

## **SPD Server (4.4.7) vs SPDE (SAS 9.1.3)**

Podstawowa odpowiedź na pytanie:

Czym różni się SPD Server od SPDE?

jest dostępna na stronie: <http://support.sas.com/rnd/scalability/spde/when.html>

Ten dokument ma za zadanie rozszerzyć tę odpowiedź o listę szczegółowych funkcjonalności różniących oba sposoby przechowywania i przetwarzania danych.

Lista różnic między wersjami odpowiednio: 4.5 vs 9.2 będzie bardzo podobna.

### **Zasadnicze różnice:**

1. SPD Server (SPDS) jest SQL-owym serwerem danych działającym w środowisku klient-serwer. Jego rozwój koncentruje się na dalszych optymalizacjach SQL-a pod kątem typowych przetwarzań jakie pojawiają się w środowisku BI.  
SPD Engine (SPDE) jest lokalnym dla sesji SAS-a enginem (komunikującym się przez system operacyjny z zasobami dyskowymi), który dziedziczy po SPD Serverze wybrane funkcjonalności tj.: sposób przechowywania danych, wielowątkowe przetwarzanie warunków where, indeksy i równoległe sortowanie „w locie”.
2. SPDS potrafi zrównoleglić więcej przetwarzań niż SPDE, jest przeznaczony dla większej grupy użytkowników, wymaga więc bogatszego sprzętowo środowiska.
3. Nowe funkcjonalności i optymalizacje w przetwarzaniu pojawiają się najpierw w SPDS-ie i (po pewnym czasie) mogą być dziedziczone przez SPDE.
4. SPD Server wymaga osobnej licencji. SPDE jest dostępne w ramach modułu BASE.

### **Czego nie ma w SPDE (co jest dostępne w SPDS-ie)?**

Wybór funkcjonalności:

5. Optymalizacji SQL-a (m.in.):
  - a. parallel group by,
  - b. parallel join, join index,
  - c. star join,
  - d. index scans (wykonania warunku na podstawie informacji w indeksie),
  - e. BI query rewrites.
6. Struktur danych dynamic cluster, które m.in.: uzależniają wydajność procesów ETL od inkrementu doładowywanych danych, a mniej od wielkości hurtowni/data martu, pozwalają na dodatkowe optymalizacje (np. wykorzystujące index MinMax), ułatwiają administrację dowolnie dużymi zbiorami.
7. Możliwości (efektywnego) ładowania danych za pomocą proc append z podmianą rekordów po kluczu unikalnym (opcja uniquesave=rep).
8. Mechanizmów administracyjnych:
  - a. Możliwości audytu dostępu do zasobów: bibliotek, zbiorów łącznie z rejestracją wykonywanych na zbiorach warunków where i SQL-i.
  - b. Możliwości monitorowania i administrowania: sesjami użytkowników, otwartymi zasobami i procesami.
  - c. Narzędzia do reorganizacji indeksów – bez konieczności ich odbudowywania.
  - d. Backup/Restore (w tym inkrementalny)
9. Mechanizmów bezpieczeństwa (m.in.) osobnej bazy użytkowników i ACL-i. W SPDS-ie można np. przypisać właścicieli dla domeny/schematu.  
Przy SPDE bezpieczeństwo jest oparte mechanizmach systemu operacyjnego lub serwera metadanych.

10. Wsparcia dla obsługi równoczesnego zapisu i odczytu oraz lokowania na poziomie rekordów. SPDE wspiera lokowanie tylko na poziomie zbioru (member-level locking). Tak więc kilku użytkowników/kilka procesów może jednocześnie otwierać ten sam zbiór tylko w trybie INPUT. Podczas pracy w trybie UPDATE dostęp ma tylko jeden użytkownik/proces. W ten sposób SPDS daje możliwość jednoczesnego UPDATE zbioru przez kilku użytkowników/procesów (record-level locking). SPDS oferuje też możliwość kolejkowania dostępu do zbioru (dynamic locking).
11. Współpracy z mechanizmem RLS (Remote Library Service).
12. SPDE nie jest wspierany przez SAS/SHARE (SPDS ma wiele analogicznych do SAS/SHARE funkcjonalności).
13. SPDE nie wspiera (m.in.):
  - a. View,
  - b. Catalogs,
  - c. Mddb,
  - d. Możliwość zmieniania hasła za pomocą proc datasets.
14. Dostępu z Javy i ODBC.
15. Inne:
  - a. Większe wsparcie w SPDS-ie dla przetwarzań formatów zdefiniowanych przez użytkownika,
  - b. Meterialized views,
  - c. Mechanizm pass-through (wraz z możliwością asynchronicznego uruchamiania przetwarzań), implicite pass-through,
  - d. Dodatkowych składni dla SQL-a pass-through: load table, copy table.

#### **W czym SPDE może okazać się lepszy?**

16. Ładowanie danych, np. przy pomocy proc append odbywa się bezpośrednio do filesystemów z pominięciem komunikacji klient-server, co może spowodować, że ładowanie pojedynczego zbioru jest szybsze. Jednak ładowanie do większych zbiorów można drastycznie przyspieszyć korzystając z technologii dynamic cluster w SPDS-ie.

#### **Współpraca SPD Servera z SPDE.**

Po spełnieniu pewnych warunków SPD Server może czytać dane zapisane w SPDE i vice versa.