



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Άσκηση 1^η

Επιτήρηση χρήσης πόρων εφαρμογών με Linux Cgroups

Εισαγωγή

Ένα κέντρο δεδομένων (data center) προσφέρει υπηρεσίες φιλοξενίας εφαρμογών στην αγορά. Από τη μία, το κέντρο δεδομένων επιθυμεί να εκμεταλλεύεται ολόκληρη την υποδομή του χωρίς αδρανείς πόρους. Από την άλλη, οι πελάτες του απαιτούν οι πόροι που έχουν παραγγείλει να είναι διαθέσιμοι με αξιοπιστία. Εάν υπερφορτωθεί με εφαρμογές το κέντρο δεδομένων τότε δεν θα μπορεί να ανταποκριθεί στις απαιτήσεις όλων των εφαρμογών.

Για την άσκηση αυτή υποθέτουμε ότι η απαίτηση μιας εφαρμογής μετράται σε χιλιοστά της υπολογιστικής ισχύος του ενός επεξεργαστή. Ένας κόμβος με 16 επεξεργαστές μπορεί να καλύψει σύνολο 16000 χιλιοστών και άρα μπορεί να εξυπηρετήσει, για παράδειγμα, 8 εφαρμογές που απαιτούν εγγύηση 2000 χιλιοστών η κάθε μία.

Εάν όμως οι 8 εφαρμογές αυτές είναι *ανελαστικές* και δεν καταναλώνουν συνεχώς το σύνολο των 2000 χιλιοστών που έχει εξασφαλιστεί για αυτές, τότε ο κόμβος μένει αδρανής. Για να το αποφύγουμε αυτό θα μπορούσαμε να βάλουμε μόνο 7 ανελαστικές εφαρμογές των 2000 χιλιοστών και 10 εφαρμογές των 200 χιλιοστών αλλά *ελαστικές*. Οι 10 ελαστικές εφαρμογές θα εκμεταλλευτούν τους επεξεργαστές μόνο όταν οι 7 δεν τους χρησιμοποιούν.

Παράδειγμα ανελαστικής εφαρμογής είναι μια βάση δεδομένων της οποίας το φορτίο διακυμαίνεται και κοστίζει πολύ στους χρήστες της η όποια καθυστέρηση στις απαντήσεις της. Παράδειγμα ελαστικής εφαρμογής είναι η επεξεργασία μεγάλου όγκου γονιδιακού υλικού η οποία μπορεί να απαιτεί χιλιάδες ώρες επεξεργασίας αλλά δεν επηρεάζεται από τις αυξομειώσεις της επεξεργαστικής ισχύος που λαμβάνει, αρκεί σε βάθος χρόνου να λάβει όσο χρειάζεται.

Για να μπορεί το κέντρο δεδομένων να επιβάλλει την προτεραιότητα κάποιων εφαρμογών στα χιλιοστά υπολογιστικής ισχύος που αναλογούν χρειάζεται ένα κατάλληλο εργαλείο. Το σύστημα cgroup του Linux είναι ένα τέτοιο εργαλείο που μπορεί να επιβάλει ποσοτικούς περιορισμούς και διαμοιρασμό υπολογιστικών πόρων.

Σκοπός της άσκησης είναι η δημιουργία προγραμμάτων που με τη χρήση του συστήματος cgroup μπορεί να εξασφαλίσει την υπολογιστική ισχύ, ρυθμισμένη σε χιλιοστά του ενός επεξεργαστή όπως παραπάνω.

Ο Δαίμων cgmond

Ως βοήθημα δίδεται ο **cgmond**, πρόγραμμα (δαίμων με την έννοια του unix) που αναλαμβάνει να εκκινεί και να παρακολουθεί εφαρμογές, καθώς και να επιβάλλει τα ελάχιστα εξασφαλισμένα χιλιοστά επεξεργαστικής ισχύος όπως δίνονται στην εκκίνηση της εφαρμογής.

Ο **cgmond** δεν ερμηνεύει τις τιμές για τα ελάχιστα εξασφαλισμένα χιλιοστά κάθε εφαρμογής, ούτε εφαρμόζει ρυθμίσεις cgroup. Αντί αυτού καλεί δύο εξωτερικά εκτελέσιμα προγράμματα, **cgmon-policy** και **cgmon-limit**, αντίστοιχα.

Κάθε φορά που δημιουργείται ή τερματίζεται μια εφαρμογή, ο cgmond εκτελεί το cgmon-policy δίνοντάς του όλες τις εφαρμογές με τις ρυθμίσεις για τα χιλιοστά τους. Το cgmon-policy παράγει τις ρυθμίσεις για το σύστημα cgroup που πρέπει να εφαρμοστούν. Στη συνέχεια, το cgmon-limit καλείται για να εφαρμόσει τις ρυθμίσεις που παρήγαγε το cgmon-policy.

Σε περίπτωση που ζητηθεί η εκκίνηση μιας εφαρμογής με χιλιοστά περισσότερα από όσα μπορεί να εγγυηθεί το σύστημα τότε προειδοποιεί και απαιτεί επιβεβαίωση.

Ζητούμενα της Άσκησης

Ζητείται η συγγραφή κατάλληλων προγραμμάτων cgmon-policy και cgmon-limit και η ανεύρεση κατάλληλων εφαρμογών ώστε να ικανοποιηθούν και να επιδειχθούν τα παρακάτω:

1. Να συμφωνούν με τις προδιαγραφές που δίδονται στις αντίστοιχες ενότητες ώστε να μπορούν να χρησιμοποιηθούν από τον cgmond.
2. Θα πρέπει μετά την εφαρμογή από το cgmon-limit των ρυθμίσεων cgroup που παράγει το cgmon-policy να εξασφαλίζονται τα ελάχιστα χιλιοστά χρήσης κάθε εφαρμογής όπως έχουν δηλωθεί κατά την εκκίνησή της.

3. Το `cgmon-policy` θα πρέπει να βαθμοδοτεί κατάλληλα την είσοδό του ώστε ο `cgmond` να αποτρέπει την εκκίνηση εφαρμογής εάν πρόκειται να κινδυνεύσουν τα εγγυημένα χιλιοστά οποιασδήποτε εφαρμογής.
4. Εάν υπάρχουν ελαστικές και ανελαστικές εφαρμογές υπό επιτήρηση να εξασφαλίζεται ότι όλοι οι επεξεργαστές του κόμβου χρησιμοποιούνται και κανείς δεν μένει αδρανής.

Εκτελέσιμο `cgmon-policy`

Συνοπτικά, το εκτελέσιμο `cgmon-policy` καλείται να μετατρέψει το μέτρο των ελάχιστων εξασφαλισμένων χιλιοστών χρήσης ενός επεξεργαστή, μέγεθος με το οποίο συναλλάσσεται ο χρήστης, σε `cpu.shares`, μέγεθος με το οποίο ρυθμίζεται το σύστημα `cgroup` για το διαμοιρασμό επεξεργαστικής ισχύος σε εφαρμογές.

Το πρόγραμμα λαμβάνει είσοδο της μορφής:

```
policy:<application name>:cpu:<value>
```

...

Όπου κάθε γραμμή αντιστοιχεί με την εφαρμογή με το όνομα *application name*. Η τιμή *value* είναι ένας θετικός ακέραιος που αντιστοιχεί στα χιλιοστά του ενός επεξεργαστή τα οποία θέλουμε να εξασφαλίσουμε για την εφαρμογή.

Κάθε γραμμή τερματίζεται με έναν χαρακτήρα `'\n'` (ascii 0xa). Τα `policy` και `cpu` είναι σταθερές και εμφανίζονται ως έχουν.

Στην έξοδό του το πρόγραμμα παράγει αντίστοιχες γραμμές:

```
score:<float>
```

```
set_limit:<application name>:cpu.shares:<value>
```

...

Τα `score`, `set_limit`, `cpu.shares` είναι σταθερές και εμφανίζονται ως έχει.

Στην πρώτη γραμμή της εξόδου, η τιμή *float* είναι μια δεκαδική τιμή με υποδιαστολή η οποία εάν είναι θετική υποδηλώνει ότι τα ελάχιστα όρια χρήσης είναι εξασφαλισμένα για όλες τις εφαρμογές, ενώ αν είναι αρνητική τότε τα όρια δεν μπορούν να τηρηθούν.

Οι επόμενες γραμμές αφορούν κάθε μία την αντίστοιχη εφαρμογή με όνομα *application name*. Η

τιμή *value* είναι θετικός δεκαδικός ακέραιος και αντιστοιχεί στη ρύθμιση `cpu.shares` του αντίστοιχου `cgroup` της εφαρμογής.

Εκτελέσιμο `cgmon-limit`

Συνοπτικά, το εκτελέσιμο `cgmon-limit` αναλαμβάνει να εφαρμόσει τις ρυθμίσεις που παρήγαγε το `cgmon-policy`, καθώς και να συντηρεί την απεικόνιση των εφαρμογών σε ιεραρχίες `cgroup` κατ'εντολή του `cgmond`.

Το εκτελέσιμο δέχεται στην είσοδο γραμμές ενός εκ των ακόλουθων τύπων:

Δημιουργία `cgroup` για μια νέα εφαρμογή:

```
create:<monitor>:cpu:<application name>
```

Κατάργηση του `cgroup` μιας εφαρμογής που έχει τερματίσει

```
remove:<monitor>:cpu:<application name>
```

Εγγραφή μιας διεργασίας στο `cgroup` μιας εφαρμογής:

```
add:<monitor>:cpu:<application name>:<process id>
```

Ρύθμιση της τιμής `cpu.shares` για το `cgroup` μιας εφαρμογής:

```
set_limit:<monitor>:cpu:<application name>:cpu.shares:<value>
```

Το αναγνωριστικό *monitor* αντιστοιχεί στο `cgroup` του `cgmond` η οποία περιέχει όλα τα `cgroups` των εφαρμογών. Με αυτόν τον τρόπο απομονώνεται ο χώρος ονομάτων των εφαρμογών από το υπόλοιπο σύστημα για την αποφυγή συγκρούσεων (π.χ. κάποιος προσπαθεί να ονομάσει μια εφαρμογή με όνομα `cgroup` που υπάρχει ήδη στο σύστημα).

Υποσύστημα `cgroup/cpu`

Το υποσύστημα `cpu` ξεκινά αναθέτοντας όλη την υπολογιστική ισχύ του συνόλου των επεξεργαστών στην κορυφή της ιεραρχίας. Κάθε `cpu cgroup` έχει μία ρύθμιση `cpu.shares`. Σε κάθε επίπεδο αθροίζονται όλες οι τιμές των `cpu.shares` των `cgroups` που ανήκουν στο επίπεδο αυτό. Η συνολική υπολογιστική ισχύς διαθέσιμη στο επίπεδο αυτό χωρίζεται σε ισάριθμα και ισοδύναμα τμήματα. Κάθε `cgroup` λαμβάνει τόσα τμήματα όσα και η τιμή `cpu.shares` που έχει.

Πρόσβαση σε βοηθητική εικονική μηχανή

Για τις ανάγκες της άσκησης σας προσφέρεται μια εικονική μηχανή. Για να τη χρησιμοποιήσετε ακολουθήστε τα βήματα:

1. Επισκεφθείτε το <https://okeanos.grnet.gr> και χρησιμοποιείστε τον ακαδημαϊκό σας λογαριασμό για να εισέλθετε. Θα σας ζητηθεί να επιλέξετε το ίδρυμά σας πριν δώσετε τον κωδικό σας.
2. Επισκεφθείτε το dashboard στην πάνω δεξιά γωνία.
3. Πατήστε στο tab *projects*
4. Επιλέξτε *join* στο *oslab2016.cslab.ntua.gr*
5. Περιμένετε την έγκριση από το προσωπικό του εργαστηρίου (θα επιβεβαιώσουν τα στοιχεία σας). Θα λάβετε ειδοποίηση στο email σας μόλις ολοκληρωθεί η εγγραφή σας.
6. Πατήστε *Cyclades* στην ζώνη στην κορυφή της σελίδας
7. Επιλέξτε να δημιουργήσετε μια νέα εικονική μηχανή
8. Επιλέξτε εικόνα (image) *Debian base*. Επισημαίνεται ότι είναι απαραίτητη η δημιουργία **IPv4**, στο αντίστοιχο βήμα, προκειμένου να υποστηρίζονται οι κλασικοί τρόποι πρόσβασης στο εικονικό μηχάνημα. Καλό είναι επίσης, να δημιουργήσετε στο αντίστοιχο βήμα και ένα κλειδί ssh, διαφορετικά **σημειώστε το password μετά τη δημιουργία**.
9. Συνδεθείτε ως χρήστης root με ssh.
10. Εκτελέστε τα επόμενα βήματα (Παράρτημα Α) χειρωνακτικά ή εκτελώντας τα ως script. Μέσα σε αυτά προσέξτε τη διεύθυνση που θα βρείτε τον κώδικα του cgmond. Ο κωδικός που απαιτείται για τη λήψη του κώδικα από το αντίστοιχο repository είναι: **g79uAFVJ3v**

Παράρτημα Α

```
# cat > setup_env.sh << HEREDOC
#!/bin/bash

STEP () {
    read -p 'Continue (Y/n) ' c
}

set -v
cd

STEP Install git and vim
apt-get install git vim python python-setuptools

STEP Enable greek UTF-8 locale
sed -i -e 's/^# *(el_GR.UTF-8)\$/\1/g' /etc/locale.gen

STEP Enable greek UTF-8 typing in shell
touch ~/.bashrc
sed -i -e '/^export LC_CTYPE=/d' ~/.bashrc
echo 'export LC_CTYPE=el_GR.UTF-8' '>>' ~/.bashrc
# Now you have to re-login to enable it

STEP Generate key for git repository

mkdir -p -m 0700 ~/.ssh
if ! [ -e ~/.ssh/cgmon_rsa_id ]; then
    ssh-keygen -b 4096 -t rsa -f ~/.ssh/cgmon_rsa_id
fi

touch ~/.ssh/config
set -i -e '/^IdentityFile ~/.ssh/cgmon_rsa_id$/d' ~/.ssh/config
echo 'IdentityFile ~/.ssh/cgmon_rsa_id' '>>' ~/.ssh/config

# Send the following public key to your course assistants
echo -e '\n\n\n\n\n'; \
echo '----- cut here -----'; \
echo -n 'command="git-shell -c \"$SSH_ORIGINAL_COMMAND\""; \
echo -n 'no-port-forwarding,no-agent-forwarding ssh-rsa '; \
cat ~/.ssh/cgmon_rsa_id.pub; \
echo '----- cut here -----'; \
echo -e '\n\n\n\n\n'

STEP Get cgmon source and start working

mv -f cgmon cgmon-$(date -Isec)
git clone git+ssh://cslab@snf-698511.vm.okeanos.grnet.gr:/git/cgmon
cd cgmon
python setup.py install
HEREDOC
# chmod +x ./setup_env.sh
# ./setup_env.sh
```