
Hodnocení bakalářské práce „Použití SQL Server Integration Services v etapě ETL budování datových skladů“

Autor bakalářské práce: Tomáš Gottwald

Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Vyšší odborná škola informačních služeb v Praze

Hodnocení práce

Tato práce má velmi vhodně zvolené téma, a to zejména k jeho aktuálnosti. Problematika „použití Microsoft SQL Server Integration Services (SSIS)“ se stává postupně více a více diskutovanou problematikou. Důvodem je to, že SSIS je nový ETL nástroj (ostrá verze byla vydána cca před rokem) s výhodným poměrem cena/výkon a s vysokými ambicemi na trhu ETL nástrojů. Tato vhodná volba tématu tedy předurčuje tuto práci k tomu, aby se stala vyhledávaným dokumentem minimálně čtenářů majících zájem o získání prvotních a/nebo základních praktických informací o SSIS. Vhodná volba tématu zároveň dává autorovi potenciál nabýt dovedností s nástrojem, který je z pohledu praxe považován za perspektivní a je na počátku svého životního cyklu. Nabyté znalosti je tedy možné dobře záručit.

Přestože je práce dle jejího názvu zaměřena na SSIS, je přímo SSIS věnována pouze/zejména kapitola 5 *Integration Services*. Zbytek práce je věnován obecnější problematice Business Intelligence, ETL a DTS. Očekávaný účel a plánovaný rozsah práce zajisté neumožňují hluboké zkoumání a detailní popis funkcionality a užití SSIS, právě proto by však měla být tato práce více soustředěna na definovaný záměr, kterým je SSIS, a více využít rozsahu práce k popisu této problematiky.

Práce se hodně věnuje popisu a vysvětlení skutečností (funkcionality, úloh, objektů, ...) (např. kapitola 5.2 *Business Intelligence Development Studio*). Na vlastní použití těchto skutečností pak už mnoho prostoru nezbylo. Za vhodnější bych považoval zaměřit se např. na použití typicky používaných úloh (SSIS tasks) v rámci ETL datových skladů a ostatním úlohám se nevěnovat.

Bez ohledu na omezený rozsah se v práci vyskytla celá řada skutečně podstatných charakteristik, které jsou pro potřeby ETL datových skladů důležité a v DTS se přitom nevyskytovaly. Jedná se např. o využití vyrovnávací paměti pro potřeby implementace stage dat, nové možnosti Business Intelligence Development Studia (BIDS), oddělení control flow a data flow, možnost zpracování výjimek, možnost práce s checkpoints a další. Svědčí to o tom, že si je autor vědom skutečností, pro něž jsou SSIS považovány za významný přínos.

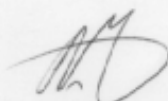
Za „vrchol“ práce považuji kapitolu 5.4 *SSIS v praxi*. Tato kapitola se věnuje konkrétnímu použití SSIS na praktické úloze. Jak již vyplývá z předchozích odstavců, uvítal bych větší zaměření práce právě na tuto část. Pozitivně však hodnotím úvahy nad možnostmi využití některých novinek SSIS v ETL pro datové sklady. Jedná se např. o využití vyrovnávací paměti pro potřeby stage dat nebo o využití úloh „Rebuild Index“, „Sort“, „Union All“, „Fuzzy Grouping“, „Slowly Changing Dimension“.

Celkově práci považuji za dostatečně kvalitní dílo z něhož jsou zřetelné podstatné přínosy SSIS pro potřeby ETL datových skladů oproti DTS a současně je navržen způsob jejich použití. Z pohledu očekávání tedy práce svůj účel naplnila. Po grafické stránce je práce zpracována úhledně, přehledně a doplněna řadou názorných obrázků. Práce současně splňuje náležitosti očekávané od odborného díla.

Náměty k diskusi při obhajobě práce

- Co chybí, co se dá vylepšovat na SSIS?

V Praze, dne 16.1.2007


Ing. Ondřej Vít