Data management system for multidirectional muon detector

Adriano Petry*[1], Fabricio Viero de Araujo[1], Miriam Pizzatto Colpo[1], Chihiro Kato[2], Juliana Bueno[3], Marlos Rockenbach da Silva[4], Lucas Ramos Vieira[3], Nikolas Kemmerich[1], Alisson Dal Lago[3], Nelson Jorge Schuch[1]

[1]Southern Regional Space Research Center - CRS/CCR/INPE - MCT in collaboration with the Space Science Laboratory of Santa Maria - LACESM/CT - UFSM, Santa Maria, Brazil.

[2]Shinshu University, Shinshu, Matsumoto, Japan.

[3]National Institute for Space Research - DGE/CEA/LAC/INPE - MCT, São José dos Campos, São Paulo, Brazil.

[4]Universidade Vale do Paraíba, São José dos Campos, Brazil.

_____

Abstract

This article describes the development and implementation of a data management system that is used for information storage and retrieval, generated by a Multidirectional Muon Detector. This instrument is in operation at the Southern Space Observatory (SSO), National Institute for Space Research (INPE), located in São Martinho da Serra, Brazil. The result of this work has enabled past and present data to be accessed in a dynamic and interactive way, through an easy-of-use Web-based data query system.

Introduction

Several ground-based space monitoring instruments have simultaneously acquired different types of data over the last decades related to atmospheric and Solar-Terrestrial phenomena. Their outputs usually have particular file formats, and information organization. Storage and querying data related to space monitoring instruments is highly important, and constitute a challenge both in scientific and operational systems. The Database Management Systems (DBMSs) emerged as a possible solution for organized storage, data access security, data consistency and integrity. Currently, these systems can provide consistent information in reasonable time, allowing data access management and maintenance, as well as speeding up data processing. A Multidirectional Muon Detector (MMD) is in operation at the Southern Space Observatory (SSO), National Institute for Space Research (INPE), located in São Martinho da Serra, Brazil. The MMD is capable of performing directional measurements of secondary cosmic rays (muon) incidence, which are created by the collision of primary cosmic rays with particles from the the Earth's atmosphere. Those measurements are important for analysis and prediction of solar phenomena, and they generate data in text files. This instrument is part of the Global Muon Detector Network (GMDN), which is an international collaboration effort developed at Shinshu University, Japan, consisting of 10 institutions from 6 countries, with real time data generation.

This paper presents the development of a database for MMD data storage, called Multidirectional Muon Detector Database (MMD-DB), as well as a Web-based interface to allow researchers to access data from the MMD-DB, using different types of queries. First the instrument is detailed, covering the muon detection technology used, data files generation and structure. After that, the DBMS implementation is described, including database modeling and data file loading process. Then, the Web-based query system is shown. At last, conclusion and future work are addressed.

The muon detection

When high-energy cosmic rays reach Earth's atmosphere, they collide with atmospheric particles and cause showers of secondary particles. These collisions may produce pions, unusual sub-atomic particles, which usually decay quickly to produce muons, neutrinos, and gamma rays [Moldwin, 2008]. Muons are elementary particles that have speed approximately equal the speed of light in a vacuum, and have a long life-time of about 2 microseconds. The counts of the incidence and direction of these particles can be obtained by measuring scintillations in a set of detectors arranged in two layers, separated by a 5 cm wide lead plate. Since muons usually have high energy capable of crossing the lead plate, directions are obtained through the analysis of the correlation between the detection of scintillations in upper and lower layer detectors, as Figure 1 illustrates. Solar activity produces coronal mass ejections (CMEs) that, if directed towards Earth, may cause geomagnetic storms on Earth, possibly damaging, for example, electronic equipments, electric power line systems and geospatial instruments. Some studies have shown that a systematic reduction on muon counts of about 2% can be detected 4-

A. Petry, F. Viero, M. Colpo, C. Kato, J. Bueno, M. Silva, L. Vieira, N. Kemmerich, A. Dal Lago, N. Schuch

8 hours ahead of a possible geomagnetic storm [Munakata et al, 2000], [Frigo et al, 2007], allowing accurate and reliable Space Weather forecasting. This forecast can be used to avoid important losses, mainly when solar activity reaches its maximum activity period.
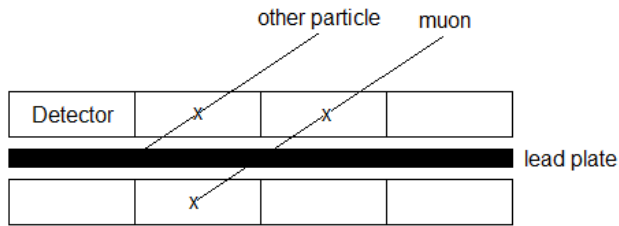

Figure 1. Schematic diagram of directional muon detection

Instrument at Southern Space Observatory

Initially, the MMD at SSO was assembled using four detectors in the upper layer and four in the bottom layer, arranged in a square (2x2x2). Between the layers, separated by a distance of 1.73 m, there was a lead plate of 5 cm width. The instrument began its operation in March 2001 providing information in nine directional channels: vertical, north, south, east, west, northeast, southeast, northwest and southwest. In December 2005, the instrument was updated to work with 28 detectors in the upper layer and 28 detectors in the lower layer, arranged in a grid of four rows and seven columns (28x7x2). The measuring system was capable of measuring 13 different directional channels. However, this large number of detectors can provide information also in several other directions, which was possible with the installation, in 2006, of a system based on Field Programmable Gate Array (FPGA) [Kilts, 2007] [Yasue, 2003]. In this new counting system, all 91 possible directions are considered.

Data production

The data produced by the MMD are generated and stored in text files on the computer installed at the SSO, which aids the instrument operation. The muon counting information is recorded every minute in a file, representing the total muon detection during that period of time. The same approach is done using time interval of ten minutes, whose information is recorded on a different file. Each one minute data file contains 60 lines, representing a total of one hour of instrument operation. Ten minutes data files contain 144 lines each, representing a total of 24 hours of instrument operation. These files are organized in one subfolder for each month, and the information structure used on both file types is the same. Data produced can be accessed remotely, using different remote access softwares. Periodically, ten minutes data files are sent to researchers at Shinshu University [Shinshu, 2011] in Japan, and they are published online at ftp://ftp.bartol.udel.edu/takao/muon_data/, including one hour data files (composed by 6 ten minute data files). Besides the data files from MMD in SSO, files from three other similar instruments, installed in the cities of Kuwait, Hobart and Nagoya, are available.

Database Modeling

Based on MMD data file structure a data modeling was developed, both conceptual and logical, in order to correctly store all relevant information and optimize data retrieval. After that, an analysis on the various database implementation technologies available was made. Among the options we decided to use PostgreSQL [PostgreSQL Group, 2011], that is an open source object-relational DBMS, with over 15 years of development. It is also extremely robust and reliable, besides being very flexible and rich in resources. It is considered object-relational since it can provide some features of object orientation, like inheritance and custom types. Studies showed that PostgreSQL provides good support to procedure-related tools, management and database administration. Regarding performance, PostgreSQL is very fast when handling large volumes of data and complex queries using multiple tables.

The database model has eight related tables using one-to-many (1: N) relationships, as shown in Figure 2. Tables "instrument_type", "instrument" and "political_location" are primarily used to identify the instrument, and also to allow inclusion of other instruments besides the MMD in operation at SSO. The other tables were specifically designed for the instrument at SSO. Each one minute, ten minutes or one hour data files with measurement data that is loaded to the MMD-DB generates a record in table "muon_file", identifying the source of information. Each line in the text files with data from measurements represents a record in the table "muon_data" and various associated records (one for each direction) in the table "muon_directional_data.

Since information about date/time of measure (table "muon_data") is separated from directional counts (table "muon_directional_data"), it is possible to store counts derived from different directional channels, allowing the addition of future directions not considered yet. In order to implement database normalization, each directional channel is identified by a record in table "directions", and each count of the table "muon_directional_data" is associated with a specific directional channel.

The barometric normalization of muon counts is performed individually for every direction, using a specific beta value. However, due to instrument ageing or adjustments, these beta values can change with time. So, each record in the table "directions" may be associated with multiple records in table "normalization_coefficients". It is also important to note that some directional counts from one hour data files already present barometric normalized counts. Thus, the "pressure_corrected" field from table "muon_directional_data" identifies if the count is normalized or not.

Information in tables "instrument_type", "instrument", "political_location", "directions" and "normalization_coefficients" are necessary for the correct database operation, but are not found in the text files being loaded. Thus, the system automatically generates these data and put them into the database, or uses them if previously entered.
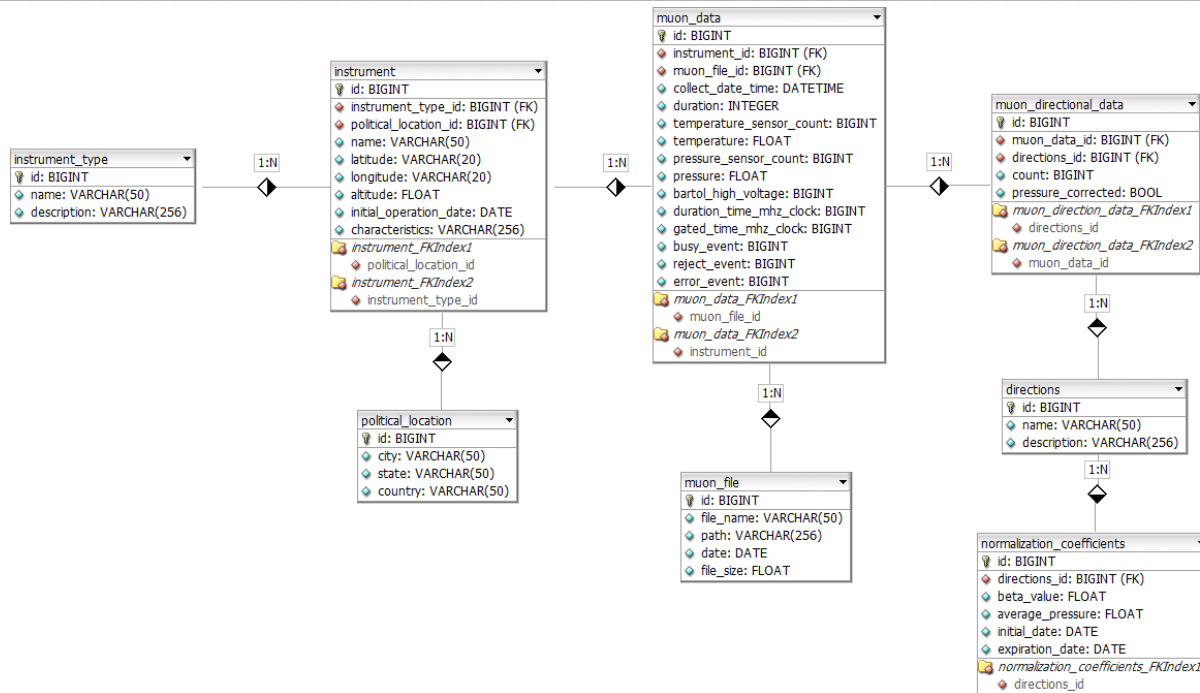
Figure 2. Database modeling for multidirectional muon detector

Data files loading

In order to properly load to database the information in the text files, it was developed a software using Java Programming Language that uses Hibernate framework [Hibernate, 2011] to access the database. This software is able to receive files containing information recorded every minute, every ten minutes or hourly data and make the reading of each piece of information, putting it in the correct place on database. Since the files are organized into one or more folders and subfolders, recursion was used in the search for valid files. Besides recursive search in folders, it was also implemented the database loading from data files available on internet, using the site ftp://ftp.bartol.udel.edu/takao/muon_data/saomartinho.
During the process of file loading, several checks are performed to ensure consistency in data being loaded. These checks basically drop values that are incompatible with the expected range or that are absurd.

Web-based query system

After database modeling and implementation, the focus became the development of a Web-based query system, so that information could be properly consulted and exploited from researchers in the area, that do not necessarily have previous knowledge about relational databases. The query system was designed to be accessed through the Web, allowing access to data dynamically and platform independent. A server, located at Southern Regional Space Research Center facility was used for both provide full access to data and regular database update. A visual interface, shown in figure 3, was constructed in order to allow data selection by different criteria without requiring the user to have advanced knowledge about relational databases. In this interface, first the user selects the initial and final date/time of interest. Besides the time interval, the user can choose an appropriate time resolution, related with the three different data files used on database loading: 1 minute, 10 minutes or 1 hour. Finally, desired directional channels are chosen as well as form of data display. It is possible to show numerical data in the browser, in text format that can be saved in the user's computer for later use, or in graphic format (shown in figure 4), where a separated chart is generated for each selected directional channel.

Figure 3. Visual interface

A. Petry, F. Viero, M. Colpo, C. Kato, J. Bueno, M. Silva, L. Vieira, N. Kemmerich, A. Dal Lago, N. Schuch
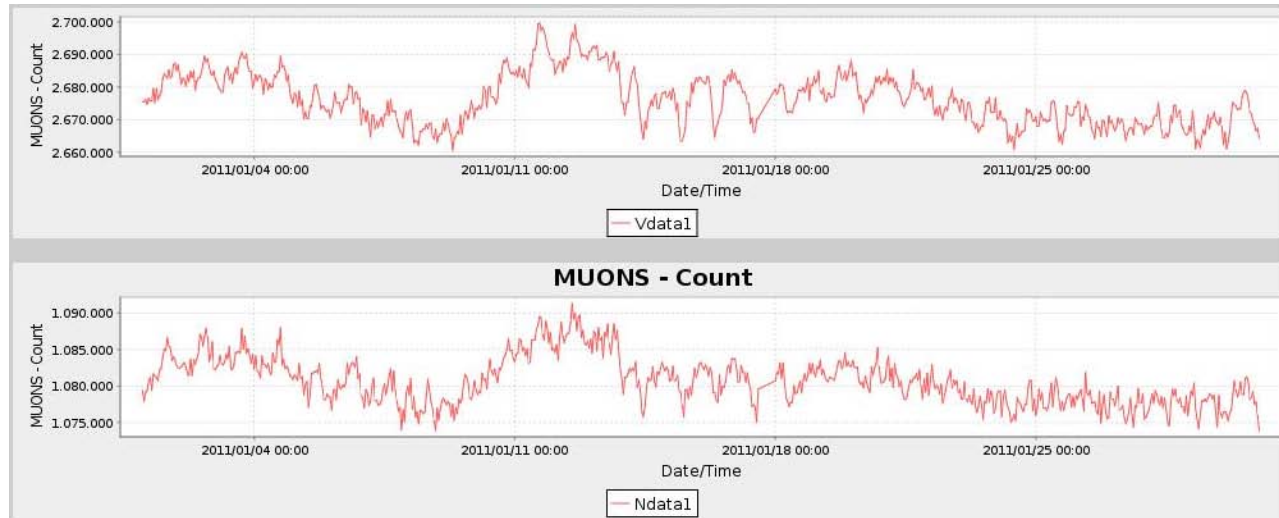
Figure 4. Example of graphic format output for the month of January 2011, directions vertical and north

When the Web-based query system development was concluded, our research group made it available online using a Web server that can be accessed at http://aquarius.lacesm.ufsm.br/site/sistema/login/index.jsp Prior to the system access, the user must provide its login and password, or first ask for it using a simple form.

Conclusion

Data management from scientific instruments is a field of great interest, since it can process information easily and lead to more qualified research. This work showed a data management system for a space monitoring instrument, composed by a relational database (MMD-DB), data file loading software and web-based query system.

It was first introduced the physics behind the muon measurements, as well as its importance. After, data production was analyzed, and database modeling was proposed. Some features about data file loading were addressed, including some details of the software developed. At the end, it was shown a web-based query system, available on the internet for the use of scientific community.

Future work may include the addition of data from other muon detectors in this system. No modification is needed on database design. However, data files loading software has to be adapted for every other instrument, since files structures are not exactly the same.

References

[1] Moldwin M. (2008), "An Introduction to Space Weather", Cambridge University Press, First Edition.

[2] Frigo, E., Savian, J. F., Silva, M. R., Dal Lago A., Trivedi, N. B., Schuch, N. J. Analysis of geomagnetic storm variations and count-rate of cosmic ray muons recorded at the Brazilian southern space observatory, Rev. Bras. Geof. vol.25 suppl.2 São Paulo, 2007.

[3] Kilts, S. Advanced FPGA Design: Architecture, Implementation, and Optimization, John Wiley & Sons, 2007.

[4] Shinshu University (2011), available at http://www.shinshu-u.ac.jp/, access on March 2011.

[5] PostgreSQL Group (2011), available at http://www.postgres.org, access on March 2011.

[6] Hibernate Relational Persistence for Java e .NET (2011), available at http://www.hibernate.org/, access on March 2011.

[7] Munakata, K., Bieber, J. W., Yasue, S., Kato, C., Koyama, M., Akahane, S., Fujimoto, K., Fujii, Z., Humble, J., Duldig, M. L., Precursors of geomagnetic storms observed by the muon detector network, Journal of Geophysical Research, v. 105, n. A12, December 2000.

[8] Yasue, S., Munakata, K., Kato, C., Kuwabara, T., Akahane, S., Koyama, M., Fujii, Z., Evenson, P., Bieber, J. W., Design of a recording system for a muon telescope using FPGA and VHDL, The 28th International Cosmic Ray Conference, 2003.

Acknowledgments