GA–A26485

CUSTOMIZABLE SCIENTIFIC WEB PORTAL FOR FUSION RESEARCH

by
G. ABLA, E.N. KIM, D.P. SCHISSEL, and S.M. FLANAGAN

JULY 2009

GENERAL ATOMICS

DISCLAIMER

GA–A26485

CUSTOMIZABLE SCIENTIFIC WEB PORTAL FOR FUSION RESEARCH

by
G. ABLA, E.N. KIM, D.P. SCHISSEL, and S.M. FLANAGAN

This is a preprint of an invited paper to be presented at the Seventh IAEA Technical Meeting on Control, Data Acquisition, and Remote Participation in Fusion Research, June 15–19, 2009, Aix-en-Provence, France, and to be published in the *Proceedings.*

Work supported by
the U.S. Department of Energy SciDAC Program
under DE-FC02-01ER25455

GENERAL ATOMICS PROJECT 30106
JULY 2009

GENERAL ATOMICS

Abstract

Web browsers have become a major application interface for participating in scientific experiments such as those in magnetic fusion. The recent advances in web technologies motivated the deployment of interactive web applications with rich features. In the scientific world, web applications have been deployed in portal environments. When used in a scientific research environment, such as fusion experiments, web portals can present diverse sources of information in a unified interface. However, the design and development of a scientific web portal has its own challenges. One such challenge is that a web portal needs to be fast and interactive despite the high volume of information and number of tools it presents. Another challenge is that the visual output of the web portal must not be overwhelming to the end users, despite the high volume of data generated by fusion experiments. Therefore, the applications and information should be customizable depending on the needs of end users. In order to meet these challenges, the design and implementation of a web portal needs to support high interactivity and user customization.

A web portal has been designed to support the experimental activities of DIII-D researchers worldwide by providing multiple services, such as real-time experiment status monitoring, diagnostic data access and interactive data visualization. The web portal also supports interactive collaborations by providing a collaborative logbook, shared visualization and online instant messaging services. The portal's design utilizes the multi-tier software architecture and has been implemented utilizing web 2.0 technologies, such as AJAX, Django, and Memcached, to develop a highly-interactive and customizable user interface. It offers a customizable interface with personalized page layouts and list of services, which allows users to create a unique, personalized working environment to fit their own needs and interests.

This paper describes the software architecture of this scientific web portal and its implementation to include deployment experiences during the 2009 DIII-D Experimental Campaign.

1. Introduction

Fusion experiments are a complex research activity due to two main reasons. The first reason is the enormous volume of data and information collected from the experiment. The research team, which includes physicists, operators and engineers, needs to analyze this large amount of data generated from each experiment within a short amount of time. The second reason is the real-time collaboration need of the distributed research team. The team members, each playing a unique role in the team's common experimental goal, tightly collaborate while they analyze the data and make decisions leading to parameter selection for the next experiment. For example, at DIII-D, each experiment pulse generates a total of about 6 GB of data, which is received from about 100,000 different diagnostic measurements and computational codes. The day's research team, typically consisting of about 30 physicists, operators and engineers, must analyze the last experiment's results and make decisions for the next experiment within a 30 minute window. In order to increase the efficacy of experimental run time, experimental data and machine status information need to be available immediately and reliable and effective communication mechanisms must be in place among team members.

There are many data analysis software and communication tools available that can be used during the fusion experiments. However, each of these tools individually address data analysis, visualization, or collaboration needs. The majority of data analysis tools are very specialized, and collaboration and communication tools are often developed as stand-alone and specialized software tools. Therefore, scientists are required to run many standalone applications in order to monitor experiment status, analyze data and communicate with their colleagues synchronously. The user experience can be very inefficient due to the need to switch among multiple application windows, perhaps even computers, when they analyze data while collaborating with their colleagues. A comprehensive tool will vastly increase the capabilities of an experimental team.

A Portal, one application program that can provide all the needed functionalities, also provides a unified interface for multiple capabilities. Developing multiple functionalities under one portal framework makes it possible to tightly integrate status monitoring, data analysis and collaboration. Collaboration capability can also be built into monitoring and data analysis applications in the portal. However, portals are not free from issues. If not carefully architected, the amount of visual output can become overwhelming due to the high volume of information generated by data analysis and collaboration tools. If too much information is presented in one screen, the usability becomes limited from a visually cluttered interface. This problem can be solved by making the portal customizable, so that individual users can choose the functional component, data, and layout based on their specific needs and interests.

A web interface for the portal development is appealing since the application runs on the web browser and there is no need for downloading software or installing it on the user's computer. The deployment of new features and updates to users is relatively effortless. Historically, developing data analysis applications on the web has been limited, since web browsers were not able to provide the highly interactive and customizable user interfaces that were available on desktop applications. Recent advances in web technologies, especially the group of interrelated web development techniques named AJAX [1], have paved the way to mimic feature-rich desktop applications on the web interface. The "*Web 2.0"*, a new group of web development tools and frameworks, offers many capabilities for web-based communication, secure information sharing, interoperability, and collaboration, as well as the foundation of customizable web page design [2].

At DIII-D, there have been efforts to provide a unified software environment for real-time experiment status monitoring, interactive data analysis, and integrated collaboration. The DIII-D web portal is the continuation of this effort, and it provides a centralized gateway for multiple software services that are needed during fusion experiments. The web portal design can be customized by the end users in multiple ways, including the layout, the number and position of the applications displayed, as well as the information filter in each application.

2. Multi-tier Software Architecture of the DIII-D Web Portal

The purpose of the DIII-D web portal is to provide researchers with a unified interface for multiple capabilities such as experiment status monitoring, data analysis and communication. Although the web portal is a rather new technology, almost all aspects of the web portal need to communicate with the existing computational infrastructure, which includes data repositories, relational databases, a plasma control system, and data analysis codes. Therefore, the design of the web portal needs to be compatible with the existing infrastructure.

The web portal is a multi-tier architecture software consisting of the following three-tiers: (1) Data tier, (2) Logic tier, and (3) Presentation tier (Fig. 1).

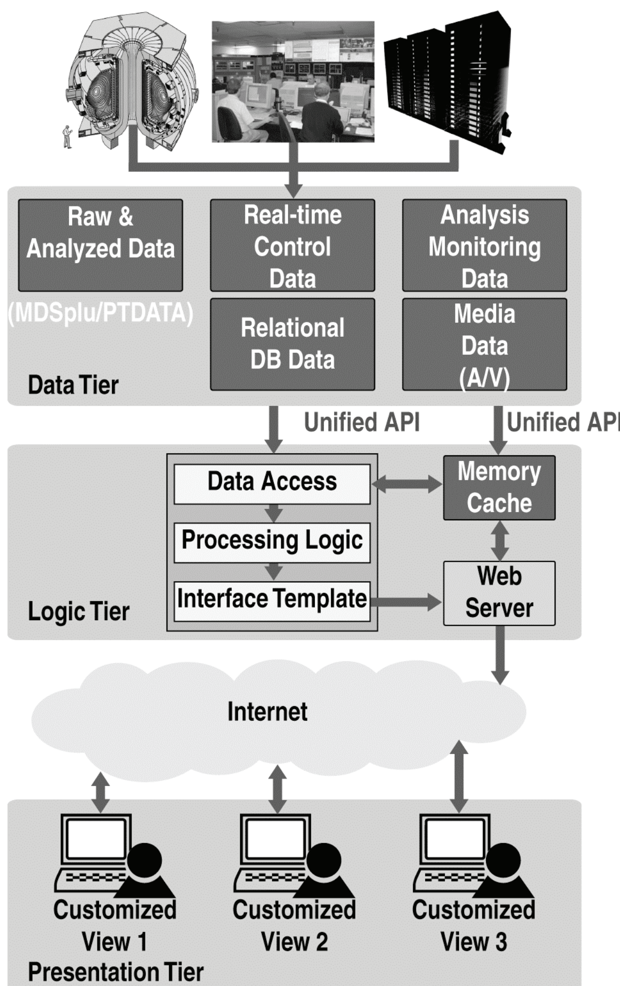

Figure 1. Multi-tier software architecture of DIII-D web-portal.

2.1 Data Tier

The Data tier provides interfaces to access a variety of data at DIII-D. There are mainly two groups of data sources. The first group of data sources consists of data servers, which include MDSplus repository and relational databases. They provide diagnostic raw data, analyzed data and meta-data. The second group of data sources consists of the tokamak control and monitoring systems which are not specialized data sources, but provide information about real time machine status, between-pulse computational data analysis status, and experiment related group activity in the control room. For example, while the real-time control servers provide hardware control requests, results and tokamak status, the media servers stream audio/video data from control room operations. The access of this group of data needs to be as fast as possible.

2.2 Logic Tier

The Logic tier processes the data acquired from the Data tier. According to the needs of each of the functionalities of the web portal, the process can be very different. For example, while the signal plotting functionality requires the generation of diagnostic data graphs, the equilibrium fitting visualization functionality requires the generation of a movie file for each pulse. The end result of each Logic tier action is to create the dynamic web page content that is needed by the presentation layer.

The Logic tier also maintains a dual purpose memory cache. The first purpose is caching a variety of data in order to minimize the number of requests to non-real-time data sources. The second purpose is for fast data sharing between multiple data sources and Logic tier components (Fig. 2).

The Logic tier utilizes Django, a python-based web framework which is based on loose Model-View-Controller design architecture [3,4].

2.3 Presentation Tier

The Presentation tier is responsible for interacting with users. It mainly receives the user requests, passes them to the Logic tier, and displays the results of Logic tier actions to the end users. The requests from the Presentation tier to the Logic tier are simply described with HTTPS GET and HTTPS POST methods. There are three types of requests: (1) Data posting, and viewing requests, (2) control requests, and (3) web-portal interface customization requests.

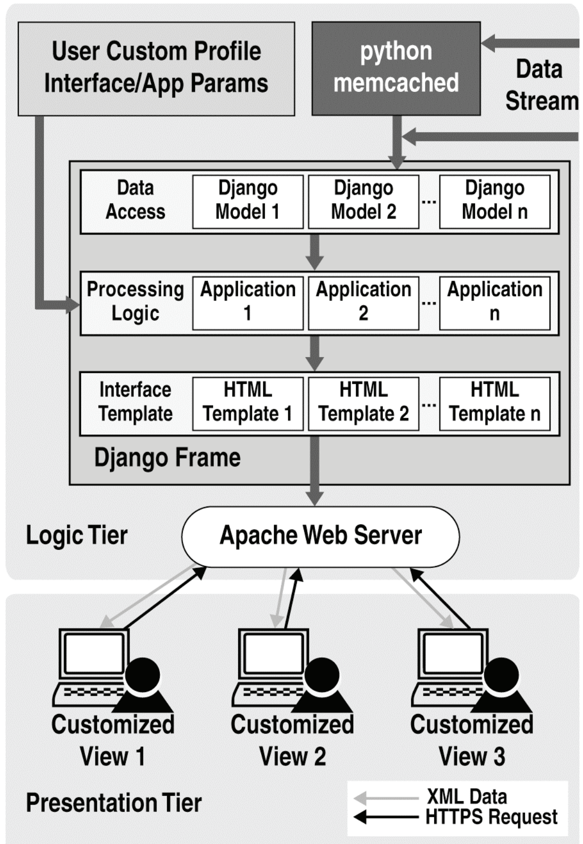

Figure 2. Detailed logic tier and presentation tier.

The Presentation tier is intentionally simple and lightweight. It doesn't perform computational tasks but simply filters and displays the Logic tier generated visual data according to user requests. It is also designed in a way that multiple presentation methods can be supported synchronously.

The customization capability is a function of the Presentation tier. AJAX technology provides a unique and efficient way to observe all changes made on the client-side, and record them from the server-side. These customization options include turning applications on/off, changing locations or updating parameters on a specific application. A list of all available applications and services are provided as links to independently turn them on or off. Once enabled, an application can easily be moved around the page by drag-and-drop (Fig. 3). When the location is redefined, it gets calculated and updated on the server-side. Some applications also offer their own customization which is discussed further in section 3. When a change is made on the client-side, JavaScript makes a HTTPS POST request to update the "User Custom Profile" database. This database includes multiple tables which share a unique key defined by Apache's authentication credentials. The web portal is password protected by Apache's authentication system with SSL. Each user is assigned with a username in order to identify oneself when using the web portal. This tightens the access security to the web portal, and also helps the system recognize the logged-in individuals in order to map each user to their "User Custom Profile". Once a user is authenticated, the web portal fetches the user customization information associated with the unique username from the "User Custom Profile" database, and dynamically draws the content of the website. Since all information in the "User Custom Profile" is stored in the database, the information is kept even after the client session is over and recalled the next time the user logs in.

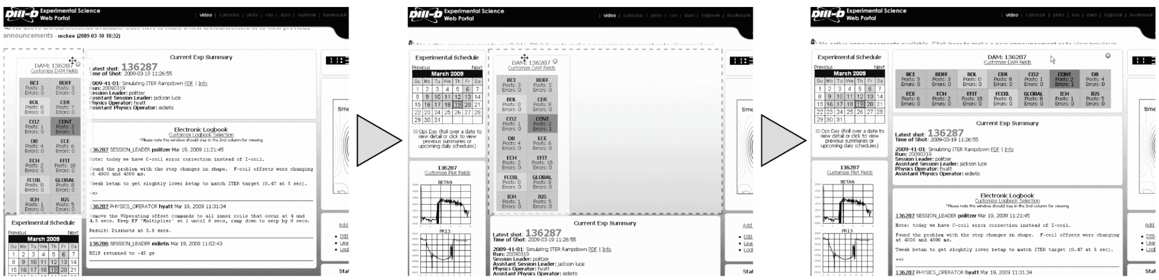

Figure 3. Application can easily be moved with drag-and-drop.

3. DIII-D Web Portal Environment

The web portal includes a number of applications and services to cover a wide variety of aspects of the DIII-D research program. The initial release includes 11 applications:

1. *Current Experiment Information* provides a summary of the experiment including information about the experimental team and the details of the proposed experiment.
2. *Data Analysis Monitoring* displays results of analysis codes in real-time.
3. *EFIT Animation* outputs an animation of the most recently calculated plasma shape data.
4. *Electronic Logbook* offers a way of organizing notes for the experimental team
5. *Experimental Schedule* provides a detailed overview of all scheduled experiments in a calendar format
6. *Signal Plots* outputs the latest plots of various measured and analyzed data.
7. *Tokamak Status Information* displays pulse information such as the status, the type and the configurations for the pulse.
8. *Video* provides a real-time view of the DIII-D control room.
9. *Bookmarks* provides a way for users to save websites.
10. *Clock* displays the current local time where the DIII-D experiments are held.
11. *Instant Global Announcement* makes announcements available to all web-portal users.

Although all applications are informative and relevant to the experiments, the order of priority and interest level varies among users. To fit the needs and interests of each user, the web portal offers three layers of customization (Fig. 4). It allows users to enable/disable an application, update the placement of the application, and certain applications allow their own customization, which manipulates the output within the application.

The fact that web portal receives large amounts of data from many different sources remains transparent to the user. The main goal is to display this abundant information in an effortless manner on the client-side. To hide the complexity from the users, each application uses source-specific methods to receive and process the data as required.

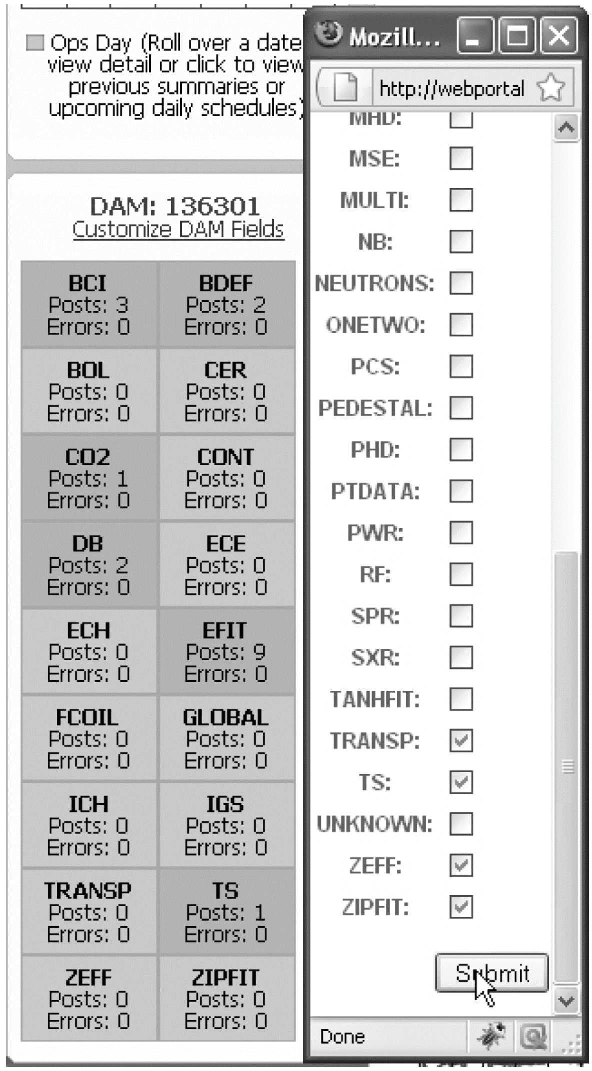

Figure 4. Data analysis monitoring application customization.

Experimental Schedule Information collects data from multiple databases in order to produce a calendar view of all experiments. The processing logic in the Logic tier utilizes Python's pymssql class to connect, execute queries and fetch the information from the MSSQL databases. Once the information is received, it parses and processes the data into an HTML table. It configures each cell with a background color for operating days, formats details as a tooltip, and attaches a link to summaries of previous experiments or detailed schedule information on upcoming ones. This allows users to easily view all the needed schedule data in a small window just over 2 inches by 3 inches.

The Data Analysis Monitor (DAM) is a monitoring system used to track the status and detect discrepancies of automated analysis codes [6]. Existing systems such as DAM can easily be configured into the web portal as long as the needed data is passed in a recognizable scheme. As DAM receives data and logs the information to a relational database, it also passes the values to Memcached in the web portal's Logic tier. Memcached is a high-performance, distributed memory object caching system which helps speed up dynamic web applications by alleviating database load [5]. The data access fetches the values from Memcached to the processing logic which creates the color coordinated HTML-based output. Since DAM has the ability to monitor hundreds of analysis codes which are categorized into around 40 unique code and diagnostic groups, the output can be overwhelming and may not be suitable for all users. Therefore the web portal application offers its own customization which provides a list of all available data analysis codes for users to select from.

4. Initial Deployment Experiences and Future Work

The DIII-D web-portal was deployed during the 2009 DIII-D experimental campaign and has been used for over four months. More than 60 scientists have viewed the web portal with approximately 35 using it on a routine basis to monitor the ongoing experiment activities. The number of web portal users not in the DIII-D control room exceeded the number of those in the control room.

To continue the idea for effortless collaboration by the means of the web, the web portal was tested to work with external web services. Since the web portal is a collection of information, it was a simple setup to provide portions of the data as a web service. This was successfully tested with the micro-blogging web service, *Twitter*. *Twitter* allows its users to send and read other users' updates. These updates are text-based posts of up to 140 characters, displayed on the user's profile page and delivered to other users who have subscribed to them. The subscribers can receive these status updates via the *Twitter* website, SMS or external applications [8]. The web portal's Instant Global Announcement which provides users to input short messages to all users' web portal interfaces was a great fit to be tested with *Twitter*. A check box option was added to the interface for submitting a new announcement message. When this option was enabled, the announcement was immediately shared by the all users logged into the web portal as well as all *Twitter* followers of DIII-D.

Because the web portal tightly integrates multiple capabilities in a simple manner and opens up room for growth on the server-side, while offering an effortless setup on the client-side, the desire is to push for greater and broader use. It has already attracted our collaborators from KSTAR and it is planned to be developed to be used for their operations.

During the initial deployment period, users have already requested new features and provided ideas for extending the web portal. As the next step, more applications such as instant messaging and visualization sharing capability will be added. A finer grained customization of portal layout also will be implemented as requested.

Another plan for the DIII-D web-portal is to improve its efficiency and scalability. Currently, the DIII-D web-portal uses a web browser's "client pull" mechanism, not the web server's "server push", to make data updates on most of the applications. This can be very inefficient as the number of users logging into the web-portal increases. Therefore, the "server push" mechanism will be implemented and the unnecessary network and serve CPU load will be minimized. As the number applications on the web-portal increase distribution of the Logic tier computational load will be considered.

Template management will also be improved. Currently, the web-portal mainly targets web browsers as clients. However, the template system can be more versatile and it is possible to dynamically create presentation templates for clients other than web browsers. One such client is as handheld device such as smart phones. Recently handheld devices have become very versatile and it is certainly possible to develop client applications and interfaces which can be utilized on these platforms. Another client that will be considered is large public displays. The large US fusion experimental control rooms have display walls that use homegrown applications to display visual data from experiments [7]. While it is possible to display web pages on the display walls, the optimum color and size for visual information as well as the interaction method on the display walls are different from normal desktop screens or smaller devices. The web-portal on the display wall will be investigated. Taken together, the web-portal can serve multiple clients with minimal development.

5. Summary

Recent developments of web technologies made it possible to build rich and interactive data analysis and collaboration applications in the Web. When these applications integrated in a Web portal environment, they provide uniform user interface for multiple capabilities. The customization capability of web portal highly increases the usability of the web portal by tailoring the functionalities and the presented data to the needs of the individual users.

The DIII-D web portal is designed and deployed by keeping the above principles in mind. It provides multiple data analysis, experiment status monitoring and collaboration capabilities via a uniform user interface. It supports several levels of customization in order to tailor the services for Scientists' needs and to create more effective user experience. The portals modular design should allow it to be relatively easily adopted by other experiments and this will be tested with an initial external deployment to support KSTAR experimental operations.

References

[1] What is AJAX?, http://www.riaspot.com/articles/entry/What-is-Ajax-

[2] What is Web 2.0, http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web20.html

[3] Django project web site, http://www.djangoproject.com, visited on May 13, 2009.

[4] Trygve M H Reenskaug, MVC XEROX PARC 1978-79, http://heim.ifi.uio.no/~trygver/themes/mvc/mvc-index.html, visited on May 13, 2009.

[5] Memcached, http://www.danga.com/memcached/, visited on January May 13, 2009.

[6] Flanagan S, *et al.,* 2004 System to monitor data analyses and results of physics data validation between pulses at DIII–D, Rev. Sci. Instrum. **75** 2043; DOI:10.1063/1.1753668, May 2004.

[7] Abla G, *et al*. 2005 Shared display wall based collaboration environment in the control room of the DIII-D National Fusion Facility, published in the Fifth Annual Workshop on Advanced Collaborative Environments.

[8] Twitter web site, http://www.twitter.com, visited on May 13, 2009.

Acknowledgment

This work was supported by the US Department of Energy SciDAC program under Cooperative Agreement DE-FC02-01ER25455.