A MySQL BASED EPICS ARCHIVER*

Christopher J. Slominski, TJNAF, Newport News, Virginia, U.S.A.

Abstract

Archiving a large fraction of the EPICS signals within the Jefferson Lab (JLAB) Accelerator control system is vital for postmortem and real-time analysis of the accelerator performance. This analysis is performed on a daily basis by scientists, operators, engineers, technicians, and software developers. Archiving poses unique challenges due to the magnitude of the control system. A MySQL Archiving system (Mya) was developed to scale to the needs of the control system; currently archiving 58,000 EPICS variables, updating at a rate of 11,000 events per second. In addition to the large collection rate, retrieval of the archived data must also be fast and robust. Archived data retrieval clients obtain data at a rate over 100,000 data points per second. Managing the data in a relational database provides a number of benefits. This paper describes an archiving solution that uses an open source database and standard off the shelf hardware to reach high performance archiving needs. Mya has been in production at Jefferson Lab since February of 2007.

MYA OVERVIEW

One deployment of a Mya system consists of one or more X86 host computers running the Linux operating system. Mya was designed to scale to the magnitude of the EPICS control system load by adding additional host computers. The Mya administrators choose a strategy of combining a small number of high performance hosts with large storage capacity, versus a larger number of low performance hosts with smaller storage capacity.

The figure below shows a Mya deployment, having three host computers. Each host has a MySQL database server, Mya software, and database tables on local storage. One host in a deployment must be designated the master host, having additional database tables which serve as a catalogue of the EPICS channels being archived by the Mya deployment. Each instance of Mya in a deployment is responsible for archiving its own subset of the control system channels that are assigned to it. It learns its assigned channels from the Mya master instance. Each of these three components of a Mya host will be briefly described in the following sections.

Mya Software

The Mya software is a single process running on the host computer. It is written in C++ to take advantage of both the speed of a compiled language and the maintainability of an object oriented design. The process creates a number of threads to fully utilize hardware capabilities. These threads run at POSIX real-time scheduling priorities to further enhance performance. Mya interacts with the database server using pre-parsed binary SQL statements to maximize database throughput.

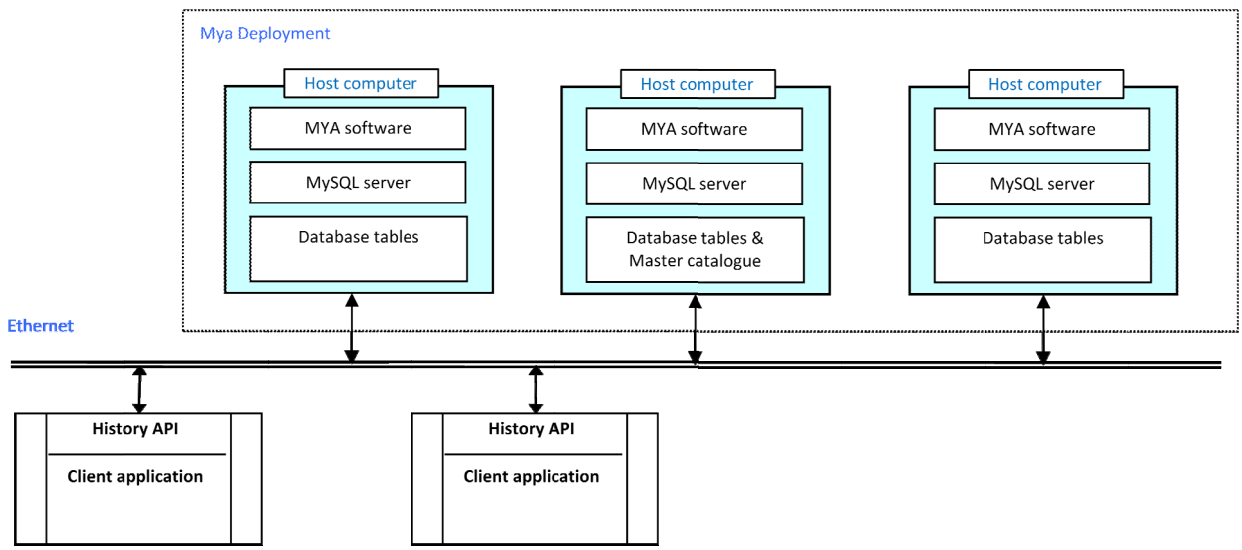

Figure 1: A three host Mya deployment with two clients accessing control system history.

The Mya software monitors the host's remaining disk capacity, automatically deleting old control system history when the available disk space becomes low. This is done to make a deployment self sufficient, not requiring administrative intervention over time.

The Mya control system history set is protected from hardware failure via a built-in backup feature. Each day Mya writes the previous day's control system updates to local files in Comma Separated Value (CSV) ASCII format. The contents of the directory are pulled away daily by an external process, which compresses the data and writes it to tape.

Clients of control system history access data using a well defined API. The API is available in the C++, Tcl, and Perl programming languages.

MySQL Server

EPICS control system history is stored in MySQL database tables. Version 5.1 or above of the server is required, as Mya takes advantage of the *table partitioning* features of MySQL, which were not available in earlier versions. Other special features of MySQL utilized by Mya include *stored procedures* and *prepared statements*. Control system history is stored in high performance *MyISAM* format tables, while metadata and control system catalogue information are stored in transaction safe *InnoDB* format tables.

Database Tables

The MySQL database tables are placed in host local storage. Each host in a Mya deployment stores the updates to its assigned channels into its local disk storage.

Deployments requiring high performance and reliability use a hardware RAID disk array configured for both striping and mirroring of data. Tests have shown that large I/O performance can be achieved with Mya when using the JFS file system as opposed to the standard Linux Ext3 file system.

SYSTEM INVENTORY

High performance archiving of an EPICS control system can be attained at a relatively low cost. This includes initial hardware investments as well as long term administrative maintenance.

When archiving needs are minimal, a very basic computer system can serve as a Mya host. The deployment can be a single host configuration and the host does not even need to be dedicated to the Mya software. Mya has run fine in the background of a simple personal workstation, archiving a few thousand control system channels with an aggregate update rate less than 1,000 per second.

To handle a large archiving load, one or more dedicated hosts may be needed, using standard off the shelf server architectures. This section describes a single Mya host deployment used at Jefferson Lab. It provides a hardware inventory, performance characteristics, and cost. Note that the trend in performance versus cost of commercial off the shelf hardware has been very favourable for a number of years, and you will likely be able to deploy a more capable Mya system than described here, for less cost.

Performance

The Mya host described in this section can sponsor 60,000 control system channels, with an aggregate update rate of up to 10,000 channel updates per second. The two 1.4TB disk partitions provide one year of history given a nominal 3,500 updates per second from the control system. It can perform at this level with enough headroom for performing its maintenance and administrative activities, as well as serving up channel history to clients.

Maintenance activities that occur simultaneously with archiving include the following:

- Adding new channels to the archive channel set.
- Monitoring disk usage to purge old history when the disk nears depletion.
- Performing incremental daily backups.
- Clipping old history of channels being archived with a limited keep span.
- Expiration of channels being archived for a limited duration.
- Recording changes to channel metadata, such as enumeration strings.

History clients open connections to the Mya channel history database at a rate of about one per minute throughout each day, with typical fetches of approximately 150,000 data values.

Hardware Overview

The example MYA unit is comprised of a Dell PowerEdge 1950 Server and a Dell PowerVault Disk Enclosure costing under $30,000. This generation of hardware is roughly half the cost of the previous generation and one can expect future generations of hardware to be reduced in cost even more.

PowerEdge 1950 ($6,000)

- 1U Rack Server configured with an internal RAID card for the system disks and an internal SAS card for the attached SAS RAID enclosure.
- Dual Quad Core Intel® Xeon® E5420, 2x6MB Cache, 2.5GHz, 1333MHz FSB
- 16GB 667MHz (4x2GB), Dual Ranked DIMMs memory.
- Two 146GB 15K RPM Serial-Attach SCSI 3Gbps disks.

PowerVault Enclosure ($23,000)

The main storage disk array is a 45 slot Serial-Attach SCSI (SAS) disk enclosure with two independent but redundant RAID controllers. The disks are laid out in two logical volumes using RAID 10 (stripe and mirror). Each of the two logical volumes has 22 disks (11 effective spindles) and provides around 1.4 TB of space.

- One PowerVault MD3000 controller Configured with two single-port controllers and 15 disk slots.
- Two PowerVault MD1000 Expansion Units each with 15 disk slots each.
- 45-146GB 15K RPM Serial-Attach SCSI 3Gbps 3.5-in HotPlug Hard Drives.

CONCLUSION

High performance and high reliability archiving of an EPICS control system can be attained with minimal investment in cost and man power, using standard off the shelf hardware and Open Source software.

REFERENCES

[1] C. Slominski, "EPICS Channel Archive Facility, Software Requirements Specification", March 2006, cjs@jlab.org.
[2] C. Slominski, "EPICS Channel Archive Facility, Design Specification", June 2006, cjs@jlab.org.
[3] C. Slominski, "Mya Administrator's Guide", June 2009, cjs@jlab.org.