UDC 004.056.3

I.S. SKARGA-BANDUROVA, M.V. NESTEROV

Technological Insitute of East Ukrainian National University, Severodonetsk, Ukraine

IMPLEMENTATION OF HIGH AVAILABILITY HEALTHCARE INFORMATION SYSTEM ARCHITECTURE

The primary challenge of developing an effective health information system is maintaining a satisfactory balance between health care information system safety and health care data and information availability. Healthcare data safety is defined, the need for establishing high availability architecture is examined, and a variety of the components of performing replication, recovery and backups for healthcare data are discussed. In addition, the components of Oracle Active Guard as technologies for guarantee high availability of data, quick loading and finding data in huge storages like health care information are outlined.

***Keywords**: healthcare information system, business intelligence, high availability, backup, data protection, Oracle, Active Data Guard, GoldenGate.*

Introduction

Today more and more information in Ukrainian health care organizations is transmitted, maintained, and stored electronically. Healthcare information systems (HIS) are becoming more common, and as we see, even primarily paper-based health care information systems contain data and information that have been created and transmitted electronically.

At the same time, the new threats to health care information have appeared and, as a consequence, show up special needs to ensure healthcare data safety and high availability to various stakeholders, such as patients, physicians, hospitals, government and the third-party institutions, with different interests and concerns.

Authors [1] divide threats to health care information systems into three categories:

- Human threats, which can result from intentional or unintentional human tampering;
- Natural and environmental threats, such as floods, fires, and power outages;
- Technology malfunctions, such as a drive that fails and has no backup.

There are different ways to decrease third categories threats through design HA architecture [2 – 4], the paper continues study to develop effective HIS configuration [5] with a look at the following topics, including examples of actual practices: Replication; Data Recovery; Backups. One aspect of this study is developing high availability architecture for implementation maintaining a satisfactory balance between HIS safety and health care data and information availability.

Availability is the degree to which a system is accessible on demand. It is measured by perception of an applications end user. And the main characteristics of high available solution are reliability, recoverability, timely error detection and continuous operation. For realization HA solution for HIS Oracle product line was used. The main concepts of Oracle's HA design principles are [6]: completeness (Minimize all planned and unplanned downtime, offer a standard validated platform for maximum availability); application oriented (protect and recover application objects, enable online application changes); scale-out model (low-cost commodity hardware, all components active in a grid infrastructure); integrating and simplicity (built-in HA with pluggable components, automatic-eliminate manual processes). According to these principles HIS system includes [7]:

- Online database for electronic health records (EHR), treatment protocols, decision making, and etc. – OLTP PROD.
- Online analytic database for reporting and other BI futures – OLAP PROD.

OLTP and OLAP Database must be open and available for users 24 hours a day, 7 days a week.

1. Replication

To find a good solution concerning data replicating two Oracle solutions were considered: Oracle Streams (OS) and Oracle GoldenGate (OGG) [8]. Table 1 presents comparing OGG and OS.

Table 1

Comparing OGG and OS

Oracle GoldenGate	Oracle Streams
Out-of-the-box flexible solution	Requires API based stored DB packages
Broad heterogeneous support	Work with Oracle Database
Excellent reference base	Few large reference
High performance in high transactional environment	Problem scaling in high-throughput environments
Reliable architecture	

The OGG technology has such advantages: Better ROI overall; OGG is lower cost to implement and maintain Enterprise-wide solution; OGG easily expands to new use cases Reliable, flexibility, high performance. That's why OGG was chosen. On fig .2 proposed schema with OGG is presented.

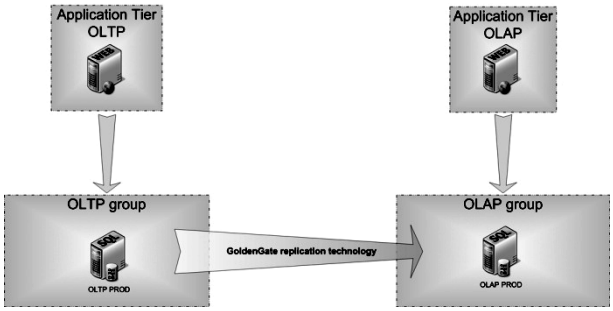

Fig. 2. Proposed OGG configuration

2. Data Recovery

For Data Recovery and availability Oracle Active Data Guard [9] was included.

Active Data Guard provides the management, monitoring, and automation software to create and maintain one or more synchronized replicas (standby databases) of a production database (primary database). An Active Data Guard standby database is an exact copy of the primary that is open read-only while it continuously applies changes transmitted by the primary database. An active standby can offload ad-hoc queries, reporting, and fast incremental backups from the primary database, improving performance and scalability while preventing data loss or downtime due to data corruptions, database and site failures, human error, or natural disaster. Also Protection, Availability, ROI, Performance, Confidence. On fig.3 such schema is deployed in detail.

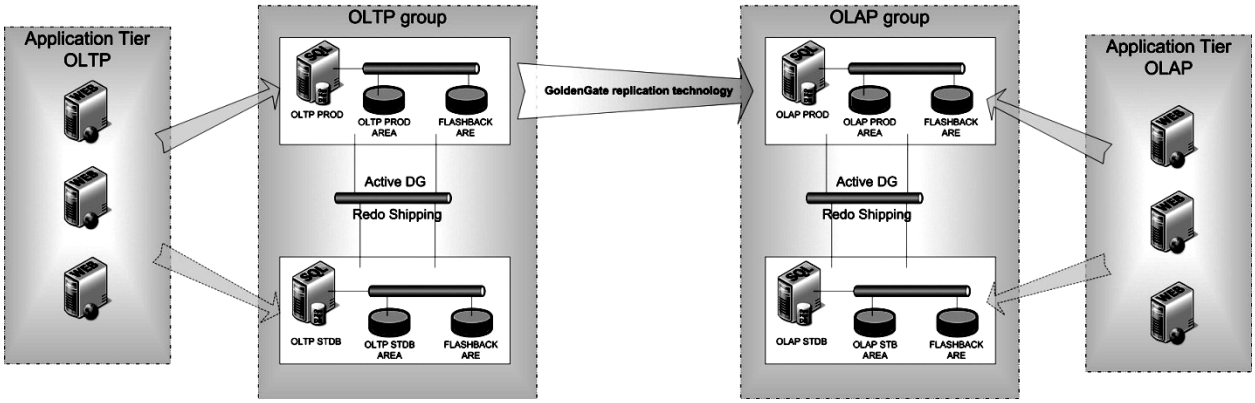

Fig. 3. Proposed schema with OGG and ADG

There are different recovery components. A number of structures and events in the database directly support backup and recovery operations.

Control Files. The control files maintain the list of database files in the database, along with a record of the most recent database backups. The control file is one of the smallest, but one of the most critical, files in the database. Recovering from the loss of one copy of a control file is relatively straightforward; recovering from the loss of only control file or all control files is more of a challenge and requires more advanced recovery techniques.

Checkpoints. The checkpoint background process controls the amount of time required for instance recovery. A checkpoint occurs automatically every time a redo log file switch occurs, either when the current redo log file fills up or when you manually switch redo log files. The DBWn processes in conjunction with CKPT routinely write new and changed buffers to advance the checkpoint from where instance recovery can begin, thus reducing the MTTR.

Redo Log Files. A redo log file records all changes to the database, in most cases before the changes are written to the data files.

To recover from an instance or a media failure, redo log information is required to roll data files forward to the last committed transaction.

Archived Redo Log Files. Usage only online redo log files ensures database protect against instance failure but not media failure. Although saving the redo log files before they are overwritten takes additional disk space and management, the increased recoverability of the database outweighs the slight additional overhead and maintenance costs.

The Flash Recovery Area. As the price of disk space drops, the difference in its price compared with tape is offset by the advantages of using disk as the primary backup medium: Even a slow disk can be accessed randomly a magnitude faster than a tape drive.

This rapid access means that any database recovery operation takes only minutes instead of hours.

3. Backups

So we have high availability EHR and BI system but if disaster occurs on PROD and STDB instances, process of recovering data is not easy when two instances are down for that reason we need a backup of database. To solve that problem we use Tape Drive as a low cost solution [10]. With backup of database on Tape Drive we can rebuild our instance form backup timestamp. To maximize database's availability, performing regularly scheduled backups must be provided.

In HIS we can make a whole backup or a partial backup. Whole backups and partial backups are known as Oracle backup strategies. The backup type can be divided into two general categories: full backups and incremental backups. Depending on whether we make database backups when the database is open or closed, backups can be further categorized into the backup modes known as consistent and inconsistent backups.

The backups can be managed using operating system and SQL commands or entirely by Recovery Manager (RMAN). Many backup types are only available using RMAN, such as incremental backups. According to backup terminology we can implement [11]:

Whole database backup. A whole database backup includes all data files and at least one control file. Online redo log files are never backed up; restoring backed up redo log files and replacing the current redo log files will result in loss of data during media recovery.

Only one of the control files needs to be backed up; all copies of the control file are identical.

Partial database backup. A partial database backup includes zero or more table spaces, which in turn includes zero or more data files; a control file is optional in a partial database backup.

Full backup. A full backup includes all blocks of every data file backed up in a whole or partial database backup.

Incremental backup. An incremental backup makes a copy of all data blocks that have changed since a previous backup. Oracle 10g supports five levels of incremental backups, from 0 to 4. An incremental backup at level 0 is considered a baseline backup; it is the equivalent of a full backup and contains all data blocks in the data files that are backed up. Although incremental backups can take less time, the potential downside is that we must first restore the baseline backup and then apply all incremental backups performed since the baseline backup.

A consistent backup. A consistent backup, also known as an offline backup, is performed while the database is not open. These backups are consistent because the SCN in the control file matches the SCN in every data file's header. Although recovering using a consistent backup requires no additional recovery operation after a failure, we reduce database's availability during a consistent backup as well as risk the loss of committed transactions performed since the consistent backup.

An inconsistent backup. Although the term inconsistent back-up may sound like something we might avoid in a database, it is a way to maintain availability of the database while performing backups. An inconsistent backup, also known as an online backup, is performed while the database is open and available to users. The backup is inconsistent because the SCN in the control file is most likely out of synch with the SCN in the header of the data files. Inconsistent backups require recovery when they are used for recovering from a media failure, but keep availability high because the database is open while the backup is performed.

Also we can provide Active Duplication on Production (PROD) database OLTP or OLAP for fast DR (fig 4). In this case we don't need redo files or log files or something else.

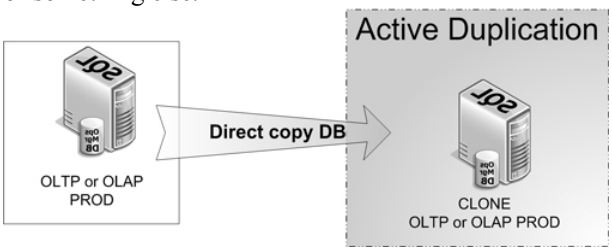

Fig. 4. Active Duplication of OLTP or OLAP PROD

But in this case we overload production database. To solve this threat we make Active Duplication on Standby instance for offload cloning the database is shown on fig. 5.

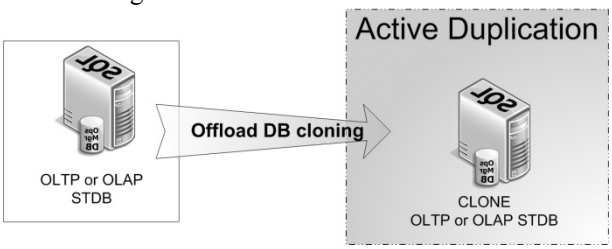

Fig. 5. Active Duplication of OLTP or OLAP STDB

And if we have lost all our working and cloned instance we can easily clone a database from backup.

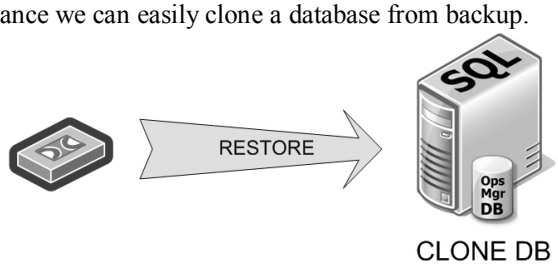

Fig. 6. Backup based duplication

And as the result we have finally high availability architecture for our EHR system and BI (fig. 7).

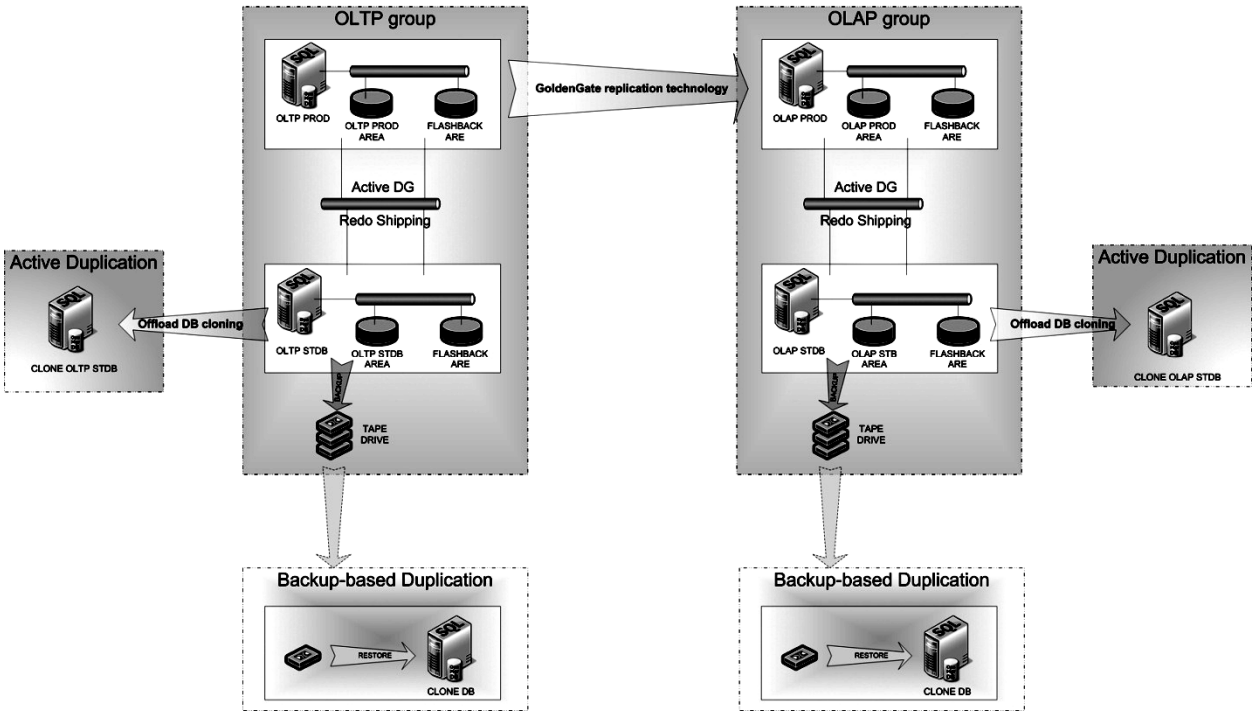

Fig. 7. HIS architecture

Conclusion

The primary challenge of developing an effective health information system is maintaining a satisfactory balance between health care information system safety and health care data and information availability. The balance between access and security should be reasonable – protecting users' rights while allowing appropriate access.

From that angle the architecture is an essential to integrating disparate applications and is a critical tool in the effort to control information technology operating costs by constraining the number of technologies supported. Presented HIS architecture can be supplemented by Oracle RAC (Real Application Cluster) for production database for more reliable solution or any other solutions with OGG and ADG.

References

1. Wager, K.A. Health care information systems : a practical approach for health care management [Text] / K.A. Wager, F.W. Lee, J.P. Glaser; foreword by L.R. Burns. – 2nd ed. – 2009. – 517 p.

2. Technology architecture guidelines for a health care system [Electronic resource] / D.T. Jones, R. Duncan, M.L. Langberg, M.M. Shabot // Proc AMIA Symp. – 2000. – P. 399 – 402. – Access mode: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2243753/pdf/procamiasymp00003-0434.pdf. – 15.03.2012.

3. Scherrer, J.R. Healthcare information system architecture (HISA) and its middleware models [Electronic resource] / J.R. Scherrer, S. Spahni // Proceeding AMIA Symp. – 1999. – P. 935 – 939. – Access mode: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2232789. – 15.03.2012.

4. Ohlmeie, N. Design and Implementation of a High Availability SIP Server Architecture [Electronic resource] / N. Ohlmeie. – Berlin: Technische Universitat, 2003 – 86 p. – Access mode: http://ohlmeier.com/work/Des_Impl_HA_SIP_Arch.pdf. – 15.03.2012.

5. Nesterov, M. Using Leading Oracle Technologies as Base Development Platform of Electronic Health Record [Text] / M. Nesterov, I. Skarga–Bandurova, A. Nesterov // Сучасні комп'ютерні системи та мережі розробка та використання: Матеріали 5-ої Міжнародної наукової конференції ACSN–2011. – Львів: НВФ «Українські технології», 2011. – С. 117 – 118.

6. Pedregal, C. Oracle Database 11g Release 2 High Availability [Electronic resource] / C. Pedregal, A. Ray // An Oracle White Paper. – Access mode: http://www.oracle.com/technetwork/database/features/availability/twp-databaseha-11gr2-1-132255.pdf. – 15.03.2012.

7. Armijo, D. Electronic Health Record Usability Evaluation and Use Case Framework [Text] / D. Armijo, Ch. McDonnell, K. Werner. – AHRQ Publication No. 09(10)–0091–1–EF, 2009. – 57 p.

8. Oracle GoldenGate 11g Release 1 Patch Set 1 (11.1.1.1) [Electronic resource]. – Access mode: www.oracle.com/technetwork/middleware/goldengate/documentation/index.html. – 15.03.2012.

9. Carpenter, L.M. Active Data Guard Hands On Lab [Electronic resource] / L.M. Carpenter, N. Kark-

hanis, J. Lee. – 2010. – 68 p. – Access mode: http://www.oracle.com/technetwork/middleware/goldengate/documentation/index.html. – 15.03.2012.

10. Encrypt, Protect and Secure Your Backup Data with Oracle Secure Backup – Technical White Pages [Electronic resource]. – Access mode: http://www.oracle.com/technetwork/database/securebackup/learnmore/osb–103–encryption–335238.pdf. – 15.03.2012.

11. Sybex OCA Oracle 10g Administration I Study Guide [Electronic resource]. – Access mode: ftp.sybex.com/4367/4367ch01.pdf. – 15.03.2012.

Поступила в редакцию 19.03.2012

Рецензент: д-р техн. наук, ст. наук. співр. В.М. Опанасенко, Інститут кібернетики ім. В.М. Глушкова, Київ, Україна.

РЕАЛІЗАЦІЯ АРХІТЕКТУРИ ВИСОКОЇ ГОТОВНОСТІ ДЛЯ МЕДИЧНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ

І.С. Скарга-Бандурова, М.В. Нестеров

В роботі визначається потреба в створенні архітектури високої готовності для інформаційної системи (ІС) охорони здоров'я, розглянуто різні компоненти для конфігурування системи і виконання основних операцій, що гарантують високу готовність медичних даних: реплікації, резервного копіювання і відновлення. Запропоновано схеми використання компонент Oracle Active Guard і GoldenGate, що гарантують доступність, швидке завантаження та пошук даних у великих сховищах, таких як ІС охорони здоров'я.

Ключові слова: медична інформаційна система, висока доступність даних, резервне копіювання даних, захист, Oracle, Active Data Guard, GoldenGate.

РЕАЛИЗАЦИЯ АРХИТЕКТУРЫ ВЫСОКОЙ ГОТОВНОСТИ ДЛЯ МЕДИЦИНСКОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ

И.С. Скарга-Бандурова, М.В. Нестеров

В работе определяется потребность в создании архитектуры высокой готовности для информационной системы (ИС) здравоохранения, рассмотрены различные компоненты для конфигурирования системы и выполнения основных операций, гарантирующих высокую готовность медицинских данных: репликации, резервного копирования и восстановления. Предложены схемы использования компонент Oracle Active Guard и GoldenGate, гарантирующих доступность, быструю загрузку и поиск данных в больших хранилищах, таких как ИС здравоохранения.

Ключевые слова: медицинская информационная система, высокая доступность данных, резервное копирование данных, защита, Oracle, Active Data Guard, GoldenGate.

Скарга-Бандурова Инна Сергеевна – канд. техн. наук, доцент, доцент кафедры компьютерной инженерии технологического института ВНУ им. В. Даля, Северодонецк, Украина, e-mail: skarga_bandurova@ukr.net.

Нестеров Максим Владимирович – аспирант кафедры компьютерной инженерии технологического института ВНУ им. В. Даля, Северодонецк, Украина, e-mail: nesxam@gmail.com.