

Application of Resource Migration on Housing Fund Archives System

Changxing Qi

Software College, Shenyang Normal University, Shenyang, Liaoning, 110031, China

e-mail: qicx@163.com

Abstract

This paper analyzes the feature and necessity of resource migration, and summarizes the commonly methods among different systems. According to the requirements of the systems in the housing fund center, an architecture model based on resource migration is designed. The process of migration, the format and specification of the data packet are described according to the architecture levels and the process. After the migration application is developed, the operation data of the system are analyzed and further suggestions are given.

Keywords: resource migration, data packet format, SQLLDR, transaction bus

Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

With the development of information technology and the popularity of the enterprise information, The explosive growth of data information poses an enormous challenge to the management of information storage [1]. At the same time, the data in the enterprise systems independently become isolated islands of information. With the construction of enterprise information for the further development, migration and sharing of resources between the various heterogeneous data and files and other resources become an urgent task in the construction of enterprise software systems.

In the Housing Fund Management Center, business management system and archivessystem have been built. However, the two systems have its own independent database systems and software applications, information with each other independent, information sharing cannot be achieved. The target of the system is to importing the data and files in the business system to the archives system, and the mapping between the original file resources with relational data must be reserved. Data migration using the form of incremental import, that is to say, every day with the business system data is updated, the updated data is imported into the archives system. Since the two systems have different data structures. The middle part of the data logic processing, particle size conversion, is an important consideration in the migration process.

Literature [2] proposed the thought of hierarchical storage management (HSM), This provides a guideline for the solution of this problem, how to migrate storage resources is a key technology in the hierarchical storage management [3].

This paper designs an application based on the heterogeneous enterprise data and information resources, realizes the functions of extracting the source data into a data packet, data logic processing and the target data import and so on. As a data migration example of the archives system, designs and realizes a migration application with open interfaces and error correction capabilities for information resources.

2. Method

Currently, data migration commonly used methods are the following:

- 1) Custom-developed program. Establish connections to the source database and target database, retrieve the data from the source database one by one and insert it into the target database, this method is less suitable for data logging or desktop database migration. This method is suitable for data migration of data-less recording or desktop database.

- 2) Import and export components of the database. Such as ORACLE import and export tools, export the objects from source database, and import the backup data files into the target database. This method is suitable isomorphic database and the structure of the database table is basically consistent.
- 3) Features of the database. Such as the use of Oracle database pipeline technology. This method is applicable to homogeneous database and the And the stability of the network environment must be guaranteed.
- 4) The third party middleware. increasing real-time data import, such as MQ of IBM.This method is applicable to the case of the data volume is not very large scale.
- 5) Processing components of database. Such as SQLLDR of ORACLE.Exported the source data in text mode, after technical processing, import the data into the target database.
- 6) Special software. Such as IBM IBM DB2 Migration Toolkit, this tool is usually the development of database vendors and expensive.
- 7) Migration software developed personalized. It is usually business-oriented personalized, and the development cycle is long.

In this system, since the file data generated every day, it is necessary to adopt an incremental approach to migrate data. In accordance with the requirements of the housing fund business system, the system must extract new data and .jpg files from business systems every day to form of the migration package, this system will learn from the above fifth method and combine with the characteristics of the system to design the software system.

3. Requirements

In the housing fund business system, the data need to be migrated to the archives system include raise funds data, other warrants data, personal loans data, loan cancel the account data, entrust bank divert data, Borrower information changes data, loan period shorted data, loan extension data, the mortgagor information changing data, Collateral changing data and so on. Each data includes relational data stored in a database table and the associated pictures. In the archives system, each business data stored in a table, and there are two tables which saves the pictures information associated with the data table, one holds summary information of the pictures, the other holds the detail information of the pictures. Therefore, during the data migration process, it must take the relevant business data aggregated into one table firstly, and this is necessary to ensure data integrity, accuracy and consistency.

4. Design

The design includes the system structure, process description, transaction bus, monitor windows,data specification, packet format,communication protocol.

4.1. System Structure

Following Figure 1 is the overall structure of the system.

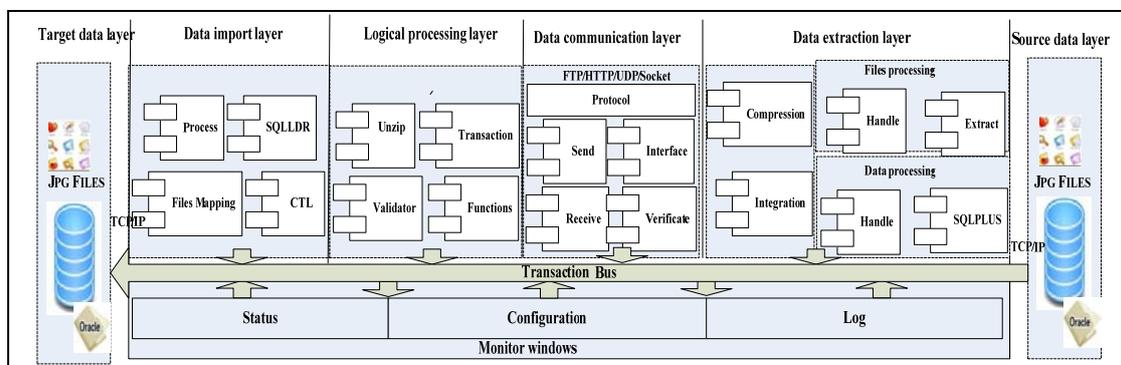


Figure 1. System Structure

4.2. Process Description

From the Figure 1, we can see that the structure contains six layers: The first is source data layer, it saves housing fund day-to-day production data and image files; the second layer is data extraction layer, it consists of two functions, one is extracting data into the txt files from database by ORACLE SQLPLUS according to the specified format, the other is copying image files: define the directory named with the archives number and copy the image files into the directory organized by date. Since production system and the archives system is located in a different location, this layer is deployed on the server of business system; the third layer is data communication layer. According to a given data transfer protocol, it transmits the data packet to the archives system; the fourth layer is logical processing layer. Its main function is to decompress packets received and complete data format validation, and given rules necessary logical processing; the fifth layer is data import layer. According to the information of the processed data and image files, it generate the CTL files to control the import process of ORACLE by SQLLDR [6]. The data in text files are imported into the archives system data table. And the pictures are copy to the server, which are saved in the directory named by current date. so the relation between the data and the pictures are built; the last layer is target database layer which saves the data and pictures of the archives system [4-6].

4.3. Transaction Bus

While the system is running, the whole process is a unified transaction including extracting data, generating data packet, communication, unzip, logical process, data import. If problems happen during any steps, the whole process will be rollback, this means the operation failed. It need to restart the process. It ensures data integrity and security during data migration process .

4.5. Monitor Windows

The system provides three monitoring interface: log monitor, exception remind monitor, dynamic running processes monitor. At the same time, the configure file can be examined to regulate the system.

4.6. Data Specification

The data that the archives system needed is classified and coded as Table 1, In the migration packet, the data files are saved as a TXT text, the first three letters of the file name are defined by the ID which have been given in the Table 1. Different encoding corresponds to different data tables.

Table 1. Data File Encoding

ID	Content
D02	raise funds data
L04	other warrants data
L08	personal loans data
L09	loan cancel the account data
L10	entrust bank divert data
L11	borrower information changes data
L12	loan period shorted
L13	loan extension data
L14	the mortgagor information changing data
L15	collateral changing data
D02	raise funds data

4.7. Packet Format

The format of migration data packet is as Figure 2.

In the target database, the tables include raise funds information, loan information and so on. And every type of data contains three tables: data table, pictures summary table and pictures detail table. In the pictures detail table, all kinds of pictures information such as photocopies of the house property certificate and identity card are recorded in it. So the packet composition are designed as follow:

- 1) All summary information of the pictures are recorded in the ScanFliesNum.txt.

- 2) All detail information of the pictures are recorded in the ScanFiles.txt.
- 3) All pictures are save in the folder which is named with archives number.

Data information as saved in the text files named with two parts:ID+yyyymmdd.txt, and the ID is defined in Table 1.

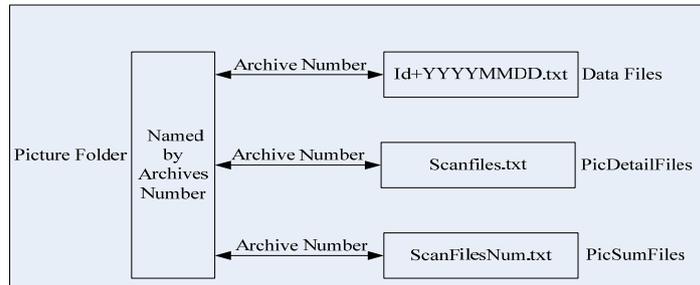


Figure 2. Data Packet Format

4.8. Communication Protocol

DataFlag.txt is the data communication interface, and composition of the file is designed as Figure 3.



Figure 3. Composition of DataFlag.txt

The packet communication process is achieved by identification of three elements in DataFlag.txt.

- 1) Flag: if the value of it is 0, it means the data packet has been received successfully. On the contrary if the value of it is 1, it means failure.
- 2) File: the name of the data packet.
- 3) Infor: the prompt information.

5. Realization

In this paper, the system is designed and realized for data migration between the archives system and the business system. The target database is ORACLE 9I. The task of the system is extracting data from source database and transmit the compressed data packets from the business system to archives system. After decompression the data are processed according to specified business rules. And in accordance with the set of insinuate relationship between source data and target data, the oracle data import control files are generated. Then according to the control file. Use the components of ORACLE's SQLLDR to perform data files importing operation [6].

5.1. Interface Modules

These modules are used to operate the application for the users.

- 1) Configure file: Set the system operating parameters and source data to a target data mapping rules, start this program (this program can also be set to started at a particular time), such as the username and password of database, the directory of the data packet, the directory of log file and control files, the directory of zip application program files.
- 2) Log files: Query the log files including whole records, include the time and operation and prompt message information.
- 3) List of datapacket: Query all data packet files which have been migrated.

- 4) SQLLDR log: Query the information that the database generated during the process of import data by ORACLE SQLLDR.
- 5) MidFiles: Query the files which generated during the process of migration, such as control files, unzip files. If the process of migration has succeed, these files will be deleted. So the function of this module is error correction reference.

5.2. Core Functions

The primary functions in the application as following:

- 1) The function of WriteSysLog is to Generate the system log files.

```
function WriteSysLog(key: string; content: string): boolean;
function WriteSysLog(key: string; content: string): boolean;
var
  dates, date: string;
  MyIniFile: TIniFile;
begin
  try
    try
      dates := FormatDateTime('yyyy-MM-DD hh:mm:ss AM/PM', Now);
      date := FormatDateTime('yyyymmdd', Now);
      dates := 'Time : ' + dates;
      MyIniFile := TIniFile.Create(IncludeTrailingBackslash(GetCurrentDir) + date
+ 'DataSysLog.ini');
      MyIniFile.WriteString(dates, key, content);
    except
      end;
    finally
      MyIniFile.Free;
    end;
    Result := true;
  end;
end;
```

- 2) The function of ZipPross is to Zip or unzip the data packet rar files, and generate the directory of the data files.

```
function ZipPross(key: integer; ZipName: string ; ): boolean;
try
  Command := RarExtPath + ' x -ep1 -y ' + DataFrom + '\ ' +
LB_Rar.Items.Strings[LB_Rar.itemindex] + ' ' + HandDataTrans + '\';
  if winexec(pchar(Command), sw_hide) <= 31 then
    begin
      RichEdit1.Lines.Add('alert' + LB_Rar.Items.Strings[LB_Rar.itemindex] + '
failure, try again!');
      UpdateDataFlag(leftstr(LB_Rar.Items.Strings[LB_Rar.itemindex], 4), '1',
LB_Rar.Items.Strings[LB_Rar.itemindex], 'unzip failure');
      WriteSysLog('file:' + LB_Rar.Items.Strings[LB_Rar.itemindex], 'unzip
failure ! ');
      exit;
    end
  else
    begin
      RichEdit1.Lines.Add('file: ' + LB_Rar.Items.Strings[LB_Rar.itemindex] + '
unzip success ! ');
      WriteSysLog('File' + LB_Rar.Items.Strings[LB_Rar.itemindex], 'unzip
success ! ');
    end;
  except
end;
```

```

    on E: Exception do
    begin
        WriteSysLog('file:' + LB_Rar.Items.Strings[LB_Rar.itemindex], 'Exception ! ' +
'Name:' + E.ClassName
        + #13#10 + 'Infor:' + E.Message);
        RichEdit1.Lines.Add('Alert : file ' + LB_Rar.Items.Strings[LB_Rar.itemindex] +
'unzip Failure ! ! ' + 'Name:' + E.ClassName
        + #13#10 + 'Infor:' + E.Message);
        exit;
    end;
end;

```

- 3) The function of WriteCTLFile is to Generate the control files to handle SQLLDR operation.

```

function WriteCTLFile(ImpFileName: string; TableName: string): Integer;
var
    Filename: string;
    f: TextFile;
    a: integer;
    s:string;
begin
    try
        Filename := DataTrans + '\ + 'CTLOperate.CTL';
        AssignFile(f, Filename);
        if FileExists(Filename) then
            Reset(f)
        else
            Rewrite(f);
            Append(f);
            Writeln(f, 'load data');
            Writeln(f, 'infile ' + "" + ImpFileName + "");
            Writeln(f, 'append into table TABLENAME');
            s:= 'fields terminated by ' + "" + "" + '( GDH ,YWLX ,';
            s:= s+' BM, MC , SL ) ' ;
            Writeln(f,s );
            CloseFile(f);
            WriteSysLog('CTLFile Generate success', ' CTLOperate.CTL');
            a := 0;
        except
            WriteSysLog(' CTLFile Generate Failure', ' CTLOperate.CTL');
            CloseFile(f);
            DELETEDFILE(DataTrans + '\ + CTLOperate.CTL');
            a := 1;
        end;
    end;
end;

```

The additional funtions:

- 1) function DataExtract(key: integer):string;
Extract data from source database and generate the data packet.
- 2) function UpdateDataFlag(Key1: string; Key2: string; Key3: string; key4: string):
boolean;
Update the elements in the DataFlag.txt file.
- 3) procedure MyCopyFile(FromDir, ToDir: string);
Transmit the image files to target directory.
- 4) function Looptar(DataCome: string): integer;
Get all files in the current directory:DataCome
- 5) function ProcessData(DataCome: string): integer;
Check and handle the data from data packet.

- 6) function UpdateFilst(Content: string): boolean;
Update filelist in the tarfilelist.txt

5.3. SQLLDR

SQLloader is one of the important tools of Oracle to migrate data between the different database system, it can migrates the data stored in text format smoothly into the oracle database by using the SQLLDR commands.It includes two parts,the one is the control files and the other is the SQLLDR commands for excuting.

The contents of the control file as follow:

- 1) load data: identifier or commands of the control file
infile'XXXX.TXT' –the name of the data txt file.
- 2) append into table TableName:append records into the database, it can include
append, insert, replace and truncate.
- 3) fields terminated by',' : field interval in file.
(id,username,password,.....j): columns in the database.

The format of SQLLDR commands for excuting as follow:

SQLLDR USERID= UserName / Password@ Sidlog= logfile CONTROL=CTLFile;

The UserName, Password and Sid is to log on the database and logfile is to record the migration and the CTLFileis the control file.

5.4. Trigger in ORACLE

In the files system, each table's primary key is completed by the sequence in Oracle; in data table, some special fields, such as date of operation, operator, and some identification field, will be set default value. The assignments of these are completed by insert triggers in ORACLE. When the system import the data into the target database, the triggers are executed.

6. Conclusion

This program can ensure the resource data migration between the archives system and the business system for the housing fund. It runs stability now and has the value of popularization. Meanwhile, with the operation of the system, the data packet increases, the speed of the system operation is also gradually slowed down, found by the analysis of the operating data, The run time of the system, in addition to the relationship with the system to the hardware configuration, but also with many elements [7]. Set the execution time of an element as t_i , and the proportion of the element as a_i , and the final time is shown as formula (1):

$$T = \sum_{i=1}^m t_i a_i \quad (1)$$

From the Analysis,the capacity and organization of the data packet itself are the most important elements to affect the speed of system running. As shown in Figure 4, with the data packet size increases, the time needed to run the system nonlinear slowed down. Meanwhile, when choose the basic data packets at the same capacity, but the number of data files of the package is different, the more numbers, the slower the system is running, as Figure 5.

As a result,to improve efficiency, it is a method to reduce data packet size or the number of txt files in the data packet. This also requires to design better algorithms to reduce data redundancy, while taking advantage of the procedure of database to improve the efficiency of the system.

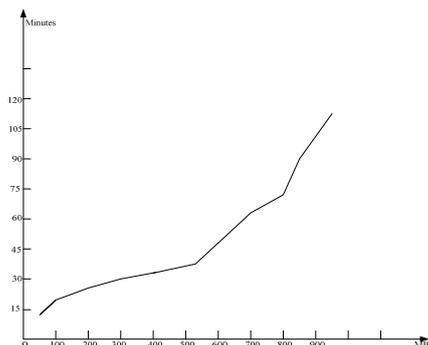


Figure 4. Relation of Time and Capacity

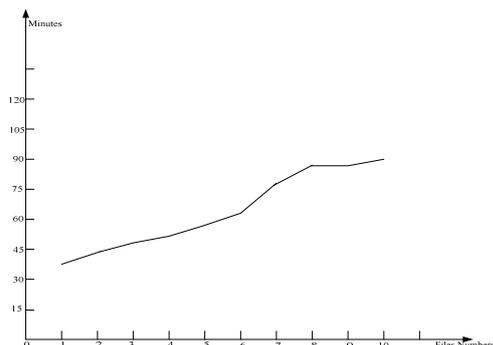


Figure 5. Relation of Time and Numbers of Files

References

- [1] Gibson T, Miller EL, Long DDEI. *Long-term file activity and inter-reference patterns*. The Proceeding of the 24th Int Conf on Technology Management and Performance Evaluation of Enterprise Wide Information Systems. Anaheim. CA:CMG. 1998: 976-987.
- [2] ZHAO Xiao-nan, LI Zhan-huai, ZENG Lei-jie. *A Hierarchical storage strategy based on block-level data valuation*. The proceeding of the 4th International conference on Net-worked Computing and Advanced Information Management. 2008: 36-41.
- [3] QI Chang-xing, DU Qing-dong. Call Center Platform Based on Application Gateway. *Journal of Shenyang Normal University (Natural Science)*. 2008; 26(40): 438-441.
- [4] CONG Hui-gang, REN Qing-dong, LI Tian-yang, YUAN Man. Implementation of Large Database Migration Tool Based on Metadata-driven. *Science Technology and Engineering*. 2011; 11(100): 2353-2356.
- [5] XiongHuaping, Li Lijiao, Chen Fuping, RongHailiang. Research And Application Of Data Migration System Betweenhuge Heterogeneous Databases. *Computer Applications and Software*. 2012; 29(7): 178-181.
- [6] ZHANG Xi. Realization on data migration function in parallel file system. *Journal of Beijing Information Science and Technology University*. 2012; 27(5): 77-80.
- [7] XuBao-yu, LIAOTao. Design of data migration time in tiered storage. *Computer Engineering And Design*. 2013; 34(2): 725-728.