An Open Platform to Teach How the Internet Practically Works

Thomas Holterbach
NANOG 78, San Francisco

Joint work with Tobias Bühler,
Tino Rellstab, and
Laurent Vanbever

ETH Zürich
How do we traditionally teach how the Internet works?
How do we traditionally teach how the Internet works?
How do we traditionally teach how the Internet works?

Which messages are exchanged?
How do we traditionally teach how the Internet works?

**theory**

**exercises**

**labs**

Which messages are exchanged?

- AS 1
- AS 0
- AS 2
- AS 4
- 82.130.64.0/21
These concepts are not sufficient to understand how the Internet *practically* works
In practice, there are **peering agreements** with **stringent SLAs**

Network operators talking during NANOG'76
In practice, there are thousands of ASes and connectivity must be monitored network-wide.
In practice, debugging can be tricky

Anybody else is experiencing packet loss since last Tuesday across the AT&T network in the L.A. area?

I'm seeing it coming from both Zayo and HE

8. ae2.cs1.lga5.us.zip.zayo.com
9. ae18.ter1.lga5.us.zip.zayo.com
10. 192.205.36.105
11. cr1.n54ny.ip.att.net
12. cgci122crs.ip.att.net
13. cgci121crs.ip.att.net
14. dvmco22crs.ip.att.net
15. slkut21crs.ip.att.net
16. la2ca21crs.ip.att.net
17. gar20.la2ca.ip.att.net
At ETH Zurich, we let the students operate their own mini-Internet,
At ETH Zurich, we let the students operate their own mini-Internet, altogether, like if they were the network operators
At ETH Zurich, we let the students operate their own mini-Internet, altogether, like if they were the network operators.
The feedback we receive from our students is very positive.

"It really allows us to apply the theoretical concepts"

"I am quite confident about many things on the Internet now"

"It is a unique project"

Quotes are from an anonymous survey.
An Open Platform to Teach How the Internet Practically Works

Thomas Holterbach
ETH Zurich
thomahol@ethz.ch

Joint work with Tobias Bühler,
Tino Rellstab, and
Laurent Vanbever

ABSTRACT
Each year at ETH Zurich, around 100 students build and operate their very own Internet infrastructure composed of hundreds of routers and dozens of Autonomous Systems (ASes). Their goal? Enabling Internet-wide connectivity. We find this class-wide project to be invaluable in teaching our students how the Internet practically works. Our students have gained a much deeper understanding of the various Internet mechanisms alongside with their pitfalls. Besides students tend to love the project - clearly the fact that all of them need to cooperate for the entire Internet to work is empowering.

In this paper, we describe the overall design of our teaching platform, how we use it, and interesting lessons we have learnt over the years. We also make our platform openly available [8].

1 INTRODUCTION
Most undergraduate networking courses, including ours [23], aim at teaching “how the Internet works”. For the instructor, this typically means painstakingly going through the TCP/IP protocol stack, one layer at a time, following a bottom-up [23] or top-down approach [13]. At the end of the lecture, students (hopefully) have learned concepts such as switching, routing, and reliable transport; together with the corresponding protocols.

Learning these concepts is not sufficient to understand how the Internet really works through or, alternatively, why it does not work for this; we think one also needs to understand the ins and outs of how the Internet is operated which includes topics such as network design, network configuration, network monitoring, and… network debugging. Understanding these topics is important as Internet operations have a huge impact on its behavior. Among others, most of the Internet downturns are due to human-induced errors [17]. Yet, undergraduate networking courses seldom include these topics. At the end of the lecture, students (hopefully) have learned concepts such as switching, routing, and reliable transport; together with the corresponding protocols.

An open platform. Given the success of our project, we have open sourced the entire platform [8] and hope that other institutions will start using it. We built our platform with three key goals in mind. First, we aimed at faithfully emulating the real Internet infrastructure. To do so, we rely on (open-source) switching and routing software implementing the most well-known protocols (e.g., STP, OSPF, BGP, etc.).

Second, while we wanted the students to learn the intricacies of how Internet operators enable Internet-wide connectivity, between any pair of IP prefixes, by transmitting IP traffic across multiple student networks. As they quickly realize though, achieving this goal is challenging and requires a truly collective effort. We found this to be empowering. The fact that all networks need to work for the Internet as a whole to work truly helps to bring together the entire classroom.

Over the years, the mini-Internet project has become a flagship piece of our networking lecture, one that the new students look forward to. Thus far, the feedback we received from the students has been extremely positive, with comments such as: “It really allows us to apply the theoretical concepts,” “I am quite confident about many things on the Internet now,” and “It is a unique project.”

Besides gaining a much deeper understanding of the various Internet mechanisms, having students build and maintain their own Internet infrastructure enables them to quickly realize the pitfalls and shortcomings behind Internet operations. Students quickly realize: (i) how fragile the Internet infrastructure is and how dependent they are on their neighbors’ connectivity; (ii) how hard it is to troubleshoot Internet-wide problems; and (iii) how difficult it is to coordinate with each other to fix remote problems. Each year, several groups of students come up with proposals (sometimes, even implementations!) to improve Internet operations. These proposals often directly relate to research topics active in our community (such as configuration verification, synthesis or active probing).

Perhaps candidly, we believe that encountering operational problems early on in their networking curriculum can also help the next-generation of network designers avoid repeating the mistakes made in the past.
Outline

1. The mini-Internet mimics the real one

2. The mini-Internet turns the students into network operators

3. The mini-Internet provides students with tools to ease operations

4. The mini-Internet provides isolation, scales and is flexible
Outline

1. The mini-Internet **mimics** the real one

2. The mini-Internet turns the students into **network operators**

3. The mini-Internet provides students with tools to **ease operations**

4. The mini-Internet provides **isolation, scales** and is **flexible**
The AS-level topology of the mini-Internet

There are 60 ASes, divided in six regions
The AS-level topology exhibits many properties found in the real Internet.
The AS-level topology exhibits many properties found in the real Internet.

There are Tier1 (○○)
The AS-level topology exhibits many properties found in the real Internet.

There are Tier1 (●), Stub (○),
The AS-level topology exhibits many properties found in the real Internet

There are Tier1 (☐), Stub (⬤) and Transit (⬤) ASes
The AS-level topology exhibits many properties found in the real Internet.

There are Tier1 (○), Stub (●) and Transit (○) ASes.

There are peer-2-peer (---).
The AS-level topology exhibits many properties found in the real Internet.

There are Tier1 (●), Stub (○) and Transit (●) ASes.

There are peer-2-peer (---) and customer-provider links (→).
The AS-level topology exhibits many properties found in the real Internet.

There are Tier1 (🔴), Stub (〇), and Transit (🔵) ASes.

There are peer-2-peer (🟢) and customer-provider links (🟠).

There are IXPs (🔵)
We build internal topologies with the technologies used in practice
We build internal topologies with the technologies used in practice
We build internal topologies with the technologies used in practice.
We build internal topologies with the technologies used in practice.
We build realistic internal topologies that require students to solve real problems
We build realistic internal topologies that require students to solve real problems.
We build realistic internal topologies that require students to solve real problems.
We build realistic internal topologies that require students to solve real problems.
We build realistic internal topologies that require students to solve real problems.
Outline

1. The mini-Internet *mimics* the real one

2. The mini-Internet turns the students into *network operators*

3. The mini-Internet provides students with tools to *ease operations*

4. The mini-Internet provides *isolation, scales* and is *flexible*
We give one transit AS and one IP prefix to each group of students
We give one transit AS and one IP prefix to each group of students
Goal: enabling Internet-wide connectivity
We give one transit AS and one IP prefix to each group of students

Goal: enabling **Internet-wide connectivity**
We give one transit AS and one IP prefix to each group of students

Goal: enabling **Internet-wide connectivity**

<table>
<thead>
<tr>
<th>Phase</th>
<th>Project start</th>
<th>1. Establishing intra-domain connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Connectivity matrix**
Besides enabling internal connectivity, students have to perform traffic engineering.
Besides enabling internal connectivity, students have to perform traffic engineering.

In the L2 network:
- e.g., custom spanning tree and VLANs

In the L3 network:
- e.g., load-balancing
Besides enabling internal connectivity, students have to perform traffic engineering.

In the L2 network
  e.g., custom spanning tree and VLANs

In the L3 network
  e.g., load-balancing
Besides enabling internal connectivity, students have to perform **traffic engineering**

**In the L2 network**
e.g., custom spanning tree and VLANs

**In the L3 network**
e.g., load-balancing

...but the traffic from ZURICH to ROMA must use this low-delay path
We give one transit AS and one IP prefix to each group of students
Goal: enabling **Internet-wide connectivity**
We give one transit AS and one IP prefix to each group of students. 

Goal: enabling **Internet-wide connectivity**
We organise a Hackathon where students gather to configure BGP sessions
We give one transit AS and one IP prefix to each group of students.

**Goal:** enabling **Internet-wide connectivity**

<table>
<thead>
<tr>
<th>Phase</th>
<th>Project start</th>
<th>1. Establishing intra-domain connectivity</th>
<th>2. Establishing inter-domain connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connectivity matrix</td>
<td><img src="image1.png" alt="Matrix" /></td>
<td><img src="image2.png" alt="Matrix" /></td>
<td><img src="image3.png" alt="Matrix" /></td>
</tr>
</tbody>
</table>
We give one transit AS and one IP prefix to each group of students.

Goal: enabling **Internet-wide connectivity**
Besides enabling BGP sessions, students have to implement routing policies.
Besides enabling BGP sessions, students have to implement routing policies.

Following business agreements, e.g., local-preference and exportation rules.

Following preferences, e.g., one provider is preferred.
Besides enabling BGP sessions, students have to implement **routing policies**

Following business agreements
e.g., local-preference and exportation rules

Following preferences
e.g., one provider is preferred
Besides enabling BGP sessions, students have to implement **routing policies**

Following business agreements
e.g., local-preference and exportation rules

Following preferences
e.g., one provider is preferred

the inbound traffic coming from a provider must preferably arrive via R2
We give one transit AS and one IP prefix to each group of students. 

Goal: enabling **Internet-wide connectivity**

<table>
<thead>
<tr>
<th>Phase</th>
<th>Project start</th>
<th>1. Establishing intra-domain connectivity</th>
<th>2. Establishing inter-domain connectivity</th>
<th>3. Configuring external routing policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connectivity matrix</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Last year, up to 96% of the ASes were connected to each other
What the students learn goes beyond just configuring some protocols.
What the students learn goes beyond just configuring some protocols

They realise that the Internet is the result of a collective effort
Students often gather to configure the network together

They realise that the Internet is fragile
A small mistake may affect the overall connectivity

They realise that the Internet can be configured more efficiently
Students often come up with automation tools
Outline

1. The mini-Internet mimics the real one

2. The mini-Internet turns the students into network operators

3. The mini-Internet provides students with tools to ease operations

4. The mini-Internet provides isolation, scales and is flexible
Operating the mini-Internet is challenging and sometimes painful.
Operating the mini-Internet is challenging and sometimes painful. Fortunately, there are tools to help.
Our students have no a priori knowledge and a limited time budget
Our students have no a priori knowledge and a limited time budget. We assist them.
Our students have no a priori knowledge and a limited time budget. We assist them.

We organise Q&A sessions every week where teaching assistants provide help.
Our students have no a priori knowledge and a limited time budget. We assist them.

We organise Q&A sessions every week where teaching assistants provide help.

We use a dedicated Slack channel where students can ask questions any time.
Our students have no a priori knowledge and a limited time budget. We assist them.

We organise Q&A sessions every week where teaching assistants provide help.

We use a dedicated Slack channel where students can ask questions any time.

If we traceroute from ROMA-host to LOND-host this is the result:

```
traceroute to 43.101.0.1 (43.101.0.1), 30 hops max, 60 byte packets
1 ROMA-host.group43 (43.104.0.2)  0.085 ms  0.009 ms  0.007 ms
2 BARC-roma.group43 (43.0.4.1)   0.030 ms  0.007 ms  0.006 ms
3 LOND-zuri.group43 (43.0.1.1)   0.042 ms  0.009 ms  0.008 ms
4 HOST-lond.group43 (43.101.0.1) 0.027 ms  0.011 ms  0.010 ms
```

From line 2 to 3 it magically jumps from barc to zuri? We don't understand. Is this a mistake?
Our students have no a priori knowledge and a limited time budget. We assist them.

We organise Q&A sessions every week where teaching assistants provide help.

We use a dedicated Slack channel where students can ask questions any time.
Our students have no a priori knowledge and a limited time budget. We assist them.

We organise Q&A sessions every week where teaching assistants provide help.

We use a dedicated Slack channel where students can ask questions any time.
Monitoring and debugging a network is tricky
Monitoring and debugging a network is tricky. We provide **monitoring and debugging tools**.

**Looking glass:** the routing table of every router is available on a web interface.

**Active probing:** the students can run ping and traceroute between any pair of ASes to test connectivity.

**DNS:** the students can use domain names instead of IP addresses.
Students are not familiar with routers and switches’ CLI
Students are not familiar with routers and switches’ CLI

We provide a documentation tailored for the mini-Internet

---

Communication Networks

Project 1: Build your own Internet

Deadline: May 3 2018 at 11.59pm

In this document, we first introduce in §1 a set of commands you may need to configure an Open vSwitch. We then show in §2 how to configure a Quagga router.

1 Configuring Open vSwitch

Open vSwitch\(^1\) [1] is one of the most popular software switches. It can typically be used in virtual environments, for instance to connect two virtual machines. When an Open vSwitch is running, a set of commands are available to check its state and configure it. To print a brief overview of the switch state and its parameters, you can use the following command:

```plaintext
> ovs-vsctl show
```

This command also tells you the VLANs each port belongs to. One port has the name of the switch and has the type `internal`. This is a local port used by the host to communicate with the switch. You do not need to use this port. To get more precise information about the status of the ports, you can use the following command:

```plaintext
> ovs-ofctl show NAME
```

where `NAME` is the name of the switch (e.g. ETH-ZURICH). To get the current configuration and all the statistics of the switch, you can get a dump of the switch database with the following command:

```plaintext
> ovsdb-client
```

For example one entry of the database could look like this:

```plaintext
> 645981b6c 0 false [] 0 false [119f5-2be8bf5] [] []
```

eth-oerl

```plaintext
{
  stp-path-cost="100"
}
```

The information is a little bit cryptic, but we have highlighted in red the important information that you may use for the assignment. For instance, this entry is for the port named `eth-oerl`, the spanning tree path cost for this link is 100, and the spanning tree protocol has turned this port in block mode. You can also find the cost of the path to the root bridge by looking at the field `stp root path cost` (not shown here). This port is also a trunk port in VLANs 10 and 20.

You can also display information about the MAC address table of the switch with the following command:

```plaintext
> ovs-appctl fdb/show NAME
```

with `NAME` the name of the switch. To change the spanning tree priority of a switch (used by the spanning tree protocol to elect the root switch), you can use the following command:

```plaintext
> ovs-vsctl set bridge NAME stp-loop-guard off
```

where `NAME` is the name of the switch. This will turn off the spanning tree loop guard on all the ports of the switch.

---

\(^1\)https://www.openvswitch.org
Students do not progress at the same speed
Students do not progress at the same speed
We ensure **minimal connectivity**

We provide redundancy in the AS-level topology
Each transit AS has two providers and two customers

We pre-configure Tier1 and Stub ASes as well as IXPs
Enough for the students to answer most of the questions
Students must configure many virtual devices
Students must configure many virtual devices
We provide tools to facilitate the remote access to the virtual devices

Two commands are enough to access a router

laptop> ssh -p 2001 root@server
g1-proxy> ./goto.sh ZURI router
Outline

1. The mini-Internet mimics the real one

2. The mini-Internet turns the students into network operators

3. The mini-Internet provides students with tools to ease operations

4. The mini-Internet provides isolation, scales and is flexible
We rely on **docker containers** to isolate the different components of the mini-Internet (hosts, switches and routers).
We rely on **docker containers** to isolate the different components of the mini-Internet (hosts, switches and routers).
We rely on **docker containers** to isolate the different components of the mini-Internet (hosts, switches and routers).

> `ssh -p 2001 root@server`

Containers are connected with virtual links.
We rely on docker containers to isolate the different components of the mini-Internet (hosts, switches and routers).

Containers are connected with virtual links.

```
> ssh -p 2001 root@server
> openvpn —config client.conf
```
We use the state of the art software suites for the switches and routers
The mini-Internet topology can easily be adapted using configuration files.
The mini-Internet topology can **easily be adapted** using configuration files

Inspired from the Internet2 topology

Inspired from the SWITCH topology
A single server with 24 cores and 256 GB of memory can handle a mini-Internet with 60 ASes

When idle, the mini-Internet uses 30% of the CPUs in average, and 58% of the memory

Under stress*, the mini-Internet uses 51% of the CPUs in average, and 65.2% of the memory

*150 iperf sessions or 15000 BGP routes
We are developing a **visualisation framework**

Framework developed by Lina Gehri, Marco Di Nard, Aedan Christie and Alexander Dietmuller
We plan to connect the **real** Internet to the mini-Internet.
We plan to connect the **real** Internet to the mini-Internet

![Diagram of network connection](image)
How to run your own mini-Internet?

1. Pull from our GitHub page
   github.com/nsg-ethz/mini_internet_project

2. Follow the documentation

3. Define your topologies

4. Run it on your server
How to run your own mini-Internet?

1. Pull from our GitHub page
   github.com/nsg-ethz/mini_internet_project

2. Follow the documentation

3. Define your topologies

4. Run it on your server

Questions ?

thomahol@ethz.ch