

## Zduplikowane rekordy baz danych

Zdarzają się w praktyce serwisowej przypadki, kiedy w katalogu bazy danych pojawiają się tajemnicze pliki posiadające rozszerzenie \*.bld. Przeglądając takie pliki dowolnym edytorem, możemy zobaczyć pogrupowane w pary, rzadziej w trójki, czwórki lub inne wartości, ciągi nic z pozoru nie znaczących liczb. Liczby te poprzedzone są z reguły angielskim komentarzem w rodzaju: The following record numbers are duplicates for: sciezka\_do\_uszkodzonego\_pliku

Opisany problem polega na powieleniu (duplikowaniu) unikalnych rekordów w plikach posiadających unikalne indeksy. Unikalność indeksów polega na tym, że w plikach takich nie może znajdować się jednocześnie więcej niż jeden rekord posiadający w określonych strukturach indeksu polach identyczne wartości. W naszym przykładzie powieleniu uległy dwa rekordy pliku kontrah.dat. Plik kontrah.dat posiada 2 unikalne indeksy, które budowane są w oparciu o kilka pól, jednak najważniejszym z nich jest pole **K:KOD**, zawierające kod kontrahenta. Jeżeli w bazie znajdują się kontrahenci o identycznych kodach, to w takim katalogu z danymi po wykonaniu indeksacji pojawi się plik o nazwie kontrah.bld, zawierający pary zduplikowanych rekordów. Przyczyną takich sytuacji jest praca na uszkodzonych plikach indeksowych, albo uszkodzenie bazy danych polegające na fizycznym poprzesuwniu zawartości rekordów w poziomie. Nie ma innej możliwości, ponieważ próba dopisania identycznego jak istniejący w bazie danych rekord z poziomu programu operacyjnego bądź z poziomu **scannera** zawsze zakończy się niepowodzeniem, jeżeli baza jest dobrze poindeksowana. Indeksacja z poziomu programu lub innym dostępnym narzędziem (**csrt**, **cfil**) po fakcie odkrycia uszkodzenia nic już niestety nie pomoże, ponieważ programy te nie potrafią same naprawić tego typu uszkodzeń.

Jak sobie radzić w takich sytuacjach. Pierwszą operacją, którą wykonujemy po odkryciu w katalogu z danymi pliku \*.bld, jest jego usunięcie - możliwe jest bowiem, że pozostał on po jakiejś wcześniejszej, usuniętej awarii bazy danych. Następnie indeksujemy bazę danych w całości lub co najmniej plik, który może zawierać zduplikowane rekordy. Jeżeli w trakcie indeksacji lub po jej zakończeniu w katalogu z danymi nie pojawią się pliki typu \*.bld, to oznacza, że wszystko jest w porządku. Jeżeli jednak tak nie jest (pliki pojawiły się), to powinniśmy przygotować **scanner (cscn)** i całą swoją wiedzę na temat struktury uszkodzonego pliku - pomocne tu mogą być pliki \*.cla do danego programu lub podręcznik do Xlinka zawierający opis struktury baz danych programów Comarch Klasyka oraz wcześniejsze doświadczenia w rozwiązywaniu tego typu problemów. Kolejną wykonywaną przez nas operacją powinno być przeniesienie zawartości pliku \*.bld na papier (wydruk lub przepisanie). W przykładowym, poprawianym przez nas pliku znajdują się następujące informacje:

The following record numbers are duplicates for:

C:\FPP\DATA\kontrah.k01

RECORD NUMBER: 120

RECORD NUMBER: 231

The following record numbers are duplicates for:

C:\FPP\DATA\kontrah.k02

RECORD NUMBER: 120

RECORD NUMBER: 231

Informacje te oznaczają, że dla indeksów kontrah.k01 i kontrah.k02 zostały zduplikowane dwa rekordy. W posiadanych materiałach szukamy teraz informacji na temat zestawu pól, które wchodzi w skład uszkodzonych indeksów. Kontrah.k01 zawiera tylko jedno pole: **K:KOD**, natomiast kontrah.k02 zawiera pola: **K:GRUPA** oraz **K:KOD**. Już z tej informacji możemy przypuszczać, poprzez które pole doszło do zaniku unikalności rekordów, ponieważ tylko pole **K:KOD** występuje w obydwu indeksach. Następnie pracując pod **scannerem** odszukujemy uszkodzone rekordy korzystając z przygotowanej wcześniej "ściągi" z listą zduplikowanych rekordów: kombinacją klawiszy **Ctrl+J** skaczemy do wybranych rekordów i wyszukujemy uszkodzenie. Sposób usunięcia awarii zależy od ustalenia przez nas przyczyny zaistniałej sytuacji. Jeżeli powielony jest rekord w całości, to naprawa uszkodzenia polega na skasowaniu jednego z nich, jeżeli natomiast rekordy zawierają różne informacje i nie można jednego z nich skasować, to musimy poprawić pole powodujące duplikację w jednym z nich w ten sposób aby zachować unikalność rekordów. Należy przy tym pamiętać, o istniejących w bazie danych relacjach pomiędzy plikami - jeżeli poprawiane pole wchodzi w relację z jakimiś innymi plikami, to poprawkami należy objąć również tamte pliki. W naszym przykładzie będą to zapisy w plikach tranhead.dat, tranelem.dat oraz upusty.dat. Po usunięciu wszystkich duplikatów, należy ponownie skasować pliki \*.bld i przeindeksować bazę. Jeżeli po indeksacji uszkodzenia nie pojawią się, to oznacza, że wszystkie działania zostały wykonane prawidłowo.

**Uwaga!** Jeżeli w pliku \*.bld lista uszkodzonych rekordów jest pogrupowana w trójki, czwórki lub inne wartości, to oznacza, że w tych właśnie trzech, czterech itd. rekordach unikalne pole zostało powielone.

Pliki \*.bld obejmujące zawartość całego lub większej części pliku \*.dat, mogą powstać również wtedy, gdy nastąpi uszkodzenie struktury logicznej pliku, tzn. gdy rekordy zostaną w nim poprzysuwane lub nadpisane z powodu nieprawidłowej pracy systemu sieciowego lub awarii sprzętowej - w takich sytuacjach bardzo często dochodzi do duplikowania rekordów, ponieważ struktury bazy nie kontrolują ani programy operacyjne, ani programy narzędziowe.

Wojciech Nowiński  
Dział Techniczny Oprogramowania