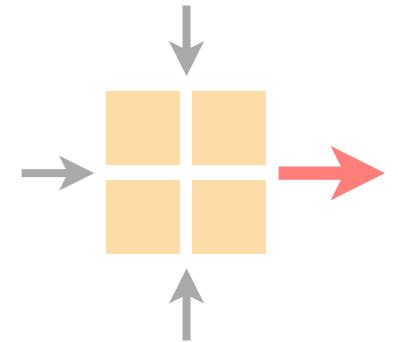


Advanced Topics in Communication Networks

Internet Routing and Forwarding



Laurent Vanbever
nsg.ee.ethz.ch

3 Nov 2020
Lecture starts at 14:15

Subsets of the materials inspired and/or coming from Olivier Bonaventure

Let me first quickly answer to your padlet questions

1. Does pure RSVP follow only IGP least cost paths to reserve resources?

1. Does pure RSVP follow only IGP least cost paths to reserve resources?

Yes.

That's why RSVP-TE has the ERO

1. Does pure RSVP follow only IGP least cost paths to reserve resources?

Yes. That's why RSVP-TE has the ERO

2. Is RSVP *really* used?

1. Does pure RSVP follow only IGP least cost paths to reserve resources?

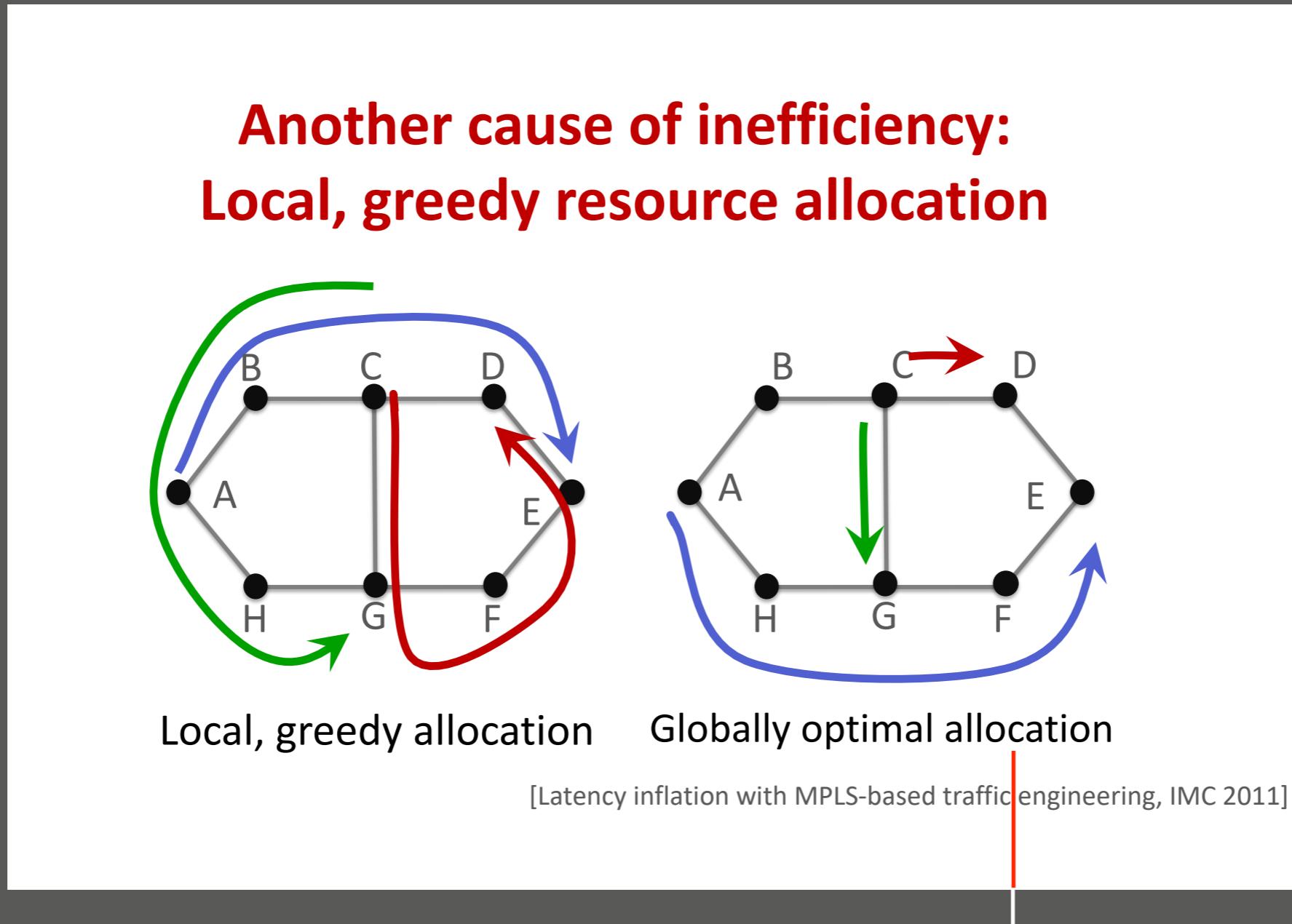
Yes. That's why RSVP-TE has the ERO

2. Is RSVP *really* used?

Yes! It's widely deployed, **but** there are a few gotchas...

Because paths are allocated **greedily**,
RSVP-TE can compute suboptimal solutions

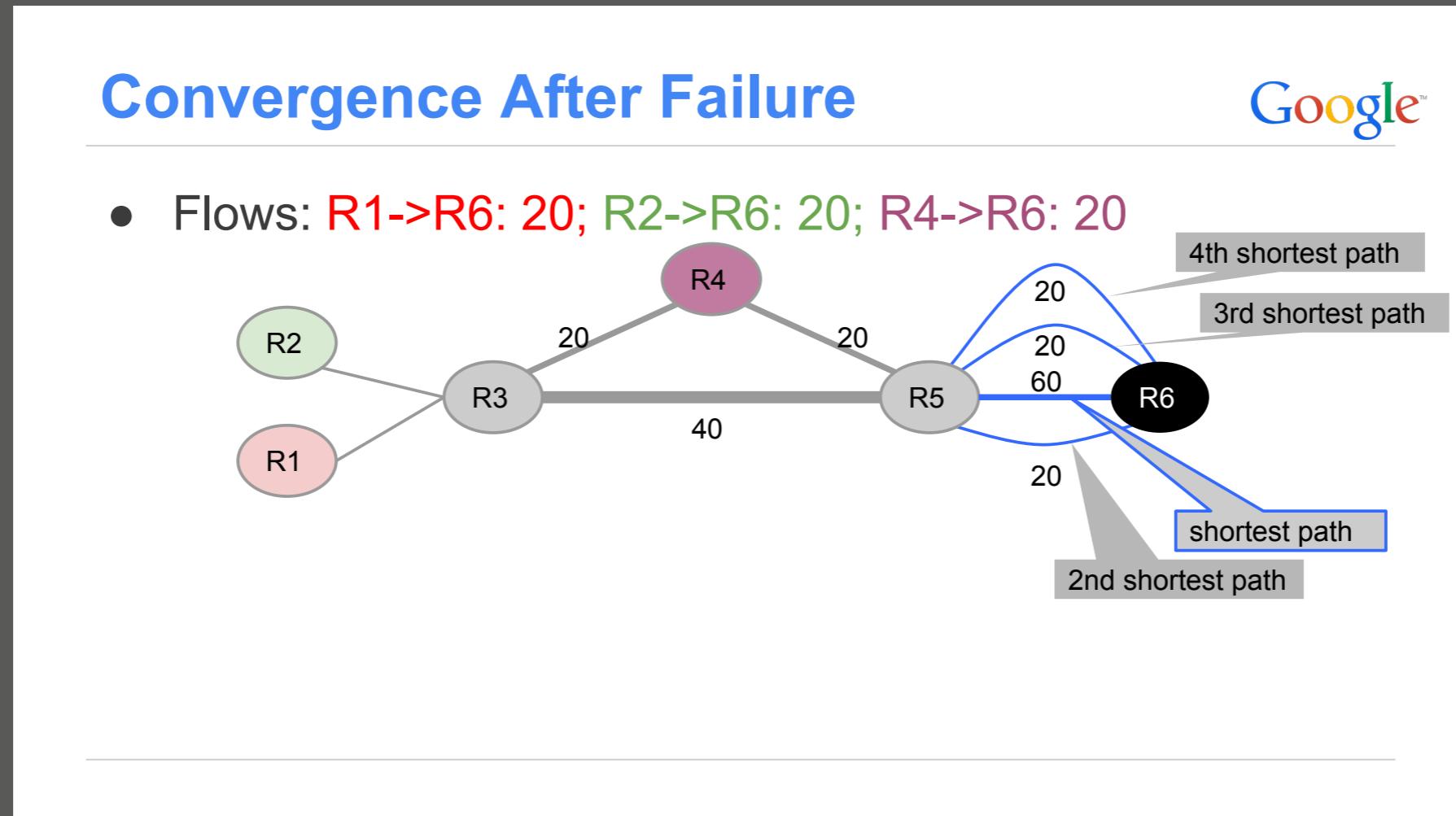
Because paths are allocated **greedily**,
RSVP-TE can compute suboptimal solutions



smaller delays and less links used!

Because paths are allocated autonomously,
RSVP-TE can take a while to converge

Because paths are allocated autonomously,
RSVP-TE can take a while to converge



Lessons Learned from B4, Google's SDN WAN [ATC 2015]

Because paths are allocated autonomously,
RSVP-TE can take a while to converge

Convergence After Failure

Google™

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

```
graph LR; R1((R1)) --- R3((R3)); R2((R2)) --- R3((R3)); R4((R4)) --- R3((R3)); R4((R4)) --- R5((R5)); R5((R5)) --- R6((R6)); R1((R1)) --- R5((R5)); R5((R5)) --- R6((R6)); R2((R2)) --- R5((R5)); R5((R5)) --- R6((R6)); R4((R4)) --- R5((R5)); R5((R5)) --- R6((R6));
```

The diagram illustrates a network topology with six routers (R1 to R6). Router R1 is at the bottom left, R2 is above it, R3 is between them, R4 is above R3, R5 is between R3 and R6, and R6 is at the bottom right. There are four main paths from R1 to R6: 1) A red path (R1-R3-R5-R6) with a capacity of 40. 2) A green path (R2-R3-R5-R6) with a capacity of 20. 3) A purple path (R4-R3-R5-R6) with a capacity of 20. 4) A blue path (R4-R5-R6) with a capacity of 20. A red arrow on the R5-R6 link is labeled '60', indicating that traffic is being rerouted through R5 instead of R4. The routers are colored: R1 is pink, R2 is light green, R3 is grey, R4 is purple, R5 is grey, and R6 is black.

Lessons Learned from B4, Google's SDN WAN [ATC 2015]

Because paths are allocated autonomously,
RSVP-TE can take a while to converge

Convergence After Failure

Google™

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

The diagram illustrates a network of six routers (R1 through R6) connected by various links. Router R6 is the destination for three flows. A red starburst indicates a link failure between R5 and R6. Router R5 has three outgoing links to R6, each with a capacity of 20. Router R5 also has a link to R4 with a capacity of 20. Router R4 has a link to R5 with a capacity of 20. Router R3 has two links to R4, both with a capacity of 20. Router R3 also has a link to R5 with a capacity of 40. Router R2 has a link to R3 with a capacity of 20. Router R1 has a link to R3 with a capacity of 20. Red dashed arrows point from R1, R2, and R4 towards R3, indicating they are finding a new path due to the failed R5-R6 link.

- R5-R6 link fails
 - R1, R2, R4 *autonomously* find next best path

Lessons Learned from B4, Google's SDN WAN [ATC 2015]

Because paths are allocated autonomously,
RSVP-TE can take a while to converge

Convergence After Failure

Google™

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

The diagram illustrates a network of six routers (R1 through R6) connected by various links. Router R1 is pink, R2 is light green, R3 is grey, R4 is purple, R5 is grey, and R6 is black. A red link connects R1 and R6. A green link connects R2 and R6. A grey link connects R3 and R4, labeled '20'. Another grey link connects R3 and R5, labeled '40'. A purple link connects R4 and R5, labeled '20'. A blue link connects R5 and R6, labeled '20'. A red starburst indicates a link failure between R5 and R6. A green arrow points from R5 to R6, labeled '20', indicating a new path being established. A red arrow points from R1 to R6, also labeled '20', indicating another path being considered. A grey box labeled 'No Traffic Engineering' is present.

- R5-R6 link fails
 - R1, R2, R4 *autonomously* try for next best path
 - R1, R2, R4 push **20** altogether

Lessons Learned from B4, Google's SDN WAN [ATC 2015]

Because paths are allocated autonomously,
RSVP-TE can take a while to converge

Convergence After Failure

Google™

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

The diagram illustrates a network of six routers (R1 through R6) connected by various links. Router R1 is pink, R2 is light green, R3 is grey, R4 is purple, R5 is grey, and R6 is black. A red link connects R1 and R6. A green link connects R2 and R6. A grey link connects R3 and R4, with a value of 20 above it. Another grey link connects R3 and R5, also with a value of 20 above it. A dark grey link connects R4 and R5, with a value of 20 above it. A thick grey link connects R5 and R6, with a value of 40 above it. A red asterisk marks a failed link between R5 and R6. A blue link connects R4 and R6, with a value of 20 above it. A green link connects R2 and R5, with a value of 20 above it. A blue link connects R5 and R6, with a value of 60 above it. A red link connects R1 and R5, with a value of 20 above it. A green link connects R1 and R6, with a value of 20 above it.

- R5-R6 link fails
 - R1, R2, R4 *autonomously* try for next best path
 - R1 wins, R2, R4 retry for next best path

Distributed Traffic Engineering Protocols

Lessons Learned from B4, Google's SDN WAN [ATC 2015]

Because paths are allocated autonomously,
RSVP-TE can take a while to converge

Convergence After Failure

Google™

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

The diagram illustrates a network of six routers (R1 to R6) connected by various links. Router R1 is pink, R2 is light green, R3 is grey, R4 is purple, R5 is grey, and R6 is black. The network has three main paths from R1 to R6: 1) A red path via R3 and R5 with a total capacity of 40. 2) A green path via R2 and R4 with a total capacity of 40. 3) A blue path via R3, R4, and R5 with a total capacity of 60. A red starburst indicates a link failure between R5 and R6. The network is labeled "Distributed Traffic Engineering Protocols".

- R5-R6 link fails
 - R1, R2, R4 *autonomously* try for next best path
 - R1 wins, R2, R4 retry for next best path
 - R2 wins this round, R4 retries again

Lessons Learned from B4, Google's SDN WAN [ATC 2015]

Because paths are allocated autonomously,
RSVP-TE can take a while to converge

Convergence After Failure

Google™

- Flows: R1->R6: 20; R2->R6: 20; R4->R6: 20

The diagram illustrates a network graph with six routers (R1, R2, R3, R4, R5, R6) and their connections. Router R1 is pink, R2 is light green, R3 is grey, R4 is purple, R5 is grey, and R6 is black. The network has the following links and capacities:

- R1 to R3: 40
- R2 to R3: 20
- R3 to R4: 20
- R3 to R5: 20
- R4 to R5: 20
- R5 to R6: 60
- R6 is a leaf node with no outgoing links.

A red starburst indicates a link failure between R5 and R6. Multiple colored arrows show different paths being considered by routers R1, R2, and R4 to reach R6. A legend at the bottom right identifies the colors: red for R1, green for R2, purple for R4, blue for R5, and grey for R3 and R6.

- R5-R6 link fails
 - R1, R2, R4 *autonomously* try for next best path
 - R1 wins, R2, R4 retry for next best path
 - R2 wins this round, R4 retries again
 - R4 finally gets third best path!

Distributed Traffic Engineering Protocols

Lessons Learned from B4, Google's SDN WAN [ATC 2015]

That's why major Wide-Area Networks (WAN) typically rely on centralized versions of RSVP-TE

SWAN: Software-driven wide area network

Ratul Mahajan

Microsoft Research

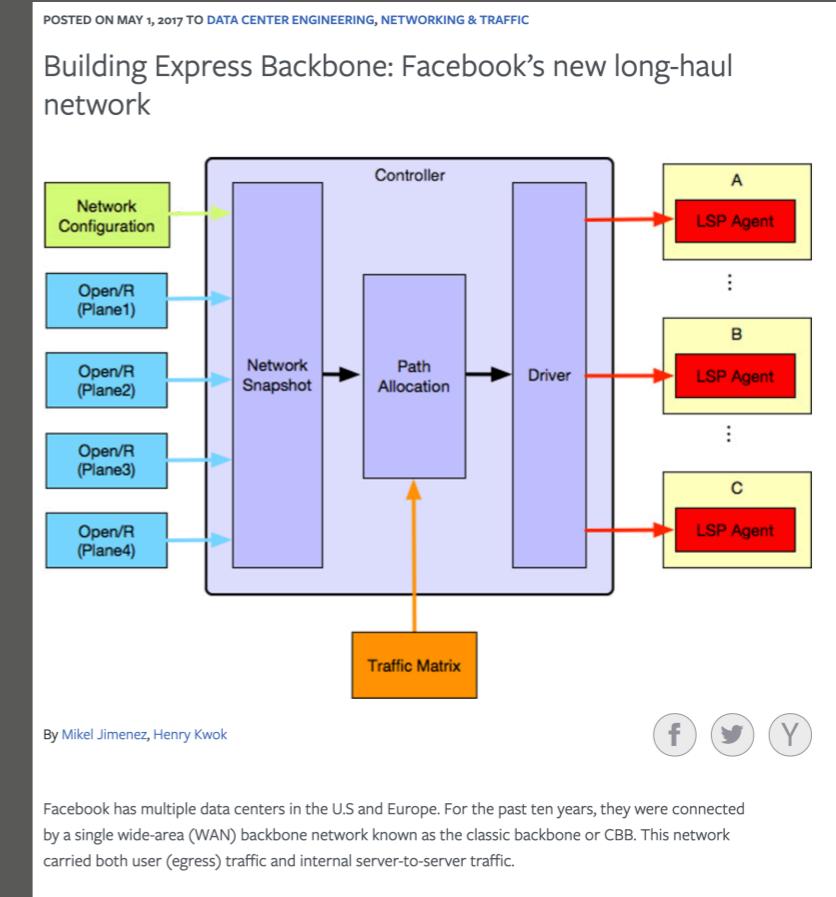


Google™

Lessons Learned from B4, Google's SDN WAN

Subhasree Mandal
July 9, 2015

Like what *you* did in the exercises!



Last week on
Advanced Topics in Communication Networks

Quality of Service

How do we manage congestion?
(locally, on a per link basis)

VPN

How do we interconnect private networks?
(across a shared infrastructure)

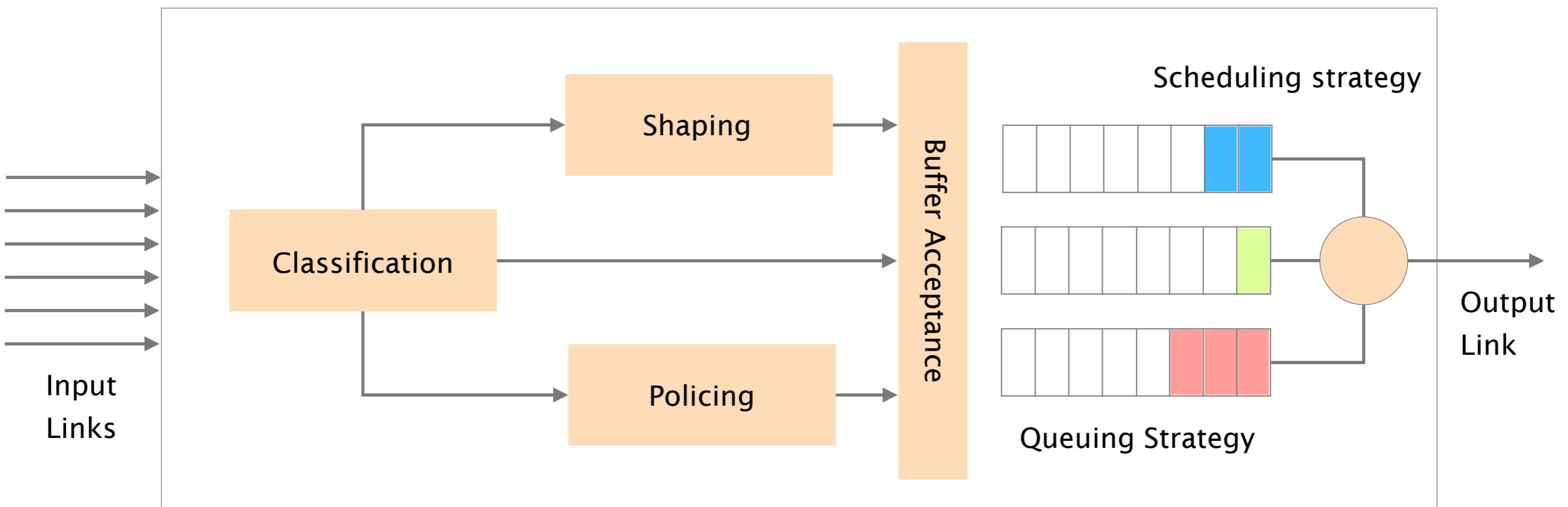
Quality of Service

VPN

How do we manage congestion?
(locally, on a per link basis)

We saw different tools to
locally manage congestion

QoS-enabled router



A quick reminder on Max. Min. FAIRNESS:

Given a set of bandwidth demands r_i and a total bandwidth C , the max-min bandwidth allocations a_i are:

$$a_i = \min(f, r_i)$$

where f is the unique value s.t. $\sum a_i = C$

Important property:

If you (a flow) don't get your full demand, no one gets more than you.

Buffer Acceptance Algorithms

Once the router has decided which queue it needs to append the packet to it.

At this time, it should decide whether to accept this packet or drop it or mark it.

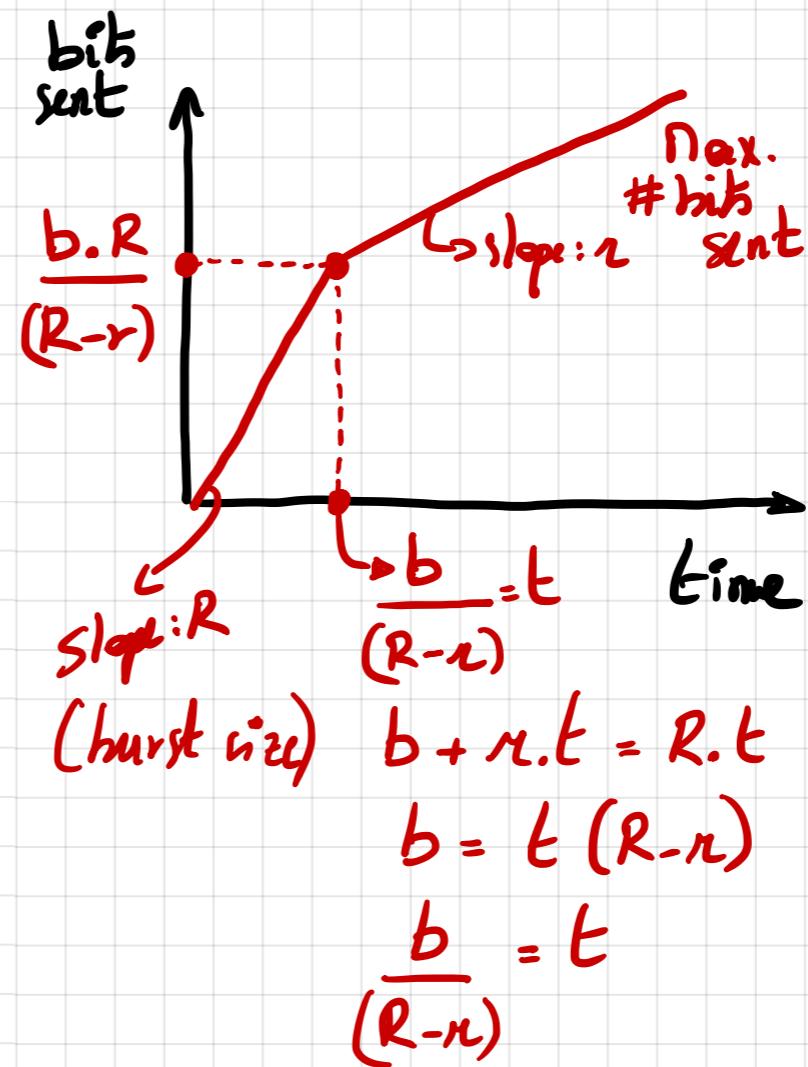
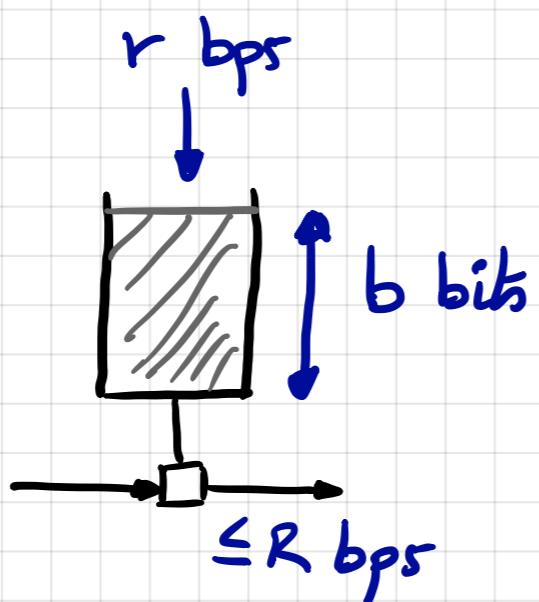
Two common examples:

1. Tail Drop: simply drops any newly arriving packets when the queue is filled to its maximum capacity.

Pros: dead simple to implement.

Cons: tends to synchronize the TCP flow when multiple packets are lost.

Token Bucket: Characterizing Burstiness



Parameters:

- $r \text{ bps}$: bucket filling rate
- $b \text{ bits}$: bucket size / max. burst size
- $R \text{ bps}$: maximum link capacity / peak rate

Enforcing traffic specification (using a TB) : (3 solutions)

1. Policing : **DROP** all data
in excess of the specification
2. Shaping : **DELAY** all data
in excess of the specification
UNTIL it obeys it.
3. Marking : **MARK** all data in
excess of the specification
and give these packets
lower priority in the network.

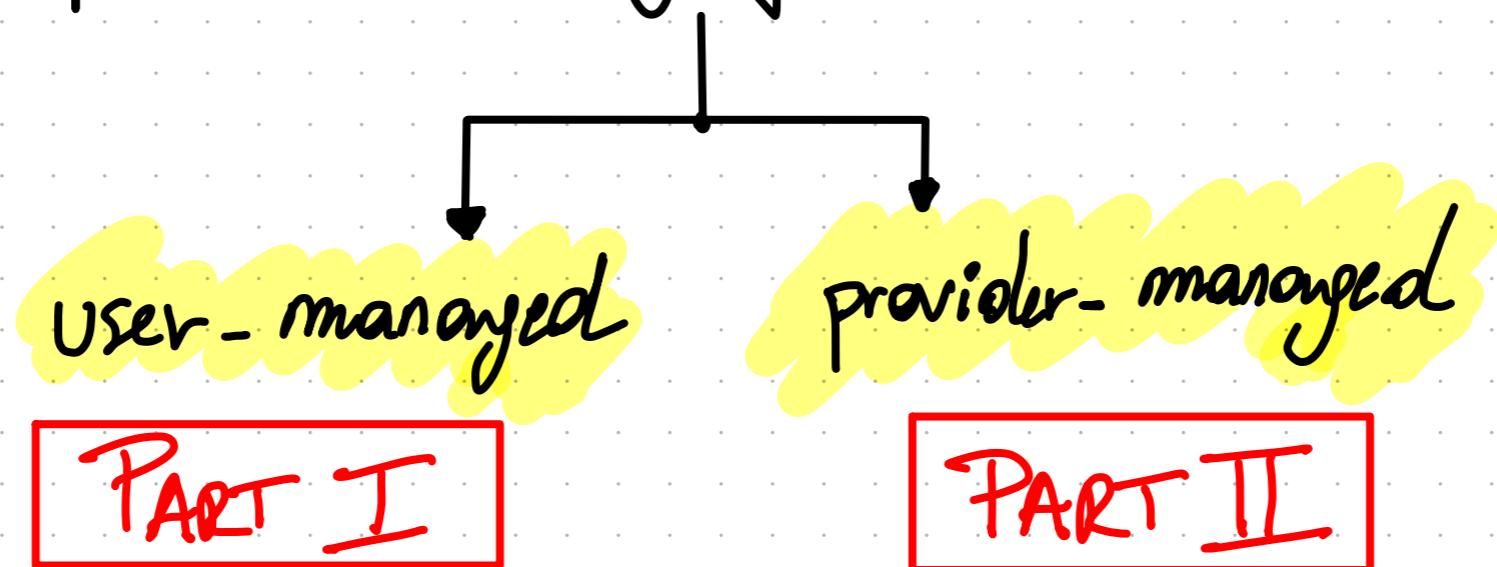
Quality of Service

VPN

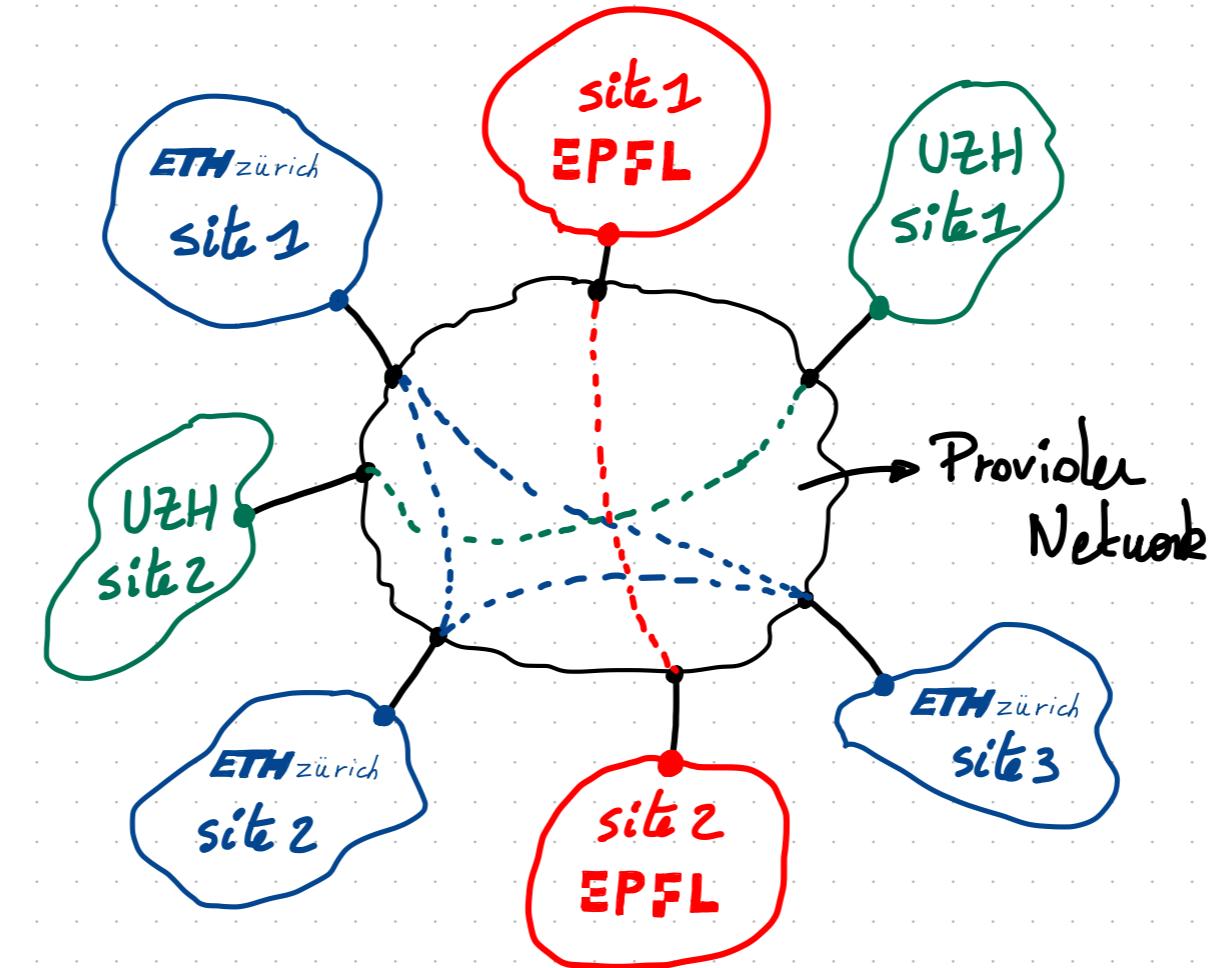
**How do we interconnect private networks?
(across a shared infrastructure)**

How do we interconnect geographically-distributed sites with the "same" privacy and guarantees as a private network?

Today's VPN solutions differ according to whom is responsible for setting them up and managing them

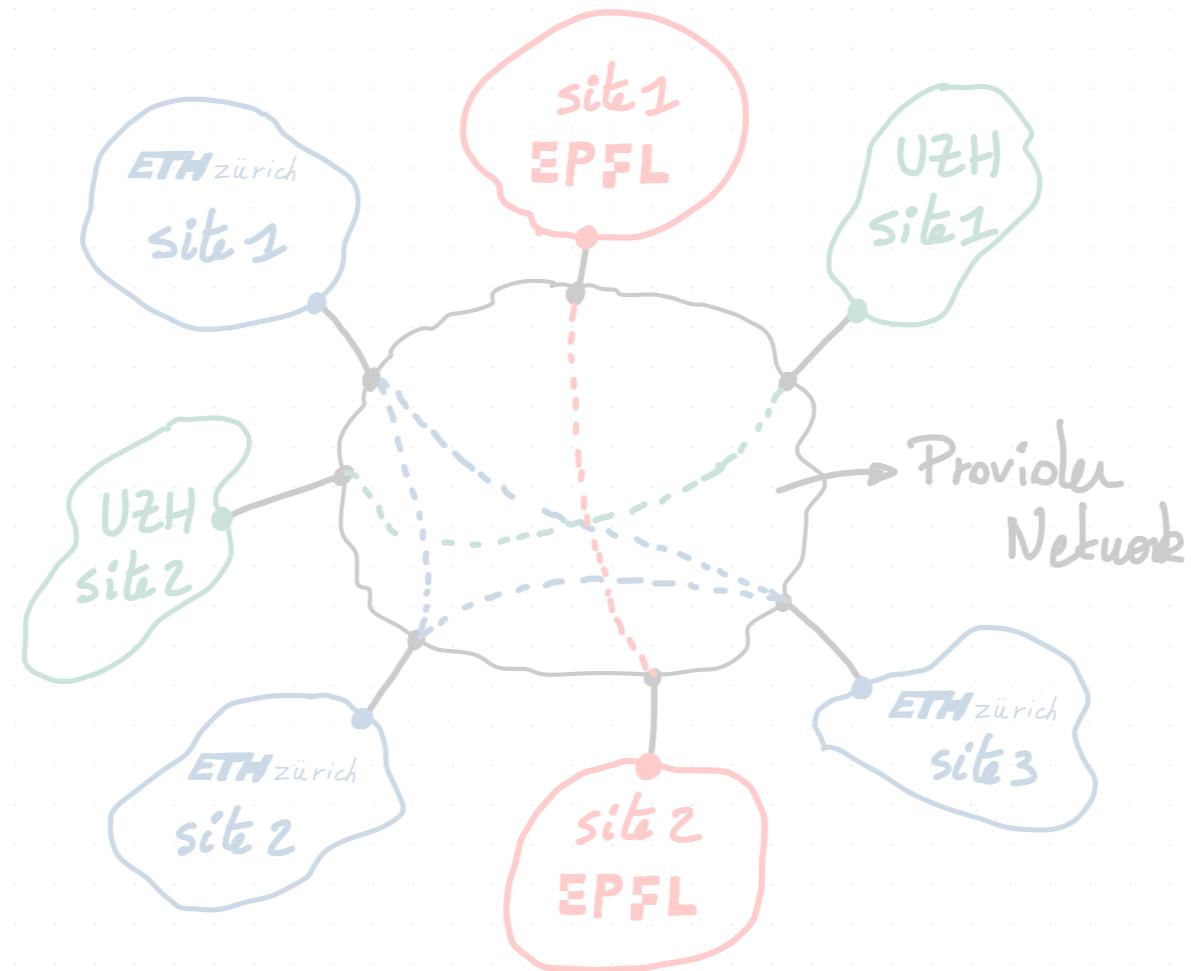


PART I: CUSTOMER-MANAGED VPNS



In a customer managed VPN, the provider is completely agnostic to the VPN service. It simply provides **dedicated physical connections** (leased lines) or IP connectivity.

PART I: CUSTOMER-MANAGED VPNS



In a customer managed VPN, the provider is completely agnostic to the VPN service. It simply provides **dedicated physical connections** (leased lines) or IP connectivity.

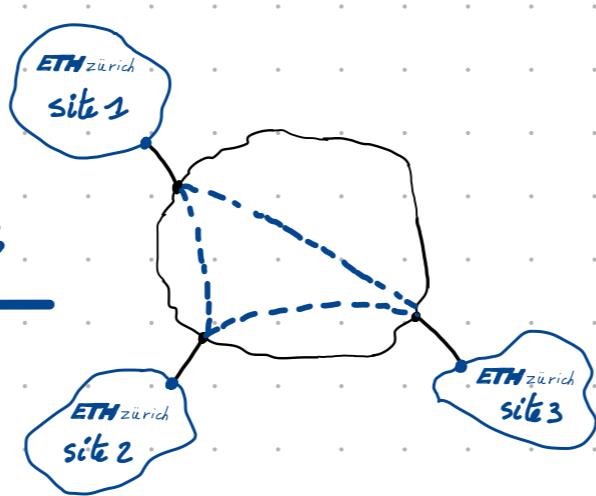
solution 1

Solution 1: LEASED LINES

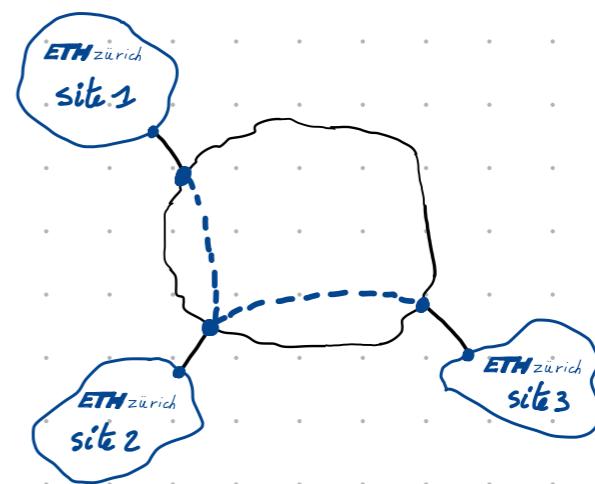
Definition: A Leased Line (LL) is a dedicated (private) connection between two geographically distant sites according to a commercial contract.

Since the connection is dedicated, the provider can provide QoS guarantees such as guaranteed bandwidth.

Typical
organizations:

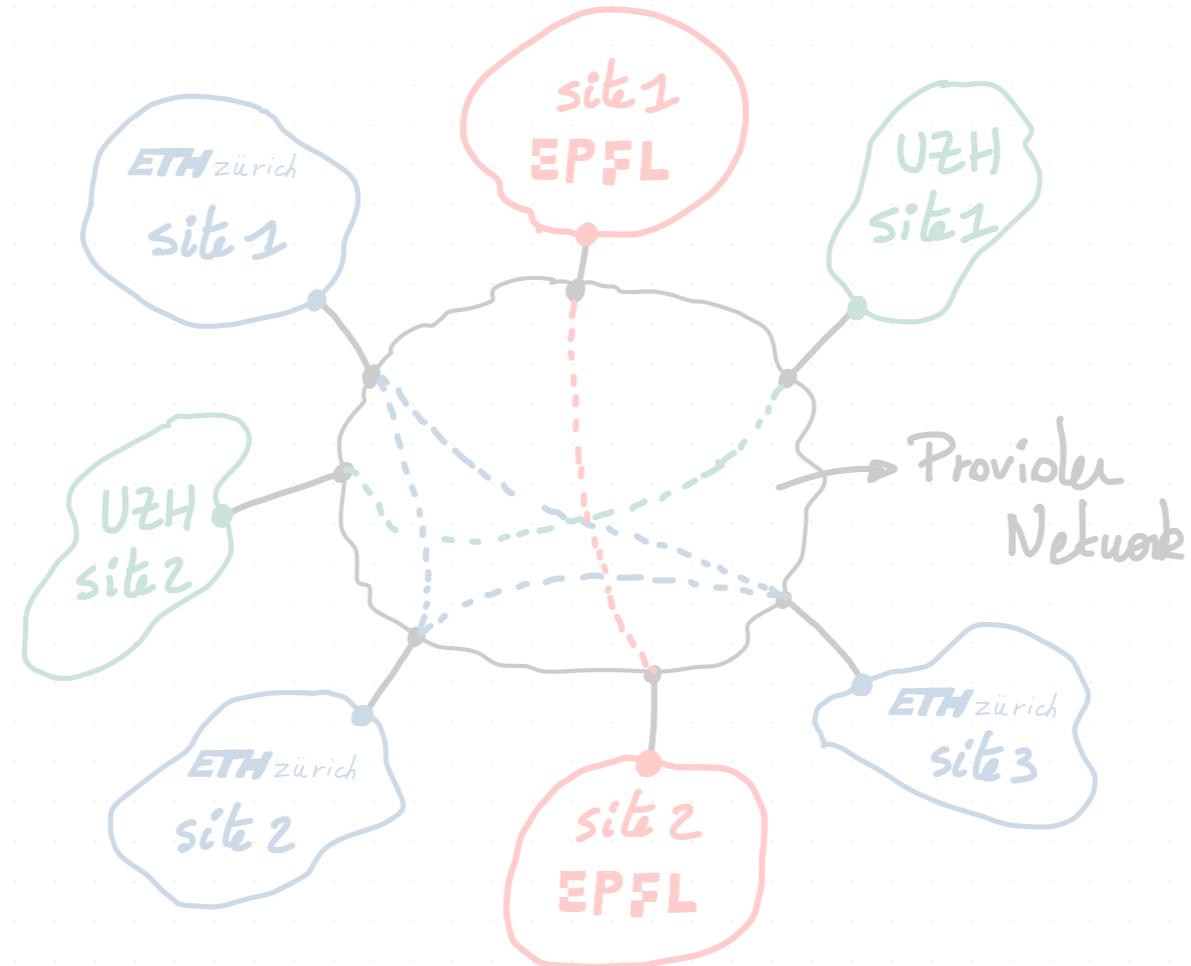


FULL-NMESH



HUB-AND-SPOKE

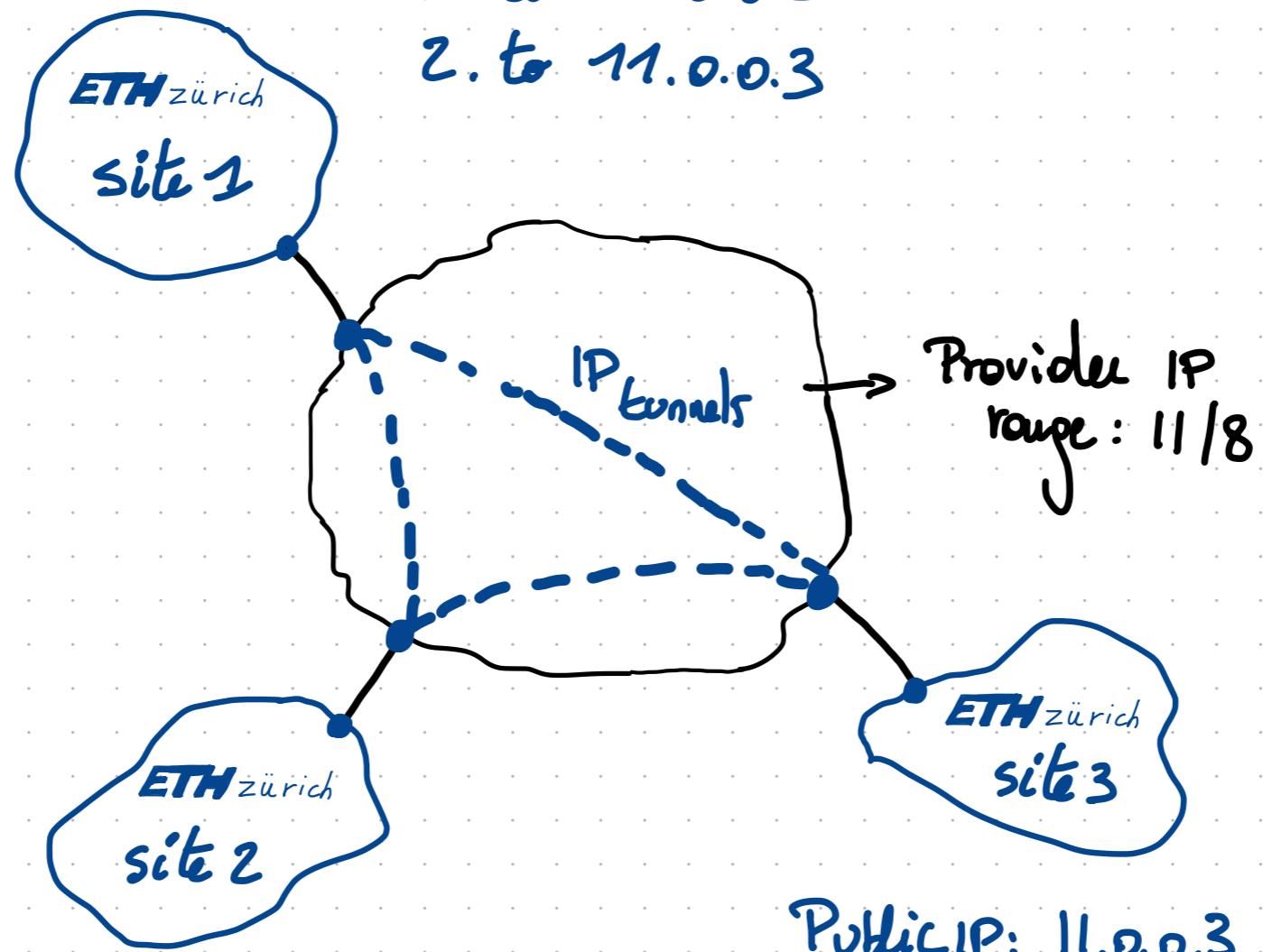
PART I: CUSTOMER-MANAGED VPNS



In a customer managed VPN, the provider is completely agnostic to the VPN service. It simply provides **dedicated physical connections** (leased lines) or IP connectivity.

solution 2

Public IP: 11.0.0.1
Private IP range: 10.1.0.0/24
IP tunnels:



Public IP: 11.0.0.1
Private IP range: 10.1.0.0/24
IP tunnels:

1. to 11.0.0.2
2. to 11.0.0.3

Provider IP range: 11/8

Public IP: 11.0.0.2
Private IP range: 10.2.0.0/24
IP tunnels:

1. to 11.0.0.1
2. to 11.0.0.3

Public IP: 11.0.0.3
Private IP range: 10.3.0.0/24
IP tunnels:

1. to 11.0.0.1
2. to 11.0.0.2

This week on
Advanced Topics in Communication Networks

Provider-managed VPNs

How do we interconnect private networks?
with the help of a provider network

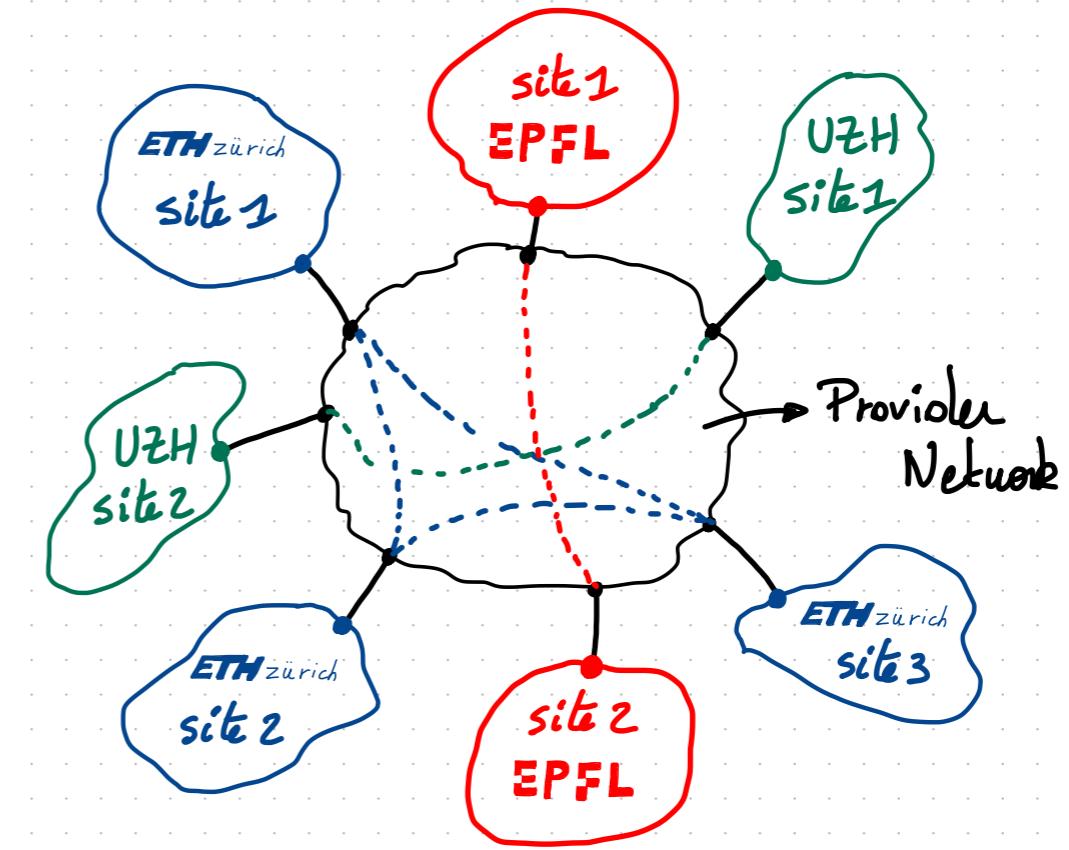
Let's switch to
05_vpn_notes.pdf

I'm resuming from

PART II: PROVIDER-MANAGED VPNS

(page 10/23)

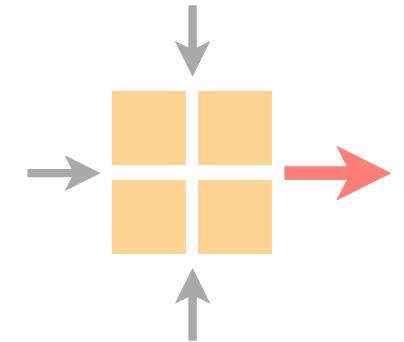
PART II: PROVIDER-MANAGED VPNs



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.

Advanced Topics in Communication Networks

Internet Routing and Forwarding



Laurent Vanbever
nsg.ee.ethz.ch

ETH Zürich (D-ITET)
27 Oct 2020