

Korištenje računarstva visokih performansi (HPC)

Petar Stipanović
pero@pmfst.hr

Primjeri naredbi označeni su ► **plavo**, dok su **crveno** istaknuti neki dijelovi.
[NEobavezni] dijelovi (opcionalni) navedeni su u uglatim zagrada.
PODESIVI dijelovi opisni su velikim slovima.

UniST-PMF
2023

1 Računalne mreže

- Što čini računalnu mrežu?
- Serveri / računalni klasteri
- Komunikacijski protokoli
- SSH protokol

2 Povezivanje sa serverom

- Serveri dostupni Odjelu za fiziku
- SSH/SFTP programi
- Upload
- SSH

3 Linux/sh

- Adrese
- Upravljanje podacima
- Rad s tekstualnim editorima
- Upravljanje programima

4 Server Unist-Phy

5 Klaster Isabella

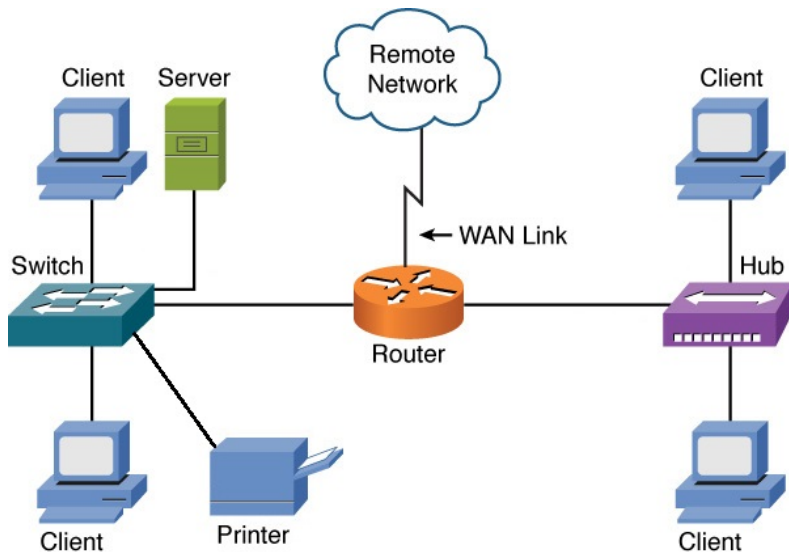
- Sustavi za upravljanje poslovima
- Tehničke specifikacije Isabelle
- Kreiranje skripte
- Pristup klasteru
- Pokretanje poslova
- Kontrola poslova

Što čini računalnu mrežu?

Računalne mreže

- **Računalnu mrežu** čini sustav međusobno povezanih računala i perifernih priključnih uređaja radi brze i jednostavne razmjene podataka.
- **Razdjelnik** (engl. *hub*) u jednoj točki spaja više računala i drugih uređaja u računalnu mrežu, dok sličan uređaj **switch** izvodi isto efikasnije jer je sposoban prepoznati port određenog računala.
- **Usmjerivač** (engl. *router*) usmjerava podatkovne pakete kroz računalnu mrežu.
- Bežičnim ili žičanim povezivanjem lokalnih mreža **(W)LAN** (eng. *(Wireless) Local Area Network*) u vrlo raširene mreže **WAN** (engl. *Wide Area Network*) omogućuje se komunikacija među računalima na području jedne države ili kontinenta.
- Skup takvih međusobno povezanih mreža čini svjetsku računalnu mrežu - **internet**.

Što čini računalnu mrežu?



Izvor: <https://apprize.info/network/comptia/1.html>

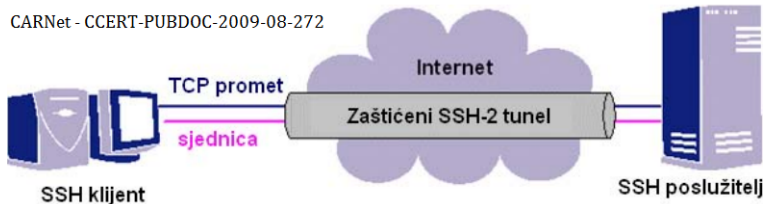
- **Server** (poslužitelj) je računalo koje drugim računalima (**klijentima**) omogućuje uslugu (e-mail, baza podataka, internet, svoje resurse. . .) ili pristup drugim uređajima u računalnoj mreži. Često je to:
 - **jedinstveno višeproseorsko računalo** gdje su svi procesori i memorija povezani kanalima matične ploče (ista arhitektura),
 - **skupina (klaster) višeprocorskih računala**, koja mogu biti ista ili heterogena, a koja su umrežena brzim LAN-om.
- **Računalnim klasterom** smatramo sustav računala koji koordinirano komuniciraju putem LAN-a. Specifične programske podrške omogućuju koordinirani zajednički rad i pretvaraju ih efektivno u cjelovit višeprocorski sustav. Možemo ih podijeliti na klastere:
 - s visokom učinkovitošću (HPC, *High Performance Computing*),
 - s visokom propusnošću (HT, eng. *High Throughput*),
 - s visokom dostupnošću (HA, eng. *High Availability*),
 - za ravnomjerno opterećenje (LB, eng. *Load Balancing*),
 - hibridnog tipa poput dodatne umreženosti s GPU procesorima (eng. *Graphics Processing Unit*)

- Komunikacijski protokoli (pravila) omogućuju uspostavljanje veze među ravnopravnim računalima, postavljanje upita ili zahtijevanje sredstava od poslužiteljskih računala...
- Najpoznatiji je **TCP/IP** protokol putem kojeg se šalju IP paketi:
 - bitovi informacije koja se prenosi - prijenosni protokol **TCP** (eng. *Transmission Control Protocol*) određuje način dijeljenja podataka na jednako velike prijenosne jedinice, dodjeljivanja rednih brojeva i nadzornih/zaštitnih instrukcija;
 - zaglavlje s adresama - mrežni protokol **IP** (*Internet Protocol*) dodjeljuje brojčane IP adrese izvorišnog i odredišnoga računala.
- **TelNet** (eng. *Telephone Network*), **FTP** (eng. *File Transfer Protocol*), **rsh** (eng. *remote shell*), **rcp** (eng. *remote copy*) i ostali protokoli šalju podatke u nekriptiranom obliku te tako postaju izloženi napadačima.
- Za razliku od spomenutih nesigurnih protokola, **SSH** (eng. *Secure Shell*) protokol prije slanja kriptira podatke i dekriptira nakon primitka te tako onemogućuje njihovo otkrivanje dok se prenose mrežom.

SSH protokol

- Osim spajanja na udaljeni poslužitelj i udaljenog izvođenja naredbi, SSH protokol omogućava i napredniju sigurnosnu komunikaciju:
 - scp – SSH inačica rcp naredbe kojom se podaci (datoteke) šalju kriptirani i zaštićeni;
 - sftp – SSH inačica FTP protokola za prijenos datoteka između računala;
 - sshfs (eng. *SSH Filesystem*) – protokol za siguran rad s datotečnim sustavom udaljenog računala.
- SSH tuneliranje omogućuje i prosljeđivanje TCP/IP prometa, primjerice X11 sjednica, odnosno grafičkih sučelja.

CARNet - CCERT-PUBDOC-2009-08-272



Povezivanje sa serverom

Za istraživačke potrebe i izvođenje praktične nastave dostupni su:

- **server UniST-Phy**

- smješten na UniST-PMF Odjelu za fiziku
- zahtjev za pristup: <https://forms.office.com/r/qHJzuMKBEJ>
- pregled zauzeća: <http://unist-phy.pmfst.hr/ganglia/>
- ime (hostname): **unist-phy.pmfst.hr**, protokol: **ssh/sftp**, port: **22**

- **računalni klaster Isabella**

- smješten na UniZG u SRCE
- zahtjev za pristup:
<https://wiki.srce.hr/display/RKI/Prijava>
- pregled trenutnog zauzeća: <http://teran.srce.hr/ganglia/>
- ime (hostname): **teran.srce.hr**, protokol: **ssh/sftp**, port: **22**

- **server mapmf**

- smješten na PMFST, autorizacija preko **AAI@EduHr**
- ime (hostname): **mapmf.pmfst.hr**, protokol: **ssh/sftp**, port: **22**

Zahtjevi za korištenje potpisuju se elektroničkim identitetom
AAI@EduHr.

SSH/SFTP programi

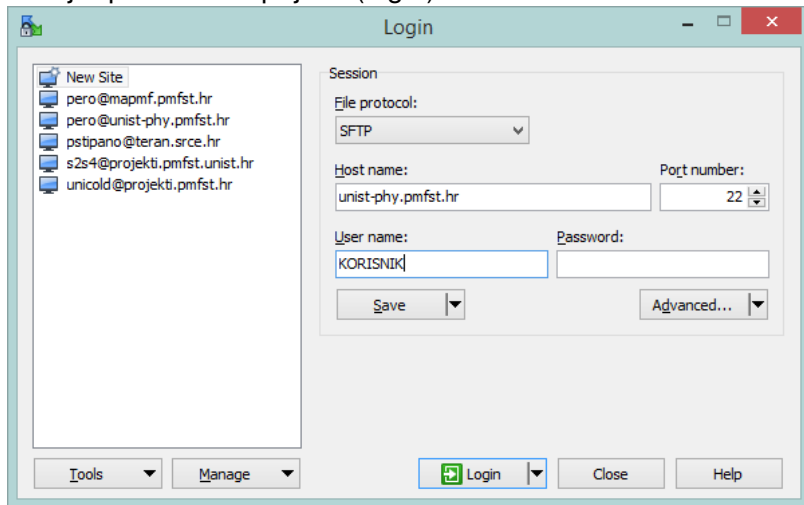
- Windows-programi za ssh/sftp pristup (port 22):
 - Bitvise SSH Client, <https://www.bitvise.com/download-area>
 - Alternativa:
 - PuTTY, <https://www.putty.org>
 - WinSCP, <https://winscp.net/eng/download.php>
- Za prijenos datoteka zgodan je WinSCP čiji terminal nema punu funkcionalnost pa je najbolji u kombinaciji s PuTTYjem.
- Pristup iz Linux ili Windows PowerShell (Win10) terminala:
 - **ssh** KORISNIK@IME.SERVERA
- Prilikom prve prijave pojavit će se poruka slična ovoj

Continue connecting to an unknown server and add its host key to a cache?

The server's host key was not found in the cache. You have no guarantee that the server is the computer you think it is.

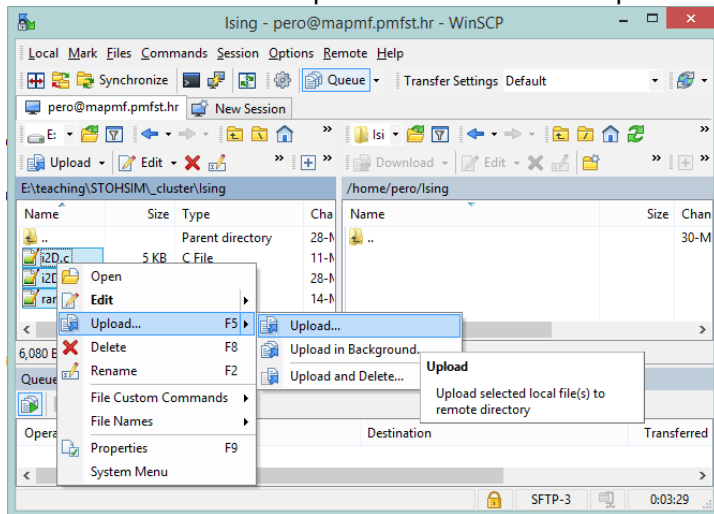
te je najzgodnije potvrditi poznavanje servera.

Primjer postavki za prijavu (login) koristeći WinSCP:



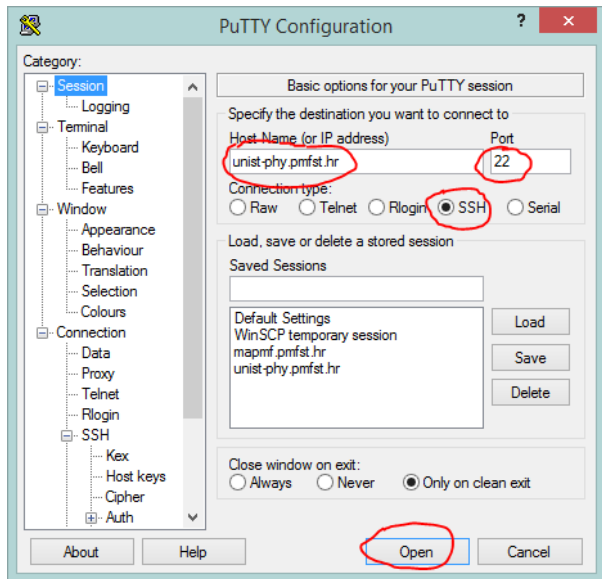
Nakon prvog upisa podataka, iste možemo pohraniti (Save) i nastaviti prijavu → Login.

U desnom WinSCP-prozoru otvorimo lokaciju na serveru, a u lijevom označimo datoteke za upload → desni klik → Upload:

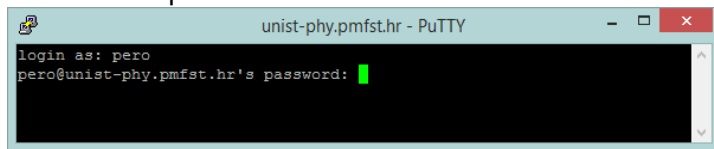


Za terminal je bolje koristiti PuTTY [Ctrl+P ili ikona (monitori+munja)].

Analogno ide prijava s PuTTYjem čije postavke također možemo sačuvati (Save) nakon

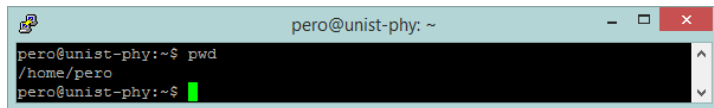


- PuTTY sam pita korisničko ime i šifru



```
unist-phy.pmfst.hr - PuTTY
login as: pero
pero@unist-phy.pmfst.hr's password: █
```

te nas nakon uspješne autorizacije odvodi u radnu mapu korisnika **/home/KORISNIK/**



```
pero@unist-phy: ~
pero@unist-phy:~$ pwd
/home/pero
pero@unist-phy:~$ █
```

- Odjavljujemo se zatvaranjem prozora ili naredbom ➤ **exit**
- S jednog se servera možemo spajati i na drugi naredbom ➤ **ssh KORISNIK@IME.SERVERA**

Primjerice Isabella ne dozvoljava pristup izvan CARNet mreže što možemo zaobići povezujući se prvo kroz ssh na mapamf.pmfst.hr, a zatim s mapmf na teran.srce.hr

Osnovne linux/sh naredbe

- Svaki korisnik vlastitu šifru može promijeniti naredbom
 - **passwd** (slijedi potvrda stare i dvostruka potvrda nove)
- Popis datoteka i mapa unutar radnog direktorija (bez i s detaljima)
 - **ls** ➤ **ls -al** (popis svih mapa i datoteka s detaljima)
- Prikaz apsolutne adrese radne mape
 - **pwd**
- Promjena radne mape
 - **cd ADRESA-MAPE**
 - **/** označava početnu nad-mapu (root) te odjeljuje imena mapa
 - adresa može biti apsolutna ili relativna u odnosu na radnu
 - **.** označava radnu mapu
 - **..** označava nad-mapu radne mape
 - **-** označava prethodnu radnu mapu
 - **~** označava početnu mapu korisnika **/home/KORISNIK/**
 - npr. ako je početna radna mapa **/home/pstipano**, ekvivalentno je:
 - ➤ **cd /sing** ➤ **cd ~/sing** ➤ **cd ./sing** ➤ **cd ./sing/**

- Stvaranje nove mape
 - `mkdir DIREKTORIJ`
- Kopiranje datoteka s adrese ŠTO u mapu ili datoteku GDJE
 - `cp [-r] ŠTO GDJE`
- Premještanje datoteke/mape ŠTO u GDJE
 - `mv [-r] ŠTO GDJE`
- Brisanje datoteke/mape ŠTO
 - `rm [-rf] ŠTO`
- Znak `*` zamjenjuje proizvoljan niz znakova te možemo koristiti u ŠTO za višestruki odabir, npr. ➤ `cp i*.txt lsing\` kopira sve txt datoteke, čije ime počinje s `i`, iz radne mape u pod-mapu `lsing`.
- Zastavica `-r` iza naredbi kopiranja, premještanja i brisanja označava rekurzivno listanje cijelog sadržaja, odnosno omogućuje dohvaćanje pod-mapa, a zastavica `-f` forsira brisanje bez dodatne potvrde, npr. ➤ `rm -r *` briše sve datoteke i pod-mape u radnoj mapi.
- Detaljnije: <https://explainshell.com/>

- Kopiranje podataka sa servera1 na server2¹
 - `scp [-r] [KORISNIK1@SERVER1:]ŠTO [KORISNIK2@SERVER2:]GDJE`
- Promjena Win -> Unix oznaka za kraj linije (EOL)
 - `cat DATOTEKA | tr '\n\r' '\n' > DATOTEKA`
- Arhiviranje datoteke/mape
 - `tar -zcvf IME-ARCHIVE.tar.gz SADRŽAJ-ZA-ARCHIVIRANJE`
 - -z = koristi program gzip
 - -c = kreiraj arhivu
 - -v = prikaži tijek arhiviranja
 - -f = koristi datoteku za pohranu
- Raspakiravanje arhive
 - `tar -zxvf IME-ARCHIVE.tar.gz [-C MAPA]`
 - -z = koristi program gzip
 - -x = raspakiraj datoteke
 - -v = prikaži tijek arhiviranja
 - -f = koristi datoteku
 - -C = promjena mape u koju će raspakirati arhivu

¹Ako su ŠTO/GDJE na serveru, najzgodnije je krenuti od početne mape ./

- Tekstualni editori: ➤ nano, ➤ joe, ➤ vi
- ➤ nano DATOTEKA stvara ili otvara txt-datoteku za uređivanje.
 - Na dnu otvorenog programa pišu naredbe.
 - Primjerice ^X predstavlja kombinaciju tipki Ctr+X što zatvara program uz upit želimo li sačuvati promjene.
- ➤ vi DATOTEKA otvara datoteku u txt-editoru „vi” u kojem pritiskom na i možemo editirati tekst, dok Esc omogućuje upis naredbi
 - :q! = zatvara datoteku bez pohrane;
 - :w = sprema promjene; :wq = sprema i zatvara datoteku;
 - :BROJ = vodi na prikaz linije BROJ, zadnje (\$) ili prve (0).
- ➤ cat DATOTEKA ispisuje sadržaj na ekran
- ➤ grep [-v] UZORAK DATOTEKA lista linije s uzorkom ili bez [-v]
- Rezultate izvršavanja naredbi možemo pohraniti u txt-datoteke upisom ➤ DATOTEKA na kraju naredbe.
- Niz naredbi možemo pohraniti u skriptu (tekstualnu datoteku) i pozvati, npr. ➤ gnuplot SKRIPTA.plt

- Standardni kompajleri za prevođenje koda u program:
 - ➤ **gfortran** KOD.f -o PROGRAM
 - ➤ **gcc** KOD.c -o PROGRAM -lm
 - ➤ **g++** KOD.cpp -o PROGRAM -lm
 - ➤ **python3** KOD.py
- ➤ **PROGRAM** (pokretanje instaliranog programa)
- ➤ **./PROGRAM** (ispred kompajliranog programa navodimo lokaciju)
- ➤ **nohup ./PROGRAM &** (nohup sprječava prekid nakon odjave)
- Ctrl+C = prekid izvršavanja programa
- ➤ **clear** (brisanje prozora - skrivanje prethodnih naredbi)
- ➤ **w** (lista logiranih korisnika i njihovih lokacija)
- ➤ **top** (lista pokrenutih procesa koja se tipkom M/P sortira prema zauzeću memorije/procesora), izlaz: **q**
- ➤ **ps [-u KORISNIK]** (prikazuje korisničke procese)
- ➤ **ps [aux -forest]** (prikazuje hijerarhijski sve sistemske procese)
- ➤ **kill [-9] PID** (gasi proces pod brojem PID)

Server Unist-Phy

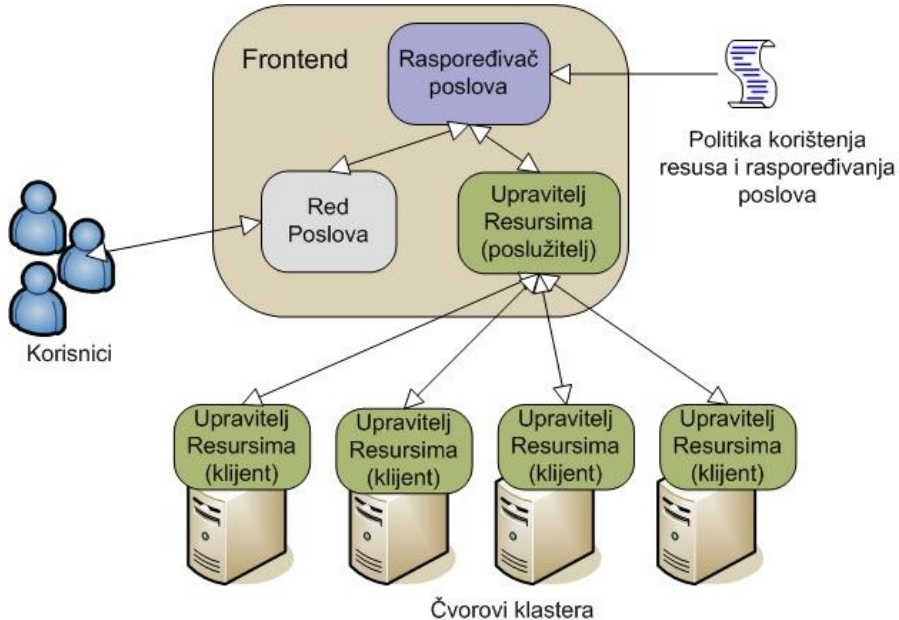
Spajanje je opisano na slajdu 9.

- Poželjno je odmah nakon prve prijave promijeniti šifru: ➤ [passwd](#)
- Višeprosorsko računalo je u testnoj fazi konfiguriranja i optimiziranja.
- Svi su portovi zatvoreni firewallom osim porta 80/tcp radi praćenja zauzeća <http://unist-phy.pmfst.hr/ganglia/> i porta 22 za pristup pa se isti dodatno nadzire i blokira korisnika i njegovu IP ako pogriješi šifru 8 puta.
- Treba posebno voditi računa o zauzeću procesora i ne zatrpiti klaster dok nema raspoređivača poslova.
- Više detalja će biti kada se završi podešavanje instaliranog raspoređivača poslova,
<https://slurm.schedmd.com/sinfo.html>

Računalni klaster Isabella

- Računalnim klasterom smatramo sustav računala koji koordinirano komuniciraju putem brze lokalne mreže (LAN).
- Specifične programske podrške omogućuju koordinirani zajednički rad i pretvaraju ih efektivno u jedinstven višeprocorski sustav.
- Računalni klaster Isabella spada u HPC (eng. *High Performance Computing*) skupinu klastera.
- Sustav za upravljanje poslovima čine 3 komponente:
 - red poslova (eng. *Queue Manager*);
 - raspoređivač poslova (eng. *Scheduler*);
 - upravitelj resursima (eng. *Resource Manager*).
- Detaljnije:

<https://wiki.srce.hr/pages/viewpage.action?pageId=8488260>



- Korisnik priprema programe i skripte na [vlastitom računalu](#).
- Korisnik sve potrebno prenosi na [pristupno računalo](#) klastera.
- Korisnik kompajlira kodove i pokreću poslove - skripte unutar kojih se nalaze instrukcije što će prenijeti na [radni čvor](#), pokrenuti tamo i koje podatke vratiti na pristupno računalo.
- Red poslova raspoređuje poslove u redove i prosljeđuje raspoređivaču poslova.
- Upravitelj resursima prikuplja podataka o stanju čvorova, pokreće i prati izvršavanje poslova.
- Raspoređivač poslova, na temelju podataka o poslovima od Reda poslova i podatke o stanju i opterećenosti resursa od Upravitelja resursima, određuje kada i gdje će se poslovi izvoditi.
- Po izvršavanju skripte Upravitelj resursima briše sve podatke na radnom čvoru.
- Korisnik može preuzeti podatke s [pristupnog računala](#) gdje ostaju dok god ih ne pobriše ili izgubi pravo pristupa.

- Korisnik pristupa klasteru i vlastitim podacima putem pristupnog poslužitelj (**teran.srce.hr**) na virtualizacijskom VMware klasteru.
- Korisnik s pristupnog poslužitelja putem sh-skripti komunicira sa sustavom za upravljanje poslova - Sun of Grid Engine (SGE).
- Zauzetost klastera vidljiva je putem sustava za nadzor - Ganglia
 - http://teran.srce.hr/ganglia/?c=Isabella&m=load_one&r=hour&s=by%20name&hc=4&mc=2
- Korisnicima je dostupno 135 računalnih čvorova koji čine jedinstven sustav od:
 - 270 CPU procesora;
 - 3.100 CPU procesorskih jezgri;
 - 12 grafičkih procesora;
 - 16 TB radne memorije;
 - 756 TB dijeljenog podatkovnog prostora.
- Detaljnije:

<https://www.srce.unizg.hr/isabella/tehnicke-specifikacije>

1. korak: kreiranje sh-skripte

Podatke prebacujemo na pristupno računalo sa instrukcijama kako će ih upravitelj resursima prebaciti na radni čvor za računanje i vratiti.

- Nakon pripreme programa, kojeg planiramo pokrenuti na klasteru, potrebno je kreirati instrukcije za SGE, odnosno niz tekst-naredbi pohraniti u SKRIPTA.sh:
 - kreiranje radne mape na radnom čvoru (SGE stvara **\$TMPDIR**), npr.
 - **mkdir \$TMPDIR/Ising**
 - kopiranje svega potrebnoga s pristupnog računala na radni čvor, npr.
 - **cp -R /home/pstipano/Ising/* \$TMPDIR/Ising**
 - postavljanje mape na radnom čvoru kao radne, npr.
 - **cd \$TMPDIR/Ising**
 - pokretanje program, npr.
 - **/home/pstipano/Ising/i2D.exe**
 - kopiranje rezultata u vlastitu mapu na pristupnom računalu, npr.
 - **cp -R \$TMPDIR/Ising/*.dat /home/pstipano/Ising/**

2. korak: pristup klasteru

Računalnom klasteru Isabella može se pristupiti samo iz CARNet mreže!

- Prijava na računalni klaster:
 - `ssh KORISNIK@teran.srce.hr`
- Ako šifru pogriješite 3 puta, vaš će račun biti blokiran.
- Prvotno dodijeljena šifra je obično nepamtljiva, a možemo je promijeniti naredbom:
 - `passwd`
- Izlazimo s ➤ `exit`

Ako niste u CARNet mreži, potrebo se je logirati na računalo koje u CARNet mreži pa obaviti pristup s njega, primjerice:

- Prijavimo se na PMFST-ov server koristeći `AAI@EduHr`
 - `ssh AAI@mapmf.pmfst.hr`
- Prijavimo se na Isabellu iz mapmf-terminala:
 - `ssh KORISNIK@teran.srce.hr`

3. korak: prijenos podataka i pokretanje poslova

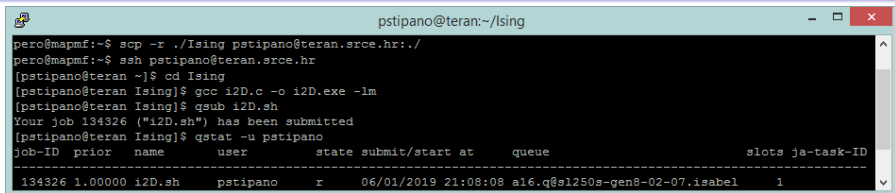
- Kodove, skripte i ulazne datoteke, koji su pripremljeni na vlastitom računalu, prenosimo na pristupno računalo (12. i 17. str.).
- Postavimo direktorij, u kojem su kod i skripta, kao radni, cd...
- Kompajliramo C/Fortran-kod
 - `gcc KOD.c -o PROGRAM -lm` ➤ `gfortran KOD.f -o PROGRAM`
- Posao šaljemo raspoređivaču poslova pozivom skripte

```
1 mkdir $TMPDIR/Ising
2 cp -R /home/pstipano/Ising/* $TMPDIR/Ising
3 cd $TMPDIR/Ising
4 /home/pstipano/Ising/i2D.exe
5 cp -R $TMPDIR/Ising/*.dat /home/pstipano/Ising/
```

- `qsub SKRIPTA.sh`
- Pregled poslova korisnika (ili svih navođenjem imena '')
 - `qstat -u KORISNIK`

Status **Q** označava posao koji čeka slobodno mjesto, **R** pokrenuti, **E** završavajući, **H** posao koji je zapeo te ga treba

4. korak: provjera trenutnog stanja poslova



```
pero@mapmf:~$ scp -r ./Ising pstipano@teran.srce.hr:./
pero@mapmf:~$ ssh pstipano@teran.srce.hr
[pstipano@teran ~]$ cd Ising
[pstipano@teran Ising]$ gcc i2D.c -o i2D.exe -lm
[pstipano@teran Ising]$ qsub i2D.sh
Your job 134326 ("i2D.sh") has been submitted
[pstipano@teran Ising]$ qstat -u pstipano
```

job-ID	prior	name	user	state	submit/start at	queue	slots	ja-task-ID
134326	1.00000	i2D.sh	pstipano	r	06/01/2019 21:08:08	a16.q@sl250s-gen8-02-07.isabel	1	

U queue-info iza @ piše adresa radnog čvora (**sl250s-gen8-05-05**), a ispred @ red posla (**a16.q**) koji u kombinaciji s job-ID (**134326**) i br. slota (**1**) omogućuje pronalazak radnog direktorija na radnom čvoru. Ako nam je baš potrebno, podatke možemo dohvatiti prije završetka:

- Pristupimo radnom čvoru ➤ **sftp sl250s-gen8-05-05**
- Otvorimo radni direktorij na čvoru ➤ **cd /scratch/134326.1.a16.q/...**
- Kopiramo podatke u radni direktorij pristupnog računala ➤ **get ŠTO**
- Vraćamo se s ➤ **exit**

Po završetku posla kreira se /home/KORISNIK/SKRIPTA.oJOBID (ispis na ekran) i /home/KORISNIK/SKRIPTA.eJOBID (popis grešaka).