

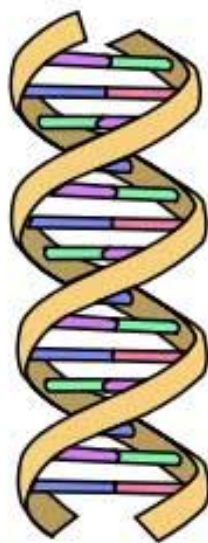
Upotreba CRO NGI u bioinformatiči i buduće potrebe

Maša Roller Milošević
Grupa za bioinformatiku
PMF, Sveučilište u Zagrebu

Sadašnja upotreba: pretraživanje baza podataka

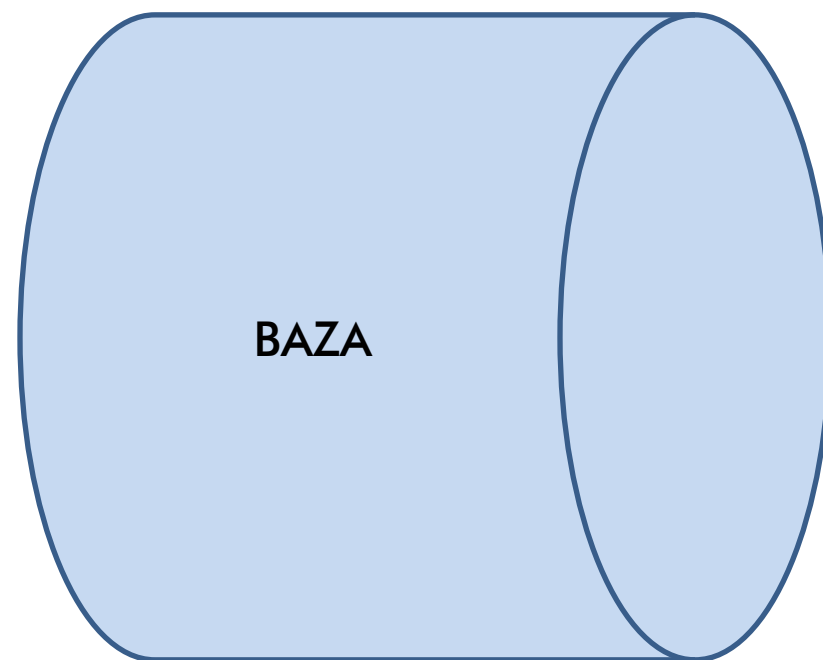
~4.5 GB

~5 GB



DNA

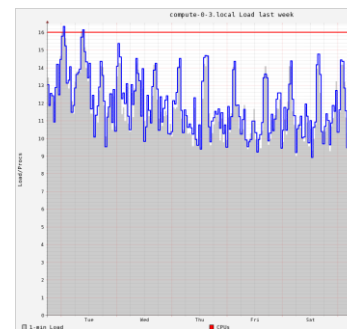
- = Adenine
- = Thymin
- = Cytosine
- = Guanine
- = Phosphate backbone



† t_{running} (1 core CPU) = 24 mjeseci

“Ljudska” Paralelizacija poslova

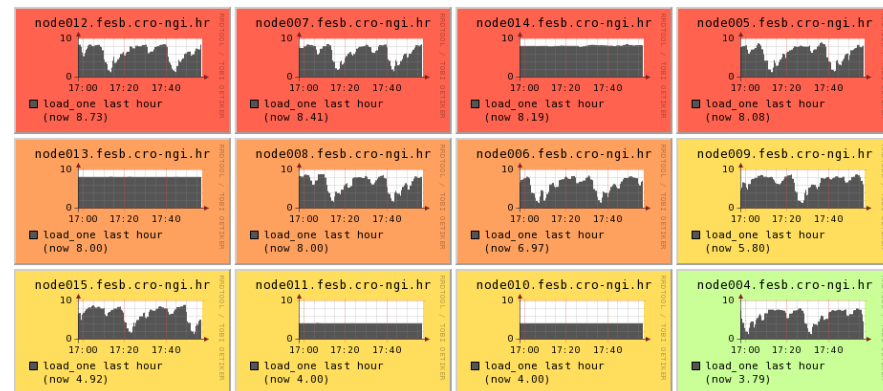
Broj CPU = x ubrzanje
 – 16 CPU = 16x brže



† **running (16 cores CPU) = 3 mjeseca**

Dijeljenje ulaznih podataka

– 12 dijelova = 12x brže



† **running (16 cores CPU) / 12 = 8 dana**

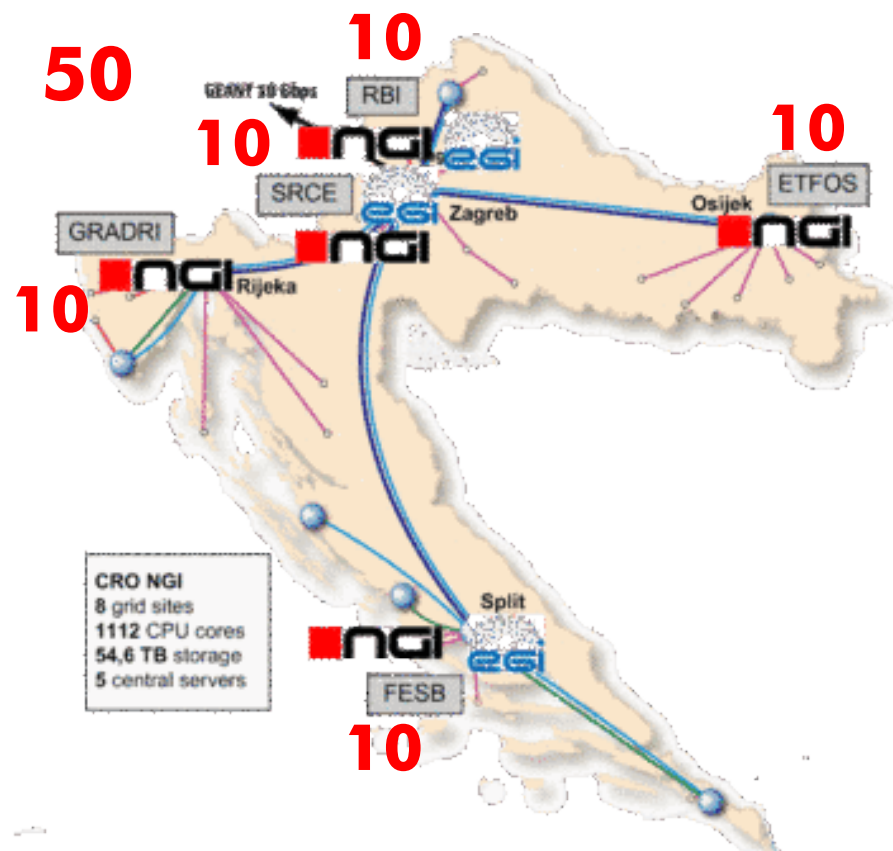
Condor-G matchmaking

Opis Sustava

- "omogućava izvođenje poslova na grid resursima bez eksplicitnog navođenja adrese udaljenog resursa"
- "dohvaća informacije o dostupnim sredstvima iz raspoloživih sustava za nadzor"
<http://wiki.cro-ngi.hr>

Što to znači u praksi

- Poslovi se odmah raspodijele po svim sjedištima
- Nema preraspodjele kako se sjedišta prazne



“Ljudski” matchmaking

input_number = 20

center = ce.srce.cro-ngi.hr/jobmanager-sge

#center = ce.etfos.cro-ngi.hr/jobmanager-pbs

#center = ce.fesb.cro-ngi.hr/jobmanager-pbs

#center = ce.gradri.cro-ngi.hr/jobmanager-sge

#center = ce.irb.cro-ngi.hr/jobmanager-sge

....

→ transfer_input_files = input/input_\$(input_number).fasta

→ transfer_output_files = OUTPUT_sponge_\$(input_number)_cog_\$(Cluster).txt

thread_number = 8

...

→ Log=condor_logs/blast_sponge_cog_\$(input_number)_\$(Cluster).log

→ Output= condor_out/blast_spongecog_\$(input_number)_\$(Cluster).out

→ Error= condor_err/blast_sponge_cog_\$(input_number)_\$(Cluster).err

...

→ Arguments= -d /home/cog8_db/cog8 -i input_\$(input_number).fasta -o

→ OUTPUT_sponge_\$(input_number)_cog_\$(Cluster).txt -p blastx -e 1e-8 -F F -w 20 -m 9 -a
→ \$(thread_number)

→ grid_resource = gt2 \$(center)

environment = "PE_MODE=single"

→ globusrs1 = (jobType=single)(count=\$(thread_number))

“Ljudsko” nadgledanje poslova

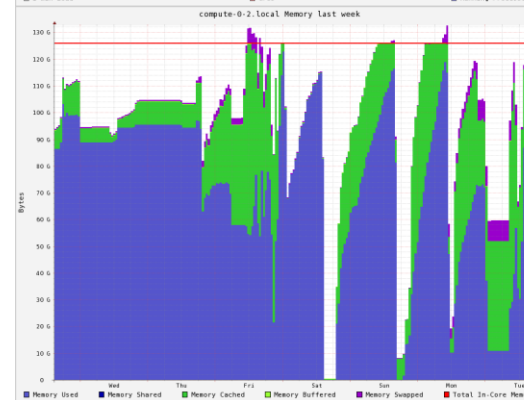
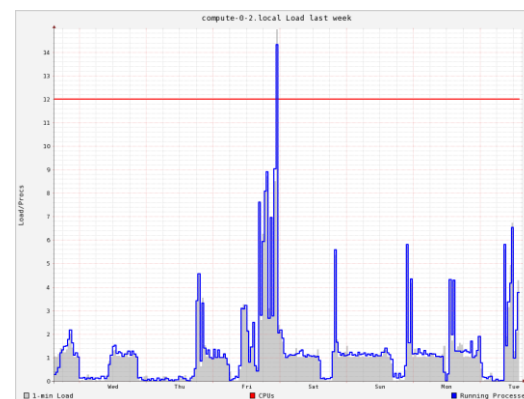
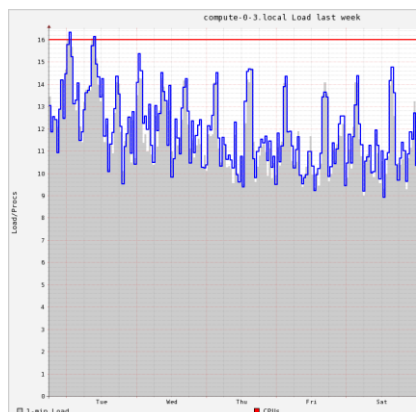
-- Submitter: mrollerm@cs1.cro-ngi.hr : <161.53.133.26:23598> : cs1.cro-ngi.hr

ID	OWNER	STATUS	MANAGER	HOST	EXECUTABLE
199478.0	mrollerm	ACTIVE	pbs	ce.fesb.cro-ngi.hr	/usr/local/ncbi/bl
199479.0	mrollerm	ACTIVE	pbs	ce.fesb.cro-ngi.hr	/usr/local/ncbi/bl
199480.0	mrollerm	ACTIVE	pbs	ce.fesb.cro-ngi.hr	/usr/local/ncbi/bl
199481.0	mrollerm	ACTIVE	pbs	ce.fesb.cro-ngi.hr	/usr/local/ncbi/bl
199482.0	mrollerm	ACTIVE	sgc	ce.srce.cro-ngi.hr	/usr/local/ncbi/bl
199483.0	mrollerm	ACTIVE	sgc	ce.srce.cro-ngi.hr	/usr/local/ncbi/bl
199484.0	mrollerm	ACTIVE	sgc	ce.srce.cro-ngi.hr	/usr/local/ncbi/bl
199485.0	mrollerm	ACTIVE	sgc	ce.srce.cro-ngi.hr	/usr/local/ncbi/bl
199486.0	mrollerm	ACTIVE	sgc	ce.srce.cro-ngi.hr	/usr/local/ncbi/bl
199487.0	mrollerm	ACTIVE	pbs	ce.fesb.cro-ngi.hr	/usr/local/ncbi/bl
199488.0	mrollerm	ACTIVE	pbs	ce.fesb.cro-ngi.hr	/usr/local/ncbi/bl
199489.0	mrollerm	ACTIVE	pbs	ce.fesb.cro-ngi.hr	/usr/local/ncbi/bl
199490.0	mrollerm	ACTIVE	pbs	ce.etfos.cro-ngi.h	/usr/local/ncbi/bl
199491.0	mrollerm	ACTIVE	pbs	ce.etfos.cro-ngi.h	/usr/local/ncbi/bl
199493.0	mrollerm	ACTIVE	pbs	ce.etfos.cro-ngi.h	/usr/local/ncbi/bl
199494.0	mrollerm	ACTIVE	pbs	ce.etfos.cro-ngi.h	/usr/local/ncbi/bl
199495.0	mrollerm	ACTIVE	pbs	ce.etfos.cro-ngi.h	/usr/local/ncbi/bl
199496.0	mrollerm	PENDING	pbs	ce.etfos.cro-ngi.h	/usr/local/ncbi/bl
199497.0	mrollerm	PENDING	pbs	ce.etfos.cro-ngi.h	/usr/local/ncbi/bl
199606.0	mrollerm	ACTIVE	sgc	ce.gradri.cro-ngi.	/usr/local/ncbi/bl

Buduće potrebe

CPU zahtjevni poslovi
CRO NGI

Memorijski zahtjevni poslovi
?



Grupa za bioinformatiku

Voditelj:

Kristian Vlahoviček



Znanstveni/ novaci/kinje:

Vedran Franke

Rosa Karlić

Tina Kokan

Vedran Lucić

Sofia Pinto

Maša Roller



Sistem administrator

Petar Jager

