

CHAPTER - 4

EMPIRICAL FINDINGS, ANALYSIS & DISCUSSION

In this chapter we have presented our findings from the case study that we have done. We have presented how data warehouse maintenance is actually carried out in an organization.

4.1 Introduction

Today, the telecom sector is expanding due to new inventions in technology. This has increased the competition in the telecom market. The telecommunication sector's aim is to satisfy their customers by providing different services efficiently. Customer satisfaction is becoming a challenging task for the telecom industry due to increased number of customers. Customer relationship management, market analysis, the evaluation of call detail records, HLR (Host Location Registry), on-line accounts and personalized telecommunication services require very efficient database support [JM00].

Bharat Sanchar Nigam Ltd. formed in October, 2000, is World's 7th largest Telecommunications Company providing comprehensive range of telecom services in India: Wire line, CDMA mobile, GSM Mobile, Internet, Broadband, Carrier service, MPLS-VPN, VSAT, VoIP services, IN Services etc. Presently it is one of the largest & leading public sector units in India.

BSNL has installed Quality Telecom Network in the country and now focusing on improving it, expanding the network, introducing new telecom services with ICT applications in villages and winning customer's confidence. Today, **it has about 46 million line basic telephone capacity, 8 million WLL capacity, 52**

Million GSM Capacity, more than 38302 fixed exchanges, 46565 BTS, 3895 Node B (3G BTS), 287 Satellite Stations, 614755 Rkm of OFC Cable, 50430 Rkm of Microwave Network connecting 602 Districts, 7330 cities/towns and 5.6 Lakhs villages.

BSNL is the only service provider, making focused efforts and planned initiatives to bridge the Rural-Urban Digital Divide ICT sector. In fact there is no telecom operator in the country to beat its reach with its wide network giving services in every nook & corner of country and operates across India except Delhi & Mumbai. Whether its inaccessible areas of Siachen glacier and North-eastern region of the country. BSNL serves its customers with its wide bouquet of telecom services.

BSNL is numerous Uno operator of India in all services in its license area. The company offers wide ranging & most transparent tariff schemes designed to suite every customer. BSNL cellular service, Cell One, has **55,140,282 2G** cellular customers and **88,493 3G** customers as on 30.11.2009. In basic services, BSNL is miles ahead of its rivals, with **35.1 million Basic Phone subscribers** i.e. 85 per cent share of the subscriber base and 92 percent share in revenue terms.

BSNL has more than 2.5 million WLL subscribers and 2.5 million Internet Customers who access Internet through various modes viz. Dial-up, Leased Line, DIAS, and Account Less Internet (CLI). BSNL has been adjudged as the NUMBER ONE ISP in the country.

BSNL has set up a world class multi-gigabit, multi-protocol convergent IP infrastructure that provides convergent services like voice, data and video through the same Backbone and Broadband Access Network. At present there are **0.6 million Data One broadband customers.**

The company has vast experience in Planning, Installation, network integration and Maintenance of Switching & Transmission Networks and also has a world class ISO 9000 certified Telecom Training Institute.

Scaling new heights of success, **the present turnover of BSNL is more than Rs.351,820 million (US \$ 8 billion) with net profit to the tune of Rs.99,390 million (US \$ 2.26 billion) for last financial year.** The infrastructure asset on telephone alone is worth about Rs.630,000 million (US \$ 14.37 billion).

The turnover, nationwide coverage, reach, comprehensive range of telecom services and the desire to excel has made BSNL the No. 1 Telecom Company of India. BSNL's primary aim is to offer top quality mobile services and promote healthy competition in the mobile market. Also BSNL aims to create value for shareholders through the serving of customers, employees, partners and the general public interest. In a long-term perspective, a strong market and customer focus, as well as a strong commitment to their employees and to society, provide the best platform for creating incremental value in their business.

To achieve this goal BSNL has been totally computerized for its day to day operations and is completely dependent on its information systems for the running of daily business. There are 5 major operational systems namely:

1. SIEBEL CRM (customer relationship management)
2. GENEVA (billing and postpaid traffic)
3. MEDIATION

i. Postpaid: For the postpaid subscribers of BSNL the traffic is rated using Mediation postpaid.

ii. Prepaid: For the prepaid subscribers of BSNL the traffic is rated using

Mediation prepaid.

4. VOMS (Voucher Management System): All the prepaid scratch cards, electronic credit transfers and easy loads are managed through VOMS.

5. MSC: The raw traffic source for BSNL is MSC source system. In this all the traffic including prepaid, postpaid, inter-connects and transit is managed.

In addition to these five major operational systems, BSNL, since the day it started its operations in BSNL Bikaner has been using an Oracle based data warehouse to strengthen its decision making process. The data warehouse is located at its central head office in Chandigarh. This is a centralized data warehouse having the main database at one location but it is in the process of conversion to a distributed data warehouse and professionals from Tera data (a subsidiary of NCR) are working on it to finish this job.

According to the data warehouse project manager at BSNL, data that is useful and helpful from all the source systems is stored in the data warehouse. This helps in the consolidation of data at a central repository. When this data is projected over a period of time a trend can easily be detected in the projections. For e.g. it can be detected which are the cell sites that are used more often and which are not. In this way a capacity planning for the cell sites can be conducted by the BSS department. Similarly activations for each area can be projected over a certain period of time and it can be made out what are the places where the growth is highest and the brand is popular. Similarly the age group in which the brand is famous is another valuable demographic which can be put to use by Marketing and Brand Management department.

There are nearly 350 employees currently accessing data warehouse including personnel from data warehouse department, operations department, business

intelligence department and business analysts group. The data warehouse department is responsible for performing tasks related to maintenance of data warehouse. All the operational systems of BSNL listed above are centralized. Data from all the operational systems is consolidated at a certain central location called COB (close of business). This COB is actually the ODS (operational data store) as described in data warehouse architecture. BSNL is storing data from several operational systems in the data warehouse. This includes all the data related to finance, call history, and subscriber's database.

Then the data warehouse is fed using the push or pull scheme depending on the operational system. Mostly the traffic or call history related data is pushed through ftp to a location which is then parsed, moved to staging, transformed and then loaded into the data warehouse. The entire subscriber's related data is extracted from the current source systems and then it is moved to staging and then transformed and loaded into the production system of the data warehouse. According to the data warehouse project manager, BSNL has nearly 350 employees using the data warehouse in these districts.

4.1.1 Query Performance

We did some experiments by which we compared the performance of query response time using our proposed model and old model on same data. The result of the both model is given below.

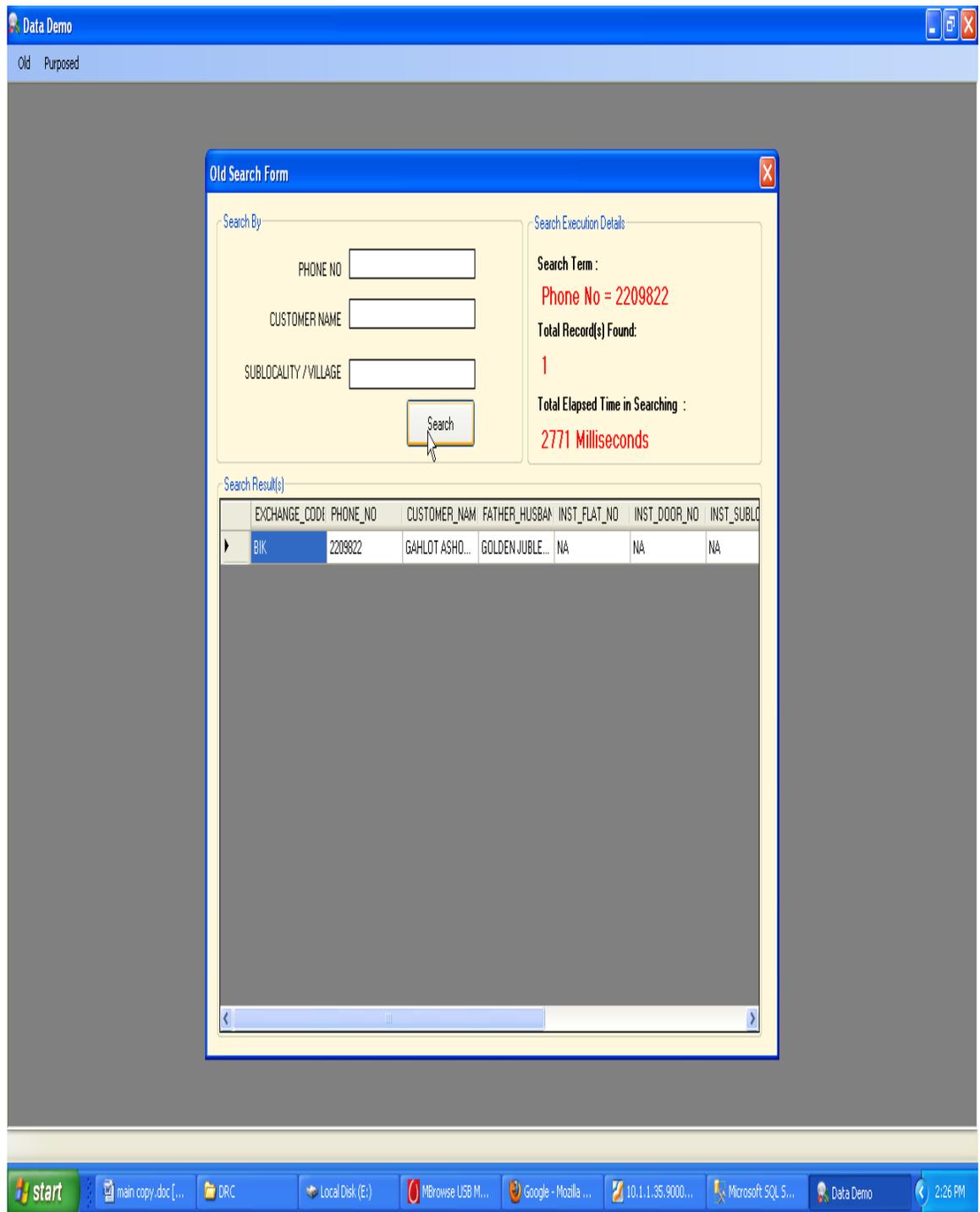


Figure 4.1 – Searching phone no. by previous technique.

The old model takes 2771 millisecond to find a telephone number. The model will go through many rounds and will take too much time. The model works on indexing concepts and it supports simple queries structure.

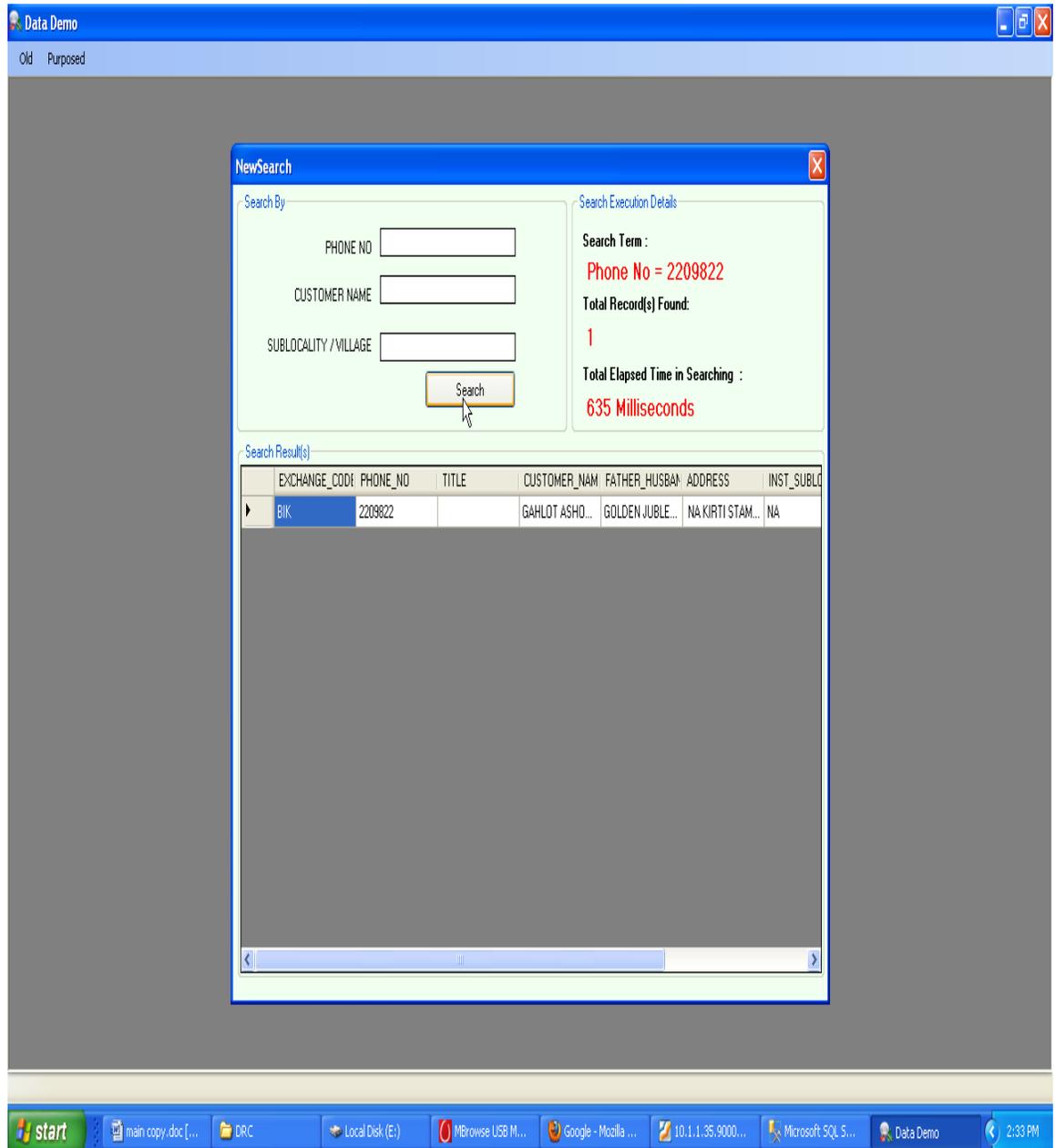


Figure 4.2 – Searching phone no. by proposed technique.

Our proposed model takes 635 milliseconds to find the same telephone number. The model will go through few rounds and will take lesser time. The model works grouping concept and it supports stored procedures. The stored procedure is a fast one because it is already precompiled. It is easy to maintain. My propose model design once and stored in executable form, so this model is quick and efficient.

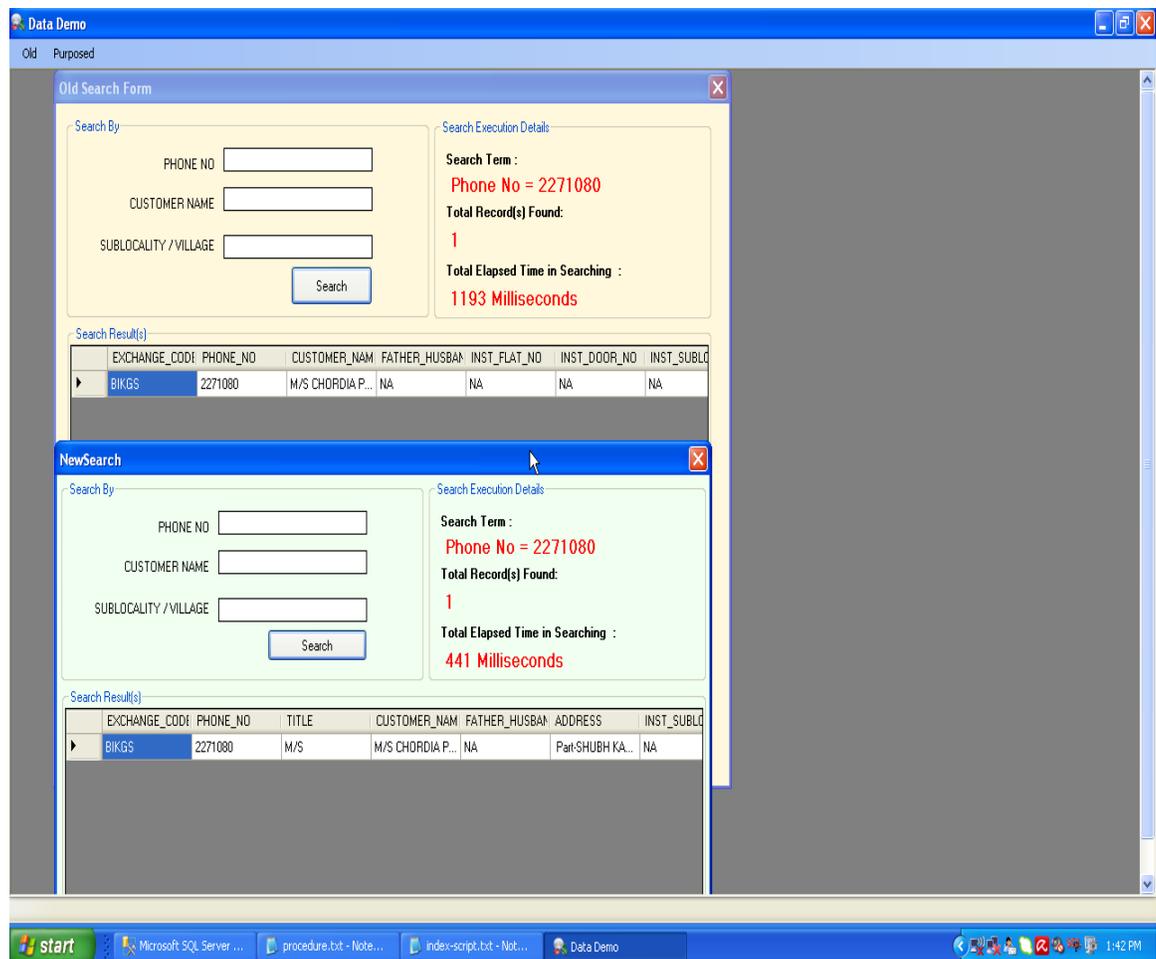


Figure 4.3 – Comparison of searching phone no. by previous technique and proposed technique.

From the figure we measure an old search form (with simple query) takes 1193 milliseconds but new propose model search take only 441 milliseconds for finding the same phone number. This near optimal speedup is explained by the fact grouping the records increases the efficiency and reduces the time for retrieval of data.

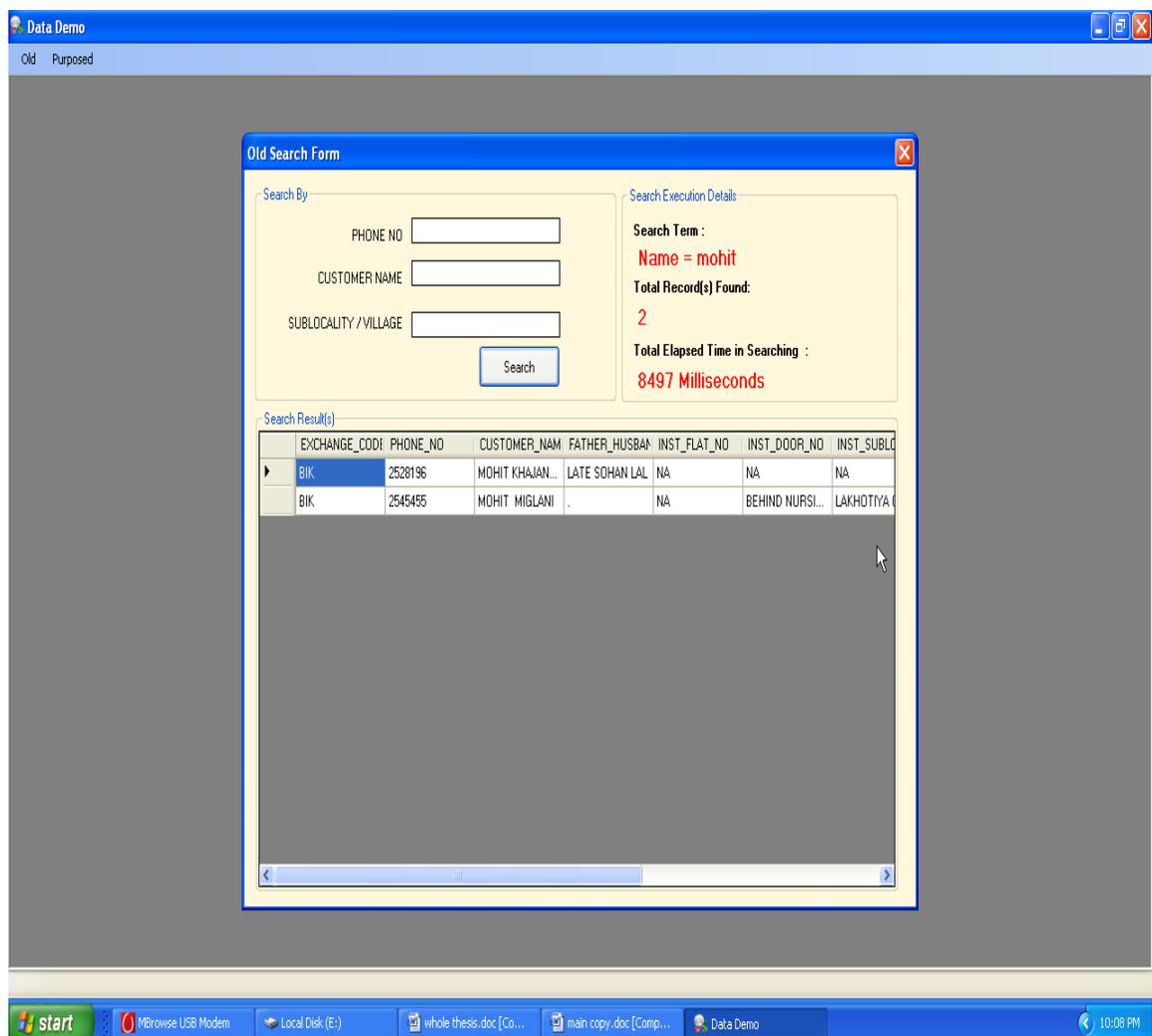


Figure 4.4 - Searching Customer Name by previous technique.

When we performed search on name query, the old search method took about 8497 milliseconds. And in the last it found 2 data according to query using indexing techniques. It took too much time in comparison to our grouping model.

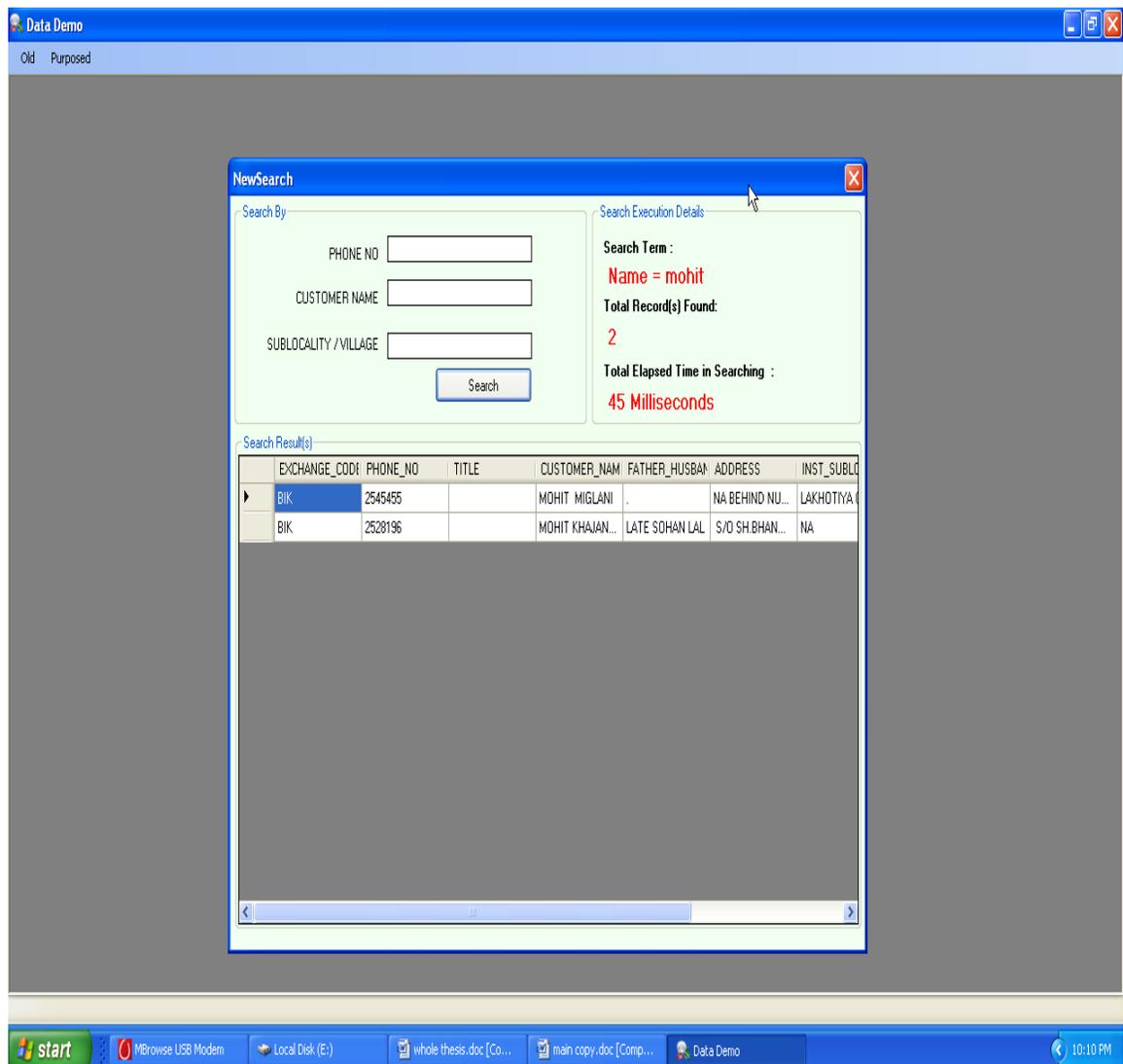


Figure 4.5 - Searching Customer Name by proposed technique.

When we performed the same query on same data using our grouping model, we found then it took 45 milliseconds which are very less in comparison to old searching methods.

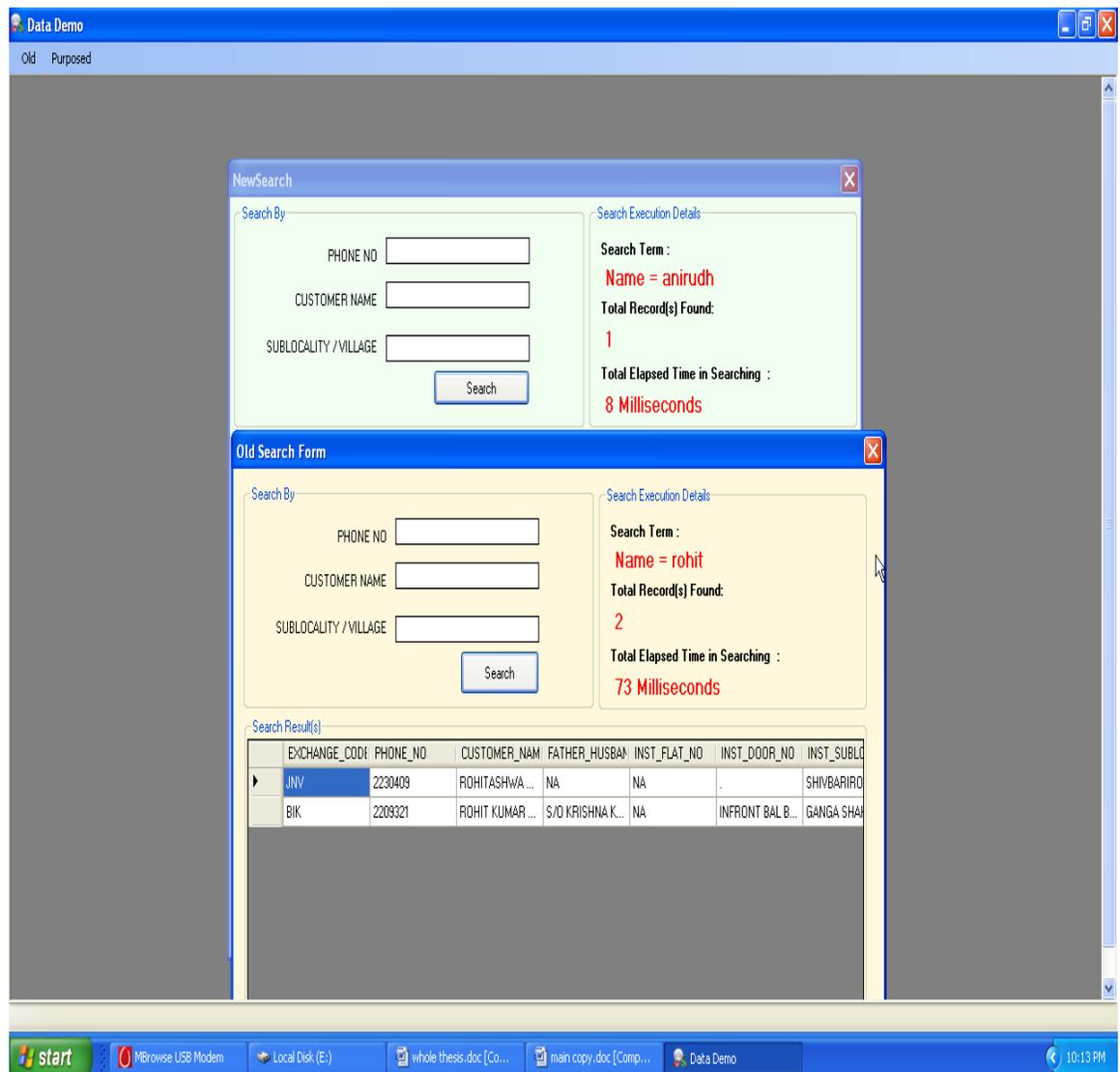


Figure 4.6 - Comparison of searching Customer Name by previous technique and proposed technique.

When we performed the same query on different data using our grouping model and indexing method, we found that our grouping model took 8 milliseconds which are very less in comparison to old searching methods which took 73 milliseconds. Our model searched 1 record whereas old searching method searched 2 records as shown in figure above.

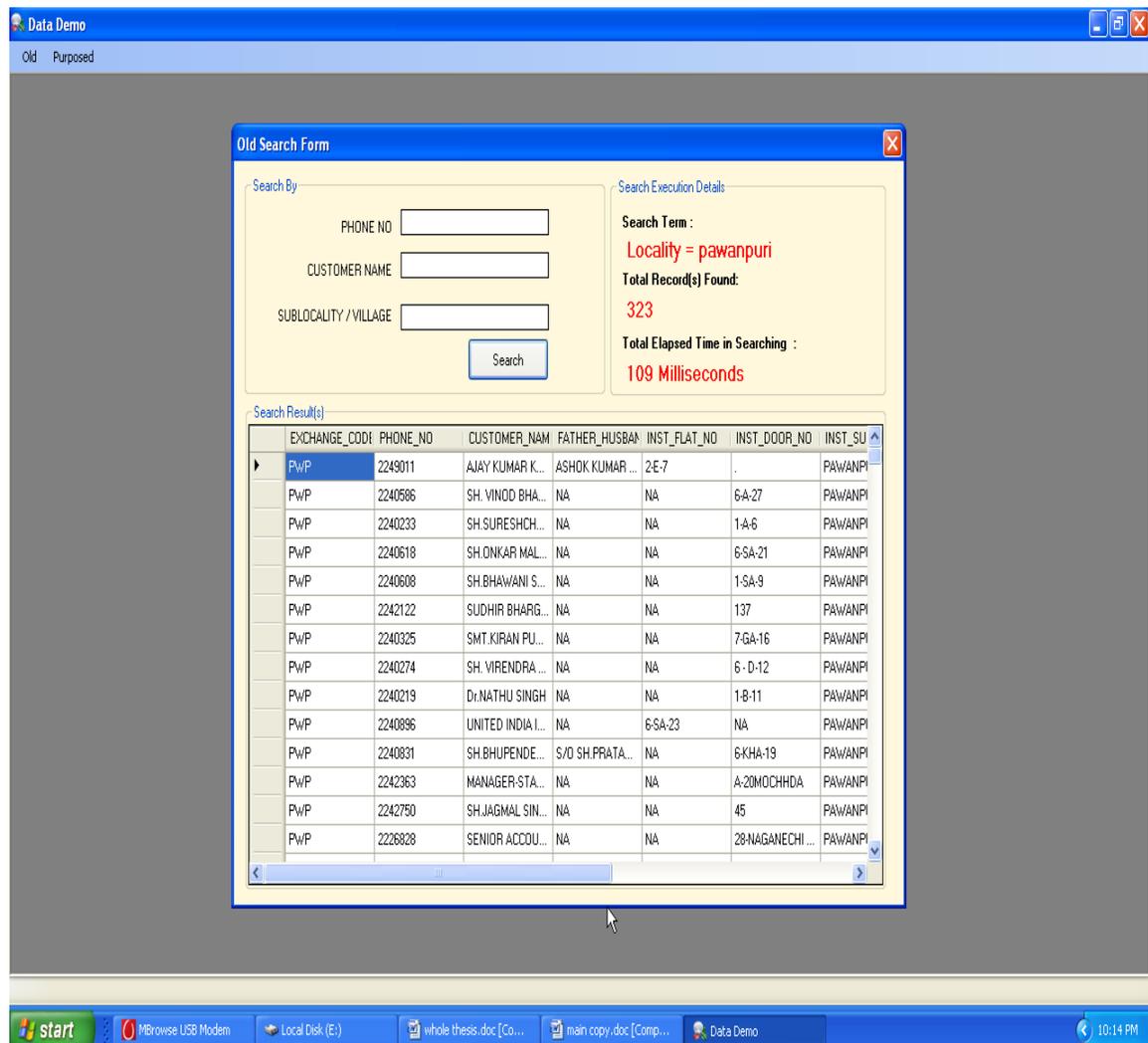


Figure 4.7 - Searching Sub locality by previous technique.

When we performed the same query on locality data using old searching model, we found then it took 109 milliseconds in finding 323 data which are much more in comparison to our method.

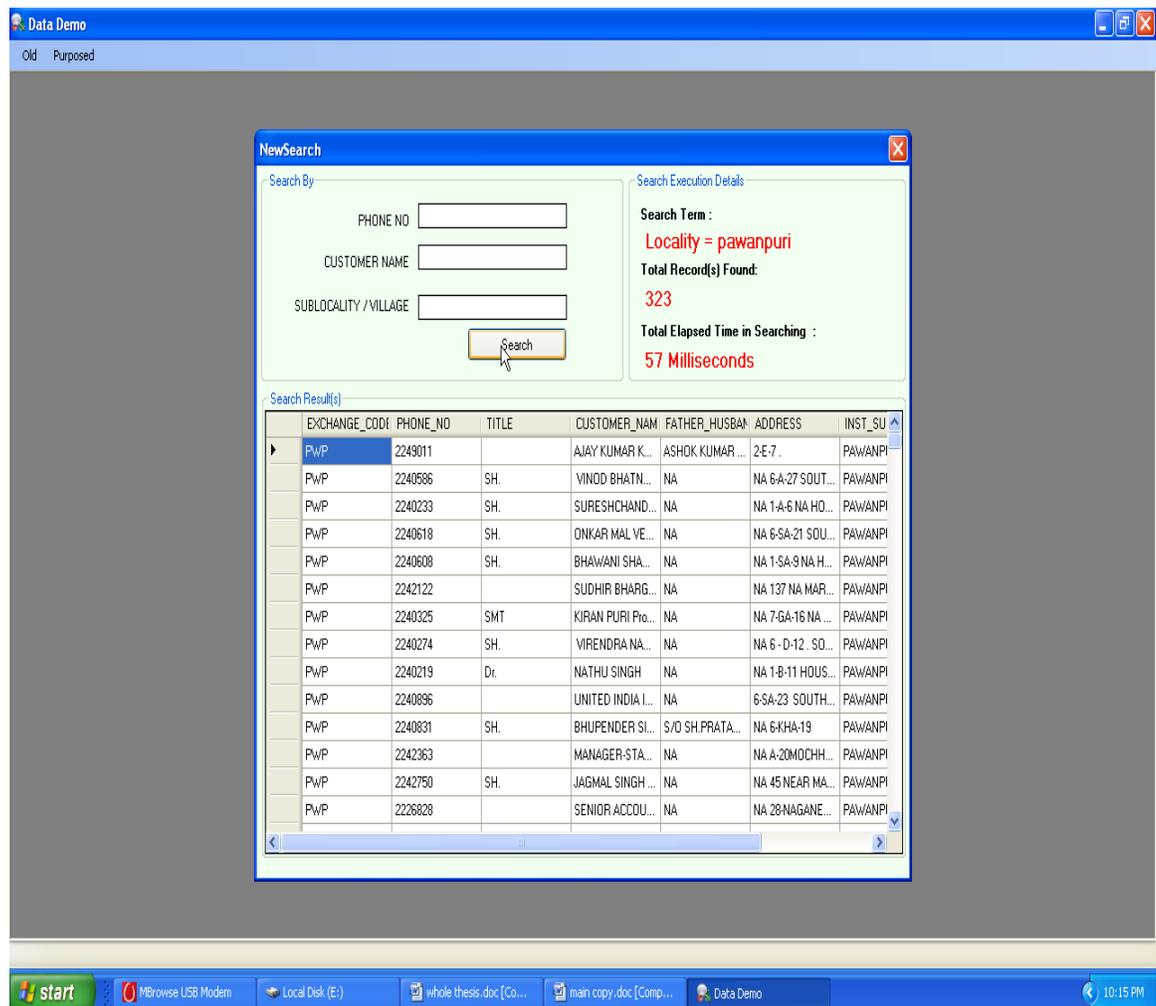


Figure 4.8 - Searching sub locality by proposed technique.

When we performed the same query on locality data using our grouping model, we found then it took 57 milliseconds which are very less in comparison to old searching methods.

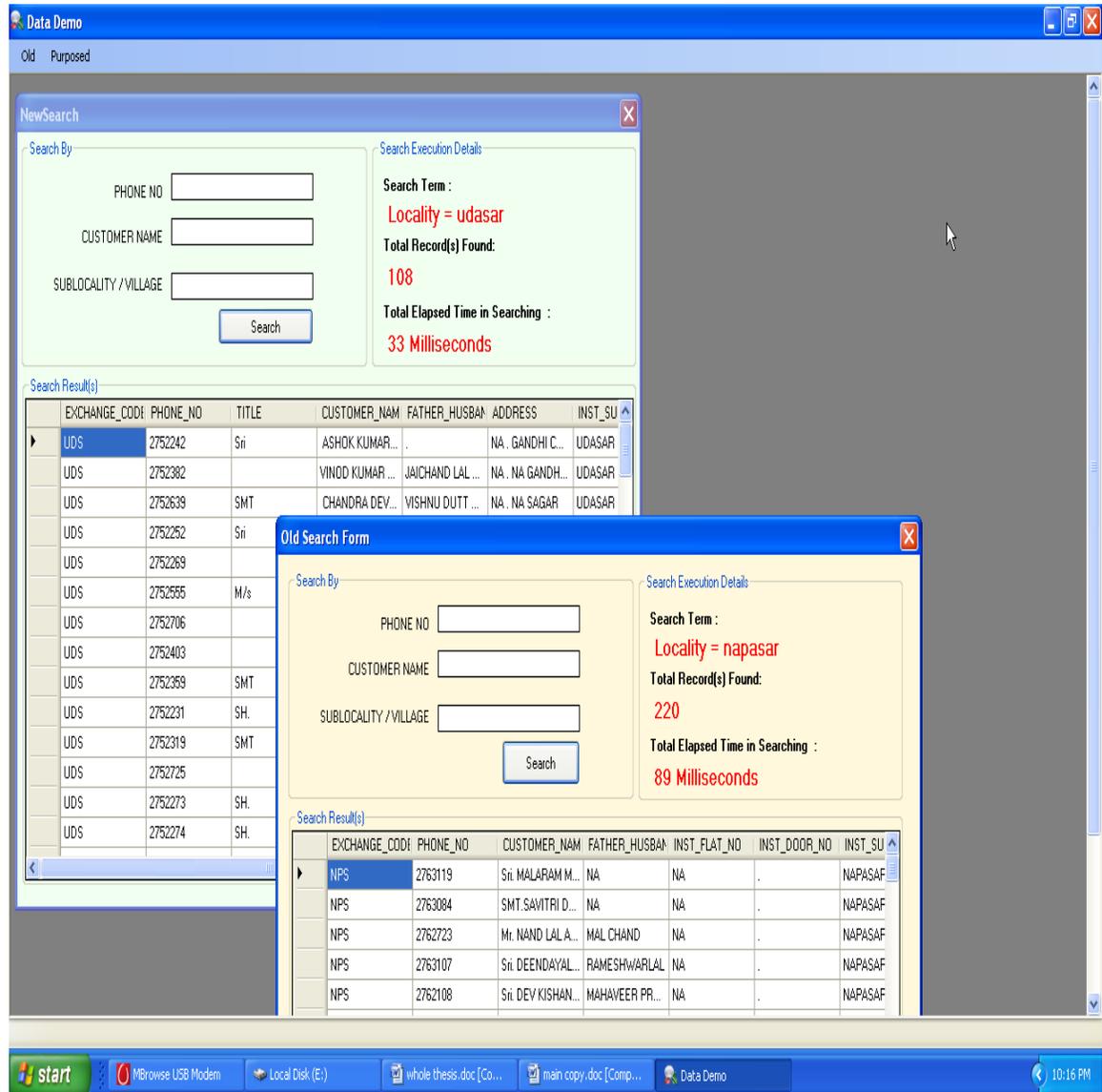


Figure 4.9 - Comparison of searching sub locality by previous technique and proposed technique.

When we performed the same query on different data using our grouping model and indexing method, we found that our grouping model took 66 milliseconds which are very less in comparison to old searching methods which took 89 milliseconds. Our model searched 216 records whereas old searching method searched 220 records as shown in figure above. So we can see that our model much-much better than old searching method

4.1.2 Coordination and Communication

Survey and interview shows that, BSNL has started its communication program by publishing a booklet related to the data warehouse. The booklet contains complete information about the purpose of the data warehouse, the scope of the project, the aim behind the implementation of data warehouse, what input is given to the system, what output can be taken from the system, and who are the responsible persons, those should be contacted to get any further information and can help in any matter regarding data warehouse. The members of the data warehouse core team are in constant contact with the data warehouse user, and they keep them informed about any development or shortcomings in the projects. Project manager further elaborated that the most difficult part in implementing a data warehouse in the company is the training of business users.

Most of the business users are technology shy. They just want the reports and analysis on paper and everything printed. Making them comfortable with the use of computer and other multimedia facilities is a very tough task. So top management always try to understand their problems and coordinate in their work if possible. Once the business users have dipped their hands in this area

they realize how important and how beneficial it is for their better understanding of the use of data warehouse.

4.1.3 Education, Training and documentation

BSNL has started a comprehensive education and training program for its employees. For this purpose they have set aside a good amount in the overall data warehouse budget. The training is conducted workshop for 5 to 15 days depend of employee's position and work.

In these workshop or training programs, each and every employees can any kind of interaction with the data warehouse. These workshops made to take up some data ware house and application related trainings and industry level courses. In which they taught application tools being used in the organization. Same goes for the managed services people who are trained on the lines of the DBA. Managed services personnel are trained locally by the vendor, while for the logical data model training the personnel are sent to the state of the art training centers of the vendors situated at ALTTC Gaziabad.

According to a Business User of BSNL, the training and education program has helped him a lot in understanding the capabilities and functionalities of data warehouse and he can now easily look for historical data that can greatly help in estimating future network and technological requirements for the business.

4.1.4 Help and support

According to the Survey and interview shows that, there is formal help and support counter created in BSNL for providing support to the data warehouse users. As the data warehouse project is in the beginning the data warehouse core team is still thinking of how to implement the help and support counter services. They are providing the 24x7 hours support to the data warehouse users. The team members provide answers to user's queries and help them if there is any technical or other problem with data warehouse usage.

BSNL have centralized IVRS (in each Zone), CTI (Computer Telephony Interface), IP EPABX ,etc. the call centers mostly one per each circle connected to the data center. The 1500 calls and the 198 calls routed to IVRS. Depending upon the number or the CLI, the call routed through the IP network to the respective help and support counter. The help and support team member have one IP phone and PC connected over an IP network to the data center. The customer data is displayed on the screen of the computer and the IP phone provide the Voice communication with the employees.

The problems encountered by the users can be of any type including report generation, usage of tools for data mining, query problems, and user management etc. The members of this team are knowledgeable in the field of data warehousing and have the ability to solve any problem related to data warehouse on their own. If the problem is beyond their control they report the problem to the Technical Team Lead. It is the responsibility of the support and backup team to report the most commonly and most frequently found problems to the Project Manager so that the root cause of the problem could be solved, whether in the form of any update to the system or development of a new

module.

They are also given training for this purpose from time to time. The help and support up team members can be contacted either by phone, email or in person at any time. The support team logs all the questions and their responses in a web based system where other users can also check and learn with their experiences. The support and team can also be contacted through a special interface on the company's local web based portal, where the users can send a memo or a note to the support team related to any functionality problem, or any enhancements for the data warehouse.

According to a member of the help and support team get a lot of questions related their problems each day and they answer to all these questions quickly, so as the business user not feel any discomfort with the data warehouse.

4.1.5 Management of data flow on network

According to the data warehouse project manager the data warehouse needs constant high speed connectivity with the network. Because it connect all SSA's, Circles and corporate office, providing connectivity to all its main exchange. All exchange dealing with customers such as JTOs, SDEs, AO and the entire management so far each SSA or circle has established network for implementing DOTSOFT and other local system so it needs high speed connection with the source systems preferably with the fiber optic link to execute the daily push and pull operation in which the files are extracted from the source or the files are thrown onto the landing server to be picked up by data warehouse through the FTP. Moreover the repository of the business application is placed on a separate server.

The following figure shows in general the exchange network and the collection

method of data. Each exchange connected to a router which is called LE router (Local Exchange router). All new technology switched such as OCB,EWSD,5ESS,AXE,shall be connected using X.25 cards and Ethernet interface (Wherever available).all CDOT exchange be connected to the LE router using CES equipment supplied by CDOT .All E10B exchange is Connected to the LE router through MTE (Magnetic Tape Emulator). Each LE Router is connected to the aggregation router through E1 Links. All the E1s coming from the different exchange be aggregated to the aggregation router. Each Aggregation router in each SSA is connected over STM-1 link to the nearest MPLS node. For redundancy purpose the connectivity established to two MPLS nodes.

Data Collection Methodology

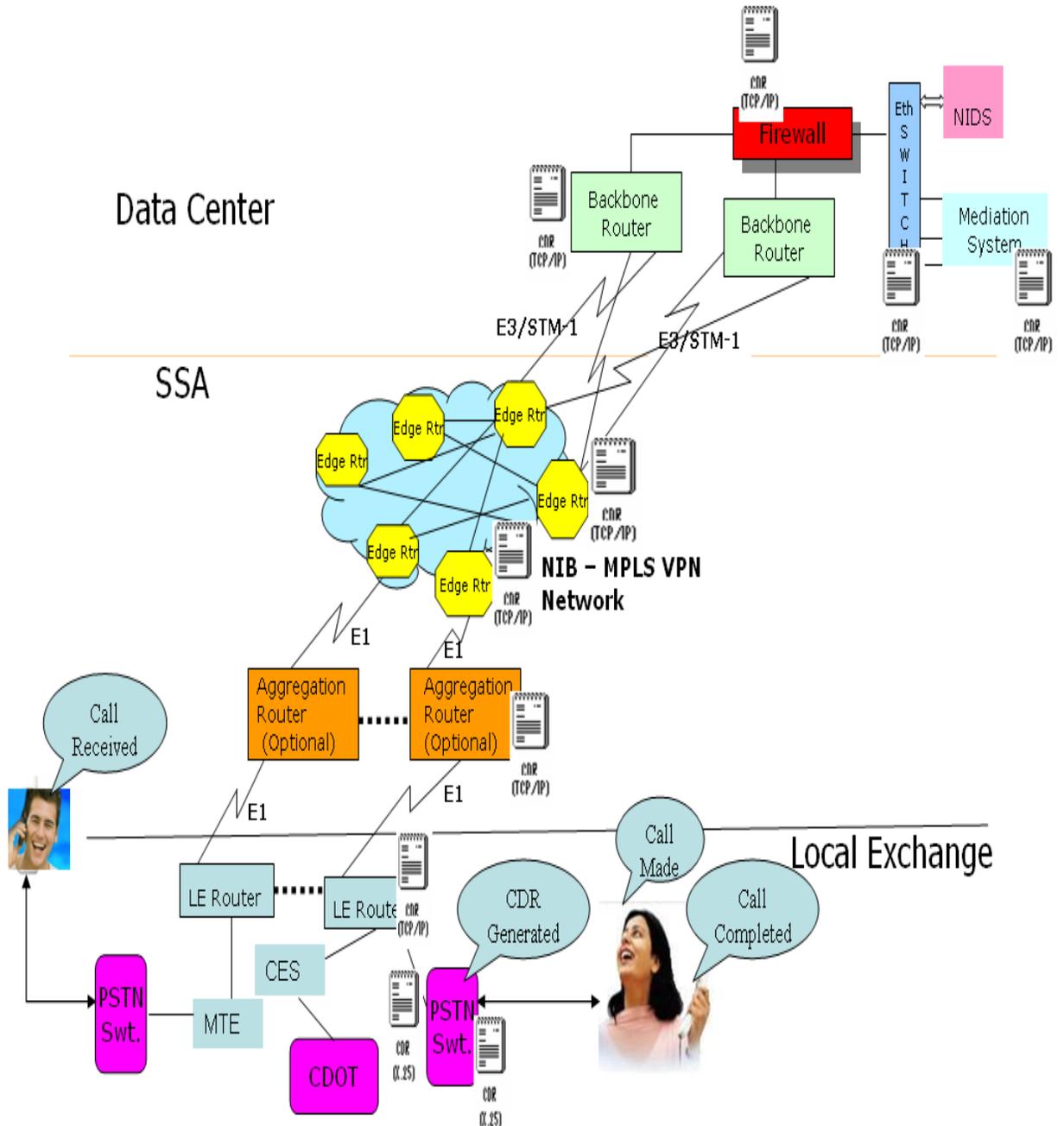


Figure 4.10 - Data Collection Methodology in BSNL

The data Center is also connected to the MPLS network presently through STM-1 links to start with .this end link enhanced to 1GBPS links or later. Thus, each exchange shall be connected to the Data Center over E1 end links and through the MPLS network the CSR Network also get connected to the Aggregation Router.

Thus, all the terminal of Commercial, TRA, FRS and Directory Enquiry which is connected to the local system is connected to the data center through the Aggregation Router.

Data Center Network Overview

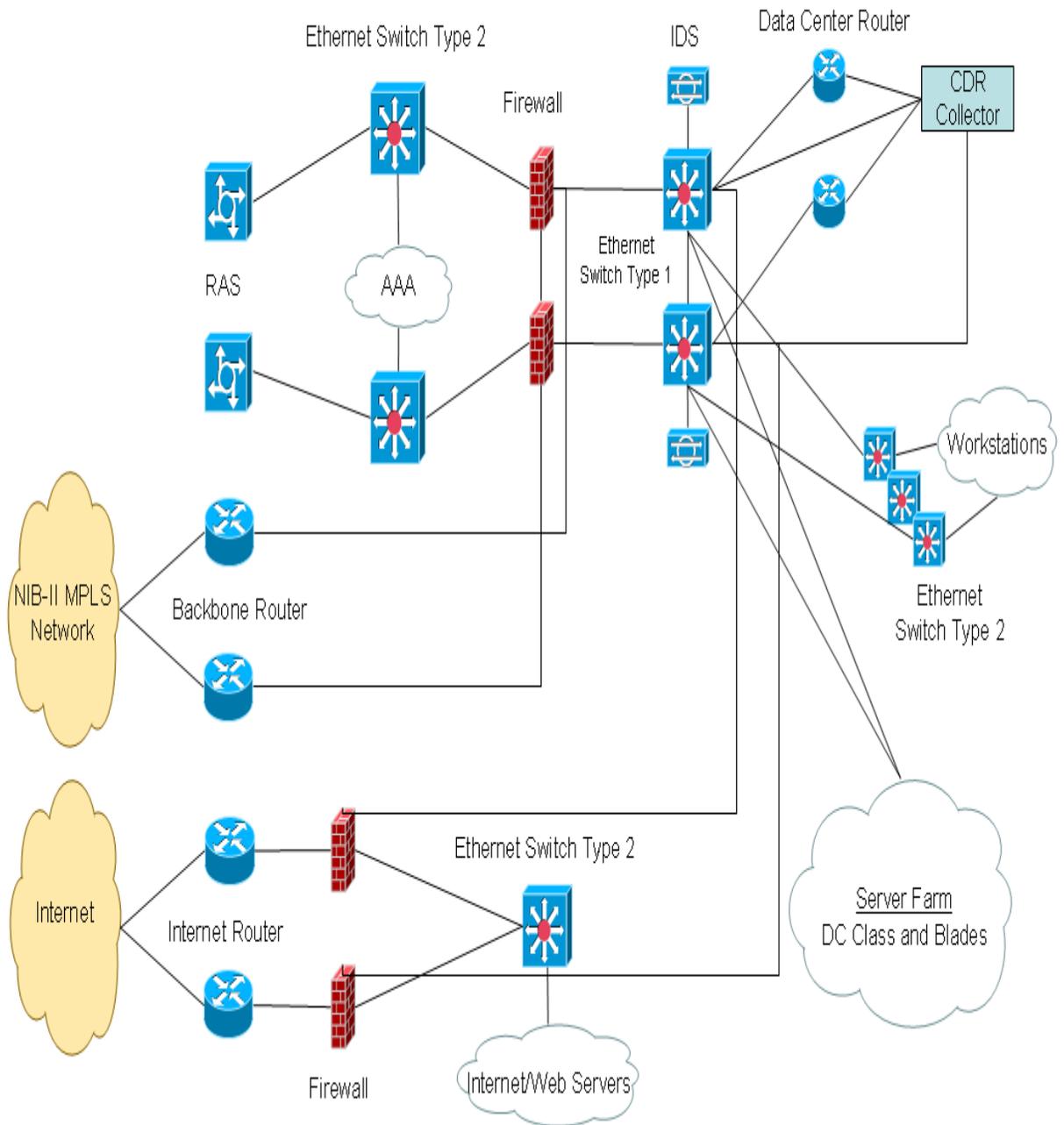


Figure 4.10 - Data Center Network Overview

Exchange router – two connections to exchange for CDR collection and service provisioning

two WAN connections.

- Aggregation routers in case of star topology.
- LAN extender/Remote router for extending LAN from LE to CSR and RLU's
- Aggregation routers connected to Edge routers of MPLS VPN.
- Central routers (optional) between Aggregation routers/LE routers and Edge routers.
- Backbone Routers (two in each DC) for connecting DC to MPLS VPN
- Data Centre routers – termination of all data traffic in the billing network. X.25 serial connectivity to CDR collectors; connected to the Backbone routers.
- Managed Ethernet switch, CSR Ethernet switch, Data centre Ethernet switches (Type I and Type II), Work station Ethernet switch at DC.
- Internet Router at DC for internet connectivity to the DC.
- RAS for connectivity for remote CSRs.

BSNL's data warehouse is not too old, its usage is not too much as well and also the data warehouse is centralized it doesn't need high maintenance of the network. Once all the links are configured and made secure the process normally runs smoothly. A team of network engineers is responsible for the maintenance of the network related tasks. This team is also responsible for the overall network of the company. The team utilizes some network monitoring tools to ensure the smooth and reliable operation of the network. Once in a week the team checks as if all the links are working properly and require any maintenance or tuning. After the up gradation of the data warehouse from centralized to a distributed there definitely be a need to upgrade the network. In

that case BSNL is planning to implement a fiber optic network.

4.1.6 Loading and Cleansing Data

According to Survey and interview shows that the data warehouse Project Manager at BSNL, ETL is one the most important and most time consuming process. There are five systems feeding the data warehouse. All these applications have a different platform than that of the data warehouse. Also there are different formats used by the operational systems and by the data warehouse. For e.g. there may be a situation where the operational systems allow the use of null values while the data warehouse does not. To take with these types of problems and ensure a consistent and reliable ETL operation BSNL has employed a specialist ETL person and given the title ETL Lead.

Professional services for the ETL function are provided by the vendors. Once all the services are in place, the data warehouse team for BSNL takes over and they are the owners of the data warehouse. ETL lead is responsible for managing all the Extract Transform and Loading related tasks. ETL chief is a expert in the field of databases having complete command over database structures, database connectivity, data extraction, data transmission, and programming.

He is responsible to give duties to members of the ETL team. Any problem that is out of control for ETL team members are reported to ETL chief who then tries to solve it himself otherwise reports it to the Technical Team Lead and ETL Team Lead are helped by the industry consultants for communications in this regard. ETL Chief works in close cooperation with the

Business Astuteness Lead. This helps both the parties in determining what level of data detail are required and which levels of aggregation are best suited for the needs of BSNL.

4.1.7 Software & Hardware

From the Survey and interview I found out that the minor hardware problems are looked after by the help and support team. Moreover incase of problems relating to the major hardware changes are looked after by the hardware vendors like Tera data (a subsidiary of NCR Corporation).

The hardware certification is done by the vendor as they are responsible for its installation, maintenance and up gradation. BSNL has adopted the policy to use hardware from the same vendor if available because of compatibility and performance issues. As far as data warehouse is concerned BSNL is using hardware from the same vendor so they are not having too much problems related to hardware. The same is the case with software. The most common thing that needed to be updated frequently is the storage media. Technical Team Lead is the person responsible for keeping an eye on the hardware resources present and the hardware resources required. If there is some shortcoming in performance due to hardware issues it is reported to the Project Manager and a decision for hardware up gradation is taken in consultation with the vendor.

From the interviews it was observed that as far as the software side is concerned, the professional services side of the vendor trains all the data warehouse resources for keeping them up to date. The business software problems are being looked after by the 3rd party whose software BSNL is using. BSNL has entered into an agreement with the data warehouse vendor

for the software up-gradation. Whenever a bug or some problem is found in the system, it is reported to the vendor. Afterwards it is the responsibility of the vendor to solve the problem. This is usually done by providing a software update for the product or by doing some troubleshooting. After every three months professionals from the vendor side also visit the data warehouse site and check its performance and determine if there are any problems or any update is required or not. The areas where problems are frequently found include data loading mechanisms and query management. The vendor also notifies if there is any need for hardware up-gradation or not.

4.1.8 Materialized view

According to data warehouse Project Manager at BSNL views are made according to the business user needs for the aggregates and for the base tables. They allow limiting the user access needs on the production tables. So once there are changes in base tables, views needed to be updated as well. Services for view maintenance are provided in the data warehouse by the vendor. It is the decision of the user to decide which views to materialize and how they be updated. They just have to use certain commands to update a view in response to changes to the underlying tables.

BSNL is using incremental view maintenance for refreshing views on tables. Instead of loading all the data from scratch the views are updated incrementally whenever the data in source systems or base tables is updated. The policy used for view maintenance depends on the type of data and the frequency of its usage. For e.g. for data that is important (a minor change can affect the results badly) the views are immediately updated while for views containing data that is not important (changes in data do not affect the

results too much) deferred view maintenance is used. Functionalities for all these operations related to view maintenance are provided by the vendor in the form of data warehouse tools.

If there is some problem found in the view maintenance it is the job of the help and support team to inform the ETL team regarding any changes in the structure of the tables. Later ETL team takes care of the problem.

4.2 Analysis & Discussions

Now I analyze my theoretical findings and the empirical findings by comparing with each other. I try to find out what are the similarities and differences between the theory and the real world system. Here again I follow our previous structure for data warehouse maintenance by using coordination and communication and training, education and documentation, help and support , management of data flow on network, software and hardware , Loading and Cleansing Data operations and materialized view as the main concepts for analysis and comparisons.

4.2.1 Query Performance

In this section techniques are implemented and tested on the basis Response time with dataset of different sizes. These techniques are full text index, cluster index, non cluster index and my propose model. The graphical representation shows all indexing techniques with respect to different datasets.

Dataset Description

The dataset which is used for analysis is taken from BSNL. The dataset of 15,00,000 records is divided into different sub datasets to identify the variation in size and calculation of execution time of different indexes.

Time Analysis

Time of various indexing techniques corresponding to different datasets has been calculated and on basis of these values, indexing techniques have been analyzed on factor like CPU time.

CPU Time Consumed Analysis

CPU time is the combination of compilation time and execution time. The performance of indexing techniques is shown graphically. It is observed that the grouping model is better than other indexes because it takes less CPU time in comparison to others. The cluster index's consumption of CPU time increased gradually as the number of records in datasets increased. It is observed that cluster index CPU time hike from 1500 ms to 5000 ms approx in range 2 to 3 lakh records. Non cluster index is in mid range and time is 35000 ms approx and Full text index is better than cluster and non cluster but my model is best all of them because it take less CPU time in comparison to all indexing.

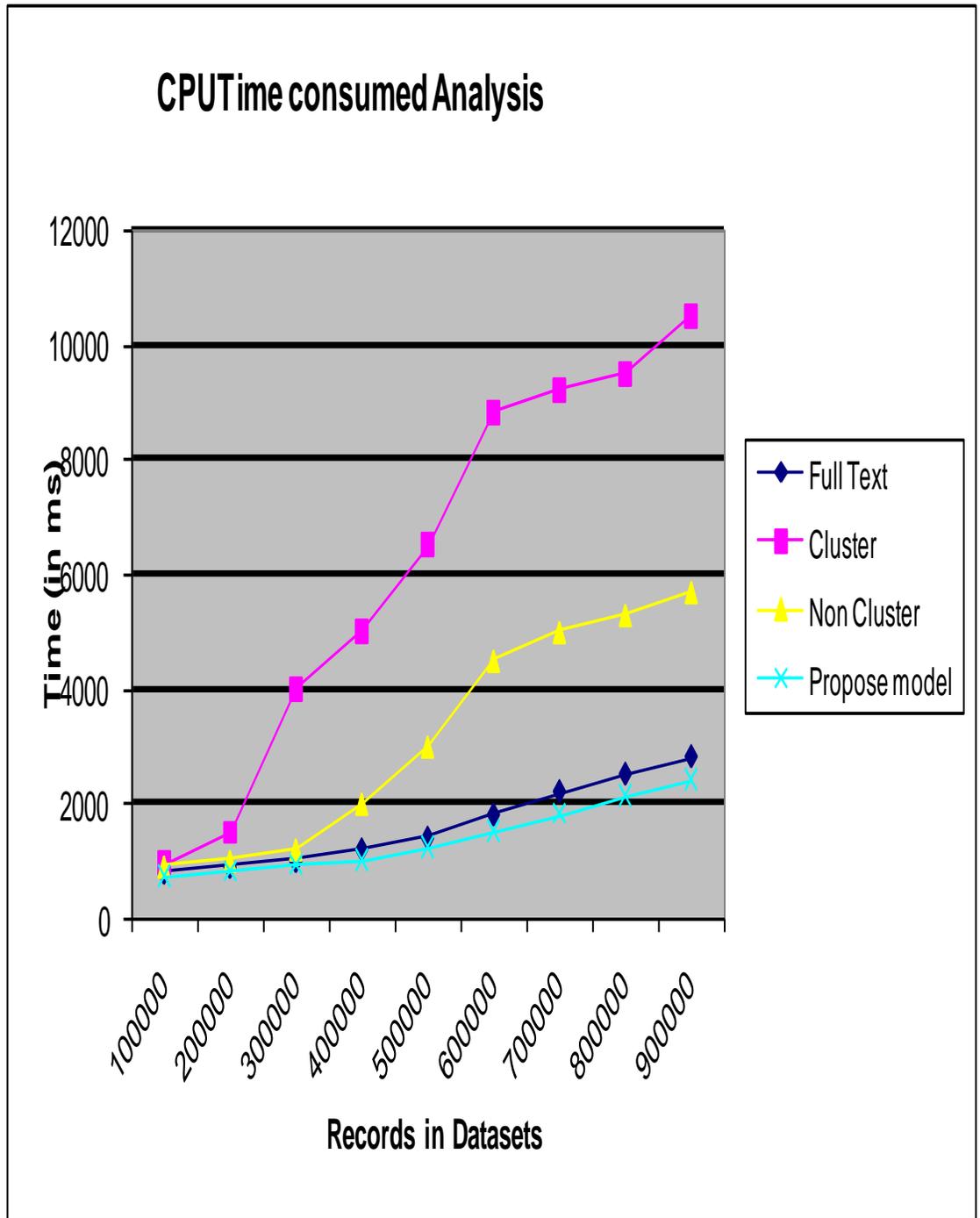


Figure 4.12 – CPU Time Consumed Analysis

Execution Time

Initially, the execution time for all indexing techniques is near to 1000 ms from 1 to 2 lakh record. Full text index execution time is below to 2000 ms but there is fast hike of time in the cluster index i.e 9000 ms approx till 9 Lakhs record. The performance of non cluster indexing is less than full text index but better than cluster index. Our grouping model is taking minimum execution time. Its maximum execution time is near to 2000 but remaining indexing is taking more than 3000 ms time.

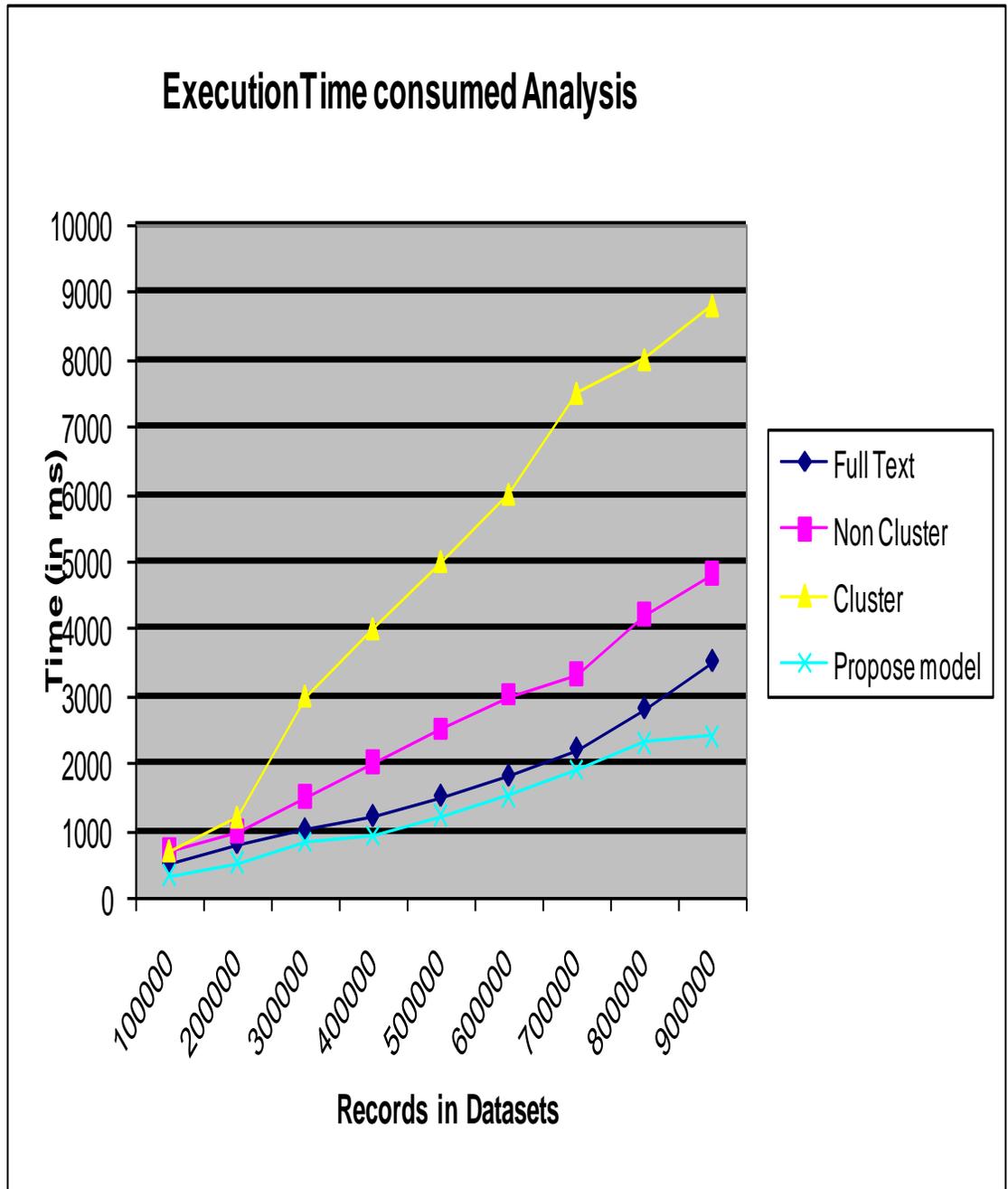


Figure 4.13 – Execution time consumed analysis

EXACT MATCH

In SQL, 'where' keyword is used for searching of exact keyword in record. The full text time consumption is increasing rapidly but on the contrary, Non Cluster index time consumption is much less nearly 400 ms approx. Cluster index gives better results in searching of exact matching of string and results in good performance even when the records in datasets increase. But here also our model takes less time in exact match query.

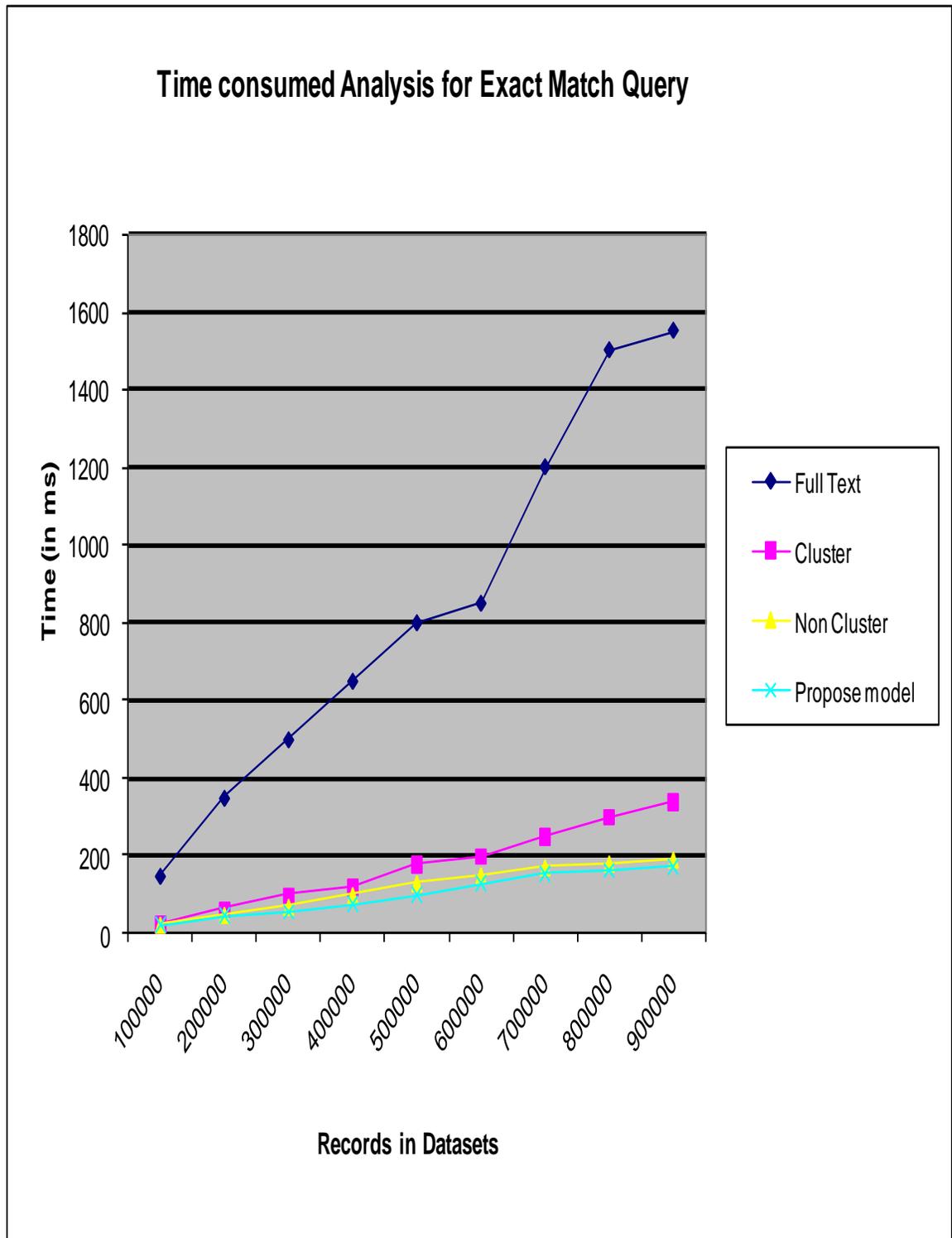


Figure 4.14 – Time consumed analysis for exact match query

4.2.2 Coordination and Communication

According to theoretical study, coordination and communication between the users of data warehouse and training of the data warehouse users are the keystone for the victory of any data warehousing project. Without proper communication and coordination, users do not know what is meant by the system and how to take results from it. In this case, the pricey data warehouse project fails and all the investments are lost.

The clear and consistent communication and coordination of company-wide warehouse goals and policies fosters employee participation on three critical issues: First, it reinforces the front-line employees' contribution of information to the warehouse. Second, it encourages information sharing to support ongoing business activities. Third, it inspires middle management to use the data warehouse to inform key stakeholders regarding decisions, and new projects.

There are nearly 350 data warehouse users in BSNL, common being the business users having little knowledge about the data warehouse. In the beginning of the data warehousing project, a large number of the users have no idea about how to operate the data warehouse and how to take decisions with the help of the data warehouse. A swap of vision, knowledge, and ideas between the industrial users and the technological persons has seriously helped business users in having a better understanding of the system and they can use the system according to their requirements and necessities. The data warehouse experts within BSNL pass on their knowledge of the system to the business users by working out the concept of coordination and communication.

General Manger further detailed that BSNL has not organize any communication programs. So to get better result there should be a separate communication program having a communication team, its own funds, principles, and resources such as facility location and digital equipment. Instead they have merged it with the education and training program. In analyze my theoretical findings and the empirical findings by comparing with each other. I try to find out what are the similarities and differences between the theory and the real world system.

They have pursue a very effortless way for the communication process by just issue a Booklet amongst the company employees containing all the information that is useful for any experts employee. But for the ordinary employee booklet is helpless in practical work if the data create problem then how can a booklet solve that problem. Secondly the members of the core team are in regular contact and coordinate the non-technical employees and junior employees by keeping them informed about the capabilities and functionalities of the data warehouse. But they not solve their regular problems that are a great problem.

4.2.2 Education, Training and Documentation

According to the General Manager, The training program implemented by BSNL is very useful. It gives a chance to each employee in the organization to gain knowledge of the tools and techniques of data ware house according to his needs. Employees who don't have the knowledge of databases or information systems can take up. Employees with some knowledge of computers and information systems can improve their skills and knowledge by taking participation in these workshops and training courses. After these

workshops and training courses they can control the data warehouse front end tools and take the desired output from the system. Top management and middle management level courses are meant for data warehouse administrators so they can better manage the data warehouse and provide the help services at any time.

As we noted in theoretical part that one day or long training at the vendor site is not enough for an average user of the data warehouse. The tools used to extract information from a data warehouse are extremely sophisticated. Often users get confused by the overload of information or forget the information before having a chance to use it. It is imperative that procedures and programs be implemented that can provide continuous help and assistance on the data warehouse and the front end tool.

The main role of the data warehouse support and protection group is the problem resolution when some problem is encountered in the data warehouse. A Help and support Center acts a coordinating body for not only collecting and logging problems with the data warehouse environment but also determining where future requirements may lay.

The training program implemented at BSNL is some like to the findings of the theoretical part . In the theory I find out that there are three levels of data warehouse staff in any organization. So BSNL should give training to their employee as their skill and their level because same training for all employees is not useful.

The training is conducted workshop for 5 to 15 days depend of employee's position and work.

In these workshop or training programs, each and every employees can any kind of interaction with the data warehouse. These workshops made to take

up some data ware house and application related trainings and industry level courses.

The training is conducted same for all the employee of the same position But problem is that in real life all employees skills are not same some employees little or no knowledge of computers and other are can operate a computer system, have basic knowledge of databases and networking and can make reports using the computer systems then same training is not useful for all the employees. It is better that give them training as their skill.

4.2.3 Help and Support

BSNL has implemented formal help and Support Counter as was found in the empirical part. One of the reasons is that currently the data warehouse is not very huge one, and only 350 employees are using it. The data warehouse Project Manager at BSNL has assigned the responsibility for providing Help and Support Counter services they have centralized IVRS (in each Zone) which have to solve all the problem of the employees.

The help and support team member have one IP phone and PC connected over an IP network to the data center. The data is displayed on the screen of the computer and the IP phone provides the Voice communication with the employees. The Members of the help and support team are trained at advanced level so they can solve any general query regarding the front end tools, data collection and report generation etc. If they can't solve any problem, it is reported to the Technical Team Lead, who decides how to act in the situation.

employees can be contacted using direct phone, email, or in person. The users can also record their reply regarding data warehouse through the local web based portal of the company. But problem is that employee can't tell their

problem in proper language on mail and phone and problem came related to data e.g. (Technical challenges moving, integrating, and transforming data from disparate environments, Short load windows, long load times, Source systems missing certain critical data) which can't be solved by phone or email.

The help and support team is functioning in a similar way found in theory, so as to help the business users, provide 24X7 supports, record and solve any problem etc. They are also taking down the troubles and their answer for future suggestion. The most universal troubles are reported to the Project Manager for improvement to the data warehouse. In the theory it is said that there should be some problem management team that should solve any problem in a combined way by helping each other. Same is the case in BSNL where there is problem management team comprising of Project Manager, Technical Team Lead and the Support Team Lead.

Regarding role of Help and Support Counter, it was found that Help and Support Counter solves a number of issues including, security and sign in, access levels for data, data quality management problems etc.

In the theoretical part it was found A Help and support Center acts a coordinating body for not only collecting and logging problems with the data warehouse environment but also determining where future requirements may lay. A Help and support Center can help to focus on these customer relationships by increasing time on customer Help and supports. In order to ensure success one needs to develop a support structure and plan an approach.

When people are using the system, the questions flow. If there are no questions than it is likely that no one is using the system. The question asked could be about the validity of the data itself, how to use the data, what calculations make sense and what levels of aggregation are valid, how to pick a report, how report measures were calculated, how to use the applications, how to change an application, and how to build and own application etc. and solve problem of the customer

4.2.4 Management of data flow on network

As the main industry of BSNL is telecommunications, it hasn't create any issue in managing the network. In the theory I found out that if the data warehouse is composed of components from different platforms, network management be a difficult part but at BSNL, as far as data warehouse is concerned all the hardware and software are from the same vendor so network management is not a complicated task but the operational systems are not from the same vendor, therefore BSNL has implemented a state of the art latest high speed network to connect the data warehouse with the source systems and the users.

This high speed link greatly reduces the time required for data loading from the source systems into the data warehouse. Secondly users get quick response from the data warehouse for their queries.

In the theoretical part it was found that there should be some person having good knowledge about networks and technology in the organization. The networking hardware is always increasing in numbers with LANs, WANs, hubs, switches, routers and multiplexers.

Users always want to access internet based data sources along with the corporate data, requiring even more bandwidth and network management resources. There should be some knowledgeable person in the organization who could handle these issues.

This is a complex area and if the staff members do not utilize the tools and associated methods frequently enough, they do not build up enough experience to become experts. So some companies find it cost effective to use outside service providers who specialize in this area to help them identify their best options and sometimes, implement the recommendations. Such firms can supply network planning, design, implementation, management and monitoring services, either remotely or on site.

Some integrated tools are required to assist data warehouse team or the network management team in monitoring the network performance At BSNL there is a team of network engineers that is responsible for managing the whole LAN of BSNL head office.

This team also takes care of the data warehouse link. The team is using state of the art network monitoring and management tools as was found in theory. Members of the network support team are qualified engineers and they can operate the tools to full extent for network monitoring.

4.2.5 Loading and Cleansing Data

Loading and Cleansing is the process which determines what data should be take out from the source systems, how it should be changed into a format that is acceptable for a data warehouse, cleaning the data that is not

useful and finally the loading of data in the data warehouse.

In the theoretical part it was found that for the work of loading and cleansing there should be an ETL architect in the business, responsible for making sure strength, survival and consistency in the Loading and Cleansing process. Loading and Cleansing is a set of processes for the provisions and protection of Business Intelligence architecture and strategy.

Time and thought are required to ensure the best architecture for its various components as well as for the selection of appropriate software tools and procedures within each component. Ongoing Business Intelligence development demands a flexible, scalable and easily maintainable environment that can only come from an architected approach. This architect should understand all the workings of the operational systems and the data warehouse. Same is the case in BSNL. Although the title for ETL architect in BSNL is Divisional Engineer who are Team Leader, but the responsibilities and duties are the same. Divisional Engineer is responsible for creating a flat and perfect data warehouse.

This engineer ensures that various functions of loading and cleansing perform as one unit. Divisional engineer with team of sub divisional engineers at BSNL is fulfilling all the duties of the ETL architect described in the literature. With the help of his support staff and sub divisional engineer the divisional engineer completes the work of loading and cleansing including recognizing data to be extracted, writing code for data extraction and communication, storing data in a temporary database, transforming data,

cleaning data, and loading data in the ETL functions reshape the relevant data from the source systems into useful information to be stored in the data warehouse.

But if source data taken from various sources is not cleanse, extracted properly, transformed and integrated in the proper way, query process these type of problem not solved by local level these are the internal problem which is hidden from the user. Without these functions, there would be no strategic information in the data warehouse. Although many loading and cleansing facilities are presented by the data ware house but still there is a requirement for ETL architect to take out these susceptible tasks.

4.2.6 Software & Hardware

BSNL has adopted the policy to keep its systems up to date, whether it be data warehouse, whether it be the network, whether it be the hardware or anything else. BSNL always tries to keep itself in line with the latest technology in confirmation with the findings of the theoretical part.

Regarding hardware issues BSNL has an MOU with the vendor who is responsible for all the hardware and software updates. There is a rising development in associations where a deal is struck for hardware or software guarantee with the vendors. Organizations are no longer keeping the hardware or software guarantee processes within the organization saving them from hire employees and spending money.

BSNL has MOU with its data warehouse vendor to provide necessary renew for its software as well. In case there is any difficulty in the system it is inform to the vendor. The vendor after checking the difficulty provides its solution by

replacement or other solution.

In the theoretical part it was initiate that the storage system should be able to read data back fast, but not at the expense of the security of payroll or other sensitive data. If see bottlenecks in current system, make sure to place requirements on both the software and hardware.

Neither one of them by themselves can solve all bottlenecks but according to Survey and interview initiate that data warehouse vendor didn't replace both hardware and software they change only one of them therefore problem of bottlenecks not solved properly and occur again after some time that affect the performance of the data ware house.

A system is balanced when the storage sub-system is capable of reading, writing and moving through the entire storage fabric – enough data to the database servers to have the CPUs adequately loaded. In other words, neither the IO capacity across the network, nor the bandwidth within the storage subsystem, or the CPUs should be a constraint to the system.

According to Survey and interview found that data warehouse vendor take a lot of time update the software and hardware so Results are gaps in available technology and software, leaving users frustrated and their needs unmet. To overcome these problems warehouses needed to get their software and hardware updated in a timely manner to avoid any shortcomings in performance.

4.2.7 Materialized view

From the interview with General Manager at BSNL I found out that there is nothing much a user can do about materialized view. Most of the functionalities related to materialized view are provided in the system by the vendor. Users just need to have knowledge about when to use those functions and when to create and refresh views. In a data warehousing environment users queries need to be very professionally and carefully written as some tables of the data warehouse are very huge and queries posted against these tables could days or weeks to complete.

Users of the data ware house should not access base tables directly; instead there should be views for those tables, which needed to be accessed. In this way users can get quick response to their queries from the data warehouse. Views are commonly built on tables that are accessed frequently and have large data sets. The commonly used strategy for materialized view by the industry is incremental where a view is updated incrementally instead of refreshing the view from scratch and re-running the view query. The use of materialized view policy depends on the type of data and its effectiveness on the overall query result.

In theoretical study originate that Materialized views provide a framework within which to collect information into the warehouse from several databases without copying each database in the warehouse. Queries on the warehouse can then be answered using the materialized views without accessing the remote databases. Provisioning, or changes, still occurs on the remote databases, and are transmitted to the warehouse as a set of modifications.

Data warehouses usually contain a very large amount of data. The main problem which evaluate in interview and survey that when any report call by higher authority which need to use highly proficient access methods and query processing techniques that not found in Materialized views. In this scenario it is very important to answer queries efficiently therefore it need to use highly efficient access methods and query processing techniques. It is an important physical design decision to decide which indices to build and which views to materialize. A data warehouse stores integrated information from multiple data sources in the form of materialized views over the source data. The data sources may be heterogeneous, distributed and autonomous.

The data change by any source, the materialized views at the data warehouse need to be updated accordingly. The process of updating a materialized view in response to the changes in the underlying source data is help and supported view maintenance. The view maintenance problem has evoked great interest in the past few years. This view maintenance in such a distributed environment gives rise to inconsistencies since there is a finite unpredictable amount of time required for propagating changes from the data sources to the data warehouse and computing view updates in response to these changes.