



# Fault Tolerant Systems

## Vysoká dostupnosť v praxi

Keď som sa konečne dostal k písaniu tohto článku a zamyslel som sa nad jeho témou, napadol mi výrok bývalého kolegu: „Dnes chcú všetci všetko a hneď. Pamätáte sa, ako to bolo pred pár rokmi? V kancelárii zazvonil telefón: 'Dobrý deň, potreboval by som hovoriť s pánom Novákom.' 'Prepáčte, kolega Novák odcestoval na pracovisko v Košiciach, skúste o týždeň, prípadne mu pošlite list.' A teraz, ak vás niekto v okamihu nezoženie na mobilnom telefóne alebo minimálne e-mailom, je z toho problém...“ Tento výrok zdôrazňoval najmä nutnosť dostupnosti človeka, napríklad servisného technika alebo obchodníka pre svojho zákazníka. Dnes už je skutočne úplne samozrejme byť dostupný pre svojich klientov, a teda chlebobdarcov 24 hodín denne a 7 dní v týždni. Na mobilnom čísle, e-mailom, prostredníctvom webovej aplikácie... Stačí vlastniť patričnú technológiu, resp. zariadenie – mobilný telefón a počítač. Stačí to? Vonkoncom nie. Naša dostupnosť klientovi bude len taká, aká bude dostupnosť komunikačného kanála, prostredníctvom ktorého s ním komunikujeme – reálne to v tomto prípade teda bude

pokrytie signálom mobilného operátora, dostupnosť jeho služby, spoľahlivosť a dostupnosť linky poskytovateľa internetového pripojenia. Slovo dostupnosť teraz budeme skloňovať vo všetkých možných pádoch.

### Čo je dostupnosť – parametre, čísla a zaklínadlá

Dostupnosť v IT terminológii je parameter určujúci schopnosť systému, aplikácie alebo riešenia byť k dispozícii prijímateľovi služby, teda používateľovi. Udáva sa v percentách a veľmi jednoducho sa dá vypočítať podľa nasledujúceho vzorca:

$$A [\%] = \text{MTBF} / (\text{MTBF} + \text{MTTR}) * 100$$

kde:

**A = Availability** = dostupnosť systému (riešenia) v percentách

**MTBF = Mean Time Between Failure** = čas medzi dvoma výpadkami služby – býva definovaný v hodinách, pričom dôvod výpadku je irelevantný. Tento parameter sa okrem vzťahu s celkovou dostupnosťou riešenia používa aj na vyjadrenie poruchovosti hardvérových komponentov a spravidla tento údaj pre každý kompo-

nent (disk, procesor, pamäť alebo server...) udáva výrobca v hodinách (typicky sú to tisícky, prípadne desiatky tisíc hodín).

**MTTR = Mean Time To Repair** = čas potrebný na opätovné obnovenie služby po výpadku (opäť udávaný v hodinách), pričom nezáleží na spôsobe a forme obnovy a na tom, koľko krokov a v ktorých úrovniach je potrebných na takéto obnovenie.

**Príklad:**                      **MTBF = 4000 hodín**  
    **MTTR = 2 hodiny**  
    **A [%] = 99,95**

Dostupnosť sa často vyjadruje vo vzťahu k perióde 1 rok:

$$A [\%] = (\text{Runtime}(\text{year}) - \text{Systemdown}(\text{year})) / \text{Runtime}(\text{year}) * 100$$

Ročnému runtimu zodpovedá 8760 hodín. Ak Systemdown je definovaný ako  $8760 * \text{MTTR} / \text{MTBF}$ , potom tieto dve výjadrenia dostupnosti nie sú identické. No ak uvažujeme, že  $\text{MTTR} \ll \text{MTBF}$ , čo je aj v praxi reálne, potom výsledky oboch výrazov sú zhodné:

$$A [\%] = (8760 - 8760 * 2 / 4000) / 8760 * 100 = 99,95\%$$

Pri výpočte celkovej dostupnosti komplexnej

infraštruktúry treba mať na zreteli, že táto infraštruktúra pozostáva nie z jedného, ale z niekoľkých elementov, navzájom prepojených buď sériovo, alebo paralelne, čo sťažuje a komplikuje teoretický výpočet a uvedené vzorce už na tento model nestačia, sú však vhodným vodidlom. Na ilustráciu: v prípade výlučne sériového napojenia stáva sa infraštruktúra nedostupnou, ak čo len jeden jej element sa stane nedostupným. V prípade výlučne paralelného napojenia stáva sa infraštruktúra nedostupnou, ak v prípade, že pozostáva z dvoch elementov, dôjde k výpadku druhého elementu, pričom prvý element je už nedostupný. Výpočet dostupnosti sériovo a paralelne spojených elementov prekračuje rozsah tohto článku, podrobnosti môžete nájsť v literatúre venujúcej sa tematike vysokej dostupnosti.

Nasledujúca tabuľka obsahuje jednoduchý prehľad dostupnosti v percentách, transformovanej na dobu výpadku počas jedného roka:

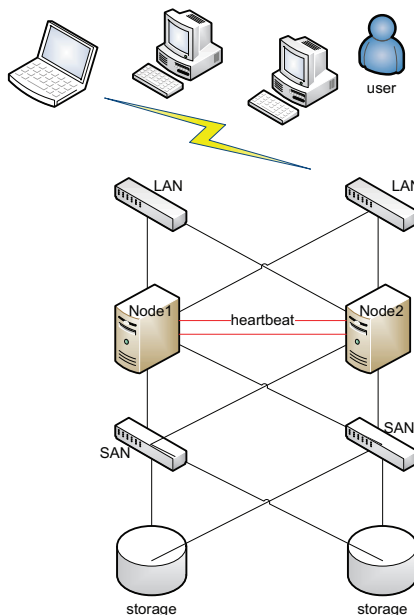
Dostupnosť %	Výpadok služby za 1 rok
99,9999	30 sekúnd
99,999	5 minút
99,99	50 minút
99,9	9 hodín
99	3,7 dní
90	37 dní

### Od hotswap disku ku geoklastrom

Dostupnosť, resp. spoľahlivosť jednotlivých hardvérových komponentov vysoko dostupnej infraštruktúry býva spravidla zabezpečená redundanciou, záložnými prvkami, a teda toleranciou týchto systémov k chybám. V prípade serverov je to spravidla redundancia všetkých kritických komponentov zariadenia – pamäťových modulov, napájacích zdrojov, chladiacich jednotiek, systémových diskov, host bus adaptérov (HBA) pripojených do SAN, pri storage systémoch je to podobne ako pri serveroch redundancia napájania, chladenia, zrkadlenie pamäte cache, disky konfigurované do tzv. RAID (Redundant Array of Independent Disks) skupín, hot-spare disky, redundancia kontrolerov, procesorov a pod. Podobne je to aj pri prepínačoch SAN a iných aktívnych prvkoch. Touto redundanciou spĺňajú systémy požiadavku No Single Point of Failure. To však nestačí, dôležité je, aby

rovnako spĺňali aj požiadavku No Single Point of Repair, teda aby chybný komponent, ktorý je v systéme redundantný, a teda systém je odolný proti jeho chybe, bol aj servisovateľný alebo vymeniteľný bez nutnosti prerušenia prevádzky tohto systému. To býva zabezpečené tzv. hot-swap funkcionalitou týchto kritických komponentov. Hot-swap disky, zdroje, ventilátory a pod. sú už dnes v serveroch a diskových poliach úplne bežná záležitosť, pri high-end systémoch a diskových poliach sú to bežne aj hot-swap procesorové dosky, kontrolery, pamäťové moduly a pod.

Keďže dostupnosť a spoľahlivosť jednotlivých systémov komplexného riešenia veľakrát nestačí, či už z dôvodu možných katastrofických vplyvov (povodne, požiare...), alebo ľudského faktora a iných vplyvov, redundanciu treba pri kritických riešeniach zabezpečiť aj na vyššej úrovni. Na to slúžia sofistikované klastrové riešenia, ktoré zabezpečia dostupnosť aplikácie, resp. služby tým, že v prípade neplánovaného alebo i plánovaného výpadku jedného alebo viacerých zdrojov výpočtového výkonu alebo kapacity zabezpečia



Na obrázku je zobrazené typické klastrové riešenie s redundanciou infraštruktúrou

automatické prepnutie dátovej komunikácie na záložné systémy. Tento proces sa v klastrovej terminológii nazýva Failover. Záložné systémy pritom počas normálnej prevádzky môžu vykonávať úplne inú činnosť, ktorá býva zvyčajne v prípade Failover scenára vypnutá. Samozrejme, dobre navrhnuté klastrové riešenie neopomína redundanciu všetkých aktívnych elementov infraštruktúry. Čo sa týka zabezpečenia dostupnosti dát, resp. zariadení na to slúžiacich, t. j. diskových polí, vo vysoko dostupnej infraštruktúre sa odporúča dáta replikovať z primárneho na záložný systém, a to buď za pomoci storage softvéru na úrovni diskového poľa, alebo tzv. host-based prostredníctvom softvéru rôznych výrobcov. Každá z týchto techník má svoje výhody aj nevýhody.

### Konzistencia dát na prvom mieste pred dostupnosťou

Aj moderné klastrové riešenia sa za určitých podmienok (typicky tzv. Split-brain, ktorý nastáva pri prerušení heartbeat komunikácie medzi jednotlivými nódami) môžu dostať do stavu, kde logika klastra nie je schopná rozoznať, ktorá klastrová noda je z hľadiska spoľahlivosti oprávnená vlastníť dátové zdroje, a teda do nich aj zapisovať, ktorá noda „žije“ a ktorá je už eliminovaná. Práve v takýchto prípadoch sa dobre navrhnuté a fungujúce klastrové riešenie zachová tak, že jednoducho eliminuje všetky nody a nedovolí zapisovať dáta ani jednej až do vyriešenia problému spôsobujúceho daný stav. Nekontrolovaný prístup klastrových nód k dátam by totiž mohol mať katastrofálne následky. V tomto prípade sa pred dostupnosťou služby uprednostňuje spoľahlivá ochrana dát, ktoré sú prostredníctvom služby poskytované.

### „Máme vysoko dostupné storage riešenie, takže už nemusíme zálohovať.“

Omyl, omyl, omyl!!! Tento výrok počuť z úst zákazníkov ešte stále dosť často. To, že investujete do sofistikovaného úložného systému, ktorý je sám osebe dostupný na úrovni 99,999 % a dokonca je replikovaný na rovnaký systém umiestnený v záložnej lokalite, ešte neznamená, že sa

PosAm

USEFUL  
TECHNOLOGIES

need+storage+useful

PosAm Infraštruktúrne riešenia<sup>UT</sup>  
[www.posam.sk](http://www.posam.sk)Hitachi Data Systems Gold Solution Partner  
Hitachi Data Systems Authorized Service Provider

konečne nemusíte zaoberať zálohovaním toho najcennejšieho, čo máte a prečo vlastne vysoko dostupnú infraštruktúru budujete, teda svojich dát. Ani to najinteligentnejšie storage riešenie nedokáže eliminovať starý známy ľudský faktor, ktorý býva stále jednou z najčastejších príčin straty dát. Úmyselne či neúmyselne zmažané dáta vám nevráti ani klaster, ani super spoľahlivý redundantný storage systém, vráti vám ich jedine dobre navrhnuté riešenie na zálohovanie dát.

### Kto potrebuje vysokú dostupnosť a kedy?

Kto vlastne potrebuje vysoko dostupné systémy a riešenia a v akej úrovni? V zásade sa dá povedať, že takéto riešenie potrebuje každý, pre koho má strata dát alebo nedostupnosť služby, ktorú poskytuje svojim klientom, výrazný dosah na jeho ďalšiu existenciu, aktuálny alebo budúci profit. V konečnom dôsledku by po finančnom vyčíslení mal byť tento dosah väčší ako sama investícia do vysoko dostupného systému, ale nemusí to byť nevyhnutne tak. Veľakrát práve nefinančné straty majú pre vlastníka dát rovnaký význam ako tie finančné. To hlavne v prípade, ak ide už dokonca o stratu dát. Vysoko dostupné systémy samy osebe neriešia problém zabezpečenia a uloženia dát, ich úlohou je garantovať k dátam prístup pokiaľ možno nepretržite, resp. počas vopred definovanej doby. Azda najjednoduchší príklad sú služby mobilných operátorov alebo vôbec poskytovateľov v oblasti komunikácie a multimédií a bankový sektor. Viete si predstaviť, akú stratu by utrpel priemerný európsky mobilný operátor v prípade, že by jeho zákazníci z dôvodu nedostupnosti služby neboli schopní komunikovať po dobu 30 minút? A viete si predstaviť, že by to trvalo 5 hodín? Skúste si to vyrátať... Alebo z iného pohľadu – mali by ste chuť ako zákazník zotrvať v banke, ktorá má permanentné problémy s dostupnosťou dnes bežne využívaného internet bankingu, prípadne nie je schopná realizovať vaše transakcie okamžite, resp. v požadovanej lehote? Objednávali by ste si tento rok letenku cez agentúru XY aj po tom, čo sa vám ju minulý rok podarilo rezervovať až na piaty pokus, pričom v deň odletu vaša rezervácia zrazu neexistovala? Alebo príklad z pracovného prostredia: Stáva sa vám, že dostanete e-mail od administrátorov vášho IT oddelenia v znení: „Dnes od 18:00 do 19:00 bude mailový server nedostupný z dôvodu údržby.“ A vy práve v tom čase čakáte dôležitý, ba až strategický e-mail od obchodného partnera a musíte naň okamžite reagovať? Teraz už viete, kto všetko vysokú dostupnosť potrebuje a približne v akom rozsahu, v akej úrovni... Implementácia systémov s vysokou dostupnosťou by

však nemala byť samoučelná a tento krok a najmä zvolená úroveň dostupnosti by vždy mali byť výsledkom rozhodnutia manažmentu korporácie na základe dôkladného zváženia a analýzy všetkých rizík a finančných dosahov nedostupnosti prevádzkovej služby. Úroveň vysokej dostupnosti má logicky výrazný vplyv na cenu implementovaného riešenia. Päťminútový výpadok platobných terminálov v supermarkete zrejme nespôsobí dramatický problém, hodinový však už môže. Hodinový výpadok systému SAP v deň uzávierky asi tiež nebude mať až taký negatívny dosah na sled účtovných procesov podniku, celodenný však už môže byť problém.

### Plánované a neplánované výpadky

Možno ste jedným z budúcich obstarávateľov nového informačného systému alebo kompletnej IT infraštruktúry a poviete si: „Takže už máme vo všetkom jasno, oboznámili sme sa s principiálnou architektúrou vysoko dostupných systémov a riešení, vieme použiť konkrétne parametre a údaje na výpočet dostupnosti riešenia, vieme celkom presne definovať naše požiadavky na dostupnosť dát a služby, teraz už len vybrať tú správnu technológiu...“ Jednu dôležitú skutočnosť som však zatiaľ spomenul len veľmi okrajovo a skryto. V predchádzajúcich odsekoch sa za faktory ovplyvňujúce dostupnosť služby akosi prirodzene automaticky považovali chyby hardvéru alebo softvéru, ako sú napríklad chyba pevného disku, procesora, pamäťového modulu a pod. Netreba však zabúdať, že výpadky systémov môžu byť takisto plánované. Typicky ide o úkony súvisiace s údržbou zariadení a aplikácií (patchovanie, update FW, preventívne výmeny alebo čistenie mechanických častí zariadení a pod.), ale takisto aj rozširovanie kapacity a výkonu a upgrade/update aplikácií. A práve tieto skutočnosti veľa zákazníkov pri výbere technológie opomína. Plánovať dostupnosť riešenia treba z dlhodobého hľadiska – v dnešnej dobe typicky v horizonte 3 až 4 rokov. Ak požadujete od dodávateľa určitú úroveň dostupnosti, tá by mala zahŕňať tak neplánované, ako aj plánované výpadky systémov. Ak vám dodávateľ sľúbi dostupnosť napríklad úložného systému na úrovni 99,999 %, vlastne vám sľubuje, že nový úložný systém nebude pre vás dostupný maximálne 5 minút v roku. No ak mesiac po úspešnej implementácii príde za vami servisný technik dodávateľa a oznámi, že treba vykonať FW update, počas ktorého budú všetky servery, ktoré nie sú pripojené k úložnému systému redundantne, nedostupné po dobu niekoľkých desiatok minút, celkom určite sa už o dostupnosti 99,999 % nedá hovoriť. Vlastne v tom okamihu zisťujete, že ste „päťdeviatkovú“ sumu zaplatili za prinajlepšom „štvordeviatkovú“ systém. To

bol príklad na plánovanú servisnú činnosť, resp. údržbu systému, podobne je to aj s rozširovaním systémov, či už ide o úložné systémy, servery, alebo upgrade a update aplikácie. Zbytočne investujete nemalé finančné prostriedky do robustných systémov, ktoré majú všetky možné aj nemožné komponenty zdvojené, ak čo i len obyčajné rozšírenie pamäte high-end unixového servera alebo rozšírenie diskového poľa a ďalší kabinet či zvýšenie kapacity jeho pamäte cache si vyžiada v tých lepších prípadoch niekoľkokrát úplnú kompletnú odstávku tohto zariadenia, prípadne jeho reboot.

### Menej je niekedy viac

Napriek tomu, že predchádzajúce odseky presvedčivo vravia za implementáciu vysokej dostupnosti do každého relevantného prostredia spíňajúceho predpoklady, nedá sa rozsah implementácie generalizovať, čo sa týka architektúry riešenia a použitých technológií. Nie v každom prípade je napríklad metroklaster riešením, ktoré zákazník potrebuje, napriek tomu, že jeho požiadavky na zabezpečenie dát jasne vravia o nutnosti replikácie a failover na záložnú lokalitu. Ak záložná lokalita nie je vybavená potrebnou infraštruktúrou na to, aby v prípade výpadku primárnej lokality poňala okrem rozbehnutia služby na serveroch a diskových poliach aj desiatky či stovky presunutých pracovníkov, ktorí aplikáciu využívajú, v tom momente je jasné, že takémuto zákazníkovi plne postačuje replikovať dáta na ich ochránenie a zabezpečenie a netreba investovať do drahej implementácie automatizovaných klastrových technológií. Takisto nesmieme zabúdať, že ani to najmodernejšie a najlepšie vysoko dostupné riešenie nie je kompletne bez nastavenia firemných procesov a bez SLA (Service Level Agreement), garantujúcej úroveň služby a reakcie servisného tímu dodávateľa v prípade výpadkov jednotlivých komponentov riešenia. Tieto všetky faktory priamo vplyvajú na TCO. Čím vyššia je úroveň dostupnosti a garancií SLA, tým vyššie sú investičné, ale aj prevádzkové náklady takejto infraštruktúry. Dôsledne teda treba zvažovať, či na beh aplikácie skutočne potrebujeme „mercedes“, a ak už ten „mercedes“ máme, či náhodou nejazdí so zatiahnutou ručnou brzdou. Ale práve teraz, v čase hospodárskej krízy, je dôkladná analýza potrieb azda už rutinou na IT oddelení každého obstarávateľa...



■ ANDREJ GURSKÝ  
Storage Solutions Specialist  
PosAm, spol. s r.o.

## PREHĽAD DODÁVATEĽOV FAULT TOLERANT SYSTEM

Názov spoločnosti	Dodávateľ		Poskytované		Referencie v SR
	Výrobca	Poskytovateľ	Služby	Produkty	
GAMO, a. s.	Nie	Áno	Áno	Áno	Stredoslovenský ústav srdcovo-cievnych chorôb Banská Bystrica, Recyklačný fond SR
Hewlett-Packard Slovakia, s. r. o.	Áno	Áno	Áno	Áno	Poskytujú na vyžiadanie
Novell Slovensko, s. r. o.	Áno	Áno	Áno	Áno	Poskytujú na vyžiadanie
PosAm	Nie	Nie	Áno	Áno	Neuvádzajú