Vysoká dostupnost systému OpenLIMS

M. Novotný

Laboratorní informační systémy patří mezi kritické aplikace zdravotnického informačního systému. Z tohoto důvodu je velmi žádoucí v rámci návrhu implementace takového systému analyzovat jeho možné chybové stavy, jejich dopad na činnost laboratorního komplementu a přijmout adekvátní opatření. Nejzávažnější dopady na chod informačních systémů nejčastěji má (hned po výpadku el. napájení) selhání počítačové techniky (dle zprávy „Průzkum stavu informační bezpečnosti ČR 2005“, DSM, Ernst&Young, NBÚ). Z pohledu systému OpenLIMS je klíčovým počítačovým prvkem databázový server, který obsahuje bezesporu nejcennější aktivum – naměřená data. Je tedy přirozené postarat se o ochranu tohoto aktiva před hrozbou jeho dočasné nedostupnosti, poškození či úplné ztráty díky selhání serveru. Tímto námětem se zabývá i následující článek.

1. SFT A HA

V oblasti spolehlivosti výpočetních systémů se často operuje se zkratkami SFT a HA.

SFT (system fault tolerant) je označení pro systémy odolné proti selhání techniky. Má-li systém splňovat požadavky SFT, nesmí obsahovat žádný kritický subsystém, jehož závada by měla vliv na normální funkci systému. Takový systém se též označuje je jako NoSPOF (No Single Point Of Failure).

Systém kontroluje všechny komponenty a jejich případné chyby musí být samoopravné. Kritéria SFT se používají pro zajištění správné a nepřerušitelné funkce u mimořádně důležitých přístrojů a systémů (autopilot, systémy pro podporu životních funkcí atd.).

HA (High Availability – vysoká dostupnost) se vztahuje k systémům, ve kterých sice může existovat zdroj kritické chyby, ale musí zde být možnost chybu v určitém časovém úseku identifikovat a opravit.

Odstranění chyby může vyžadovat zásah zvenčí a chyba komponenty může mít přechodný vliv na normální funkci systému. Míra vysoké dostupnosti se vyjadřuje dobou nedostupnosti služeb systému za jeden rok vztaženým k jednomu roku (viz. Tab. 1).

V případě non-stop laboratorních provozů je jistě oprávněný požadavek na nejméně „třídevítkové“ HA (99,9 %).

Tab.1: Vyjádření míry HA

Parametr	Výpadek [hod / rok]	Dostupnost [hod]	Míra HA
SFT	0	8760	100.000 %
HA	1	8759	99.989 %
HA	8	8752	99.909 %
HA	32	8728	99.635 %

Jak již bylo v úvodu zmíněno, omezujeme se zde pouze na databázový server, ačkoli HA systému OpenLIMS mohou snížit poruchy i jiných komponent zejména aplikačního serveru a v případě geograficky rozprostřeného systému i WAN (zpravidla Internet). Jedno důležité upozornění: **zajištění vysoké dostupnosti databázového serveru v sobě implicitně nezahrnuje úplnou ochranu datového aktiva před totální ztrátou**. Důvodem je to, že příčiny totální ztráty dat (disaster) mohou spočívat i jinde než v selhání počítačové techniky (živelná katastrofa, teroristický útok, selhání energetické infrastruktury apod.). Komplexně se tímto tématem zabývá specializovaný obor informační bezpečnosti často označovaný jako „Disaster recovery“.

2. ŘEŠENÍ HA V SYSTÉMU OPENLIMS

Z architektury i použitých technologií systému OpenLIMS vyplývají i technická opatření směřující k zajištění vysoké dostupnosti. Jsou jimi:

a) redundantní technické komponenty serveru
b) „fail-over“ klastr serverů (redundance na úrovni serveru)
c) on-line replikace databáze

Všechna opatření mají své výhody a nevýhody a lze je i vhodně kombinovat.

2.1. REDUNDANTNÍ KOMPONENTY SERVERU

Zkušenost ukazuje, že nejčastější závadou serveru je chyba diskového media a dále pak porucha napájecího zdroje. Databázový server pro OpenLIMS je standardně navržen jako značkový server s těmito redundantními komponentami:

a) interní napájecí zdroj (s výjimkou malých provozů – 1 laboratoř do deseti klientů),
b) RAID – redundantní diskové pole podporujícím algoritmus RAID1 (zrcadlení disků) a RAID1+0 (zrcadlení disků a následné rozprostření bloků dat přes dvě a více zrcadlených dvojic),

c) externí záložní zdroj UPS pro případ neočekávaného výpadku el. napájení.

V případě rozsáhlých provozů (zdravotnická zařízení s více laboratořemi) se k těmto komponentám doplňují ještě zdvojené vnitřní ventilátory a zdvojené síťové adaptéry, které umožní automaticky překlenout i výpadek aktivního prvku LAN a systém monitoringu HW komponent a OS. Samozřejmostí je i lokální zálohovací systém umožňující při správném používání alespoň částečnou obnovu ztracených či poškozených dat nazpět v časových bodech.

Výhody tohoto přístupu jsou zřejmé – poměrně vysoká efektivnost. S relativně nízkými pořizovacími náklady lze dosáhnout velmi uspokojivé míry HA zejména v případě malých a středních provozů (do 50 klientů). Za důležitý považujeme fakt, že daný počítač je svou konstrukcí určen pro nepřetržitý provoz a že jeho technická kvalita, kompatibilita s OS a servis jsou výrobcem garantovány.

Nevýhodou tohoto řešení je neschopnost eliminovat rizika vyplývající z méně pravděpodobných scénářů nedostupnosti datových aktiv (narušení db nebo OS, porucha procesoru či základní desky). Závažnou nevýhodou u velkých laboratorních provozů může být nemožnost provádět plánovanou údržbu serverů, zejména instalaci oprav a nových verzí systémového a aplikačního SW (OS, ovladače, databázový systém, nová verze aplikačního modulu apod.) bez přerušení tohoto provozu.

Charakteristika:
Řešení vhodné pro malé a středně velké laboratoře (do 30 klientů) za předpokladu dobře zvládnutého procesu zálohování a obnovy dat ze záložních medií. Typická doba odstávky provozu po závažné havárii serveru 1 až 2 dny dle okolností (podmínky supervizní smlouvy, přítomnost odborného personálu v místě zákazníka apod.).

2.2. „FAIL-OVER" KLASTR SERVERŮ

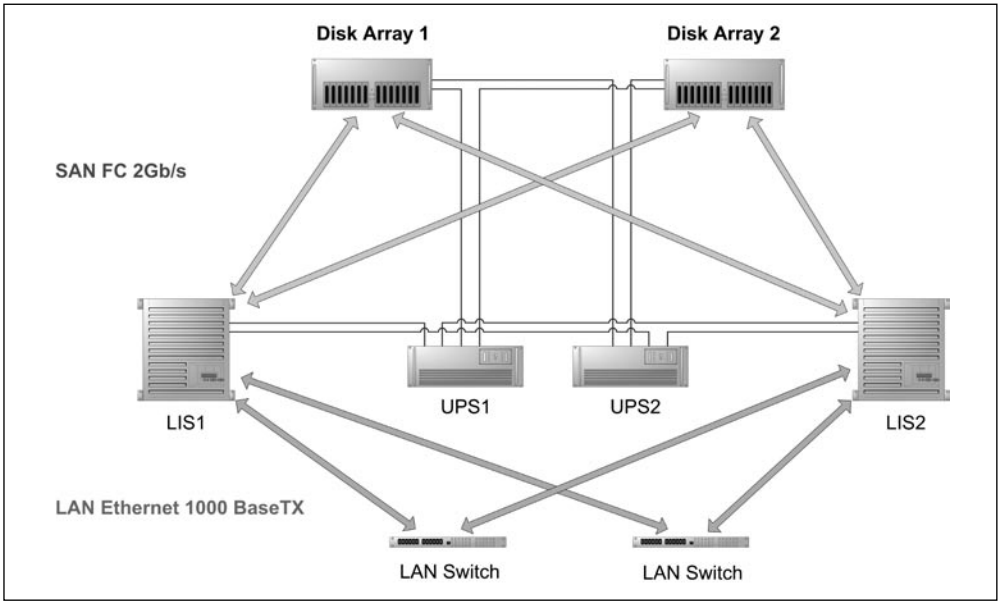
Jde o seskupení dvou nebo více serverů připojených ke sdílenému externímu diskovému poli datovým kanálem na bázi SAN (alternativně SCSI), které se vůči klientovi jeví jako jeden virtuální server. Uzly (servery) v klastru tvoří tzv. kvórum a jsou schopny detekovat HW poruchu nebo selhání OS popř. databázového serveru kteréhokoliv svého člena. Tato technologie překoná výpadek HW a OS konkrétního serveru nebo jeho služby tím, že po detekci závady se automaticky převedou klastrové zdroje zhavarovaného serveru (diskový prostor, IP adresu, db stroj popř. jiné služby) na jiný server v klastru. Síťový klient následně automaticky obnoví (pokud to daný aplikační protokol podporuje) své připojení k těmto zdrojům tak, jak byly využívány před pádem serveru (otevřené soubory, tiskové fronty apod.). Toto řešení má však některá omezení:

1) Nedojde k automatickému obnovení TCP spojení a kontinuálnímu běhu aplikace OpenLIMS, protože toto používaný db stroj MS SQL nepodporuje (neuchovává si seznam a stav otevřených TCP spojení po restartu).
2) K nastartování db stroje na náhradním serveru v klastru nedojde ihned, ale je nutno počítat alespoň s 2 minutovou prodlevou mezi vznikem problému na serveru a jeho překonáním.

Samotný klastr lze navrhnout jako SPOF (Single Point Of Failure) nebo NoSPOF. V případě nasazení OpenLIMS i do velkých provozů o variantě NoSPOF (viz obr. 1) neuvažujeme, protože se jeví jako ekonomicky neadekvátní. Tedy navrhujeme 2 uzlové (serverové) klastry s jedním diskovým polem typu SAN (volitelně Ultra320 SCSI), které však je redundantní co do svých komponent (řadič, interní sběrnice, disky, napájecí zdroje a zálohovaná paměť CACHE). Spoléháme zde na docílení redundance na úrovni serverů.

Výhodou tohoto přístupu je vyšší míra HA zejména s ohledem na plánované odstávky serverů z důvodu údržby. S výhodou je také možné nasadit OpenLIMS do klastrového prostředí tam, kde již toto prostředí existuje zejména v kombinaci s NIS StaproAkord (též OS MS Windows a db systém MS SQL). Tím lze dosáhnout lepšího využití HW prostředků (každý systém může běžet na „svém serveru" a v případě poruchy jednoho serveru se běh systému přesune na „zdravý server"). Zároveň se využije obchodní statut firmy Stapro jakožto Microsoft GP, který opravňuje zákazníka obdržet neomezené množství serverových licencí MS SQL Serveru pro zmíněné aplikace fy Stapro s.r.o. zdarma. Další významnou výhodou zejména pro velmi velké provozy může být klonování databáze za provozu za účelem následného zálohování či archivace. Toto se děje vlastními prostředky diskového pole SAN.

Nevýhodou tohoto řešení jsou vysoké pořizovací a provozní náklady zejména v případech, kdy je klastr vyhrazen pouze pro laboratorní provoz a jeho prostředky tak nejsou dostatečně efektivně využity. Většinou se totiž primární i záložní server dimenzuje stejně, protože hlavní investici představuje zpravidla externí diskové pole. Poněkud paradoxně nevýhodným také může být již zmiňovaný automatismus překonání závady, který může někdy nevhodně maskovat vadu HW serveru nebo chybu v OS či aplikaci. To je však řešitelné nasazením monitoro-



Obr. 1: Schema zapojení NoSPOF klastru.

vacích nástrojů s automatickým zasíláním poplašných zpráv. Toto řešení (stejně jako předchozí) si neumí poradit s náhodnou poruchou databáze např. v důsledku závady diskového pole nebo havárie služby MS SQL. Obecnou nevýhodou je pak poměrně značná technologická složitost (a tím i zranitelnost) celého systému.

Charakteristika:
Vhodné v kombinovaném provozu s NIS či jinými zdravotnickými aplikacemi v rámci velkých zdravotnických zařízení. Typická doba odstávky provozu při havárii serveru – jednotky minut.

2.3. ON-LINE REPLIKACE DATABÁZE

Vychází z možností databázového produktu MS SQL server 2005. Tento produkt disponuje více metodami, jak „rozmnožit" původní databázi na primárním (provozním) databázovém serveru na jeden nebo více serverů prostřednictvím sítě LAN bez nutnosti sdíleného diskového pole. Z množství dostupných metod volíme variantu Transakční replikace, jejímž charakteristickým rysem je asynchronní přenos transakčních poznámek po lokální počítačové síti z originální databáze na databázi záložní. Asynchronnost přenosu umožňuje to, že primární server nečeká, až se daná transakce zreplikuje na záložní databázi, a proto záložní server může být dimenzován pro znatelně nižší datovou propustnost (tedy starší nebo levnější). Záložní server je pasivní z pohledu klientů MS SQL, ale může sloužit pro zcela jiné účely. Ve chvíli, kdy dojde k selhání databázového serveru na primárním serveru z jakýchkoli příčin, je potřeba manuálně nebo poloautomaticky provést přesměrování klientů MS SQL na záložní databázový server. V případě systému OpenLIMS to znamená uvést databázi na záložním serveru do plně provozního stavu a aplikační a komunikační servery přesměrovat pomocí předem připraveného konfiguračního souboru na záložní server.

Výhody tohoto přístupu jsou kombinací předchozích dvou variant – s relativně nízkými pořizovacími náklady lze dosáhnout velmi uspokojivé míry HA i v případě selhání celého serveru. Navíc servery mohou být umístěny v různých lokalitách propojených LAN a ani několikaminutový výpadek konektivity záložního serveru není v tomto případě důvodem k selhání replikace. V rámci nácviku postupu při havárii lze plánovaně a dočasně převést provoz na záložní server a provést údržbu serveru primárního a po té vrátit vše do původního stavu.

Nevýhody spočívají v nutnosti ne zcela triviální konfigurace replikace na úrovni db serveru. Jistým rizikem je možnost ztráty určité části transakční historie v případě „nevhodné havárie" MS SQL serveru na primárním serveru. Jako nevýhodná se většinou bude jevit i potřeba manuálního zásahu pro převedení provozu na záložní server a zpět. Ten lze ale předem popsat a do jisté míry zautomatizovat. Nicméně provedení tohoto kroku se neobejde bez účasti odborného technického personálu. Tuto nevýhodu poněkud zmírňuje

fakt, že pokud dojde k tak závažné havárii, která vyžaduje převedení provozu na záložní server, je vždy nutné odborné posouzení jejích příčin a následná realizace nápravných opatření. Tedy nutnost vyřešit havárii serveru za účasti odborníka může být v praktickém provozu laboratorního systému výhodnější než automatický mechanismus obnovy.

Charakteristika:
Řešení vhodné pro středně velké a velké laboratoře za předpokladu dobře zvládnutého procesu obnovy provozu ze záložního serveru. Typická doba odstávky provozu po závažné havárii serveru desítky minut dle okolností (monitoring serverů, zpracované postupy, přítomnost odborného personálu v místě zákazníka apod.).