Parallel Processing
Lecture 3, Lab 1

Dr. Guy Tel-Zur

~~Version 26/11/2014~~
~~Version 04/04/2015~~
~~Version 13/04/2015~~
~~Version 10/11/2015~~
~~Version 28/03/2016~~
~~Version 27/03/2017~~
~~Version 30/10/2018~~
~~Version 10/03/2019~~
~~Version 01/11/2020~~
Version 12/11/2022

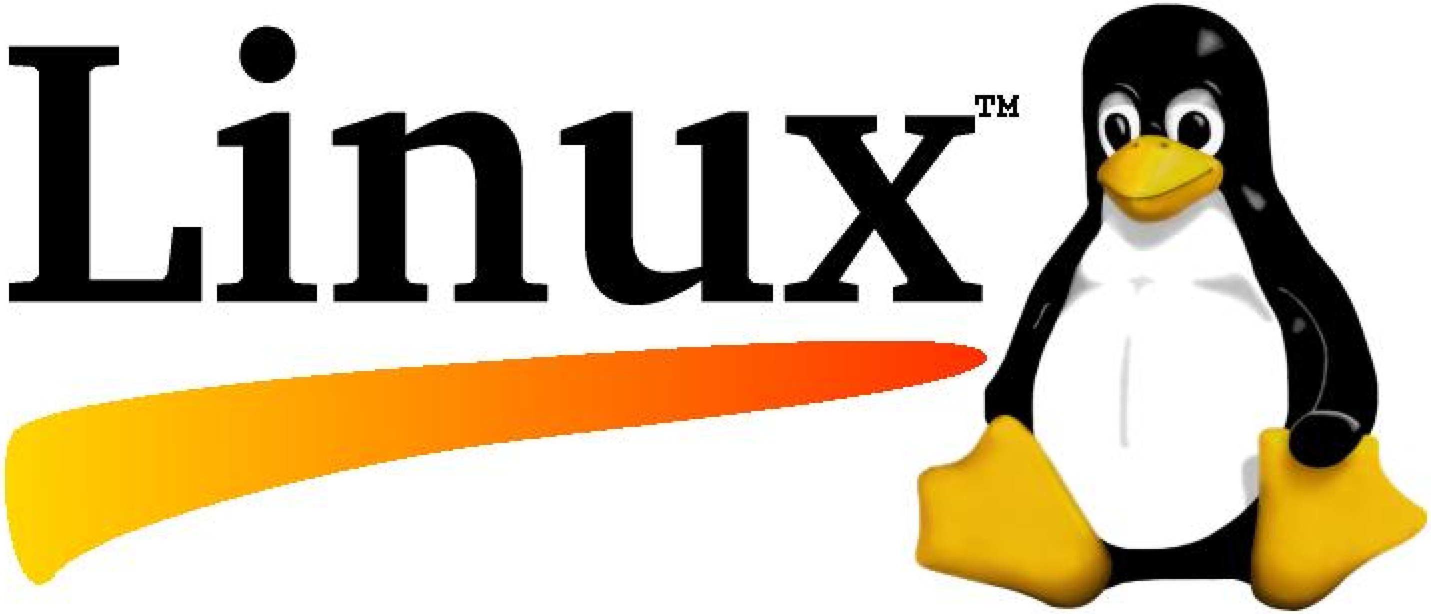
מטרות השיעור

- התחברות ועבודה במחשב הוירטואלי ובקלאסטר המקבילי.
- תרגול ביצוע משימות בסיסיות תחת מערכת ההפעלה Linux
- הרצת תכניות מקביליות בסיסיות המשתמשות ב-MPI
- שימוש ב- Profilers: jumpshot, scalasca

יעדים

- סביבת העבודה: מחשבים וירטואליים וקלאסטר מחלקתי (התחברות מרחוק)
- יישור קו בנושא לינוקס
- הכרות ראשונית עם MPI
  - פיתוח קוד: כתיבה ועריכת קוד בעורך
  - קימפול
  - ניפוי שגיאות
  - הרצה
  - ניתוח התוצאה

Linux



Tutorial: http://www.ee.surrey.ac.uk/Teaching/Unix/

Basic Linux Commands – 1/5

Enter the system	login: *username* password: *passwd*
Exit the system	*exit*
Text editors	*pico*, *vi*, *(x)Emacs*, *gedit*… or Edit on Windows then transfer file using *ftp*
C Compiler	*gcc –o file file.c*

Basic Linux Commands – 2/5

	DOS	Linux
See files	dir	*ls* *ls -l*
Copy files	copy	*cp*
Erase files	del	*rm*

Basic Linux Commands – 3/5

	DOS	Linux
Make directory	mkdir	*mkdir*
Remove directory	rmdir	*rmdir*
More/Rename	rename	*mv*
OS version	ver	*uname -a*

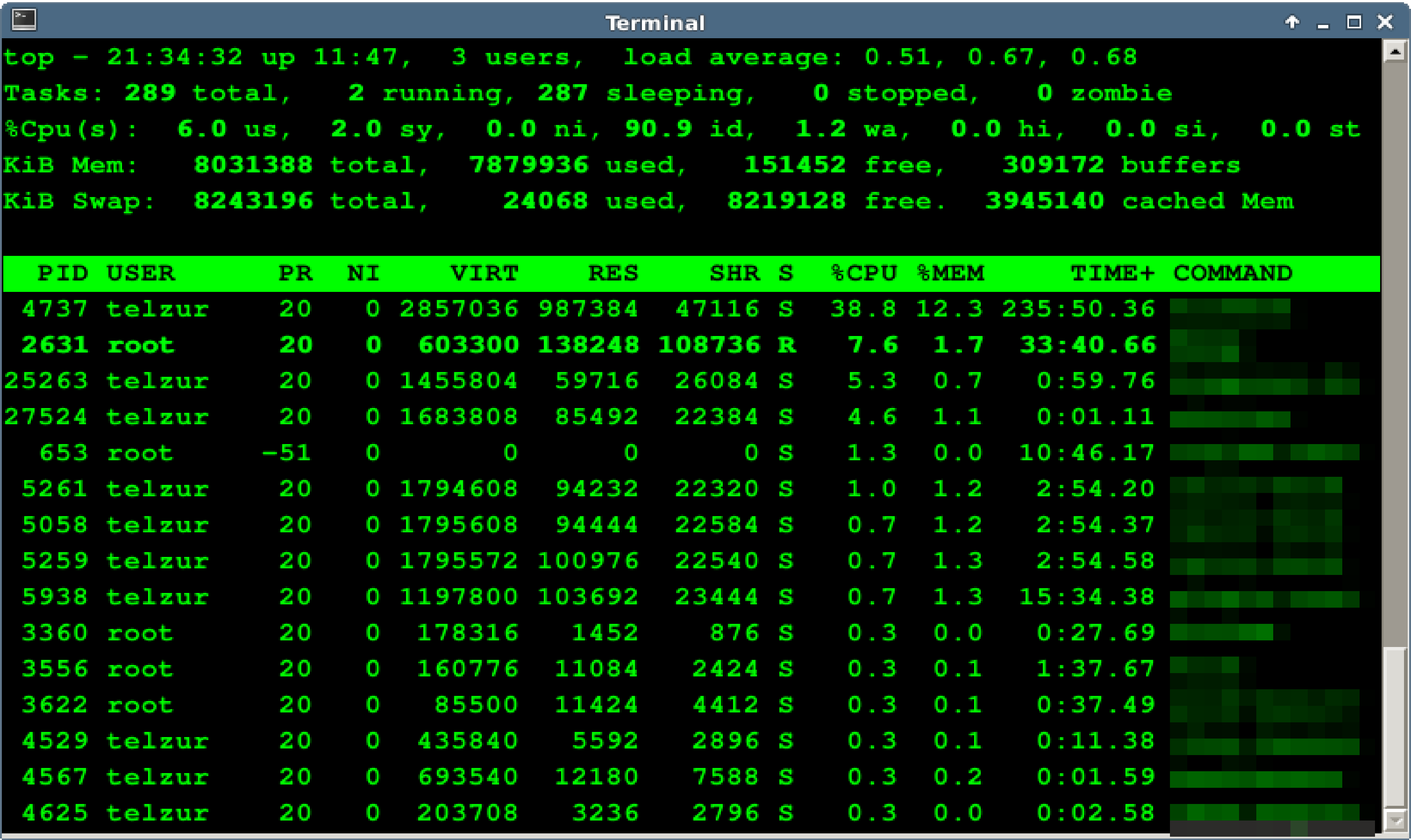
Basic Linux Commands – 4/5

- Getting help: *man* topic
- Look at the contents of a file: *cat*, *more*,*head* and *tail*
- Quit from *man* or *more*: *q*
- Where am I? *pwd*
- Clear the screen: *clear*

Basic Linux Commands – 5/5

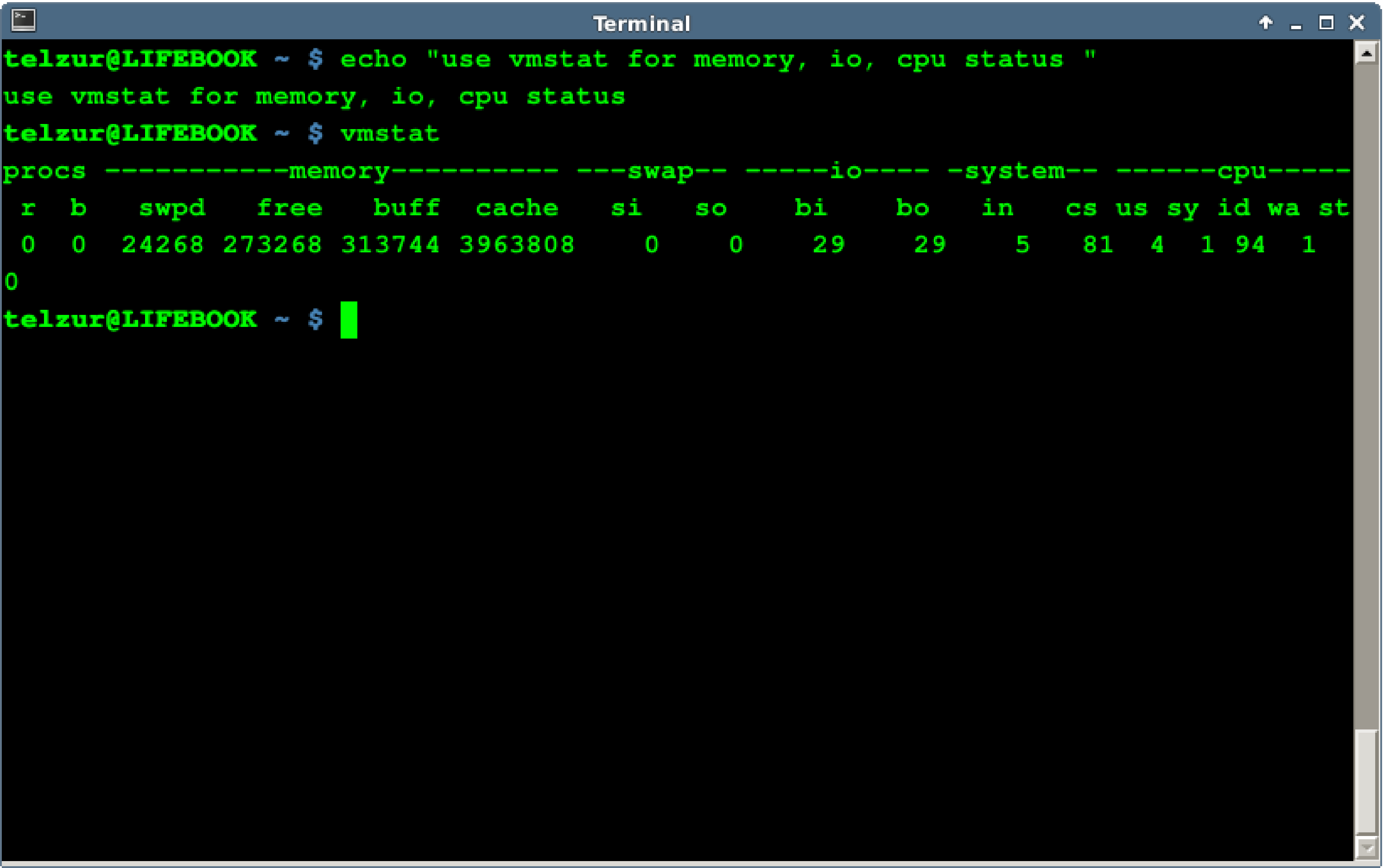
- Redirection: >, >>
- Pipe: |
- telnet
- ftp
- ping
- chmod
- chown

What is now running on my computer? `top`



```
top - 21:34:32 up 11:47,  3 users,  load average: 0.51, 0.67, 0.68
Tasks: 289 total,   2 running, 287 sleeping,   0 stopped,   0 zombie
%Cpu(s):  6.0 us,  2.0 sy,  0.0 ni, 90.9 id,  1.2 wa,  0.0 hi,  0.0 si,  0.0 st
KiB Mem:   8031388 total,  7879936 used,   151452 free,   309172 buffers
KiB Swap:  8243196 total,    24068 used,  8219128 free.  3945140 cached Mem

  PID USER      PR  NI    VIRT    RES    SHR S  %CPU %MEM     TIME+ COMMAND
 4737 telzur    20   0 2857036 987384  47116 S  38.8 12.3 235:50.36 [redacted]
 2631 root      20   0  603300 138248 108736 R   7.6  1.7  33:40.66
25263 telzur    20   0 1455804  59716  26084 S   5.3  0.7   0:59.76
27524 telzur    20   0 1683808  85492  22384 S   4.6  1.1   0:01.11
  653 root     -51   0       0      0      0 S   1.3  0.0  10:46.17
 5261 telzur    20   0 1794608  94232  22320 S   1.0  1.2   2:54.20
 5058 telzur    20   0 1795608  94444  22584 S   0.7  1.2   2:54.37
 5259 telzur    20   0 1795572 100976  22540 S   0.7  1.3   2:54.58
 5938 telzur    20   0 1197800 103692  23444 S   0.7  1.3  15:34.38
 3360 root      20   0  178316   1452    876 S   0.3  0.0   0:27.69
 3556 root      20   0  160776  11084   2424 S   0.3  0.1   1:37.67
 3622 root      20   0   85500  11424   4412 S   0.3  0.1   0:37.49
 4529 telzur    20   0  435840   5592   2896 S   0.3  0.1   0:11.38
 4567 telzur    20   0  693540  12180   7588 S   0.3  0.2   0:01.59
 4625 telzur    20   0  203708   3236   2796 S   0.3  0.0   0:02.58
```

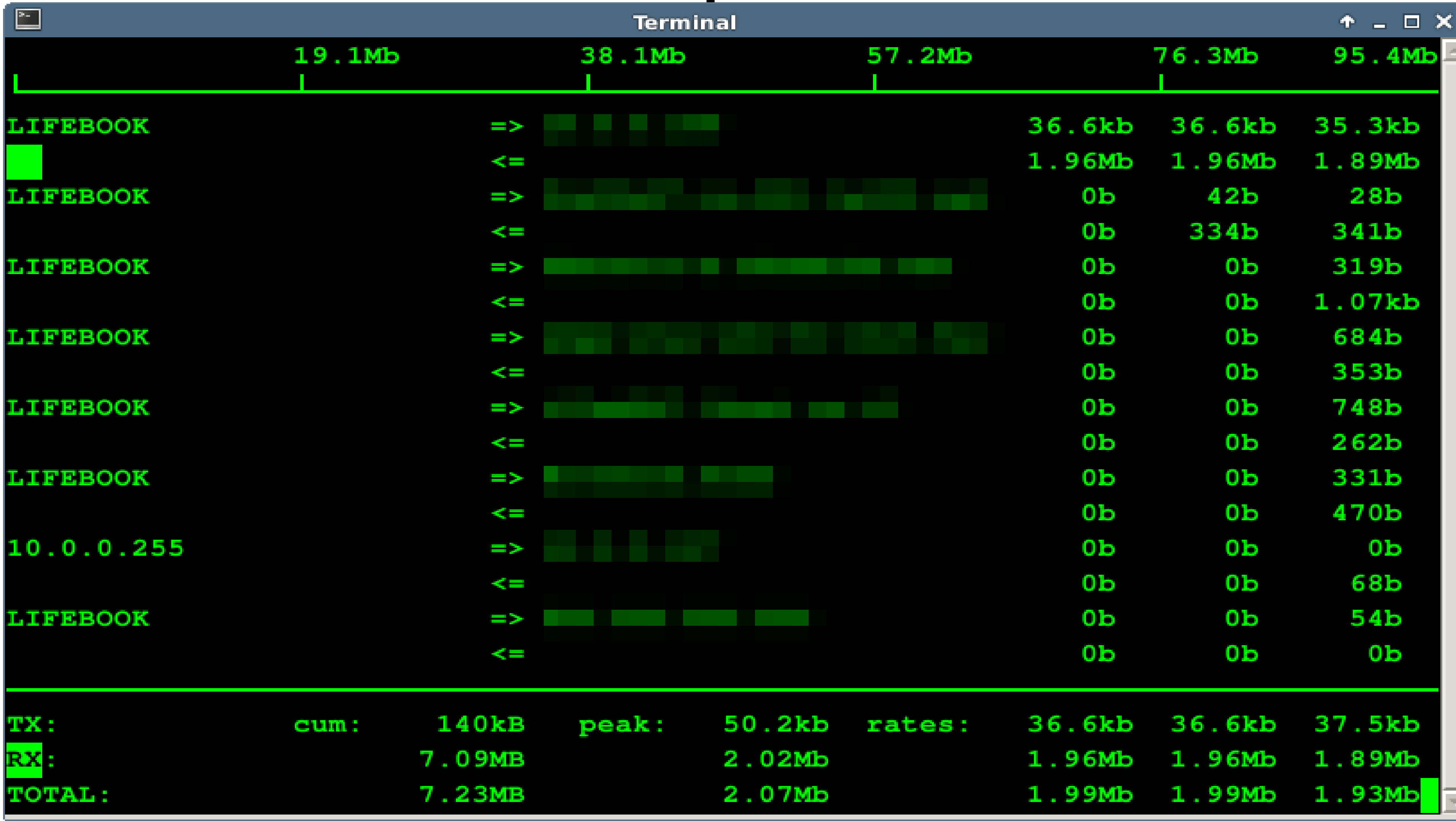
vmstat



```
telzur@LIFEBOOK ~ $ echo "use vmstat for memory, io, cpu status "
use vmstat for memory, io, cpu status
telzur@LIFEBOOK ~ $ vmstat
procs -----------memory---------- ---swap-- -----io---- -system-- ------cpu-----
 r  b   swpd   free   buff  cache   si   so    bi    bo   in   cs us sy id wa st
 0  0  24268 273268 313744 3963808    0    0    29    29    5   81  4  1 94  1
0
telzur@LIFEBOOK ~ $ 
```

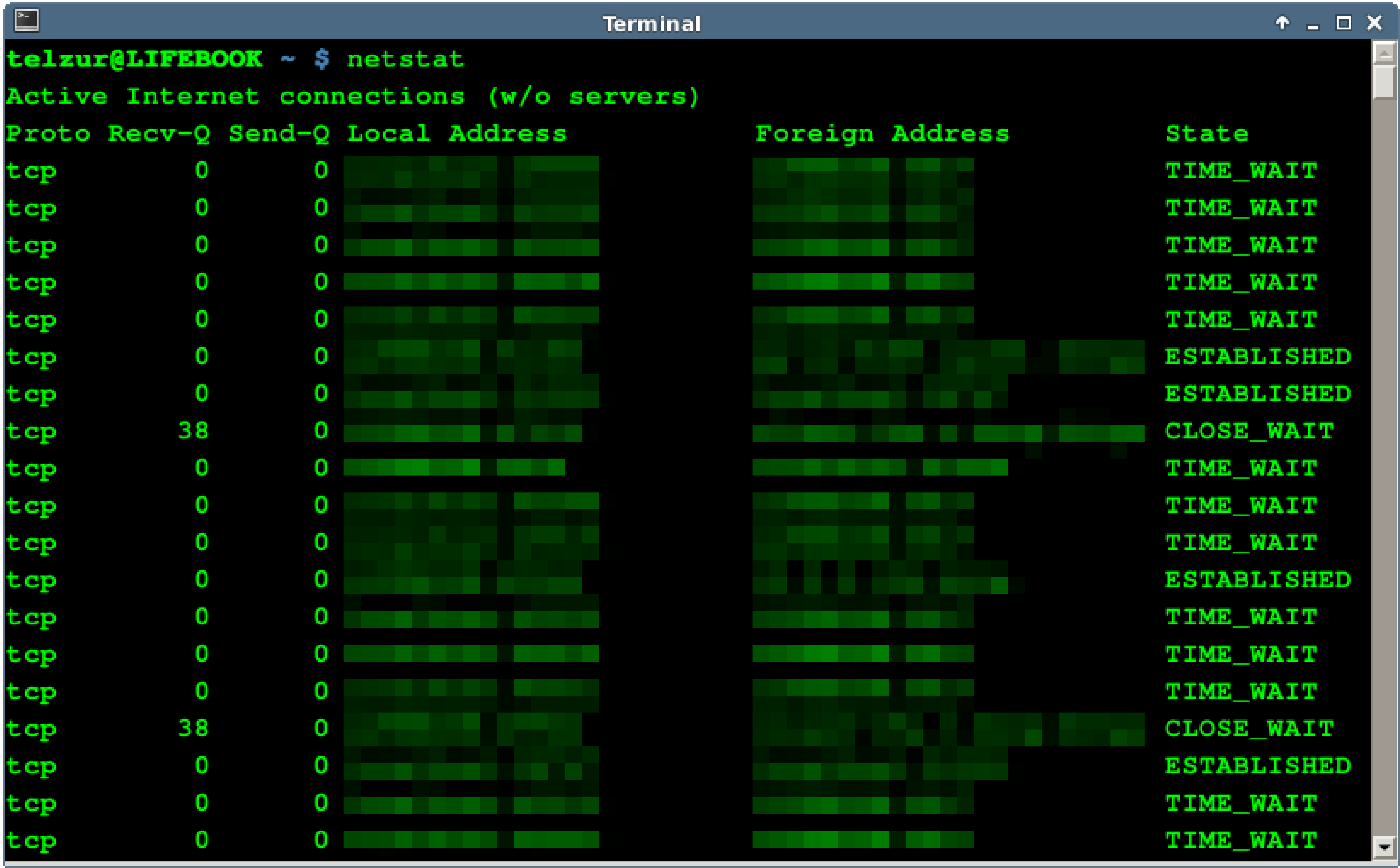
What is now running on my network interface card?

```
sudo iftop -i wlan0
```



```
                19.1Mb          38.1Mb          57.2Mb          76.3Mb     95.4Mb
└───────────────┴───────────────┴───────────────┴───────────────┴──────────
LIFEBOOK                   => [redacted]                 36.6kb  36.6kb  35.3kb
                           <=                            1.96Mb  1.96Mb  1.89Mb
LIFEBOOK                   => [redacted]                     0b     42b     28b
                           <=                                0b    334b    341b
LIFEBOOK                   => [redacted]                     0b      0b    319b
                           <=                                0b      0b  1.07kb
LIFEBOOK                   => [redacted]                     0b      0b    684b
                           <=                                0b      0b    353b
LIFEBOOK                   => [redacted]                     0b      0b    748b
                           <=                                0b      0b    262b
LIFEBOOK                   => [redacted]                     0b      0b    331b
                           <=                                0b      0b    470b
10.0.0.255                 => [redacted]                     0b      0b      0b
                           <=                                0b      0b     68b
LIFEBOOK                   => [redacted]                     0b      0b     54b
                           <=                                0b      0b      0b
──────────────────────────────────────────────────────────────────────────────
TX:             cum:   140kB   peak:   50.2kb  rates:   36.6kb  36.6kb  37.5kb
RX:                   7.09MB           2.02Mb           1.96Mb  1.96Mb  1.89Mb
TOTAL:                7.23MB           2.07Mb           1.99Mb  1.99Mb  1.93Mb
```

netstat

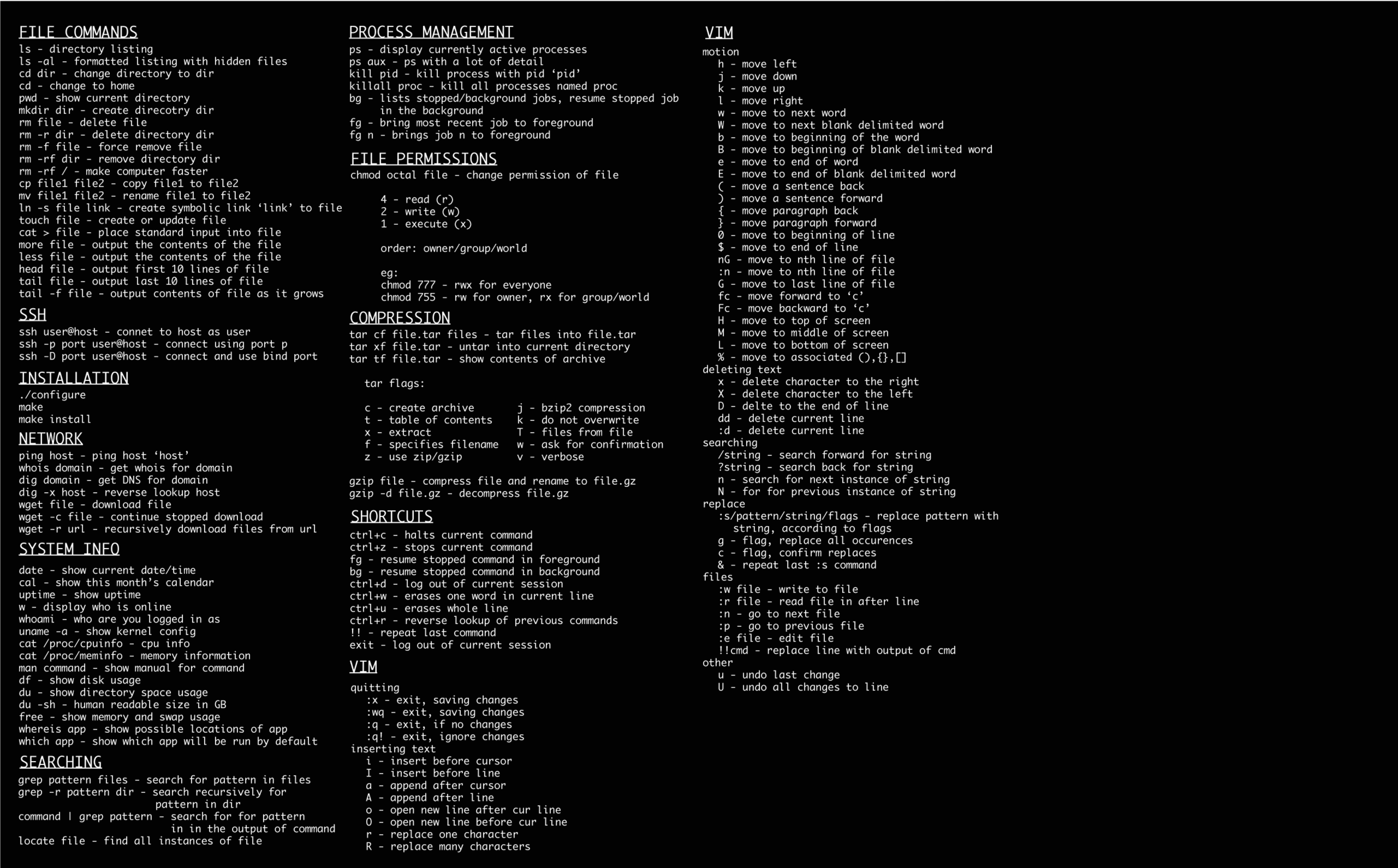


```
telzur@LIFEBOOK ~ $ netstat
Active Internet connections (w/o servers)
Proto Recv-Q Send-Q Local Address           Foreign Address         State
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              ESTABLISHED
tcp        0      0 [redacted]              [redacted]              ESTABLISHED
tcp       38      0 [redacted]              [redacted]              CLOSE_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              ESTABLISHED
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp       38      0 [redacted]              [redacted]              CLOSE_WAIT
tcp        0      0 [redacted]              [redacted]              ESTABLISHED
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
tcp        0      0 [redacted]              [redacted]              TIME_WAIT
```

Linux FAQ

http://www.ctssn.com/linux/linuxfaq.html

http://www.linuxstall.com/wp-content/uploads/2012/01/linux-command-line-cheat-sheet.png



FILE COMMANDS

```
ls - directory listing
ls -al - formatted listing with hidden files
cd dir - change directory to dir
cd - change to home
pwd - show current directory
mkdir dir - create direcotry dir
rm file - delete file
rm -r dir - delete directory dir
rm -f file - force remove file
rm -rf dir - remove directory dir
rm -rf / - make computer faster
cp file1 file2 - copy file1 to file2
mv file1 file2 - rename file1 to file2
ln -s file link - create symbolic link ‘link’ to file
touch file - create or update file
cat > file - place standard input into file
more file - output the contents of the file
less file - output the contents of the file
head file - output first 10 lines of file
tail file - output last 10 lines of file
tail -f file - output contents of file as it grows
```

SSH

```
ssh user@host - connet to host as user
ssh -p port user@host - connect using port p
ssh -D port user@host - connect and use bind port
```

INSTALLATION

```
./configure
make
make install
```

NETWORK

```
ping host - ping host ‘host’
whois domain - get whois for domain
dig domain - get DNS for domain
dig -x host - reverse lookup host
wget file - download file
wget -c file - continue stopped download
wget -r url - recursively download files from url
```

SYSTEM INFO

```
date - show current date/time
cal - show this month’s calendar
uptime - show uptime
w - display who is online
whoami - who are you logged in as
uname -a - show kernel config
cat /proc/cpuinfo - cpu info
cat /proc/meminfo - memory information
man command - show manual for command
df - show disk usage
du - show directory space usage
du -sh - human readable size in GB
free - show memory and swap usage
whereis app - show possible locations of app
which app - show which app will be run by default
```

SEARCHING

```
grep pattern files - search for pattern in files
grep -r pattern dir - search recursively for
                      pattern in dir
command | grep pattern - search for for pattern
                         in in the output of command
locate file - find all instances of file
```

PROCESS MANAGEMENT

```
ps - display currently active processes
ps aux - ps with a lot of detail
kill pid - kill process with pid ‘pid’
killall proc - kill all processes named proc
bg - lists stopped/background jobs, resume stopped job
     in the background
fg - bring most recent job to foreground
fg n - brings job n to foreground
```

FILE PERMISSIONS

```
chmod octal file - change permission of file

     4 - read (r)
     2 - write (w)
     1 - execute (x)

     order: owner/group/world

     eg:
     chmod 777 - rwx for everyone
     chmod 755 - rw for owner, rx for group/world
```

COMPRESSION

```
tar cf file.tar files - tar files into file.tar
tar xf file.tar - untar into current directory
tar tf file.tar - show contents of archive

  tar flags:

  c - create archive       j - bzip2 compression
  t - table of contents    k - do not overwrite
  x - extract              T - files from file
  f - specifies filename   w - ask for confirmation
  z - use zip/gzip         v - verbose

gzip file - compress file and rename to file.gz
gzip -d file.gz - decompress file.gz
```

SHORTCUTS

```
ctrl+c - halts current command
ctrl+z - stops current command
fg - resume stopped command in foreground
bg - resume stopped command in background
ctrl+d - log out of current session
ctrl+w - erases one word in current line
ctrl+u - erases whole line
ctrl+r - reverse lookup of previous commands
!! - repeat last command
exit - log out of current session
```

VIM

```
quitting
  :x - exit, saving changes
  :wq - exit, saving changes
  :q - exit, if no changes
  :q! - exit, ignore changes
inserting text
  i - insert before cursor
  I - insert before line
  a - append after cursor
  A - append after line
  o - open new line after cur line
  O - open new line before cur line
  r - replace one character
  R - replace many characters
```

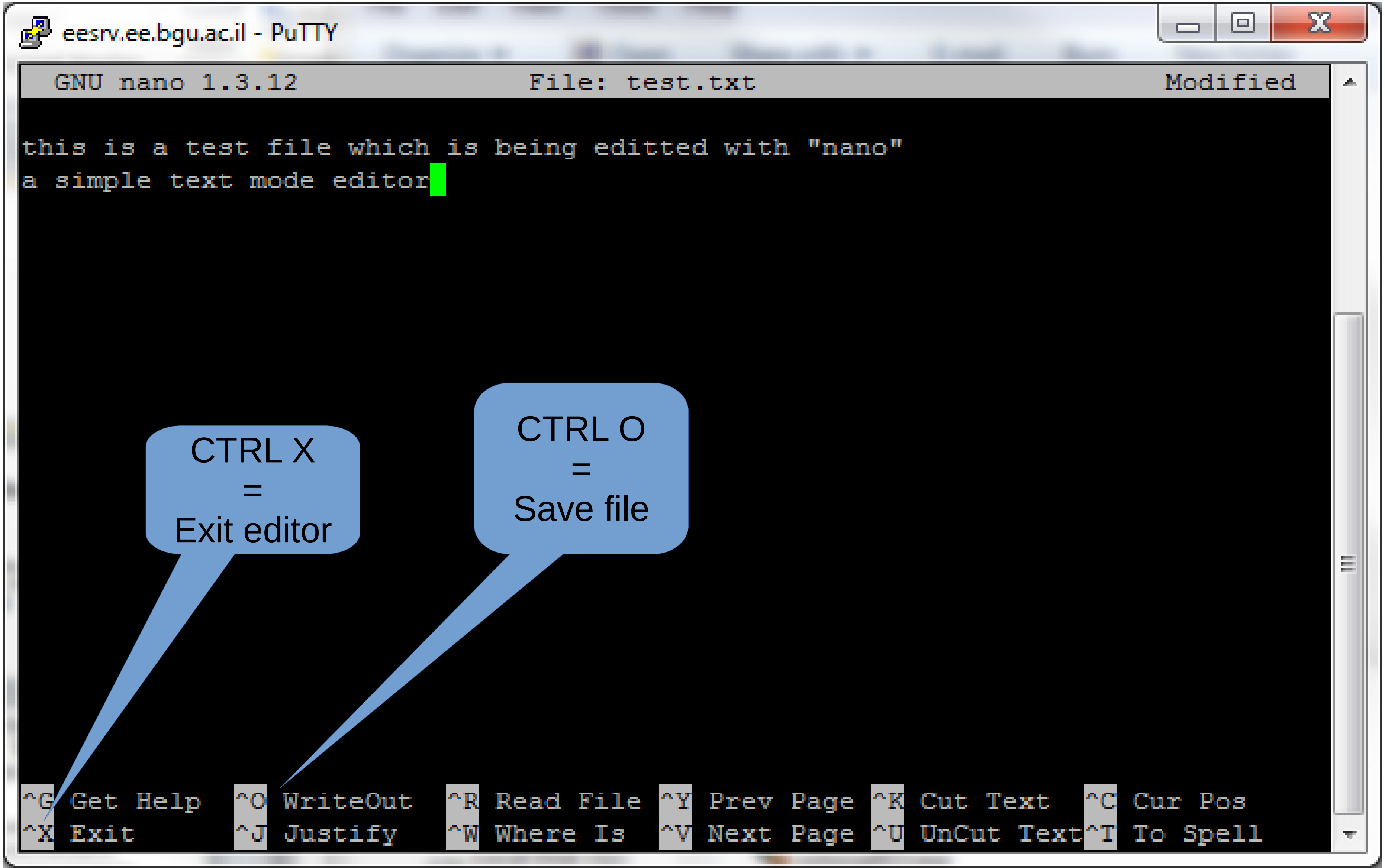
VIM

```
motion
  h - move left
  j - move down
  k - move up
  l - move right
  w - move to next word
  W - move to next blank delimited word
  b - move to beginning of the word
  B - move to beginning of blank delimited word
  e - move to end of word
  E - move to end of blank delimited word
  ( - move a sentence back
  ) - move a sentence forward
  { - move paragraph back
  } - move paragraph forward
  0 - move to beginning of line
  $ - move to end of line
  nG - move to nth line of file
  :n - move to nth line of file
  G - move to last line of file
  fc - move forward to ‘c’
  Fc - move backward to ‘c’
  H - move to top of screen
  M - move to middle of screen
  L - move to bottom of screen
  % - move to associated (),{},[]
deleting text
  x - delete character to the right
  X - delete character to the left
  D - delte to the end of line
  dd - delete current line
  :d - delete current line
searching
  /string - search forward for string
  ?string - search back for string
  n - search for next instance of string
  N - for for previous instance of string
replace
  :s/pattern/string/flags - replace pattern with
    string, according to flags
  g - flag, replace all occurences
  c - flag, confirm replaces
  & - repeat last :s command
files
  :w file - write to file
  :r file - read file in after line
  :n - go to next file
  :p - go to previous file
  :e file - edit file
  !!cmd - replace line with output of cmd
other
  u - undo last change
  U - undo all changes to line
```

Linux Cheat Sheet

https://www.google.com/search?q=linux+cheat+sheet+pdf

nano



```
  GNU nano 1.3.12                File: test.txt                        Modified

this is a test file which is being editted with "nano"
a simple text mode editor




















^G Get Help   ^O WriteOut   ^R Read File  ^Y Prev Page  ^K Cut Text   ^C Cur Pos
^X Exit       ^J Justify    ^W Where Is   ^V Next Page  ^U UnCut Text^T To Spell
```

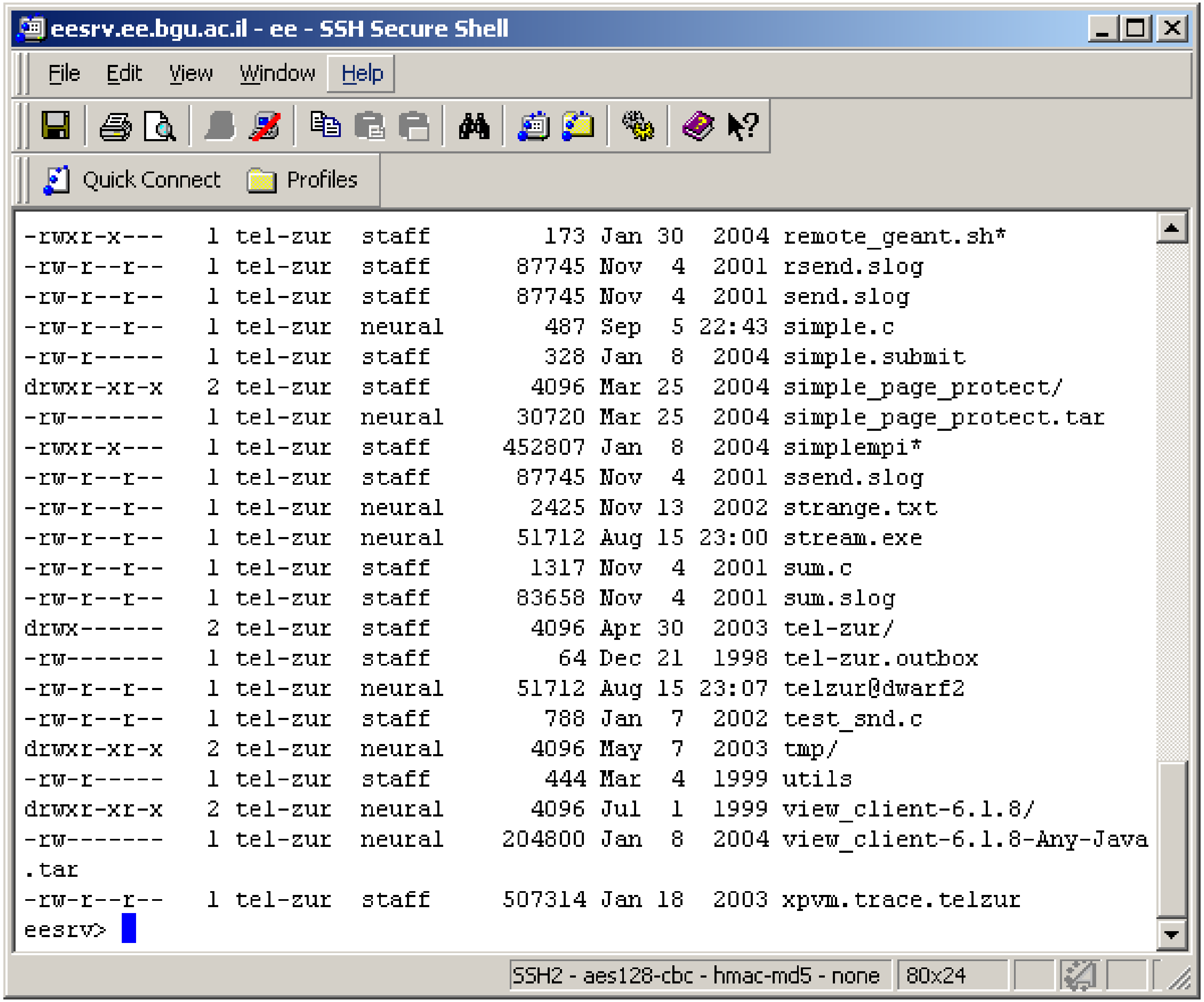
CTRL X
=
Exit editor

CTRL O
=
Save file

Other text editors

- Vi, Vim
- Pico
- Nano
- Emacs/Xemacs
- Nedit (very friendly)
- Gedit
- code::blocks
- Code
- atom
- sublime
- Eclipse IDE (PTP)

הגנה על הפרטיות – הרשאות גישה לקבצים


eesrv.ee.bgu.ac.il - ee - SSH Secure Shell
File
Edit
View
Window
Help
Quick Connect
Profiles
SSH2 - aes128-cbc - hmac-md5 - none
80x24

```
-rwxr-x---   1 tel-zur  staff      173 Jan 30  2004 remote_geant.sh*
-rw-r--r--   1 tel-zur  staff    87745 Nov  4  2001 rsend.slog
-rw-r--r--   1 tel-zur  staff    87745 Nov  4  2001 send.slog
-rw-r--r--   1 tel-zur  neural     487 Sep  5 22:43 simple.c
-rw-r-----   1 tel-zur  staff      328 Jan  8  2004 simple.submit
drwxr-xr-x   2 tel-zur  staff     4096 Mar 25  2004 simple_page_protect/
-rw-------   1 tel-zur  neural   30720 Mar 25  2004 simple_page_protect.tar
-rwxr-x---   1 tel-zur  staff   452807 Jan  8  2004 simplempi*
-rw-r--r--   1 tel-zur  staff    87745 Nov  4  2001 ssend.slog
-rw-r--r--   1 tel-zur  neural    2425 Nov 13  2002 strange.txt
-rw-r--r--   1 tel-zur  neural   51712 Aug 15 23:00 stream.exe
-rw-r--r--   1 tel-zur  staff     1317 Nov  4  2001 sum.c
-rw-r--r--   1 tel-zur  staff    83658 Nov  4  2001 sum.slog
drwx------   2 tel-zur  staff     4096 Apr 30  2003 tel-zur/
-rw-------   1 tel-zur  staff       64 Dec 21  1998 tel-zur.outbox
-rw-r--r--   1 tel-zur  neural   51712 Aug 15 23:07 telzur@dwarf2
-rw-r--r--   1 tel-zur  staff      788 Jan  7  2002 test_snd.c
drwxr-xr-x   2 tel-zur  neural    4096 May  7  2003 tmp/
-rw-r-----   1 tel-zur  staff      444 Mar  4  1999 utils
drwxr-xr-x   2 tel-zur  neural    4096 Jul  1  1999 view_client-6.1.8/
-rw-------   1 tel-zur  neural  204800 Jan  8  2004 view_client-6.1.8-Any-Java
.tar
-rw-r--r--   1 tel-zur  staff   507314 Jan 18  2003 xpvm.trace.telzur
eesrv>
```

gathering information

Kernel version: `uname –a OR uname -r`

CPU information: `more /proc/cpuinfo`

Memory Information: `more /proc/meminfo`

```
lsblk – Block devices
lscpu – CPU information
lshw – HW information
lsusb – UBS buses and devices
lspci – HW configuration
```

Connecting to the cluster from Windows machines

Part 1: connecting as a text terminal

- PuTTY
- ssh.com
- cygwin

Part 2: connecting with X (graphics mode)

- X2go
- cygwin xwin server
- (VNC)

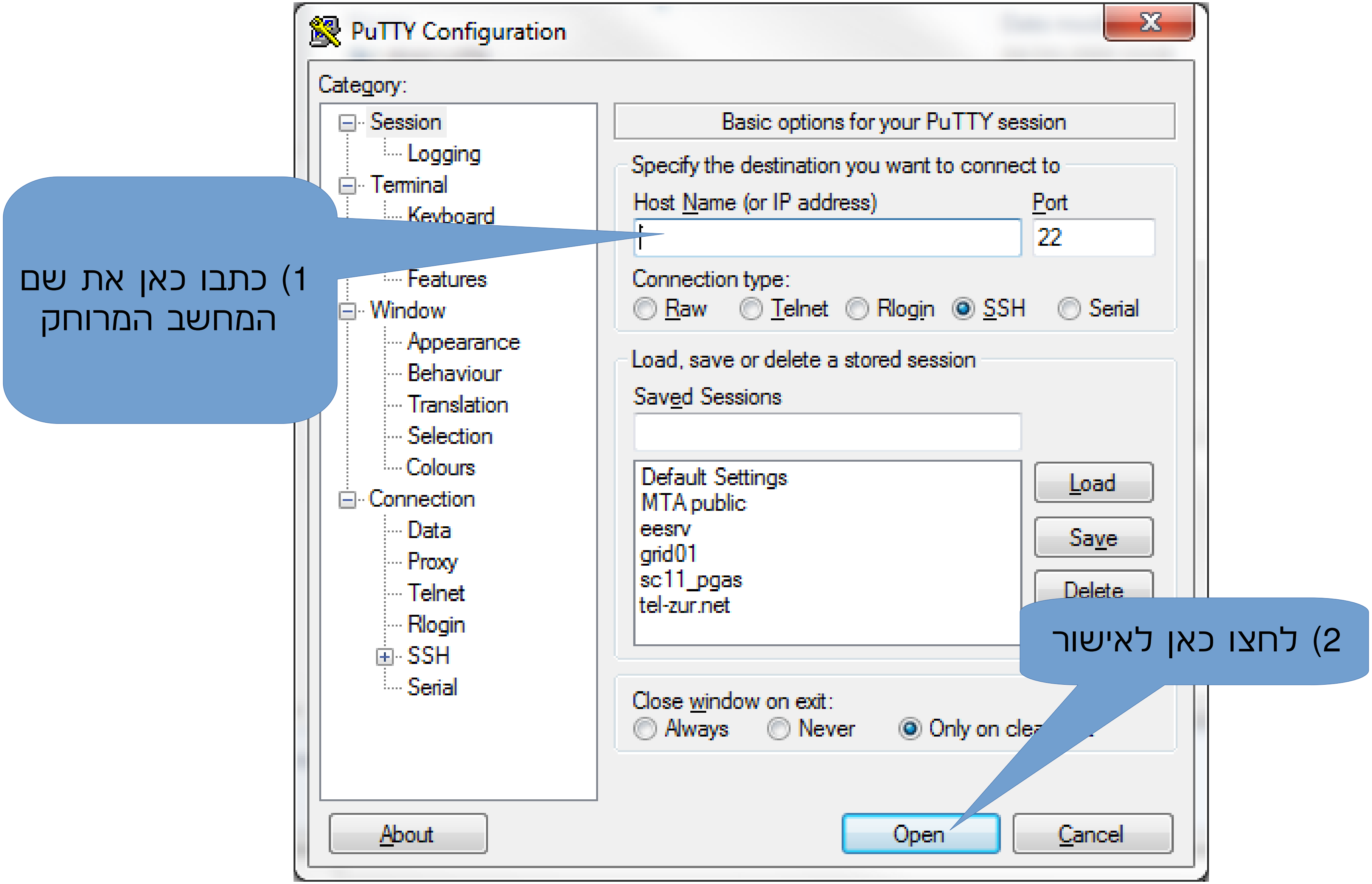
Connecting to a remote node

- Secured: SSH

SSH client from:
http://www.ssh.com/support/downloads/
PuTTY:
http://www.chiark.greenend.org.uk/~sgtatham/putty/

Please download Putty!!!!!

Putty



PuTTY Configuration

Category:

- Session
  - Logging
- Terminal
  - Keyboard
  - Features
- Window
  - Appearance
  - Behaviour
  - Translation
  - Selection
  - Colours
- Connection
  - Data
  - Proxy
  - Telnet
  - Rlogin
  - SSH
  - Serial

Basic options for your PuTTY session

Specify the destination you want to connect to

Host Name (or IP address):

Port: 22

Connection type: ○ Raw ○ Telnet ○ Rlogin ● SSH ○ Serial

Load, save or delete a stored session

Saved Sessions

Default Settings
MTA public
eesrv
grid01
sc11_pgas
tel-zur.net

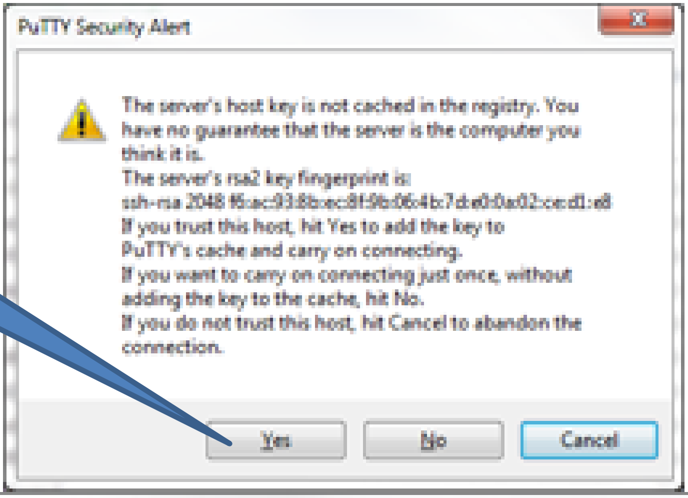
Close window on exit: ○ Always ○ Never ● Only on clea

1) כתבו כאן את שם המחשב המרוחק

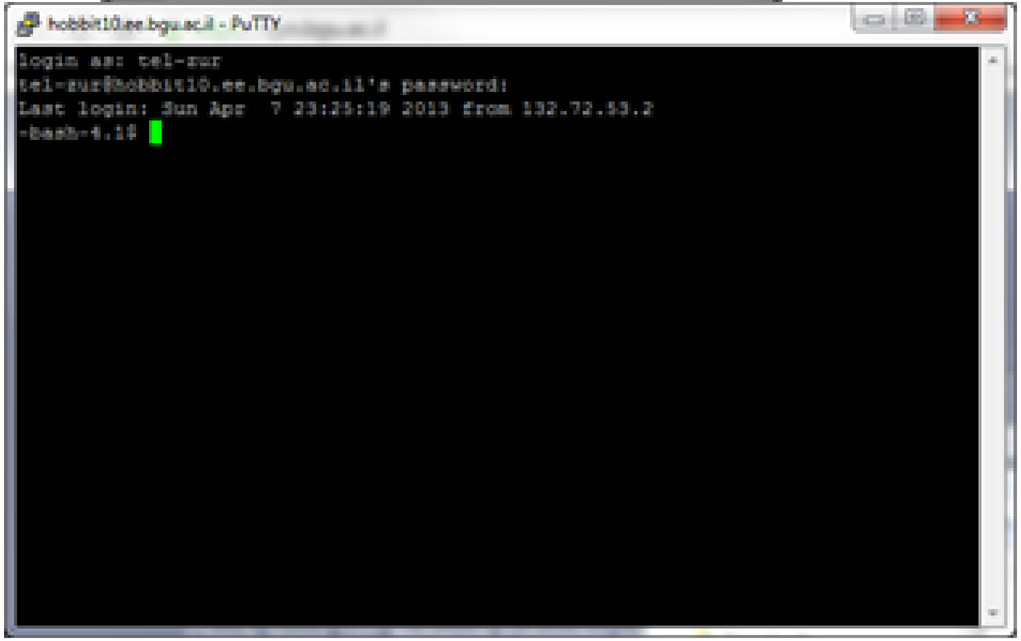
2) לחצו כאן לאישור

Upon first connection

Click on “yes”
מופיע רק
בהתחברות
הראשונה



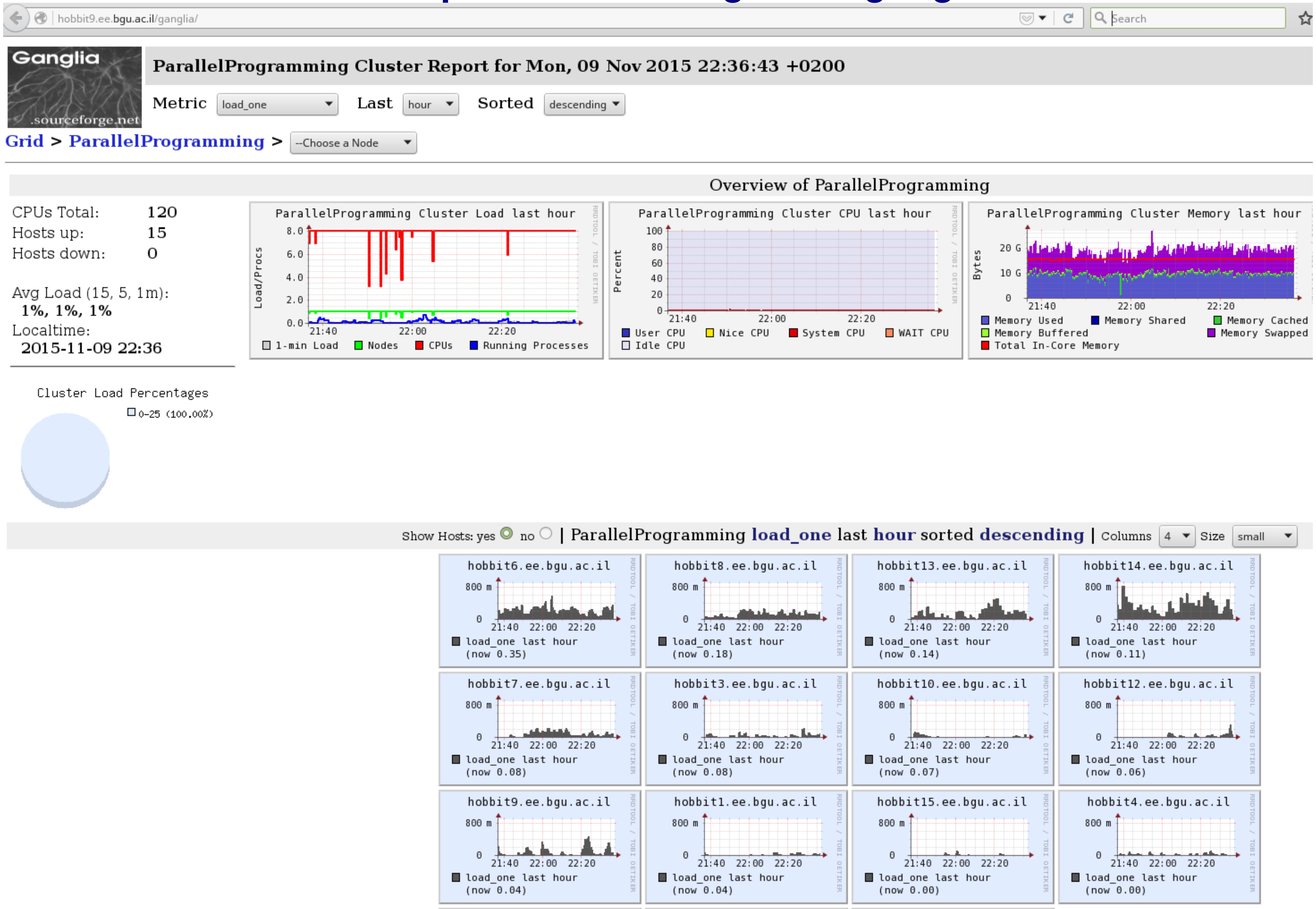
The server's host key is not cached in the registry. You have no guarantee that the server is the computer you think it is.
The server's rsa2 key fingerprint is:
ssh-rsa 2048 f6:ac:93:8b:ec:8f:9b:06:4b:7d:e0:0a:02:ce:d1:e8
If you trust this host, hit Yes to add the key to PuTTY's cache and carry on connecting.
If you want to carry on connecting just once, without adding the key to the cache, hit No.
If you do not trust this host, hit Cancel to abandon the connection.



```
login as: tel-zur
tel-zur@hobbit10.ee.bgu.ac.il's password:
Last login: Sun Apr  7 23:25:19 2013 from 132.72.53.2
-bash-4.1$ 
```

The Educational Cluster

http://hobbit9.ee.bgu.ac.il/ganglia/



hobbit9.ee.bgu.ac.il/ganglia/

Search

Ganglia .sourceforge.net

ParallelProgramming Cluster Report for Mon, 09 Nov 2015 22:36:43 +0200

Metric load_one Last hour Sorted descending

Grid > ParallelProgramming > --Choose a Node

Overview of ParallelProgramming

CPUs Total:	120
Hosts up:	15
Hosts down:	0

Avg Load (15, 5, 1m):
1%, 1%, 1%

Localtime:
2015-11-09 22:36

ParallelProgramming Cluster Load last hour
Load/Procs
8.0
6.0
4.0
2.0
0.0
21:40
22:00
22:20
□ 1-min Load 🟩 Nodes 🟥 CPUs 🟦 Running Processes
ParallelProgramming Cluster CPU last hour
Percent
100
80
60
40
20
0
21:40
22:00
22:20
🟦 User CPU 🟨 Nice CPU 🟥 System CPU 🟧 WAIT CPU
□ Idle CPU
ParallelProgramming Cluster Memory last hour
Bytes
20 G
10 G
0
21:40
22:00
22:20
🟦 Memory Used 🟦 Memory Shared 🟩 Memory Cached
🟩 Memory Buffered 🟪 Memory Swapped
🟥 Total In-Core Memory
Cluster Load Percentages
□ 0-25 (100.00%)
RRDTOOL / TOBI OETIKER

Show Hosts: yes ● no ○ | ParallelProgramming **load_one** last **hour** sorted **descending** | Columns 4 Size small

The Educational Cluster

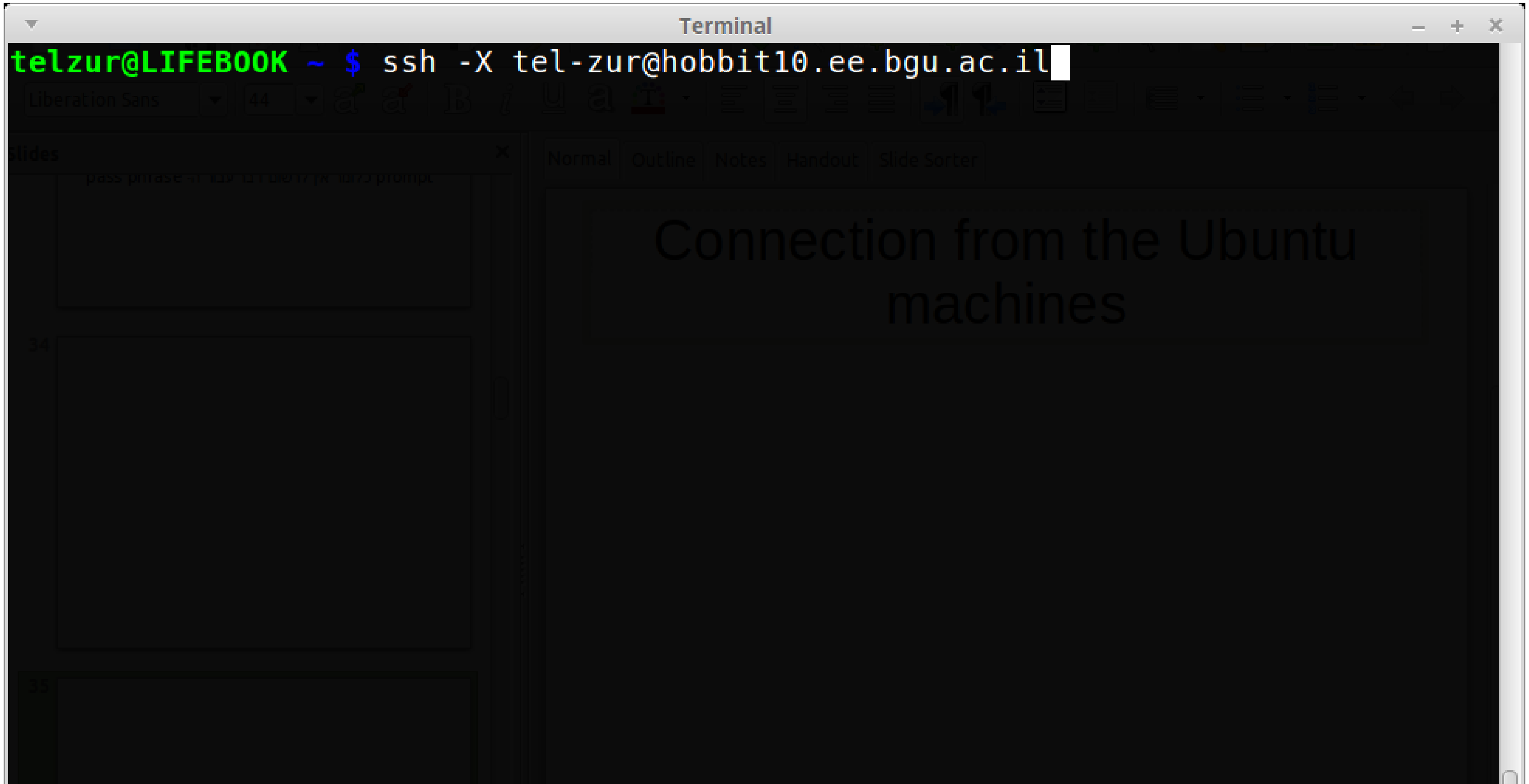
Hobbit

hobbitX.ee.bgu.ac.il

Where X=1,2,3..,15

Connecting to the Cluster from Linux machines

Connection from the Ubuntu machines (from lab-330!): restart the machine and select Ubuntu at the boot manager

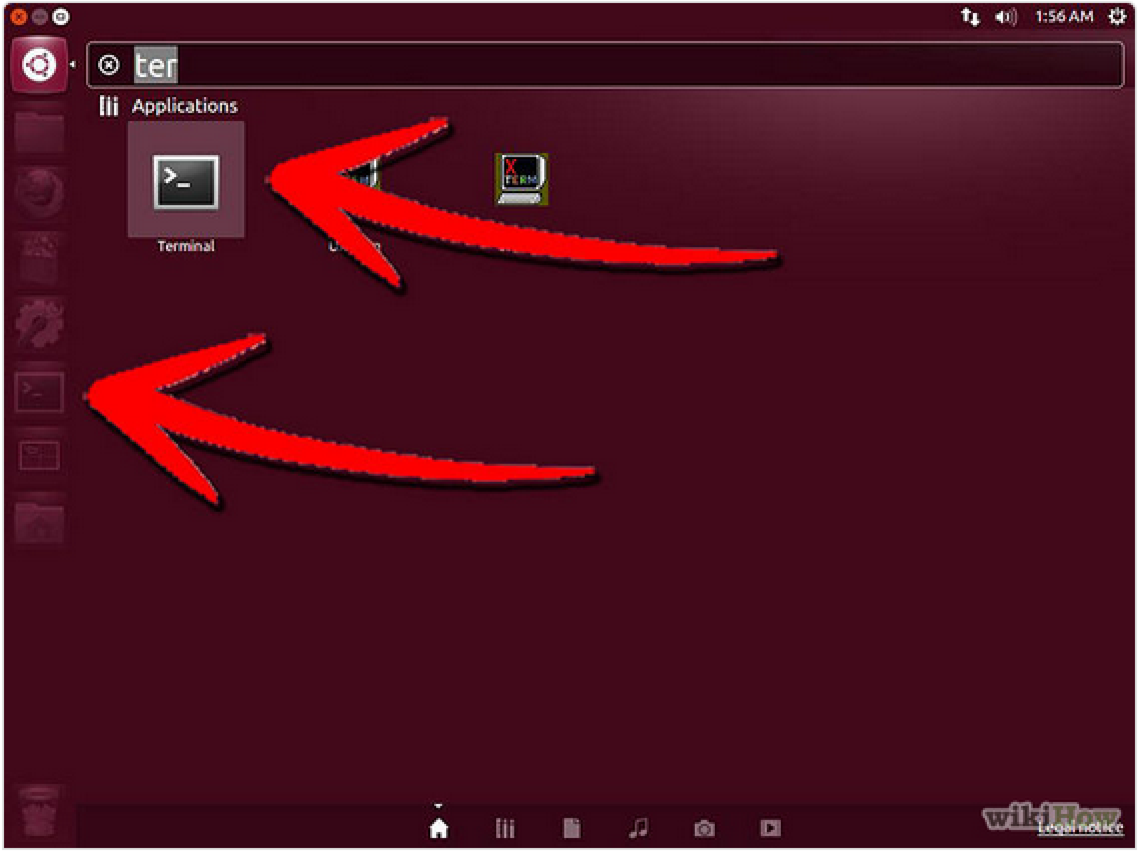


```
telzur@LIFEBOOK ~ $ ssh -X tel-zur@hobbit10.ee.bgu.ac.il
```

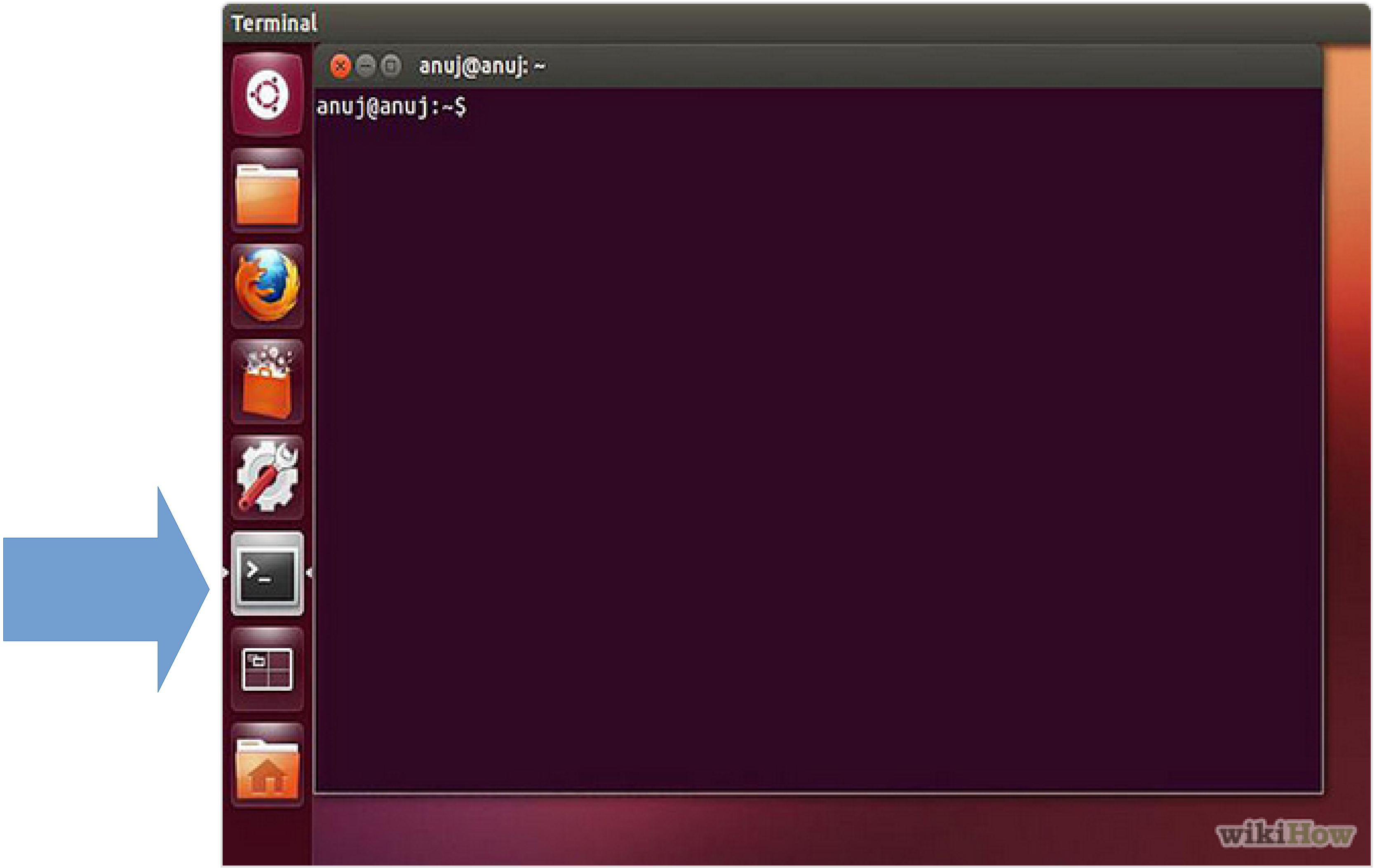
Open a terminal in Ubuntu

Use the keyboard shortcut Ctrl+Alt+T

- You also may be able to access the terminal by going to the side bar on the left, or by searching for it in the Dashboard, which is located at the top in the left side bar.



Open a terminal in Ubuntu



```
anuj@anuj:~$
```

```
ssh -X username@hostname
```

תמיכה במוד גראפי

Use your user name and password and select the hobbit you want to connect

If you can’t connect with ssh, type:

```
unsetenv LD_LIBRARY_PATH
```

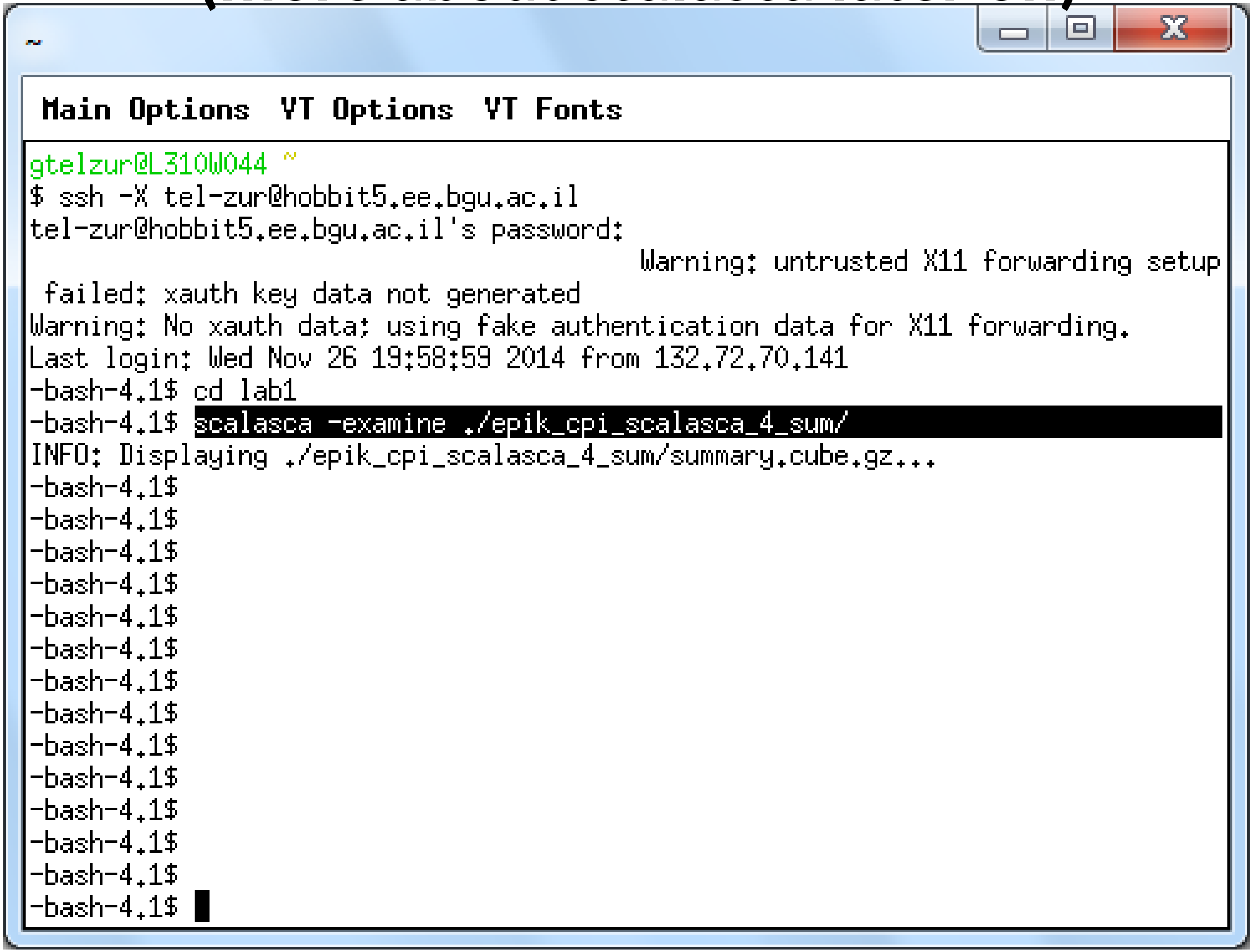
נדרש חיבור מקדים ב- VPN

בעת ההתחברות הראשונה יש ללחוץ מספר פעמים ENTER עד שמקבלים את ה- Shell prompt
לדוגמה: <csh

In case of login troubles

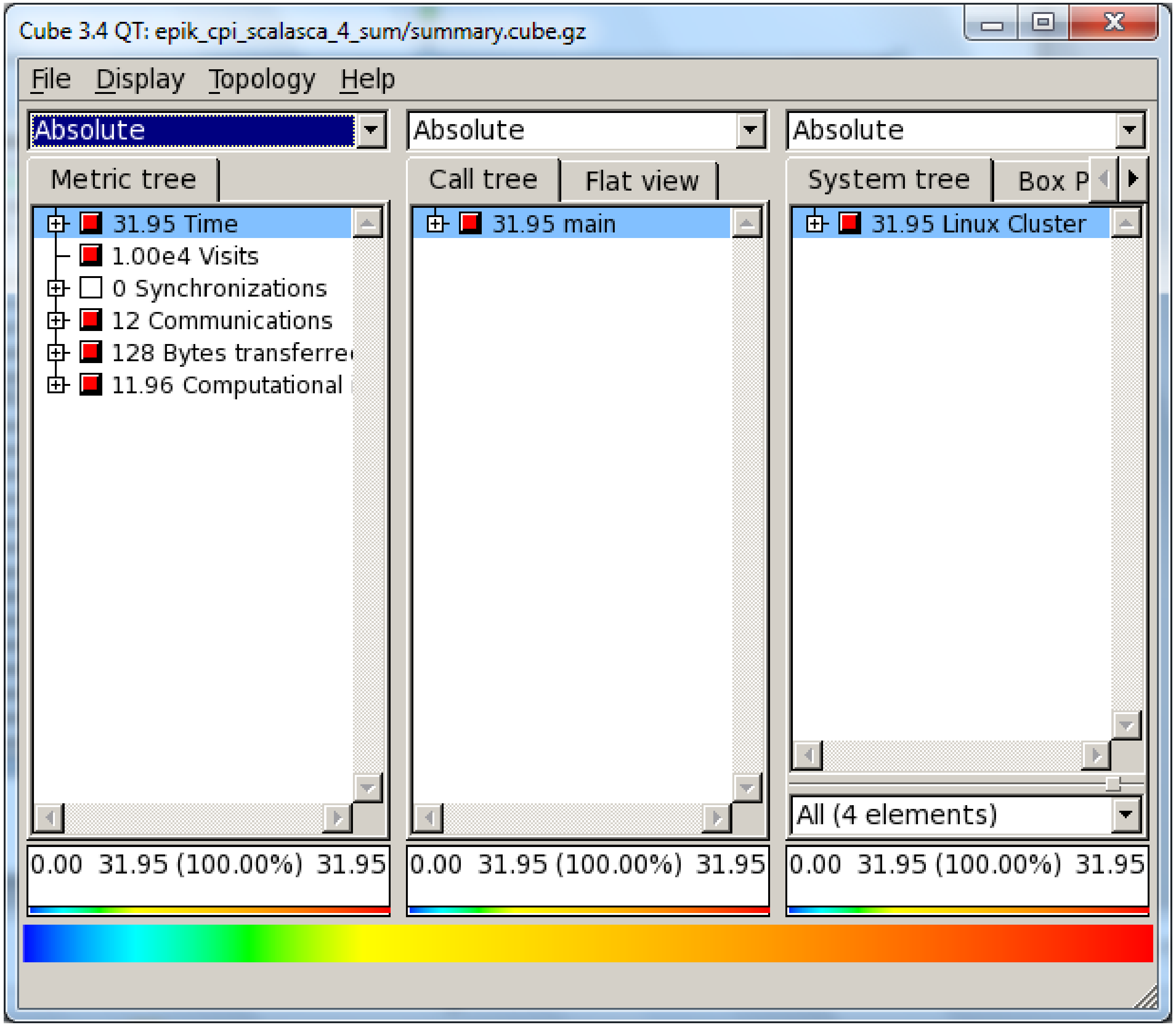
- אם אינכם מצליחים לעשות SSH בין מחשבי ההוביט ללא צורך בסיסמא יש למחוק את המחיצה **.ssh**
בצורה הבאה: rm -Rf .ssh
- אח"כ יש לצאת ולהכנס שוב. עתה בכל עצירה של המחשב יש ללחוץ <enter> בלבד עד שמקבלים את ה- prompt כלומר אין לרשום דבר עבור ה- pass phrase

Now you have X, lets try scalasca for example (more about scalasca later on)



```
gtelzur@L310W044 ~
$ ssh -X tel-zur@hobbit5.ee.bgu.ac.il
tel-zur@hobbit5.ee.bgu.ac.il's password:
                                          Warning: untrusted X11 forwarding setup
 failed: xauth key data not generated
Warning: No xauth data; using fake authentication data for X11 forwarding.
Last login: Wed Nov 26 19:58:59 2014 from 132.72.70.141
-bash-4.1$ cd lab1
-bash-4.1$ scalasca -examine ./epik_cpi_scalasca_4_sum/
INFO: Displaying ./epik_cpi_scalasca_4_sum/summary.cube.gz...
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
-bash-4.1$
```

The graphical app will pop up



- Password-less SSH:

http://telzur.blogspot.co.il/2006/03/password-less-ssh-connection.html

Files Transfer

- FileZilla
- Secure FTP (sftp)
- Cyberduck
- From a linux host: scp….

The syntax is:

```
scp ./file user@host:/path/to/directory/.
```

לביצוע במחשבים הוירטואלים ובהוביטים

תרגיל מס' 1

- התחבר לאחת מההתחנות תוך שימוש ב- ssh
- כתוב תכנית מחשב קצרה כגון : Hello World
- בצע קומפילציה:

```
gcc –o hello_world hello_world.c
```

- הרץ את התכנית ושמור הפלט:

```
./hello_world  > hello.txt
```

- בדוק את הפלט על-ידי:

```
more hello.txt
```

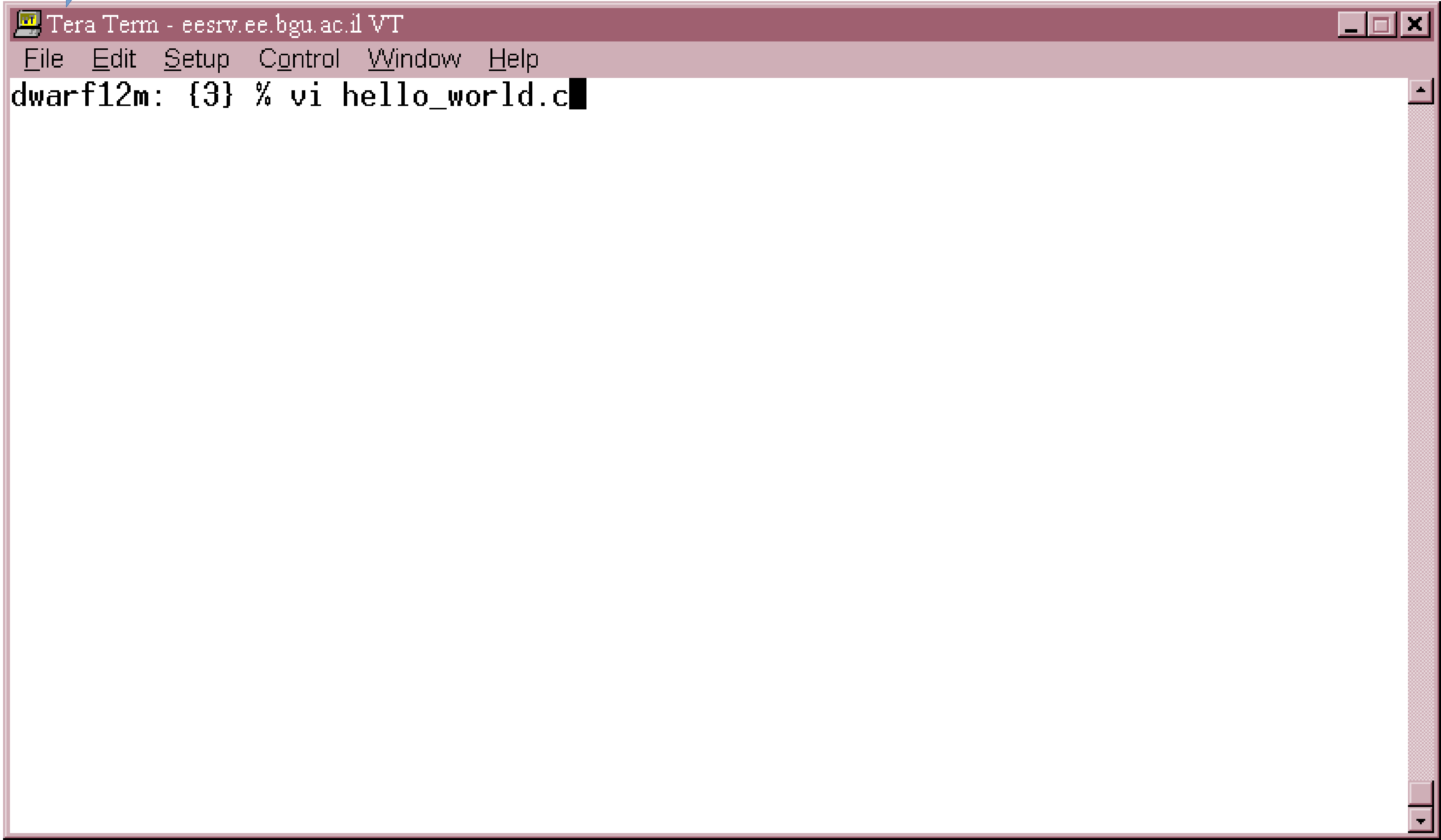
The GNU compiler

http://gcc.gnu.org/

- `gcc filename.c`
  - Will produce an executable "`a.out`"
- `gcc –o runme filename.c`
  - Will produce an executable "runme"
- Optimization: `gcc –O3 –o runme filename.c`
- `gcc –c filename.c` will produce an object file "filename.o"

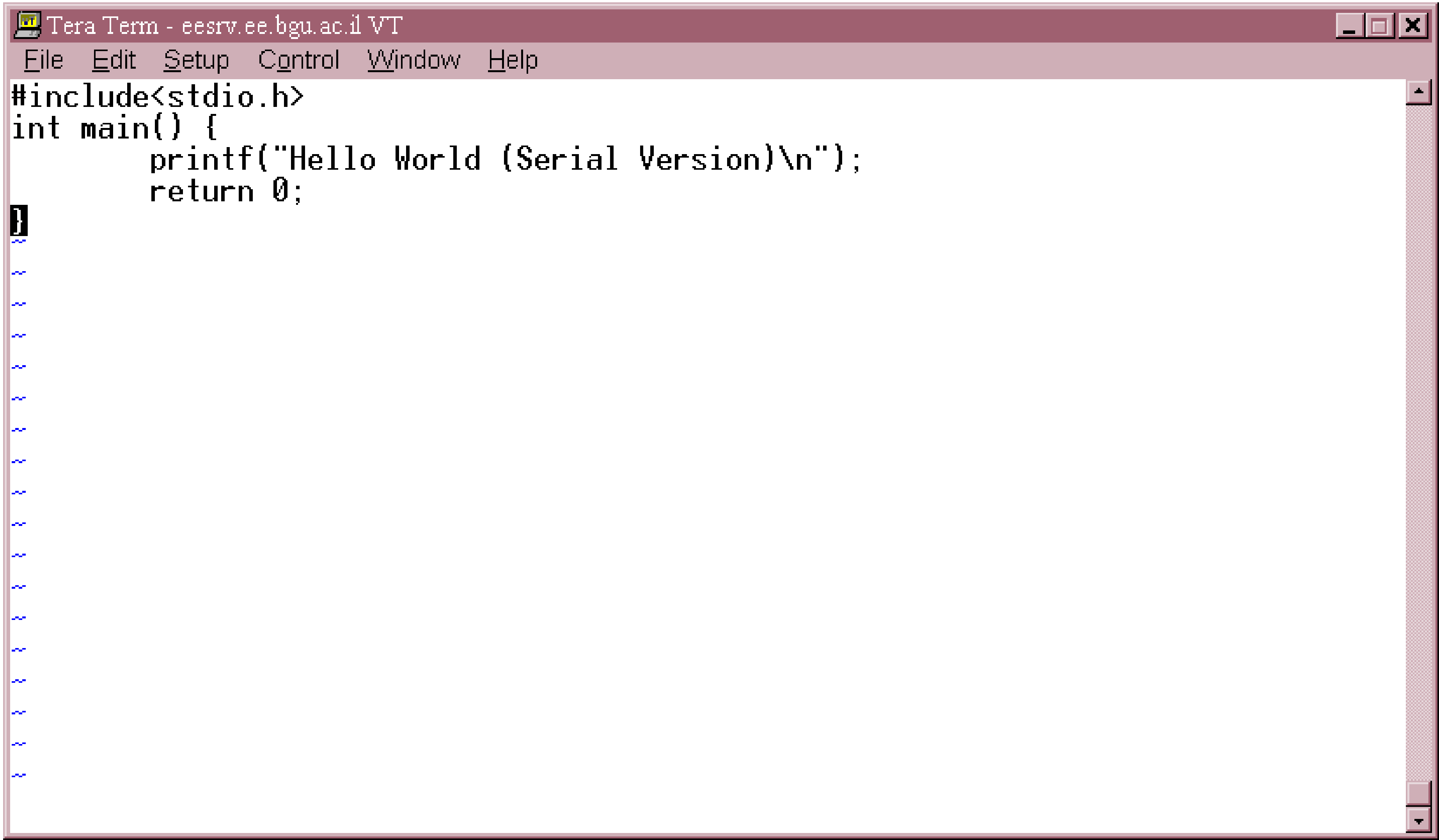
Terminal window

פתרון תרגיל מס' 1 שימוש בעורך- vi


Tera Term - eesrv.ee.bgu.ac.il VT
File Edit Setup Control Window Help

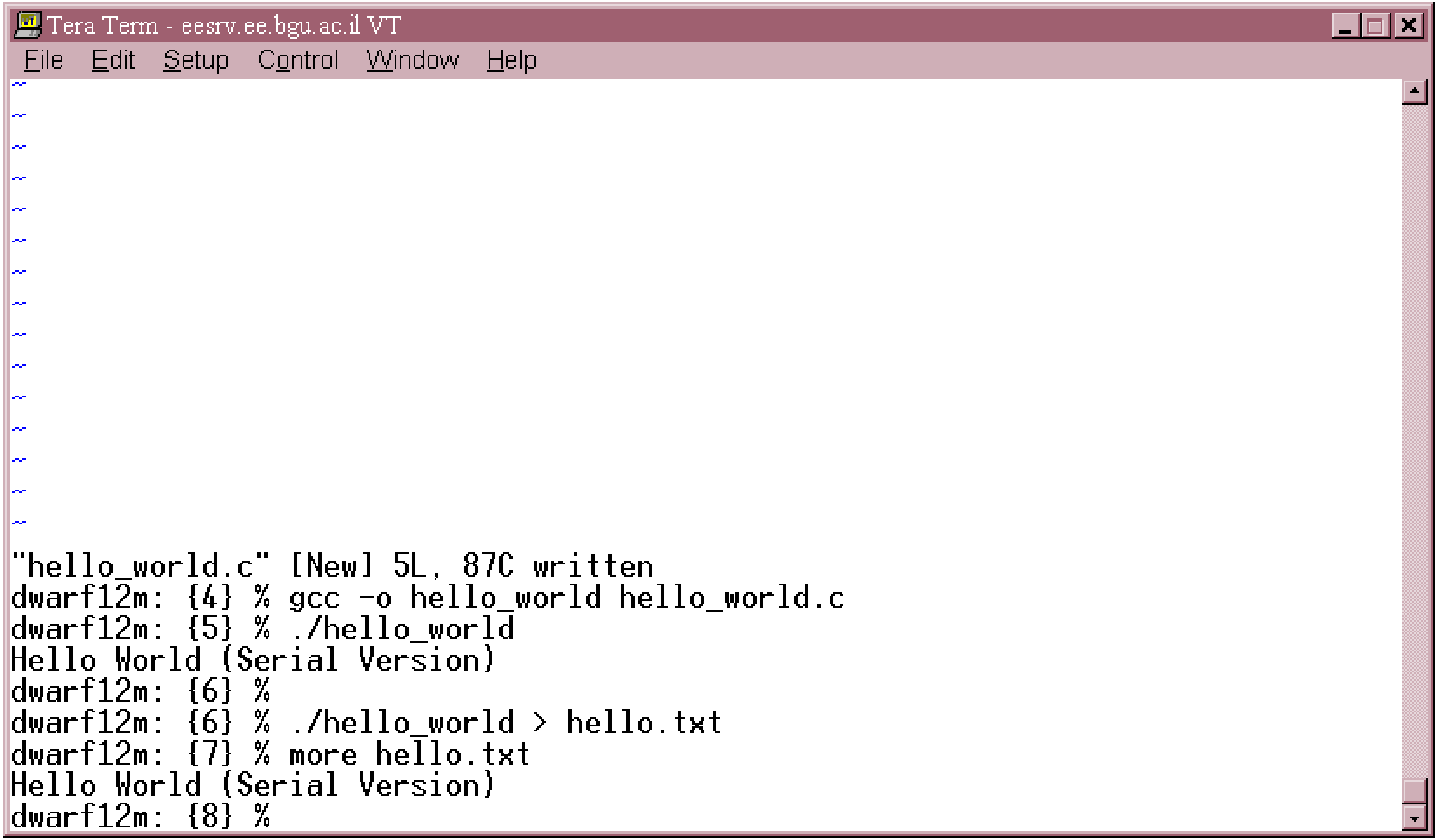
```
dwarf12m: {3} % vi hello_world.c
```

פתרון תרגיל מס' 1 – 2/3


Tera Term - eesrv.ee.bgu.ac.il VT
File Edit Setup Control Window Help

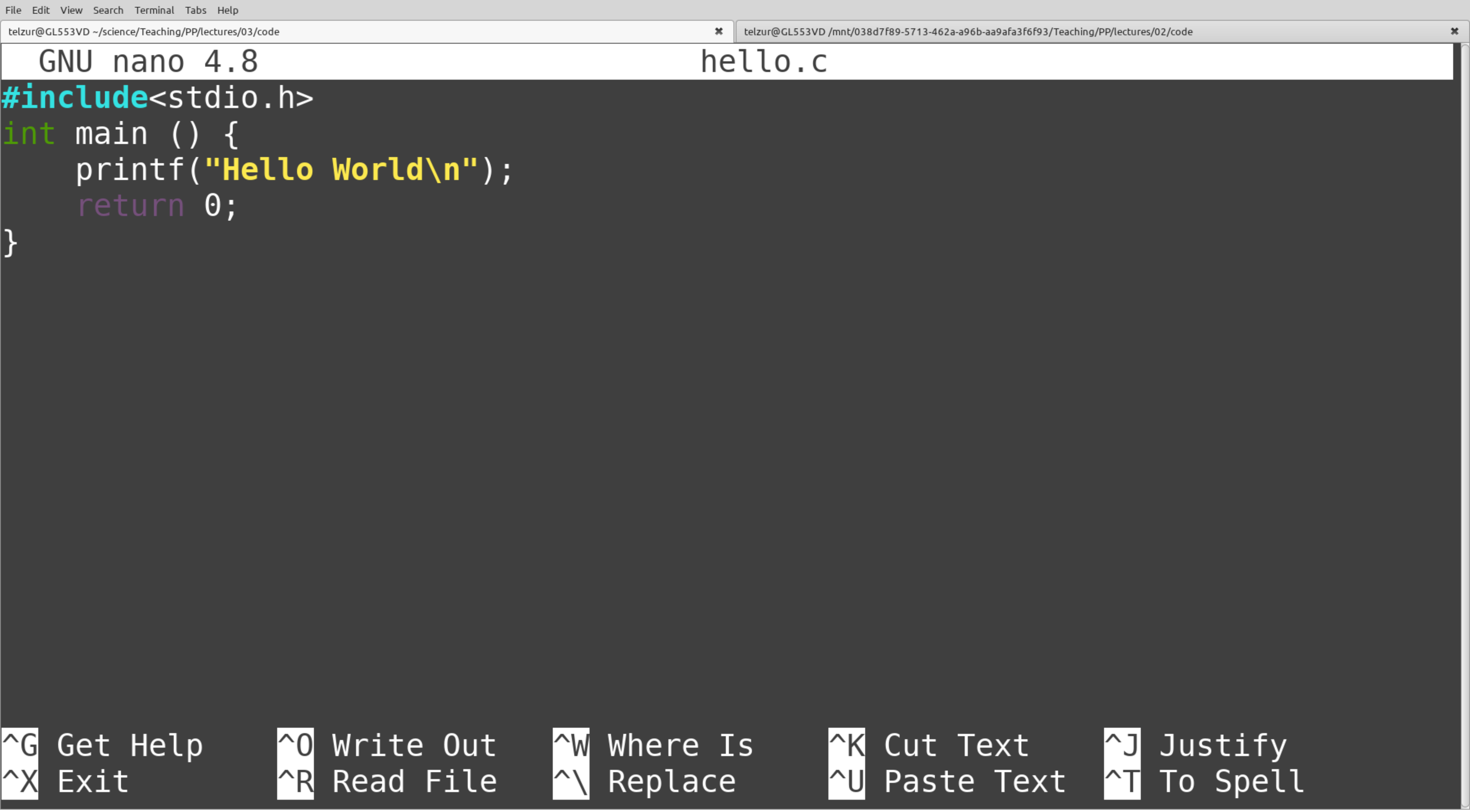
```
#include<stdio.h>
int main() {
        printf("Hello World (Serial Version)\n");
        return 0;
}
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
```

פתרון תרגיל מס' 1 – 3/3


Tera Term - eesrv.ee.bgu.ac.il VT
File Edit Setup Control Window Help

```
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
"hello_world.c" [New] 5L, 87C written
dwarf12m: {4} % gcc -o hello_world hello_world.c
dwarf12m: {5} % ./hello_world
Hello World (Serial Version)
dwarf12m: {6} %
dwarf12m: {6} % ./hello_world > hello.txt
dwarf12m: {7} % more hello.txt
Hello World (Serial Version)
dwarf12m: {8} %
```

שימוש בעורך nano



File Edit View Search Terminal Tabs Help
telzur@GL553VD ~/science/Teaching/PP/lectures/03/code
telzur@GL553VD /mnt/038d7f89-5713-462a-a96b-aa9afa3f6f93/Teaching/PP/lectures/02/code

```
  GNU nano 4.8                            hello.c
#include<stdio.h>
int main () {
    printf("Hello World\n");
    return 0;
}
```

```
^G Get Help     ^O Write Out    ^W Where Is     ^K Cut Text     ^J Justify
^X Exit         ^R Read File    ^\ Replace      ^U Paste Text   ^T To Spell
```

Message Passing Interface

Quick Reference in C

```
#include <mpi.h>
```

Blocking Point-to-Point

Send a message to one process. (§3.2.1)

```
int MPI_Send (void *buf, int count,
    MPI_Datatype datatype, int dest, int
    tag, MPI_Comm comm)
```

Receive a message from one process. (§3.2.4)

```
int MPI_Recv (void *buf, int count,
    MPI_Datatype datatype, int source, int
    tag, MPI_Comm comm, MPI_Status *status)
```

Count received data elements. (§3.2.5)

```
int MPI_Get_count (MPI_Status *status,
    MPI_Datatype datatype, int *count)
```

Wait for message arrival. (§3.8)

```
int MPI_Probe (int source, int tag,
    MPI_Comm comm, MPI_Status *status)
```

Related Functions: MPI_Bsend, MPI_Ssend, MPI_Rsend, MPI_Buffer_attach, MPI_Buffer_detach, MPI_Sendrecv, MPI_Sendrecv_replace, MPI_Get_elements

Non-blocking Point-to-Point

Begin to receive a message. (§3.7.2)

```
int MPI_Irecv (void *buf, int count,
    MPI_Datatype, int source, int tag,
    MPI_Comm comm, MPI_Request *request)
```

Complete a non-blocking operation. (§3.7.3)

```
int MPI_Wait (MPI_Request *request,
    MPI_Status *status)
```

Check or complete a non-blocking operation. (§3.7.3)

```
int MPI_Test (MPI_Request *request, int
    *flag, MPI_Status *status)
```

Check message arrival. (§3.8)

```
int MPI_Iprobe (int source, int tag,
    MPI_Comm comm, int *flag, MPI_Status
    *status)
```

Related Functions: MPI_Isend, MPI_Ibsend, MPI_Issend, MPI_Irsend, MPI_Request_free, MPI_Waitany, MPI_Testany, MPI_Waitall, MPI_Testall, MPI_Waitsome, MPI_Testsome, MPI_Cancel, MPI_Test_cancelled

Persistent Requests

Related Functions: MPI_Send_init, MPI_Bsend_init, MPI_Ssend_init, MPI_Rsend_init, MPI_Recv_init, MPI_Start, MPI_Startall

Derived Datatypes

Create a strided homogeneous vector. (§3.12.1)

```
int MPI_Type_vector (int count, int
    blocklength, int stride, MPI_Datatype
    oldtype, MPI_Datatype *newtype)
```

Save a derived datatype (§3.12.4)

```
int MPI_Type_commit (MPI_Datatype
    *datatype)
```

Pack data into a message buffer. (§3.13)

```
int MPI_Pack (void *inbuf, int incount,
    MPI_Datatype datatype, void *outbuf,
    int outsize, int *position, MPI_Comm
    comm)
```

Unpack data from a message buffer. (§3.13)

```
int MPI_Unpack (void *inbuf, int insize,
    int *position, void *outbuf, int
    outcount, MPI_Datatype datatype,
    MPI_Comm comm)
```

Determine buffer size for packed data. (§3.13)

```
int MPI_Pack_size (int incount,
    MPI_Datatype datatype, MPI_Comm comm,
    int *size)
```

Related Functions: MPI_Type_contiguous, MPI_Type_hvector, MPI_Type_indexed, MPI_Type_hindexed, MPI_Type_struct, MPI_Address, MPI_Type_extent, MPI_Type_size, MPI_Type_lb, MPI_Type_ub, MPI_Type_free

Collective

Send one message to all group members. (§4.4)

```
int MPI_Bcast (void *buf, int count,
    MPI_Datatype datatype, int root,
    MPI_Comm comm)
```

Receive from all group members. (§4.5)

```
int MPI_Gather (void *sendbuf, int
    sendcount, MPI_Datatype sendtype, void
    *recvbuf, int recvcount, MPI_Datatype
    recvtype, int root, MPI_Comm comm)
```

Send separate messages to all group members. (§4.6)

```
int MPI_Scatter (void *sendbuf, int
    sendcount, MPI_Datatype sendtype, void
    *recvbuf, int recvcount, MPI_Datatype
    recvtype, int root, MPI_Comm comm)
```

Combine messages from all group members. (§4.9.1)

```
int MPI_Reduce (void *sendbuf, void
    *recvbuf, int count, MPI_Datatype
    datatype, MPI_Op op, int root, MPI_Comm
    comm)
```

Related Functions: MPI_Barrier, MPI_Gatherv, MPI_Scatterv, MPI_Allgather, MPI_Allgatherv, MPI_Alltoall, MPI_Alltoallv, MPI_Op_create, MPI_Op_free, MPI_Allreduce, MPI_Reduce_scatter, MPI_Scan

Groups

Related Functions: MPI_Group_size, MPI_Group_rank, MPI_Group_translate_ranks, MPI_Group_compare, MPI_Comm_group, MPI_Group_union, MPI_Group_intersection, MPI_Group_difference, MPI_Group_incl, MPI_Group_excl, MPI_Group_range_incl, MPI_Group_range_excl, MPI_Group_free

Basic Communicators

Count group members in communicator. (§5.4.1)

```
int MPI_Comm_size (MPI_Comm comm, int
    *size)
```

Determine group rank of self. (§5.4.1)

```
int MPI_Comm_rank (MPI_Comm comm, int
    *rank)
```

Duplicate with new context. (§5.4.2)

```
int MPI_Comm_dup (MPI_Comm comm, MPI_Comm
    *newcomm)
```

Split into categorized sub-groups. (§5.4.2)

```
int MPI_Comm_split (MPI_Comm comm, int
    color, int key, MPI_Comm *newcomm)
```

Related Functions: MPI_Comm_compare, MPI_Comm_create, MPI_Comm_free

MPI_Comm_test_inter, MPI_Comm_remote_size, MPI_Comm_remote_group, MPI_Intercomm_create, MPI_Intercomm_merge

Communicators with Topology

Create with cartesian topology. (§6.5.1)

```
int MPI_Cart_create (MPI_Comm comm_old,
    int ndims, int *dims, int *periods, int
    reorder, MPI_Comm *comm_cart)
```

Suggest balanced dimension ranges. (§6.5.2)

```
int MPI_Dims_create (int nnodes, int
    ndims, int *dims)
```

Determine rank from cartesian coordinates. (§6.5.4)

```
int MPI_Cart_rank (MPI_Comm comm, int
    *coords, int *rank)
```

Determine cartesian coordinates from rank. (§6.5.4)

```
int MPI_Cart_coords (MPI_Comm comm, int
    rank, int maxdims, int *coords)
```

Determine ranks for cartesian shift. (§6.5.5)

```
int MPI_Cart_shift (MPI_Comm comm, int
    direction, int disp, int *rank_source,
    int *rank_dest)
```

Split into lower dimensional sub-grids. (§6.5.6)

```
int MPI_Cart_sub (MPI_Comm comm, int
    *remain_dims, MPI_Comm *newcomm)
```

Related Functions: MPI_Graph_create, MPI_Topo_test, MPI_Graphdims_get, MPI_Graph_get, MPI_Cartdim_get, MPI_Cart_get, MPI_Graph_neighbors_count, MPI_Graph_neighbors, MPI_Cart_map, MPI_Graph_map

Communicator Caches

Related Functions: MPI_Keyval_create, MPI_Keyval_free, MPI_Attr_put, MPI_Attr_get, MPI_Attr_delete

OHIO SUPERCOMPUTER CENTER

LAM & MPI Information

1224 Kinnear Rd.
Columbus, Ohio 43212
614-292-8492

lam@tbag.osc.edu

http://www.osc.edu/lam.html
ftp://ftp.osc.edu/pub/lam

Error Handling

Related Functions: MPI_Errhandler_create, MPI_Errhandler_set, MPI_Errhandler_get, MPI_Errhandler_free, MPI_Error_string, MPI_Error_class

Dynamic Processes

Spawn a process. (MPI-2)

```
int MPI_Spawn (char prog[], char *argv[],
    int maxprocs, MPI_Info info, int root,
    MPI_Comm parents, MPI_Comm *children,
    int errs[]);
```

Related Functions: MPI_Spawn_multiple, MPI_Ispawn, MPI_Ispawn_multiple, MPI_Port_open, MPI_Port_Close, MPI_Accept, MPI_Connect, MPI_Name_publish, MPI_Name_unpublish, MPI_Name_get, MPI_Iaccept, MPI_Iconnect, MPI_Info_create, MPI_Info_set, MPI_Info_get, MPI_Info_get_valuelen, MPI_Info_get_nkeys, MPI_Info_get_nthkey, MPI_Info_dup, MPI_Info_free, MPI_Info_delete

Environmental

Determine wall clock time. (§7.4)

```
double MPI_Wtime (void)
```

Initialize MPI. (§7.5)

```
int MPI_Init (int *argc, char ***argv)
```

Cleanup MPI. (§7.5)

```
int MPI_Finalize (void)
```

Related Functions: MPI_Get_processor_name, MPI__Wtick, MPI_Initialized, MPI_Abort, MPI_Pcontrol, MPI_Get_version

Constants

Wildcards (§3.2.4)

```
MPI_ANY_TAG, MPI_ANY_SOURCE
```

Elementary Datatypes (§3.2.2)

```
MPI_CHAR, MPI_SHORT, MPI_INT, MPI_LONG,
MPI_UNSIGNED_CHAR, MPI_UNSIGNED_SHORT,
MPI_UNSIGNED, MPI_UNSIGNED_LONG, MPI_FLOAT,
MPI_DOUBLE, MPI_LONG_DOUBLE, MPI_BYTE,
MPI_PACKED
```

Reserved Communicators (§5.2.4)

```
MPI_COMM_WORLD,MPI_COMM_SELF,MPI_COMM_PARENT
```

Reduction Operations (§4.9.2)

```
MPI_MAX, MPI_MIN, MPI_SUM, MPI_PROD,
    MPI_BAND, MPI_BOR, MPI_BXOR, MPI_LAND,
    MPI_LOR, MPI_LXOR
```

LAM Quick Reference

Session Management

Confirm a group of hosts.

```
recon -v <hostfile>
```

Start LAM on a group of hosts.

```
lamboot -v <hostfile>
```

Terminate LAM.

```
wipe -v <hostfile>
```

Hostfile Syntax

```
# comment
<hostname> <userid>
<hostname> <userid>
...etc...
```

Compilation

Compile a program for LAM / MPI.

```
hcc -o <binary> <source> -I<incdir>
    -L<libdir> -l<lib> -lmpi
```

Processes and Messages

Start an SPMD application.

```
mpirun -v -s <src_node> -c <copies>
    <nodes> <program> -- <args>
```

Start a MIMD application.

```
mpirun -v <appfile>
```

Appfile Syntax

```
# comment
<program> -s <src_node> <nodes> -- <ar
<program> -s <src_node> <nodes> -- <ar
...etc...
```

Examine the state of processes.

```
mpitask
```

Examine the state of messages.

```
mpimsg
```

Cleanup all processes and messages.

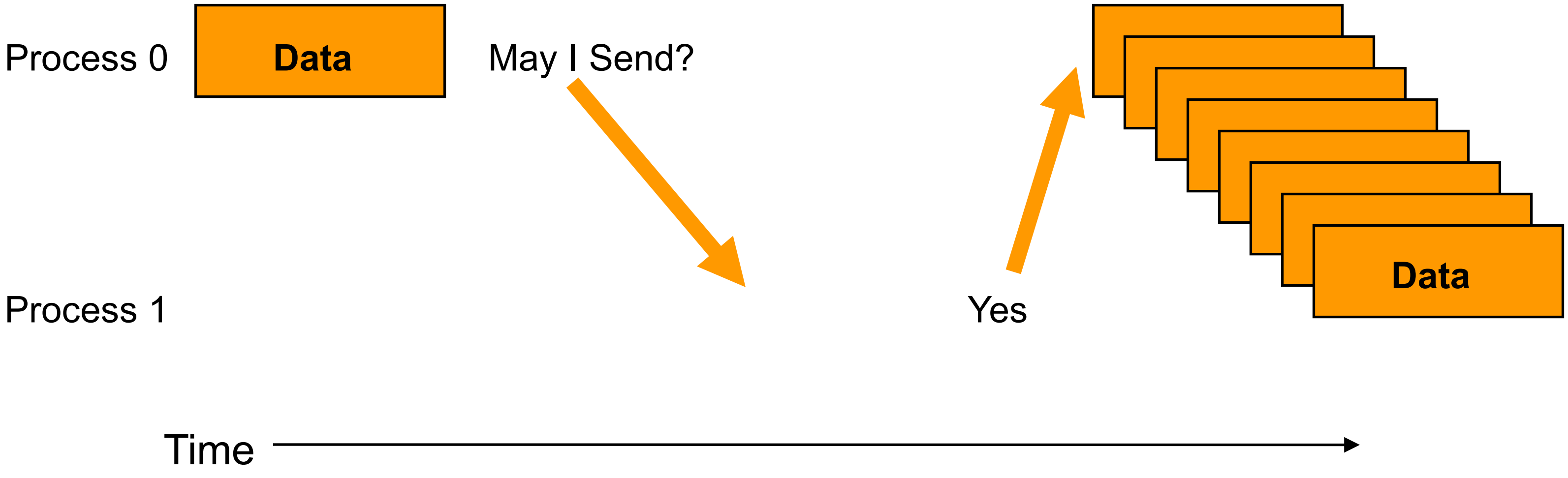
```
lamclean -v
```

MPI Quick Reference Card:

`http://web.eecs.utk.edu/~dongarra/WEB-PAGES/SPRING-2006/mpi-quick-ref.pdf`

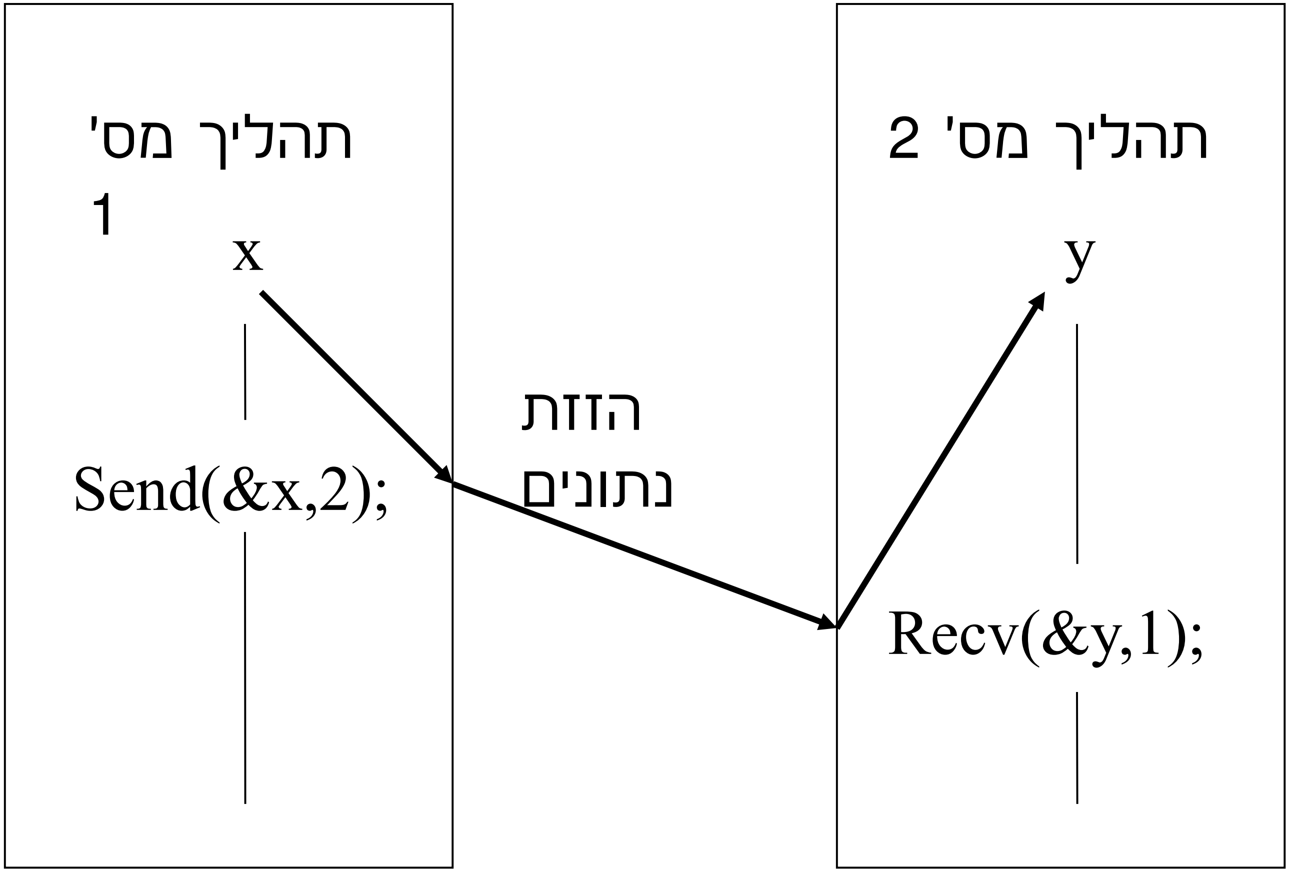
What is message passing?

- Data transfer through messaging



- Requires a **<u>sender</u>** and a **<u>receiver</u>** cooperation

Point to Point:
Basic Send/Receive



תהליך מס'
1

x

Send(&x,2);

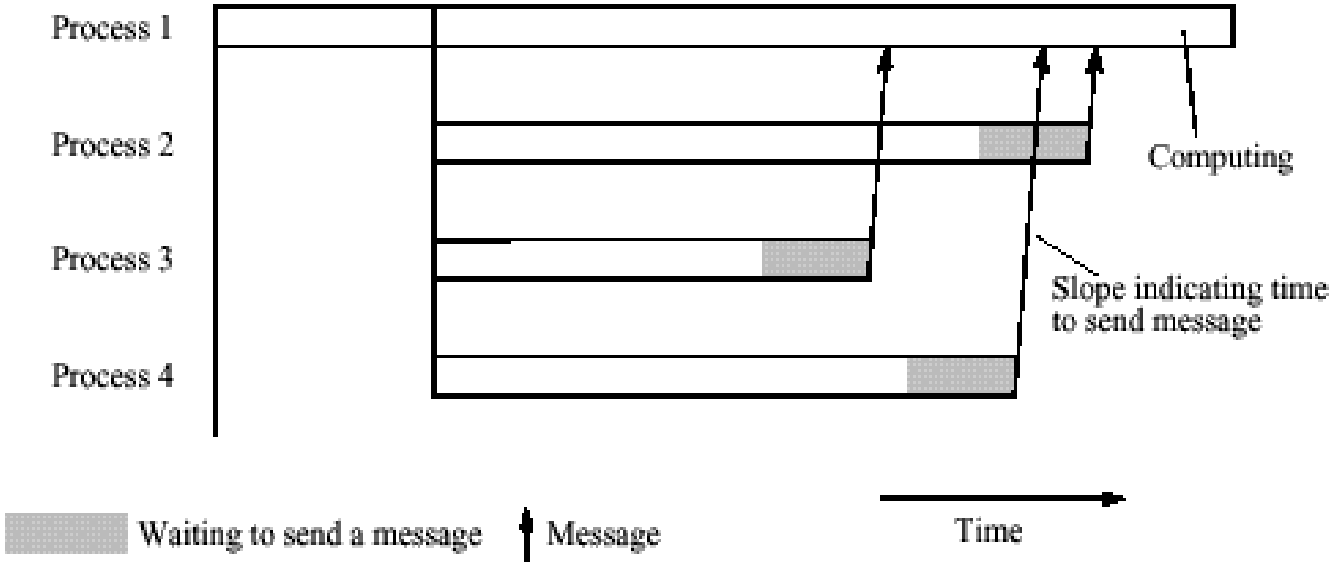
הזזת
נתונים

תהליך מס' 2

y

Recv(&y,1);

Space-Time Diagram of a Message-Passing Program



Process 1
Process 2
Process 3
Process 4
Computing
Slope indicating time to send message
Waiting to send a message
Message
Time

MPI - Message Passing Interface API

- MPI is a standard not an implementation
- Popular implementations are LAM and MPICH
- MPICH is installed under ***/usr/local/mpich***
- Always put in the code: `#include “mpi.h”`
- Compilation: `mpicc –o filename file.c`
- Execution: `mpirun –np N filename`
- Help: `man mpirun`

MPI Naming Conventions

`MPI_Xxxxx(parameter,...)`

Example: `MPI_Init(&argc,&argv)`

The First 4 Functions of MPI

- `MPI_Init`
- `MPI_Finalize`
- `MPI_Comm_size`
- `MPI_Comm_rank`

- …and don’t forget the
  `#include “mpi.h”`

The First 4 Functions Syntax

- **int MPI_Init(int argc, char *argv[])**
- **int MPI_Finilize()**
- **int MPI_Comm_size(MPI_Comm comm, int *size)**
- **int MPI_Comm_rank(MPI_Comm comm, int *rank)**

MPI Communicator

A **communicator** is a handle representing a group of processors that can communicate with one another.

The **communicator** name is required as an argument to all point-to-point and collective operations.

The **communicator** specified in the **send** and **receive** calls must agree for communication to take place.

Processors can communicate only if they share a **communicator**.

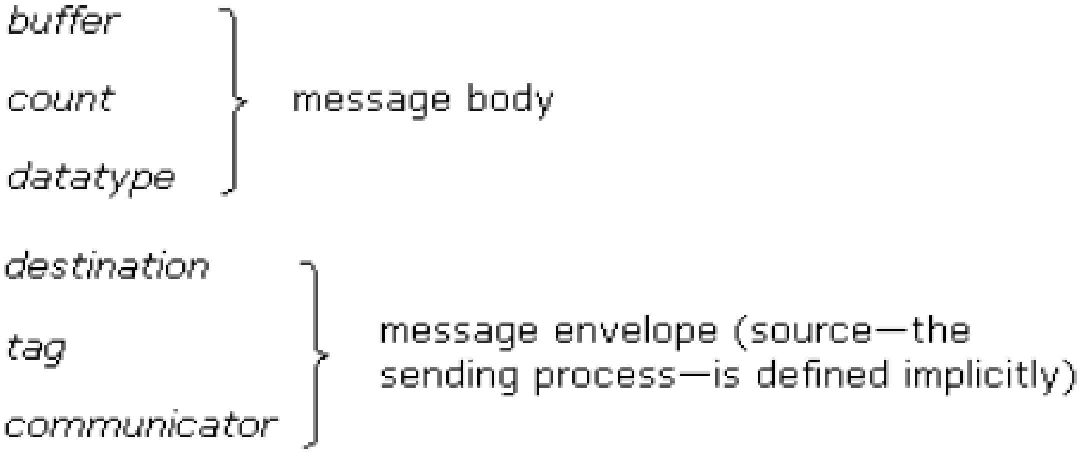
Basic *Point to Point* Functions

- **MPI_Send**
- **MPI_Recv**
- **MPI_Send**(void *buf, int count, MPI_Datatype datatype, int dest, int tag, MPI_Comm comm);

- **MPI_Recv**(void *buf, int count, MPI_Datatype datatype, int source, int tag, MPI_Comm comm, MPI_Status status);

MPI_Send

```
int MPI_Send(void *buf, int count, MPI_Datatype dtype,
int dest, int tag, MPI_Comm comm);
```



buffer

count

datatype

message body

destination

tag

communicator

message envelope (source—the
sending process—is defined implicitly)

תרגיל מס' 2

- הרצת תכנית קצרה ב- MPI:
Hello_World מקבילי
- כתוב תכנית בה כל מחשב יאמר שלום ויודיע את מספר התהליך שלו בריצה, לדוגמה:

```
Hello world from process 1 of 2
```

פתרון תרגיל מס' 2

```
// see more examples: /usr/local/mpich/examples

#include <stdio.h>
#include "mpi.h"
int main( argc, argv )
int  argc;
char **argv;
{
    int rank, size;
    MPI_Init( &argc, &argv );
    MPI_Comm_size( MPI_COMM_WORLD, &size );
    MPI_Comm_rank( MPI_COMM_WORLD, &rank );
    printf( "Hello world from process %d of %d\n",
rank, size );
    MPI_Finalize();
    return 0;
}
```

פתרון תרגיל מס' 2

- Note on syntax:

```
int main( argc, argv )
int  argc;
char **argv;
```

<u>Is Equivalent to:</u>

```
int main(int argc, char *argv[])
```

לאמץ את
הפורמט הזה

...and then...

```
MPI_Init(&argc,&argv);
```

Create a “machinefile”

```
nano ./machinefiles
```

Enter machines’ names, e.g.:

```
hobbit1
hobbit2
hobbit3
hobbit4
```

Save and exit

השמות המעודכנים של המחשבים
hobbit1 .. hobbit15

Our machine file

- Specify In the machine file short computers' names, e.g.

```
hobbit1
Hobbit2
```

and not:

```
hobbit1.ee.bgu.ac.il
hobbit2.ee.bgu.ac.il
```

Hello World - Execution

```
% mpicc -o helloworld helloworld.c
% mpirun -np 4 helloworld
```

Or

```
mpirun –np 4 –machinefile ./machinefile
 ./helloworld
Hello world from process 0 of 4
Hello world from process 3 of 4
Hello world from process 1 of 4
Hello world from process 2 of 4
%
```

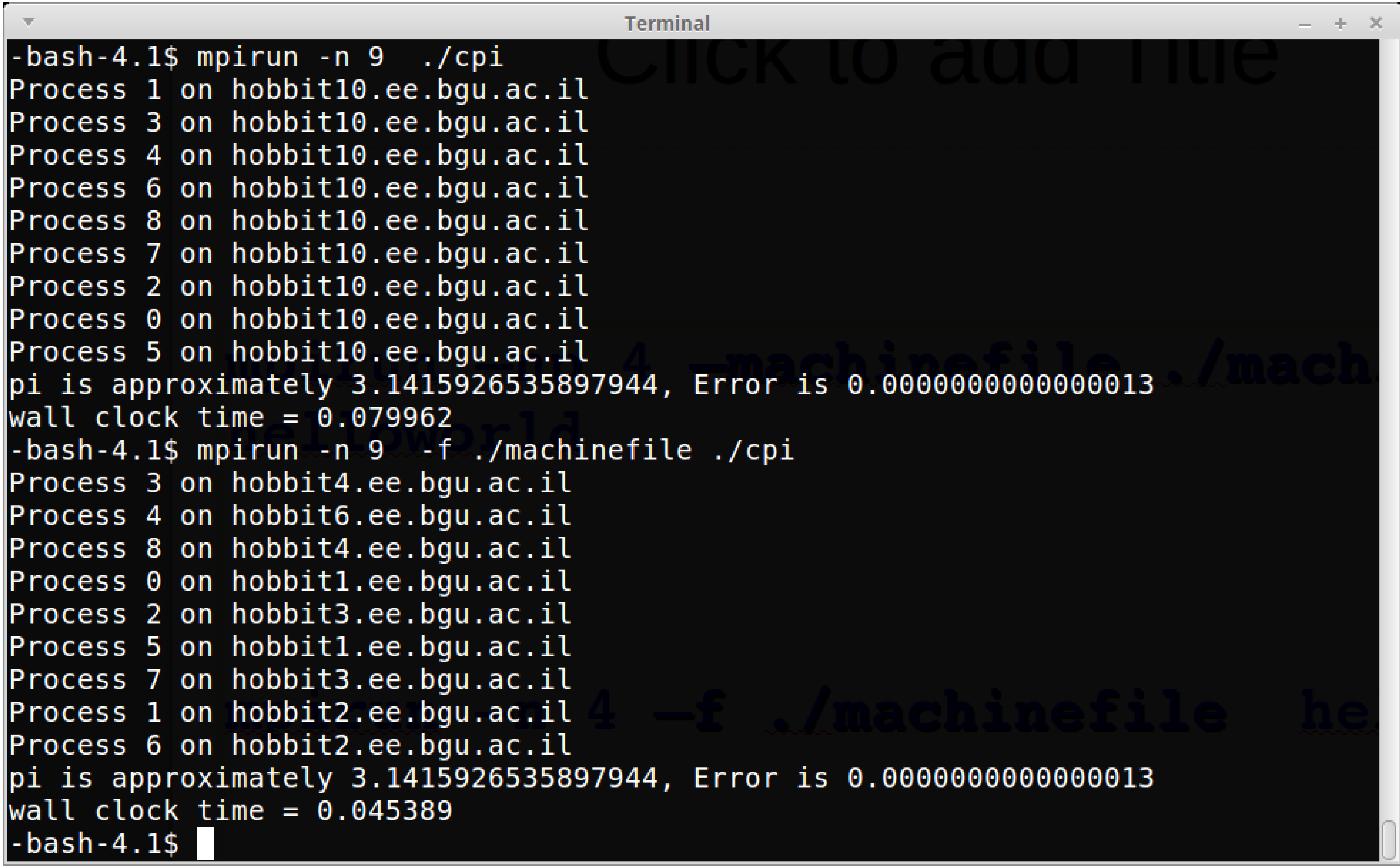
```
mpirun –np 4 –machinefile ./machinefile
helloworld
```

אפשר גם:

```
mpirun –n 4 –f ./machinefile
./helloworld
```

דוגמה בשקף הבא

mpirun ↔ mpiexec

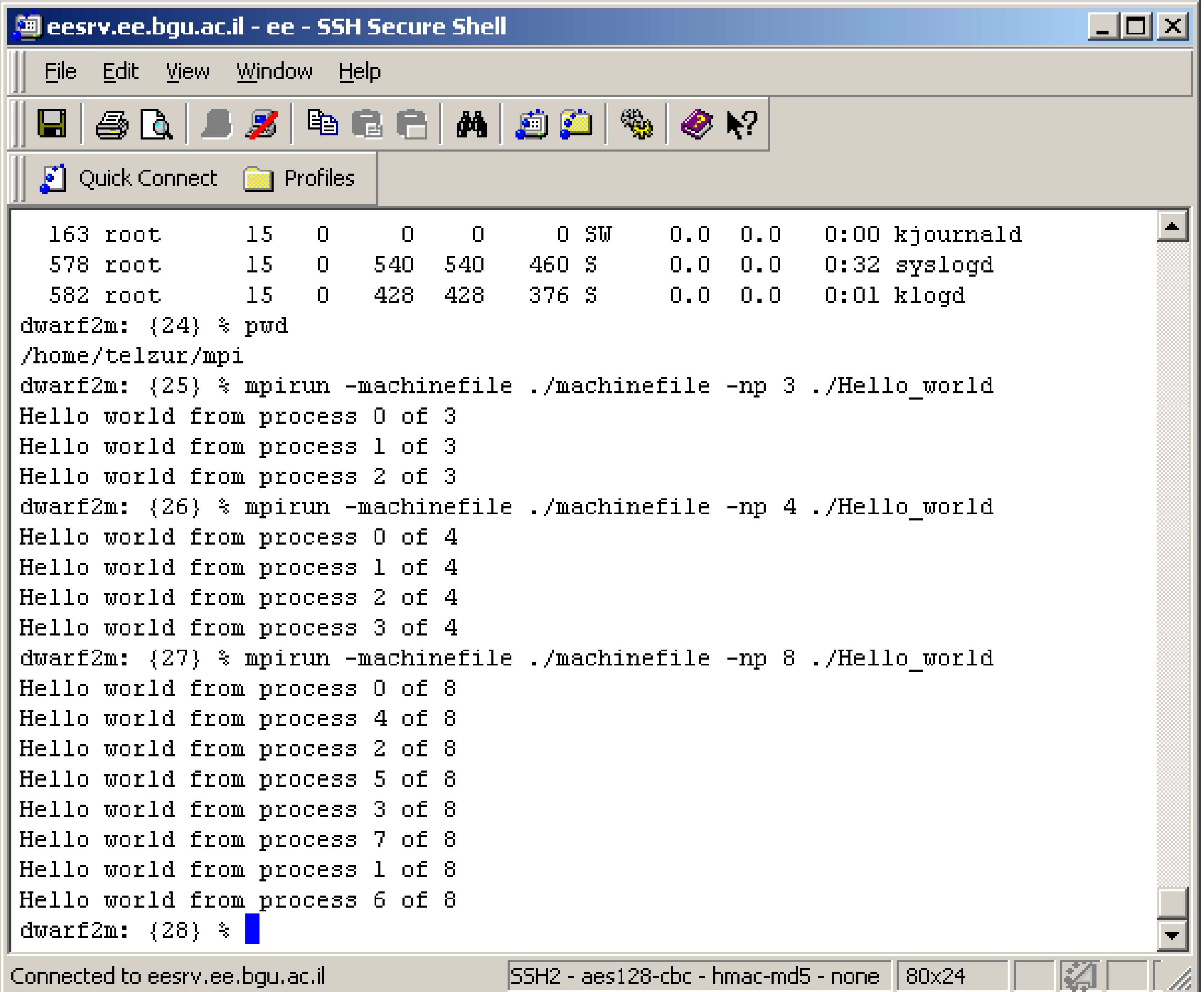


```
-bash-4.1$ mpirun -n 9  ./cpi
Process 1 on hobbit10.ee.bgu.ac.il
Process 3 on hobbit10.ee.bgu.ac.il
Process 4 on hobbit10.ee.bgu.ac.il
Process 6 on hobbit10.ee.bgu.ac.il
Process 8 on hobbit10.ee.bgu.ac.il
Process 7 on hobbit10.ee.bgu.ac.il
Process 2 on hobbit10.ee.bgu.ac.il
Process 0 on hobbit10.ee.bgu.ac.il
Process 5 on hobbit10.ee.bgu.ac.il
pi is approximately 3.1415926535897944, Error is 0.0000000000000013
wall clock time = 0.079962
-bash-4.1$ mpirun -n 9  -f ./machinefile ./cpi
Process 3 on hobbit4.ee.bgu.ac.il
Process 4 on hobbit6.ee.bgu.ac.il
Process 8 on hobbit4.ee.bgu.ac.il
Process 0 on hobbit1.ee.bgu.ac.il
Process 2 on hobbit3.ee.bgu.ac.il
Process 5 on hobbit1.ee.bgu.ac.il
Process 7 on hobbit3.ee.bgu.ac.il
Process 1 on hobbit2.ee.bgu.ac.il
Process 6 on hobbit2.ee.bgu.ac.il
pi is approximately 3.1415926535897944, Error is 0.0000000000000013
wall clock time = 0.045389
-bash-4.1$ 
```

machinefile

If no *machinefile* is specified you are running on the local machine

If a *machinefile* exists you are running on the machines specified in the file


eesrv.ee.bgu.ac.il - ee - SSH Secure Shell
File Edit View Window Help
Quick Connect
Profiles

```
  163 root      15   0     0    0     0 SW     0.0  0.0   0:00 kjournald
  578 root      15   0   540  540   460 S      0.0  0.0   0:32 syslogd
  582 root      15   0   428  428   376 S      0.0  0.0   0:01 klogd
dwarf2m: {24} % pwd
/home/telzur/mpi
dwarf2m: {25} % mpirun -machinefile ./machinefile -np 3 ./Hello_world
Hello world from process 0 of 3
Hello world from process 1 of 3
Hello world from process 2 of 3
dwarf2m: {26} % mpirun -machinefile ./machinefile -np 4 ./Hello_world
Hello world from process 0 of 4
Hello world from process 1 of 4
Hello world from process 2 of 4
Hello world from process 3 of 4
dwarf2m: {27} % mpirun -machinefile ./machinefile -np 8 ./Hello_world
Hello world from process 0 of 8
Hello world from process 4 of 8
Hello world from process 2 of 8
Hello world from process 5 of 8
Hello world from process 3 of 8
Hello world from process 7 of 8
Hello world from process 1 of 8
Hello world from process 6 of 8
dwarf2m: {28} %
```

Connected to eesrv.ee.bgu.ac.il

SSH2 - aes128-cbc - hmac-md5 - none

80x24

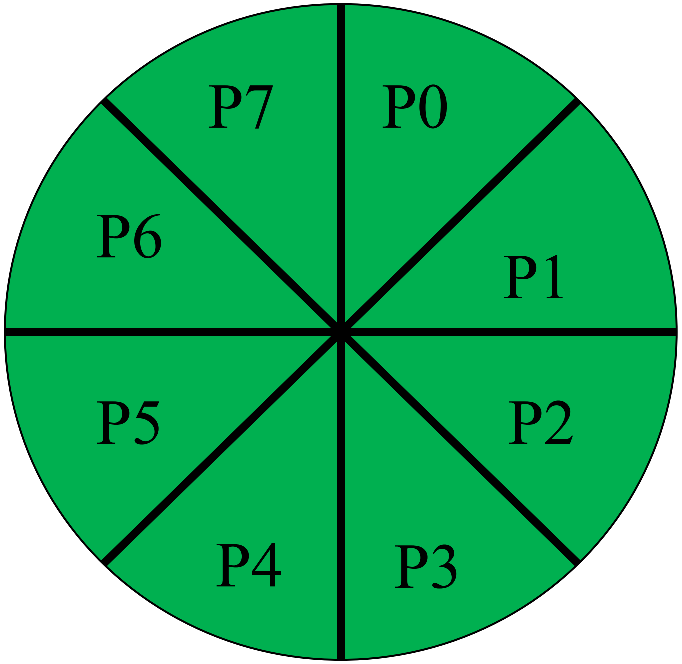
תרגיל מס' 3: חישוב

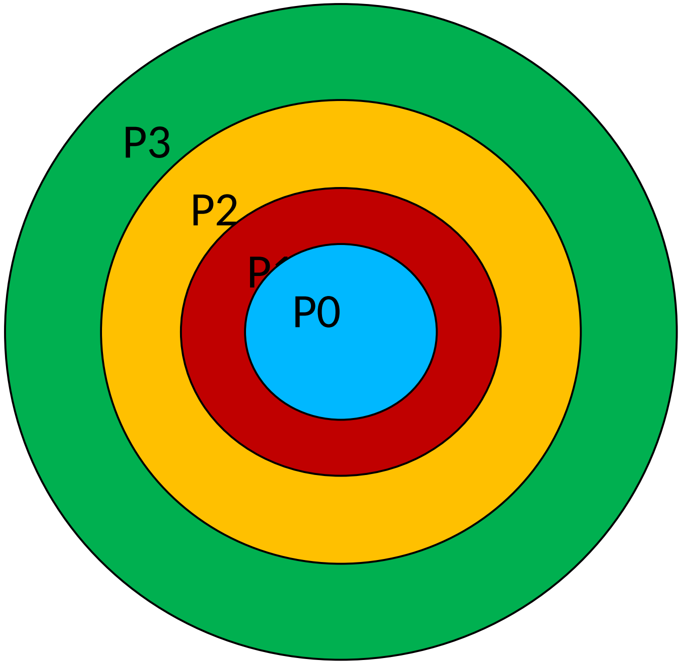
- חישוב באמצעות אינטגרציה

$$\int_0^1 \frac{1}{1+x^2}dx = \arctan(x)|_0^1 = \arctan(1) - \arctan(0) = \arctan(1) = \frac{\pi}{4}$$

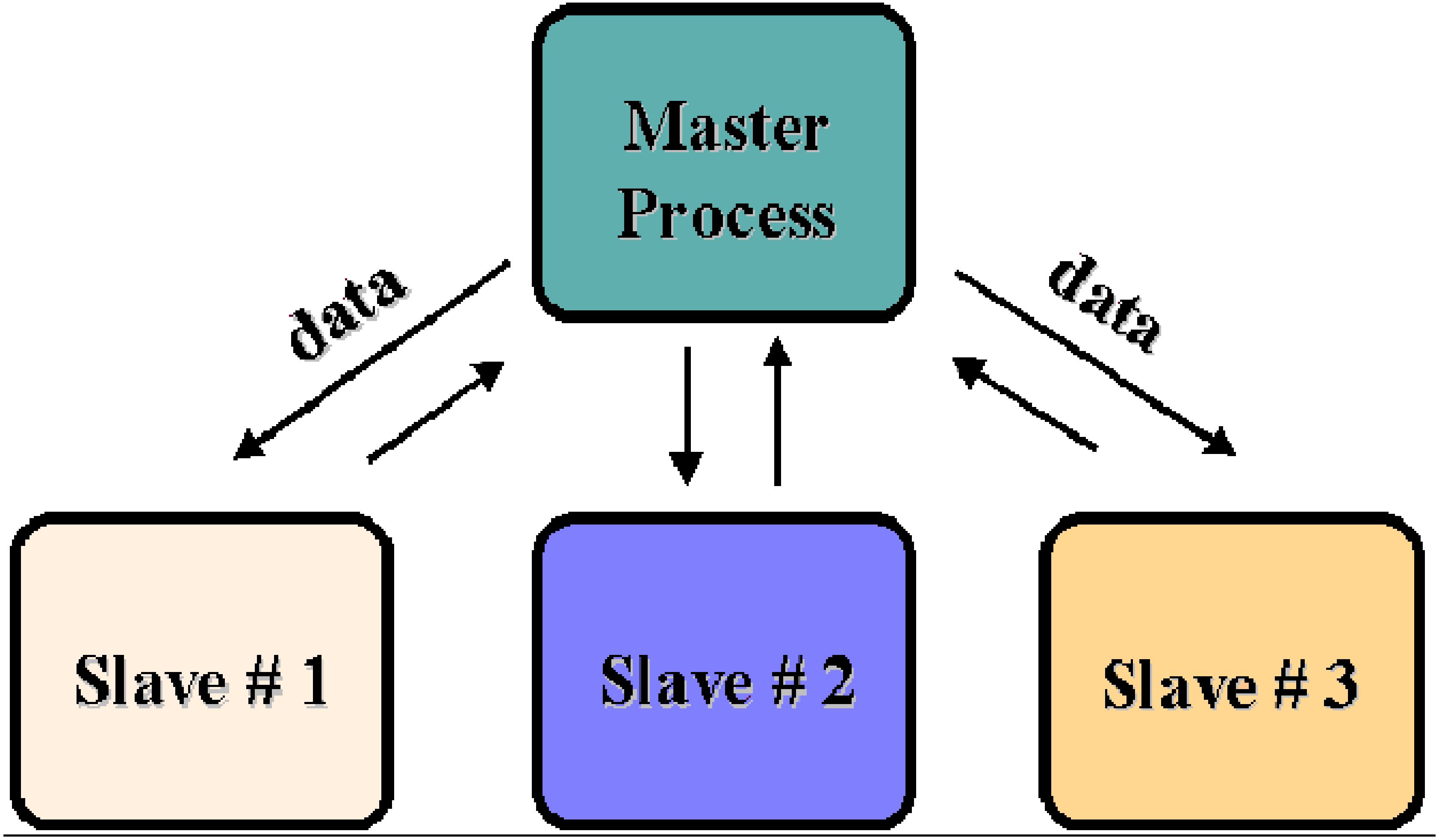
- נבצע סכימה על הפונקציה
- $f(x)=4/(1+x^2)$ בין 0 ל- 1 על-ידי חלוקת התחום ל-n מלבנים

Domain Decomposition, There is more than one way to decompose the circle. Which one will you choose?





Master/Workers



פתרון תרגיל מס' 3

- בפתרון השתמשנו בפונקציה למדידת זמן
- הנקראת : **MPI_Wtime**

- ראה דוגמא לפתרון התרגיל תחת:

/usr/local/mpich/examples/cpi.c

ניתן להוריד את התכנית מאתר הקורס וגם מהכתובת:

http://www.mcs.anl.gov/research/projects/mpi/usingmpi/examples/simplempi/cpi_c.htm

Copy this program to a location under your home directory where you have a write permission! Execute it from there

פתרון תרגיל מס' 3

```
#include "mpi.h"
#include <stdio.h>
#include <math.h>
double f(double a)
{
    return (4.0 / (1.0 + a*a));
}

void main(int argc, char *argv[])
{
    int done = 0, n, myid, numprocs, i;
    double PI25DT = 3.141592653589793238462643;
    double mypi, pi, h, sum, x;
    double startwtime, endwtime;
    int  namelen;
    char processor_name[MPI_MAX_PROCESSOR_NAME];
```

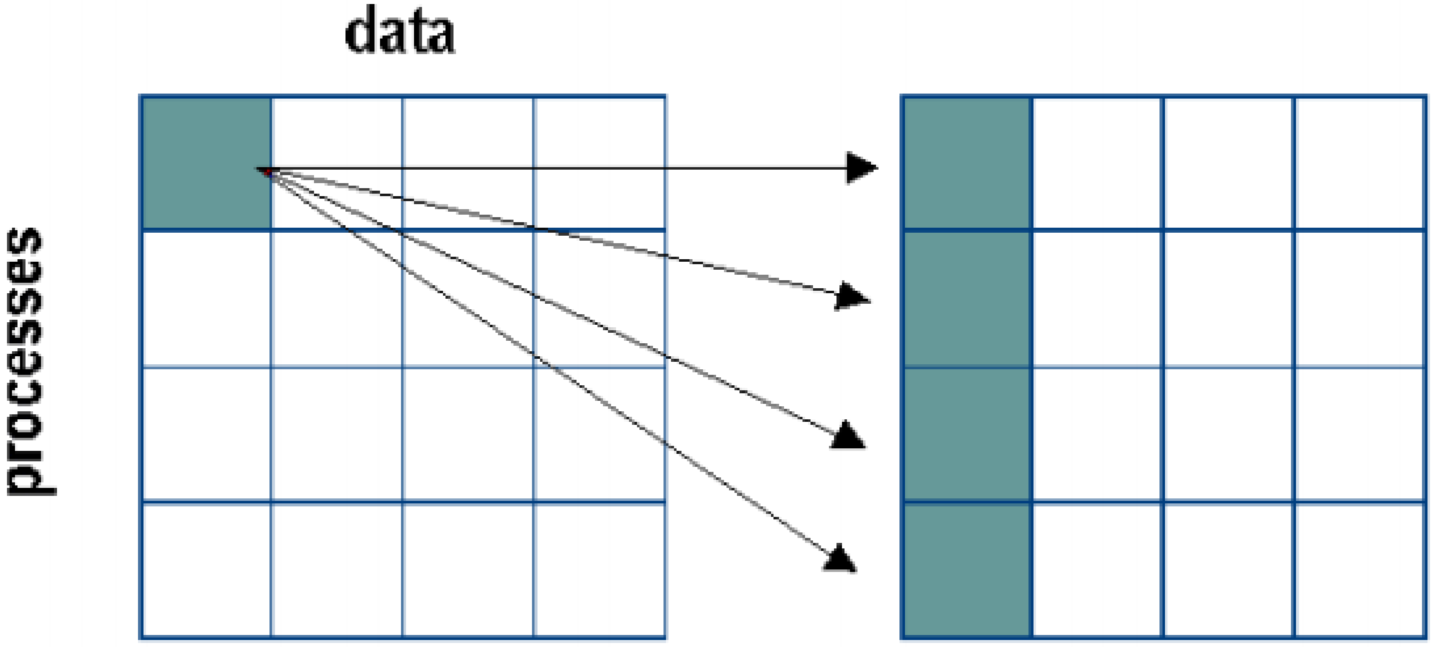
פתרון תרגיל מס' 3– המשך

```
    MPI_Init(&argc,&argv);
    MPI_Comm_size(MPI_COMM_WORLD,&numprocs);
    MPI_Comm_rank(MPI_COMM_WORLD,&myid);
    MPI_Get_processor_name(processor_name,&namelen);

    fprintf(stderr,"Process %d on %s\n",myid,
processor_name);
    fflush(stderr);

    n = 0;
```

פתרון תרגיל מס' 3 – המשך

```
while (!done)
    {
        if (myid == 0)
       {
   printf("Enter the number of intervals: (0 quits)
");
   fflush(stdout);
   scanf("%d",&n);
   startwtime = MPI_Wtime();
        }
        MPI_Bcast(&n, 1, MPI_INT, 0, MPI_COMM_WORLD);
        if (n == 0)
            done = 1;
        else
        {
```

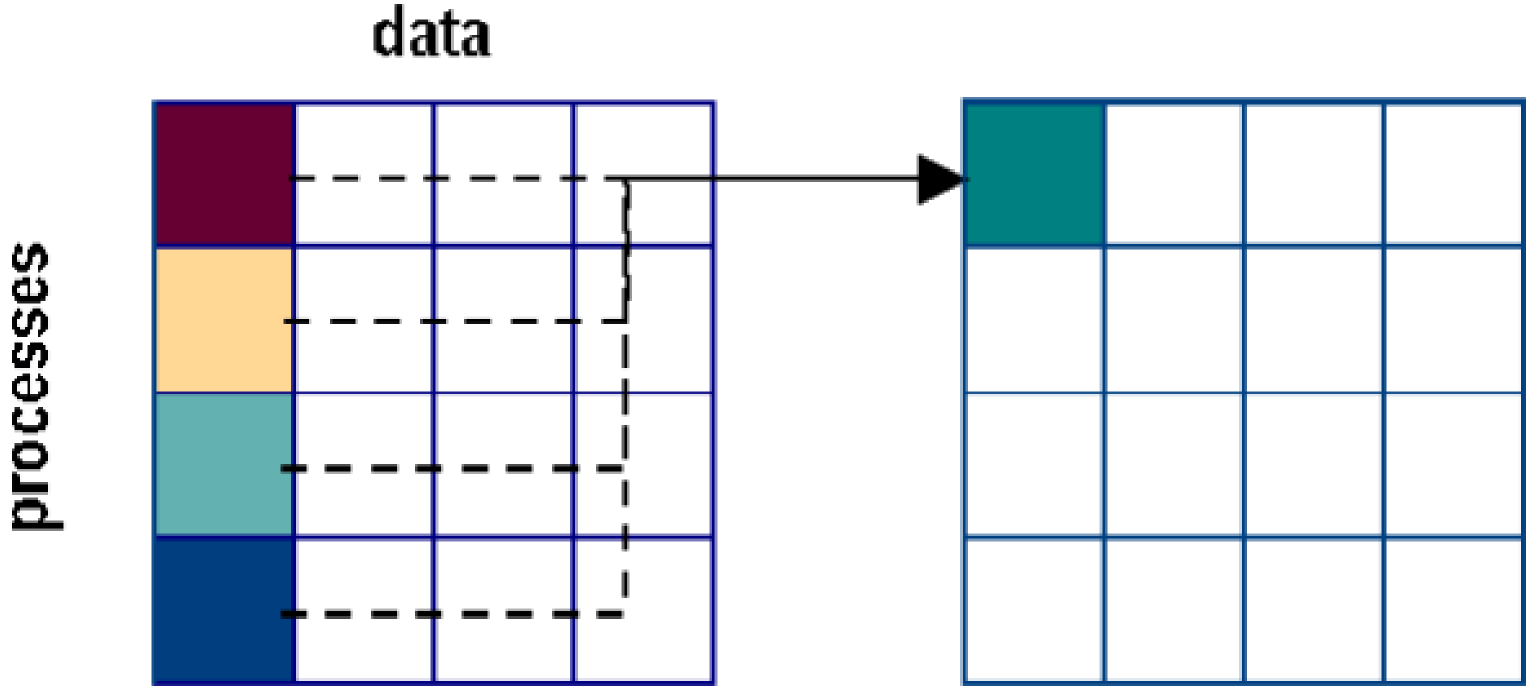
MPI **Broadcast** operation



פתרון תרגיל מס' 3 – המשך

```
h   = 1.0 / (double) n;
sum = 0.0;
for (i = myid + 1; i <= n; i += numprocs)
{
    x = h * ((double)i - 0.5);
    sum += f(x);
}
mypi = h * sum;
MPI_Reduce(&mypi, &pi, 1, MPI_DOUBLE,MPI_SUM,
0,MPI_COMM_WORLD);
```

MPI **reduction** operation



פתרון תרגיל מס' 3 – המשך

Basic *Collective* Functions

- **MPI_Bcast**
- **MPI_Reduce**
- The exact syntax:
- **MPI_Bcast**`(void *buf, int count, MPI_Datatype datatype, int root, MPI_Comm comm);`

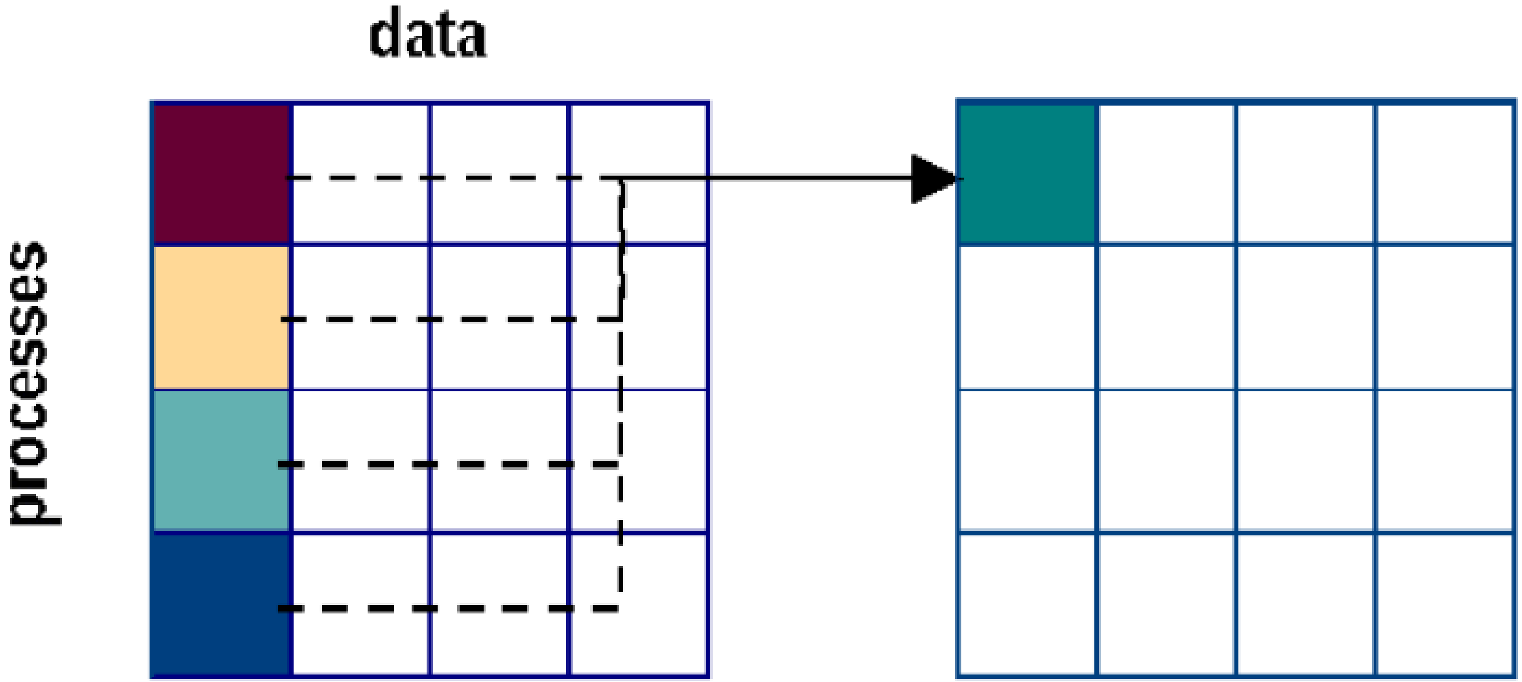
- **MPI_Reduce**`(void *sendbuf, void *recvbuf, int count, MPI_Datatype datatype, MPI_Op op, int root, MPI_Comm comm);`

פתרון תרגיל מס' 3 – המשך

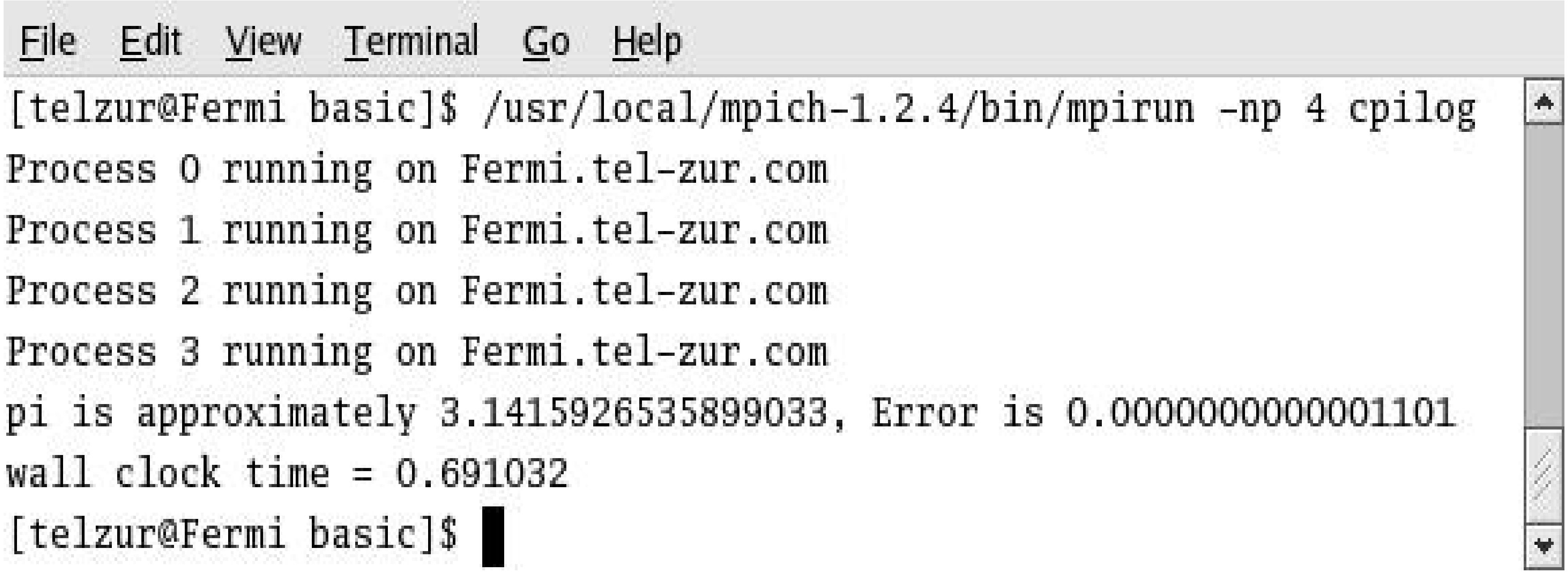
```
if (myid == 0)
   {
   printf("pi is approximately %.16f, Error is %.16f\n",
   pi, fabs(pi - PI25DT));
   endwtime = MPI_Wtime();
   printf("wall clock time = %f\n",endwtime-startwtime);
   }   /* end of if    */
  }    /* end of while */
    MPI_Finalize();
}      /* end of main  */
```

פתרון תרגיל מס' 3 – המשך

MPI Reduce



הרצת 4 תהליכים



```
[telzur@Fermi basic]$ /usr/local/mpich-1.2.4/bin/mpirun -np 4 cpilog
Process 0 running on Fermi.tel-zur.com
Process 1 running on Fermi.tel-zur.com
Process 2 running on Fermi.tel-zur.com
Process 3 running on Fermi.tel-zur.com
pi is approximately 3.1415926535899033, Error is 0.0000000000001101
wall clock time = 0.691032
[telzur@Fermi basic]$
```

How to execute `cpi.c`?

- `mkdir mpi`
- `cd mpi`
- Download cpi.c from the course website using wget:

```
wget http://tel-zur.net/teaching/bgu/pp/cpi.c
```

- `mpicc –o cpi cpi.c`
- `create a hostfile`
- `mpirun –np 4 –machinefile ./machinefile ./cpi`
- מטלות:
  - התחברות ב SSH
  - יצירת מחיצה + עריכת קובץ ושמירתו
  - קימפול עם MPI על-ידי mpicc
  - הרצת התכנית cpi על-ידי mpirun

Jumpshot (hobbits)

עדכון: לא פועל
על ההוביטים.
להפעיל את התכנה
מהמכונה הוירטואלית

- `mpecc -o cpi ./cpi.c -mpilog`

or

- `mpicc -o cpi ./cpi.c –llmpe –lmpe`

- **Then convert the `clog2` file by:**

```
clog2TOslog2 ./file_name.clog2
```

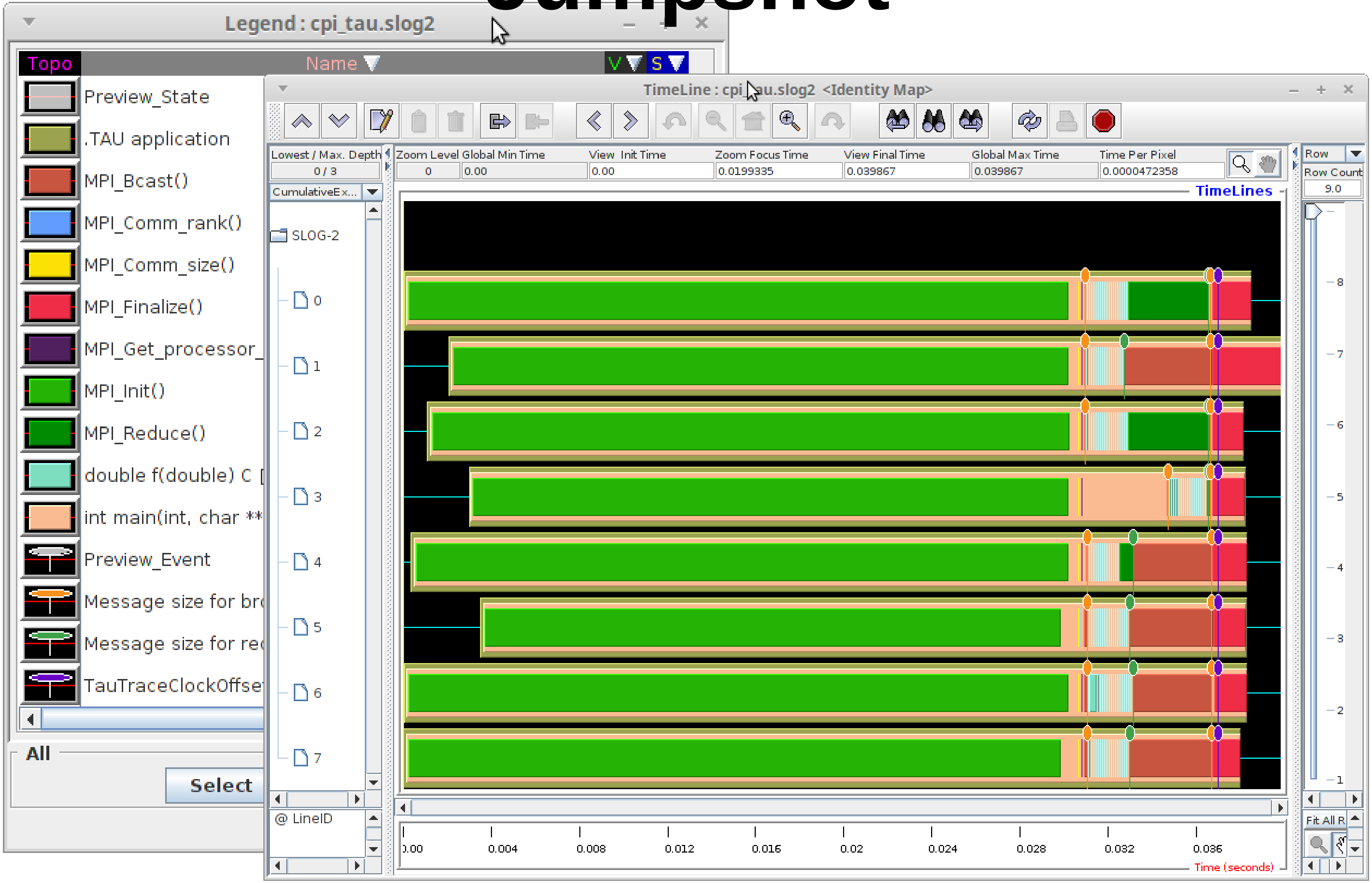
- **Finally view the profiling:**

```
jumpshot file_name.slog2
```

Jumpshot (VM)

```
tau_cc.sh -o cpi_tau ./cpi.c # compile with profiling and tracing support

mpirun -np 8 ./cpi_tau       # execute and generate profiling

paraprof                     # paraprof visualizing

tau_treemerge.pl             # merge the mpi tasks profiling files into combined files

tau2slog2 tau.trc tau.edf    # prepare slog2 file

jumpshot ./tau.slog2         # view tracing in jumpshot
```

Jumpshot



Lowest / Max. Depth: 0 / 3

Zoom Level	Global Min Time	View Init Time	Zoom Focus Time	View Final Time	Global Max Time	Time Per Pixel
0	0.00	0.00	0.0199335	0.039867	0.039867	0.0000472358

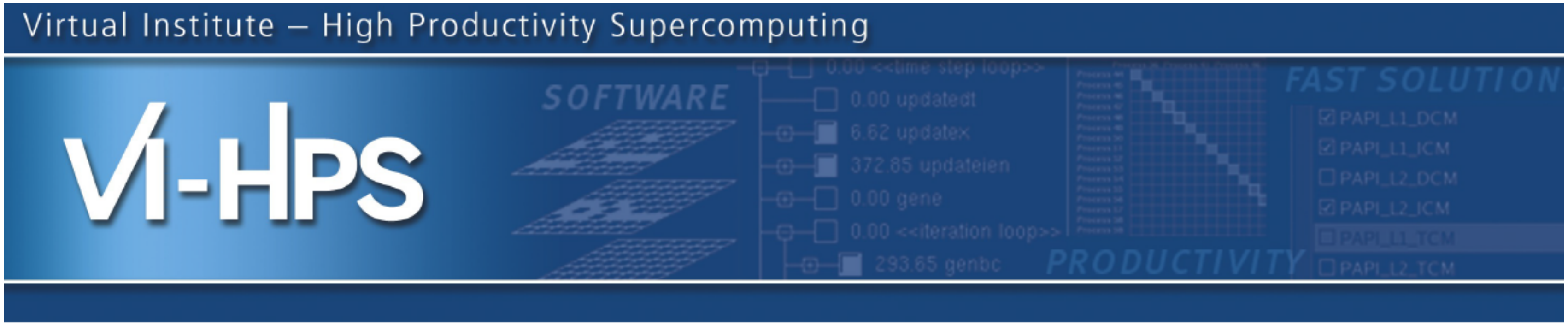
Row Count: 9.0

Scalasca

- Is installed on the *hobbits*

- VI-HPS *http://www.vi-hps.org/*
- LiveDVD: http://www.vi-hps.org/training/live-iso/
- יש להוריד ולהתקין תוכנה זו כ"אורחת" בתוכנת הוירטואליזציה Virtualbox
- Quick Reference Guide: http://apps.fz-juelich.de/scalasca/releases/scalasca/1.4/docs/QuickReference.pdf

Next slides are from:



scalasca

Scalable performance analysis of large-scale parallel applications

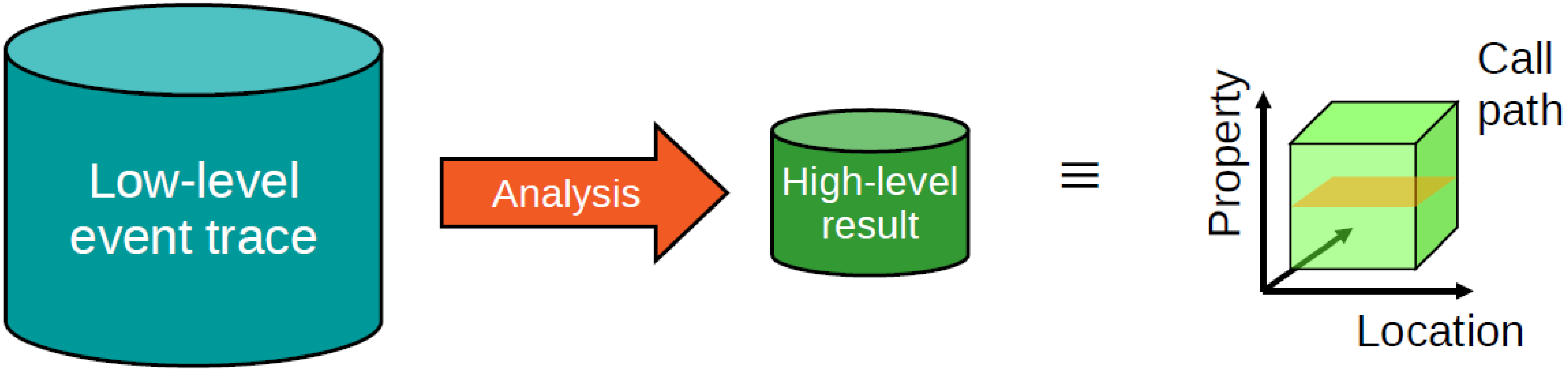
Brian Wylie & Markus Geimer
Jülich Supercomputing Centre
scalasca@fz-juelich.de

August 2012

- Profile analysis
  - Summary of aggregated metrics
    - per function/callpath and/or per process/thread
  - Most tools (can) generate and/or present such profiles
    - but they do so in *very* different ways, often from event traces!
  - e.g., gprof, mpiP, ompP, ***Scalasca***, TAU, Vampir, ...
- Time-line analysis
  - Visual representation of the space/time sequence of events
  - Requires an execution trace
  - e.g., Vampir, Paraver, JumpShot, Intel TAC, Sun Studio, ...
- Pattern analysis
  - Search for event sequences characteristic of inefficiencies
  - Can be done manually, e.g., via visual time-line analysis
  - or automatically, e.g., KOJAK, ***Scalasca***, Periscope, ...

Automatic trace analysis

- Idea
  - Automatic search for patterns of inefficient behaviour
  - Classification of behaviour & quantification of significance



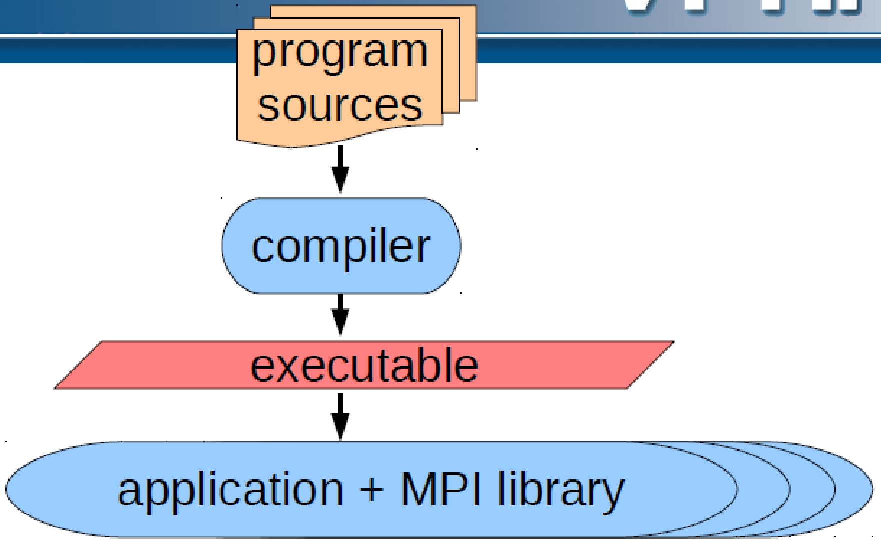
  - Guaranteed to cover the entire event trace
  - Quicker than manual/visual trace analysis
  - Parallel replay analysis exploits memory & processors to deliver scalability

Scalasca features

- Open source, New BSD license
- Portable
  - Cray XT, IBM BlueGene, IBM SP & blade clusters, NEC SX, SGI Altix, SiCortex, Solaris & Linux clusters, ...
- Supports parallel programming paradigms & languages
  - MPI, OpenMP & hybrid OpenMP+MPI
  - Fortran, C, C++
- Integrated instrumentation, measurement & analysis toolset
  - Automatic and/or manual customizable instrumentation
  - Runtime summarization (aka profiling)
  - Automatic event trace analysis
  - Analysis report exploration & manipulation

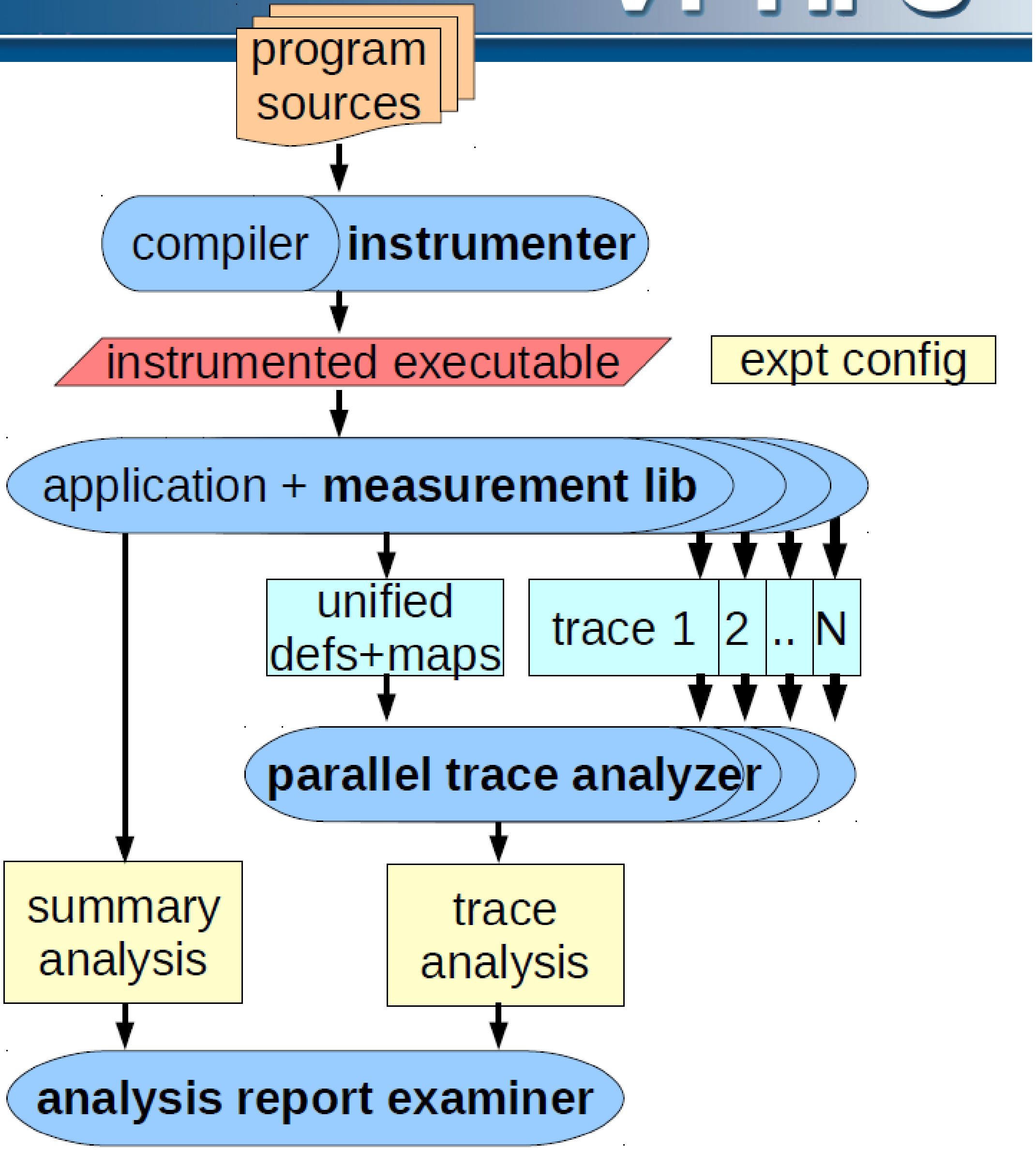
Generic MPI application build & run

- Application code compiled & linked into executable using MPICC/CXX/FC
- Launched with MPIEXEC
- Application processes interact via MPI library



Scalasca toolset components

- Scalasca instrumenter = SKIN
- Scalasca measurement collector & analyzer = SCAN
- Scalasca analysis report examiner = SQUARE

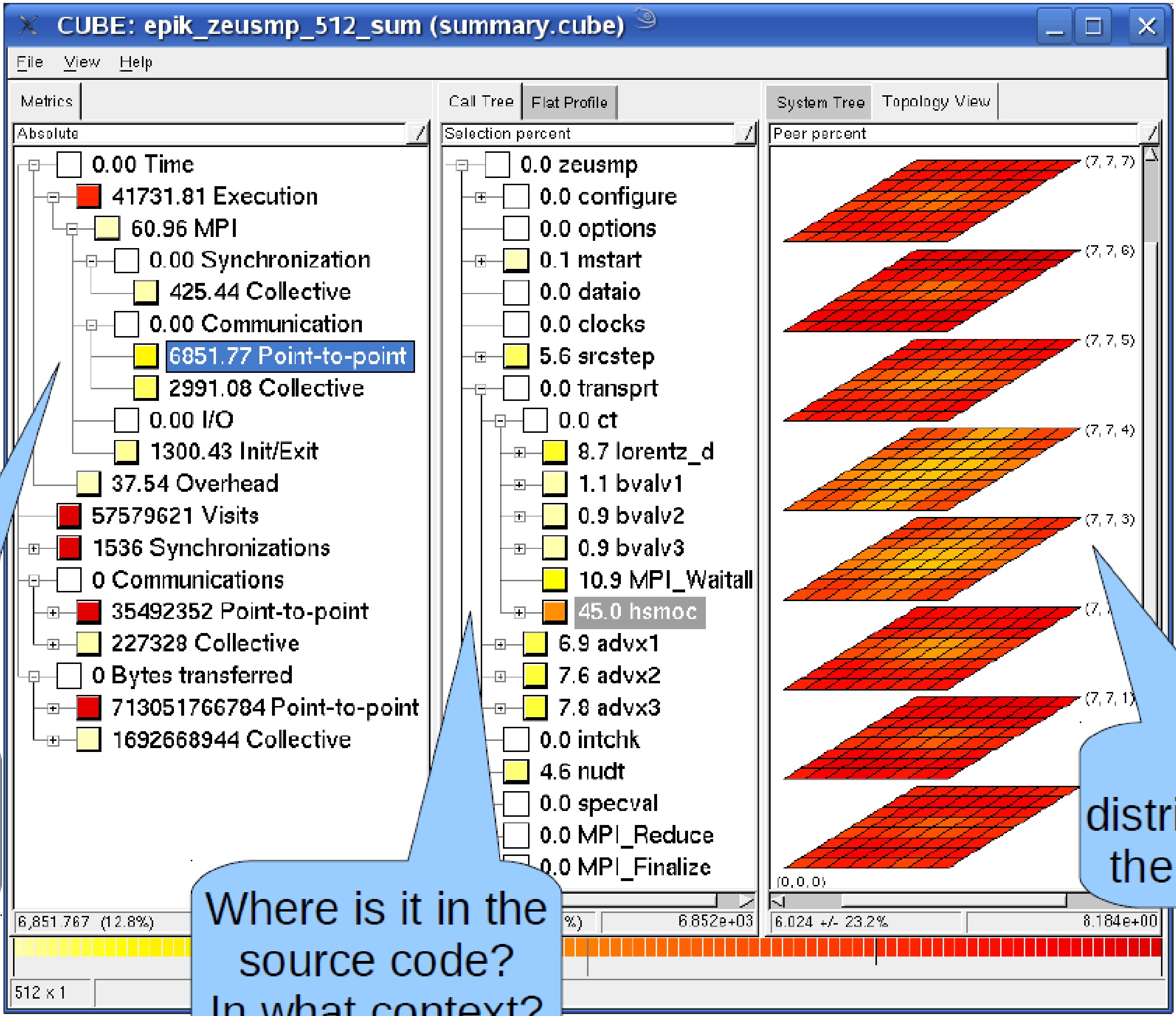


scalasca

- One command for everything

```
% scalasca
  Scalasca 1.4
  Toolset for scalable performance analysis of large-scale apps
  usage: scalasca [-v][-n] {action}
  1. prepare application objects and executable for measurement:
     scalasca -instrument <compile-or-link-command>    # skin
  2. run application under control of measurement system:
     scalasca -analyze <application-launch-command>    # scan
  3. post-process & explore measurement analysis report:
     scalasca -examine <experiment-archive|report>     # square

  [-h] show quick reference guide (only)
```

Scalasca analysis report explorer (summary)



```
0.00 Time
  41731.81 Execution
    60.96 MPI
      0.00 Synchronization
        425.44 Collective
      0.00 Communication
        6851.77 Point-to-point
        2991.08 Collective
      0.00 I/O
      1300.43 Init/Exit
    37.54 Overhead
57579621 Visits
1536 Synchronizations
0 Communications
  35492352 Point-to-point
  227328 Collective
0 Bytes transferred
  713051766784 Point-to-point
  1692668944 Collective
```

```
0.0 zeusmp
  0.0 configure
  0.0 options
  0.1 mstart
  0.0 dataio
  0.0 clocks
  5.6 srcstep
  0.0 transprt
    0.0 ct
      8.7 lorentz_d
      1.1 bvalv1
      0.9 bvalv2
      0.9 bvalv3
      10.9 MPI_Waitall
      45.0 hsmoc
    6.9 advx1
    7.6 advx2
    7.8 advx3
  0.0 intchk
  4.6 nudt
  0.0 specval
  0.0 MPI_Reduce
  0.0 MPI_Finalize
```

6,851.767 (12.8%)

%) 6.852e+03

6.024 +/- 23.2% 8.184e+00

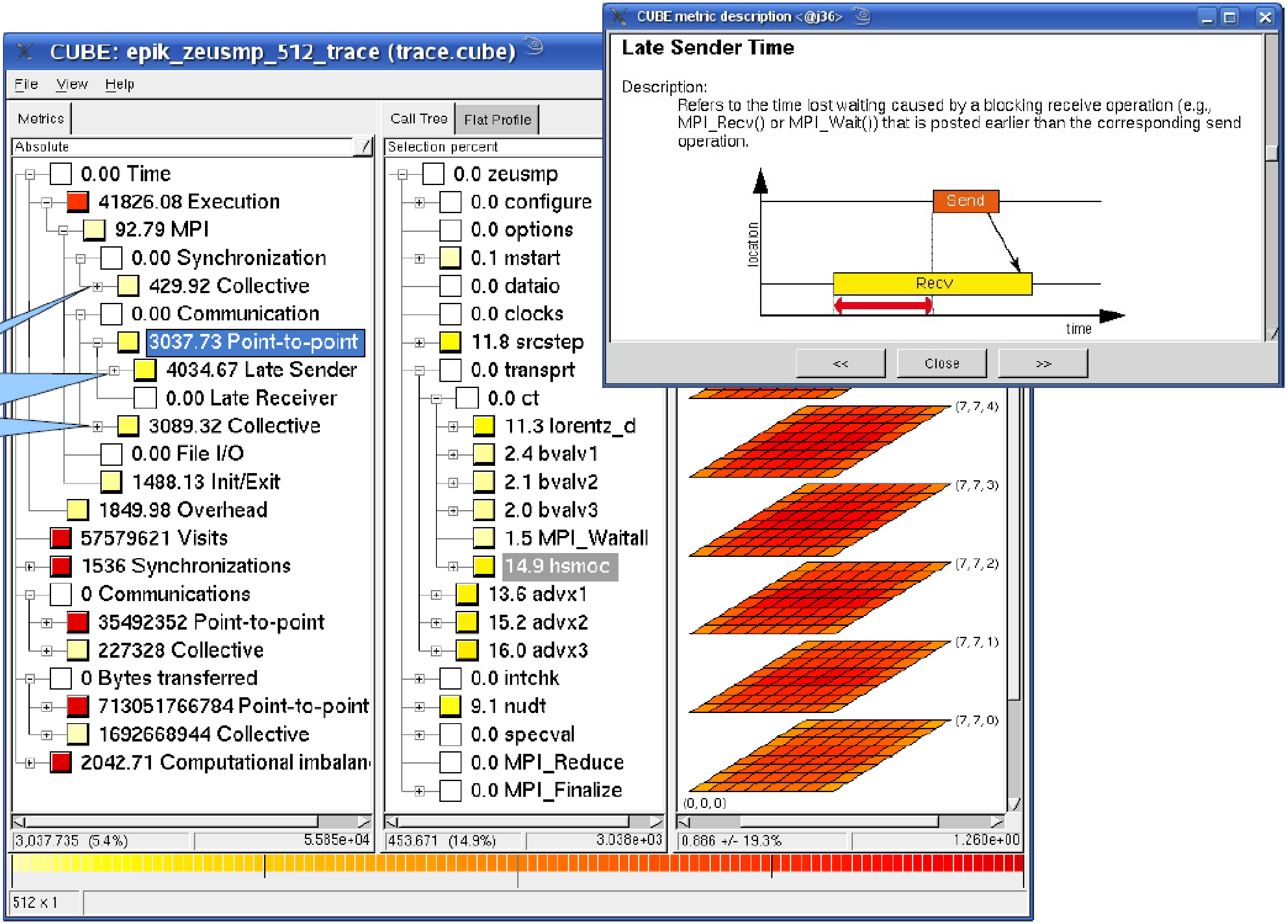
512 x 1

What kind of performance problem?

Where is it in the source code? In what context?

How is it distributed across the processes?

Scalasca analysis report explorer (trace)



Absolute

- 0.00 Time
  - 41826.08 Execution
    - 92.79 MPI
      - 0.00 Synchronization
        - 429.92 Collective
      - 0.00 Communication
        - 3037.73 Point-to-point
          - 4034.67 Late Sender
          - 0.00 Late Receiver
        - 3089.32 Collective
      - 0.00 File I/O
      - 1488.13 Init/Exit
  - 1849.98 Overhead
- 57579621 Visits
- 1536 Synchronizations
- 0 Communications
  - 35492352 Point-to-point
  - 227328 Collective
- 0 Bytes transferred
  - 713051766784 Point-to-point
  - 1692668944 Collective
- 2042.71 Computational imbalan

Selection percent

- 0.0 zeusmp
  - 0.0 configure
  - 0.0 options
  - 0.1 mstart
  - 0.0 dataio
  - 0.0 clocks
  - 11.8 srcstep
  - 0.0 transprt
    - 0.0 ct
      - 11.3 lorentz_d
      - 2.4 bvalv1
      - 2.1 bvalv2
      - 2.0 bvalv3
      - 1.5 MPI_Waitall
      - 14.9 hsmoc
    - 13.6 advx1
    - 15.2 advx2
    - 16.0 advx3
  - 0.0 intchk
  - 9.1 nudt
  - 0.0 specval
  - 0.0 MPI_Reduce
  - 0.0 MPI_Finalize

3,037.735 (5.4%) 5.585e+04

453.671 (14.9%) 3.038e+03

0.686 +/- 19.3% 1.260e+00

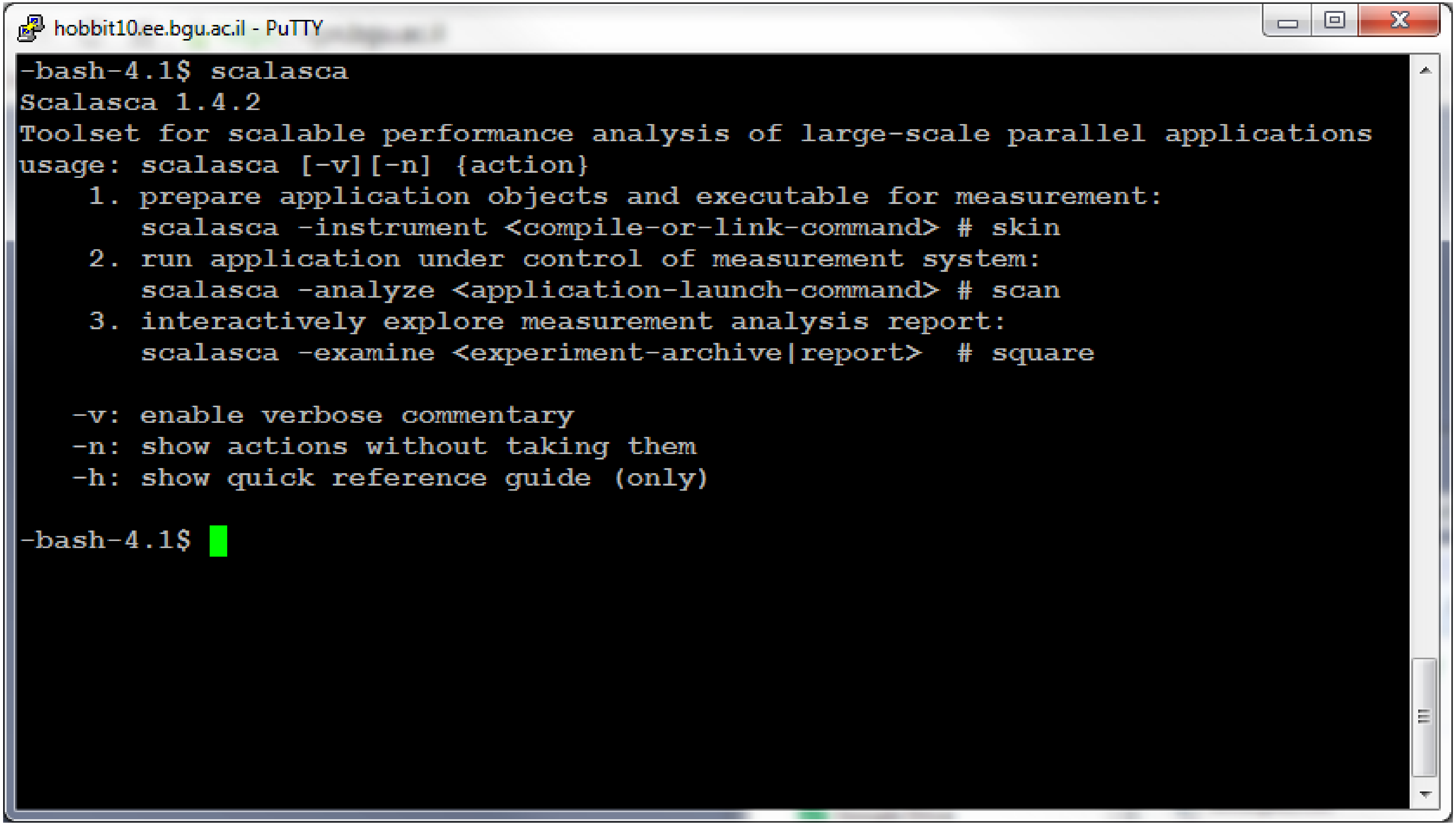
512 x 1

Late Sender Time

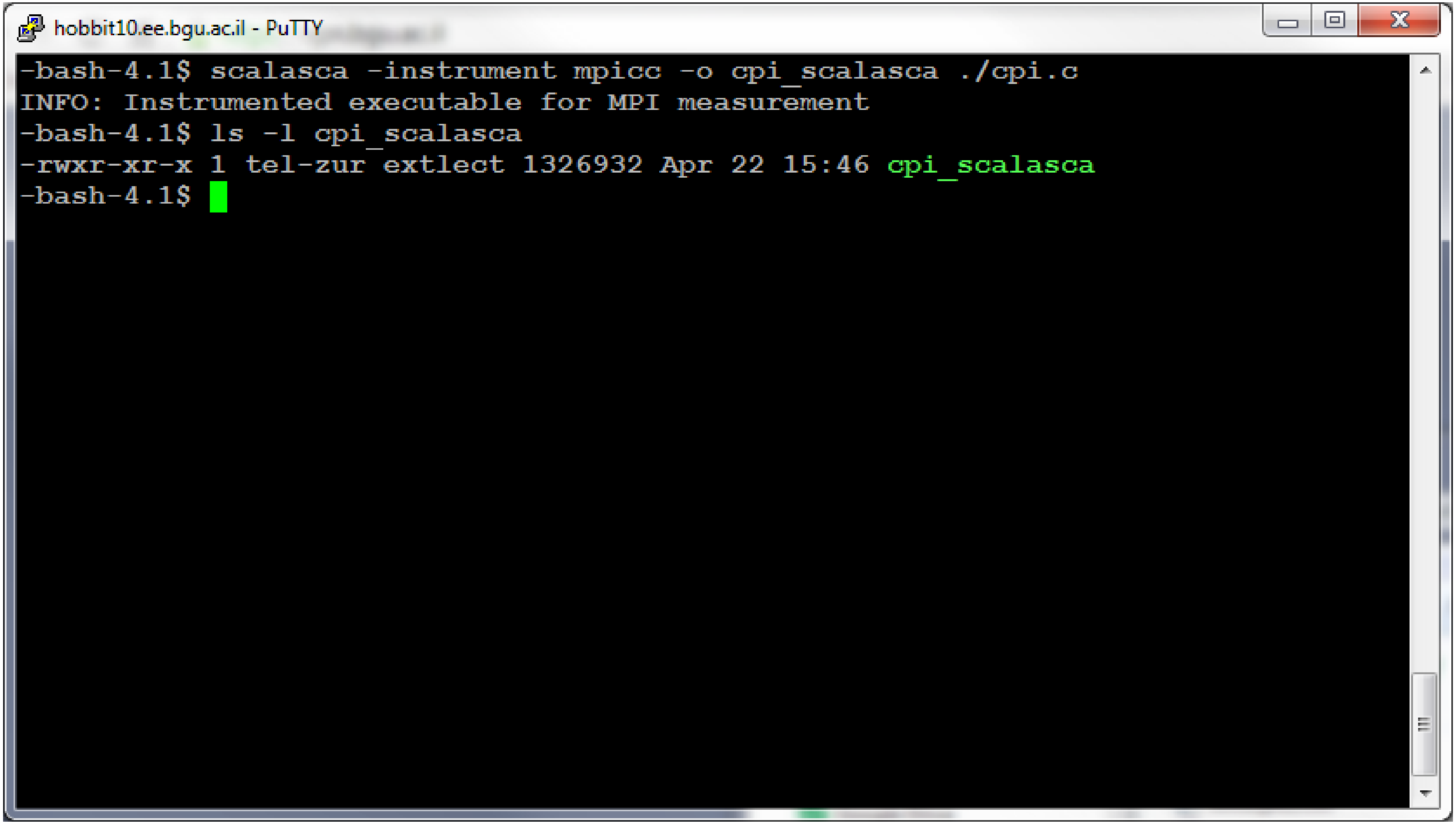
Description:

Refers to the time lost waiting caused by a blocking receive operation (e.g., MPI_Recv() or MPI_Wait()) that is posted earlier than the corresponding send operation.

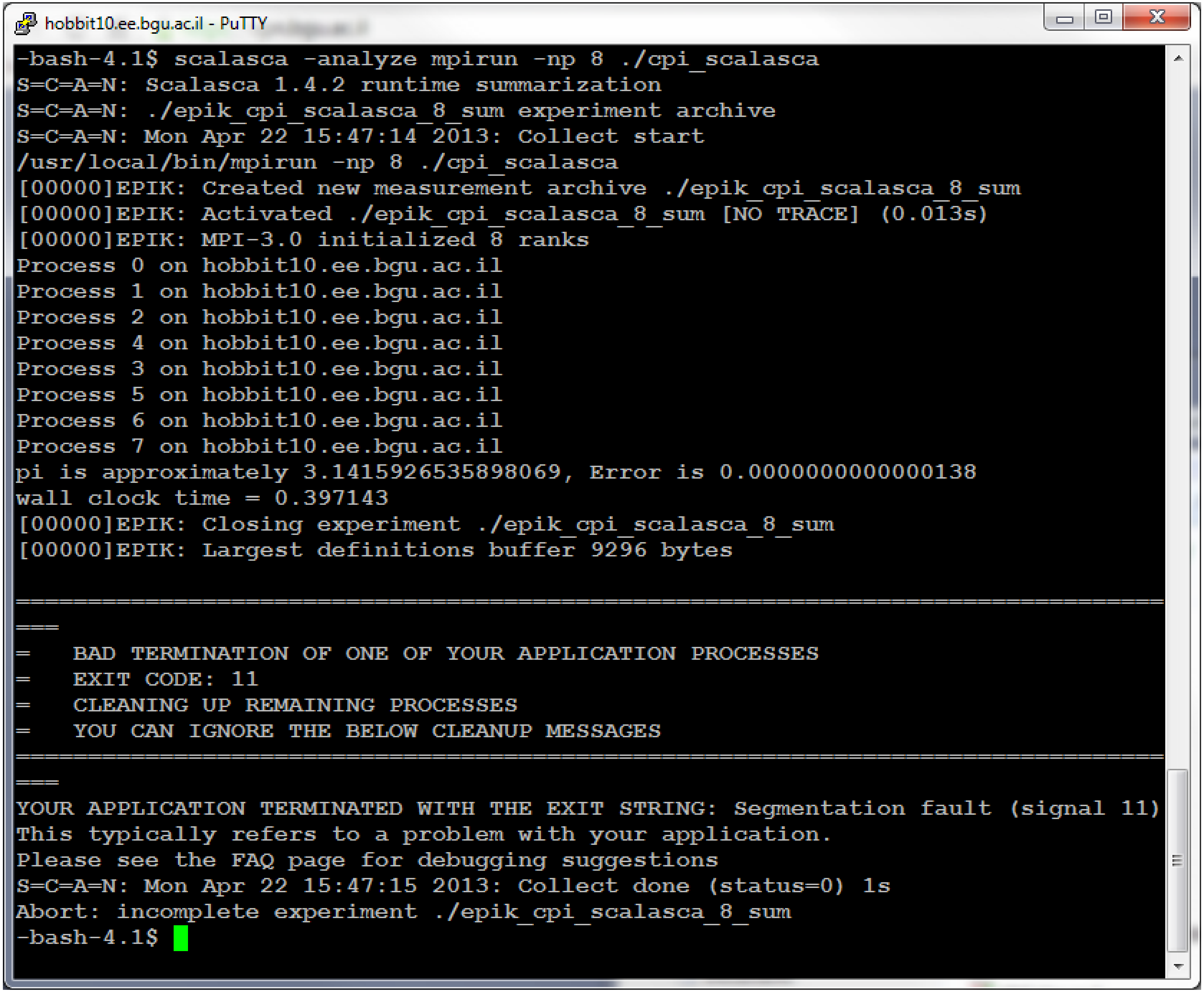
Additional metrics determined from trace



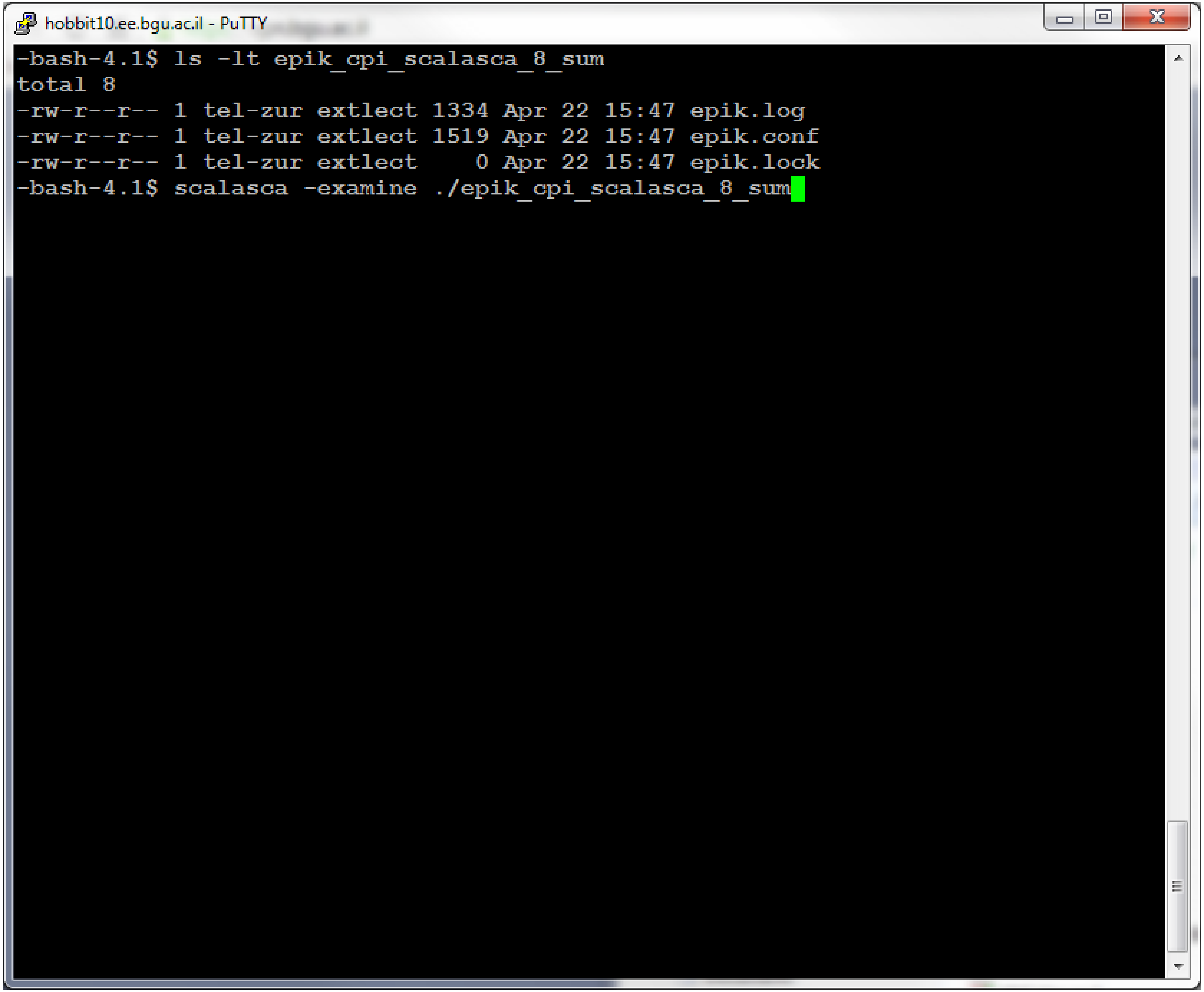
```
-bash-4.1$ scalasca
Scalasca 1.4.2
Toolset for scalable performance analysis of large-scale parallel applications
usage: scalasca [-v][-n] {action}
    1. prepare application objects and executable for measurement:
       scalasca -instrument <compile-or-link-command> # skin
    2. run application under control of measurement system:
       scalasca -analyze <application-launch-command> # scan
    3. interactively explore measurement analysis report:
       scalasca -examine <experiment-archive|report>  # square

   -v: enable verbose commentary
   -n: show actions without taking them
   -h: show quick reference guide (only)

-bash-4.1$
```



```
-bash-4.1$ scalasca -instrument mpicc -o cpi_scalasca ./cpi.c
INFO: Instrumented executable for MPI measurement
-bash-4.1$ ls -l cpi_scalasca
-rwxr-xr-x 1 tel-zur extlect 1326932 Apr 22 15:46 cpi_scalasca
-bash-4.1$
```



```
-bash-4.1$ scalasca -analyze mpirun -np 8 ./cpi_scalasca
S=C=A=N: Scalasca 1.4.2 runtime summarization
S=C=A=N: ./epik_cpi_scalasca_8_sum experiment archive
S=C=A=N: Mon Apr 22 15:47:14 2013: Collect start
/usr/local/bin/mpirun -np 8 ./cpi_scalasca
[00000]EPIK: Created new measurement archive ./epik_cpi_scalasca_8_sum
[00000]EPIK: Activated ./epik_cpi_scalasca_8_sum [NO TRACE] (0.013s)
[00000]EPIK: MPI-3.0 initialized 8 ranks
Process 0 on hobbit10.ee.bgu.ac.il
Process 1 on hobbit10.ee.bgu.ac.il
Process 2 on hobbit10.ee.bgu.ac.il
Process 4 on hobbit10.ee.bgu.ac.il
Process 3 on hobbit10.ee.bgu.ac.il
Process 5 on hobbit10.ee.bgu.ac.il
Process 6 on hobbit10.ee.bgu.ac.il
Process 7 on hobbit10.ee.bgu.ac.il
pi is approximately 3.1415926535898069, Error is 0.0000000000000138
wall clock time = 0.397143
[00000]EPIK: Closing experiment ./epik_cpi_scalasca_8_sum
[00000]EPIK: Largest definitions buffer 9296 bytes

================================================================================
===
=   BAD TERMINATION OF ONE OF YOUR APPLICATION PROCESSES
=   EXIT CODE: 11
=   CLEANING UP REMAINING PROCESSES
=   YOU CAN IGNORE THE BELOW CLEANUP MESSAGES
================================================================================
===
YOUR APPLICATION TERMINATED WITH THE EXIT STRING: Segmentation fault (signal 11)
This typically refers to a problem with your application.
Please see the FAQ page for debugging suggestions
S=C=A=N: Mon Apr 22 15:47:15 2013: Collect done (status=0) 1s
Abort: incomplete experiment ./epik_cpi_scalasca_8_sum
-bash-4.1$
```



```
-bash-4.1$ ls -lt epik_cpi_scalasca_8_sum
total 8
-rw-r--r-- 1 tel-zur extlect 1334 Apr 22 15:47 epik.log
-rw-r--r-- 1 tel-zur extlect 1519 Apr 22 15:47 epik.conf
-rw-r--r-- 1 tel-zur extlect    0 Apr 22 15:47 epik.lock
-bash-4.1$ scalasca -examine ./epik_cpi_scalasca_8_sum
```

That's it!