

RSYNC protokol, nástroj

- Michal Fiala <fiala@mfiala.net>
- http://mfiala.net/download/x36mti_rsync.tgz



Co je rsync?

- platformě nezávislý nástroj pro synchronizace datových uložišť
- využívá algoritmu takéž nazvaného rsync od kodéra Andrew Tridgella
- prvotním záměrem bylo minimalizovat datový tok při synchronizaci – šetřit přenosové pásmo
- synchronizace postavena na síťovém modelu klient-server, kdy server naslouchá na TCP portu 873 – lze však využít i rsh

Jádro algoritmu rsync 1

- soubory A, B
- Problém zní, chci aby B obsahovala data jako A
 - zkopírovat A do B ? (lokálně OK)
- soubory A, B jsou na různých strojích, vzájemně propojené pomalou linkou
 - komprese a přesun ? (lepší než nic, malá akc.)
 - přesunout pouze rozdíl, ale jak zjistit rozdíl, když nejsou na stejném uložišti?

Jádro algoritmu rsync 2

- problém řeší rsync protokol
- efektivně zjistí, které části souboru A je zapotřebí přenést tak, aby se soubor B mohl zaktualizovat do podoby souboru A
- jak na to? nejdříve rekapitulace problému
 - stroj α drží soubor A, stroj β drží soubor B
 - A a B jsou podobné soubory
 - mezi α a β je pomalá linka
 - soubor B chceme zaktualizovat dle souboru A

Jádro algoritmu rsync 3

- β rozdělí soubor B do série neopakujících se datových bloků o konstantní velikosti S, poslední blok může být menší nežli S
- β pro každý z bloků vypočítá 2 hashe (silný hash – 128b, slabý rotující hash – 32b)
- β pošle všechny hashe straně α
- α se pokusí nálezt v celém A bloky specifické zaslányými hashi (libovolné posunutí, rychlé díky rotujícímu hashi)
- α posílá β sekvencí instrukcí jak sestavit soubor B (ukazatel na existující blok dat v B nebo data)
- β nyní disponuje instrukcemi dle kterých lze poskládat soubor B – získal je během jednoho roundtripu

Celý algoritmus rsync - role

- doposud jsme se zabývali pouze synchronizací dvou přesně daných souborů
- jak probíhá synchronizace celého adr. stromu?
- role v komunikaci
 - klient: inicializuje spojení
 - server: vzdálený rsync proces nebo systém, ke kterému se klient připojuje (rsyncd, ssh)
 - démon: rsync démon, který čeká na příchozí požadavky od rsync klientů

Celý algoritmus rsync - role

- pokračování rolí
 - vzdálený shell: jeden nebo více procesů, které zprostředkovávají komunikaci
 - vysílač: rsync proces, který má přístup ke zdrojovému souboru (vzor)
 - přijímač: rsync proces, který má přístup k cílovému uložišti
 - generátor: proces, který zastřešuje souborovou logiku na straně přijímače

Celý alg. rsync - inicializace

- Inicializace
 - spuštění rsync klienta
 - spojení se serverem (roura, síťový socket)
 - roura – rsh (ssh), lokální synchronizace
 - síťový socket – rsync -> rsyncd
 - domluva na verzi protokolu rsync

Celý alg. rsync – seznam souborů

- Seznam souborů
 - obsahuje cestu k souborům, vlastnictví uživatele a skupiny, práva, velikost, čas poslední modifikace (checksum)
 - během sestavování odesílá vysílač položky příjemci, po dokončení se seznam setřídí dle názvu souborů a opatří indexy
 - příjemce provede fork, nově vytvořeny proces je generátor

Celý alg. rsync – generátor

- pracuje se soubory na straně přijímače, porovnává a provádí příslušnou činnost
- pokud je definováno --delete, odstraní příslušné soubory
- sekvenčně prochází seznam souborů a určuje, zdali je možné je přeskočit – kontrola času modifikace, velikosti (checksum)
- nepřeskakují se adresáře, speciální soubory, symbolické linky
- soubory, které se nepřeskočily jsou označeny za bazové
- z bazových souborů se vytvoří hashe a odešlou vysílači

Celý alg. rsync – vysílač

- viz jádro algoritmu rsync, tedy stručně
- nalezne dle zasláního indexu synchronizovaný soubor
- pokusí se nálezt v celém souboru bloky specifické zasláními hashi
- vysílač zašle přijímači sekvenci instrukcí jak sestavit synchronizovaný soubor
- na konci komunikace v rámci jednoho souboru vysílač provede hash celého souboru a zašle přijímači

Celý alg. rsync – přijímač

- obdrží instrukce od vysílače, jak sestavit synchronizovaný soubor
- otevře bazový soubor ale také vytvoří dočasný soubor
- do dočasného souboru zapisuje data dle instrukcí
- zkontroluje hash, pokud 2x za sebou chybný u stejného souboru, rsync končí s chybou, jinak nastavení atributů a přesun na místo bazového souboru
- vysílač a přijímač opakují operace přes všechny nepřeskočené položky seznamu souborů

Klady

- optimalizace velikosti přenášených dat (protokol, komprese)
- kompletní synchronizace uložště různou formou (vlastníctví, práva, čas, symlink, hardlink, spec. soub.)
- kontrola integrity přenášených dat
- vysoká efektivita díky protokolu s nízkou režíí
- nástroj rsync je výtečně konfigurovatelný – silný a robustní pomocník

Zápory

- rsync není klasickým protokolem (žádná hlavička, žádné tělo, žádná přesná specifikace obsahu)
- tok Bytů, sémantika dle prováděné akce
- špatně dokumentovatelný, rozšířitelný, pozorovatelný
- náročný na operační paměť na straně vysílače i přijímače (seznam souborů, průměrně 100B na položku)
- náročný na CPU výkon u vysílače – počítání hashů
- náročný na IO operace u přijímače – práce s dočasnými a bázovými soubory

Jednoduché příklady 1

- `rsync [OPTION]... SRC [SRC]... DEST`
- `rsync -av --progress /src/foo /dest`
 - Inkrementální synchronizace v rámci jednoho systému
 - `a = -rlptgoD`
(archive = recursive link perms times group owner devices)
 - `v = verbose`
 - synchronizace celého adresáře `foo` včetně jeho samotného

Jednoduché příklady 2

- `rsync -avz --progress -delete \`
`/src/foo/ host::mname/dest`
 - synchronizace mezi dvěma hosty
 - z = compress
 - komunikace rsync -> rsyncd přes síťové sockety
 - synchronizace pouze obsahu adresáře foo

Jednoduché příklady 3

- `rsync -a --delete-after \`
`--rsh='ssh -p55055' /backup \`
`buser@bhost.cz:/backup/part1`
 - vhodný zálohovací rsync příkaz
 - úplné zrcadlo
 - spojení pomocí ssh, vhodné doplnit klíčovou autentizací

Šetření času

- ÚKOL: máme několik stanic stejné hw konfigurace, jak je co nejrychleji všechny přivést do použitelného stavu?
 - kompletně připravit jeden stroj, nakonfigurovat a spustit rsync demona, vyexportovat /, umožnit přístup přes účet root, provést restrikcí na úrovni packet filtering
 - na ostatních strojích pouze provést inicializaci disků, vše ostatní přenést pomocí rsync, nainstalovat boot loader a změnit konfliktní nastavení (sít', hostname)

Rsync v akci

- služba rajce.idnes.cz – poskytování statických dat
 - data jednoho uživatele jsou uloženy v jednom adresáři
 - uložení obsahuje cca 55 000 uživatelů
 - uložení obsahuje cca 1,5TiB dat, průměrná velikost souboru řádově v desítkách KiB
 - velké množství adresářů a souborů o malé velikosti, minimální změny
 - služba běží na jednom výkonem serveru, možnost pouze horizontální škálovatelnosti

Rsync v akci 2

- Úkol: přemigrovat uživatelská data na nové dynamické uložení tak, aby to mělo minimální dopad na provoz služby
- (I) `nice -n5 rsync -a /stare /nove`
- (II) `nice -n5 rsync -a --size-only --whole-file /stare /nove`
- (III) `find /stare -type d -maxdepth 1 | while read f; do
 nice -n5 rsync -a --size-only --whole-file /stare/${f} /nove
done`
- (IV) `find /stare -type d -maxdepth 1 | while read f; do
 nice -n5 cp -a /stare/${f} /nove
done`

Rsync

Děkuji za pozornost
Dotazy?