Advanced Topics in Communication Networks

Internet Routing and Forwarding

Laurent Vanbever

nsg.ee.ethz.ch

10 Nov 2020

Lecture starts at **14:15**

Subsets of the materials inspired and/or coming from Pierre Francois, Olivier Bonaventure, Jennifer Rexford

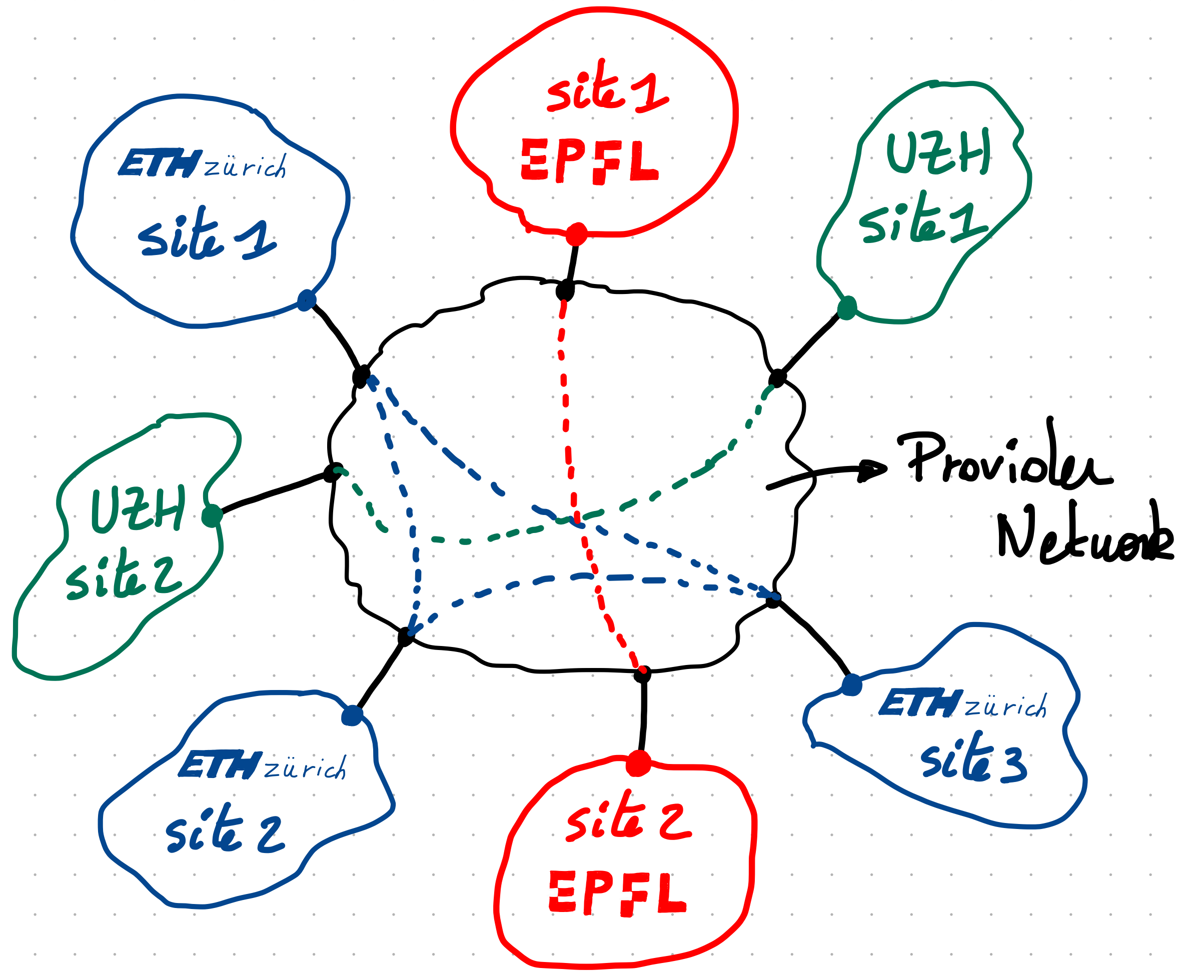
Last week on

Advanced Topics in Communication Networks

Virtual Private Networks

How do we interconnect private networks
across a shared infrastructure?

Part II: Provider-Managed VPNs



In a provider-managed VPN, the customer this time is agnostic to the service. For the customer, it is like its different sites are directly connected together through the provider.

There are 2 main problems to solve

1 Routing:

1.1. What kind of routes do we need where?

1.2. How do we distribute these routes?

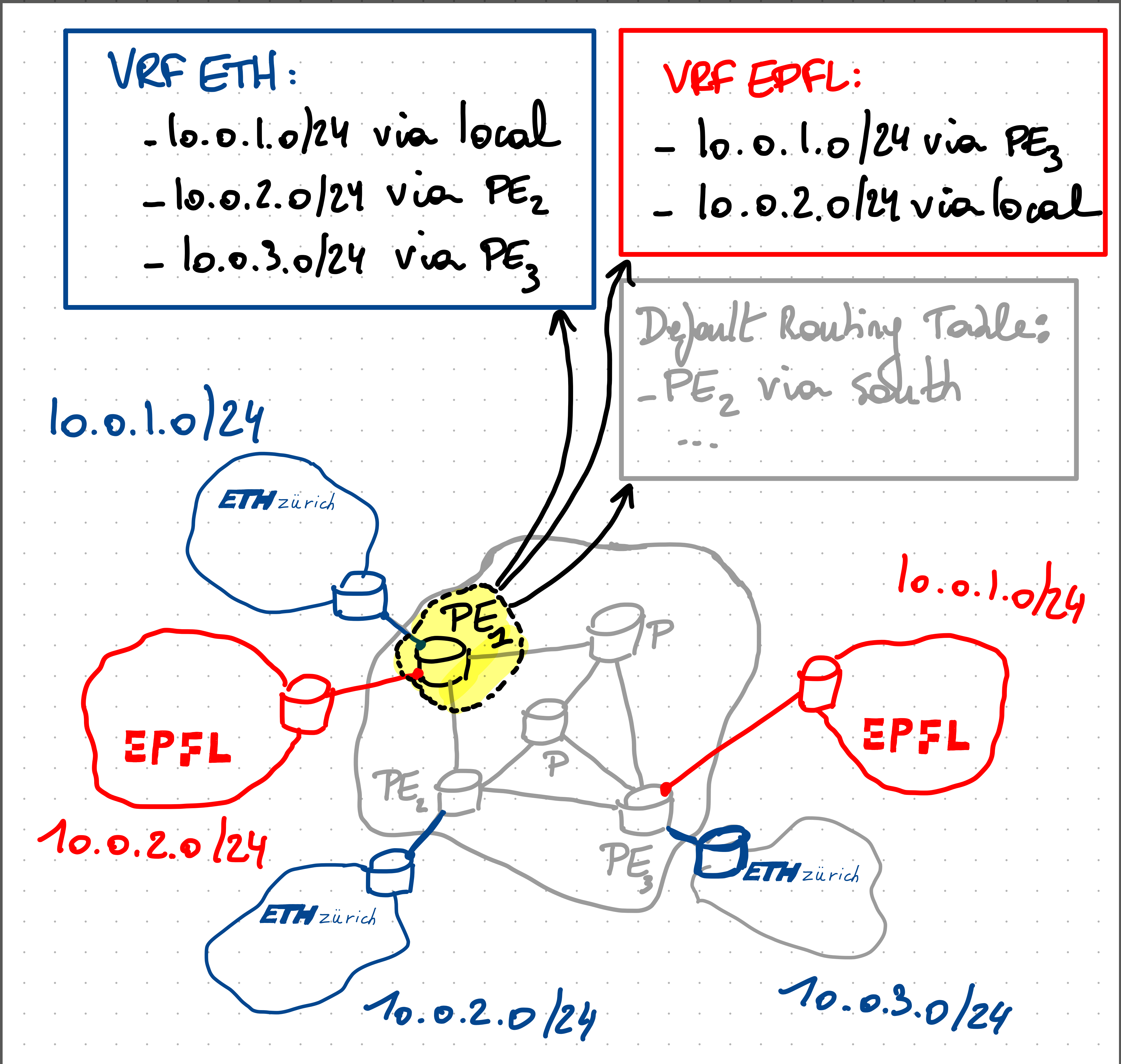
1.3. How do we deal with conflicting route information?

2 Forwarding:

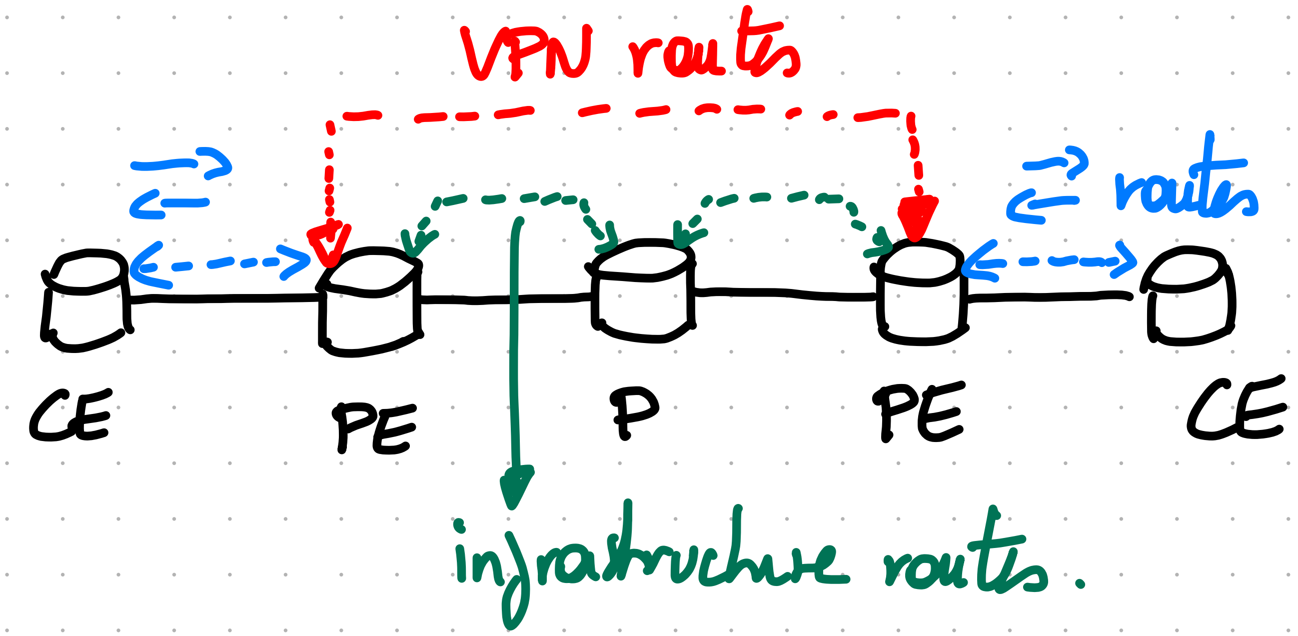
2.1. How do we forward traffic in such a network?

PE routers: In addition to the default routing table containing backbone routes, PE routers maintain one dedicated routing table for each VPN it is attached to.

This table is known as a VRF: VPN Routing and Forwarding table


VRF ETH:
- 10.0.1.0/24 via local
- 10.0.2.0/24 via PE2
- 10.0.3.0/24 via PE3
VRF EPFL:
- 10.0.1.0/24 via PE3
- 10.0.2.0/24 via local
Default Routing Table:
- PE2 via south
...
10.0.1.0/24
ETH zürich
PE1
P
EPFL
10.0.1.0/24
EPFL
P
PE2
10.0.2.0/24
PE3
ETH zürich
ETH zürich
10.0.2.0/24
10.0.3.0/24

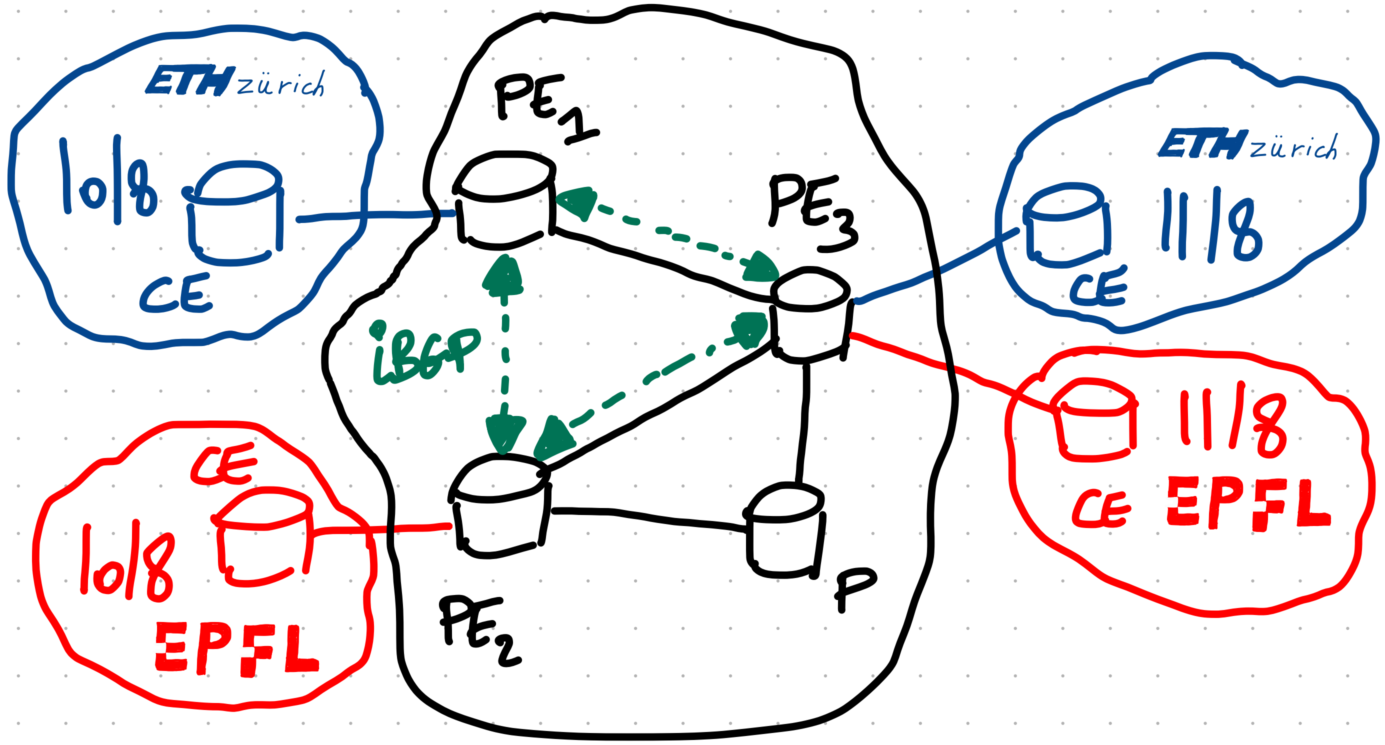
2)



Each PE must receive two types of routing info:

1. per-VPN routes reachable through local CE routers and remote PE router.

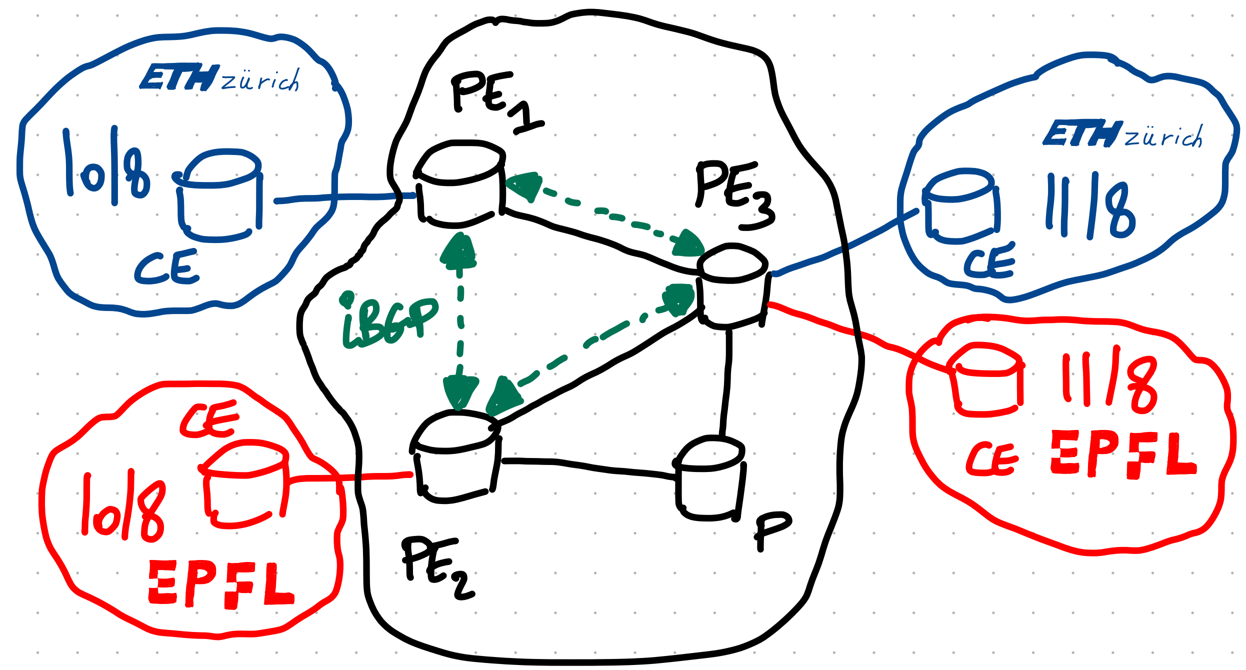
2. internal ISP routes to reach other PE router.



Problem 1: Since ≠ VPNs can use overlapping IP space, how do PEs distinguish between 10/8 advertised by ETH and 10/8 advertised by EPFL?

Problem 2: How do we ensure that the CEs of ETH and EPFL only learn the routes pertaining to their own VPN.

FORWARDING IN A BGP/MPLS VPN:



2.1. How do we forward traffic in such a network?

<u>Insight</u>:
- CE routers send pure IP pkt.
- PE routers encapsulate these IP packets with MPLS labels.

Each PE pushes two labels:
1) an outer label which identifies the next-hop PE;
2) an inner label which identifies the VRF to use in the remote PE.

This week on

Advanced Topics in Communication Networks

Fast Convergence

How do we *quickly* retrieve connectivity
upon *sudden* failures?

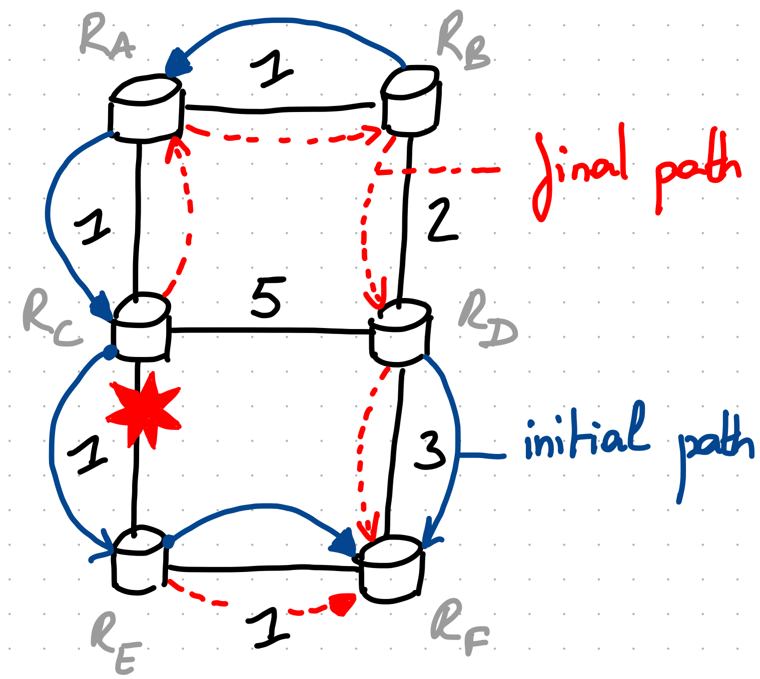
Let's switch to

06_fast_convergence_notes.pdf

Advanced Topics in Communication Networks

L8: Fast Convergence / 10.11.2020

Prof. Laurent VANBEVER - nsg.ee.ethz.ch

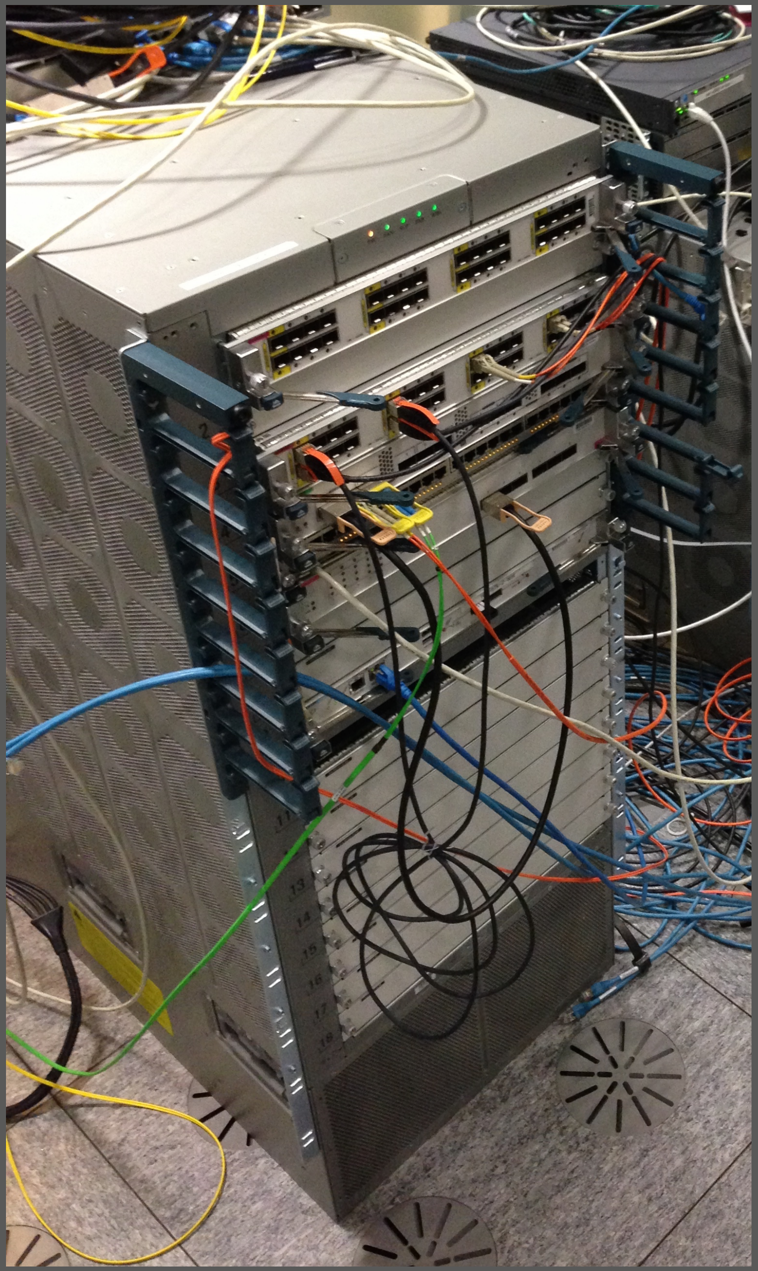


How do we retrieve connectivity as fast as possible after a link or a node failure?

↳ goal: <50 ms

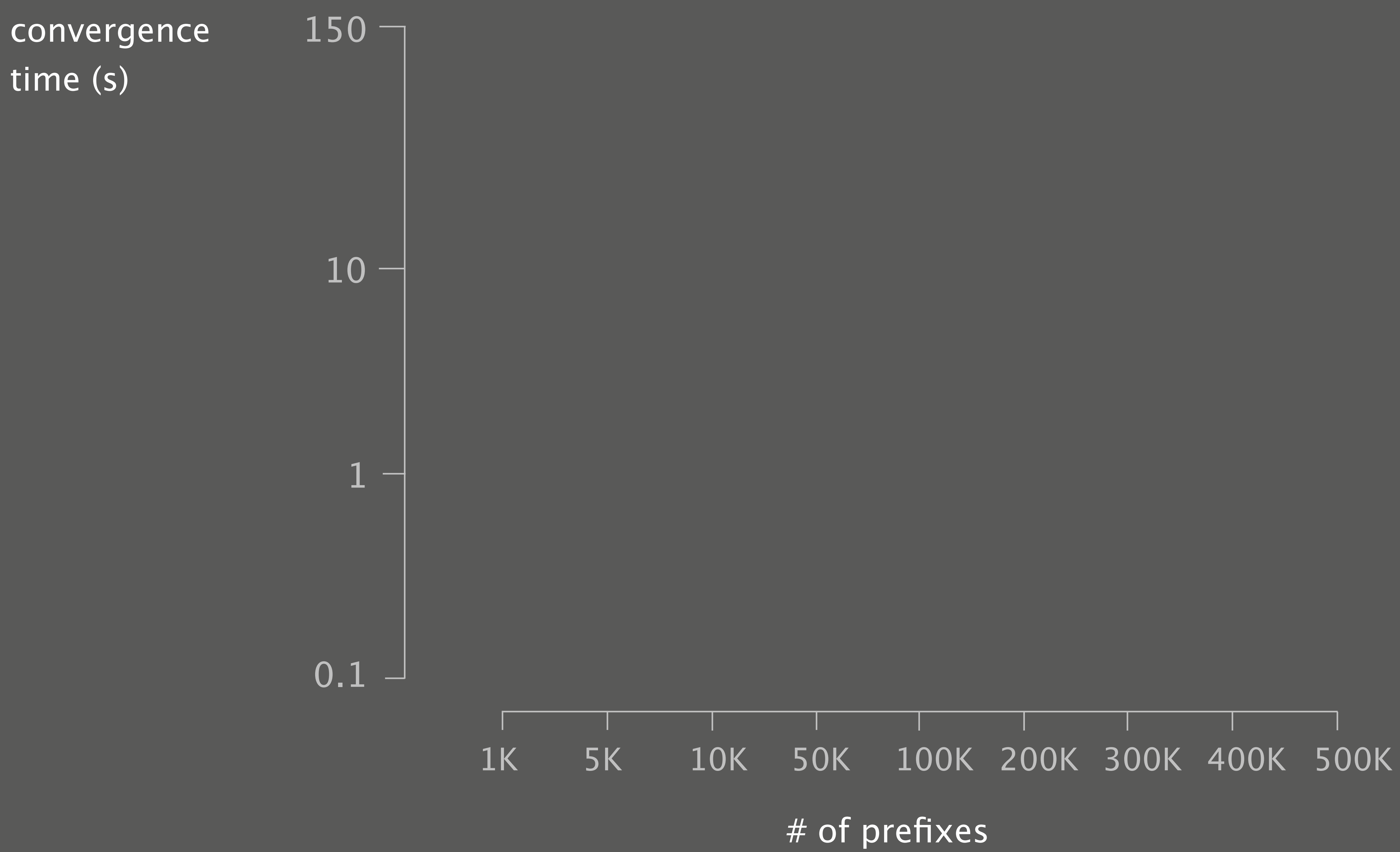
We measured how long it takes for

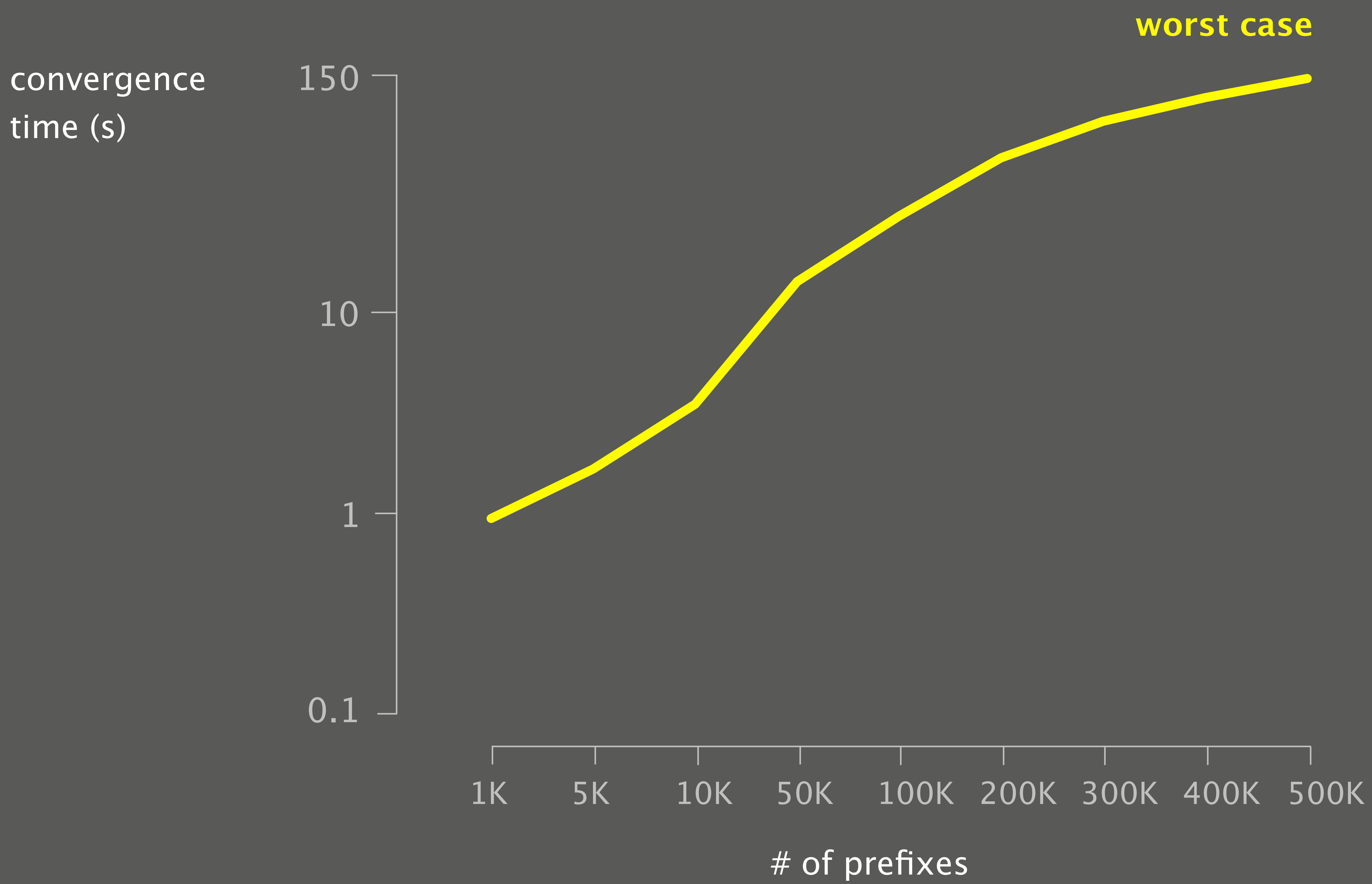
an ETH router to converge

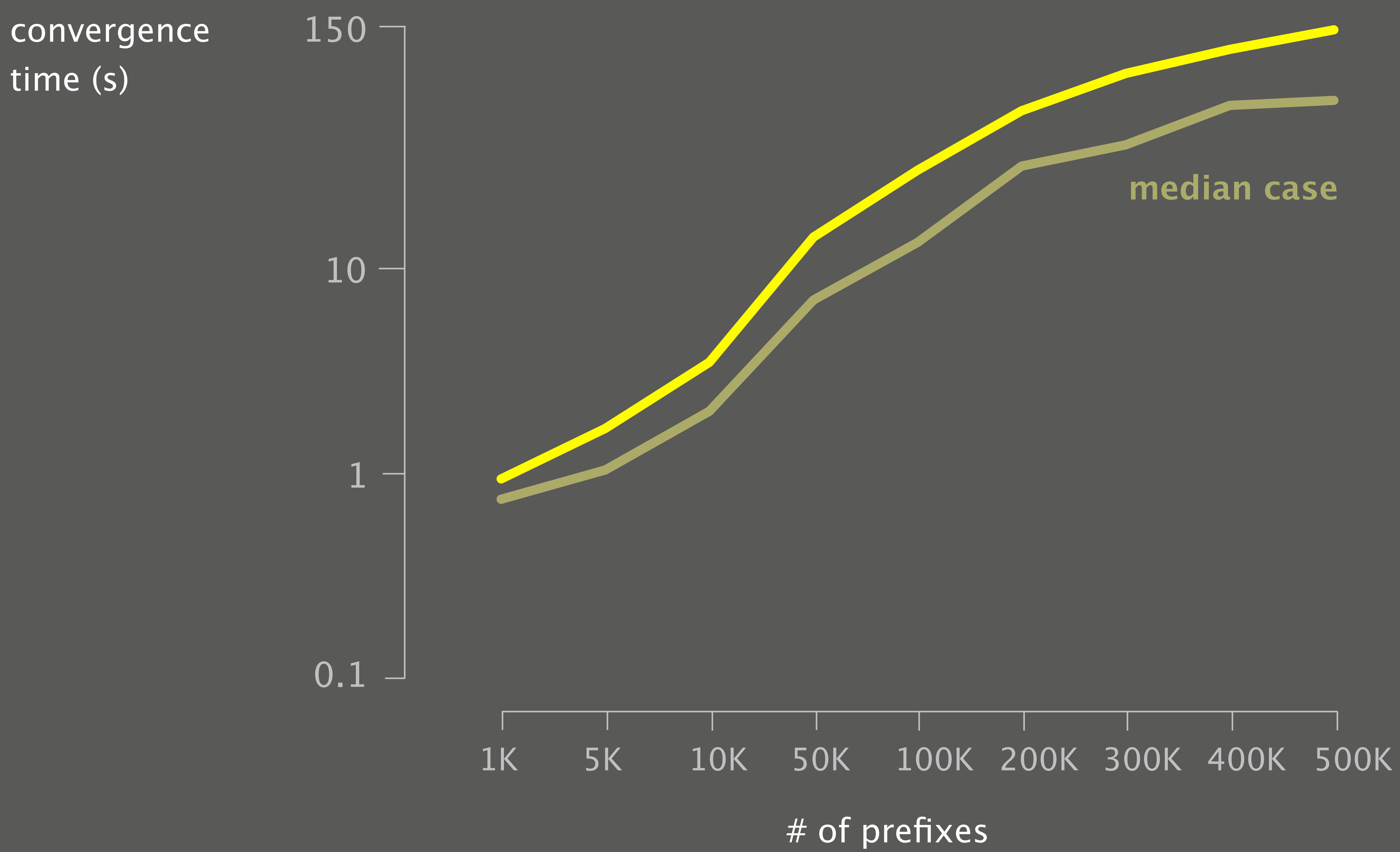


Cisco Nexus 9k

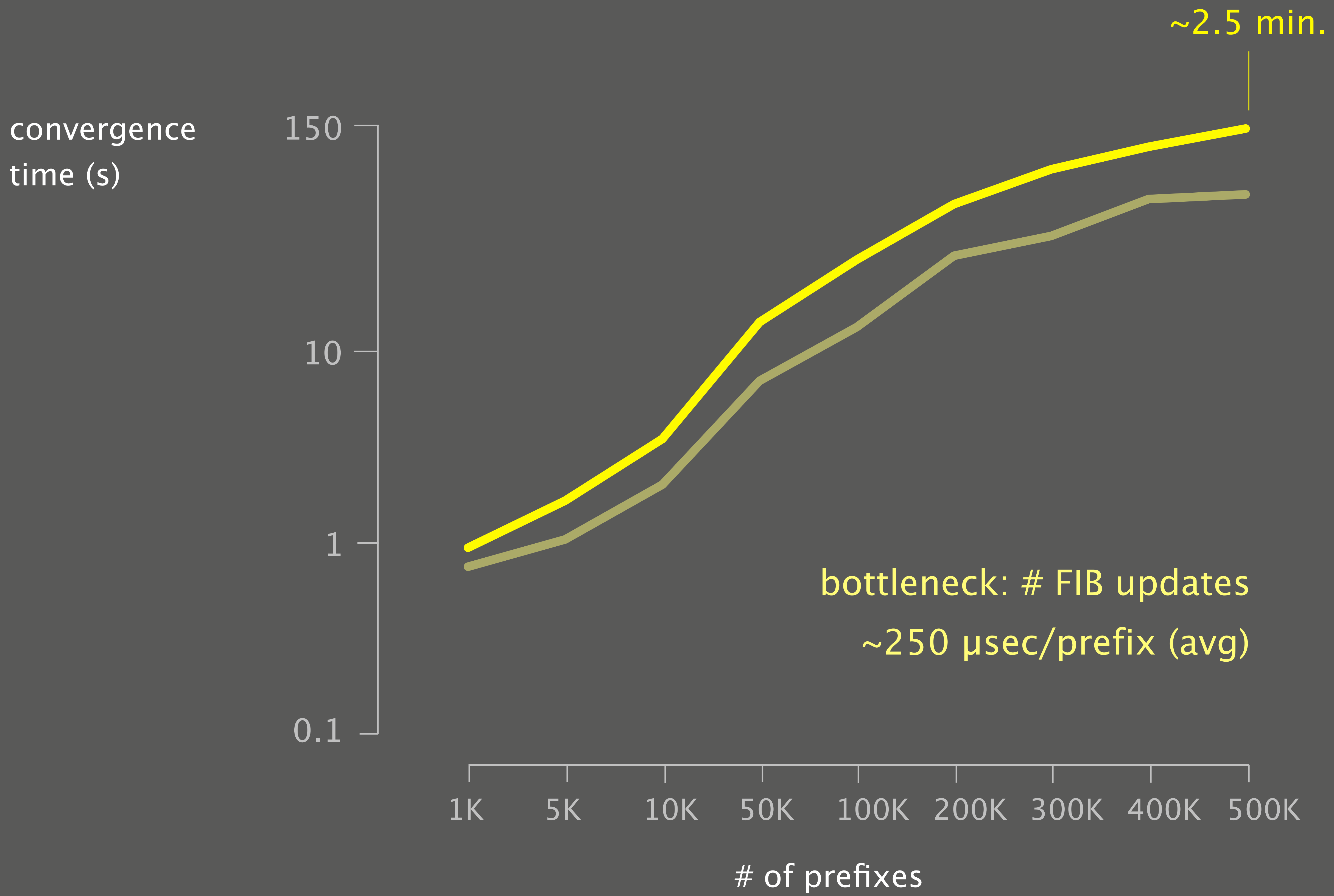
ETH's recent routers

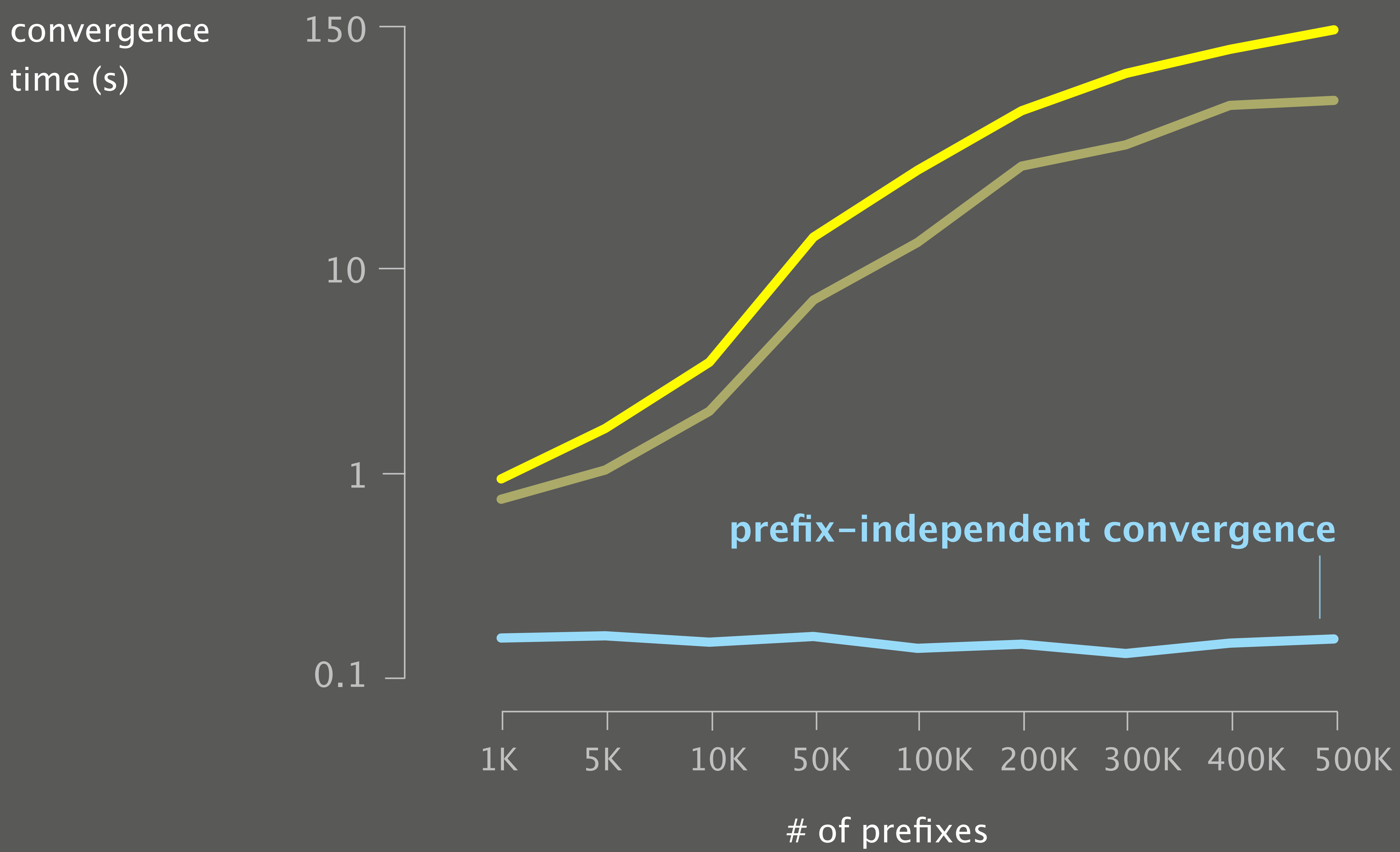

convergence
time (s)
150
10
1
0.1
1K
5K
10K
50K
100K
200K
300K
400K
500K
# of prefixes


worst case
convergence
time (s)
150
10
1
0.1
1K
5K
10K
50K
100K
200K
300K
400K
500K
# of prefixes


convergence
time (s)
150
10
1
0.1
median case
1K
5K
10K
50K
100K
200K
300K
400K
500K
# of prefixes

Traffic can be lost for several minutes!




convergence
time (s)
150
10
1
0.1
prefix-independent convergence
1K
5K
10K
50K
100K
200K
300K
400K
500K
# of prefixes

In practice,

LFA coverage is highly topology-dependent

Network	# links	LFA coverage (% links)
Abilene	28	42
GEANT	72	66
ISP1	114	54
ISP2	26	15
ISP3	265	65

In practice,

LFA coverage is highly topology-dependent

Network	# links	LFA coverage (% links)	
Abilene	28	42	*ring-based* network design
GEANT	72	66	
ISP1	114	54	
ISP2	26	15	
ISP3	265	65	*highly-meshed* network design

Advanced Topics in Communication Networks

Internet Routing and Forwarding

Laurent Vanbever

nsg.ee.ethz.ch

ETH Zürich (D-ITET)

10 Nov 2020