Assignment 1:
Delays, Throughput, and Paths

Deadline:
Professor: Maria Papadopouli
TA: Katerina Lionta
mailing list: hy335a-list@csd.uoc.gr
TA: klionta@csd.uoc.gr

Part 1: Network delays, packet and circuit switching, throughput (50 points)

1.1 Υποθέστε ότι οι χρήστες μοιράζονται μία ζεύξη 2 Mbps. Επίσης, υποθέστε ότι κάθε χρήστης μεταδίδει συνεχώς με ρυθμό 1 Mbps όταν κάνει μετάδοση, αλλά κάθε χρήστης μεταδίδει μόνον 20% του χρόνου.

A. Εάν η ζεύξη υποστηρίζει TDM, πόσους χρήστες μπορεί να εξυπηρετήσει; Αιτιολογήστε την απάντησή σας. ***(2 points)***
Απάντηση
Αν η ζεύξη είναι 2 Mbps και κάθε χρήστης μπορεί να μεταδώσει με ρυθμό 1Mbps, τότε η μεταγωγή κυκλώματος μπορεί να υποστηρίζει:
2 Mbps/1Mbps = 2 χρήστες

B. Εάν η ζεύξη υποστηρίζει FDM, πόσοι χρήστες μπορούν να εξυπηρετηθούν ταυτόχρονα; Αιτιολογήστε την απάντησή σας. ***(2 points)***
Απάντηση
Αν η ζεύξη είναι 2 Mbps και κάθε χρήστης μπορεί να μεταδώσει με ρυθμό 1Mbps, τότε η μεταγωγή κυκλώματος μπορεί να υποστηρίζει:
2 Mbps/1Mbps = 2 χρήστες

C. Συγκρίνετε την απόδοση της μεταγωγής κυκλώματος (circuit switching) με την απόδοση του TDM και του FDM ***(2 points)***
Απάντηση
Δεν έχουν διαφορά στην απόδοση και τα δυο μπορουν να υποστηριξουν 2 χρηστες που μεταδιδουν με 1 Mbps αν έχουν διαθέσιμη μια ζεύξη 2 Mbps.

D. Για το υπόλοιπο αυτού του προβλήματος, υποθέστε ότι χρησιμοποείται μεταγωγή πακέτου (packet switching). Γιατί δεν θα υπάρχει επι της ουσίας καθυστέρηση ουράς (queue delay) πριν την ζεύξη, εάν δύο ή λγότεροι χρήστες μεταδίδουν ταυτόχρονα;

Γιατί θα υπάρχει καθυστέρηση ουράς (queue delay) αν τρεις χρήστες μεταδίδουν ταυτόχρονα; ***(3 points)***

Απάντηση

Αν δύο ή λιγότεροι χρήστες μεραδίδουν ταυτόχρονα τότε ο μέγιστος ρυθμός μετάδοσης που μπορούν να φθάνουν είναι 1 και 2 Mbps αντίστοιχα, τα οποία είναι ≤ από το ρυθμό μετάδοσης της ζεύξης. Αν μεταδίδουν ταυτόχρονα τρεις ή παραπάνω χρήστες θα ξεπεράσουν τον ρυθμό μετάδοσης της ζεύξης και θα υπάρξει καθυστέρηση.

E. Βρείτε την πιθανότητα να μεταδίδει κάποιος συγκεκριμένος χρήστης. ***(3 points)***

Απάντηση

Αφού η πιθανότητα να μεταδίδει ένας χρήστης είναι 20%, η πιθανότητα να μεταδίδει ένας χρήστης σε μια τυχαία χρονικη στιγμή ειναι 0.2.

F. Θεωρήστε ότι στέλνετε ένα πακέτο από έναν υπολογιστή προέλευσης (source host) σε έναν υπολογιστή προορισμού (destination host) μέσω μίας σταθερής διαδρομής. Απαριθμήστε τις συνιστώσες της καθυστέρησης για την καθυστέρηση end-to-end. Ποιες από αυτές τις καθυστερήσεις είναι σταθερές και ποιές μεταβλητές; ***(2 points)***

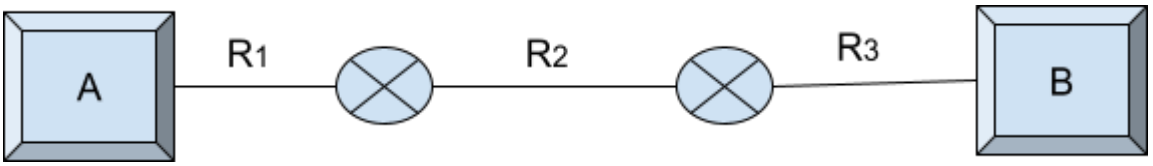
Απάντηση

- d_{proc}: Καθυστέρηση επεξεργασίας κόμβου, σταθερή
- d_{prop}: Καθυστέρηση διάδοσης, είναι σταθεή για δεδομένη απόσταση μεταξύ των end hosts και τύπο μέσου
- d_{queue}: Καθυστέρηση ουράς, είναι μεταβλητή αφού εξαρτάται από τη συμφόρηση του δικτύου
- d_{trans}: Καθυστέρηση μετάδοσης, είναι σταθερή για δεδομένο bandwidth

1.2 Υποθέστε ότι ο υπολογιστής A θέλει να στείλει ένα μεγάλο αρχείο στον Υπολογιστή B. Η διαδρομή από τον υπολογιστή A μέχρι τον Υπολογιστή B έχει τρεις ζεύξεις, με ρυθμούς R_1=500 kbps, R_2= 2 Mbps και R_3=1 Mbps.

A. Σχεδιάστε το set up του προβλήματος. ***(2 points)***

Απάντηση



B. Υποθέτοντας ότι δεν υπάρχει άλλη κίνηση στο δικτύο, ποια είναι η διεκπεραιωτική ικανότητα (throughput) για τη μεταφορά του αρχείου; ***(3 points)***

Απάντηση

$\min\{R_1, R_2, R_3\} = R_1 = 500$ kbps

C. Υποθέστε ότι το αρχείο έχει μέγεθος 4 εκατομμύρια bytes. Πόσο θα διαρκέσει η μεταφορά του αρχείου στον Υπολογιστή B; ***(3 points)***

Απάντηση

1 byte = 8 bits
time = 4Mb/throughput = 4Mb/500kbps = 4.000.000*8/500.000 = 320/5 = 64s

D. Επαναλάβετε τις (B) και (C), αλλά τώρα χρησιμοποιήστε τιμή R_2=100 kbps. ***(5 points)***

Απάντηση

(A) min{R_1, R_2, R_3} = R_2 = 100 kbps

(B) time = 4.000.000*8/100.000 = 320s

1.3 Θεωρήστε μία εφαρμογή που μεταδίδει δεδομένα με σταθερό ρυθμό (π.χ. ο αποστολέας παράγει μία μονάδα N-bit δεδομένων κάθε k μονάδες χρόνου, όπου το k είναι μικρό και σταθερό). Επίσης, όταν αρχίσει μία τέτοια εφαρμογή, θα συνεχίσει να εκτελείται για ένα σχετικά μεγάλο χρονικό διάστημα. Απαντήστε στις παρακάτω ερωτήσεις, δικαιολογώντας με λίγα λόγια την απάντησή σας:

A. Γι΄αυτήν την εφαρμογή θα ήταν πιο κατάλληλο ένα δίκτυο μεταγωγής πακέτου (packet switching) ή ένα δίκτυο μεταγωγής κυκλώματος (circuit switching); Γιατί; ***(4 points)***

Απάντηση

Μια εφαρμογή που μεταδίδει δεδομένα με σταθερό ρυθμό και εκτελείται για μεγάλα χρονικά διαστήματα δουλεύει καλύτερα με ένα δίκτυο μεταγωγής κυκλώματος (circuit switching) αφού προσδίδει:

- Εγγυημένο εύρος ζώνης (bandwidth)
- Σταθερή καθυστέρηση, κάθε χρήστης έχει σταθερό bandwidth κάθε χρονική στιγμή
- Αποφυγή συμφόρησης (congestion), δεν υπάρχει ανταγωνισμός με άλλα πακέτα

B. Υποθέστε ότι χρησιμοποιείται ένα δίκτυο μεταγωγής πακέτου (packet switching) και ότι η μόνη κίνηση μέσα σ΄αυτό το δίκτυο πρόερχεται από εφαρμογές όπως αυτήν που περιγράψαμε παραπάνω. Ακόμη, υποθέστε ότι το άθροισμα των ρυθμών δεδομένων των εφαρμογών είναι μικρότερο από τις χωρητικότητες (bandwidth) κάθε μιας από τις ζεύξεις. Χρειάζεται κάποια μορφή ελέγχου συμφόρησης (congestion control); Γιατί; ***(4 points)***

Απάντηση

Ο έλεγχος συμφόρησης δεν είναι απαραίτητος, οι λόγοι είναι οι εξής:

- Αφού το άθροισμα των ρυθμών δεδομένων των εφαρμογών είναι μικρότερο από τις χωρητικότητες των ζεύξεων δεν θα δημιουργούνται μεγάλες ουρές, τα πακέτα θα περνάνε από τη ζεύξη μόλις φθάνουν.
- Αφού τα δεδομένα έρχονται με σταθερό ρυθμό και όχι κατα ριπές πάλι οδηγούμαστε στο συμπέρασμα ότι δεν θα υπάρχουν ουρές.

1.4 Θεωρήστε ένα σενάριο όπου οι χρήστες παράγουν δεδομένα με ρυθμό 100 kbps όταν είναι απασχολημένοι, αλλά παράγουν δεδομένα μόνο με πιθανότητα p=0,1. Υποθέστε ότι η ζεύξη όπου μεταδίδουν οι χρήστες είναι 1 Gbps.

A. Ποιος είναι ο μέγιστος αριθμός N χρηστών που μπορούν να υποστηρίζονται ταυτόχρονα με μεταγωγή κυκλώματος (circuit switching); ***(5 points)***

Απάντηση

Στη μεταγωγή κυκλώματος ο κάθε χρήστης καταλαμβάνει ένα ποσοστό της ζεύξης είτε έχει να στείλει δεδομένα είτε όχι. Αν ο κάθε χρήστης στέλνει με ρυθμό 100 kbps και υπάρχει στη διάθεσή μας μια ζεύξη 1 Gbps τότε ο μέγιστος αριθμός χρηστών που μπορεί α υποστηρίζει το σύστημα είναι:

$$N = 1\ Gbps/100\ kbps = 1*10^9\ bps/\ 100*10^3\ bps = 1*10^4 = 10.000\ \text{χρήστες}$$

B. Θεωρήστε τώρα μεταγωγή πακέτου (packet switching) και πληθυσμό χρηστών M. Δώστε έναν τύπο (συναρτήσει p, M, N) για την πιθανότητα περισσότεροι από N χρήστες να στέλνουν δεδομένα. ***(7 points)***

Απάντηση

- M: Ο συνολικός αριθμός χρηστών στο δίκτυο μεταγωγής πακέτων
- N: Ο συνολικός αριθμός χρηστών στο δίκτυο μεταγωγής κυκλώματος
- p: Η πιθανότητα κάποιος χρήστης να στείλει δεδομένα μια συγκεκριμένη χρονική στιγμή

Η πιθανότητα ένας χρήστης να στείλει δεδομένα είναι p και η πιθανότητα να μην στείλει είναι 1-p, επομένως η πιθανότητα να στέλνουν k χρήστες από τους M, θα είναι, χρησιμοποιώντας binomial distribution:

$$P(K=k) = \binom{M}{k} * p^k * (1-p)^{M-k}$$

Τώρα η πιθανότητα να στέλνουν πάνω από N χρήστες ταυτόχρονα, είναι:

$$P(K>N) = 1 - P(K \leq N) = 1 - \sum_{k=0}^{N} \binom{M}{k} * p^k * (1-p)^{M-k}$$

Part 2: Ping, Traceroute, Wireshark (50 points)

!ΠΡΟΣΟΧΗ:

- **Τρέξτε τις εντολές ping και traceroute (tracert για windows) με το flag -4**
- **Οι χρήστες των windows πρέπει να ανοίγουν το command prompt και το wireshark ως administrators**

2.1

A. Τρέξτε ένα ping προς ένα υπολογιστή της google, παραθέστε το screenshot του output στην αναφορά και εξηγήστε τα διάφορα fields. **(5 points)**
Απάντηση

```
C:\WINDOWS\system32>ping -4 google.com

Pinging google.com [142.250.187.142] with 32 bytes of data:
Reply from 142.250.187.142: bytes=32 time=39ms TTL=116
Reply from 142.250.187.142: bytes=32 time=40ms TTL=116
Reply from 142.250.187.142: bytes=32 time=37ms TTL=116
Reply from 142.250.187.142: bytes=32 time=44ms TTL=116

Ping statistics for 142.250.187.142:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 37ms, Maximum = 44ms, Average = 40ms
```

B. Ανοίξτε το wireshark και τρέξτε ένα ping που στέλνει ακριβώς 5 πακέτα προς έναν υπολογιστή της google.
   a. Παραθέστε screenshot απο το output του ping καθώς και την εντολή που τρέξατε **(5 points)**
   Απάντηση

```
C:\WINDOWS\system32>ping -4 -n 5 google.com

Pinging google.com [142.250.187.142] with 32 bytes of data:
Reply from 142.250.187.142: bytes=32 time=45ms TTL=116
Reply from 142.250.187.142: bytes=32 time=38ms TTL=116
Reply from 142.250.187.142: bytes=32 time=38ms TTL=116
Reply from 142.250.187.142: bytes=32 time=41ms TTL=116
Reply from 142.250.187.142: bytes=32 time=38ms TTL=116

Ping statistics for 142.250.187.142:
    Packets: Sent = 5, Received = 5, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 38ms, Maximum = 45ms, Average = 40ms
```

   b. Σταματήστε το capturing στο wireshark και παραθέστε ένα screenshot από το output του wireshark που δείχνει μόνο τα πακέτα που έστειλε ως απάντηση ο υπολογιστής της google. Πως το βρήκατε; Εξηγήστε οτιδήποτε ασυνήθιστο στο output. **(5 points)**

Απάντηση

Από το screenshot βλέπουμε ότι η destination ip είναι η 142.250.187.142, άρα στο filter bar του wireshark θα γράψουμε ip.src == 142.250.187.142, για να δούμε όλα τα πακέτα που μας έστειλε ως απάντηση ο υπολογιστής της google. Αφού λάβαμε απαντήσεις και για τα πέντε πακέτα θα υπάρχουν πέντε πακέτα στο wireshark με αυτή την ip σαν source.

ip.src == 142.250.187.142

No.	Time	Source	Destination	Protocc	Lengt	Info
687	43.915043	142.250.187.142	192.168.2.18	ICMP	74	Echo (ping) reply id=0x0001, seq=13/3328, ttl=116 (request in 686)
689	44.921637	142.250.187.142	192.168.2.18	ICMP	74	Echo (ping) reply id=0x0001, seq=14/3584, ttl=116 (request in 688)
693	45.948640	142.250.187.142	192.168.2.18	ICMP	74	Echo (ping) reply id=0x0001, seq=15/3840, ttl=116 (request in 692)
702	46.965194	142.250.187.142	192.168.2.18	ICMP	74	Echo (ping) reply id=0x0001, seq=16/4096, ttl=116 (request in 700)
736	47.980071	142.250.187.142	192.168.2.18	ICMP	74	Echo (ping) reply id=0x0001, seq=17/4352, ttl=116 (request in 732)

2.2

A. Περιγράψτε με λίγα λόγια πως με τη βοήθεια του traceroute ανακαλύπτουμε το path προς ένα απομακρυσμένο host. **(10 points)**

Απάντηση

Τρέχοντας την εντολή traceroute και το όνομα του υπολογιστή ή την ip του, μπορούμε να δούμε πόσα hops μεσολαβούν. Το traceroute χρησιμοποιεί το ttl (time-to-live) metric, το οποίο είναι ένας αριθμός που δίνεται σε κάθε πακέτο και κάθε φορά που το πακέτο περνάει από ένα hop αυτός ο αριθμός μειώνεται κατα ένα, όταν το ttl γίνει 0 οι routers απορρίπτουν το πακέτο. Όταν ξεκινάει το traceroute στέλνει πρώτα ένα πακέτο με ttl=1, φτάνει στο πρώτο hop όπου το ttl του πακέτου μηδενίζεται, έπειτα στέλνει ένα πακέτο με ttl=2 κοκ, μέχρι να φτάσει στον τελικό προορισμό.

B. Τρέξτε ένα traceroute με προορισμό έναν υπολογιστη της google, παραθέστε το screenshot στην αναφορά και περιγράψτε το output, προσπαθήσε να δώσετε τις πιθανές αιτίες σε οτιδήποτε ασυνήθιστο στο output. **(10 points)**

Απάντηση

Από το screenshot παρακάτω μπορούμε να παρατηρήσουμε τα εξής:

- Το πακέτο πέρασε ο 9 hops για να φτάσει στον προορισμό του
- Το 3ο hop πιθανόν για privacy reasons δεν μας έδωσε πληροφορίες για την ip του
- Η απάντηση για το τελευταίο πακέτο έφτασε πιο νώρις από ότι για τα 5ο, 6ο, 7ο, 8ο, αυτό πιθανόν να οφείλεται σε συμφόρηση στους συγκεκριμένους routers
- Οι χρόνοι το πακέτων συγκριτικά στα 3 διαφορετικά runs είναι περίπου ίδιοι

```
C:\WINDOWS\system32>tracert -4 google.com

Tracing route to google.com [142.250.187.110]
over a maximum of 30 hops:

  1     4 ms     4 ms     4 ms  vodafone.station [192.168.2.1]
  2    24 ms    21 ms    22 ms  loopback2004.med01.dsl.hol.gr [62.38.0.170]
  3     *        *        *     Request timed out.
  4    23 ms    25 ms    22 ms  ae3-100-ucr1.atm.cw.net [195.89.103.89]
  5    39 ms    39 ms    39 ms  ae24-xcr1.sof.cw.net [195.2.16.5]
  6    40 ms    48 ms    39 ms  209.85.168.146
  7    40 ms    40 ms    42 ms  216.239.62.49
  8    39 ms    39 ms    39 ms  142.251.52.83
  9    38 ms    38 ms    38 ms  sof02s44-in-f14.1e100.net [142.250.187.110]

Trace complete.
```

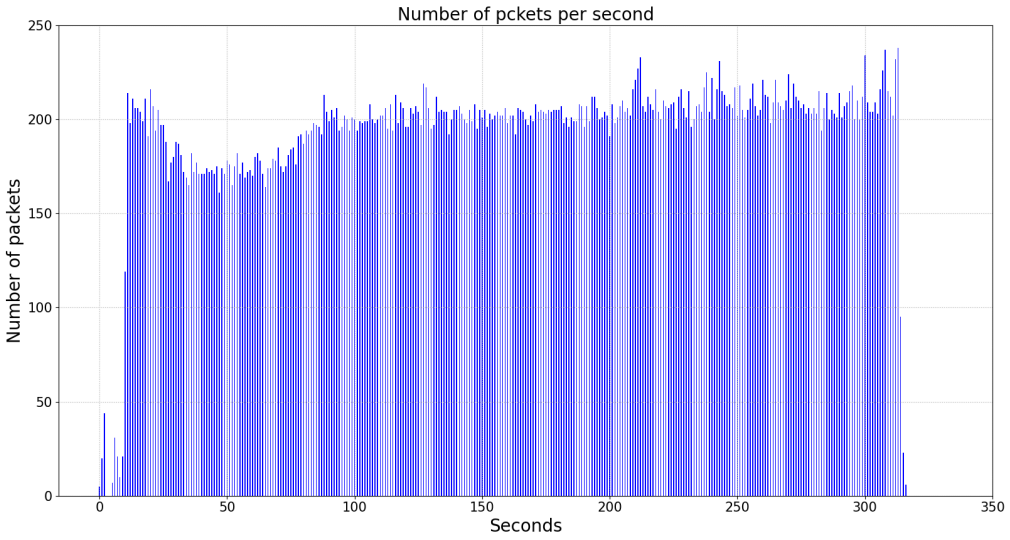
2.3 Σε αυτή την άσκηση σας δίνεται ένα .csv αρχείο το οποίο είναι ένα output από ένα wireshark run CS335a, χρησιμοποιώντας python υπολογίστε τα παρακάτω:

A. Υπολογίστε την συνολική ποσότητα δεδομένων (σε bytes) που έγιναν capture. **(5 points)**
Απάντηση
23,291,256 bytes

B. Βρείτε τις πρώτες 5 IPs από τις οποίες στάλθηκαν τα περισσότερα δεδομένα. **(5 points)**
Απάντηση

IPs	Bytes
173.0.68.4	12412218
192.168.1.133	10835407
23.102.24.114	10195
192.229.145.200	6462
04.40.141.105	6352

C. Χωρίστε τα δεδομένα ανα δευτερόλεπτο και φτιάξτε ένα bar plot που δείχνει το πλήθος των πακέτων που έγιναν capture ανα δευτερόλεπτο. **(5 points)**
Απάντηση



Για ευκολία χρησιμοποιήστε την βιβλιοθήκη pandas για να ανοίξετε το csv και επεξεργαστήτε το ως dataframe.

Format Guidelines

- Είναι σημαντικό κάθε διάγραμμα να παρουσιάζει καθαρά τους τίτλους και τις μονάδες του άξονα x και y. Θα πρέπει επίσης να περιλαμβάνει λεζάντες (legends) και πλέγμα (grid).
- Κάθε εικόνα (Figure), η οποία μπορεί να περιλαμβάνει ένα ή περισσότερα διαγράμματα, θα πρέπει να έχει αριθμό και μια λεζάντα (caption) που να περιγράφει ξεκάθαρα τα κύρια αποτελέσματα των διαγραμμάτων.

Η αναφορά σας θα πρέπει να έχει την εξής μορφή:

- Γραμματοσειρά: Arial ή Times New Roman
- Μέγεθος γραμματοσειράς κύριου κειμένου: 11pt
- Τίτλοι: 12pt/έντονη γραφή (bold)
- Λεζάντα εικόνας: Η πρώτη γραμμή της λεζάντας πρέπει να είναι σε μέγεθος γραμματοσειράς 10pt και έντονη γραφή (bold), και να περιλαμβάνει τα κύρια ευρήματα. Μια περιγραφή κάθε διαγράμματος πρέπει να ακολουθεί.

Submission

- Συγκεντρώστε την αναφορά σας σε ένα ενιαίο αρχείο PDF
- Συμπεριλάβετε τον Python κώδικά σας (μαζί με ένα αρχείο readme) και το αρχείο CSV.
- Συμπιέστε τα σε ένα αρχείο .zip.
- Στείλτε το στο email: **klionta@csd.uoc.gr** με θέμα: **335a_assign1_AM** (οι παραδόσεις με διαφορετικό θέμα δεν θα γίνονται δεκτές).
- Για οποιαδήποτε απορία στείλτε στη λίστα **hy335a-list@csd.uoc.gr**