

Izgradnja clustera visoke dostupnosti na Linuxu

Vrste clustera

Dvije namjene clustera:

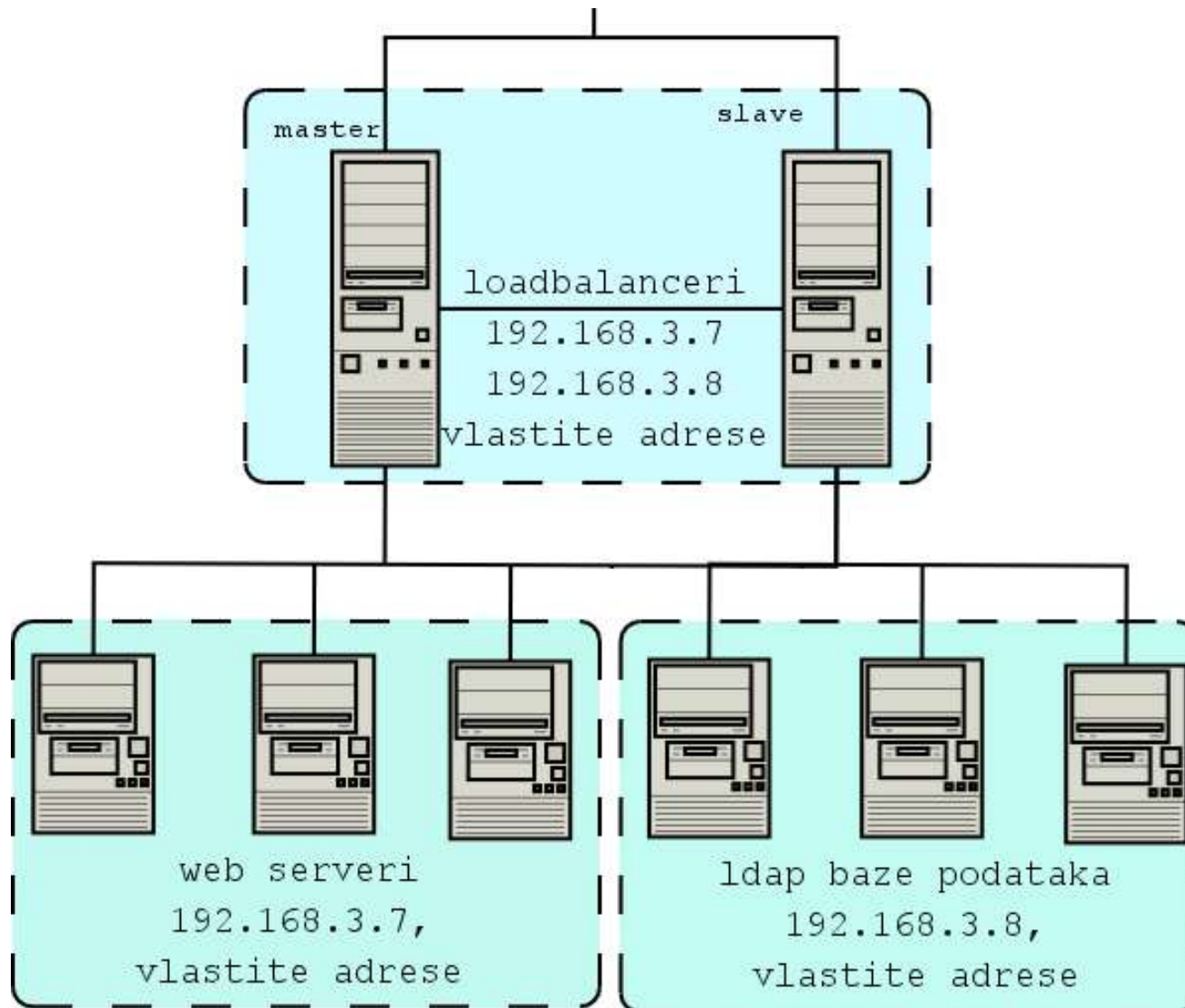
- **visoka dostupnost (high availability)**
 - težnja na dostupnosti usluge, upotreba loadbalancera za failover servisa
 - primjer: Google, Hotmail
- **visoke performanse / paralelno procesiranje**
 - težnja za što bržom obradom informacija
 - primjer: Clusteri za izradu vremenskih prognoza

Potrebne aplikacije

- **IPVS** za linux kernel ($\geq 2.4.23$ ili 2.6 kerneli)
- **ipvsadm** – upravljanje virtualnim serverom
- **heartbeat** – za failover loadbalancera
- **ldirectord** – failover servisa i upravljanje ipvsadmom

Projektiranje clustera

- Primjer jednostavne konfiguracije



Projektiranje clustera

- **Odabir metode za loadbalancing**
 - **NAT** - bez podešavanja računala clustera, svi zahtjevi prema serverima u clusteru idu i vraćaju se preko loadbalancera
 - **TUNNELING** - u posebnim situacijama, potrebno podešavanje računala clustera, sporije, samo dolazni zahtjevi na računala u clusteru idu preko loadbalancera
 - **DIRECT ROUTING** – najbolje performanse, potrebno podešavanje računala clustera, samo dolazni zahtjevi na računala u clusteru idu preko loadbalancera. Ovu metodu koristimo u primjeru.

Projektiranje clustera

- **Loadbalanceri, 2 komada**

- failover računala korištenjem heartbeata, samo je jedan aktivan u jednom trenutku
- master loadbalnacer ima podignute obje adrese koje se dijele (192.168.3.7, 192.168.3.8)

- **Web serveri, 3 komada**

- failover https servisa putem ldirectord/ipvsadm kombinacije
- prvi server može obrađivati dvostruko više podataka od slijedeća dva, dvostruko je snažniji
- svaki web server uz vlastitu adresu ima i adresu 192.168.3.7, ne odgovara na arp zahtjeve

Projektiranje clustera

- **Ldap serveri, 3 komada**
 - failover ldap servisa putem ldirectord/ipvsadm kombinacije
 - adrese ldap servera nisu važne prema van, već samo prema web serverima, svaki ldap server uz vlastitu adresu ima i adresu 192.168.3.8, ne odgovara na arp zahtjeve
 - replikacija ldap baza odvija se neovisno o loadbalancerima, na vlastitim ip adresama

Podešavanje clustera – loadbalanceri

- **heartbeat konfiguracija**

- heartbeat loadbalancera preko serijske veze
- failover dvije ip adrese i pokretanje ldirectorda

<i>/etc/ha.d/ha.cf</i>	<i>/etc/ha.d/authkeys</i>	<i>/etc/ha.d/haresources</i>
logfile /var/log/ha-log logfacility local0 keepalive 2 deadtime 30 warntime 10 initdead 120 serial /dev/ttyS0 baud 19200 node master node slave	<i>auth 1</i> <i>1 crc</i>	master 192.168.3.7/24 ldirectord master 192.168.3.8/24 ldirectord

Podešavanje clustera – loadbalanceri

- **Idirectord konfiguracija**
 - **loadbalancing i failover https-a** na web računalima, za https je potrebno osigurati perzistentnost konekcije. Također je potrebno osigurati da na prvo računalo ide polovica svih web upita (dvostruko je snažnije).
 - **loadbalancing i failover Idapa** na Idap računalima

Podešavanje clustera – loadbalanceri

/etc/ha.d/conf/ldirectord.cf

```
checktimeout=3
checkinterval=6
quiescent=no
autoreload=yes
callback="/etc/ha.d/sync"
```

```
virtual=192.168.3.7:443 masq virt_https
  real=192.168.3.20:443 masq 2 "check.txt","OK" Web1
  real=192.168.3.21:443 masq 1 "check.txt","OK" Web2
  real=192.168.3.22:443 masq 1 "check.txt","OK" Web3
service=https
checktype=negotiate
request="check.txt"
receive="OK"
scheduler=wrr
persistent=1500
netmask=255.255.255.255
protocol=tcp
```

```
virtual=192.168.3.8:389 masq virt_ldap
  real=192.168.3.24:389 masq 1 "", "" ldap1
  real=192.168.3.25:389 masq 1 "", "" ldap2
  real=192.168.3.26:389 masq 1 "", "" ldap3
service=ldap
checktype=connect
scheduler=wrr
netmask=255.255.255.255
protocol=tcp
```

Podešavanje clustera – serveri

- **dodavanje loopback adresa**

za direct routing metodu potrebno je dodati loopback adrese na svako računalo u clusteru koje ima servis nad kojim se radi loadbalancing na loadbalancerima

- **Na linux računalima:**

```
ifconfig eth0:0 192.168.3.7 netmask 255.255.255.255
```

```
echo 1 >/proc/sys/net/ipv4/conf/lo/hidden
```

(potreban patch ako nije mandrake ili suse kernel)

- **na windows operacijskim sustavima:**

Instaliranje loopback mrežnog uređaja i dodavanje adrese s potrebnom mrežnom maskom

Ostali problemi

- Dijeljenje datotečnog sustava:
 - Zajednički medij za pohranu podataka i globalni datotečni sustav (gfs, opengfs,...)
 - rsync, nfs, samba
- Sesije:
 - Održavanje perzistentnosti spajanja
- Programiranje za cluster
 - Povećana složenost programiranja