

# Data Management along with Post analysis and Data Retrieval Facility for Negative Neutral Beam Development: A Survey

<sup>1</sup> Aakansha Saxena, <sup>2</sup> Himanshu Tyagi, <sup>3</sup> Prof. Krunal Panchal

<sup>1</sup> Student of Gujarat Technological University,

<sup>2</sup> Scientist-SC at Institute for Plasma Research,

<sup>3</sup> Assistant Professor at L. J. Institute Of Engineering & Technology,

<sup>1</sup> Department of Computer Engineering,

<sup>1</sup> L. J. Institute of Engineering and Technology, Gujarat Technological University, Ahmedabad, Gujarat, India

**Abstract** - NNBD(Negative neutral beam development) is a research community in IPR(Institute for Plasma Research) involved in development of high energy neutral beams for heating the plasma in order to achieve fusion temperatures in magnetic confinement reactors called as tokamak. It includes three research set ups- 1] Robin 2] Twin source 3] INTF. These systems are operated by indigenously developed Data Acquisition and Control system (DACS). These facilities operate in pulsed mode in which they operate for a fixed duration. The data acquired for each of this experimental pulses (or shots) is stored separately. A huge amount of database will be generated from these facilities which are an important asset for the research community. The main aim is to develop a prototype facility with a mature user interface which facilitates post data analysis, data visualization and data retrieval as per user requirements along with secure web based data access facility.

**Index Terms** - NNBD, DAMS, tokamak, jwebscope, XINETD, Office of research integrity, Network attached storage, HDF5, IPR(Institute for plasma research).

## I. INTRODUCTION

For addressing the R& D activities in area of Negative Neutral Beam Injector, in IPR, Negative Neutral Beam development program has been started. As a part of the program, three test-beds as follows, [1] ROBIN, [2] Twin Source (TS) and [3] Indian Test Facility (IN-TF) has been planned. Out of the three, the ROBIN- signal RF driver based 100 KW, 1 MHz negative ion source test-bed is already commissioned and in operational stage. Now Twin Source- an inductively coupled two RF driver based 180 KW, 1 MHz negative ion source experimental setup is initiated with the objective of understanding the physics and technology of multi-driver coupling. Twin Source (TS) also provides an intermediate platform between operational ROBIN and eight RF drivers based Indian test facility IN-TF. The IN-TF facility aims to make significant contributions to the physics and technology of upcoming NB systems for reactors. It is foreseen that continuation of a facility of the kind proposed would bring the Indian NB program to a focus with the participation of the international community. The data acquired for each of this experimental pulses (or shots) is stored separately. A huge amount of database will be generated from these facilities which are an important asset for the research community.

Individual each database is to be taken and develop a prototype facility with a mature user interface which facilitates post data analysis. Apart from that it is necessary to decide an appropriate database for storing the generated data and then develop program for data retrieval as per user requirements. This facility must also include a secure web based data access facility to facilitate the data access to any other experimentalist from remote location for analysis purpose.

The rest of this paper is organized as follows. Section II describes the background theory. Section III gives the detail of literature survey and some action plan based on survey. Finally, some conclusions are presented in section IV.

## II. BACKGROUND THEORY

Database Access and management system is a computer software application that interacts with the user, other applications and the database itself to capture and analyse data. A general purpose DAMS is designed to allow the definition, creation, querying, update and administration of databases.

It is one of the core areas addressed by the Office of Research Integrity (ORI) in its responsible conduct of research initiative. It represents a significant investment of time and conduct of research effort by the Principal Investigator (PI) of a research project. these issues related to data management:-

- **Data Ownership** : This pertains to who has the legal rights to the data and who retains the data after the project is completed, including the PI's right to transfer data between institutions.
- **Data Collection**: This pertains to collecting project data in a consistent, systematic manner (i.e., reliability) and establishing an ongoing system for evaluating and recording changes to the project protocol (i.e., validity).
- **Data Storage**: This concerns the amount of data that should be stored -- enough so that project results can be reconstructed.

- **Data Protection:** This relates to protecting written and electronic data from physical damage and protecting data integrity, including damage from tampering or theft.
- **Data Retention:** This refers to the length of time one needs to keep the project data according to the sponsor's or funder's guidelines. It also includes secure destruction of data.
- **Data Analysis:** This pertains to how raw data are chosen, evaluated, and interpreted into meaningful and significant conclusions that other researchers and the public can understand and use.
- **Data Sharing:** This concerns how project data and research results are disseminated to other researchers and the general public, and when data should not be shared.
- **Data Reporting:** This pertains to the publication of conclusive findings, both positive and negative, after the project is completed.

A DAMS consists of the following three elements:

- **Physical database:** the collection of files that contain the data
- **Database engine:** the software that makes it possible to access and modify the contents of the database
- **Database scheme:** the specification of the logical structure of the data stored in the database

### Functions of a DAMS

- A DAMS provides automated methods to create, store and retrieve data
- A DAMS reduces data redundancy and inconsistency.
- A DAMS allows for concurrent access by multiple users, each with their own specific role.
- A DAMS increases security and reliability.
- Backup and recovery: processes to back-up the data regularly and recover data if a problem occurs.
- Integrity: database structure and rules improve the integrity of the data.

### III. LITERATURE SURVEY

Since the system that is to be implemented is totally a new system because the test beds are in developing phase. There are other architecture like Compass Database system, MAST's Integrated Data Access Management system: IDAM, with reference to them the system will be developed. The findings from the research paper are given below-

Sr. No.	Paper Tittle	Technologies Used	Advantages	Disadvantages
1.	Integrated data acquisition, storage, retrieval and processing using the COMPASS Database (CDB).	Relational Database: MYSQL, Fileformat: HDF5, OS: Centos Processor: Intel(R) Xeon(R) CPU E5506 @ 2.13 GHz Core functionality: Python	Low overhead, open, fast and easily extensible and scalable. Data consistency & persistency.	Less Secure.
2.	MAST's Integrated Data Access Management system: IDAM.	XML-SQL-relational database,PostgreSQL File format: legacy and modern OS: Linux, 32 bit dual processor 2 Ghz Fedora Core. Core functionality: C/C++	Record data quality values at the signal level, excellent level of performance, Data in legacy and modern formats. Local and remote data sources.	Network overhead was larger at 100–150 ms.
3.	A web based MDSplus data analysis and visualization system for EAST.	Data visualization: Webscope Data storage and access system: MDSplus in MYSQL database. Core functionality: Java	Faster subsampled remote access due to slice storage mechanism, Higher performance as compared to eastscope, jscope	No metadata search function. To access and visualize the date, Researchers have to access MDSplus database by jScope-setting up and configuration difficult.
4.	The upgrade of the J-TEXT experimental data access and management system.	Data storage and access: MDSplus Data visualization: jWebscope along with jscope configuration file Core functionality: ASP .Net	Automatic segmentation reading strategy. No need of setting up configuration at client side for remote access.	No automatic post analysis and preprocessing task facility available.
5.	HIDS: A new web-based data access system.	Data access: via RESTFUL web service with MDSplus as backend. Core functionality: Python	Faster response time and decreased load on the primary data server Shared with collaborators	No proper treatment of data processing filters as resources, allowing authorised users to upload

			via a URL . The Django REST Framework provides serialisation of data and parsing for the requested content type.	custom data filters.
--	--	--	---	----------------------

**IV. ACTION PLAN**

A complex data handling system for the three testbeds Robin, Twin source and INTF operated by IPR is to be developed. The system, called DAMS(Database access and management system), integrates different data sources as an assortment of data acquisition hardware and software from different vendors is used. DAMS is vendor and platform independent and it can be easily scaled and distributed. The data is directly stored and retrieved using a standard NAS(Network Attached Storage), hence independent of the particular technology; the description of the data(the metadata) is recorded in a relational database. Database structure is general and enables the inclusion of multi-dimensional data signals in multiple revisions (no data is overwritten). This design is inherently distributed as the work is off-loaded to the clients. Both NAS and database can be implemented and optimized for fast local access as well as secure remote access.

This has to be implemented in Python language; bindings for Java, C/C++, IDL and Matlab are provided. An automated data post-processing server is a part of CDB. Based on dependency rules, the server executes, in parallel if possible, prescribed post-processing tasks. When New version of signal is stored or error: No overwriting, new revision is stored so full history is saved that assures consistency and persistency.

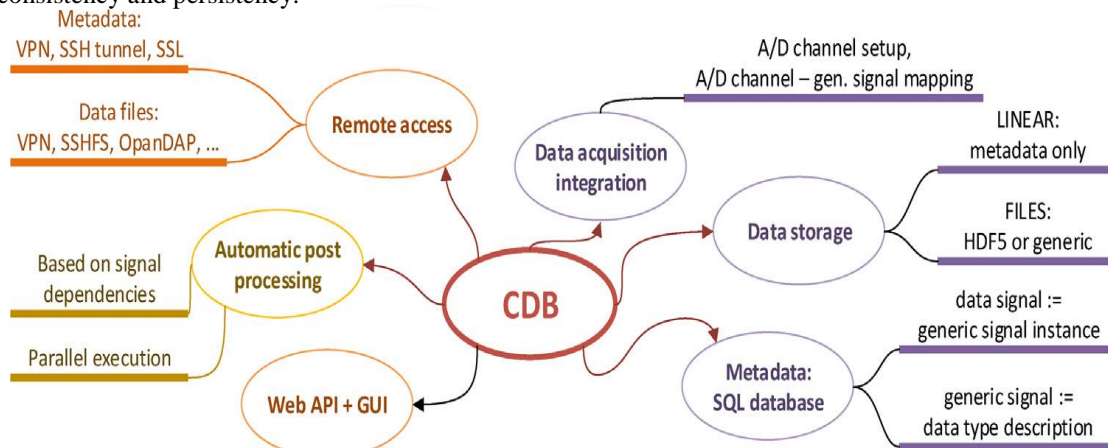


Fig: An Overview Diagram of DAMS components

**V. CONCLUSION**

The proposed paper briefly discussed a critical survey of existing literature on Database access and management system. This paper provides the concepts of DAMS along with its working and the tools and technologies that are to be used in developing a DAMS in a tabular format.

**VI. REFERENCES**

- [1] J.Urban, J.Pipek, M.Hron, F.Janky, R.Paprok, M.Peterka, A.S. Duarte “Integrated data acquisition, storage, retrieval and processing using the COMPASS Database (CDB)” Fusion Engineering and Design 89 (2014)PP: 712–716 ©Elsevier B.V.
- [2] F.Yang, B.J.Xiao “A web based MDSplus data analysis and visualization system for EAST” Fusion Engineering and Design 87 (2012)PP: 2161– 2165 ©Elsevier B.V.
- [3] C.Yang, M.Zhang, W.Zheng, R.Liu, G.Zhuang “The upgrade of the J-TEXT experimental data access and management system” Fusion Engineering and Design 89 (2014)PP: 726–730 ©Elsevier B.V
- [4] K.Behler, H.Blank, A.Buhler, R.Drube, H.Eixenberger, K.Engelhardt, A.Lohs, R.Merkel, G.Raupp, W.Treutterer “Update on the ASDEX Upgrade data acquisition and data management environment” Fusion Engineering and Design 89 (2014) PP: 702–706 ©Elsevier B.V.
- [5] D.G.Pretty, B.D.Blackwell, the H-1 Team “H1DS: A new web-based data access system” Fusion Engineering and Design 89 (2014)PP: 731–735 ©Elsevier B.V.
- [6] D.G. Muir, L.Appel, N.J.Conway, A.Kirk, R.Martin, H. Meyer, J.Storrs, D.Taylor, N.Thomas-Davies, J.Waterhouse “MAST’s Integrated Data Access Management system: IDAM” Fusion Engineering and Design 83 (2008)PP: 406–409 ©Elsevier B.V.