

open source

Linux backupi

Rano i često (1)

Slogan cijele *backup* industrije (*early-and-often*) kaže nam da kopiju moramo napraviti uvijek prije kvara. Osim vremenske procjene, moramo dobro odabrati i medij i metodu, a Linux ima veliku mogućnost izbora na tom polju

Piše: Ivan Capan

U prošlom je broju opisan *backup*, njegovi principi, mediji, softver, sve što ga sačinjava. Kako ne bismo ponavljali već rečeno, posvetit ćemo se nekim njegovim specifičnostima koje se odnose na Linux. Kako je on dosta orijentiran serverskim primjenama, tako je iz Unix svijeta naslijedio neke koncepte, poput spremanja na trake i pripadajućeg softvera. Linux je snažan u skriptiranju, pa ćemo tako dati temelj za izradu vlastitih skripti koje će se u velikom broju slučajeva ponašati bolje nego veliki *backup* programi. Istodobno ćemo dobiti uvid od čega se točno sastoji naša sigurnosna kopija podataka, a upravo to bi kod njenog vraćanja u slučaju havarije moglo biti presudno, jer se onda očekuje trenutno reagiranje, a ne proučavanje dokumentacije.

Kilometri trakica

Jedan od najstarijih medija za *backup* su magnetske trake raznih formata, koje su i danas standard na mnogim serverima. Najjeftinije i najslabijih performansi su DAT (Digital audio tape) trake, originalno zamišljene za spremanje audiozapisa, dok se nije razvio DDS format zapisa digitalnih podataka. Postoje još i QIC, Travan, DLT (Digital linear tape), pa AIT, te LTO i njen Ultrium format traka. Zadržat ćemo se u donjem cjenovnom segmentu, jer će za kućne i manje serverske zahtjeve DAT trake sasvim dobro poslužiti, pogotovo kad znamo da njihov kapacitet ide do 20 GB nekomprimirano (DDS-4), a brzine se kreću oko 2-3 MB/s. Ta brzina je moguća zbog dijagonalnog zapisa podataka na traku, poput VHS-a, u kojem glava ne miruje, već se rotira velikom brzinom, dok je linearna brzina trake relativno mala. Podaci se tako zapisuju helikoidno, u ukošene trake.

Prednost traka je i trajnost podataka, pod uvjetom da su uredno uskladištene i da se ne nađu u magnetskom polju. Često se spominje podatak od deset godina trajnosti. Glavni nedostaci su brzina prijenosa podataka i linearni pristup, u kojem se glava ne može direktno pozicionirati na konkretni podatak na traci, jer ga ne može locirati. U sustavima gdje je svaka minuta nedostupnosti jako skupa, to može predstavljati problem. Vraćanje cijelog *backupa*, pa i jedne jedine datoteke dugo traje.

Tar

Korištenje traka oduvijek se svodilo na upotrebu softvera znanog pod imenom *tar* (*tape archive*) koji je danas daleko najrašireniji format pakiranja datoteka. Uređaji se povezuju preko SCSI kontrolera, a njihova oznaka je `/dev/nst0` ili `/dev/st0`. Obje se oznake odnose na isti uređaj, a razlika je u tome da potonji zna da se



treba premotati na početak nakon završetka operacije. Program za kontrolu naziva se `mt`. Za provjeru trake pišemo `"mt -f /dev/st0 status"`. Traka će se sama premotati na početak i možemo ju početi koristiti. Ako želimo zapisati nešto na uređaj, koristit ćemo `tar` na sljedeći način: `"tar cpvf /dev/st0 /home"`. Za provjeru sadržaja trake: `"tar tvf /dev/st0"`. Za otpakiranje sadržaja na disk: `"tar xvf /dev/st0"`. Za dohvat jedne datoteke, na kraj prethodne naredbe dodat ćemo kompletnu putanju do nje. Ta procedura može potrajati iznimno dugo, jer se moraju slijedno pročitati sve datoteke prije nje. Sve ove operacije premotavat će traku na početak, dakle bit će moguća samo jedna arhiva po jednoj traci, što je u redu ako je tako dimenzionirana. Ako želimo da nam se arhive nastavljaju jedna iza druge jer su manje, pisat ćemo ime uređaja kao `nst0`, pa će traka ostati na mjestu nakon završetka neke operacije. Po arhivama se pomičemo s naredbom `"mt -f /dev/nst0 fsf 1"` što će preskočiti prvu i pozicionirati traku na početak druge `tar` arhive. Ako pak imamo podataka koji su veći od jedne trake, opcijama tara dodajemo slovo `M` kako bi se arhiva nastavljala. Istu opciju koristimo i kod vraćanja podataka.

Na kraju snimanja pokrećemo naredbu `"mt -f /dev/st0 rewoff"` koja će traku premotati i izbaciti iz uređaja. Jako ju je bitno označiti nekom jedinstvenom oznakom, a također i naredbom kojom je arhiva stvorena, kako kasnije ne bismo imali problema kod vraćanja. Moguće ju je malom mehaničkom sklopkom

zaštititi i od pisanja. Ako ne želimo kod svake naredbe pisati oznaku uređaja, možemo napraviti simboličku poveznicu: `"ln -s /dev/st0 /dev/tape"`.

dump

Osim tara, često se za zapis podataka koriste još i `dump` i `cpio`. `Dump` je svojedobno došao na loš glas jer je čak i Linus Torvalds izrazio svoje negativno mišljenje o njemu. Njegova je mana što je jako osjetljiv na promjene na "živom" sustavu, odnosno particiji koju sprema jer radi direktno s uređajem, zaobilazeći Linux rutine za odgođeno spremanje podataka (*write cache*). Ako se neka datoteka promijenila u nekom trenutku, `dump` će ju pogrešno spremići čak i neko vrijeme nakon što je promjena završena, jer se još nisu sve promjene stigle



▲ Što je jedan tvrdi disk prema zastrašujućim ormarima od par desetaka ili stotina istih?