Teeme mõisted selgeks: mis on replikatsioon ja Continuous Data Protection

10 aastat tagasi - 06.09.2016 Autor: Ivar Pikker

? Tihti räägitakse andmevarunduses replikatsioonist ja Continuous Data Protectionist. Kuid mis need täpsemalt on ja kuidas või mis olukordades seda vaja läheb?

! Nende kahe mõiste puhul võib pidevalt kohata puterdamist ja segamini ajamist. Sisuliselt on tegemist täiesti erinevate tehnikatega. Replikatsioon tähendab et mingid andmed peegeldatakse kas siis uni-direction (ainult ühes suunas) või bi-direction (mõlemas suunas, sünkroniseering) teisel asukohal olevate andmetega. Continuous Data Protection (CDP) aga tähendab reaalajas varundust.

Replikatsiooni võib laias laastus jagada kaheks - *low-level* sektorpõhiseks (terve ketas korraga) ja *high-level* failipõhiseks. Sektorpõhist replikatsiooni tarkvara on rohkem esindatud. Liidriks võiks pidada Veritast. Veritase replikatsioonitarkvara alustas nimega Volume Manager, siis muutis Storage Foundation-iks ja nüüd on InfoScale nime all. Veritas Volume Replicator on kogu tarkvarapaketi üks osa. Veritase replikatsioonitarkvara on väga töökindel ja pika ajalooga. Võrdluseks näiteks Veritase NetBackup-i ei saa kuidagi töökindlaks tooteks pidada. Veritase Replicator töötab kas sünkroonselt (kettale kirjutuse lõpu teade saabub opsüsteemile peale mõlemale kettale kirjutust), või asünkroonselt (opsüsteemile teade saabub kohe peale lokaalsele kettale kirjutust ja remote kettale kirjutus jätkub). Kõike saab detailselt seadistada.

Failipõhine replikatsioon aga on hoopis teine maailm ja kohati isegi komplitseeritum. Failipõhine replikatsioon on samuti kas *uni-direction* ehk ühesuunaline või *bi-direction* ehk kahesuunaline. Üldjoontes käib asi nii. Kõigepealt tehakse algne sünkronisatsioon/replikatsioon, nagu ka sektorpõhise puhul. Remote failiserver saadab üle võrgu kataloogipuu listi koos failide suuruse, muutmisaja ja parameetritega. Lokaalne failiserver võrdleb seda enda listiga ning otsib muudatused mis tuleks teha remotes (kopida fail, kirjutada üle, kustutada, muuta parameetrid). Kopimine toimub block-levelil faili suhtes, seega klassikaline delta-copy. Edasi kontrollib service opsüsteemi poolset failide kirjutust/kustutust. Mingit open-file manageri või VSS-i ei kasutata - oodatakse seni, kuni fail suletakse. Remote failiserverisse kopitakse failid kõigepeat ajutisse kausta ning

sealt alles õigesse kausta. Võimsaim toode selles vallas on firma Attunity tarkvara RepliWeb. See on tähelepanuväärselt mahukas tarkvaraplatform, lehekülgede kaupa seadistusi, igat liigutust saab detailselt seadistada. RepliWeb on uni-direction spetsiaalse download liidesega (enda muudatusteks saadakse andmed eraldi serverist) ja võimaldab replikeerida mitmesse failiserverisse korraga. Unikaalne võimalus on tal salvestada kõik replikatsiooni toimingud (kaasa arvatud failide kustutused) ning soovi korral hiljem "rollbackida" remote failiserveri kaustad ajas tagasi. Peale selle omab ta veel mooduleid tuntumate aplikatsioonide replikeerimiseks (webiserver jne), siis kõrgemal aplikatsioonide levelil kui seda on failisüsteem.

Teine failipõhine replikaator on Peer Software´i "PeerSync". See on väiksem tarkvaraplatform, kuid pisut teise kasutuseesmärgiga, rohkem paralleelseks dokumenditöötluseks või grupitööks mõeldud. Tema unikaalne omadus on failide lukustamine teistes failiserverites kui ühes toimub dokumenditöötlus. Midagi sarnast Sharepointi dokumenditöötluse reeglistikule.

Continuous Data Protection (CDP) aga on hoopis varunduse teema ja pole midagi ühist replikatsiooniga. CDP töötab ainult failipõhiselt, ei kasuta VSS-i, ei kasuta shedulingut. Kunagi leiutati selline mõiste nagu "near CDP", siis alternatiiviks "true CDP". On üks selline lõbus ütlemine nagu "poolrase", viidates et see pole võimalik. Vot see "near CDP" ongi see poolrase. Kuna CDP sai mõni aeg tagasi väga populaarseks tehnikaks, siis marketingikaalutlustel tahtsid kõik firmad näidata et ka nende toode võimaldab CDP-i, kuigi tegelikult mitte. Nad üritasid näidata et varundus toimub väga lühikeste ajavahemike järelt. Loomulikult pole siin tegemist mitte millegi muuga kui tavalise *scheduling*iga, mis on lihtsalt pandud tihemini tööle. CDP-iga aga pole siin midagi pistmist.

Teine hämamise koht seisneb *block-level*i teemal. Kui keegi räägib *block-level*ist, mis pole mitte failipõhine *delta-copy*, vaid *low-level* failisüsteemi sektorite põhine ning nimetab seda CDP-iks, siis tegelikult pole tegemist mitte CDP-iga vaid replikatsiooniga. Näiteks Microsofti Data Protection manager (tänaseks lõpetatud projekt) on sisuliselt sektorpõhine replikatsioon koos tavalise varunduse shedulinguga. Ka Veritase Volume Replicator teeb täpselt sama kui kasutada FlashSnapi. Klassikaline CDP varundab faili ainult selle sulgemisel, seni ootab. Ka failide kustutused fikseeritakse. Saab seadistada seda, kui kaua hoitakse kustutatud ja vanemaid versioone. Puudub *point-in-time* ketta esitus - iga fail on sõltumatu teistest. Üks esimesi pioneere selles on kunagine 1998. aastal loodud firma Storactive toode LiveBackup, siis peale omaniku vahetust Atempo LiveBackup nime all ja täna ASG Live-Navigator. Selle käest-kätte jooksutamise

tõttu on see pisut varju jäänud toode.

Teine, palju tuntum toode on Symanteci Desktop and Laptop Option (DLO), täna Veritase nime all. See on LiveBackupiga üllatavalt sarnane ja arvata võib et siin on pisut ideid piilutud teise tagant. Aga Veritas DLO on väga hea ja töökindel toode, väga mugav ja praktiline ning paindliku seadistusega.

Protsess töötab madala prioriteediga ja pole karta et CPU-d hakkaks koormama või teist programme segama, isegi kui palju faile on järjekorras ootel. Ainus viga, et agent serveri opsüsteemis ei tööta, server ise aga küll, isegi AD domeeni nõuab (klient mitte). Ei oska öelda, miks selline veider piirang. Kui keegi teab paremat softi, mis ka serverist varundaks, siis oleks põnev teada saada.

- Lahendused

- Salvestusseadmed
- Serverid

Pilt

