

LabSO 2022

Laboratorio Sistemi Operativi - A.A. 2021-2022

dr. Andrea Naimoli

Informatica [0514G] - 2 - Scienze e Tecnologie Informatiche (LT)
andrea.naimoli@unitn.it

dr. Michele Grisafi

Ingegneria informatica, delle comunicazioni ed elettronica (LT)
michele.grisafi@unitn.it

Presentazione



Requisiti

- Uso basilare Linux da shell
- Utilizzo editor per codice
- Conoscenza fondamentali file-system (file, cartelle, permessi) e processi
- Teoria di *Sistemi Operativi*

Obiettivo

Dimestichezza nell'ideazione e realizzazione di applicazioni binarie complete o almeno singole componenti impostate con il linguaggio C utilizzabili su shell bash in un sistema Linux debian/ubuntu che implementino metodi di IPC.

Note

Terminologia e contenuti sono contestualizzati al corso: eventuali approssimazioni o semplificazioni di contenuti e significati sono legati a tale scopo (e quindi i concetti non sono da considerarsi esaustivi in senso assoluto). Alcuni concetti sono solo toccati marginalmente e/o descritti nel momento che si incontrano

Elementi

- terminal & bash
 - docker
 - gcc
 - make
 - linguaggio C:
 - fondamenti
 - costrutti
 - system-calls
 - IPC
-

Corso

Presentazione contenuti principalmente tramite “slides” commentate direttamente con esposizione concetti.

- Punti fondamentali
- Semplici schematizzazioni / grafici
- Esemplicazioni di schermate (screenshot)

Possono essere mostrate semplici interazioni tramite screen-sharing.

È richiesto disporre di un sistema che consenta di utilizzare le applicazioni presentate per poter fare prove in modo interattivo.

Esame

Possibili quesiti sulla teoria di laboratorio ed esercizi di programmazione in C e bash.

Voto che fa media al 50% con la teoria. Voto minimo 18 per poter passare il corso.

Terminal & bash

—

Terminal

Il terminale (o *terminal*) è l'ambiente testuale di interazione con il sistema operativo.

Tipicamente è utilizzato come applicazione all'interno dell'ambiente grafico ed è possibile avviarne più istanze, pur essendo anche disponibile direttamente all'avvio (in questo caso normalmente in più istanze accessibili con la combinazione CTRL+ALT+Fx).

L'interazione avviene utilizzando un'applicazione specifica in esecuzione al suo interno comunemente detta SHELL.

Bash

SHELL di riferimento per l'interazione con il sistema.

Propone un prompt per l'immissione diretta di comandi da tastiera e fornisce un feedback testuale. È anche possibile eseguire sequenze di comandi preorganizzate contenute in file testuali (*script* o *batch*). A seconda della modalità (diretta/script) alcuni comandi possono avere senso o meno o comportarsi in modo particolare.

L'insieme dei comandi e delle regole di composizione costituisce un linguaggio di programmazione orientato allo scripting.

Bash

Esistono numerose shell. Bash è una delle più utilizzate e molte sono comunque simili tra loro, ma hanno sempre qualche differenza (e anche comandi analoghi possono avere opzioni o comportamenti non identici).

Tipicamente - almeno in sessioni non grafiche - al login un utente ha associata una shell particolare.

Una shell può essere o meno di login e può essere o meno interattiva

SHELL	login no	login sì
interattiva no	non-login non-interactive	login non-interactive
interattiva sì	non-login interactive	login interactive

POSIX

Portable Operating System Interface for Unix: è una famiglia di standard IEEE. Nel caso delle shell in particolare definisce una serie di regole e di comportamenti che possono favorire la portabilità (che però dipende anche da altri fattori del sistema!).

La shell *bash* soddisfa molti requisiti, ma presenta anche alcune differenze ed “estensioni” per agevolare almeno in parte la programmazione. (v. costrutti per confronti logici)

Comandi interattivi

La shell attende un input dall'utente e al completamento (conferma con INVIO) lo elabora.

Per indicare l'attesa mostra all'utente un PROMPT (può essere modificato).

Fondamentalmente si individuano 3 canali:

- input (tipicamente la tastiera), canale 0, detto `stdin`
- output standard (tipicamente il video), canale 1, detto `stdout`
- output errori (tipicamente il video), canale 2, detto `stderr`

(il numero del canale è quello standard se non si varia)

Struttura generale comandi

Solitamente un comando è identificato da una parola chiave cui possono seguire uno o più “argomenti” opzionali o obbligatori, accompagnati da un valore di riferimento o meno (in questo caso hanno valore di “flag”) e di tipo posizionale o nominale. A volte sono ripetibili.

```
ls -alh /tmp
```

gli argomenti nominali sono indicati con un trattino cui segue una voce (stringa alfanumerica) e talvolta presentano una doppia modalità di riferimento: breve (tipicamente voce di un singolo carattere) e lunga (tipicamente un termine mnemonico)

app -h	app --help
--------	------------

Termini

- l'esecuzione dei comandi avviene “per riga” (in modo diretto quella che si immette fino a INVIO) (*)
- un “termine” (istruzione, argomento, opzione, etc.) è solitamente una stringa alfanumerica senza spazi
- spaziature multiple sono solitamente valide come singole e non sono significative se non per separare termini
- è possibile solitamente usare gli apici singoli o doppi per forzare una sequenza come termine singolo
- gli spazi iniziali e finali di una riga collassano
- le righe vuote sono ignorate

(*) per i file batch (gli “script”) l'argomento sarà trattato in modo approfondito in un blocco successivo

Commenti

È anche possibile utilizzare dei **commenti** da passare alla shell. L'unico modo formale è l'utilizzo del carattere # per cui esso e tutto ciò che segue fino al termine della riga è considerato un commento ed è sostanzialmente ignorato.

Alcuni comandi fondamentali

- clear
- pwd
- ls
- cd
- wc
- date
- cat
- echo
- alias/unalias
- test
- read
- file
- chown
- chmod
- cp/mv
- **help**
- type
- grep
- truncate
- function

Canali in/out

I comandi possono essere “builtins” (funzionalità intrinseche dell’applicazione shell utilizzata) o “esterni” (applicazioni eseguibili che risiedono su disco).

Ogni comando lavora su un insieme di canali (*) e in particolare uno per l’output standard e uno per l’output degli errori (normalmente entrambi vanno a video)

```
ls # mostra output della cartella corrente
```

```
ls not-existent-item #mostra un messaggio d'errore
```

(*) File Descriptors: l’argomento sarà trattato in modo approfondito in un blocco successivo

Redirezionamento di base

I canali possono essere redirezionati (anche in cascata):

- `ls 1>/tmp/out.txt 2>/tmp/err.txt`
- `ls not-existent-item 1>/tmp/all.txt 2>&1`

Redirezionamento di base

- `'<'`: `command<file.txt` invia l'input al comando (`file.txt` read-only):
`mail -s "Subject" rcpt < content.txt` (anziché interattivo)
- `'<>'`: come sopra ma `file.txt` è aperto in read-write (raramente usato)
- `source>target`: `command 1>out.txt 2>err.txt`
redireziona *source* su *target*:
 - *source* può essere sottinteso (vale 1)
 - *target* può essere un canale (si indica con `&n`, ad esempio `&2`)
- `'>|'`: si comporta come `>` ma forza la sovrascrittura anche se bloccata nelle configurazioni

Redirezionamento di base

- `>>` : si comporta come `>` ma opera un *append* se la destinazione esiste
- `<<` `<<<` : permettono di utilizzare testi come here-doc

Esistono molte varianti e possibilità di combinazione dei vari operatori di redirezionamento.

Un caso molto utilizzato è la “soppressione” dell’output (utile per gestire solo side-effects, come ad esempio il codice di ritorno), ad esempio:

`type command 1>/dev/null 2>&1` (per sapere se *command* esiste)

Ambiente e variabili

- La shell può utilizzare delle variabili per memorizzare e recuperare valori
- I valori sono generalmente trattati come stringhe o interi: sono presenti anche semplici *array* (vettori di elementi)
- Il formato del nome è del tipo `^[_[:alpha:]][_[:alnum:]]*$`
- Per set (impostare) / get (accedere) al valore di una variabile:
 - Il set si effettua con `[export] variabile=valore (*)`
 - Il get si effettua con `$variabile` o `${variabile}` (sostituzione letterale)

(*) lo “scope” è generalmente quello del processo attuale: antepoendo “export” si rende disponibile anche agli eventuali processi figli

Ambiente e variabili

- La shell opera in un ambiente in cui ci sono alcuni riferimenti impostabili e utilizzabili attraverso l'uso delle cosiddette “variabili d'ambiente” con cui si intendono generalmente quelle con un significato particolare per la shell stessa (v. esempi di seguito)
- Tra le variabili d'ambiente più comuni troviamo ad esempio:
SHELL, PATH, TERM, PWD, PS1, HOME
- Essendo variabili a tutti gli effetti si impostano ed usano come le altre.

Variabili di sistema

Alcune variabili sono impostate e/o utilizzate direttamente dal sistema in casi particolari. Se ne vedranno alcune caratteristiche degli script/batch, ma intanto in modalità diretta se ne usano già diverse:

- SHELL : contiene il riferimento alla shell corrente (path completo)
- PATH : contiene i percorsi in ordine di priorità in cui sono cercati i comandi, separati da “:”
- TERM : contiene il tipo di terminale corrente
- PWD : contiene la cartella corrente
- PS1 : contiene il prompt e si possono usare marcatori speciali
- HOME : contiene la cartella principale dell’utente corrente

Esecuzione comandi e \$PATH

- Quando si immette un comando (o una sequenza di comandi) la shell analizza quanto inserito (*parsing*) e individua le parti che hanno la posizione di comandi da eseguire: se sono interni ne esegue le funzionalità direttamente altrimenti cerca di individuare un corrispondente file eseguibile: questo è normalmente cercato nel file-system solo e soltanto nei percorsi definiti dalla variabile PATH a meno che non sia specificato un percorso (relativo o assoluto) nel qual caso viene utilizzato esso direttamente.

Dall'ultima osservazione discende che per un'azione abbastanza comune come lanciare un file eseguibile (non “installato”) nella cartella corrente occorre qualcosa come: `./nomefile`

Array

- Definizione: `lista=("a" 1 "b" 2 "c" 3)`
- Output completo: `${lista[@]}`
- Accesso singolo: `${lista[x]}` (0-based)
- Lista indici: `${!lista[@]}`
- Dimensione: `${#lista[@]}`
- Set elemento: `lista[x]=value`
- Append: `lista+=(value)`
- Sub array: `${lista[@]:s:n}` (from index s , length n)

Variabili \$\$ e \$?

Le variabili \$\$ e \$? non possono essere impostate manualmente (la stessa sintassi lo impedisce dato che i nomi sarebbero \$ e ? non utilizzabili normalmente):

- \$\$: contiene il PID del processo attuale (*)
- \$? : contiene il codice di ritorno dell'ultimo comando eseguito

(*) sarà approfondito il concetto di PID in sezioni successive

Esecuzione comandi e parsing

- La riga dei comandi è elaborata con una serie di azioni “in sequenza” e poi rielaborata eventualmente più volte.
- Tra le azioni vi sono:
 - Sostituzioni speciali della shell (es. “!” per accedere alla “history” dei comandi)
 - Sostituzione variabili
 - Elaborazione subshell
- Sono svolte con un ordine di priorità e poi l’intera riga è rielaborata (v. Subshell)

Concatenazione comandi

È possibile concatenare più comandi in un'unica riga in vari modi con effetti differenti:

- `comando1 ; comando2` concatenazione semplice: esecuzione in sequenza
- `comando1 && comando2` concatenazione logica “and”: l'esecuzione procede solo se il comando precedente non fallisce (codice ritorno zero)
- `comando1 || comando2` concatenazione logica “or”: l'esecuzione procede solo se il comando precedente fallisce (codice ritorno NON zero)
- `comando1 | comando2` concatenazione con piping (v. prox)

Operatori di piping (“pipe”): | e |&

La concatenazione con gli operatori di piping “cattura” l’output di un comando e lo passa in input al successivo:

- `ls | wc -l` : cattura solo stdout
- `ls |& wc -l` : cattura stdout e stderr

NOTA. il comando `ls` ha un comportamento atipico: il suo output di base è differente a seconda che il comando sia diretto al terminale o a un piping (*)

(*) internamente sfrutta `isatty(STDOUT_FILENO)` (si approfondirà in seguito)

Subshell

- È possibile avviare una subshell, ossia un sotto-ambiente in vari modi, in particolare raggruppando i comandi tra parentesi tonde:
`(...comandi...)`
- Spesso si usa “catturare” l’output standard (stdout) della sequenza che viene così sostituito letteralmente e rielaborato e si può fare in due modi:

`$ (...comandi...)` oppure `` ...comandi... ``

Nota: attenzione al fatto che le parentesi tonde sono utilizzate anche per definire *array*

Esempio alcuni comandi e sostituzione subshell

```
echo "/tmp" > /tmp/tmp.txt ; ls $(cat /tmp/tmp.txt)
```

i comandi sono eseguiti rispettando la sequenza:

- `echo "/tmp" > /tmp/tmp.txt` crea un file temporaneo con `"/tmp"`
- `ls $(cat /tmp/tmp.txt)` è prima eseguita la subshell:
 - `cat /tmp/tmp.txt` genera in stdout `"/tmp"` e poi con sostituzione:
 - `ls /tmp` mostra il contenuto della cartella `/tmp`