



arnes 



NIxOS
National Initiatives for Open Science in Europe

Arnesova prenovljena superračunalniška gruča

Peter Kacin in Blaž Česnik, Arnes

Mreža znanja 2021, od 23. do 30. novembra

V OMREŽJU
GĚANT SMO
POVEZANI TUDI
SUPERRAČUNALNIKI.

SUPERRAČUNALNIKI
DELAMO VES ČAS.
TIHO BRNIMO,
A GOVORIMO V EN GLAS:
RAČUNAJTE NA NAS!

SLOVENSKA
SUPER EKIPA
SE IMENUJE
SLING.

PRAVKAR
PRERAČUNAVAM DNK.
A LAHKO RAČUNAM
MALO NATE?

POMAGAL SEM
IZRAČUNATI HIGGSOV
BOZON IN TUDI
TOKRAT BOM Z
VESELJEM SODELOVAL!

$$m_h^2 = (m_h^2)^2 (m_h^2) + \frac{3g^2 A_h^2}{32\pi^2 m_h^2} (m_h^2 + 2m_h^2 + m_h^2) + \frac{3g^2 m_h^2}{64\pi^2 m_h^2} \left(m_h^2 \ln \frac{A_h^2}{m_h^2} - 2m_h^2 \ln \frac{A_h^2}{m_h^2} - m_h^2 \ln \frac{A_h^2}{m_h^2} \right)$$

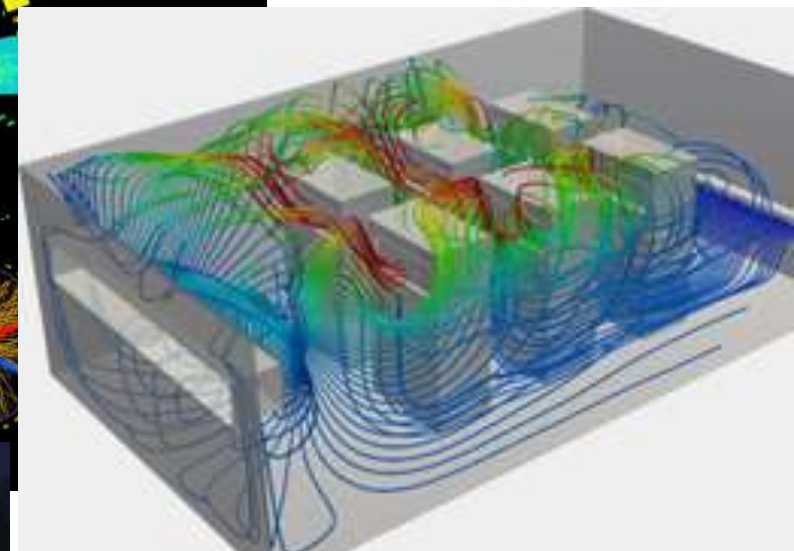
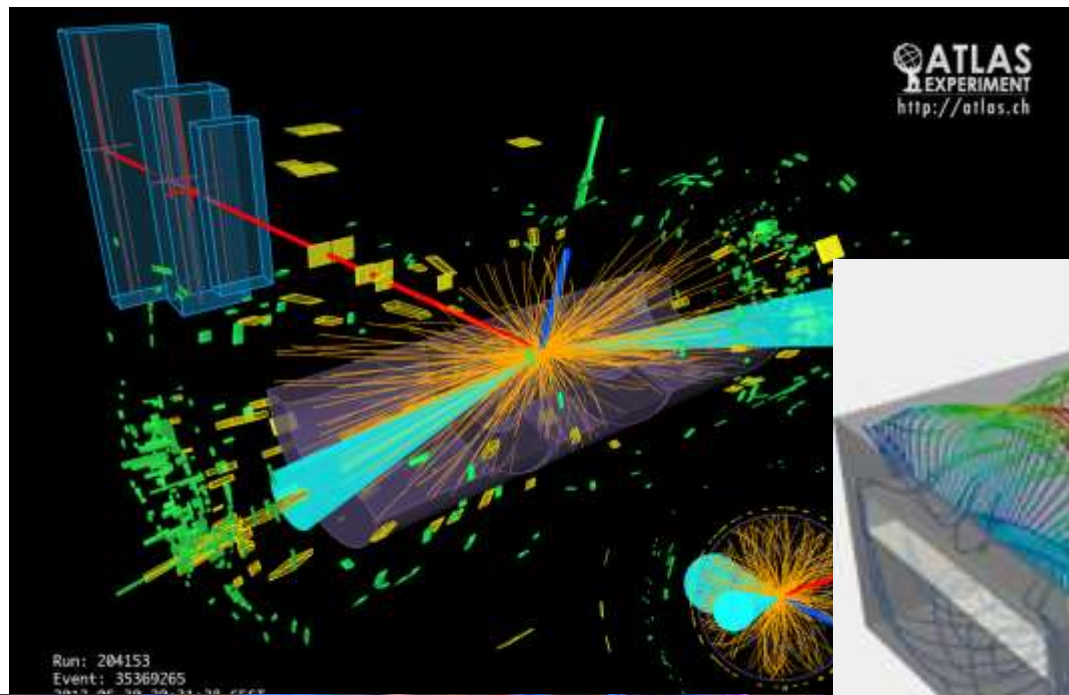


Uporaba

- Računsko zahtevni problemi
- Vzporedno računanje

Na voljo vsem slovenskim raziskovalcem

Nekaj 1000x zmogljivejši od osebnega računalnika



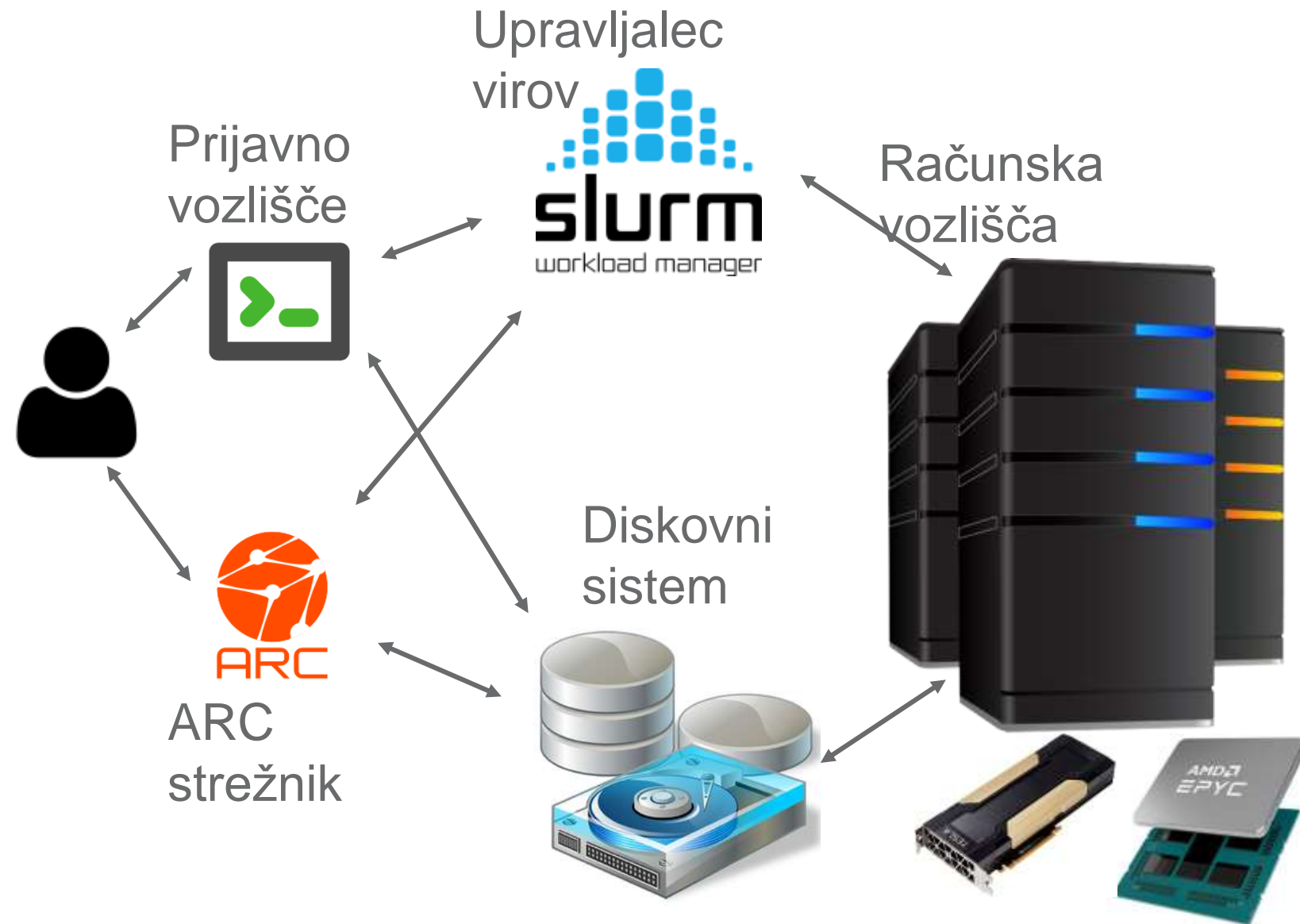
support@sling.si

doc.sling.si



**Posodobitev Arnesovega
superračunalnika – kaj je
novega?**

Arhitektura



Način dostopa



Prijavno vozlišče

- SSH prijava(s ključem)
- Neposredna interakcija z upravljalcem virov SLURM



Vmesna programska oprema ARC

- ARC odjemalec
- SiGNET certifikat
- XRSL datoteka

Diskovni prostor

Porazdeljen diskovni sistem:

- Odpornost na napake strojne opreme
- Visoke hitrosti prenosa podatkov

4 stare strežnike zamenjali z **10 novimi**:

- skupaj 280 HDD in 60 SSD diskov

• Skupaj uporabnega prostora

- **1PB HDD**
- **100TB SSD**

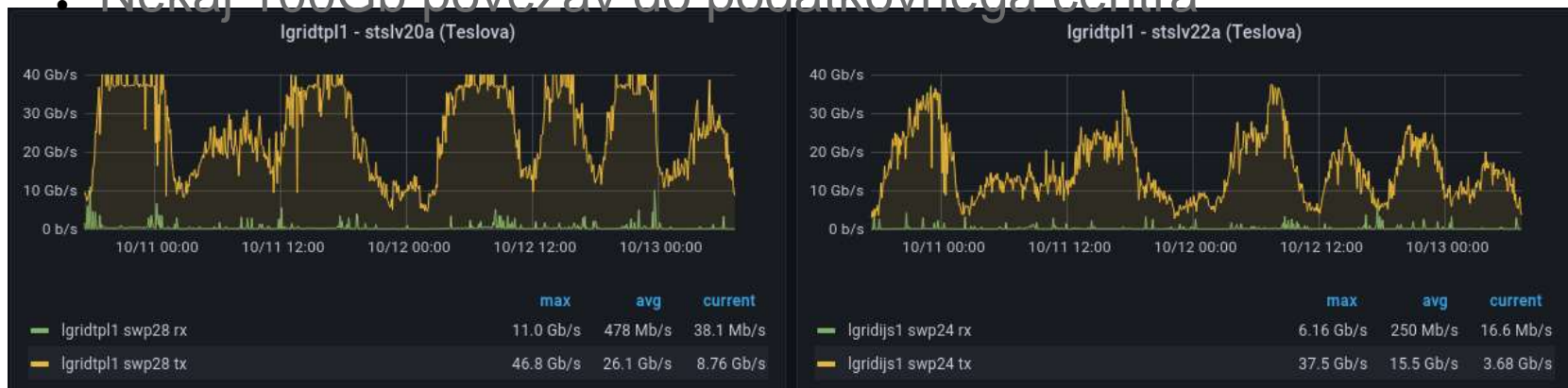


Omrežje



- Glavni infrastrukturni strežniki 2x 25Gb
- Workerji 10Gb
- DC Teslova \leftrightarrow DC Arnes
 - trenutno 2x40Gb
 - Plan 2x 100Gb

- Nekaj 100Gb povezav do podatkovnega centra



Računska vozlišča - “Workerji”

Posodobitev 2021:

- 4256c/8512t
- 48x NVIDIA V100s
- 62 CPU workerjev
 - 64c/128t
 - 256GB RAM
- 24 GPU workerjev
 - 2xNVIDIA V100s

Skupaj v gruči:

- več kot 10000 jeder(niti)
- skoraj 30TB pomnilnika
- 48 pomožnih grafičnih pospeševalnikov

Skupna teoretična zmogljivost slabih **600 TFLOPS**





Proces vključitve računskega vozlišča v gručo HPC

Avtomatizacija


- Ročna namestitvev = zamudna in draga.
- Avtomatizacija = hitro in učinkovito.
- Zagon preko omrežja.

Dodeljevanje identitete računskemu vozlišču

- Vnos DNS zapisa:
 - Pretvorba IP naslova v spletni naslov.
 - 153.5.72.111 → wn111.arnes.si.
- Vnos DHCP zapisa:
 - Dodelitev IP naslova mrežni kartici.
 - 00:90:fa:aa:b2:c4 → 153.5.72.111.
 - Konfiguracija parametrov za zagon OS-a preko omrežja.

Foreman

- Orodje za avtomatizirano upravljanje serverjev.
- OS na računskem vozlišču predstavi vozlišče Foremanu.
- Namestitev osnovnega OS-a.
- Datoteka kickstart in particijska datoteka.



```
Discovery status
Status: SUCCESS
Primary NIC: 00:0A:F7:6F:47:A0
Primary IPv4: N/A
Primary IPv6: N/A
Discovery server: https://153.5.73.73
Endpoint type: foreman
Latest server response:
201:
{"last_report":"2021-10-14T12:49:27.289Z","build":false,"id":660,"name
Kernel command line:
BOOT_IMAGE=boot/fdi-image/vmlinuz0
<Resend> <Status> <Facts> <Logs> <SSH> <Reboot>
```

FOREMAN Default Organization Default Location Admin User

Discovered Hosts

Filter Search Select Action

Name	Model	IP Address	CPUs	Memory	Disk Count	Disks Size	Manufacturer	Dns_hostname	Inventory_disk	Discovery_bootif	Nmprimary_dhcp4_option_filename	Location	Organization	Subnet	Last Facts Upload	Actions
win074	ProLiant DL380 Gen9	153.572.74	2	252 GB	2	280 GB	HP	win074.arnes.si	pci-0000:03:00.0-scsi-0:1:0:0	00:90:fa:43:87:d4	grub2/grubx64.efi	Default Location	Default Organization	HPC_network (153.572.0/23)	2 minutes ago	Provision
win077	ProLiant DL380 Gen9	153.572.77	2	252 GB	2	280 GB	HP	win077.arnes.si	pci-0000:03:00.0-scsi-0:1:0:0	00:00:c9:cc:a2:de	grub2/grubx64.efi	Default Location	Default Organization	HPC_network (153.572.0/23)	6 minutes ago	Provision
win076	ProLiant DL380 Gen9	153.572.76	2	252 GB	2	280 GB	HP	win076.arnes.si	pci-0000:03:00.0-scsi-0:1:0:0	00:00:c9:c8:fc:d8	grub2/grubx64.efi	Default Location	Default Organization	HPC_network (153.572.0/23)	10 minutes ago	Provision
win075	ProLiant DL380 Gen9	153.572.75	2	252 GB	2	280 GB	HP	win075.arnes.si	pci-0000:03:00.0-scsi-0:1:0:0	00:90:fa:54:44:8a	grub2/grubx64.efi	Default Location	Default Organization	HPC_network (153.572.0/23)	10 minutes ago	Provision
win073	ProLiant DL380 Gen9	153.572.73	2	252 GB	2	280 GB	HP	win073.arnes.si	pci-0000:03:00.0-scsi-0:1:0:0	00:90:fa:43:87:c8	grub2/grubx64.efi	Default Location	Default Organization	HPC_network (153.572.0/23)	10 minutes ago	Provision
win072	System x3650 M5-	153.572.72	2	251 GB	1	278 GB	LENOVO	win072.arnes.si	pci-0000:15:00.0-scsi-0:2:0:0	00:0e:1e:a1:2b:20	grub2/grubx64.efi	Default Location	Default Organization	HPC_network (153.572.0/23)	4 minutes ago	Provision

```

129
130 # IPv4
131 if (subnet4 && !subnet4.dhcp_boot_mode?) || @static
132   network_options.push("--bootproto static")
133   network_options.push("--ip=#{iface.ip}")
134   network_options.push("--netmask=#{subnet4.mask}")
135   network_options.push("--gateway=#{subnet4.gateway}")
136 elsif subnet4 && subnet4.dhcp_boot_mode?
137   network_options.push("--bootproto dhcp")
138 end
139 if subnet4
140   nameservers.concat(subnet4.dns_servers)
141   network_options.push("--mtu=#{subnet4.mtu}") if subnet4.mtu.present?
142 end
143
144 # IPv6
145 if rhel_compatible && os.major >= 6
146   if (subnet6 && !subnet6.dhcp_boot_mode?) || @static6
147     network_options.push("--ipv6=#{iface.ip6}/#{subnet6.cidr}")
148     network_options.push("--ipv6gateway=#{subnet6.gateway}")
149   elsif subnet6 && subnet6.dhcp_boot_mode?
150     if host_param_true?('use-slaac')
151       network_options.push("--ipv6 auto")
152     else
153       network_options.push("--ipv6 dhcp")
154     end
155   end
156   if subnet6

```

```

1 <<#
2 name: Default GRID
3 snippet: false
4 model: Ptable
5 os family: Redhat
6 organizations:
7   - Default Organization
8 locations:
9   - Default Location
10 <<#
11 zerombr
12 clearpart --all --initlabel
13
14 << if @host.pxe_loader.include?('UEFI') <<#
15
16 part /boot/efi --fstype="efi" --ondisk=sda --size=200 --fsoptions="umask=0077,shortname=efi"
17 << if (@host.operatingsystem.family == 'Redhat' && @host.operatingsystem.major.to_i > 7) <<#
18 part /boot --fstype="xfs" --ondisk=sda --size=1024
19 << else <<#
20 part /boot --fstype="ext4" --ondisk=sda --size=1024
21 << end <<#
22 << else @host.pxe_loader.include?('BIOS') <<#
23
24 part /boot --fstype=ext4 --size=1024 --ondisk=sda

```

Puppet

- Orodje za upravljanje programske opreme.
- Recepti za namestitev in konfiguracijo modulov.
- Uporaba testov za zagotavljanje integritete.

```
1 #CF worker
2 node 'will.arnes.si' {
3
4   class { '::profile::base':
5     sshd_users => ['grid', 'jost-root-workerscli'],
6     node_exporter_labels => {
7       'node_role' => 'worker-cpu',
8     },
9   }
10  include ::profile::puppet_agent
11
12  class { '::firewalloff': }
13  class { '::custom::package':
14    packages => ['wget', 'vim'],
15  }
16
17  class { '::repository::repel':
18    debuginfo_enabled => true,
19  }
20  class { '::repository::cernvm': }
21  class { '::repository::egl_trustanchors': }
22  class { '::repository::elastic': }
23  class { '::repository::ceph': }
24  class { '::repository::arnes_hpc': }
25  #class { '::repository::mariadb':
26    # version => '10.4',
27  #}
28  class { '::repository::elrepo':
29    archive => true,
30  }
31
32  class { '::grid_ceph': }
33
34  class { '::autofs': }
35
36  class { '::irqbalance': }
37
38  class { '::auge': }
39
40  class { '::singularity': }
41
42  class { '::cvmfs': }
43
44  class { '::slurm::client': }
45  class { '::nhc': }
46
47  #class { '::...': }
```

Sistem za razporejanje delovnih nalog - SLURM

- Dodajanje računskega vozlišča v konfiguracijo SLURM.
- Pričetek komunikacije glavnega in računskega vozlišča.

```
NodeName=wn105 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn106 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn107 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn108 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn109 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn110 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn111 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn112 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn113 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn114 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn115 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn116 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn117 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn118 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn119 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn120 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
NodeName=wn121 CPUs=128 Boards=1 SocketsPerBoard=1 CoresPerSocket=64 ThreadsPerCore=2 RealMemory=256000 Feature=amd,rome
```


Spremljanje stanja - monitoring

Prometheus

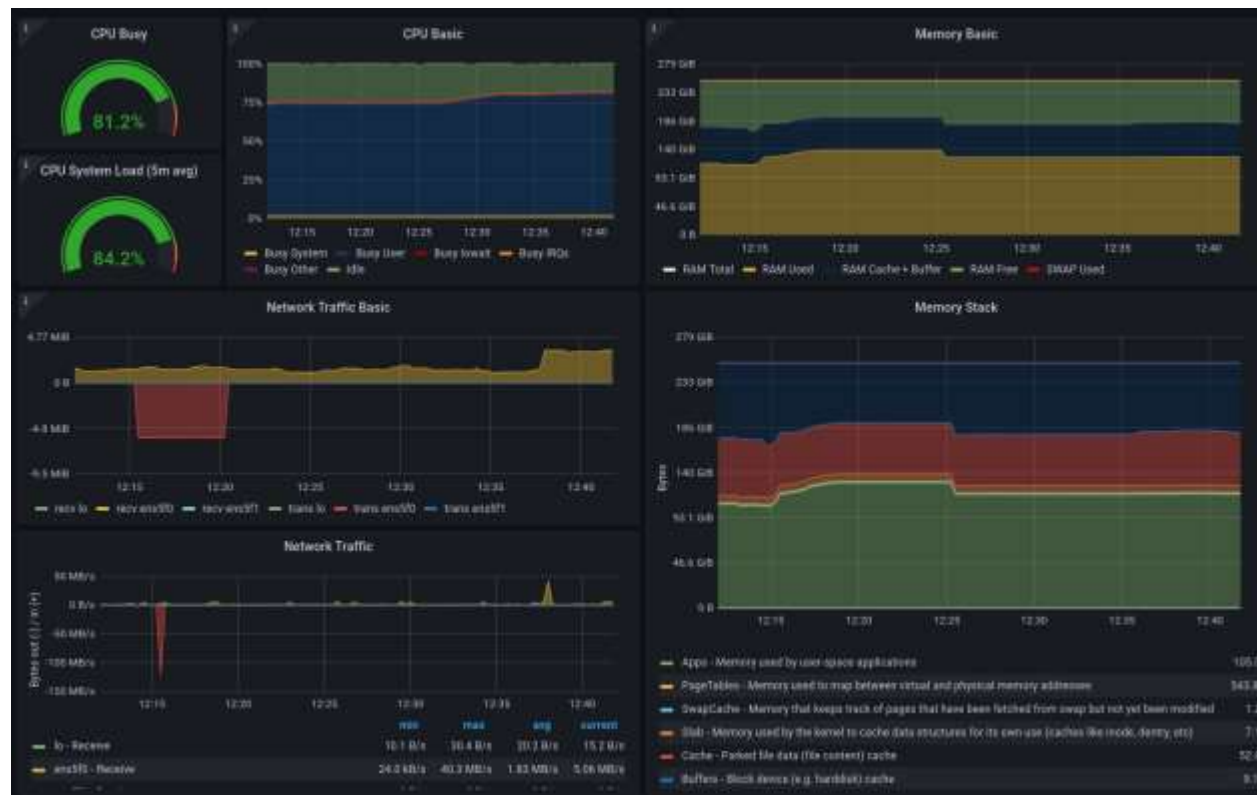
Zbiranje metrik.

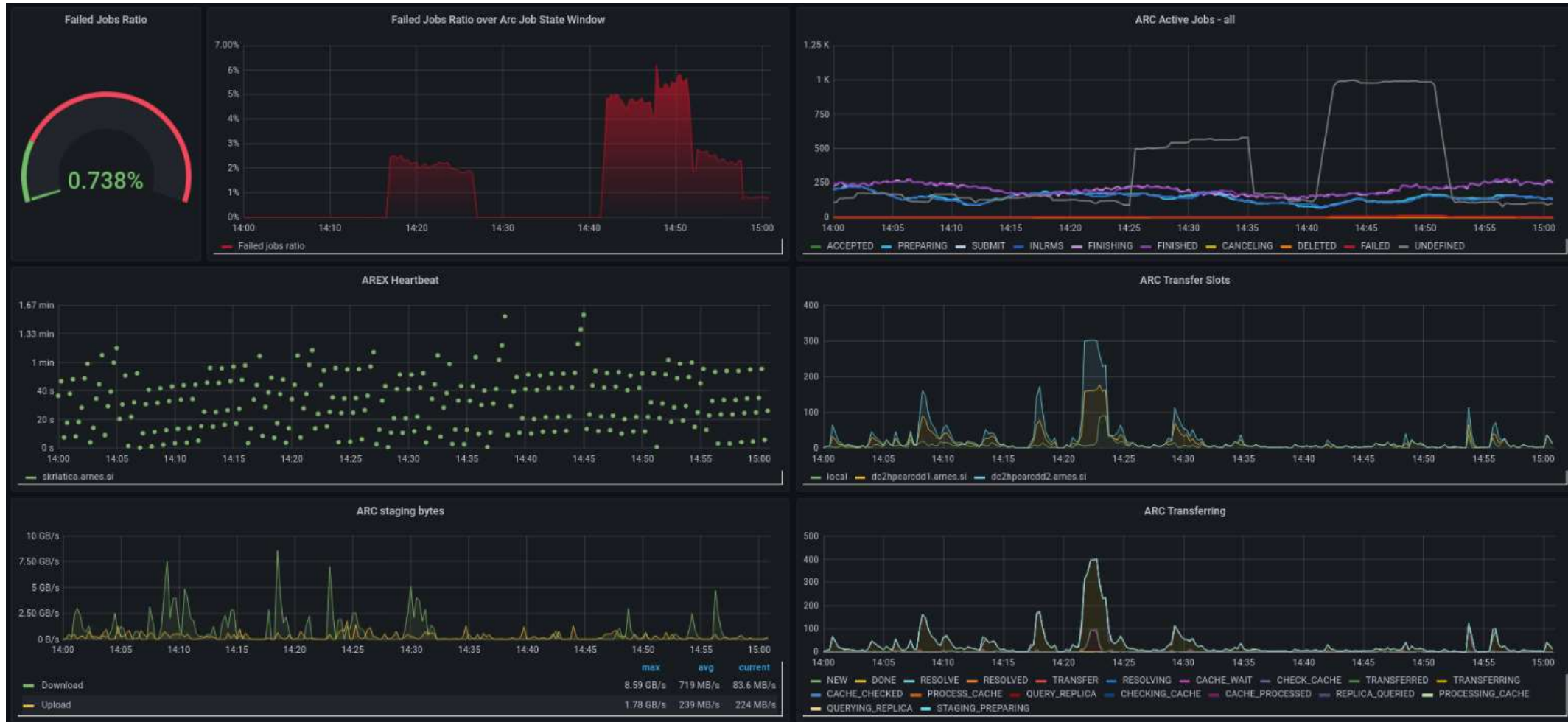
Grafana

Vizualizacija z uporabo metrik.

Nagios

Alarmni sistem.





Prijetno uporabo in hvala za pozornost!
peter.kacin@arnes.si, blaz.cesnik@arnes.si

support@sling.si

doc.sling.si

arnes 