

## Übung 9: Routing praktisch, Congestion control

2020-01-08

---

### 1 Routing praktisch

Anmerkung: Die erste Aufgabe ist nur zur Erinnerung

#### 1. Virtualisierungsumgebung

In späteren Übungsblättern werden Sie unter Linux Netzadministratorsaufgaben lösen. Dazu brauchen Sie zwei Dinge:

- (a) eine Virtualisierungsumgebung: Installieren Sie auf Ihrem Rechner<sup>1</sup> z.B. Virtualbox <https://www.virtualbox.org> oder ähnlich (VMWare, Parallels, o.ä.). Die Aufgaben werden mit Virtualbox entwickelt und getestet, so dass sich diese Anwendung anbieten dürfte.
- (b) Vagrant (<https://www.vagrantup.com>): Definition der Konfiguration virtueller Machine in einem kompaktem Format.

Beides steht für alle gängigen Betriebssysteme zur Verfügung. Installieren Sie das und testen Sie, ob Sie ein sog. Vagrantfile erzeugen und starten können. Insbesondere sollte das Vagrantfile im Verzeichnis `0_single_machine` von <http://groups.uni-paderborn.de/fg-karl/lehre/ws1920/cn/uebung/vagrants.tgz> funktionieren. Sie sollten sich mit `vagrant ssh web` dort anmelden können.

Machen Sie sich ggf. mit der Kommandozeile vertraut. Hinweise: <http://wiki.bash-hackers.org/scripting/tutoriallist>, <http://linuxcommand.org>, <https://ryantutorials.net/linuxtutorial/>, <https://www.codecademy.com/learn/learn-the-command-line>.

#### 2. Zwei Rechner

---

<sup>1</sup>Wenn Ihnen kein geeigneter Rechner zur Verfügung steht, kontaktieren Sie bitte den Übungsleiter.

- (a) Starten Sie das Vagrantfile in a\_web\_db.
  - (b) Loggen Sie sich auf beiden Rechnern ein (`vagrant ssh web` bzw. `vagrant ssh db`).
  - (c) Finden Sie die IP-Adressen der Rechner mit `ifconfig` heraus. (Hinweis: man `ifconfig`, ggf. man `man`).
  - (d) Prüfen Sie, ob Sie mit `ping` den jeweils anderen Rechnern erreichen können.
  - (e) Installieren Sie auf dem Rechner db das Programm `netcat`. (Hinweis: `sudo bash, apt-get install netmap, https://nmap.org/ncat/`).
  - (f) Erstellen Sie eine Textdatei und nutzen `ncat`, um diese Datei auf Port 8080 zum Herunterladen anzubieten. (Oder: Nutzen Sie die Ausgaben von `/bin/date`)
- Lösung:** Ein daytime Server:
- Server: `ncat -l -p 8080 --send-only --exec /bin/date`
  - Client: `ncat 192.168.56.102 8080`
- (g) Nutzen Sie auf dem Rechner web z.B. `curl`, `wget` oder auch `ncat` selbst, um diese Datei zu laden.

### 3. IP-Forwarding konfigurieren

Nutzen Sie das Vagrantfile in b\_one\_router, um drei virtuelle Maschine mit Vagrant zu starten. Melden Sie sich mit `vagrant ssh` auf jeder der drei Maschinen an.

Versuchen Sie, von der Maschine web einen ping auf die Maschine db mit IP-Adresse 192.168.2.102 zu machen. Das sollte nicht gelingen! (Pings zwischen router und web bzw. router und db sollten in beide Richtungen gehen – probieren Sie das aus und schauen sich davor und danach jeweils die arp-Caches an – Kommando `arp`.)

Ihre Aufgabe ist es, dieses ping von web zu db (und umgekehrt) zu ermöglichen, in dem Sie die drei VMs richtig konfigurieren. Hierzu ein paar Hinweise:

- (a) Recherchieren Sie, welche Änderung in der Konfiguration des Routers Sie vornehmen müssen, damit diese VM prinzipiell Pakete weiterleitet.

**Lösung:** Forwarding aktivieren.

- Permanent: in `/etc/sysctl.conf` die Einträge `net.ipv4.ip_forward=1` und `net.ipv6.conf.all.forwarding=1` einkommentieren.
- Just once: `sysctl -w net.ipv4.conf.all.forwarding=1` (Hinweis: man `sysctl`)

- (b) Zusätzlich müssen Sie beim router Regeln eintragen, damit Pakete zwischen den beiden Ethernet-Interfaces enp0s8 und enp0s9 weitergereicht werden. (Hinweis: die Benennung der Interfaces ist eine Ubuntu-Besonderheit.) Hierzu ist das Werkzeug iptables oder, moderner, nft geeignet.

Überlegen Sie sich, warum Sie (bei iptables) zwei Regeln brauchen.

**Lösung:** sudo iptables -A FORWARD -i enp0s8 -o enp0s9 -j ACCEPT  
sudo iptables -A FORWARD -i enp0s9 -o enp0s8 -j ACCEPT

- (c) Trotzdem wird das ping noch nicht gehen. Welche Änderungen müssen Sie noch bei web und db durchführen, damit das gelingt? Warum müssen Sie diese Änderung auf *beiden* Rechnern durchführen?

**Lösung:** web: sudo route add -net 192.168.2.0 netmask 255.255.255.0 gw 192.168.1.2

db: sudo route add -net 192.168.1.0 netmask 255.255.255.0 gw 192.168.2.2

Hinweis: Es ist kein DNS konfiguriert; Sie müssen also mit IP-Adressen arbeiten.

## 2 Congestion control

### 1. Greedy-Verhalten schadet sich selbst

Nehmen Sie an, Sie haben ein einfaches Ring-Netz wie in der folgenden Abbildung 1 gezeigt.

Machen Sie folgenden Annahmen:

- Alle Sender senden mit der gleichen Rate  $\lambda$  und möchten zwei Links in Uhrzeigerrichtung benutzen (Knoten  $i$  möchte also an Knoten  $i + 2$  senden, modulo Anzahl der Knoten).
- Alle Links haben eine feste Datenrate  $c$ . Ist ein Link überlastet (ist die Summe der eingehenden Raten größer als  $c$ ), so verteilt der Link seine Datenrate zwischen seinen beiden Eingängen proportional zu den eingehenden Datenraten. Überzählige Pakete werden verworfen.

Beispiel: Link 3 bekommt von seinem Vorgänger Link 2  $3/4c$  zur Weiterleitung, von seinem lokalen Knoten 3  $c/2$  zur Weiterleitung. Dann wird der Verkehr von Link 2 mit  $\frac{3/4c}{3/4c+1/2c} \cdot c = 4/5c$  weitergeleitet ( $\frac{1-4/5}{3/4}c$  geht also verloren). Entsprechend werden von Knoten 3s  $1/2c$  nur  $1/5c$  weitergeleitet.

Berechnen Sie, in Abhängigkeit von  $\lambda$ , mit welcher Datenrate ein Knoten  $i$  zu seinem Empfänger  $i + 1$  senden kann.

Hinweis: Eine Fallunterscheidung nach  $\lambda \leq \frac{c}{2}$  bzw.  $\lambda > \frac{c}{2}$  ist sinnvoll.

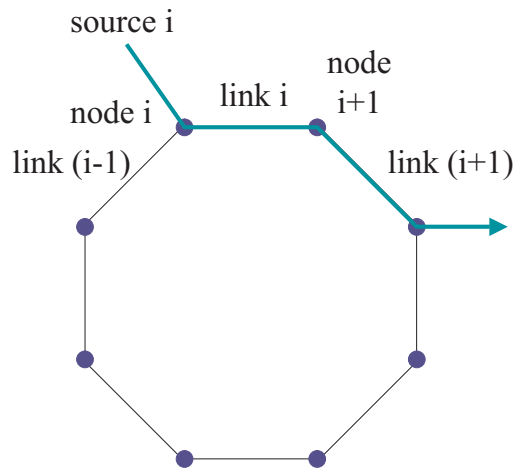


Abbildung 1: Ring-Netz für congestive collapse, nach [http://ica1www.epfl.ch/PS\\_files/LEB3132.pdf](http://ica1www.epfl.ch/PS_files/LEB3132.pdf)

**Lösung:** Die beste Erklärung findet sich in Abschnitt 1.1.1 von [https://ica1www.epfl.ch/PS\\_files/LEB3132.pdf](https://ica1www.epfl.ch/PS_files/LEB3132.pdf)

## 2. Fehlerrate und TCP-Durchsatz

Wie groß darf die Bitfehlerrate eines Links höchstens sein, wenn Sie einen Durchsatz von 10 Gb/s erreichen möchten? Nehmen Sie an, dass

- die Datenrate des Links unbeschränkt ist,
- die RTT 100 ms beträgt,
- die MSS 1500 Bytes ist.

Nutzen Sie die Formel auf Folien 45, Kapitel 7. Geben Sie auch das entsprechende CWND an.

Ist dies realistisch? Optische Links erreichen typischerweise Bitfehlerraten von bis zu  $10^{-13}$ .

**Lösung:** Durchsatzformel:  $D \leq \frac{1.22 \cdot \text{MSS}}{\text{RTT} \sqrt{p}}$

Auflösen nach Paketfehlerrate  $p$  in Werte einsetzen:  $p = \left( \frac{1.22 \cdot \text{MSS}}{\text{RTT} \cdot D} \right)^2$

oder  $p = \left( \frac{1.22 \cdot 12.000}{0.1 \cdot 10^{-10^9}} \right)^2 = (1.46410^{-5})^2 \approx 2.14 \cdot 10^{-10}$

Ein entsprechende Bitfehlerrate wäre dann:  $1 - p = (1 - \text{BER})^{12.000} \leftrightarrow \log(1 - \text{BER}) = \frac{1}{12.000} \log(1 - p) \leftrightarrow \text{BER} = 1 - e^{\frac{1}{12.000} \log(1-p)} \approx 1 - 7.8 \cdot 10^{-15}$

Es fehlen also zwei Zehnerpotenzen.

Das CWND müsste dem Datenraten-Delay-Produkt entsprechen, also  $0.1s \cdot 10 \cdot 10^9 \text{bit/s} = 10^9 \text{bit} = 125\text{MB}$

### 3. CWND-Größe und Slowstart

Wie lange dauert es, bis die für 10 Gb/s notwendige CWND-Größe erreicht ist (Annahme: Slowstart, MSS 1500 Byte, RTT 100 ms)?

**Lösung:** Soll-CWND: Bandwidth-Delay-Produkt, also hier 10 Gbit/s mal 50 ms =  $10^{10} \cdot 100 \cdot 10^{-3} = 10^9$  bits = 1 Gbits. Oder in Paketen: 125.000 Pakete.

Slowstart verdoppelt die CNWDs einmal pro RTT. Wir sehen also CWNDs von 1, 2, 4, 8, ... Wir suchen also das kleinste ganzzahlige  $k$  so dass  $2^k > 125.000$  ist. Das ist  $k = 17$ .

Wir brauchen also 17 RTTs, also 1.7 Sekunden.

In der Zeit übertragen wir ca.  $1 + 2 + 4 + \dots + 2^{17} = 2^{18} - 1$  viele Pakete zu je 1500 Bytes. Das sind ca. 375 Mbyte.

### 4. TCP SACK

Wie sieht das genaue Format für TCP SACK aus? Welche Information ist in einem ACK enthalten?

**Lösung:** See <https://tools.ietf.org/html/rfc2018>, Section 3

### 5. TCP-Varianten

Nutzen Sie verschiedene TCP-Varianten, um eine größere Datei zu laden. Sie können unter Linux eine Liste von Optionen hier finden: `/proc/sys/net/ipv4/tcp_available_congestion_control`. Mit dem Kommando `sysctl net.ipv4.tcp_congestion_control` und der gewünschten TCP-Variante als Parameter können Sie zu einer anderen TCP-Variante als default wechseln.

Hinweis: `man sysctl`. Mit `setsockopt()` kann man auch aus einem Anwendungsprogramm heraus, pro TCP-Verbindung, auswählen.