

# 인터넷기반 실시간 원격 데이터 분석 시스템 설계 및 구현

<sup>0</sup>김중만\*, 김지환\*, 이성구\*, 안세영\*\*, 박미경\*\*, 김홍기\*\*\*

\*한신대학교 정보시스템공학과, \*\* (주)수엔지니어링컨설팅, \*\*\* (주)창민테크

{forbliss, aclouds, sklee}@hanshin.ac.kr, \*\*{syahn, mkpark}@waterenc.com, \*\*\*rainbow@changmin.co.kr

## Design and Implementation of a Real-Time Remote Data Analysis System based on Internet

<sup>0</sup>JongMan Kim\*, JiHwan Kim\*, SungKoo Lee\*, SeYoung Ahn\*\*, MiKyoung Park\*\*, HongKi Kim\*\*\*

\*Dept. of Information System Engineering, Hanshin University

\*\*Water Engineering and Consulting, Inc.

\*\*\*Changmin Tech. Co., Ltd.

### 요 약

지리적으로 분산된 컴퓨터의 정보 교환을 위해 탄생된 인터넷은 네트워크 응용 프로그램을 개발하기 위한 시스템 개발자들이 로컬 컴퓨터에서 발생하는 실시간(real-time) 데이터의 중앙관리를 위한 통신 수단을 제공한다. 본 논문은 지리적으로 분산된 하/폐수 처리장에서 발생하는 다양한 종류의 실시간 데이터에 대한 효과적인 중앙 관리와 이러한 데이터를 처리, 분석하여 필요한 정보를 즉시 각 처리장에 제공함으로써 처리장 사이의 정보 공유는 물론 감독기관의 각 처리장에 대한 효율적인 운영관리를 위한 인터넷 기반 원격 데이터 분석 시스템(RRDAS)의 설계 및 구현에 대해 논한다.

### 1. 서 론

최근 우리 나라의 환경 산업은 지속적으로 발전하여 왔으며, 지역적으로 분산된 환경 기초 시설이 급격히 증가함에 따라 이러한 기초 시설들로부터 발생하는 데이터의 효율적이고 효과적인 운영 및 관리에 대한 관심은 크게 증가되었다. 특히, 지리적으로 분산된 하/폐수 처리장에서 발생하는 다양한 종류의 실시간 데이터에 대한 중앙 관리와 이러한 데이터를 처리, 분석하여 필요한 정보를 즉시 각 처리장에 제공함으로써 처리장 사이의 정보 공유는 물론 감독기관의 각 처리장에 대한 원격 제어를 손쉽게 할 수 있는 실시간 원격 제어 시스템에 대한 필요성이 요청되고 있다[1].

컴퓨터사이의 정보 교환을 위한 통신 수단으로써 탄생된 인터넷과 함께 Java 프로그래밍 언어는 시스템 개발자들에게 분산 네트워크 응용 프로그램을 개발하기 위한 수단을 제공한다 [2][3]. 특히, 네트워크 프로그래밍 언어로써 개발된 Java는 TCP/IP 프로토콜을 기반으로 하는 인터넷에서 프로그래머들에게 실시간 응용 프로그램을 개발하기 위한 상위 수준의 API(Application Programming Interface)를 제공한다. 본 논문에서 구현된 RRDAS 시스템은 Java기반 네트워크 시스템으로써 실시간으로 수집되는 처리장 데이터의 네트워크 통신과 수집을 위한 (1) 데이터수집기 시스템, 처리장 데이터의 중앙관리를 위한 (2) 데이터관리 시스템, 실시간 자료는 물론 과거 자료의 시각화를 통한 (3) 시각화 시스템, 영역 지식과 분석 결과의 비교를 위한 (4) 의사결정 지원 시스템과 같은 4개의 서브 시스템으로 구성된다.

본 논문은 다음과 같이 구성된다. 2장은 RRDAS를 구성하는 전체적인 구조 및 설계를 묘사한다. 3장에서는 시스템개발환경을 기술하고, 4장에서는 구현내용을 다룬다. 마지막으로 5장에서 결론 및 추후 연구과제에 대해 언급한다.

### 2. RRDAS 설계

#### 2.1 시스템 구성도

효과적인 실시간 데이터의 중앙 관리와 의사 결정을 통한 원

격 제어를 위해 RRDAS 시스템의 전체적인 구성도는 [그림 1]과 같다. 실시간 데이터 수집기는 각 처리장에서 발생하는 데이터를 인터넷을 통해 프로젝트 데이터베이스에 저장한다. 시각화 시스템에서는 용이한 사용자 분석을 위해 저장된 과거 데이터는 물론 실시간 데이터의 시각화가 수행된다. 환경 전문가의 추론능력을 프로그래밍화한 의사결정지원 시스템은 영역 지식 데이터베이스의 지원과 함께 원격 제어를 위한 결과를 도출한다. 이 과정에서 발생한 결과 또한 프로젝트 데이터베이스에 저장된다.

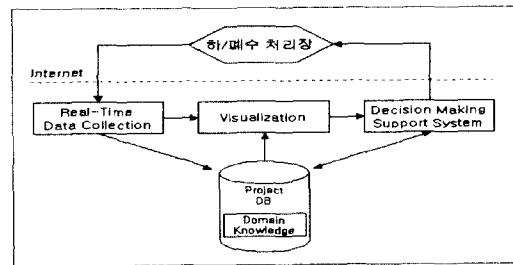


그림 1 시스템 구성도

#### 2.2 GUI 설계

RRDAS 시스템 사용자는 이론적으로 숙련된 환경분야 전문가가 아닌 관리자가 대상이므로 간편성과 효율성 측면을 고려한 입력 GUI 설계가 필요하다. 또한, 본 시스템의 출력 GUI는 사용자가 각 처리장에서 발생하는 상황을 신속하게 이해하고, 처리장에서의 비정상적인 동작을 인지하기 위해 실시간 데이터와 과거 정상적으로 발생된 데이터들과 쉽게 비교 분석할 수 있는 시각화를 위한 설계가 필요하다. RRDAS는 사용자의 다양한 플랫폼에서 일관된 look-and-feel을 제공하고, 효율적인 GUI 프로그래밍 능력을 제공하는 Java JFC 패키지가 이용되었으며, 신속한 비교 분석을 위한 시각화 작업을 위해 Sitraka Software에서 제공하는 JClass Chart 패키지를 이용하였다[4].

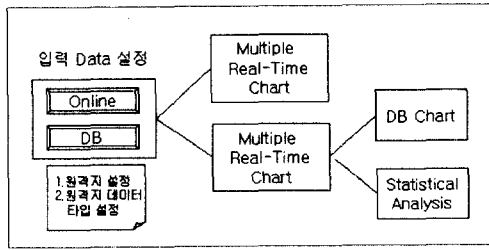


그림 2 입/출력 GUI

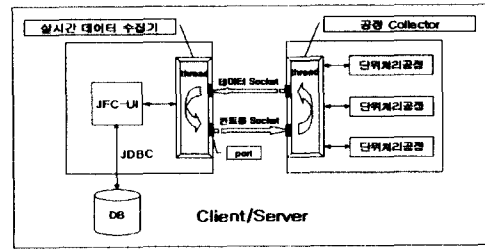


그림 3 RRDAS 개발환경

[그림 2]는 RRDAS 시스템에서 설계된 5가지 형태의 입/출력 GUI를 보인다. 본 시스템에서 관리와 분석을 위해 사용되는 변수들은 TCP/IP 기반의 네트워크 설정을 통해 이루어진다. 따라서 사용자는 원격지에서 실시간으로 변화하는 데이터를 수집하기 위해 원격지 설정과 설정된 원격지 데이터의 타입을 결정하는 단계를 필요로 한다.

네트워크 설정 후, 사용자는 실시간 분석을 위해 필요한 시각화 변수들을 설정할 수 있는 GUI와 실시간 데이터를 데이터베이스에 저장하고 다양한 검색 기능을 제공하는 GUI를 이용한다. 실시간 변화하는 데이터의 분포를 관찰하고, 이에 따른 의사결정을 쉽게 하기 위해 데이터를 시각화해서 볼 수 있는데, 이를 위해서 사용자는 분석하기 위한 데이터를 다양한 시각화 과정을 통하여 관찰할 필요가 있다. 'Multiple Real-Time Chart'는 시각화를 위한 차트설정 및 데이터 수집 설정 등의 사용자 정의 메뉴가 포함된다. 또한, 저장된 데이터 혹은 과거 데이터의 시각화를 통한 용이한 분석을 위해 검색된 결과로부터 다양한 사용자 요청을 만족하는 GUI 설계가 필요하다. 예를 들면, 'Statistical Analysis'는 주어진 기간 내에 발생한 과거 데이터의 통계분석을 위한 다양한 메뉴들을 제공한다.

2.3 데이터베이스 설계

RRDAS의 프로젝트 데이터베이스는 처리장의 수많은 단위 공정에서 발생하는 데이터, 기계 장비의 상태 데이터, 시스템 자체에서 발생하는 중간 데이터, 의사 결정 과정동안 발생하는 데이터, 환경기초 분야의 영역 지식을 위한 데이터를 저장하기 위해 설계된다. Java 응용 프로그램인 RRDAS는 ODBC 데이터베이스인 프로젝트 데이터베이스를 연결하고 제어하기 위해 JDBC API를 이용하였으며 ODBC 데이터베이스 연동을 위해 JDBC-ODBC 드라이버를 사용했다[5].

3. RRDAS 개발환경

인터넷을 통해 각 처리장에서 발생하는 데이터를 관리자 컴퓨터에 연결하기 위해 TCP/IP 프로토콜 기반 Java 네트워크 프로그래밍을 가능하게 하는 'Socket'을 이용한다. 특히, 패킷 지향의 네트워크 데이터 통신보다 안정적인 서비스를 제공하는 Java Socket 과 Multithreading 기능은 RRDAS와 같은 신뢰할 수 있는 실시간 데이터 통신을 필요로 하는 시스템에 적절하다. RRDAS의 서브시스템들은 모두 Java로 구현되었다. Java 언어 자체에서 제공되는 다음과 같은 특징들은 RRDAS와 같은 시스템을 구현하기 위해 적합하다[6].

- Platform-Independent : 사용자 혹은 처리장에서 다양한 컴퓨터환경에 독립적인 GUI 프로그래밍
- Object-Oriented : 용이한 소프트웨어 재사용 컴포넌트(예, JClass 패키지) 기반의 프로그래밍
- Network-Oriented, Multithreaded : 신뢰할 수 있는 실시간 데이터 통신을 위한 프로그래밍

[그림 3]은 RRDAS의 client/server 개발환경을 보인다. 사용자 컴퓨터와 처리장간의 네트워크 데이터통신은 Multithreading을 사용하여 「실시간 데이터 수집기」와 「공정 Collector」 간에 실시간으로 이루어진다. 사용자는 IP주소와 적절한 port번호로 임의의 하수처리장과 연결을 시도하는데, 이때 Socket을 통해 데이터들을 수집한다. 또한 원격지에서 시스템 사용자는 자신이 원하는 데이터 타입을 설정하고, 온라인 데이터를 DB에 저장하며 또한 저장된 데이터를 검색·분석하여 결과를 해당 처리장으로 보냄으로써 원격제어를 수행할 수 있다. 본 시스템의 데이터 통신은 양방향 통신으로 처리장과 사용자 컴퓨터 각각은 Client/Server의 역할을 동시에 수행한다.

4. RRDAS 구현내용

실시간 데이터를 수집하기 위하여 임의의 처리장에서 가동되고 있는 소켓 서버에 연결할 수 있는 인터페이스를 가지며 이 인터페이스를 통해 사용자는 처리장 컴퓨터의 IP 주소와 Port번호에 의해 네트워크 연결을 가능하게 한다. 이때, Java 'ServerSocket' 클래스는 처리장 컴퓨터의 한 스레드로 수행되고 있어야 한다. [그림 4, 5]은 실시간 데이터의 시각화 예제를 보인다. 특히 [그림 4]의 인터페이스로부터 사용자는 다양한 차트 선택, 차트 이름과 데이터수집 시간설정, 실시간 입력 변수들을 설정된 차트에 연결하는 방법을 제공한다. [그림 5]은 사용자에게 의해 주어진 시간에 따라 '분석데이터1', '분석데이터2' 라고 불리는 2개의 차트를 각각 'Plot Chart' 그리고 'Scatter Plot Chart' 타입으로 그린다.

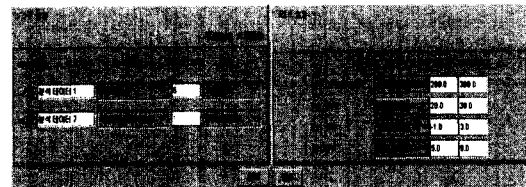


그림 4 차트 형태 설정 화면

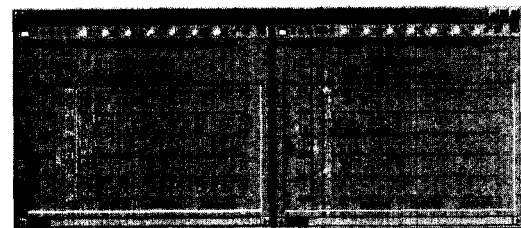


그림 5 실시간 데이터 시각화

[그림 6,7,8]는 데이터베이스 검색 인터페이스와 관련된 RRDAS 시스템의 결과 화면을 보인다.

예를 들면, 과거에 저장된 '온라인' 데이터중 사용자가 '펌프 유입' 단위공정으로 부터 발생한 '유량', '온도', 'PO4-P', 'SS' 값들에 대한 데이터를 검색한 결과는 [그림 7]에서 보여진다. 결과 데이터로부터 사용자는 주어진 날짜 사이에 저장된 데이터의 검색은 물론 Drag & Drop 기능에 의해 선택된 데이터들의 시각화(차트) 및 통계 결과를 통해 용이한 분석을 할 수 있다. 신속한 데이터베이스 검색과 통계 처리를 위하여 본 논문에서 주로 사용된 데이터 구조는 HashTable 과 Vector이다.

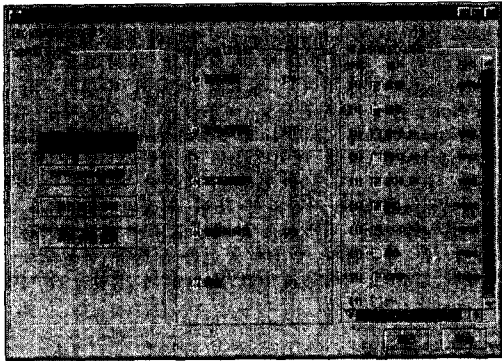


그림 6 사용자 질의 화면

2001-06-18 12:01:30.0	76.0	21.0	1.122	21.0
2001-06-18 12:01:30.0	71.0	31.0	1.003	23.0
2001-06-18 12:01:30.0	72.0	32.0	1.325	21.0
2001-06-18 12:01:40.0	73.0	33.0	1.004	23.0
2001-06-18 12:01:42.0	73.0	13.0	1.006	22.0
2001-06-18 12:01:43.0	76.0	13.0	1.122	22.0
2001-06-18 03:31:16.0	80.0	27.0	1.004	23.0
2001-06-18 03:31:17.0	74.0	24.0	1.212	23.0

그림 7 데이터베이스 검색 결과

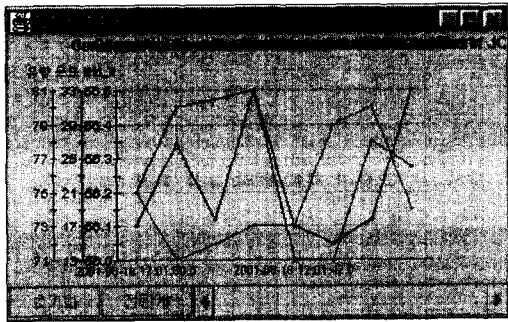


그림 8 데이터베이스 시각화

[그림 9]는 Visualization 프로그래밍을 위해 사용된 'JClass Chart'에서의 클래스 계층구조와 RRDAS에서 개발된 클래스들간의 연관관계를 보인다. 시각화를 위해 재사용하고 있는 JClass 패키지에는 차트를 위해 두 가지 중요한 세부 package를 제공하는데, 시각화를 위한 차트관련 클래스들과 데이터 소스 관련 클래스들이 그것이다. RRDAS 시스템은 실시간 데이터와 DB관련 시각화 구현을 위해 MultiChart 클래스를 재사용하

였고(OnlineChart, ZommXChart), DataSource를 구현하기 위해 JCDefaultDataSource와 JDBCDataSource를 재사용하였다 (DetailDataSource, DBSource). 이들로부터 'composition' 관계에 의해 ProcessChart는 실시간 데이터의 시각화를 위한 클래스이며, DetailChartFrm는 데이터 검색과 과거 데이터의 시각화를 위한 주된 클래스가 된다.

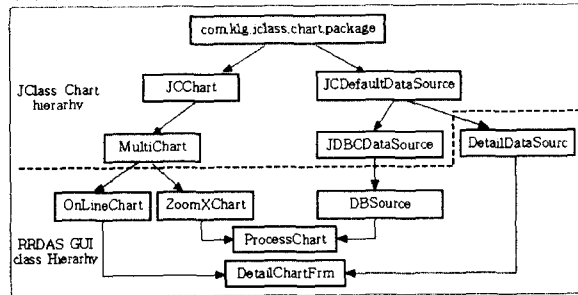


그림 9 시각화 클래스 계층구조

### 5. 결론 및 향후과제

본 논문에서는 지역적으로 산재된 산업 기초 시설에서 발생되는 데이터들의 실시간 처리를 통한 효율적인 원격 제어를 가능하게 하는 RRDAS 시스템에 대해 언급했다. 이 시스템은 사용자에게 history 자료와 실시간 자료의 시각화를 통한 분석의 용이함을 제공할 뿐만 아니라 의사 결정 지원 시스템의 도움으로 지속적으로 각 처리장의 최적 운영 조건을 제시할 수 있다.

향후 과제로서 각 처리장의 공정 프로세스의 모델링과 이러한 모델링을 이용하여 원거리에서 실제 처리장에서 발생될 수 있는 결과를 예측하는 시뮬레이션 프로그램의 개발은 RRDAS 시스템의 사용자에게 유용한 서브 시스템일 것이다. 모델링/시뮬레이션 프로그램과 본 논문에서 언급된 서브 시스템들과의 유기적인 협력관계에 의해 사용자는 시간과 공간의 흐름에 따라 변화하는 과정은 물론, 여러 가지 다른 변수를 이용한 시뮬레이션을 실행함으로써 해당 처리장에서의 프로세스를 보다 더욱 잘 이해할 수 있다. 현재 RRDAS의 의사 결정 지원 시스템의 추론 능력과 영역 지식을 위한 데이터베이스의 보다 완벽한 구현을 위해 환경 기초 분야 전문지식에 대한 많은 연구가 필요로 될 것이다.

### 6. 참고문헌

- [1] H. Praehofer, "Object Oriented Modeling and Configuration of Simulation Program," <http://citeseer.nj.nec.com/praehofer96object.html>
- [2] 김형일, 이승룡, "Java 기반 SCADA 시스템 설계 및 구현," 정보처리학회 추계 학술발표 논문집, p.951-p.954, 1998.
- [3] J. Farley, "Java Distributed Computing," O'Reilly, 1998
- [4] <http://www.sitraka.com/software/jclass/jclasschart.html>
- [5] <http://java.sun.com/j2se/1.3/docs/api/index.html>
- [6] WH. M. Deitel and P. J. Deitel, "Java How To Program," Prentice Hall, 1999.