

Pro moderní platformy datových skladů je dnes prakticky samozřejmostí, že obsahují funkce hardwarové správy pro vyvažování zátěže I/O, diskových úložišť a CPU, potažmo paměti.

Novější dodavatelé se ve snaze nabídnout přidanou hodnotu oproti etablovaným výrobcům začali soustřeďovat na problematiku optimalizace storage, na to ovšem reagoval i zbytek trhu a dnes se těmto otázkám věnují prakticky všichni. Úložné systémy jsou proto dnes dobře optimalizovány pro nasazení v datových skladech a nabízejí účinné a výkonné metody komprese, jakož i podporu metod pro umístění dat podle frekvence použití.

V rámci snahy o optimalizaci lze pozorovat také rostoucí důraz na využití postupů a metod, které poskytnou lepší odezvu systému. Postupně tak lze pozorovat například větší důraz na využití datových tržišť (data marts), která představují aplikačně specifické zdroje analytických dat obvykle pro užší skupiny uživatelů. Mnozí výrobci je začali opětovně ve větší míře využívat k tomu, aby odsunuli část zátěže z hlavních systémů a zvýšili tak výkon celého řešení.

Zajímavý potenciál mají také tzv. sloupcově orientované databáze, které ukládají obsah do sloupců namísto do záznamů (tj. řádků). Ty nabízejí zajímavé výhody zejména pro datové sklady a agregační úlohy a mohou být využity jako platformy pro data marts, ale v některých případech i jako hlavní platforma pro data warehouse.

Ke klíčovým technologiím současnosti začínají patřit také in-memory databáze a in-memory analýza. Tyto přístupy využívají namísto pevných disků jako primární úložiště paměť počítače nebo serveru, což umožňuje podstatně zrychlit odezvu a rychlost zpracování dat. Tento pokrok umožnil především postupný pokles cen pamětí a rozšíření 64bitových systémů, které jsou schopny adresovat takový objem operační paměti, do něž lze bez problémů nahrát všechna data, jež mají být analyzována, nebo i celou databázi. Do budoucna se tomuto trendu budou podřizovat i datové modely a služby a podle expertů by rozšíření in-memory technologií mohlo znamenat i to, že analytické a transakční systémy začnou v budoucnu využívat stejnou databázi. Dnes tento přístup nachází stále větší ohlas v oblasti business intelligence, kde umožňuje nebývale rychlé zpracování dat na operační úrovni pro okamžitou podporu rozhodování.

Správné řešení

Důležitost technologie, která je základem pro analýzu podnikových dat, dnes stoupá s tím, jak roste zájem o aplikace typu business intelligence (BI), data miningu a dalších nástrojů pro analýzu.

Business intelligence je dnes ve firmách využívána ve stále větším měřítku jak pro podporu strategického rozhodování nejvyššího vedení společnosti, tak pro zefektivnění běžného provozu a zlepšování obchodních výsledků.

Nástroje pro data mining již umožňují provádět komplexní analýzy strukturovaných i nestrukturovaných dat a v posledních letech umožňují provádět také predikční analýzy pro odhad budoucího vývoje nebo analýzu textových, sémantických a obrazových dat. Kromě toho se začaly uplatňovat při analýzách dat ze sociálních sítí.

Koncoví uživatelé ve firmách mají k dispozici jednoduše použitelné nástroje, jež jim dovolují snadno získat informace pro rozhodování, zacílení marketingových kampaní, zlepšení obchodních výsledků, zvýšení spokojenosti zákazníků nebo obecně posílení konkurenceschopnosti.

Pro úspěšnou implementaci datového skladu a nástrojů pro vyhodnocování trendů a podporu inteligentního rozhodování a prodeje je důležité správně definovat cíl a priority, zvolit správné řešení splňující specifické požadavky prostředí a konkrétních procesů a nalézt rovnováhu mezi náklady a přínosy, aby bylo možné dlouhodobě optimalizovat TCO a dosáhnout rozumné návratnosti investic.

Petr Velecký

Kvalitu prověří testování

Jedním z důležitých předpokladů pro úspěšné vybudování datového skladu nadále zůstává důsledné testování konceptu (proof of concept), které potvrdí vhodnost a kvalitu navrhovaného řešení. Podstatou problému je, že v každé firmě pracuje datový sklad s jinými typy dat v rámci specifického prostředí a procesů, takže nelze obecně říci, který systém je pro danou aplikaci tím nevhodnějším. Aby bylo možné realizovat co nejlepší řešení a dosáhnout jeho maximální hodnoty a návratnosti investic, doporučují odborníci provést v rámci fáze výběru databázového systému intenzivní testování pro užší seznam výrobců, kteří se jeví jako potenciálně vhodní. V rámci testování konceptu je přitom vhodné použít pokud možno reálná data extrahovaná ze skutečných zdrojů, tj. z provozních systémů, a spolupracovat s co největším počtem koncových uživatelů, kteří vytvoří pracovní zátěž blížící se produkčnímu nasazení.

Z hlediska výkonu je potřeba si uvědomit, že od současných systémů datových skladů se očekává, že budou provádět současně několik typů úloh, mezi něž patří nahrávání dávek, reporting, Online Analytical Processing (OLAP), data mining, provozní business intelligence nebo kontinuální načítání dat v reálném čase. Proto by systémy měly být při testování a hodnocení posuzovány také s ohledem na výkon při smíšených zátěžích, aby bylo možné účinně plánovat jejich zatížení. V opačném případě budou v průběhu doby s růstem objemu dat a s přidáváním nových úloh neúměrně narůstat administrační náklady a ohrožovat další rozvoj systému.

zaj

Pozvánka na Symposium InterSystems 2011

Desátý ročník mezinárodní konference pro vývojáře aplikací, integrátory systémů, vedoucí pracovníky IT v obchodních společnostech a IT manažery ve zdravotnictví.

„Dosáhněte špičkového výkonu“

6. a 7. června 2011

Konferenční centrum
AV ČR Zámek Liblice

InterSystems
Symposium 2011

www.InterSystems.cz

ROZHOVOR
S PŘEDSTAVITELEM
INTERSYSTEMS PŘINESEME
V ZÁŘÍJOVÉM VYDÁNÍ.

EK005707