



ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΜΑΘΗΜΑ: ΛΕΙΤΟΥΡΓΙΚΑ ΣΥΣΤΗΜΑΤΑ
ΑΚΑΔΗΜΑΪΚΟ ΕΤΟΣ: 2015-2016

Εργασία στο Bash Scripting

3 ατόμων

Βοηθοί μαθήματος: 1. Παρασκευάς Ελευθέριος (eparaskei@csd.auth.gr)
2. Φούντα Αντιγόνη-Μαρία (founanti@csd.auth.gr)

Προθεσμία: **23/11/2015 (23:59)**

Εισαγωγή

Το bash είναι ένα από τα πιο διαδεδομένα κελύφη (unix shells) και αποτελεί το καθιερωμένο κέλυφος σε πολλές διανομές Linux. Στόχος της εργασίας είναι να εξοικειωθείτε με τη χρήση του, ώστε να είστε σε θέση να αυτοματοποιήσετε εργασίες αναπτύσσοντας bash scripts.

Στην παρούσα εργασία καλείστε να αναπτύξετε bash script το οποίο:

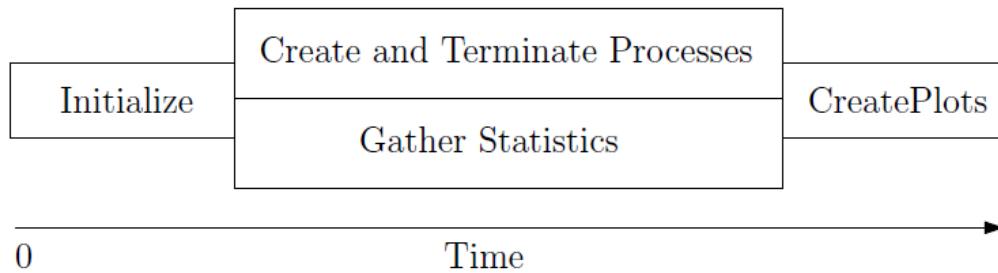
1. Θα εκκινεί τον browser Chromium, θα διαβάζει μία λίστα από urls και θα ανοίγει αντίστοιχο αριθμό tabs,
2. Θα καταγράφει τους πόρους του συστήματος που καταναλώνονται από τις διεργασίες του Chromium,
3. Θα φροντίζει για τον διαδοχικό τερματισμό όλων των διεργασιών του προαναφερθέντος browser, και τέλος
4. Θα δημιουργεί γραφικές παραστάσεις όπου θα απεικονίζεται η χρήση των πόρων του συστήματος κατά την εκτέλεση του Chromium.

Αναλυτική Περιγραφή Εργασίας

Το script το οποίο θα αναπτύξετε θα πρέπει να εκτελεί τις παρακάτω ανεξάρτητες λειτουργίες:

1. προετοιμασία εκτέλεσης script (initialize)
2. εκκίνηση και τερματισμός διεργασιών του Chromium
3. καταγραφή στατιστικών στοιχείων
4. δημιουργία και αποθήκευση γραφικών παραστάσεων

Για κάθε μία από αυτές τις λειτουργίες θα πρέπει να αναπτυχθεί και μία αντίστοιχη συνάρτηση. Στο Σχήμα 1 παρουσιάζεται η πορεία εκτέλεσης του script. Παρατηρήστε ότι οι συναρτήσεις που αφορούν στην εκκίνηση και τερματισμό των διεργασιών και στην καταγραφή των στατιστικών στοιχείων εκτελούνται παράλληλα. Ο παραλληλισμός συναρτήσεων του bash μπορεί να επιτευχθεί πολύ εύκολα με την εκτέλεσή τους στο παρασκήνιο (background).



Σχήμα 1: Η πορεία εκτέλεσης του script

Προετοιμασία Εκτέλεσης

Η συνάρτηση που αφορά στην προετοιμασία εκτέλεσης του script (initialize) θα πρέπει να τερματίζει τυχόν υπάρχουσες διεργασίες του Chromium, ώστε αυτές να μην επηρεάσουν την καταγραφή των στατιστικών στοιχείων. Στη συνάρτηση αυτή μπορείτε επίσης να τοποθετήσετε εντολές που θα εξασφαλίζουν τη σωστή λειτουργία του script σας (π.χ. δημιουργία ή καταστροφή προσωρινών αρχείων), αν αυτό απαιτείται.

Διαχείριση Διεργασιών

Η συνάρτηση που αφορά στην διαχείριση των διεργασιών θα πρέπει να εκτελεί διαδοχικά τα παρακάτω:

1. Εκκίνηση μίας διεργασίας του Chromium (απλή κλήση του εκτελέσιμου χωρίς κανένα όρισμα)
2. Ανάγνωση της λίστας με τα urls από το αρχείο url.in
3. Για κάθε url στη λίστα:
 - a. Κλήση του Chromium με όρισμα το αντίστοιχο url (παράδειγμα: chromium-browser www.google.com)
 - b. Αναμονή 5 δευτερολέπτων (sleep)
4. Αναμονή 30 δευτερολέπτων (resting time)
5. Δημιουργία λίστας με όλα τα pids που σχετίζονται με τις διεργασίες του Chromium
6. Τερματισμός (kill -9) της νεότερης διεργασίας. Ως νεότερη θεωρείται η διεργασία που δημιουργήθηκε τελευταία.
7. Αναμονή 5 δευτερολέπτων
8. Επανάληψη από το βήμα 5 έως ότου δεν έχει απομείνει καμία διεργασία που να ανήκει στον Chromium
9. Ενημέρωση συνάρτησης καταγραφής στατιστικών για την ολοκλήρωση της καταγραφής (προαιρετικά - αν απαιτείται)

Hint: Προσέξτε ώστε τα pids να αφορούν **μόνο** τις διεργασίες του Chromium!

Καταγραφή Στατιστικών Στοιχείων

Η συνάρτηση καταγραφής στατιστικών στοιχείων θα πρέπει ανά 0.5 δευτερόλεπτο να δημιουργεί μία λίστα με τις διεργασίες που ανήκουν στον Chromium και να υπολογίζει τα παρακάτω στατιστικά:

- πλήθος διεργασιών που ανήκουν στον Chromium
- μέγιστο πλήθος threads
- μέσο πλήθος threads ανά διεργασία
- συνολική κατανάλωση μνήμης από όλες τις διεργασίες (RSS - Resident Set Size) σε MB
- μέγιστη κατανάλωση μνήμης (RSS) ανά διεργασία σε MB
- μέσο πλήθος voluntary context switches ανά διεργασία
- μέσο πλήθος non-voluntary context switches ανά διεργασία

Τα στοιχεία που καταγράφονται θα πρέπει να αποθηκεύονται σε ένα αρχείο, όπου κάθε γραμμή αντιστοιχεί στα στατιστικά που καταγράφονται σε μία χρονική στιγμή. Κάθε γραμμή θα ξεκινά με τη χρονική στιγμή της καταγραφής (δευτερόλεπτα από την έναρξη του προγράμματος) και θα ακολουθούν τα υπόλοιπα στατιστικά χωρισμένα με κενό.

Ενδεικτικό παράδειγμα αρχείου καταγραφής στατιστικών:

```
0 1 1 1.00 4.64 4.64 5.00 5.00
0.5 7 37 7.71 249.83 78.92 63.57 84.85
1.0 12 38 5.58 414.65 92.13 84.25 88.50
1.5 9 38 7.77 389.35 103.88 169.66 137.88
2.0 10 38 7.90 481.10 130.35 205.90 166.70
```

Η πρώτη τιμή της κάθε γραμμής αντιστοιχεί στην χρονική στιγμή της καταγραφής, ενώ οι υπόλοιπες στα στατιστικά στοιχεία που καταγράφηκαν. Παρατηρήστε ότι δημιουργείται μία νέα γραμμή ανά 0.5 δευτερόλεπτο.

Η συνάρτηση θα πρέπει να τερματίζει μετά τον τερματισμό της συνάρτησης διαχείρισης διεργασιών (δηλαδή όταν δεν έχει απομείνει καμία διεργασία του Chromium στο σύστημα).

Δημιουργία Γραφικών Παραστάσεων

Για τη δημιουργία των γραφικών παραστάσεων μπορείτε να χρησιμοποιήσετε το πρόγραμμα gnuplot. Το gnuplot μπορεί να εκτελέσει scripts δημιουργίας γραφικών παραστάσεων. Συνεπώς, προτείνεται να προ-δημιουργήσετε ένα script του gnuplot το οποίο θα λαμβάνει το αρχείο με τα στατιστικά που παράγεται από τη συνάρτηση καταγραφής στατιστικών και θα δημιουργεί τις αντίστοιχες γραφικές παραστάσεις.

Ενδεικτικά παρέχεται ένα script που δημιουργεί μία γραφική παράσταση με δύο καμπύλες. Ο άξονας x αντιστοιχεί στον χρόνο καταγραφής (στήλη 1 του αρχείου) και ο y στις επόμενες δύο στήλες (στήλη 2 και 3).

```
set xlabel "time from start (s)"
set ylabel "count"
set autoscale
set term png
set output "test.png"
plot "data.out" using 1:2 with lines title "", "data.out" using 1:3 with
lines
```

Αν αποθηκεύσετε το παραπάνω αρχείο με όνομα `myscript.gr`, μπορείτε να το εκτελέσετε με την εντολή `gnuplot myscript.gr`. Το αρχείο με τα στατιστικά θα πρέπει να ονομάζεται `data.out` και να βρίσκεται στον ίδιο φάκελο με το `script`. Παρατηρήστε στο `script` την εντολή `plot "data.out" using 1:2 with lines title ""`. Η εντολή αυτή επιτρέπει τη δημιουργία μίας γραφικής παράστασης από τα δεδομένα του αρχείου `data.out`, χρησιμοποιώντας τη στήλη 1 ως τετμημένη και τη στήλη 2 ως τεταγμένη. Περισσότερες από μία γραφικές παραστάσεις μπορούν να δημιουργηθούν στο ίδιο σχήμα με τον τρόπο που παρουσιάζεται στο `script` (επανάληψη της εντολής στην ίδια γραμμή χωρίς το `plot` και χωρισμένη με κόμμα). Η γραφική παράσταση αποθηκεύεται στο αρχείο `test.png` σύμφωνα με την προτελευταία εντολή.

Παρατηρήσεις

- Μπορείτε να θεωρήσετε ότι όσες διεργασίες περιλαμβάνουν στην εντολή εκτέλεσής τους τη φράση `chromium-browser` ανήκουν στον Chromium, ακόμη και αν έχουν διαφορετικό όνομα εκτελέσιμου (π.χ. `/usr/lib/chromium-browser/chro`).
- Δεν μπορείτε να βασιστείτε στο `pid` μίας διεργασίας για να εντοπίσετε τη νεότερη διεργασία, αφού ενδέχεται μία παλαιότερη διεργασία να έχει αριθμητικά μεγαλύτερο `pid` από μία νεότερη. Πρέπει να ταξινομήσετε τις διεργασίες με βάση τον χρόνο εκτέλεσής τους στο σύστημα (execution time).
- Μπορείτε να αντλήσετε όλες τις πληροφορίες σχετικά με τα στατιστικά των διεργασιών από το `proc filesystem (procfs)` που παρέχει το Linux. Πιο συγκεκριμένα, στο αρχείο `/proc/pid/status` περιλαμβάνονται όλα τα στατιστικά που θα χρειαστείτε, σχετικά με τη διεργασία με αντίστοιχο `pid`.
- Η κλήση του Chromium με όρισμα ένα url οδηγεί σε δημιουργία ενός καινούριου tab στο υπάρχον παράθυρο. Παρόλα αυτά ενδέχεται, υπό ορισμένες συνθήκες, να δημιουργούνται ξεχωριστά παράθυρα. Αυτό δεν επιφέρει καμία αλλαγή στον τρόπο προσέγγισης της εργασίας και μπορείτε να αγνοήσετε αυτή την ενδεχομένη συμπεριφορά.
- Καθώς τερματίζετε τις διεργασίες του Chromium, ενδέχεται να επαναδημιουργούνται κάποιες απο αυτές. Αυτό δεν δημιουργεί πρόβλημα, αφού μετά από ορισμένο πλήθος προσπαθειών τερματισμού, θα τερματιστούν όλες οι διεργασίες χωρίς πρόβλημα.

Παραδοτέα και Διαδικαστικά

Θα πρέπει να παραδώσετε

α) τον πηγαίο κώδικα του/των `script/s` που θα αναπτύξετε, καθώς και

β) ένα ενδεικτικό αρχείο `url.in` με 10 urls. **Ο κώδικας του script θα πρέπει να είναι επαρκώς σχολιασμένος.**

Επίσης, θα πρέπει να εκτελέσετε δύο φορές το `script` και να παραδώσετε

γ) τα αρχεία με τις γραφικές παραστάσεις και

δ) τα αρχεία με τα στατιστικά.

Στην πρώτη εκτέλεση δεν θα έχετε καμία αλληλεπίδραση με τον browser, ενώ στη δεύτερη θα πρέπει να επισκεφτείτε μερικά επιπλέον sites χειροκίνητα κατά τη διάρκεια του `resting time` (χρόνος μεταξύ της δημιουργίας των διεργασιών και του τερματισμού των διεργασιών).

Διαδικαστικά

- Στην πρώτη γραμμή μετά τη δήλωση του bash interpreter, να υπάρχει σχόλιο με τα **ονοματεπώνυμα και τα ΑΕΜ** των φοιτητών/φοιτητριών που συμμετείχαν.
- Η εργασία θα γίνει σε ομάδες **3 ατόμων** και πρέπει να παραδοθεί μέσω email στους βοηθούς του μαθήματος έως και τη **Δευτέρα 23 Νοεμβρίου 2015 (23:59)**. Να στείλετε την εργασία **και στους δύο βοηθούς** (με την εισαγωγή του ενός στο πεδίο της κοινοποίησης).
- Τα mail σας να έχουν θέμα **"OS_2015_P1_"** και να ακολουθούνται από τα ΑΕΜ της ομάδας (πχ OS_2015_P1_2001_2002_2003).

Χρήσιμες Πηγές

1. Πληροφορίες σχετικά με το procfs: <http://en.wikipedia.org/wiki/Procfs> και <http://www.tldp.org/LDP/Linux-Filesystem-Hierarchy/html/proc.html>
2. Γρήγορη εισαγωγή στο gnuplot <http://people.duke.edu/~hpgavin/gnuplot.html>
3. Πληροφορίες σχετικά με τον τρόπο λειτουργίας του Chromium <http://www.chromium.org/developers/design-documents/multi-process-architecture> και <http://www.chromium.org/developers/design-documents/process-models>
4. bash <http://www.tldp.org/LDP/abs/html/>
5. grep http://tldp.org/LDP/Bash-Beginners-Guide/html/sect_04_02.html
6. awk <http://www.gnu.org/software/gawk/manual/gawk.html>
7. bc (bench calculator) <http://www.basicallytech.com/blog/?/archives/23-command-line-calculations-using-bc.html>