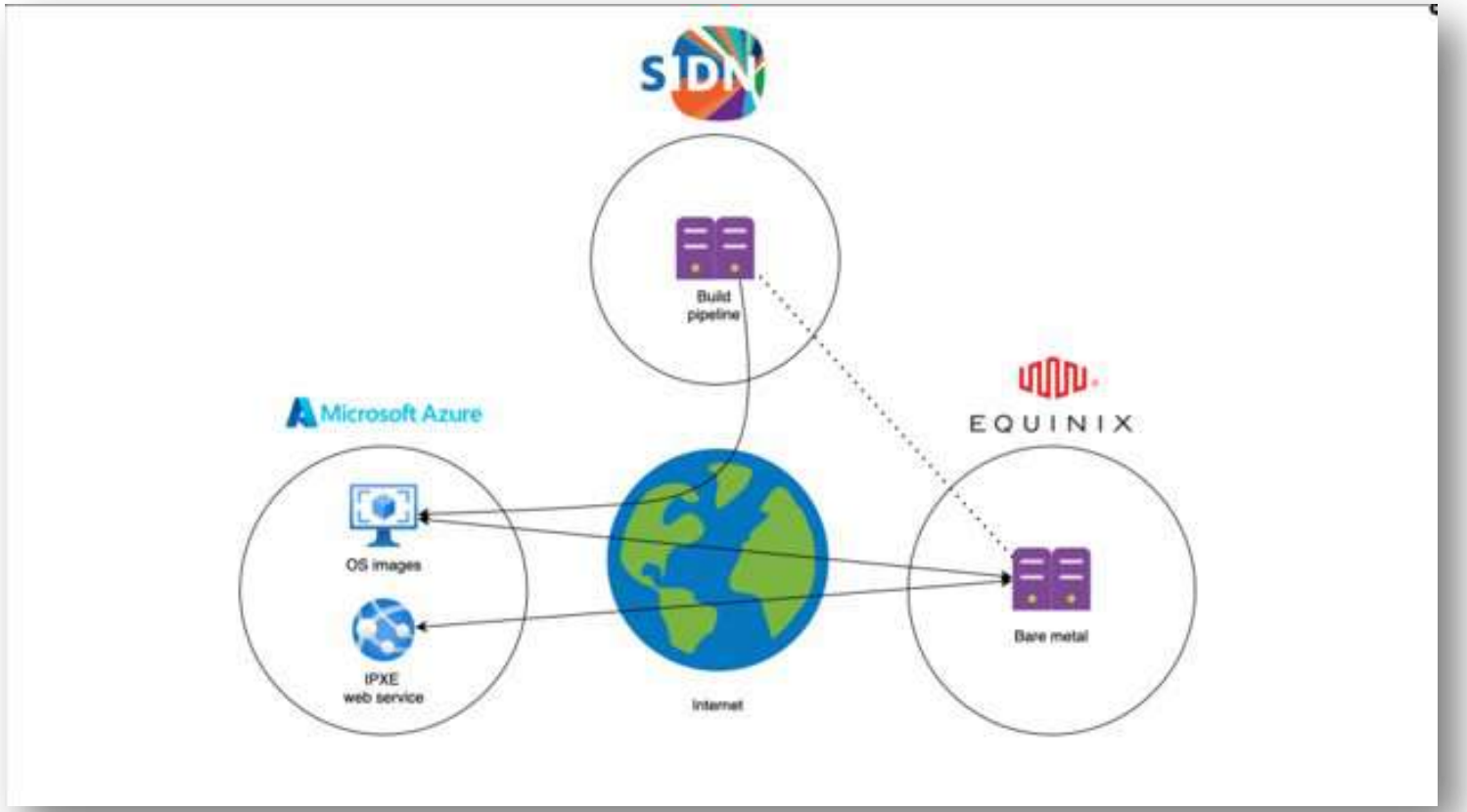


One-click uitrol op BMaaS

- Handmatige uitrol
- Binnen 10 minuten een actief systeem

One-click uitrol op BMaaS*

* Bare-Metal-as-a-Service



Samenvatting

Automatisering van:

- Bouwen en testen van een eigen image
- Dat image uitrollen op BMaaS
- Controles voor livegang
- BGP aanzetten
- Zone bijhouden
- Monitoring, statistieken en data verzameling
- De CI/CD pijplijn zelf

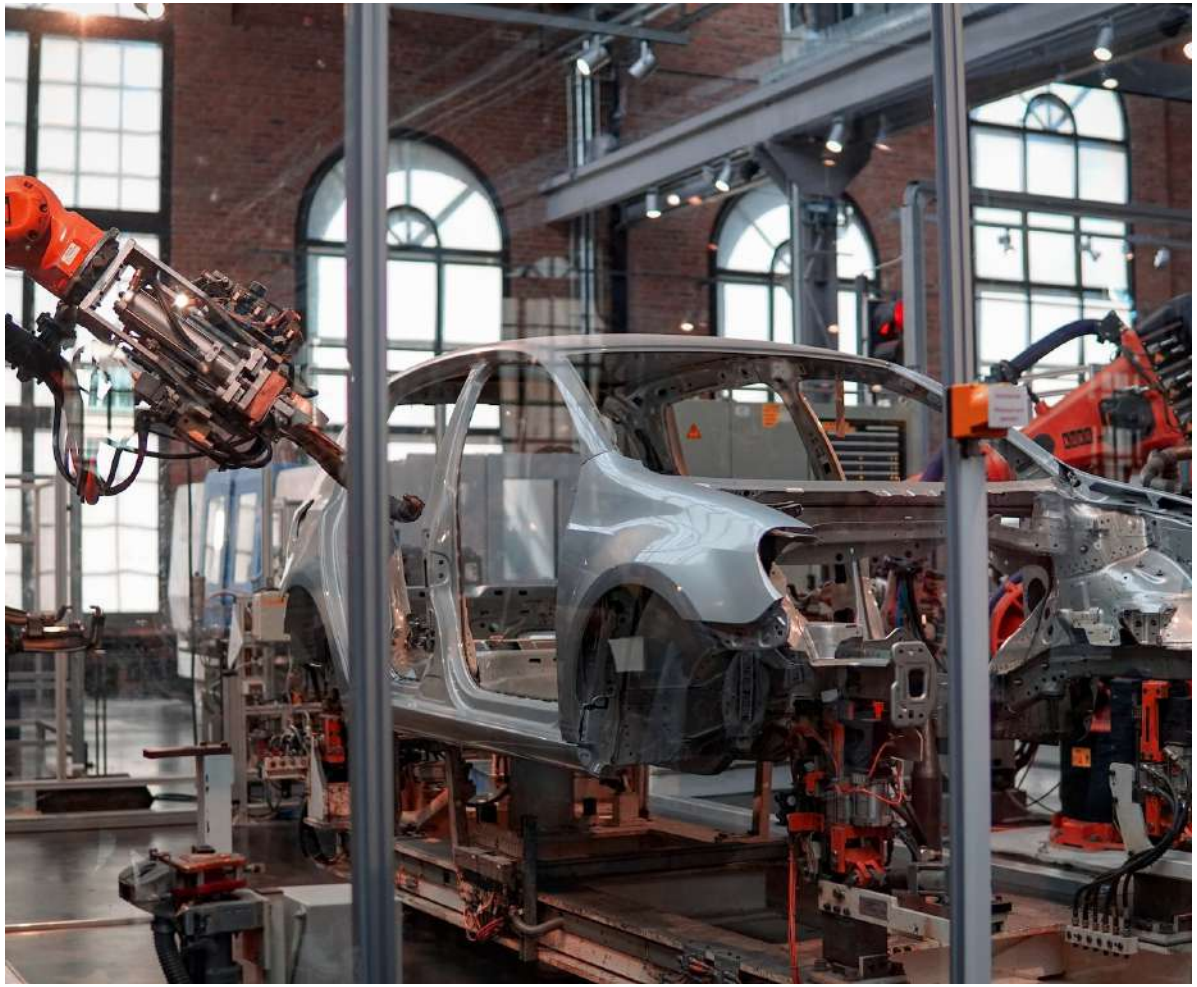


Foto door Lenny Kuhne op Unsplash

Toekomstige ontwikkelingen

- Optimaliseren PCAP-verwerking
- Opnieuw bekijken van de gebruikte software voor alle stappen
- Verbeteren parallel gebruik
- Meer testen
- Statistieken gebruiken voor het optimaliseren (*data-driven*)
- Automatische uitrol
- Beschikbaar maken voor andere TLD's

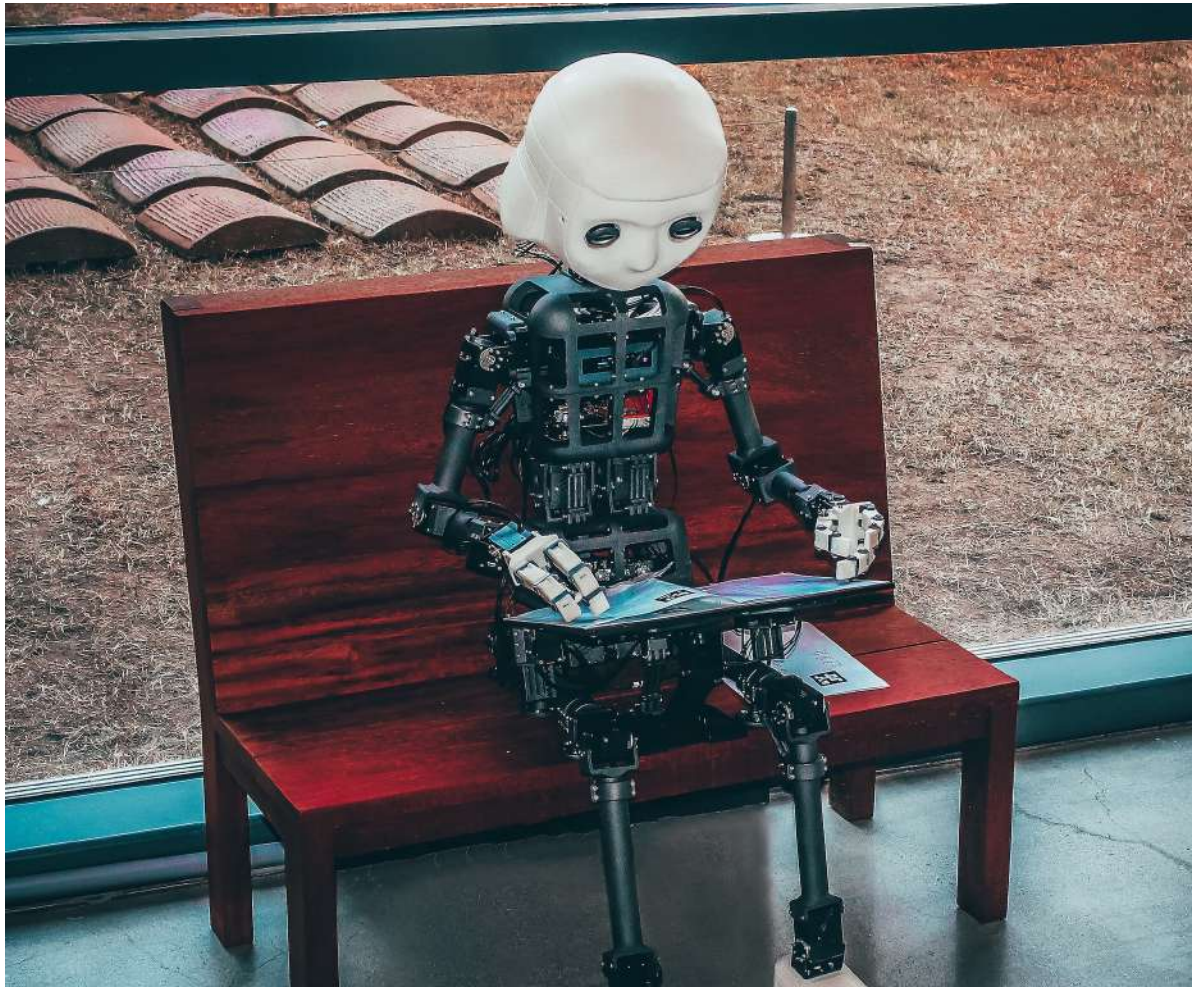
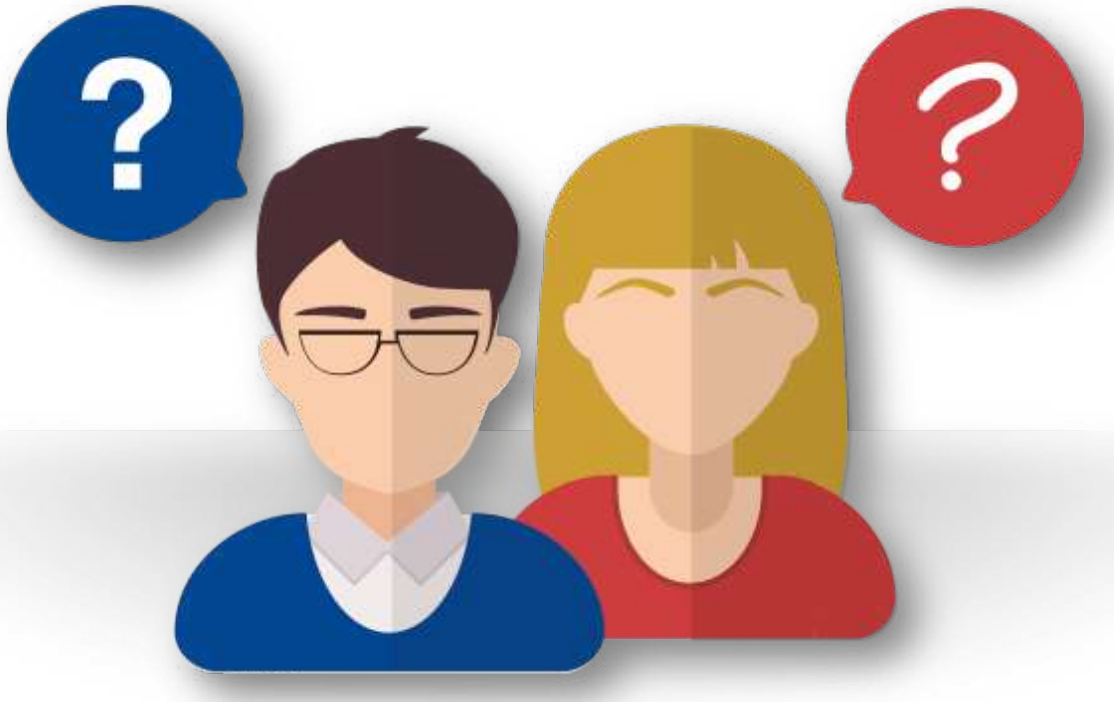


Foto door Andrea De Santis op Unsplash

Vragen, discussie



Bedankt!

SIDN LABS