

Panorama storage 9.2: BASE, SPD Server 4.5, SPDE, SAS/Accessy i In-Database

Sławomir Bokiniec, sas

Agenda

Co nowego w storage @ SAS 9.2

- BASE
- SAS/Accessy
- In-Database
- SPD Server 4.5
- SPDE

BASE

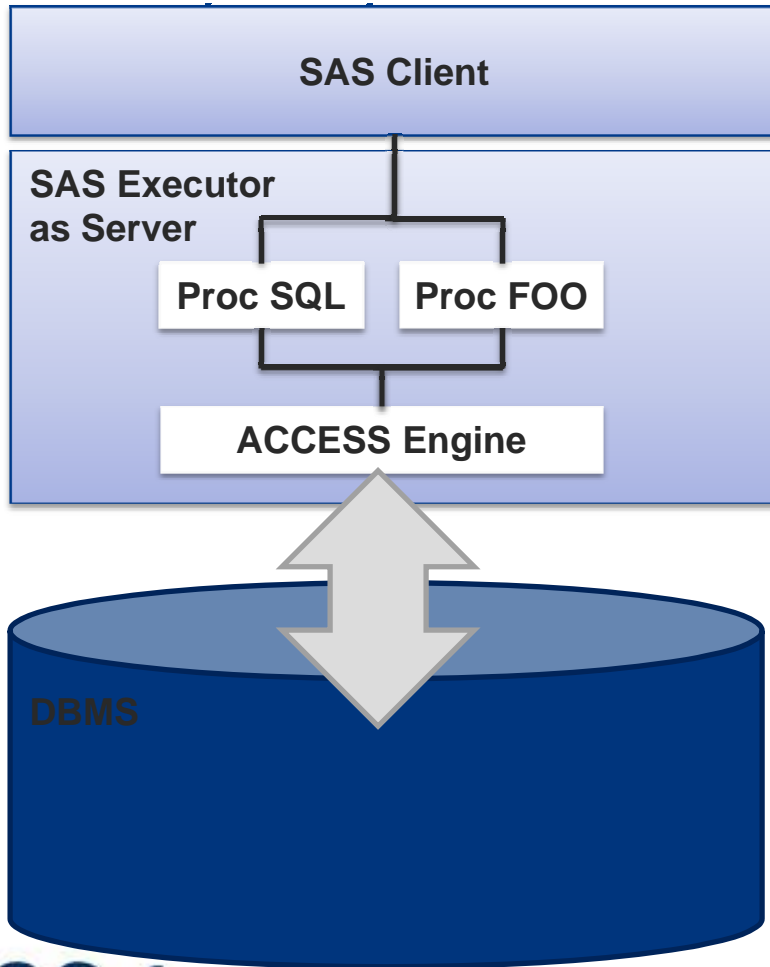
Przypomnienie wybranych nowości:

- ponad 25 nowych formatów i ponad 15 nowych informatów
- ponad 45 nowych funkcji i call routines
- ponad 40 nowych opcji systemowych
- sortownie słownikowe (linguistic) w procedurze sort
- procedura transpose z wieloma zmiennymi w instrukcji id
- Hash Object
 - Wsparcie dla nieunikalnych kluczy
 - statystyka frequency w funkcji Find()
- proc FCMP
 - tworzenie funkcji na podstawie kodu 4gl

Przypomnienie wybranych nowości w SAS/ACCESS 9.2 for Relational Databases

- Przez implicit pass-through można przekazywać nowe funkcje sasowe. Dodatkowo są dostarczane opcje pozwalające definiować mapowanie funkcji sasowych na DBMS'owe
- Opcja biblioteki DBMSTEMP= pozwala na tworzenie tymczasowych tablic w DBMS'ie wykorzystując składnię specyficzną dla DBMS
- Opcje:
 - biblioteki AUTHDOMAIN=
 - systemowa DBIDIRECTEXEC=, w tym DELETE
- Wsparcie dla SAS/Access'ów w aplikacjach sasowych: DI Studio, Enterprise Guide

Proc SQL + ACCESS Engine



- Proc SQL Passthru
 - Explicit (EP)
 - Implicit (IP)

Celem jest przesłanie jak najwięcej przez implicit passthru do serwera.

- Proc FOO & Engine
 - SAS Data Set API mapowane do SQL przez Engine
 - Ruzultaty to często “fat pipe”

Advanced Options [X]

Pre-Assign | Input | Output | Input/Output | Connection | Locking | **Optimization** | Other Options

Block insert buffer size:

Block read buffer size:

Block update buffer size:

Pass DELETE to the DBMS:

Whether to use indexes:

Whether to check for null keys when generating where clauses:

Multi data source optimization:

Do not use extra WHERE clause for updates:

Whether to create a spool file for two-pass processing:

Threaded DBMS access:

Pass functions to the DBMS that match those supported by SAS:

Pass generated SELECT SQL to the DBMS

DBMS processing:

Exceptions to DBMS processing:

Available items:
 NOFUNCTIONS
 NOMULTIOUTJOINS
 NOGENSQL
 NOWHERE

Selected items:

OK Cancel Help

In-Database

Idea

- Jeszcze bardziej przybliżyć przetwarzanie generowane przez SAS-a do danych w DBMS eliminując przesyłanie danych

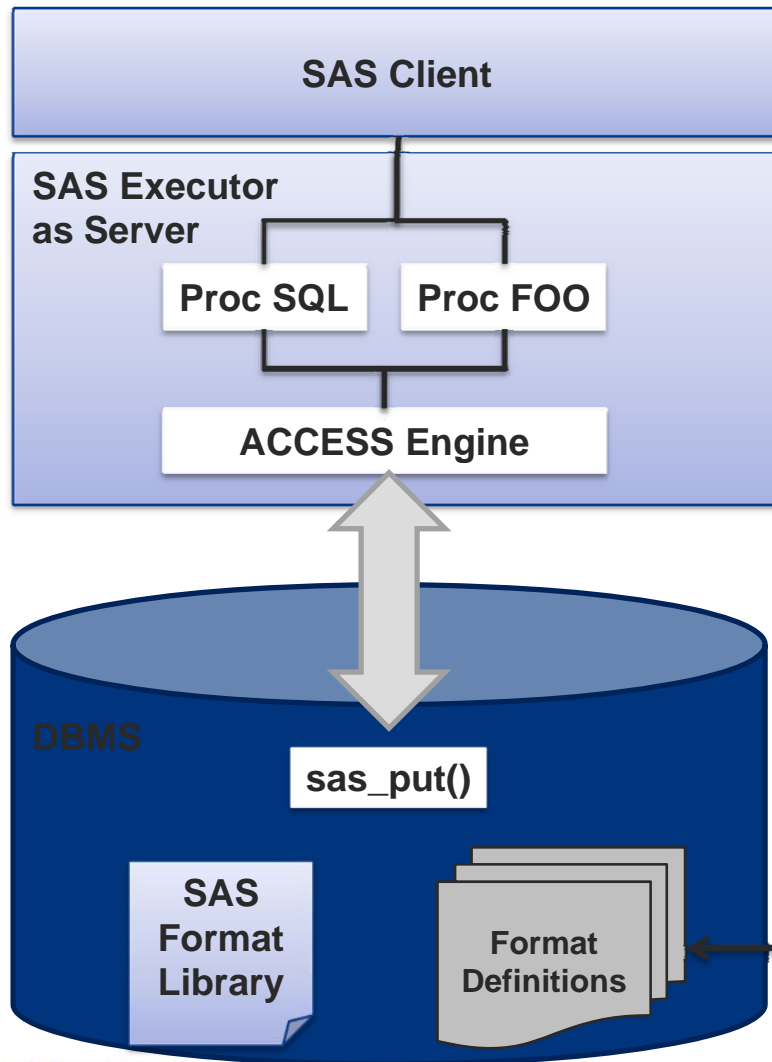
Wybrane inicjatywy

- Enhanced SAS Base Procedures: FREQ, RANK, MEANS, SUMMARY (w M2 do Teradaty)
- Accelerated Procedures for SAS/STAT: REG, VARCLUS, PRINCOMP, SCORE (w M2 do Teradaty)
- Format Publishing (już dostępny z Teradatą)

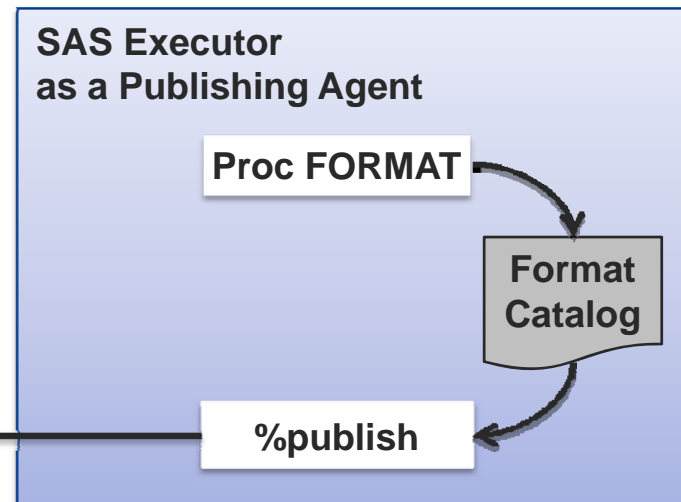
In-Database wybrane inicjatywy c.d.

- Optimized Score Code (z SAS Enterprise Minera, działa już z Teradată)
- W kolejnych maintenance m.in.
 - REPORT, TABULATE, SORT (nodupkey, dupout=)
 - Accelerated Analytical procedures: CORR, CANCELL, FACTOR, TIMESERIES (ETS)
 - SAS stored process invoked as TD stored procedure

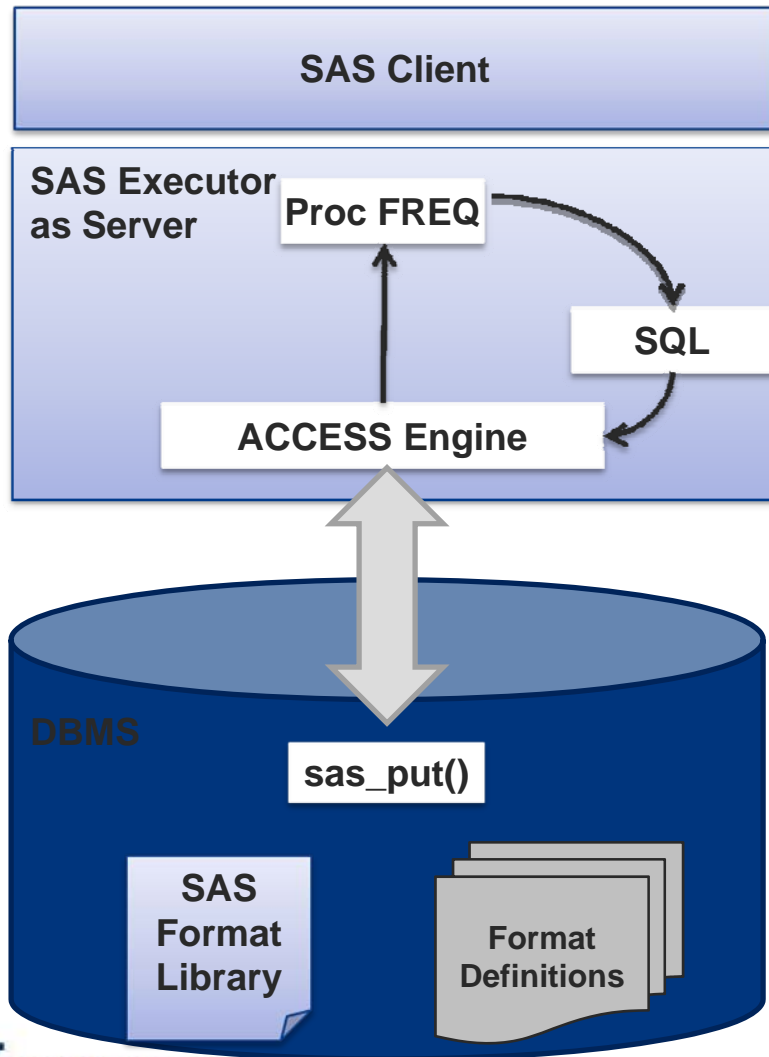
Publishing Formats



- %publish zapisuje formaty z Proc FORMAT do DBMS w jej wewnętrznym formacie
- SAS_PUT() wykorzystuje formaty w SQL-u DBMS
- Mniej przesłanych danych!



Proc Enablement with SQL Pushdown



- Aplikacja SAS generuje PROC FREQ
- Proc FREQ generuje SQL do wstępnej agregacji
Jeżeli są wykorzystywane formaty, jest wykorzystywana funkcja sas_put.
- SQL jest wykonywany po stronie bazy danych
- Proc FREQ czyta wyniki SQL i uzupełnia analizę

Procedury, które wyliczają statystyki opisowe wykorzystują ten schemat.

SPD Server 4.5

- Wersje 4.4.x – dużo nowych funkcjonalności
 - Dynamic clusters, indeksy MinMax, Plug-in do SMC, transformacje w DI Studio, Dynamic Locking, przechowywanie haseł w LDAP, Parallel join z Parallel Group By, Star Join, Index Scan, where costing, user defined formats, ...
- Integracja z najnowszymi rozwiązaniami SAS 9.2 m.in.
 - ESD, proces instalacji
- Implicit pass-through potrafi przekazać funkcję do SPDS-a
- W M2 SPDS-a (M3 do SAS-a)
 - table constraints – nowy sposób na zapewnienie row level security
 - user class low/high/med – osobne spdsserv.parm
 - dalszy rozwój możliwości plug-ina SMC: SPD Server Management

Wybrane nowości w SPDE

- Nowe lub wzbogacone opcje zbioru
 - COMPRESS: CHAR/BINARY
 - ENCRYPT
 - IDXBY
- Nowe opcje biblioteki
 - ACCESS=READONLY
- Wzbogacone opcje systemowe
 - MINPARTSIZE

BASE vs SPDE vs SPD Server

- BASE: podstawowy engine dla dowolnego przetwarzania sasowego, np. ETL-i.
- SPDE: dla pojedynczych użytkowników, dostęp do dowolnie dużych zbiorów dający m.in. wydajniejsze I/O (w przypadku posiadania odpowiednio wydajnej pamięci masowej), bardziej efektywne indeksy, zrównoleglone przetwarzanie warunków where.
- SPD Server (dodatkowo w stosunku do SPDE i BASE): dostęp klient-serwer, zwiększony poziom bezpieczeństwa, efektywniejsze ładowanie, przechowywanie i administracja danymi (w strukturze dynamic cluster), zrównoleglone m.in. agregacje i joiny z procedury SQL, optymalizacja star join.

Prezentacja nie uwzględnia

- SAS OLAP – w osobnej prezentacji
- SAS Grid
- SAS Table Server

DZIĘKUJĘ ZA UWAGĘ

SAS® forum
POLSKA 2009 **warszawa**