

Zwiększanie wydajności przetwarzania danych. Hash Tables

Przyspieszenie wykonywania operacji oraz zwiększenie wydajności procesów przetwarzania to niezwykle pożądane i poszukiwane efekty pracy programistów. Często operacje takie jak łączenie tabel, wyszukiwanie podzbiorów, agregacja wymagają optymalizacji wykorzystywanych zasobów, w tym także pamięci. W SAS 9 pojawia się alternatywa dla standardowych technik wykonujących powyższe operacje - obiekt typu hash table. Technologia hash table pozwala na optymalizację szybkości przetwarzania danych w sytuacjach łączenia, przeszukiwania tabel oraz agregacji danych. Technika ta kładzie nacisk na optymalne wykorzystanie pamięci poprzez tworzenie dedykowanej struktury kluczy przechowywanej w pamięci, co pozwala na szybki odczyt wartości kluczy. Uczestnicy szkolenia wykonują praktyczne ćwiczenia łączenia zbiorów, wyciągania podzbiorów oraz wyliczania prostych statystyk na zbiorach przy pomocy hash table. Wydajność rozwiązań wykorzystujących hash table, na tle dotychczasowych metod w 4GL (formaty, indeksowanie, procedury agregujące) jest tematem przeprowadzanych przez uczestników testów. W szkoleniu prezentowana jest także praktyczna implementacja technologii hash table w narzędziu SAS Data Integration Studio. Ponadto uczestnicy dowiedzą się jak skonstruowany jest obiekt hash table i co sprawia, że przy jego zastosowaniu można tworzyć bardzo wydajnie działające kody. Ćwiczenia zawarte w szkoleniu oparte są o ciekawe przykłady zastosowania obiektu hash table (np. w rachunku materiałowym).

Poruszane tematy:

- Koncepcja i filozofia hash table,
- Składnia i proste przykłady użycia,
- Agregacja danych z wykorzystaniem hash table,
- Porównanie wydajności różnych metod łączenia danych,
- Prezentacja hash table w transformacji Lookup w SAS Data Integration Studio,
- Ciekawe zastosowania hash table.

Wymagania wstępne:

- Przetwarzanie danych w Systemie SAS (cz. II plus)

Struktura szkolenia:

- wykład 20%,
- ćwiczenia 80%

Liczba dni: 1

Wykorzystywane oprogramowanie:

- SAS Base (v9)
- SAS Data Integration Studio