DESIGNING AN INTEGRATED LIBRARY MANAGEMENT AND RETRIEVAL SYSTEM USING OPEN SOURCE TOOLS

Dr. Sukumar Mandal*

Sushanta Kumar Roy**

Abhijit Chakrabarti***

Abstract

The paper discusses the application of open source tools for the libraries in different cluster on Ubuntu operating system. Open source tools are essential in integrated library management and retrieval system for housekeeping operations and information retrieval system. This research work explores the twelve innovative open source tools on the basis of global recommendations and local requirements in different college libraries. How to use these tools in the college libraries both for the data management and backup and restoration. However, there are many tools are available in the open source environment but this research work select the most comprehensive in the level authority and bibliographic on different databases. These open source tools are to be easily managed the digital and library resources in the libraries.

Keywords: Open source tools, Integrated Library System, and Integration

* **Assistant Professor, Department of Library and Information Science, The University of Burdwan**

** **Library and Information Assistant, Indian Institute of Science Education and Research Kolkata**

*** **Librarian, Banwarilal Bhalotia college, Asansol**

1.0   Introduction

Open source is more important for college libraries. Open source tools is one of the important concept in automated and digital library system. All the tools are to be access from the web to manage and maintain the library resources. It is source code available in online environment. Backup and restoration is also possible from these open source tools (Calvert, 2015). Basically there are six types of domain specific cluster including integrated library system cluster, digital media archiving cluster, content management system cluster, learning content management system cluster, federated search system cluster and community communication interaction cluster (Breeding, 2014). These clusters can be developed by using the open source software and open source tools. Housekeeping operations and information retrieval system is also performed by the open source tools. But this research paper only explores the important innovative open source tools which easily manage the domain specific cluster for designing and developing an integrated information management and retrieval system for libraries. There are many open source tools are available in online environment but this research work select only the most comprehensive tools and it helps in different areas like installation and configuration of domain specific software for college libraries. College librarians can easily manage the different tasks through these open source software. Privilege control is also possible from the graphical user interface of SQLyog and here librarians are the super user for managing the library operations effectively and efficiently. Content management system and learning content management system is manage by open source tool which increase the use of information resources in college library to retrieved the right information at the right time.

1.1   Objectives

Actually unix was started in the university and academic purpose because its source code available in online environment. Since much of the development of Internet technologies took place within the walls of universities and research facilities, Unix became the operating system that was used for this development. The main objectives of this research paper are explained in the following ways :

(i)   To explore the barcode generate in college libraries on housekeeping operations.

(ii)   To provide the data backup, restoration and migration from one system to another on Ubuntu operating system.

(iii) To managed the authority data and bibliographic data through open source tools.

(iv)   To retrieved and access the data through discovery tool and data transfer tools in Ubuntu OS.

These objectives are the most important in the research paper to setup the single list of parameters and it also possible to integrated all the open source tools in a single window based interface for the housekeeping operations and information retrieval system. In this paper only select these tools because open source software gives the user the right to use, copy, distribute, examine, change and improve the software architecture in different levels like barcode generation, spine label creation, file transfer and access, server maintenance on Ubuntu operating system.

1.2   Methodology

Open source tools are to be selected on the basis of global recommendations and local requirements in the college libraries. All the softwares are managed on Ubuntu operating system because it gives more security, user-friendly and reliability. There are several studies about how free and open source software is developed, who takes part in the development and their motives for developing this kind of software. One observation made is the high level of creativity in development projects for free and open source software. Matured level softwares are to be selected in the following areas:

Parameters	Tools
Barcode generation	Glabels
Data backup and restoration	SQLyog, automysqlbackup and Postgresql
Data migration from one format to another	MarcEdit
Authority management	OCLC Dewey Cutter Program

FRBR and RDA	RIIMF
Discovery tool for the advanced users	VuFind
File access or run from the server machine	Telnet
File transfer	FTP
Offline communication in the campus	Squirrelmail or roundcubemail
System backup and installation	Remastersys

Table – 1: Tools selections in college libraries

1.3 Open Source Tools Used in Libraries

Now, in modern age increasing the use of open source software in college libraries. All the softwares are open because its source code available and customize with maintain this software time to time from a particular software home page. The important tools are to be represents are as follows :

(i) Glabels

This is the barcode generation tool (See Figure – 1) in integrated library system. There are two ways to generate the barcode like direct export barcode in excel format from the Koha databases and on the otherhand it also possible to generate the barcode from the excel sheet under the open office in Ubuntu operating system. So as to its new innovative tool that can manage the barcode with size, position and color.
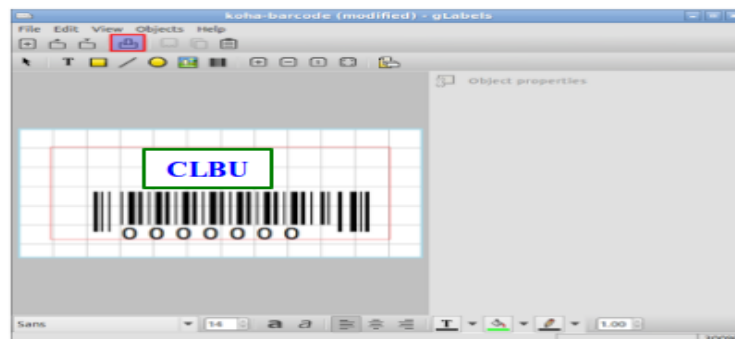

Figure -1: Glabels in barcode generation

(ii)   SQLyog

It is the graphical user interface tool (See Figure – 2) and integrated with the MySQL superuser. This is the most powerful tool which manage the database and helps to the college librarians for backup and restoration of a particular databases. It is also known as open database connectivity for data transfer and migration. Backup all the data related with domain specific cluster by using this tool both user and librarian interface.
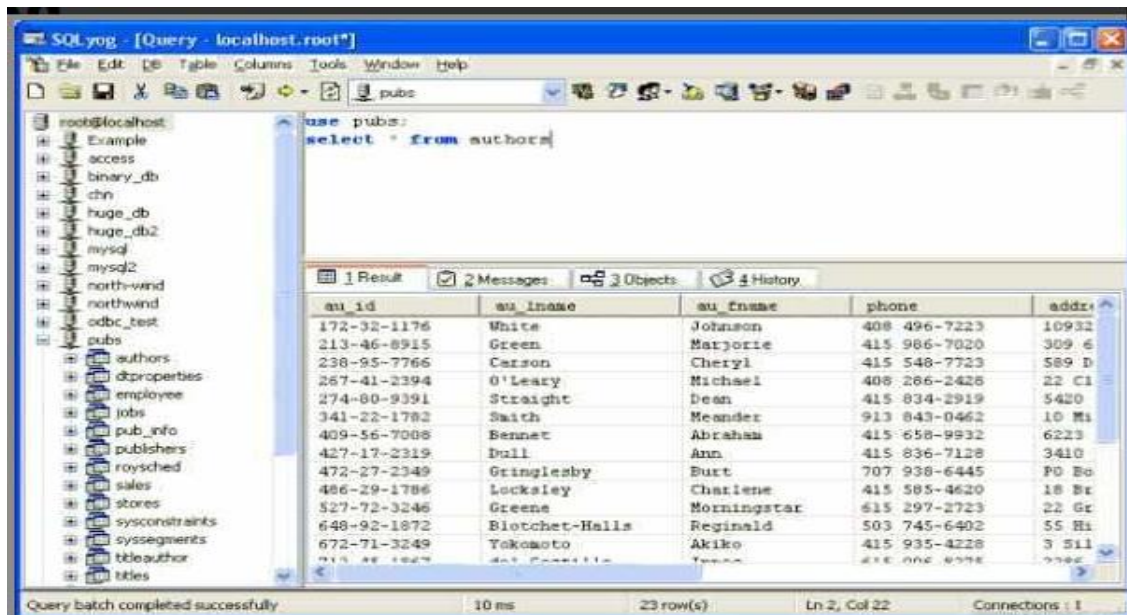

Figure – 2 : SQLyog interface for data backup and restoration

(iii)   Automysqlbackup

Database backup is one of the important task in domain specific cluster. Automatically backup the databases just open the home folder in Ubuntu and here backup all the softwares in logically and systematically also (See Figure - 3). Database backup is possible in three intervals like daily, weekly and monthly for MySQL database. College librarians can easily backup the databases by using the automysqlbackup tool because its very easy and systematic. Open the terminal from the application and just write the command sudo automysqlbackup and press here enter. Now, the database will be stored under the home directory in automysqlbackup folder.
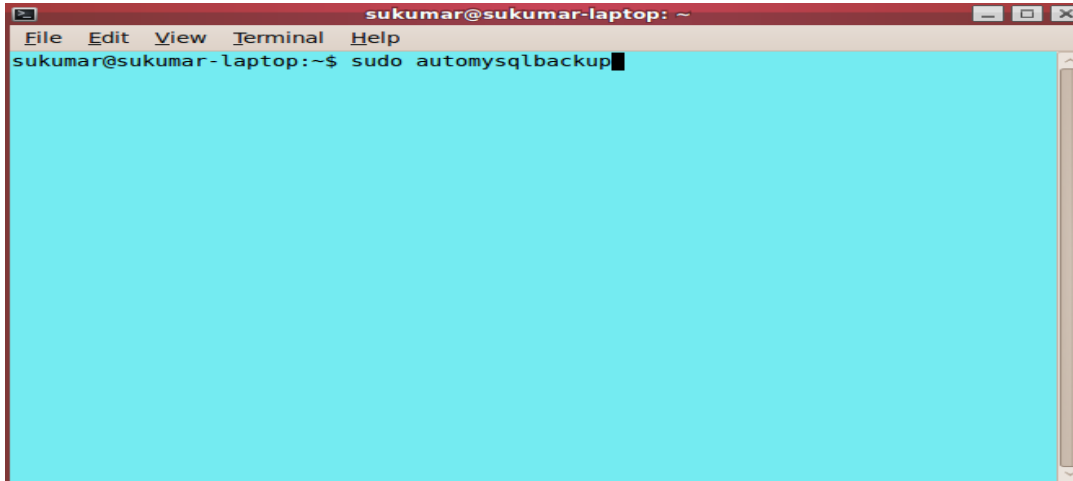
Figure – 3 : Automysqlbackup interface

(iv)   MarcEdit

MarcEdit (See Figure - 4) is developed by Terry Reese during 1999 for the management of bibliographic data. Data conversion is also possible from one format to another format. Customized the necessary fields and sub-fields from the bibliographic and authority format. MARC file is created from the ISO file as CDS/ISIS Winisis and also convert the .mrk format to .mrc format. Data migration is also possible by using this tool from one system to another. Excel data is imported in Koha by the marceditor which helps to users for searching and downloading the bibliographic information. College librarians can easily convert the file from iso to marc format and other format also.

Figure – 4 : MarcEdit interface for data migration

(v) OCLC Dewey Cutter Program

Cutter number is one important aspects in bibliographic information. It is a open source software that can manage the cutter numbers in two ways like four-figure cutter tables and cutter-Sanborn four-figure table based on the input text. The Figure – 5 is represents the OCLC Dewey Cutter program interface for college libraries. It is nicely works both the operating system like Windows and Linux and it enhances the classification efficiency under the tag of 082.
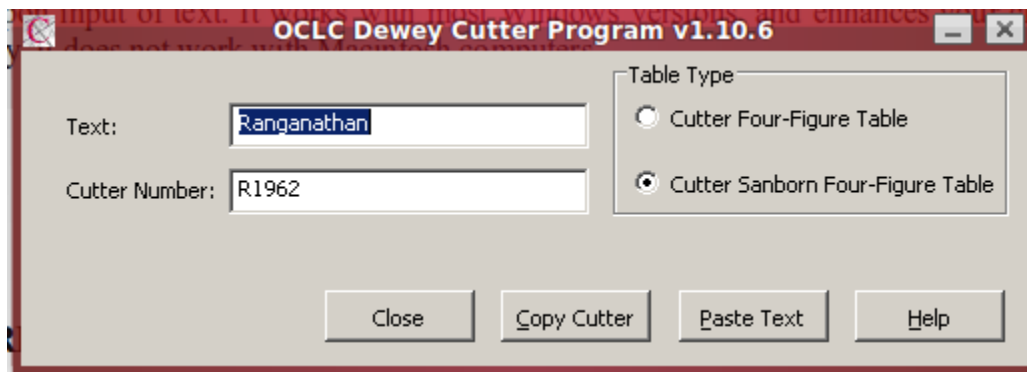

Figure – 5 : OCLC Dewey Cutter Program interface

(vi) RIMMF

It is a open source tool for management of RDA related documents in library automation. RIMMF stands for RDA in Many Metadata Formats. It is support the different formats like RDF, XML, MARC and etc. It also helps to exported the excel formats. New record of any items against in work, expression, manifestation, item, person, family, corporate body, concept, object, event, place and etc. Apart from these it also manage the FRBR standards by using this tool. This is the new innovative tool for college libraries that can helps to import the authority records from the VIAF and from the other authority control tools. Visualized is better, simple and comprehensive in different item types. WEMI concept is to be presents in the Figure – 6 for a particular entity in a library.
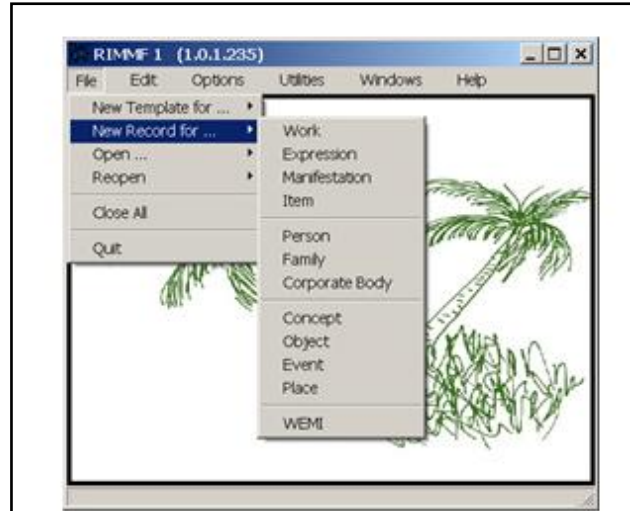
Figure – 6 : RIMMF interface for RDA

(vii)   PostgreSQL

It is an open source object relational database management system and used in backup and restoration. This tool is greatly helps to connect the two domain specific software like DSpace and NewGenLib. The primary function is to store the data securely for retrieving of information to the college users efficiently and effectively. It can handle the large amount of data both for single machine or other machine if the computer is connected to the LAN. Triggers is the important due it fully support and attached to other tables. Update triggers are also nice performing in this section to execute the programs and set of conditions. The Figure – 7 is represents the postgresql interface in college libraries for data backup and restoration in different domain specific cluster.
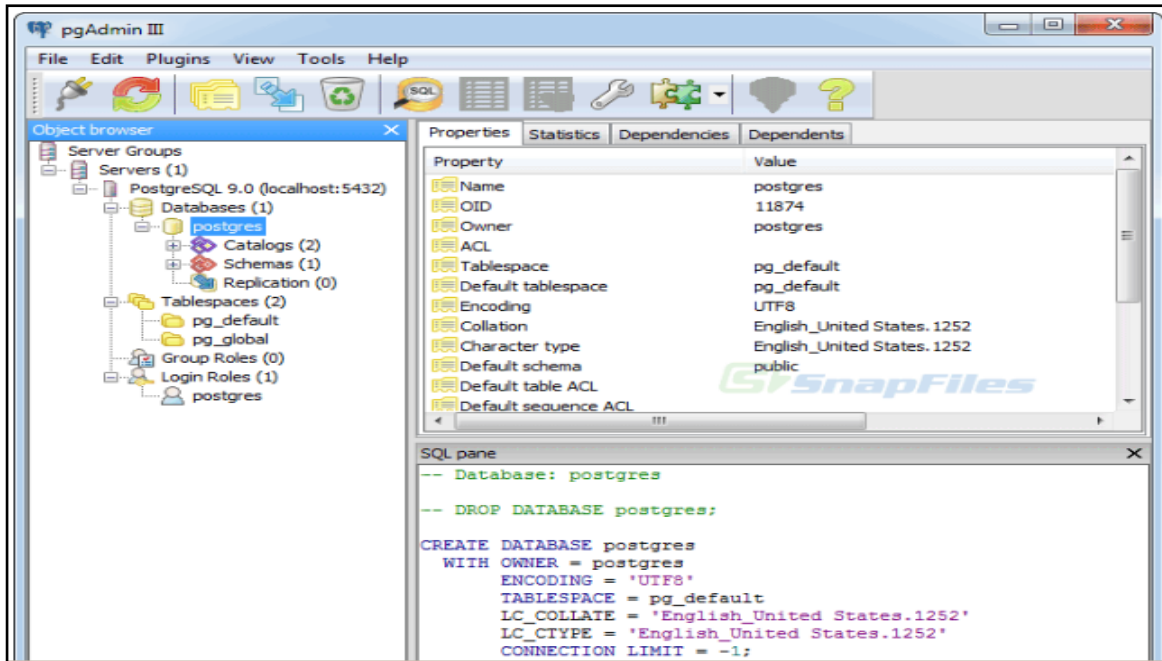
Figure – 7 : Postgresql interface for databack and restoration

(viii)   VuFind

VuFind is an open source discovery tool developed by Villanova University in 2010. Search and browsing the bibliographic information from the traditional OPAC into the VuFind. This interface is very much simple, sophisticated and google like interface. Keywords searching is very flexible because retrieved the most relevant information which available in Koha databases or other databases both for integrated library system and digital library system (Carter, 2009). While most commonly used for searching catalog records, VuFind can be extended to search other library resources including but not limited to: locally cached journals, digital library items, and institutional repository and bibliography (Chickering & Yang, 2014). It supports the multilingual interface in user as well as admin interface. Faceted search results that allow users to narrow items by format, call number, language, author, genre, era, region, and more (Yang & Wagner, 2010). The Figure – 8 represents the discovery interface of VuFind for browsing capability. Personal organization and annotation of resources through favorites lists, texting, e-mailing, tagging, and commenting features (Elleroa, 2014). A patron centered library is extensible; it remains flexible and reactive to user needs. Technology is a tool, not an end, so

librarians explore new technology in a climate where experimentation is encouraged through sound management techniques (Hofmann & Yang, 2012). Access services, document delivery and ILL operations are being redesigned to improve the user experience (Hoeppner, 2012). It support the Solr indexing tool as discovery layer services for the libraries. SolrMarc is used to import MARC metadata to the "biblio" index on Solr used by VuFind. The import process is controlled by the settings in the import/marc.properties file under the directory of /usr/local/vufind/import in VuFind installation directory. Sometimes, it is necessary to perform special data manipulation beyond the capabilities of the built-in SolrMarc functions. Apache Solr, an open source search engine, offers amazing performance and scalability to allow for VuFind to respond to search queries in milliseconds time. It has the ability to be distributed if you need to spread the load of the catalog over many servers or in a server farm environment. Solr is bundled as the built-in search in many applications such as CMS systems. The major Hadoop distributions from Cloudera, Hortonworks and MapR all bundle Solr as the search engine for their Big Data platforms. Solr is supported as an end point in various data processing frameworks and Enterprise integration frameworks.
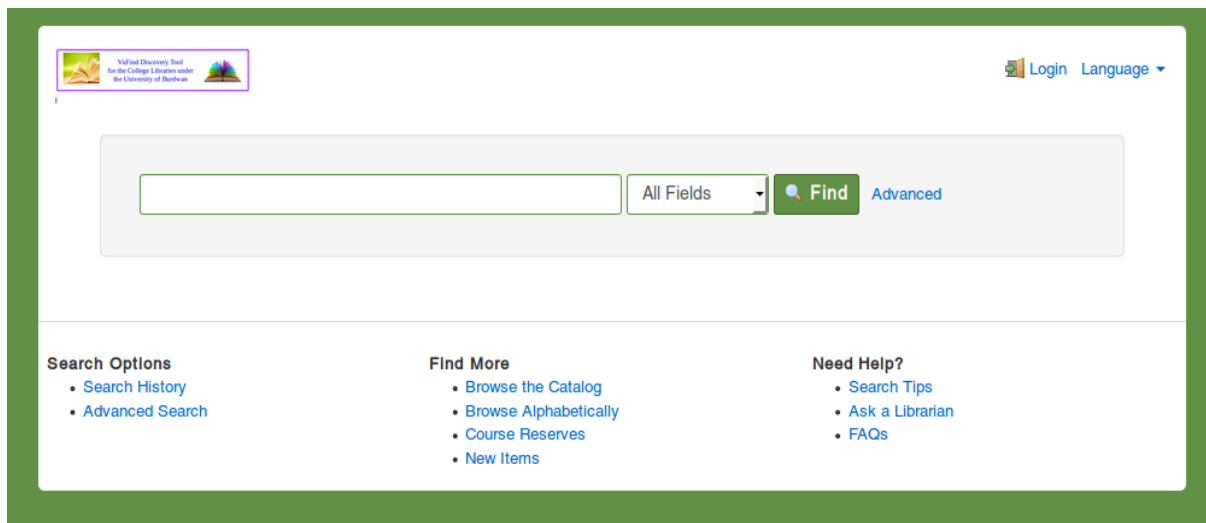

Figure – 8 : VuFind discovery tool interface

(ix)   Squirrelmail

Squirrelmail is an open source tool developed by Nathan and Luke Ehresman in 1999 and written in php programming language. It is support the LAMP architecture because its support the cross-platform. Offline mail is possible within the campus or colleges when Internet is not

present in the college campuses. But here only required the LAN connectivity to transfer the message and files from one computer to another. It is also known as client-server architecture that can easily manage the IMAP server. Multilingual information is to be transfer due its fully support the Unicode based standards. It can be run through the Mozilla Web browser and Google chrome also. File attachment is also possible during the mail sending to the friends or other persons. The Figure – 9 is represents the squirrelmail interface of college libraries in domain specific cluster which can helps the community communication interactions (Levine-Clark, 2014).
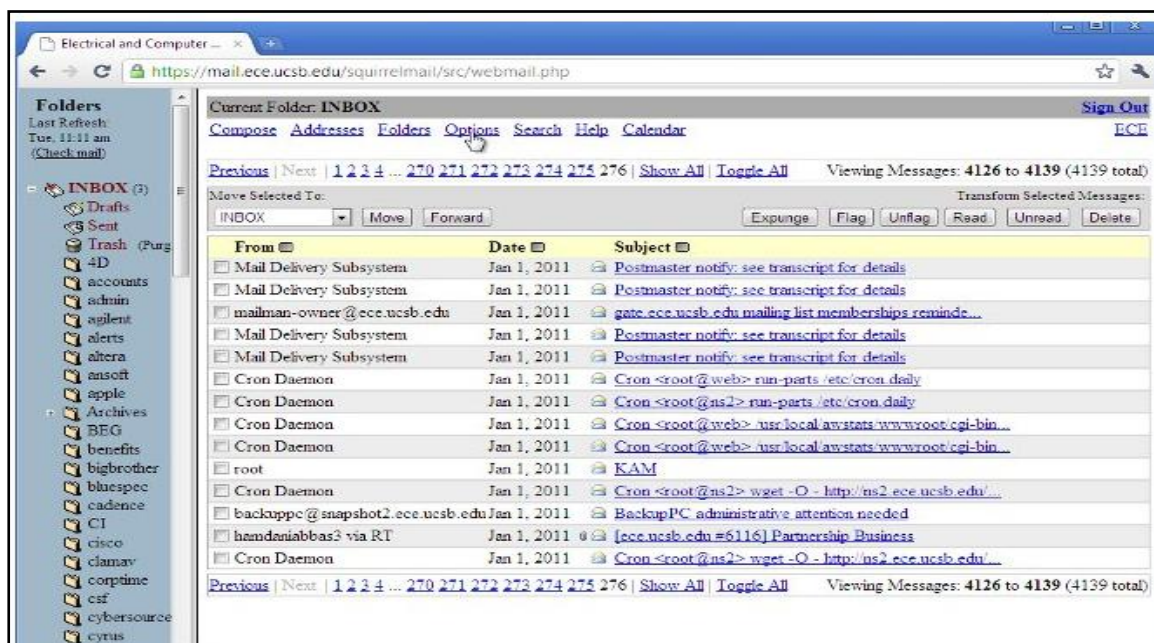

Figure – 9 : Squirrelmail interface of domain specific cluster

(x)   Telnet

Telnet is a telecommunication network based on client server protocol which is more reliable in domain specific cluster to access the files from the server machine. TCP/IP is the king protocol of the Internet that can transfer and manage the network resources in the college libraries. Other software is to be run and access by using this tool and the Figure – 10 is represents the telnet command on terminal with machine IP address. VuFind discovery tool is also run by using the terminal from the client machine and virtual terminal connection.

Figure – 10 : Telnet interface in domain specific cluster

(xi)   File Transfer Protocol

FTP stands for file transfer protocol and used for transfer the files from one computer to another. Get the files from the server machine if its connected to the local area networking based on the client server architecture. It also support the hypertext transfer protocol which fix the bug for small ephemeral transfer. The Figure – 11 is represent the file transfer protocol interface in the domain specific cluster for community communication interaction. FTP has a stateful control connection which maintains a current working directory and other flags, and each transfer requires a secondary connection through which the data are transferred. In "passive" mode this secondary connection is from client to server, whereas in the default "active" mode this connection is from server to client.
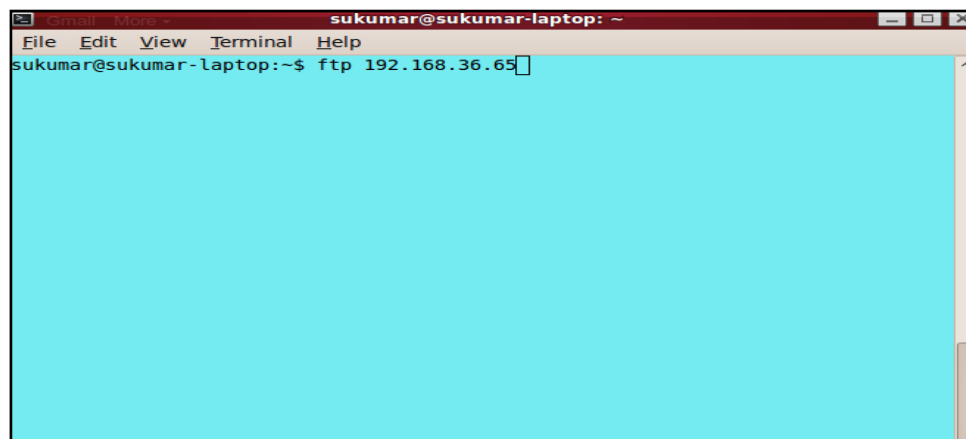

Figure – 11 : File transfer protocol interface in domain specific cluster

(xii)   Remastersys

Remastersys is also an important open source live DVD creation tool in this framework to provide the all files and users data. It is support both the operating system like Windows and Ubuntu, so as to it is known as cross-platform. It is also found in two formats like .deb and wget format but this research work select the second option because this is very easy to create the ISO image file in Ubuntu (http://remastersys.org/). Backup can be made on the basis of size of the files if file size larger then only backup the user data (i.e super user) and on the other hand it will select the whole backup file size is small. Settings is also an important aspects in live DVD format first name the live DVD (e.g CLBU Live DVD) and it will stored in the specific location on Ubuntu (/home/remastersys/CLBU.iso) for buruning and run the system. Burning is also possible from the option of bracero after that run and successfully installed the Live DVD on Ubuntu. All the college libraries can run their necessary cluster by using the whole system in a single window based interface.
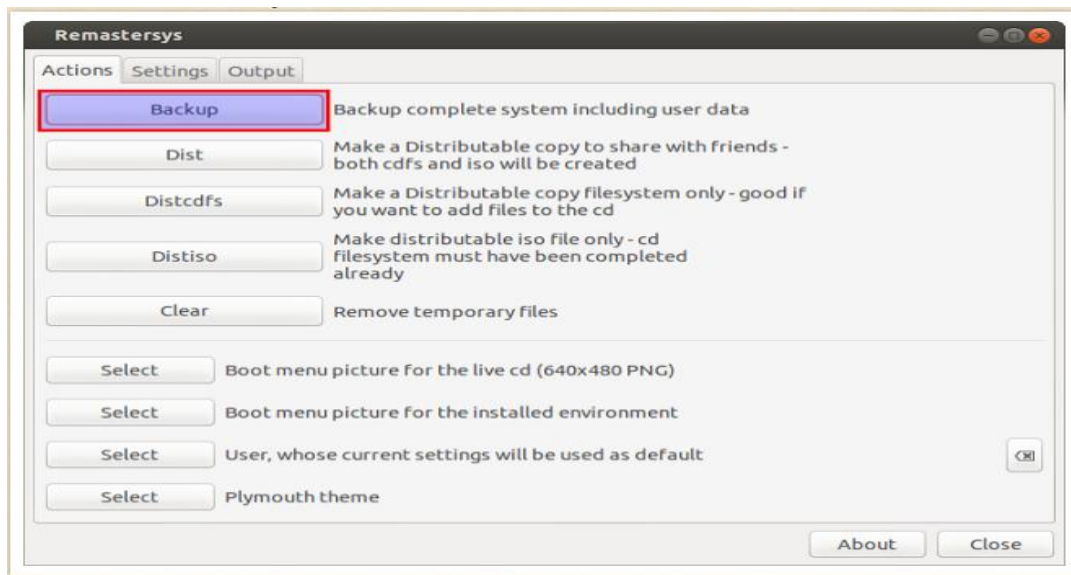

Figure – : Remastersys backup on Ubuntu

1.4   Integration in Ubuntu Operating System

Integration facilities is one of the important aspects in Ubuntu operating system. This research work successfully integrated the above mentioned open source tools in Ubuntu to provide the

better services in libraries. The most comprehensive parameters are to be selected and integrated in housekeeping operations and information retrieval system for the users as well as library professionals. After creation the files through remastersys and its possible to installation and configuration of other target libraries. Data migration is also an important tasks in automated and digital library system and this can be done through MarcEditor. Most comprehensive tasks is Backup and restoration in integrated library management and retrieval system this can be achievd through SQLyog. Another difference of Telnet from a raw TCP session is that Telnet is not 8-bit clean by default. 8-bit mode may be negotiated, but high-bit-set octets may be garbled until this mode was requested, and it obviously will not be requested in non-Telnet connection. Explicit FTPS is an extension to the FTP standard that allows clients to request FTP sessions to be encrypted. This is done by sending the "AUTH TLS" command. The server has the option of allowing or denying connections that do not request TLS. The SSH file transfer protocol (chronologically the second of the two protocols abbreviated SFTP) transfers files and has a similar command set for users, but uses the Secure Shell protocol (SSH) to transfer files.

1.5   Findings

The findings of this research paper are explained in the following ways :

(i)   It is easily possible data conversion and data migration from one system to another through MarcEditor.

(ii)   Data backup and restoration is possible through two important tools like SQL yog and automysqlbackup.

(iii)   Author mark is managed by the OCLC Dewey Cutter program under the MARC tag 082.

(iv)   RDA and Vocabulary is managed by the RIIMF in college libraries and it also support the FRBR entity relationship model.

(v)   Remastersys helps to create the live DVD both for data and user related documents.

(vi)   File access and transfer from one computer to another computers are managed by the telnet and ftp in college libraries.

(vii) Offline community communication is possible through squirrelmail in the college campus.

(viii) Bibliographic and authority data is managed and send to the users mobile through VuFind discovery tool.

(ix) Glabels can managed the barcode in college libraries and it also managed the patron card layout.

(x) Postgresql helps to achived the backup in digital media archiving cluster like DSpace and this can be easily possible through open source tool.

1.6   Conclusion

The domain specific cluster is one of the important things in college libraries. Technology is inevitable and indispensable and in this respect one important quotes is given by the Arthur C. Clarke : "Any sufficiently advanced technology is indistinguishable from magic.". All the tools of domain specific cluster are important because it can easily manage the housekeeping operations and information retrieval system in college libraries. Backup and restoration is possible through these tools which can increase the library services for the students and teachers also. The college librarians can easily manage and maintain the domain specific cluster in different areas including integrated library system, digital media archiving cluster, content management system, learning content management system, community communication interaction and federated search system also. Search, browse and retrieve the relevant information to the users from one computer to another computer by using the backup and restoration tool like SQLyog and postgresql. So, as it can conclude that these tools are to easily and efficiently solved the problem of different advanced level tasks in libraries.

References

Breeding, Marshall. (2014). Library Resource Discovery Products: Context, Library Perspectives, and
 Vendor Positions. *Library Technology Reports*, 50 (1). Available at: http://dx.doi.org/ 10.5860/ltr.50n1
Calvert, Kristin. (2015). Maximizing Academic Library Collections: Measuring Changes In Use Patterns
 Owing to EBSCO Discovery Service. *College and Research Libraries*, 76 (1): 81-99. Available

online:
http://dx.doi.org/10.5860/crl.76.1.81

Carter, Judith (2009). Discovery: What Do You Mean by That? *Information Technology & Libraries* 28 (4), 161–63, http://dx.doi.org/10:6017/ital.v28i4.3326.

Chickering, F. William, and Sharon Q. Yang. (2014). Evaluation and Comparison of Discovery Tools: An

Update. *Information Technology and Libraries*, 33 (2): 5-30. Available at:
http:// dx.doi.org/10.6017/ital.v33i2.3471

Clarke, Arthur C. (October 1945). Extra-Terrestrial Relays. *Wireless World (Iliffe and sons, Ltd.),* 51 (10): 305–308.

Elleroa, Nadine P. (2014). Exploring Library Discovery Positions: Are They Emerging or Converging?" *Journal of Web Librarianship*, 8 (4): 331-348. Available at:
http://dx.doi.org /10.1080/19322909.2014.963778

Hoeppner, Athena. (2012). The Ins and Outs of Evaluating Web-Scale Discovery Services. *Computers in*

Libraries, 32 (3): 6-11. Available at: http://www.infotoday.com/ cilmag/apr12/ Hoeppner-Web-Scale-Discovery-Services.shtml

Hofmann, Melissa A. and Yang, Sharon Q. (2012). Discovering' What's Changed: A Revisit of the

OPACs of 260 Academic Libraries. *Library Hi Tech,* 30 (2), 253–74.

Levine-Clark, Michael. (2014). Access to Everything: Building the Future Academic Library Collection. *portal: Libraries and the Academy*, 14 (3): 425-437. Available at:
http://dx.doi. org/10.1353/pla.2014.0015

Remastersys (2016). *Remastersys live dvd.* Retrieved from http://remastersys.org/. (Accessed on 15th July 2016)

Yang, Sharon Q. and Wagner, Kurt (2010). Evaluating and Comparing Discovery Tools: How Close Are We towards Next Generation Catalog? *Library Hi Tech,* 28 (4), 690–709.