

Stanisław CIEŚLA

Politechnika Śląska, Instytut Informatyki

## **DOSTĘP DO BAZ DANYCH W SIECI INTERNET — MECHANIZM INTERNET DATABASE CONNECTOR**

**Streszczenie.** W artykule przedstawiono charakterystykę i sposoby użycia przykładowego mechanizmu dostępu do baz danych w sieci Internet z wykorzystaniem techniki WWW. Omówiono aspekty bezpieczeństwa tego mechanizmu i przedstawiono przykładowe rozwiązanie.

## **ACCESS TO DATABASES WITH INTERNET NETWORK — INTERNET DATABASE CONNECTION MECHANISM**

**Summary.** This paper describe characteristic and usage of database access mechanism through Internet by means of WWW technique. Security aspect of access is also reported and example solution is shown.

### **1. Idea zdalnego dostępu do baz danych**

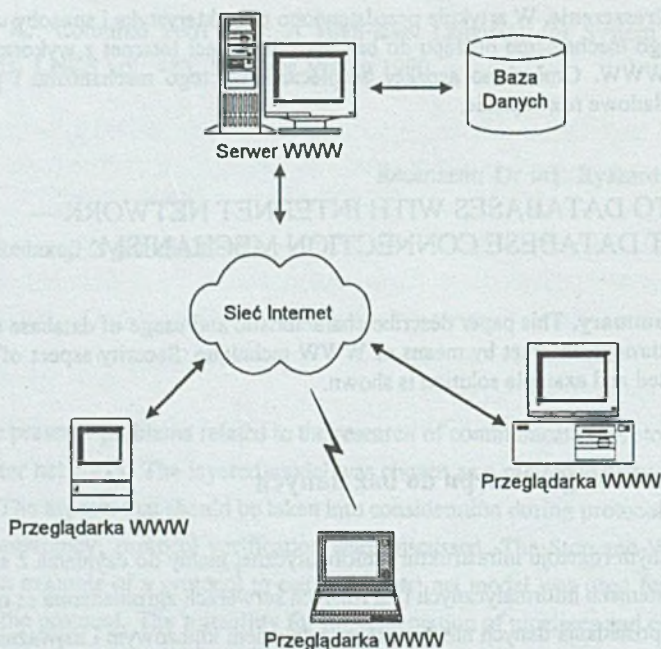
Przy obecnym rozwoju infrastruktury informatycznej mamy do czynienia z sytuacją, gdzie w różnych systemach informatycznych i na różnych serwerach zgromadzone są ogromne ilości danych. Fakt posiadania danych nie jest już zagadnieniem kluczowym i najważniejszym. Dużo większe zapotrzebowanie jest aktualnie na systemy udostępniające dane jak najszerszemu gronu odbiorców. Opracowywane są aktualnie różnorakie systemy dostępu do serwerów baz danych w sieciach komputerowych. Tworzenie takich systemów wymaga dużych nakładów finansowych i zabiera czas. Wymaga również stworzenia odpowiednich narzędzi i zbudowania specjalnej infrastruktury.

Analizując problem zdalnego dostępu do baz danych, należy także uwzględnić możliwość wykorzystania istniejących narzędzi, mechanizmów i infrastruktury dostosowując je do speł-

nienia wymaganych specyfikacji. System zdalnego dostępu do baz danych powinien umożliwiać następujące operacje:

- a) udostępnianie informacji zawartych w bazie danych
- b) wstawianie, uaktualnianie i usuwanie informacji z bazy danych
- c) wykonywanie innych poleceń języka manipulowania danymi, na przykład języka SQL.

Możliwą platformą dla takiego systemu zdalnego dostępu do baz danych może być sieć Internet. Zaletami tego rozwiązania jest wyeliminowanie nakładów na infrastrukturę techniczną (połączenia), wykorzystanie istniejących łączy oraz serwerów WWW do publikowania informacji, wykorzystanie dowolnych komputerów wyposażonych w przeglądarki WWW jako stacji roboczych pełniących funkcje terminali, wykorzystanie języka HTML i usługi HTTP w transmisji danych od serwera do klienta.



Rys. 1. Idea dostępu do bazy danych w sieci Internet  
Fig. 1. Idea of database access through Internet network

Jednym z zagadnień, które należy przy takich założeniach przedyskutować, jest skonstruowanie mechanizmów dostępu do serwerów baz danych i opracowanie sposobu interakcji użytkownika z bazą danych poprzez sieć Internet.



Przykładowym mechanizmem, w oparciu o który można taki dostęp do bazy danych w sieci Internet zrealizować, jest mechanizm Internet Database Connector firmy Microsoft.

## 2. Charakterystyka mechanizmu Internet Database Connector

Mechanizm IDC jest komponentem serwera WWW firmy Microsoft o nazwie Internet Information Server pracującego na platformie Windows NT [3]. Jest on dostarczany jako element oprogramowania tego serwera WWW i nie wymaga żadnych dodatkowych narzędzi, aby zapewnić serwerowi możliwość współpracy z bazą danych. Mechanizm IDC umożliwia publikowanie informacji zawartych w bazie danych w łatwy i szybki sposób poprzez dynamiczne tworzenie stron WWW. Nie wymaga od administratora serwera WWW oraz autora stron wiedzy na temat wymyślnych technik programowania. Sposób tworzenia stron umożliwiających operowanie na bazie danych jest prosty, przejrzysty i intuicyjny. Pozwala on także na dużą dowolność w stosowaniu bazy danych, bo mechanizm IDC wykorzystuje jako bezpośredni sprzęg z bazą danych technikę „otwartego dostępu do baz danych” ODBC [2].

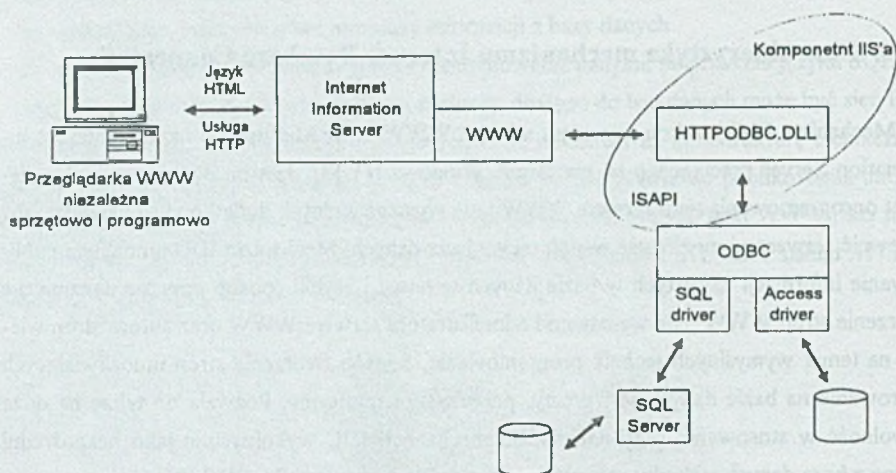
## 3. Zasada funkcjonowania mechanizmu IDC

Mechanizm IDC jest tak skonstruowany, że po stronie klienta może występować dowolna przeglądarka WWW zainstalowana na dowolnej platformie sprzętowej. Przeglądarka na stacji klienta komunikuje się poprzez usługę HTTP z wykorzystaniem języka HTML z serwerem WWW. Serwer WWW w tym przypadku to Internet Information Server (IIS) firmy Microsoft zainstalowany na serwerze Windows NT.

Jedną z usług IIS'a jest usługa WWW, z którą to komunikuje się przeglądarka użytkownika przysyłając zapytania dostępu do bazy danych i otrzymując ich rezultaty w formie stron w języku HTML [3 i 4]. Usługa WWW komunikuje się z aplikacją HTTPODBC.DLL, będącą sercem całego mechanizmu. Aplikacja ta napisana jest z wykorzystaniem funkcji Internet Server API opracowanych do oprogramowywania serwera WWW firmy Microsoft.

Aplikacja HTTPODBC.DLL przetwarza zapytanie dostępu do bazy danych i przekazuje je do mechanizmu ODBC zainstalowanego na serwerze Windows NT, otrzymując w zamian rezultaty tegoż zapytania w formie rekordów. Z rezultatów tych aplikacja HTTPODBC.DLL tworzy dynamicznie strony WWW i przesyła je do usługi WWW serwera IIS. Bezpośrednia interakcja z bazą danych dokonuje się za pomocą mechanizmu ODBC z wykorzystaniem spec-

jalizowanego dla danej bazy danych sprzęgu. Ogólny schemat funkcjonowania mechanizmu IDC ilustruje rysunek 2.



Rys. 2. Komponenty mechanizmu IDC służące do połączenia z bazą danych  
Fig. 2. Components of IDC for connecting to databases

## 4. Interakcja z bazą danych

### 4.1. Wykonywanie zapytań i publikowanie rezultatów

Wykonywanie operacji na bazie danych za pomocą mechanizmu Internet Database Connector odbywa się w oparciu o dwa typy zbiorów umieszczonych na serwerze. Jeden z tych zbiorów o rozszerzeniu *.IDC* to skrypt zawierający informacje niezbędne do wykonania operacji na bazie danych, drugi o rozszerzeniu *.HTX* zawiera format strony prezentującej rezultaty operacji [3]. Skrypt opisujący operację wykonywaną na bazie danych zawiera komplet informacji potrzebnych do nawiązania połączenia z bazą danych za pomocą mechanizmu ODBC.

Zawiera on nazwę źródła danych, nazwę użytkownika, jego hasło, nazwę zbioru *.HTX*, w którym mają zostać umieszczone uzyskane rezultaty operacji, oraz frazę zawierającą zapis operacji na bazie danych w języku SQL.

Przykładowy skrypt *query.IDC* przedstawiono poniżej:

```
Datasource: WWWSql
Username: Adam
Password: alumat
Template: result.htx
SQLStatement:
```



```
+SELECT year, month, FROM pubs.dbo.dates
+WHERE debet_account = 'YES'
```

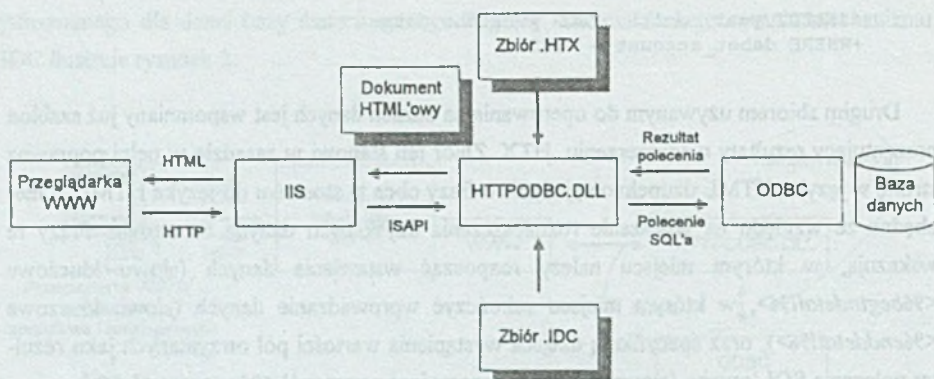
Drugim zbiorem używanym do operowania na bazach danych jest wspomniany już szablon prezentujący rezultaty o rozszerzeniu .HTX. Zbiór ten stanowi w zasadzie w pełni poprawną stronę w języku HTML uzupełnioną jedynie o frazy obce w stosunku do języka HTML, a niezbędne ze względu na wskazanie rozmieszczenia uzyskanych danych na stronie. Frazy te wskazują, w którym miejscu należy rozpocząć wstawianie danych (słowo kluczowe `<%begindetail%>`), w którym miejscu zakończyć wprowadzanie danych (słowo kluczowe `<%enddetail%>`), oraz specyfikują miejsca wystąpienia wartości pól otrzymanych jako rezultat polecenia SQL'owego (słowa kluczowe oznaczające nazwy pól `<%nazwa_pola%>`).

Przykładowy zbiór *result.HTX* przedstawiono poniżej:

```
<HTML>
<BODY>
<HEAD ...>
...
Rezultaty zapytania
<%begindetail%>
Za miesiąc <%month%> w roku <%year%>
<%enddetail%>
...
</BODY>
</HTML>
```

Aplikacja HTTPODBC.DLL wykorzystuje oba te zbiory do operowania na bazie danych, działając w oparciu o bardzo prostą metodę. Zbiór rekordów stanowiących rezultat operacji jest uzyskiwany poprzez skrypt z frazą SQL. Natomiast dynamicznie tworzona strona w języku HTML, która ma zostać przesłana do przeglądarki użytkownika, jest przygotowana w oparciu o wzorec strony umieszczony w zbiorze .HTX, oraz wykorzystując uzyskany zbiór rezultatów zapytania.

Technika generowania strony wygląda następująco: cała strona ze zbioru .HTX jest przepisywana bez zmian, aż do wystąpienia słowa kluczowego `<%begindetail%>`, od tego miejsca wszystkie polecenia języka HTML aż do wystąpienia frazy kończącej `<%enddetail%>` przepisywane są tyłkrotnie, ile uzyskano rekordów w rezultacie zapytania. Innymi słowy, wszystko, co jest ograniczone słowami kluczowymi `<%begindetail%>` i `<%enddetail%>`, interpretowane i przepisywane jest dla każdego rekordu ze zbioru rezultatów. Powiązanie z danymi pochodzącymi z rekordów uzyskuje się przez umieszczenie w wybranym miejscu nazwy pola obudowanej symbolami `<% i %>`; w miejsce tych pól aplikacja generując stronę wstawi wartości tych pól pochodzące z aktualnie przetwarzanego rekordu.



Rys. 3. Schemat realizacji zapytania dostępu do bazy danych  
Fig. 3. Diagram of processing database request

Oczywiście, podany wyżej przykład, na podstawie którego mechanizm IDC został scharakteryzowany, jest bardzo prosty i ogranicza się jedynie do wykonania frazy *SELECT*, przez co uzyskuje się efekt opublikowania danych znajdujących się w bazie danych na stronie WWW; nie jest to jednak jedyna możliwość. Mechanizm IDC pozwala na wykonywanie także innych poleceń języka SQL zgodnych z podzbiorem tego języka określonym dla mechanizmu ODBC, dzięki temu możliwe jest spełnienie pozostałych postulatów zdalnego dostępu do baz danych wymienionych w rozdziale pierwszym. Chodzi o wykonywanie operacji uaktualniania, usuwania i dodawania rekordów do bazy danych.

Istnieje także możliwość wykonania bardziej skomplikowanych operacji na bazie danych, które są trudne bądź niemożliwe do wykonania ze względu na ograniczoność składni języka SQL mechanizmu ODBC. Żeby tę niedogodność ominąć, można użyć funkcji ODBC pozwalającej na przekazanie nietypowego polecenia bezpośrednio do źródła danych, z pominięciem jego interpretacji. Pozwala to na użycie mechanizmów charakterystycznych dla danego serwera bazy danych, a nie implementowanych przez ODBC. Skomplikowane operacje na bazie danych można także uzyskać przez wywołanie wykonania procedury zapisanej w serwerze bazy danych, o ile oprogramowanie serwera na taką możliwość pozwala.

#### 4.2. Wywoływanie zapytań i ich parametryzowanie

W poprzednim rozdziale przedstawiono ogólny sposób konstruowania poleceń dostępu do bazy danych na serwerze, podane przykłady ograniczały się jedynie do wykonania stałych fraz języka SQL. Mechanizm IDC pozwala także na parametryzowanie wykonywanych zapytań przez użytkownika korzystającego z przeglądarki. Jednak przed omówieniem tego zagadnienia należy wspomnieć o wywoływaniu poleceń dostępu do bazy danych na stronie WWW [3 i 4].



Wywołanie skryptu ze strony WWW w przeglądarce uzyskuje się przez wykorzystanie sposobu, w jaki przeglądarka interpretuje wywoływane powiązania (*links*). Otóż, jeśli na stronie WWW użytkownik wywoła któreś z połączeń przez wskazanie frazy `<A HREF = "http://server/.../zbiór.xxx"`, przeglądarka kieruje wywołanie do wskazanego serwera pozostawiając jemu interpretację zbioru *zbiór.xxx*. Dzięki takiemu mechanizmowi na stronie WWW wysyłanej do przeglądarki użytkownika można umieścić dowolny zbiór, który jest interpretowany przez serwer wysyłający daną stronę. Taka otwartość zapewni, że każda przeglądarka będzie w stanie „wykonać” polecenie, które w formie powiązania zostało umieszczone na stronie. „Wykonać” – to znaczy przesłać do serwera z żądaniem zinterpretowania i ewentualnego wykonania.

Dokładnie w ten sposób zdecydowano się na wywoływanie poleceń dostępu do bazy danych w mechanizmie IDC. Najprostsze wywołanie stałego zapytania do bazy danych ze strony WWW ma postać:

```
<A HREF="http://server/scripts/query.idc?>Naciśnij  
by wywołać zapytanie</A>
```

Znak zapytania po nazwie skryptu oznacza, iż zakończyła się nazwa zbioru i nastąpią po niej wartości ewentualnych przekazywanych parametrów skryptu w formacie *parametr = wartość*. Jest to pierwszy sposób przesłania parametrów ze strony WWW do serwera dla wykonywanego skryptu. Celem użycia parametrów w poleceniu SQL'owym należy frazę *SQLStatement* skryptu *.IDC* zmodyfikować przez wstawienie w odpowiednie miejsca nazw parametrów otoczonych symbolami `<% i %>` jak w poniższym przykładzie:

```
SQLStatement:  
+SELECT name FROM pubs.dbo.users WHERE id = <%search_id%>
```

Wprowadzenie takiego parametru na stronie WWW może się odbywać przez umieszczenie na niej obiektu wprowadzającego dane, typu na przykład pole tekstowe i użycie jego zawartości przy wywoływaniu skryptu za pomocą powiązania. Obiekty do wprowadzania danych są elementem języka HTML.

Innym sposobem przekazania parametrów do skryptu jest użycie struktury języka HTML zwanej formularzem. W strukturze tej wyspecyfikowane są obiekty służące do wprowadzenia danych od użytkownika oraz akcja (przez nazwę wywoływanego zbioru), jaką ma wykonać serwer po otrzymaniu z przeglądarki porcji danych wprowadzonych przez użytkownika.

Wykorzystanie tej struktury na potrzeby mechanizmu IDC polega na nazwaniu obiektów służących do wprowadzania danych, tak jak nazwane są parametry w skrypcie, oraz wskazaniu skryptu jako akcji związanej z danym formularzem. Poniżej przedstawiony jest przykładowy zapis na stronie HTML.

```
<FORM METHOD="POST" ACTION="scripts/query.idc">
```

```
...  
Wprowadź identyfikator: <INPUT NAME="search_id">  
...  
<INPUT TYPE="SUBMIT" VALUE="Wyświetl dane"  
</FORM>
```

W powyższym przykładzie po naciśnięciu przez użytkownika klawisza z napisem *Wyświetl dane* wywołany zostanie na serwerze skrypt *query.idc*, w którym parametr *search\_id* otrzyma wartość zgodną z tą, którą wpisał użytkownik w polu tekstowym na stronie WWW w swojej przeglądarce.

## 5. Aspekty bezpieczeństwa dostępu do danych

Ważnym aspektem rozważanego zagadnienia jest zabezpieczenie baz danych przed niepożądanym dostępem. Mechanizm Internet Database Connector pozwala na ochronę dostępu do danych poprzez kilka cech w swojej konstrukcji.

Podstawową zaletą jest przetwarzanie danych po stronie serwera; użytkownik za pomocą przeglądarki jedynie wywołuje żądania dostępu i ewentualnie wprowadza ich parametry, natomiast serwer poprzez informacje w skrypcie *.IDC* łączy się z bazą danych podając odpowiednio użytkownika i hasło. Zarówno nazwa użytkownika, jak i hasło oraz fraza SQL'a są ukryte przed użytkownikiem przez umieszczenie ich na przykład w osobnym katalogu serwera i opatrzenie atrybutem zezwalającym jedynie na wykonanie. Tak więc potencjalny napastnik nie jest w stanie podejrzeć zawartości skryptu, nie pozna więc nazwy użytkownika i hasła dostępu do bazy danych.

Podobnie ma się sprawa z użytkownikami domeny Windows NT, w której zainstalowany jest Internet Information Server. Oni to w odróżnieniu od użytkowników pochodzących z sieci Internet mają możliwość dostępu do serwera i jego zasobów, czyli potencjalnie mogą odczytać zawartość skryptu. Jednak tu także system jest chroniony, gdyż do katalogów zawierających dane dla serwera WWW (strony HTML, skrypty, rysunki) domyślnie prawa dostępu mają tylko administratorzy domenowi i specjalny użytkownik *IUSR\_computername* [3], poprzez którego konto serwer WWW operuje zasobami serwera NT [1]. Użytkownikowi temu serwer WWW zmienia automatycznie i bardzo często hasło, tak więc penetracja zasobów serwera poprzez to konto nie jest możliwa. Jednocześnie wszystkie aplikacje współpracujące z serwerem WWW posiadają automatycznie prawa wykonywania jedynie przez wymienionego wcześniej specjalnego użytkownika „internetowego”. Z kolei można powiedzieć, że istnieje zagrożenie włamania do systemu operacyjnego poprzez operowanie z sieci Internet na zasobach serwera NT jako użytkownik *IUSR\_computername*. Tutaj także system jest chroniony, gdyż



domyślnie użytkownik ten posiada jedynie prawa odczytu i wykonania w zasobach związanych z serwerem, natomiast nie posiada żadnych praw do pozostałych zasobów systemu [1].

Powyższe spojrzenie obrazuje wyraźne odseparowanie użytkowników sieci Internet od zasobów serwera i domeny Windows NT, jak i podobne działania w drugą stronę, to jest ograniczenie dostępu użytkownikom systemu Windows NT do zasobów używanych przez serwer WWW i jego usługi.

## 6. Przykład zastosowania

Omawiany w powyższym opracowaniu mechanizm IDC został użyty przeze mnie do skonstruowania eksperymentalnego serwera WWW służącego jako platforma udostępniania, akwizycji i przetwarzania danych o szkodach górniczych na ciągach drogowych. Pomysł systemu informatycznego realizującego powyższe zadanie został zgłoszony przez Katedrę Komunikacji Lądowej Wydziału Budownictwa Politechniki Śląskiej w Gliwicach.

Jako platformę do skonstruowania nakreślonego powyżej systemu informatycznego zaproponowałem sieć Internet i narzędzia oraz usługi WWW. Mechanizm IDC w konstruowanym serwerze stanowił czynnik zapewniający zróżnicowany i zależny od uprawnień dostęp wszystkim zainteresowanym do bazy danych. Spełnił on wystarczająco swoją rolę umożliwiając zrealizowanie wszystkich założonych funkcji, jakie baza danych winna była spełniać. Poniżej przytaczam funkcje realizowane przez ten eksperymentalny serwer w oparciu o mechanizm Internet Database Connector:

- a) logowanie do systemu i weryfikacja użytkowników,
- b) kontrola dostępu do zasobów bazy danych na różnym poziomie w zależności od uprawnień użytkowników znanych systemowi,
- c) ochrona stron WWW przed niepowołanym dostępem, uniemożliwienie „przeskoczenia” niektórych stron bez weryfikacji użytkownika,
- d) opracowanie formularzy wprowadzających dane i parametryzujących wydawane zapytania, także w oparciu o dane już wprowadzone do bazy (na przykład listy wyboru w zapytaniu tworzone w oparciu o pozycje już istniejące w bazie),
- e) prezentowanie danych pochodzących z bazy danych w formie liczbowej (tabelarycznie) oraz graficznej (wykresy, diagramy – z użyciem mechanizmów ActiveX) oraz w postaci rysunków – z użyciem mechanizmów ActiveX i VBScript [2].

## 7. Podsumowanie

Przedstawiony mechanizm Internet Database Connector dostępny wraz z serwerem WWW firmy Microsoft stanowi prosty i łatwy w użyciu interfejs zapewniający dostęp do baz danych użytkownikom sieci Internet. Poniżej chciałbym przedstawić dostrzeżone przeze mnie wady i zalety tego mechanizmu.

### a) Zalety

- jest to oprogramowanie dostępne bezpłatnie tak jak cały serwer WWW (Internet Information Server) firmy Microsoft,
- oprogramowanie tego mechanizmu, tworzenie skryptów dostępu do bazy danych i ich używanie są stosunkowo proste i łatwe do nauczenia się nawet przez osoby nie posiadające praktyki w pisaniu programów,
- powiązanie z bazą danych przy użyciu mechanizmu ODBC pozwala na dowolny wybór stosowanej bazy danych,
- udostępniane bazy danych są przy odpowiedniej konfiguracji wystarczająco chronione przed niepożądanym dostępem zarówno przy próbie włamania z sieci Internet, jak też przed penetracją z sieci lokalnej,
- mechanizm działa w pełnej zgodności ze standardem języka HTML, tak więc oprogramowanie i sprzęt u klienta korzystającego ze zdalnego dostępu do bazy danych są całkowicie dowolne i niezależne od jakiegokolwiek producenta.

### b) Wady

- mechanizm IDC jest zintegrowany z serwerem WWW firmy Microsoft, co wymusza stosowanie serwera tej właśnie firmy,
- dotkliwą niedogodnością jest konieczność ograniczenia liczby poleceń języka SQL wykonywanych w skrypcie tylko do jednego, co wymusza tworzenie serii skryptów wykonujących skomplikowane operacje lub oprogramowywanie do tego celu serwera bazy danych<sup>1</sup>,
- niezbyt elastyczny sposób przekazywania parametrów z przeglądarki u użytkownika do serwera i dalej do i z aplikacji pisanych z użyciem Internet Server API, w tym do HTTPODBC.DLL'a stanowiącego jądro mechanizmu IDC<sup>1</sup>.

Pomimo licznych wad i ograniczeń sądzę jednak, że ze względu na swoją prostotę i łatwość stosowania mechanizm IDC będzie coraz powszechniej stosowany w prostych serwerach WWW i z całym powodzeniem polecam go do takich zastosowań.

---

<sup>1</sup> Niedogodność usunięta w IDC 2.0 dostępnym wraz z Internet Information Server dystrybuowanym bezpłatnie jako część Windows NT Server 4.0.



Jednocześnie dogłębne poznanie tego mechanizmu i wykorzystanie jego otwartości może pozwolić na stworzenie bardzo ciekawych rozwiązań niekoniecznie służących stricte do publikowania informacji, ale na przykład jako narzędzia do dynamicznego tworzenia stron WWW.

## LITERATURA

- [1] Windows NT system Guide. Microsoft Corporation 1995.
- [2] Microsoft TechNet. Volume 4, Issue 8, August 1996. Redmont 1996.
- [3] Internet Information Server, Installation Guide. Microsoft Corporation 1996.
- [4] Internet Information Server Planning Guide. Microsoft Corporation 1996.

Recenzent: Dr inż. Ryszard Winiarczyk

Wpłynęło do Redakcji 18 listopada 1996 r.

## Abstract

This article, discusses access to database through Internet using mechanism Internet Database Connector and is divided into seven parts.

The first one discusses ideology of access to remote database from Internet's point of view using HTTP protocol and HTML language for this purpose.

The second part describes characteristic of Internet Database Connector method and possibility of using this method of Internet server. Interface between IDC and database by means of ODBC connectivity is also described.

The third section discusses an idea of how IDC works. Web browsers submit request to the Internet server by using HTTP. The Internet server responds with a document formatted in HTML. Access to databases is accomplished through a component of Internet Information Server called the Internet Database Connector. The IDC, main component Httpodbc.dll, is an ISAPI DLL that uses ODBC to gain access to databases.

The fourth part describes the way to contact with databases. The IDC uses two types of files to control how the database is accessed and how the output Web page is constructed. These files are Internet Database Connector (.IDC) files and HTML extension (.HTX) files.

The next part discusses the security aspects of using database access via Internet, and secure data from being local accessed by user.

The last part shows an example of database server using IDC to control database access, and describe possibility of extraordinary using IDC with Web publishing solution (for example to prepare dynamically creating page with graphics).