Real-Time Custom Processing and Delivery of Large ECG Data Sets

MA Paracha, SN Mohammad, PW Macfarlane, JM Jenkins

University of Michigan, Ann Arbor MI, USA

Abstract

A tool for rapid custom-processing large data sets of ECG signals has always been in demand in research communities. To facilitate this requirement, a link must be established between the numerical processing engine and the data storage system. In our research, we have established an extensive collection of ECG signals, known as the Glasgow Data, which is accessible to subscribers over the Internet (see related Computers in Cardiology paper 'Implementation of Web Database for ECG, M.A. Paracha, S.N. Mohammad, P.W. Macfarlane, J.M. Jenkins'). Meaningful analysis, such as specific plots and images of patient ECG signals, are stored as images in this large database. Initially, these images are generated in a numerical processing and scientific visualization tool such as Matlab.

The key to rapid deployment of results from these processing tools is the facility directly, in real-time, to provide results to online users of the database. This requires connectivity between the processing engine and the database back-end. In this design we use open-source modules to establish this link, and demonstrate usability in the field of electrocardiology. With over 3300 individual patients registered in the Macfarlane Glasgow database, we are able to run researcher-specific algorithms over the entire domain of signal data and provide results instantly over the web-based system. This allows researchers to mold and study the data according to their own requirements and thus transform the data into useful knowledge, which should play a vital role in their respective research.

1. Introduction

Custom processing of large data sets plays a vital role in research studies. This paper describes the development of a system that provides access to highly valuable data (a large collection of normal ECGs) and gives an opportunity to the researcher to customize it in real time. This will allow the investigator to focus entirely on the research and the system will provide the useful results to collaborators and subscribing viewers through internet services [1]. This system will connect the database, which is comprised of over 3300 normal ECGs from neonates, infants, adolescents, and adults with the Matlab numerical analysis software package. Thus the researcher can export and import the data between Matlab and the relational database (MySQL), use the Matlab computational and analytical tools to process the data, covert the data into useful information, and store that information back into the same or another database. The system provides the feature of importing data from multiple databases during a single session and of obtaining the required results using two different data sources irrespective of their geographical location. This reduces overhead associated with the data processing and can be used to share the data between investigators to get fruitful results.

2. Materials and Methods

A System of Systems (SoS) approach was taken to design and integrate the various components comprising this product. The tools employed are the following:

-Matlab by MathWorks (Natick, Massachusetts) for numerical analysis [2]
-Matlab Database Toolbox by MathWorks [2]
-MySQL Database Server by MySQL Inc. (Seattle, Washington) [3]
-Generic FTP and terminal clients

The connection of MySQL database with Matlab using Database Toolbox requires the setup of a data source. A data source consists of data that one wants the toolbox to access. One can setup more than one data source with each having a unique identifier (source name).

In early 2003, the Glasgow data source was used as a testbed for implementing this system. Using powerful MySQL functions, the entire pediatric dataset was partitioned into various age groups and the ST-T segments of the ECG were averaged to determine at which age groups the ECG differed between males and females [4]. Significant overhead of copying files and importing native formats into Matlab were eliminated by directly accessing the data tunneled through the Database toolbox middleware directly into Matlab. Consequently, plots of outputs and results were written and associated with patients directly into the database without the need

for any file conversion utilities and manual data entry. This allowed the results of the research experiment to be available instantly on our collaborative web portal front-end while Matlab computations resided in the back-end.

Figure 1 illustrates the overall high-level architecture of the system. Similar to the tiered system discussed in our *Implementation of Web Database for ECG* publication [1], the Matlab and Database Toolbox components essentially constitute a first and second tier application, running in unison with the first and second tiers of the web services application. The third tier, which is the MySQL Database engine and the data itself, is a shared layer.

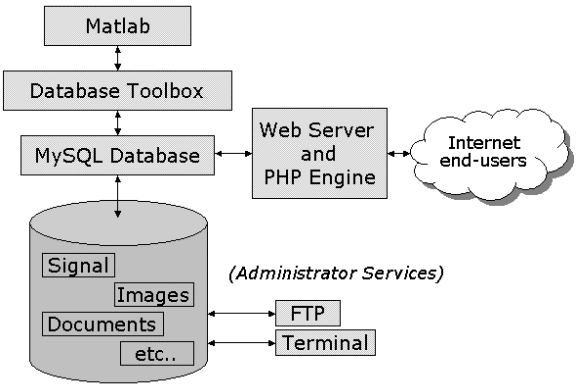

Figure 1. System of Systems view, Flow Diagram. The "Web Server and PHP Engine / Internet end-users" module was developed as part of a separate research project published in the current Computers in Cardiology proceedings [1].

3. Results

Ability to access patient database tables directly and to analyze related files (signal data, images, documents, etc.) directly from Matlab reduces analysis time by eliminating the need to condition the files for import into Matlab.

This feature allows numerical analysis results of patient data to be directly stored and updated into the database, making them available instantly to other services, such as our Web-based portal.

4. Conclusions

The System of Systems approach allows our research team to develop value-added tools that reduce turn-around time for results. The availability of an integrated web-based environment provides an easy-to-use interface through which multiple groups can collaboratively investigate the rich collection of patient data.

Although the system is currently running on a prototype computer system in our lab, we intend to further develop these tools and provide them to the wider research community. Using the same tiered approach, numerous tools can be designed and developed to access the valuable data stored on the servers. For example, legacy applications such as Ann Arbor Electrogram Libaries Viewer (AAELVIEW), which allows viewing of recorded signal files, requires the user to know beforehand the patient-to-file mapping [5]. In the proposed three-tiered architecture, the tool could be migrated to reading patient data directly through the database third tier, and provide the end-user with value-added features such as viewing other patient data through the same interface.

Essentially, the proposed developments reported in this paper serve to provide value-added tools to the research community and increase productive output. Implementations and extensions of this system can provide forums for development of advanced collaborative research environments which enhance the researcher's experience.

Acknowledgements

This work was supported in part by National Science Foundation grant BES9825125, Guidant Corporation, and an equipment grant from Intel Corporation.

References

[1] M.A. Paracha, S.N. Mohammad, P.W. Macfarlane, J.M. Jenkins, Implementation of Web Database for ECG. *Computers in Cardiology* 2003. Thessaloniki, Greece. (In press.)
[2] MathWorks web site: http://www.mathworks.com
[3] MySQL web site: http://www.mysql.com
[4] Biersack M, Mohammad S, Jenkins J, Macfarlane P, "When Do Boys Differ From Girls?" International Society for Computerized Electrocardiology 2003, Salt Lake City Utah, USA. (In press, *Journal of Electrocardiology*.)
[5] Ann Arbor Electrogram Libraries, 230 East Ontario, Suite 1105, Chicago, IL 60611. http://www.electrogram.com

Address for correspondence:

Professor Janice M Jenkins
Dept. of Electrical Engineering and Computer Science,
and Dept. of Biomedical Engineering
Medical Computing Laboratory
University of Michigan
Ann Arbor MI 48109-2122
jenkins@umich.edu