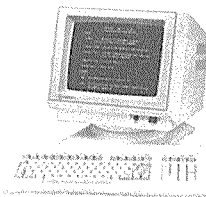


吳 吉 祿  
韓國電子通信研究所  
컴퓨터기술연구단장 / 工博

# 한글 관계형 데이터 베이스 관리체계 개발



## 1. 머리말

행정전산망에 사용될 전산기의 요구사항 중에는 효율적 데이터 관리기능이 포함된다.' 이는 행정전산망에서 다루어야 할 자료의 양이 방대함을 예상할 수 있기 때문이다. 예를 들어, 주민업무에서는 41.4기가(Giga) 바이트 그리고 부동산 업무에서 26.6 기가 바이트가 소요 예상되는 순수 데이터 베이스 크기이다. 여기서 순수 데이터 베이스 크기라 함은 작업 공간이나 데이터 베이스 관리체계(Database Management System : DBMS)를 사용함에 따라 추가로 요구되는 저장공간 등을 제외한 것이다.

그리고 현재 행정전산망 구축 중이거나 계획 중인 주민, 부동산, 고용, 자동차, 경제통계 및 통관에 관한 여섯개 업무의 성질을 분석한 결과, 평균 약 40%의 업무 처리 형태가 온라인 처리일 것으로 예상됨으로 강력한 온라인 트랜잭션 관리기능이 중요하게 부각된다.

한국전자통신연구소에서는 행정전산망용 주전산기 개발 사업을 주관하고 있으며 그 일환으로 관계형 DBMS 개발연구를 수행하고 있다. 본 과제에서는 행정전산망에서 필요한 DBMS를 개발 공급하고 데이터 베이스에 관련된 기술을 지원하는 것을 주목적으로 한다. 본 과제를 계기로 데이터 베이스 관련 기술을 축적하고 앞으로 연구 활동을 위한 기반 기술 습득과 국제적 연구 동향을 파악하여 본과제가 끝나는 1991년부터는 관련 분야에서 선진 연구를 담당할 수 있는 제반 조건을 갖춘다.

첫해엔 기존하는 관계형 DBMS 중 하나를 선정하여 한글 자료 처리 기능을 추가하고 행정전산용 주전산기의 도입기종인 Tolerant 사의 Eternity 시스템 상에서의 성능 개선을 하여 응용 프로그램 개발을 담당하는 데이터통신(주)에 제공한다. 첫해의 연구 개발 내용은 다음과 같으며 아울러 올해는 연구 개발 환경을 구축하며 다음해부터 개발될 DBMS 기능적 모습 결정과 그에 필요한 요소 기술을 습득한다.

- 한글 처리의 범위를 포함한 한글 관계형 DBMS의 모습 결정,
- 한글 자료 표현 및 운영 체제와의 연관 관계 결정,
- 선정된 DBMS의 분석과 한글 처리 기능 지원을 위한 수정,
- 한글 사용자 지침서 작성 및 발간.

2차년도부터 사용자 인터페이스 분야의 메뉴 시스템, 보고서 작성기, 화면폼을 이용한 질의기(Query By Form), C 프로그래밍 언어 접속기 등을 자체 개발하며 온라인 트랜잭션 처리 지원 기능을 연구 개발한다. 또한 행정전산망 주전산기 목표 시스템의 특성인 다중 처리를 이용한 성능 개선이 연구 내용의 주축을 이룰 것이다. 마지막 해엔 개발된 DBMS를 목표 시스템에 이식하여 성능 평가를 거쳐 현장 시험을 한다.

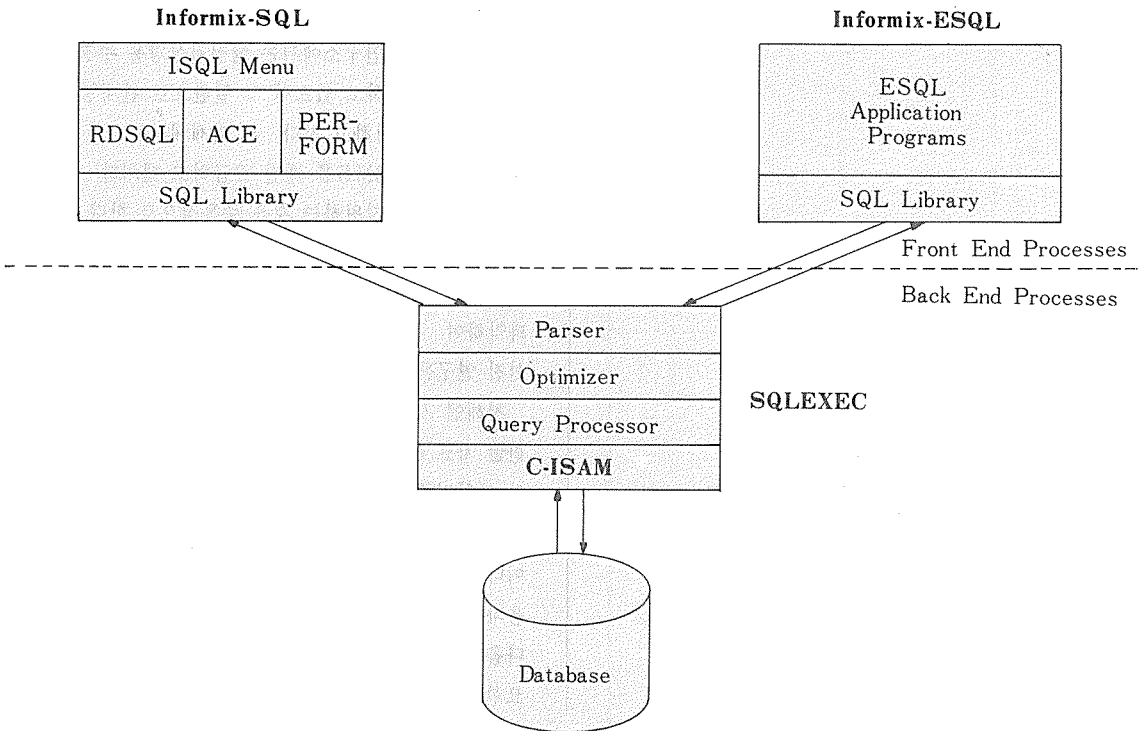
본고에서는 일차년도에 수행할 한글 관계형 DBMS 개발에 관해 다루고자 한다. 머리말에 이어서 2절에선 선정된 DBMS를 소개하고 3절에선 그의 한글화 범위를 설명한다. 4절에선 한글 처리 기능 설계 방법을 제시하고, 현재의 구현 상황을 5절에서 간략히 설명한다. 그리고 6절에서 앞으로 개발 방향을 밝히면서 본고를 맺는다.

## 2. 선정 DBMS 소개

데이터통신(주)에 선정된 DBMS는 Informix Software Inc.에서 개발된 관계형 DBMS인 Informix-SQL (ISQL, 2.10판)이다.<sup>2)</sup> 이 제품은 자료 질의어로는 국제공업 표준으로 거의 채택된 SQL(Structured Query Language)를 지원하며, 자료접근법으로는 ISAM(Indexed Sequential Access Method)을 바탕으로 하는 C-ISAM(3.0판)을 제공하는데 C-ISAM의 외부 인터페이스는 유럽과 미국의 11개 컴퓨터 회사들이 서로 소프트웨어의 호환성을 지원하기 위해 조직된 X/OPEN<sup>3)</sup>에서 채택된 제품이다. C 언어에서 데이터 베이스 응용을 내포할 수 있게 Embedded-SQL/C(ESQL/C, 2.10판)도 제공한다.

### 가. Informix-SQL 구조

ISQL의 구조는 <그림 1>에서 보여지는 것과 같이 Front-end와 Back-end 프로세스로 구성된 이중 프로세스 구조를 갖는다. Front-end 프로세스는 사용자가 여러 형태의 질의문을 만들고 데이터 베이스상에서 다른 수행들(보고서 작성, 화면 폼 사용한 자료 입출력, 대화



<그림 1> Informix-SQL 구조

식 질의)을 할 수 있는 환경을 제공한다. Back-end 프로세스에서는 Front-end에서 넘어온 사용자 요구를 검사, 요구의 정당성을 검증하며 화일 관리와 모든 데이터 베이스의 실제 입출력을 담당한다.

Front-end와 Back-end 프로세스는 서로 파이프를 통하여 메시지 패킷을 주고받는다. Front-end 패킷은 Back-end를 위한 자료 접근 요구를 포함하여 Back-end 패킷은 요구된 명령어 수행의 결과(데이터와 요구된 상태 메시지들)를 반환한다.

### 나. 모듈별 기능 설명

#### 1) Front-ends

가) RDSQL : SQL이 모체인 ISQL 질의어로 아래와 같은 언어들로 구성

- 자료 정의 언어
- 자료 처리 언어
- 자료 제어 언어

나) ACE : 관계형 보고서 작성기로 아래와 같은 2개의 프로그램으로 구성

- Saceprep : 보고서 양식을 해석한다.
- Sacego : 해석된 보고서를 작성한다.

다) PERFORM : 화면 폼에 의한 데이터 베이스의 자료 입력과 QBF (Query by Form) 형태의 질의를 지원하는 부속 시스템이다.

- Sformbld : 시스템 편집기로 만든 폼 정의 화일을 컴파일한다.
- Sperform : 화면 폼을 이용하여 자료입력과 QBF 형태의 질의를 수행

#### 2) Back-ends

가) Parser : 입력 질의문 스트링의 구문을 해석한다.

나) Optimizer : Parser의 출력으로 부터 질의를 처리하는 최적의 순서를 결정하는 Linear Tree를 만든다.

다) Query Processor : Optimizer의 Linear Tree 순서로 데이터 베이스의 테이블을 접근한다. 그리고 출력 컬럼을 반환한다.

라) C-ISAM : DBMS의 물리적 계층의 일을 수행하는 부분으로 자료접근 뿐만 아니라 다음과 같은 일도 처리해 준다.

- 색인 화일을 읽고 쓰는 일
- 디스크 화일을 읽고 쓰는 일

-감사 트래일과 트랜잭션 Logging(고장 회복 지원 기능)

-동시성 제어(테이블과 Row-Level 잠금)

-테이블과 데이터 베이스의 사용자 접근 제어

## 3. 한글 처리 범위

국내 관계형 DBMS의 활용 실태의 사례조사 결과<sup>4)</sup> 개인용 컴퓨터 계열에서 몇개의 DBMS의 한글화가 되어있는 상태이다. 소형 컴퓨터 이상 계열의 경우는 관계형 DBMS가 보급 초기이며 제대로 한글화가 이루어진 것은 거의 없다. 외국의 경우, '80년대 초부터 관계형 DBMS가 상용화되어 현재는 본격적으로 사용되고 있는 실정에 비추어보아 국내에서도 자체 개발에 앞서 완벽한 한글 처리가 가능하도록 DBMS의 한글화가 시급하다. 본고에서 다루는 한글화된 ISQL의 한글 지원 기능의 범위는 다음 여섯가지로 요약할 수 있다.<sup>5)</sup>

### 가. 한글 자료형태 제공

선정 DBMS에서 제공하는 기본 자료형태(Data Type)에 한글 자료형태(Hchar)를 추가하여 기본 자료형태를 확장한다. 한글 자료형태는 (한글문자 | 영문자 | 숫자 | 심볼)의 총합으로 널(Null)을 포함한다.

### 나. 한글 자료 처리 함수 지원

한글 자료형태로 정의된 테이블 컬럼의 값인 한글 자료를 이용하는 모든 오퍼레이션들이 오류없이 수행될 수 있도록 한글 자료 처리 함수를 구현, 지원한다.

### 다. 한글 식별자 사용

선정 DBMS에서 식별자로 쓰이는 영문 이름을 한글 이름으로 쓸 수 있도록 한다. 여기에는 데이터 베이스 이름, 테이블 이름, 컬럼 이름 그리고 기타 DBMS에서 사용되는 보고서나 화면 폼 명세 화일(Screen Form Specification File) 이름 등이 포함된다.

### 라. 한글 메뉴 제공

선정 DBMS에서 제공하는 모든 메뉴를 사용자의 이해를 돕기위해 한글 메뉴로 대체한다.

### 마. 메시지의 한글화

선정 DBMS에서 제공하는 영문 메시지(상황 메시지, 도움 메시지, 오류 메시지 등등)를 한글화 한다.

### 바. 한글 사용자 지침서

위에서 채택한 한글화의 범위를 포함시켜 개발될 IS-QL의 동작 기능을 설명하는데는 기존하는 사용자 지침서의 한글화가 요구된다. 현재, C-ISAM, Informix-SQL 그리고 ESQL/C 영문 사용자 지침서의 한글화가 완성되었다<sup>6), 7), 8)</sup>

## 4. 한글 처리 기능 설계

### 가. 한글 코드 선택.

일차년도 과제 목표인 선정 DBMS의 한글화를 수행하는데 한글 코드로 2바이트 완성형 코드가 선택되었다. 이는 한글 DBMS를 올릴 도입 기종 시스템의 운영체제가 완성형 코드를 채택하여 개발 중이기 때문이다. 완성형 코드를 선택함으로써 우려되는 몇 가지 문제점들은 뒤에서 설명하였다.

### 나. 한글 내부 표현

2바이트 완성형 코드를 시스템 내부 코드로 받아들일 때 한글 자료를 어떻게 저장할 것인가에 대한 사항으로 <표 1>과 같이 표현하여 시스템에서 자료를 사용할 때와 보조 기억장치에 자료를 저장할 때의 자료 표현에 일관성을 유지하여 효율적인 관리 및 처리를 할 수 있게 한다.

### 다. 한글 자료 형태

테이블에서 어떤 컬럼이 한글 영문 혼용이든지 한글 전용으로 사용되면 그 필드의 자료 형태를 Hchar로 지정하여 사용한다. 컬럼이 Hchar(n) 형태로 선언되었다면 실제 길이는 (2\*n)으로 잡히지며 n 문자까지 가질 수 있다.

### 라. 한글 자료 처리 지원 함수

1) 한글 1바이트의 타입을 결정하는 마크로 :

- 한글인가?
- 현재 가리키는 바이트가 한글 첫번째 바이트인가?
- 현재 가리키는 바이트가 한글 두번째 바이트인가?

<표 1> Hchar 자료 형태인 자료 표현법

한영문 혼용 형태 (Hchar Type)의 코드 지원			
	입출력 및 통신시	주 기억 공간에서	자료저장시 (보조 기억장치에서)
영문	ASC II	ASC II	ASC II
한글	2 바이트 완성형	2 바이트 완성형	2 바이트 완성형

2) 한글 스트링 처리에 필요한 함수 :

- 스트링의 문자수 계산 및 비교, 검색
- 임의의 스트링에 대하여 주어진 문자수(바이트수)에 해당하는 바이트 수(문자수) 계산
- 임의의 갯수의 문자 Copy/Concatenate

### 마. 원천 프로그램 수정 원칙

1) 한글 기능 추가 부분은 원천 코드에 #Ifdef HANGUL 형식을 써서 조건 컴파일을 하여 영문 판의 Super Set이 되도록 한다.

2) 한글 1문자와 영어 1문자를 동일 수준에서 처리한다.

- 한글일 경우 커서는 한글 첫번째 바이트에 위치하게 한다.
- 삽입, 삭제, 대치의 경우 한글 2바이트가 한 문자로 처리될 수 있도록 한다.

## 5. 구현 상황

관계형 DBMS 개발 과제를 수행하는 일차년도인 현재, DBMS에서의 한글화 범위와 한글 기능 설계 방법을 연구, 결정하였고, 이를 선정 DBMS에 구현했을 때 필요한 사용자 지침서를 완성하였다.

한글 기능 설계방법을 결정하는데 발생했던 문제점으로 한글 코드를 결정하는 것이었는데 이는 한글 처리 지원이 가능한 UNIX 운영체제가 2바이트 완성형 코드를 선택하여 구현되었고 (예 : SUN 한글 O.S., Eternity O.S. 인 한글 TX), 또한 행정전산망용 주전산기 개발 스펙에 한글 코드를 2바이트 완성형 코드로 정하였기 때문에 완성형 코드를 일차년도 과제의 한글 코드로 택하였다. 그러나 완성형 코드 선택에는 몇가지 미해결의 숙제를 안고 있는데 예를 들면 코드 테이블에 없는 글자 지원, 음소 단위의 처리 불가능, 한글 코드로 인한 사용할 수 없는 특수문자 등으로 이에

대한 해결안이 연구되어야 할 것이다.”

또한 데이터 베이스 전문 분야의 용어 통일이 문제로 대두되었다. 일차적으로는 정보과학회가 발행한 “컴퓨터용어대사전”과 한국전자통신연구소에서 발행한 “전기통신용어사전”을 참고하고 그리고도 모자라는 것은 자체적으로 용어의 한글화를 꾀하였다.<sup>10)</sup>

선정 DBMS의 한글화 구현은 원천 코드를 도입, 분석을 완료하여 분석 검토 모임을 가졌고 현재 한글 기능을 원천 코드위에서 설계, 코딩 중이다. 일차년도가 끝나는 시점인 1988년 6월엔 도입기종인 Eternity 시스템에 이식하여 데이터통신(주)에 공급할 계획이다.

## 6. 맺음말을 대신하여

본 한글 관계형 DBMS 개발 연구 과제를 수행하면서 마주친 중요한 고려 사항은 다음과 같다.

○UNIX 운영체제에서 개발된 DBMS의 성능 문제.

○DBMS를 사용하는 온-라인 트랜잭션 지원 기능.

○자체 개발의 범위를 최대한 넓혀 DBMS의 국산화.

행정전산망용 주전산기의 도입기종이나 자체개발 목표 시스템은 모두 UNIX 운영체제를 채택하고 있다.

그래서 일어나는 첫번째 문제는 잘 알려진 문제로 특히 UNIX의 화일 관리시스템을 사용할 경우 일어나는 문제이다. 최근에 새로 나오는 DBMS들을 살펴보면 UNIX의 화일 시스템을 포기하고 직접 디스크 입출력을 사용하는 경향이다. 예를 들면, UNIFY의 EXPRESS I/O, INFORMIX-TURBO, Sybase, 그리고 TUXEDO System/D 등이다. 이들은 모두 Raw I/O 기능을 사용하여 직접 디스크 입출력을 함으로써 DBMS의 효능을 높이는데 중점을 두었다. 우리도 향후 목표 DBMS를 개발하는데 이 방법을 반드시 참고하고 채택하여야 할 것이다.

앞의 머리말에서도 언급하였듯이 행정전산망에서 온-라인 트랜잭션 처리 지원이 필수적으로 요구된다. 이는 행정전산망 뿐만 아니라 일반 상업적 응용에서도 많이 요구됨으로 연구 개발 결과의 상품성을 고려해서도 앞으로 DBMS 개발과 함께 중점적으로 연구되어야 한다. 외국의 선진 기업에서도 그 개발이 이제 끝나 새로운 상품으로 소개하기 시작한 상태이다. 예를 들면 영국의 ROOT Business Systems Ltd. 사의 Uni-TECS나 AT & 사의 TUXEDO System/T 등이 이

제 개발을 막 끝내고 판매에 들어갈 예정이다.

마지막으로 행정전산망용 DBMS를 개발하면서 국산화를 반드시 실현하여 DBMS 분야의 기술 자립을 이루어야 하는데, 그 일의 양과 어려움은, 한예로 Informix-SQL 시스템의 원천 코드 양이 C 언어로 18만 줄에 해당함을 보임으로써 설명된다. 그래서 본 과제의 주 담당 부서인 한국전자통신연구소 뿐만 아니라 여러 대학교와 관련 업체와도 긴밀한 협조체계를 이루어 공동 협력 개발이 되어야할 것이다.

## 참고 문헌

1. 데이터통신(주), “행정전산망 주전산기 요구사항 및 사양 (안)”, 데이터통신(주), 1987. 2.
2. Informix Software Inc. “INFORMIX-SQL : Relational Database Management System”, Reference Manual, July 1987.
3. X/OPEN, X/OPEN Portability Guide, Vol. 4, Data Management”, Elsevier Science Publishers B. V., January 1987.
4. 김문자, 허대영, 김창석, 홍동권, 김명준, 오길록, “데이터 베이스 관리체제에서 한글 처리 기능에 관한 연구”, 과학기술처 '86 특정연구결과 발표회 논문집, pp. 119-123, 1987. 7.
5. 이애영, 허대영, 김문자, 김승겸, “HISQL의 한글 지원 기능에 대한 제안”, TD87-1700-28, 행정전산망용 주전산기 개발본부, 한국전자통신연구소, 1987. 8.
6. 이진수, 김창석, 김명준, “HC-ISAM 사용자 지침서”, TD87-1760-53, 행정전산망용 주전산기 개발본부, 한국전자통신연구소, 1988. 1.
7. 이애영, 허대영, 조완섭, 김문자, 김승겸, 김창석, 손지수, “HINFORMIXSQL 사용자 지침서”, TD87-1760-54, 행정전산망용 주전산기 개발본부, 한국전자통신연구소, 1988. 1.
8. 손정규, 김종민, 송근태, 오석균, 김명준, “INFORMIX-HESQL/C 사용자 지침서”, TD87-1760-64, 행정전산망용 주전산기 개발본부, 한국전자통신연구소, 1988. 2.
9. 조완섭, “DB팀의 한글 코드에 대한 고려 사항”, MM87-WSCHO-03, S/W 공학 연구실, 한국전자통신연구소, 1987. 8.
10. 이애영, “용어의 한글화”, TD87-1700-27, 행정전산망용 주전산기 개발본부, 한국전자통신연구소, 1987. 6.