

open source

Linux Clusteri

Debian Cluster Components

Clustering je metoda spajanja nekoliko računala da rade kao jedno s velikim performansama. Možemo se uvjeriti kako i naši inženjeri znaju razviti projekt koji može parirati stranima

Piše: Ivan Capan

Clusteri u računalnom svijetu su nakupine računala koja zajednički rade neki posao. Ona se u tu svrhu povezuju, umrežavaju, programiraju se zajedničke aplikacije, a konfiguracija se odvija centralizirano. Obično razlikujemo dvije vrste *cluster*a. U serverskom svijetu *cluster* mogu činiti dva ili više računala, od kojih je samo jedno dostupno izvana, a ostala ga mogu zamijeniti u slučaju kvara. To su takozvani *high availability* (HA) *cluster*i koji ne zahtijevaju značajnije planiranja. Nisu namijenjeni povećavanju performansi, nego samo neprekidnoj raspoloživosti, pri čemu svaka minuta nedostupnosti može značiti velike gubitke u prodaji.

U raznim znanstvenim aplikacijama gdje je potrebno obraditi ogromnu količinu podataka koristi se drugi tip *clusteriranja*, u kojem se velik broj računala u serverskoj farmi (pojam kojim se obično opisuje veća skupina računala) predstavlja kao jedan, višestruko većih performansi. Takvi *cluster*i ponekad su sastavljeni čak i od običnih računala, kojima je već istekao "rok trajanja" i nisu više dovoljno brza za neke uredske ili *desktop* zadaće. Oni sasvim dobro mogu poslužiti u *clusteru*, jer slabije performanse nadoknađuju svojim velikim brojem. Upravo takva računala (od nekoliko stotina ili tisuća čvorova) u sprezi s Linuxom tvore neka od najbržih računala na svijetu, koja bi, da su izgrađivana od komercijalnih rješenja, koštala red veličine više.

Beowulf mitologija

Često se za Linux *cluster* koristi izraz Beowulf, no to nije neka konkretna tehnologija, kako se često misli, već gore opisan način povezivanja računala. Postoje dvije smjernice pri izgradnji Beowulf *cluster*a, MPI i PVM. Prvi, tj. *Message Passing Interface*, komunikacijski je protokol za razmjenu podataka između čvorova koji svaki vrte dio posla nekog paralelnog programa (koji zna podijeliti posao u više relativno neovisnih cjelina). Kako biste to vizualizirali, najbolje je zamisliti primjer kodiranja WAV datoteke u MP3. Svaki od čvorova može dobiti komad datoteke na obradu. Svi će čvorovi djeliće MP3 datoteke vratiti glavnom čvoru, koji će ih spojiti.

PVM (*Parallel Virtual Machine*) omogućuje mreži mnogih računala da postane procesorsko-memorijska jedinka. Program koji pokrećemo vidi to kao jedno računalo.



▲ Cluster na IRB-u sastavljen je od klasičnih PC kućišta i off-the-shelf dijelova

Domaća pamet - DCC

Nezadovoljni postojećim rješenjima i izborom distribucija (rpm baziranih), stručnjaci Instituta Ruđer Bošković u Zagrebu odlučili su napraviti vlastite prilagodbe, bazirane na Debian GNU/Linuxu. Tako se on dodatkom DEB paketa vlastite izrade može pretvoriti u glavni čvor u *clusteru*, odnosno instalacijski server. U kolekciji 15 paketa zajedničkog naziva Debian Cluster Components nalazi se nekoliko vlastitih alata te nekoliko paketa čija je namjena konfiguracija već postojećih. Od ostalih paketa vrijedi spomenuti Torque, sustav za raspodjelu procesa po čvorovima, čijem su razvoju serijom zakrpa pridonijeli i IRB-ovi inženjeri. Neke od komponenti sustava su i paket Ganglia za nadzor, DHCP, LDAP i TFTP serveri.

Sustav koji se time izgrađuje je po prirodi MPI: centralni čvor zaprima zahtjeve za obradom poslova te ih raspodjeljuje čvorovima. U svakom trenutku zna koliko koji čvor ima resursa na raspolaganju (procesori, memorija, prioritet poslova koji se već obrađuju) i kad nađe neki koji je slobodan, pošalje mu posao na izvršavanje.

Mrežna instalacija

No, osim odličnih mogućnosti *clusteriranja*, ono što ovaj paket izdvaja je mogućnost instaliranja velikog broja računala preko mreže s centralnog servera. Nakon što se na server instaliraju potrebni paketi, potrebno je pripremiti zasebnu sliku operacijskog sustava koji će

se poslati klijentu. Zasad je podržan samo Debian. Nakon što se on konfigurira u dovoljnoj mjeri, u njega instaliraju svi potrebni paketi; dovoljno je da nekom računalu u mreži zadamo (u BIOS-u) da se pokreće preko mreže. On će pronaći server, s njega preuzeti sliku cijelog svog operacijskog sustava i prebaciti je na lokalni tvrdi disk. Već pri idućem *restartanju* podići će se novi sustav, kao da smo proveli zadnjih par sati u ručnoj instalaciji.

Iako je svrha ovog načina instalacije prvotno bila podizanje *cluster*a, gdje bi ručna instalacija većeg broja računala trajala danima, može se prilagoditi bilo kakvoj svrsi, na primjer u računalnim učionicama ili raznim *desktop* i uredskim strojevima. Na serveru je moguće pripremiti nekoliko različitih slika sustava, zavisno od potreba.

Glavne komponente sustava već su pogonu na tri *cluster*a, a u siječnju bi trebala izići verzija 0.1. Trenutna faza razvoja je pisanje dokumentacije, a status se može provjeriti na www.irb.hr/en/cir/projects/dcc.

Rezultat ovog opsežnog *open source* projekta zapravo je plod entuzijazma nekolicine ljudi. Tim s IRB-a čine Nikola Pavković i Valentin Vidić u ulozi razvijatelja te dr. Karolj Skala kao voditelj projekta. Iako u ovom trenutku ne postoji izvor financijskih sredstava, u susret DCC-u je izašla tvrtka Pop&Pop, izradivši logotip projekta.