An Oracle White Paper
April 2013

VocaLink
Oracle Exadata Technical Case Study

Overview

VocaLink designs, builds and operates world-class payment systems and ATM switching platforms. Using advanced technology, expert consultancy and unique insights into dealing with high volumes of payment transactions, VocaLink can boost efficiency, cut costs and increase the range of payment products available to customers – with total reliability. VocaLink also manages the core infrastructure for major national payment systems and has delivered three new clearing implementations in recent years in sterling, krona and euros. VocaLink's switching platform connects over 65,000 ATMs and its payment platform processes over 10 billion payment transactions with a value of £4.9 trillion.

VocaLink has successfully deployed a new national payments system in a unique partnership with a country in Europe. The new system represents a cost-effective technology renewal for a national banking system and offers a platform for continuing payment innovation. The innovative business model implemented by the new system generates economies of scale that permit more cost-effective payment services to be delivered to customers in shorter timescale.

The new electronic payment system is deployed on Oracle Database running on Oracle Exadata Database Machine and the Oracle Maximum Availability Architecture (MAA). This case study describes how VocaLink has used Exadata and MAA to achieve the high levels of performance, data availability and data protection required by mission critical online transaction processing systems.

"Our critical electronic payments service has been live on Exadata since early 2011 with 100% uptime. The service reliably processes the transfer of billions of Euros per week and achieves subsecond response times for online enquiries."
Martin McGeough — Database Technical Architect, VocaLink

Intended Audience

This paper reviews VocaLink's use of Oracle Exadata Database Machine and provides configuration details and benefits specific to the deployment being discussed. Readers are assumed to have experience with Oracle Database, familiarity with the Oracle Maximum Availability Architecture framework (MAA), and a general technical understanding of Oracle Exadata Database Machine. When referenced in this paper, in-depth background on these topics will be deferred as they are covered in other documentation and technical white papers available on the Oracle Technology Network[1]. Where applicable, links to additional technical information are provided in footnotes that accompany each section. Appendix A also provides a list of the Oracle acronyms used in this paper.

Introduction

VocaLink[2] has been a pioneer in electronic payment systems establishing a proven track record of innovation spanning a period of 40 years. It has evolved from a domestic UK supplier to a large-scale international provider of modern payment services. VocaLink securely processes over 10 billion payments a year. On a peak day its payment platform processes over 90 million transactions and its switching technology powers the world's busiest network of over 65,000 ATMs

VocaLink and Oracle

Beginning in 2000, VocaLink initiated a programme to renew its Bacsservice (Bankers Automated Clearing System). The goal was to replace the heritage mainframe technology with a highly scalable, highly available and modular infrastructure that would reduce costs while simultaneously improving

[1] http://www.oracle.com/technetwork/database/exadata/index.html
[2] http://www.vocalink.com/

performance. Bacs is an intrinsic part of the UK economy. It is used by tens of millions of consumers and thousands of organizations to simplify payments and manage cash flow more efficiently.

VocaLink conducted a thorough evaluation of potential technology providers for its next generation architecture and selected the following strategic partners:

- BEA (now Oracle) was chosen to provide the Application Server software
- Sun (now Oracle) was chosen to provide the hardware.
- Oracle was chosen to provide the database software due to the introduction of Oracle Real Application Cluster (Oracle RAC), Oracle Data Guard and Oracle Recovery Manager (RMAN) Oracle RAC provided high availability and scalable performance and Data Guard provided a reliable disaster recovery solution.
- Oracle Clusterware, Oracle Automatic Storage Management (ASM), and Oracle Flashback technologies were added to VocaLink's HA architecture following the release of Oracle Database 11g. These additional capabilities combined with using Oracle documented HA best practices, completed VocaLink's implementation of the Oracle Maximum Availability Architecture (MAA).

The subsequent success of the Bacs technology refresh programme using Oracle Database established a standard High Availability (HA) architecture that has been implemented in other strategic VocaLink systems, including

- Faster Payments Service - the first application of the VocaLink Real-Time Payments Platform
- A new electronic payment service for a national banking community in Europe

With the renewal of every service, the lessons learned, new configurations, enhancements, features and improvements are all captured within VocaLink's standard build and are applied to the next service that undergoes renewal. This process emphasizes continuous, low risk, improvement – essential to VocaLink maintaining its leadership position. This same discipline was applied to their migration to Oracle Exadata Database Machine.[3]

Electronic Payment Services for a National Banking System

VocaLink recently introduced the next-generation of their electronic payment services designed to support national banking systems outside of the UK. VocaLink's first commercial deployment of the new system also required a migration from their previous system infrastructure to Exadata Database Machine. The service automates payments for banking and corporate communities and also provides

[3] http://www.oracle.com/us/products/database/exadata/overview/index.html

access to SWIFT. VocaLink's systems are located in two UK data centers that are 100 kilometers apart. A high-level view of parties participating in the electronic payment service is provided in Figure 1.

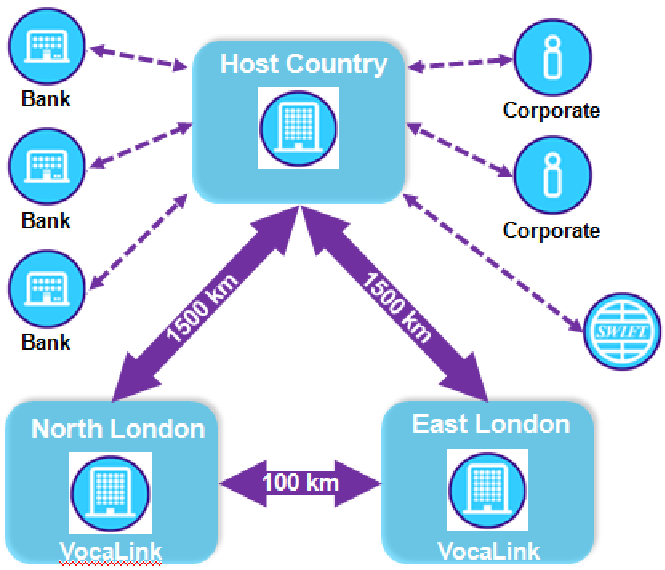

FIGURE 1: VOCALINK ELECTRONIC PAYMENT SERVICE - GEOGRAPHIC LOCATIONS

The electronic payment service is a very sophisticated system that executes multiple payment cycles during a single day. Each payment cycle has stringent Service Level Agreements (SLAs).The consequences of downtime are severe and include substantial financial penalties.

Why Migrate to Exadata Database Machine

The new electronic payment service involved a major step change in performance in order to allow real time enquires against transaction data whilst at the same time processing large bulk payment volumes in batch cycles. An unmovable deadline for introducing the new service combined with an increase in the projected payment volumes added to the challenge, requiring VocaLink to find a solution that could cope with the increased payment volumes without the need for application design changes. It was quickly determined that the ability to run Oracle applications unchanged on Exadata, combined with Exadata's unique performance features – Smart Flash Cache and Smart Scan, would enable VocaLink to meet its very strict service levels and deadline for implementation.

Service Level Agreements

There are many different performance SLA's for the different stages of the payment processing cycle. Each SLA in each stage of the processing cycle must be achieved while also meeting an average 0.54 second response time SLA for real-time enquiries via webservices.

Additional VocaLink SLAs for availability and data protection are:

- High Availability (HA) - no single point of failure can cause loss of service – this applies to both batch (payment processing) and online (webservices).
- Data Protection – There is zero tolerance for data loss regardless of outage, be it due to a local failure or a natural disaster. Recovery Point Objective (RPO) = 0.
- Disaster Recovery (DR). Complete site failover in a disaster event must complete in under 15 minutes. Recovery Time Objective (RTO) = 15 minutes.

Migration

VocaLink's previous environment was implemented on a pair of SUN M5000's in each of the two data centers. Each system was configured with 2 x Dual core 2.5 MHz CPU's and 96 GB of RAM running Solaris SPARC version 10 Update 8. The Oracle Database was an 11.1.0.7 two-node Oracle RAC cluster. Storage was provided by a dedicated EMC CLARiiON SAN.

The migration was performed while the database was still less than 500GB in size (it has since grown to 1TB with expected growth to 4.5TB). The high level migration steps were:

- Oracle Transportable Tablespaces was used to export the tablespaces to a DBFS filesystem on the Exadata system.
- The tablespaces were converted from Solaris SPARC format to Linux format and then imported into the Exadata database.
- During this process the database was also upgraded from 11.1.0.7 to 11.2.0.2.

See *Oracle MAA Best Practices: Migration to Exadata Database Machine* for more information on different options for migrating to Exadata.[4]

[4] http://www.oracle.com/technetwork/database/features/availability/xmigration-11-133466.pdf

VocaLink Architecture on Oracle Exadata Database Machine

Exadata MAA Architecture

The electronic payment system architecture overview is provided in Figure 2. At a high level it is comprised of:

- Client tier: Proxy using Apache, MQ, and FTP
- Application tier: Weblogic and MQ
- Database tier: Oracle Database 11g (11.2.0.2)

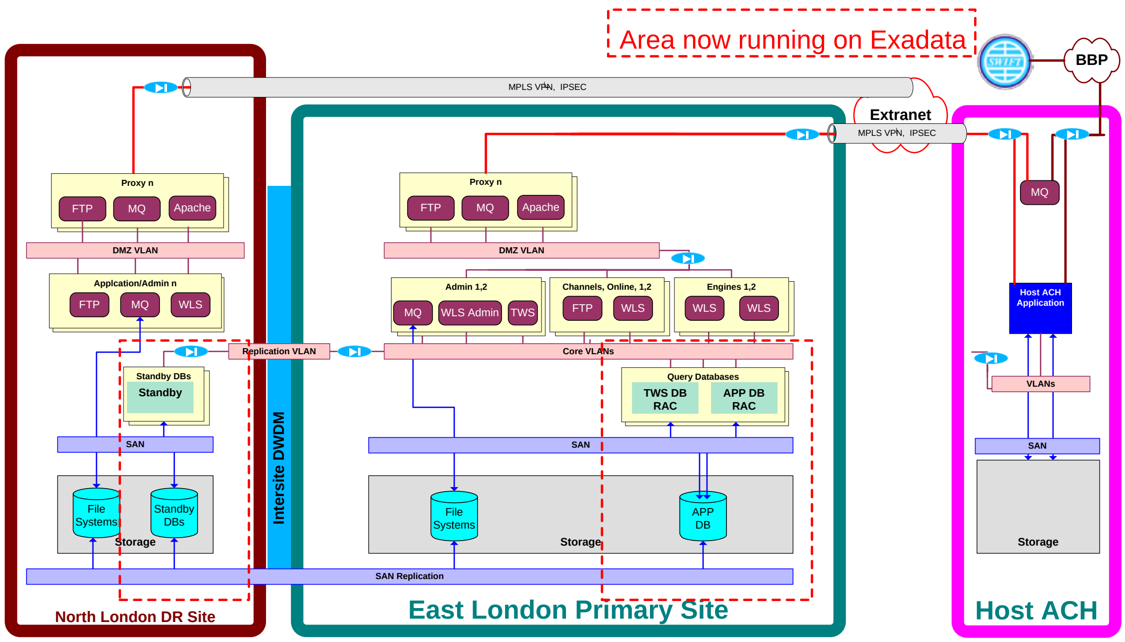

FIGURE 2 – ARCHITECTURE OVERVIEW

Additional configuration details are provided in the following sections.

Primary Database - HA

The primary database is configured as a 2-node Oracle RAC[5] database running on an Exadata Quarter Rack system. Local HA is provided by Oracle RAC and redundant components internal to Exadata. VocaLink's database was 1TB at the time this paper was written, and is projected to grow to 4.5 TB.

---

[5] www.oracle.com/goto/rac

The primary database is also configured with:

- Flashback Database[6] using a flashback retention period of 24 hours.
- There are two ASM[7] diskgroups, APP_DAT (data) and APP_FRA (Fast Recovery Area) configured using ASM normal redundancy. ASM normal redundancy is an MAA best practice when a Data Guard standby database is available to protect against multiple failures. The disk groups are sized at 6TB and 3TB respectively. Database online redo logs are placed on both the APP_DAT and APP_FRA diskgroups for redundancy reasons.
- VocaLink follows MAA Best Practice for optimal corruption prevention, detection and repair.[8] VocaLink has configured:
  - `DB_BLOCK_CHECKSUM` – to validate physical consistency
  - `DB_BLOCK_CHECKING` – to validate logical intra-block consistency
  - `DB_LOST_WRITE_PROTECT` – to detect silent corruptions and prevent propagation to standby database.
  - Active Data Guard automatic block repair. If a block corruption is detected at the primary database, Oracle will automatically repair the corruption using a valid copy obtained from the Active Data Guard standby database. The repair is transparent to the application and the user. Notification that a block was repaired is provided in the database alert log. Similarly, if a corruption is detected at the standby database, it is automatically repaired using a valid copy obtained from the primary.

The workload generated by the electronic payment system is a mix of online transactions and batch processing. Data is loaded 24x7 from data files, MQ, and web services. Data files used for loads are stored in Oracle SecureFiles[9] as a binary data file (blob). All settlement activity is processed during business hours. Database workload is characterized by a high volume of small inserts.

Secondary Database – Active Data Guard Physical Standby

A similarly configured Active Data Guard physical standby database runs on a separate Exadata Quarter Rack located in Data Center 2, 100 kilometers (sixty miles) from the primary site.

Data Guard is configured in Maximum Availability mode with synchronous redo transport providing zero data loss protection in the event of any single failure. The impact to primary database performance due to round-trip network latency and synchronous redo transport is within tolerance of

---

[6] www.oracle.com/goto/flashback

[7] www.oracle.com/technetwork/database/index-100339.html

[8] www.oracle.com/technetwork/database/focus-areas/availability/maa-datacorruption-bestpractices-396464.pdf

[9] www.oracle.com/technetwork/database/features/secure-files/index.html

VocaLink's performance SLA's. The Data Guard `net_timeout` parameter is set to a value of 60 seconds; establishing the maximum amount of time that the primary will wait for acknowledgement from the standby database to confirm that data is protected before proceeding with the next transaction. This prevents a standby or network outage from impacting the availability of the primary database when the primary is otherwise healthy. Standby redo log files at the standby database have been placed on the highest performance ASM disk group (DATA), to further minimize the impact of synchronous transmission on primary database performance.

Active Data Guard Automatic Block repair is used to automatically repair on-disk corruptions detected by Oracle, transparent to the user and application. Automatic repairs have been observed in VocaLink test environments, but they have not been observed in production as of the date this case study was written.

A custom caching application implemented at the DR site is able to access the Active Data Guard Standby Database (the standby database is open read-only while it receives and applies changes received from the primary). This keeps reference data used by the application tier at the standby site up-to-date, enabling fast application activation if a failover is required.

Data Guard Broker[10] provides a simple command-line interface to enable the Data Guard management of all primary/standby databases from any system that is part of a Broker configuration. Manual role transitions are executed via Broker command (switchover for planned transitions, failover for unplanned). The Broker also enables the complete set of Data Guard monitoring and administration capabilities available with Oracle Enterprise Manager.

See the Oracle Technology Network for more information on Oracle Active Data Guard.[11] See *Disaster Recovery Best Practices for Exadata Database Machine* for more information on the specific use of Data Guard with Exadata.[12]

Application Server Tier

Mid-Tier application servers are SUN x440 servers, Solaris 10 x64, JDK1.6 and WLS10.3.

Oracle WebLogic servers at each site are configured to connect to the local database server.[13] Load-balanced multipool configurations are used to reduce the amount of manual intervention required after a database failover or switchover.

---

[10] http://docs.oracle.com/cd/E11882_01/server.112/e17023/toc.htm

[11] www.oracle.com/goto/dataguard

[12] www.oracle.com/technetwork/database/features/availability/maa-wp-dr-dbm-130065.pdf

[13] www.oracle.com/technetwork/middleware/weblogic/overview/index.html

Network

Primary and standby systems are connected via GigE and two DWDM links. Round-trip network latency between primary and standby sites (RTT) is between 2 and 3 milliseconds.

Given that VocaLink is using SYNC redo transport, the SDU size has been increased to its maximum value of 32767. Note that from Data Guard 11g onward, it is no longer necessary to increase SDU size beyond the default when using Data Guard ASYNC redo transport or for when Data Guard automatically resolves archive log gaps using the ARCH process.

VocaLink set their TCP send/receive buffer sizes for Data Guard communication at 10MB for the most efficient utilization of their network. MAA best practice requires setting TCP send/receive buffer size to a value equal to 3x the bandwidth delay product (BDP) or 10MB, whichever is larger.[14] The BDP calculation, (network bandwidth) x (round trip network latency), for VocaLink's configuration yielded a value of 1125000 bytes, below the 10MB minimum.

Monitoring and Management:

VocaLink is currently using Oracle Enterprise Manager Grid Control 11g and the available plugins to monitor the Exadata servers[15]. An upgrade to Enterprise Manager Cloud Control 12c is planned to take advantage of new enhancements for monitoring Exadata.[16] Following the implementation of Cloud Control 12c, VocaLink will also implement Oracle Configuration Manager (OCM) to automatically upload configuration details to multiple CSI's.

Exachk automated health-check for Exadata is also run regularly on all of VocaLink's Exadata systems.[17]

Automatic Service Request (ASR)[18] is being investigated along with a separate activity to implement Ops Center at VocaLink.

---

[14] http://docs.oracle.com/cd/E11882_01/server.112/e10803/config_dg.htm#CEGEADFC

[15] https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1110675.1

[16] www.oracle.com/technetwork/oem/enterprise-manager/overview/index.html

[17] https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1070954.1

[18] www.oracle.com/asr

"Active Data Guard blurs the line between high availability (HA) and disaster recovery (DR). VocaLink is able to execute a complete site failover with approximately 8 minutes of total downtime, much faster than our previous DR architecture. Data Guard Standby-First Patching is also instrumental in reducing downtime for planned maintenance." **Martin McGeough — Database Technical Architect, VocaLink**

Data Guard Role Transitions (switchover/failover)

Failover/switchover processes are controlled by a job scheduler (Tivoli Workload Scheduler) due to external dependencies above the Application and Database tiers. This also controls the Data Guard Broker[19] and the application activation processes. All role transitions include network, proxy, application and database tiers.

Complete site failover (for unplanned outages) has been timed at approximately 8 minutes

- Database failover consumes 5 minutes.
- Application tier failover takes an additional 3 minutes.

Flashback Database is implemented to allow fast reinstantiation of a failed primary database as a standby database for the new primary (no restore required). The time required for complete re-instantiation is a function of the time required for two activities

- Time spent flashing back the original primary database to insure it is in the past of the SCN where the standby became the new primary. This is very fast given that VocaLink is using synchronous transport.
- Time spent resynchronizing the new standby with all transactions that have occurred at the new primary database. This time will vary based upon how long it takes to repair the fault that triggered the failover in the first place.

VocaLink has only had to perform a failover as part of a test procedure. Tests allow 30 minutes of processing at the new primary after failover before the original primary was brought back online. The process of complete reinstatement and resynchronization of the old primary as a standby database for the new primary is completed in less than 10 minutes.

Complete site switchover (in support of planned maintenance) has been timed at less than 10 minutes. Switchover takes slightly longer to allow for the orderly transition of application clients from the original primary to the new primary, something that is not required for failover when there is a sudden outage of the original primary database.

---

[19] http://www.oracle.com/pls/db112/to_toc?pathname=server.112/e17023/toc.htm

VocaLink conducts annual failover tests. Switchover operations are conducted more frequently to minimize downtime during planned maintenance. Hardware maintenance and many types of software maintenance are performed first on the standby system. The only downtime for such maintenance is the time required for the switchover operation to complete – less than 10 minutes from above, regardless of the time spent on the maintenance activity.

Backup Strategy

VocaLink currently performs full backups on a nightly basis using Oracle Recovery Manager (RMAN). This will change to use an incremental-based backup strategy once the database grows significantly past its current 1TB size.

Backups are performed at both the primary and the standby sites. VocaLink considered the option of offloading the primary database and performing backups only at the standby (primary and standby databases are exact physical replica's, backups are interchangeable). VocaLink chose the more conservative approach of backing up both primary and standby for the added security of having redundant backups taken from two different sources and to have a local backup at each site for the fastest possible restore if needed.

For optimal performance and data protection RMAN[20] is used to perform on-disk backups to the Fast Recovery Area on Exadata storage. Backups are then run off to tape to conserve disk space. RMAN scripts also provide automated management of archive log files.

Under normal circumstances, VocaLink will use Data Guard's automatic gap resolution to resynchronize the standby database following network or standby server outages. In the event of an extended outage, VocaLink will use an RMAN incremental backup in place of Data Guard's automated solution to more quickly resynchronize a standby database if it has fallen far behind the primary. The RMAN incremental process for catching up a standby database is:[21]:

- Determines the last SCN applied at the standby.
- Perform a fast incremental backup of the primary from that SCN. Note that RMAN Block Change Tracking is used for fast incremental backups.
- The incremental is used to restore the standby database to a point in time close to the current time on the primary.
- Data Guard then automatically performs the final resynchronization using redo from transactions that had committed since the incremental backup was taken.

---

[20] www.oracle.com/goto/rman

[21] http://docs.oracle.com/cd/E11882_01/server.112/e25608/rman.htm#CIHIAADC

While VocaLink has not had to use this process with their Exadata Data Guard configuration, they have successfully utilized this process in production on one of their non-Exadata systems.

Planned Maintenance

VocaLink's general approach to Exadata patching is to implement Exadata Bundle Patches on a half yearly basis and to assess the remainder of their software stack on a similar cycle unless there is a recommendation that requires a faster implementation due to the risk profile.

Functional and performance tests are run with Bundle Patches over a course of five days in isolation from production on an Exadata test environment. Successful testing preceded the implementation of Exadata Bundle Patch 4 and Bundle Patch 10 over the past year with no negative impact to production. Bundle Patch 13 scheduled in the first quarter of calendar 2012. BP17 Dec 2012

Likewise the Exadata Storage Server software is run through a similar test cycle in the Exadata test environment, though the test period was extended to eight days to accommodate additional HA testing. The length of test cycles depends upon the extent of the change being tested, for example VocaLink expects to have a significantly longer test cycle when upgrading to a future patchset of Oracle Database 11.2.

Once testing is complete using the Exadata test system, VocaLink uses Standby-First patching when possible as the preferred method for introducing patches and other planned maintenance into the production environment. Standby-First patching uses Data Guard switchover to reduce downtime and risk when installing patches and performing other planned maintenance in the production environment.

The general approach to VocaLink's use of Standby-First Patching is:

- First install a patch at the physical standby while production continues to run unchanged on the primary system.
- After the patch is deployed Data Guard will automatically resynchronize the patched physical standby database with the primary.
- When satisfied that there has been sufficient validation of the patch at the standby database, a Data Guard switchover is executed to transition production to the standby.
- The only database downtime typically required by the Standby-First patching process is the time required to execute a switchover. If a patch also requires SQL scripts or commands, then the downtime period will be slightly longer while these are run as part of a post-installation process on

the new primary (after the switchover has been executed but before the database is open for new connections).

See My Oracle Support Note 1265700.1 for more information and best practices on Standby-First patching[22].

In addition to consultation with Oracle's Exadata Solutions Support Center staff (ESSC), VocaLink utilizes the following key Oracle Support notes when planning maintenance activities:

- Note 888828.1: Latest updates and patching recommendations for Exadata Database Machine.[23]
- Note 1262380.1: Best practices for testing/implementing recommended patches.[24]
- Note 1070951.1: MAA Health Check - exachk[25]
- Note 1110675.1: Exadata Database Machine monitoring and the Automated Service Request (ASR).[26]

Post-Production Experience

VocaLink has concluded the following from its Exadata experience:

- VocaLink has found Exadata to be completely resilient. The only outages VocaLink have experienced thus far have been for planned maintenance. For example, there was an event where a disk corruption caused one of the primary database nodes to fail. The internal resiliency of Exadata combined with Oracle RAC prevented an unplanned outage; production continued to run on the second instance as designed. VocaLink ultimately decided to switch production to the Data Guard standby in order to restore the corrupt database node on the original primary and thus eliminate the perceived risk exposure.
- VocaLink have not experienced any SLA breaches due to Exadata. Exadata has proven to be fully resilient and Exadata performance has been consistent in meeting the various payment processing SLA's while also maintaining the average response time to Websevices queries to under 0.54 seconds

[22] https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1265700.1
[23] https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=888828.1
[24] https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1262380.1
[25] https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1070951.1
[26] https://support.oracle.com/CSP/main/article?cmd=show&type=NOT&id=1110675.1

Conclusion and Lessons Learned

As a general recommendation, VocaLink encourages Exadata users to keep their installations as vanilla as possible and to follow Oracle MAA Best Practices[27]. Doing so enabled VocaLink to benefit not only from its own experience with Exadata, but also from the collective experience of the extensive Exadata installed base.

Secondly, VocaLink found it crucial to invest in an Exadata test system to enable thorough testing in isolation from the production environment. This is not unique to Exadata, it is a standard best practice for any mission critical platform. The Exadata test system not only reduced the risk of introducing change to the production environment, it enabled thorough evaluation and the determination of which Exadata features and best practice are most beneficial for a given application. For example, VocaLink's applications have been iteratively tuned over the last year to further improve the performance by making use of Exadata Smart Flash Cache and Smart Scan functionality. Smart Flash Cache has primarily been used to consistently achieve SLAs for Websevices response time as volume has grown. Smart Scan has been used to meet and exceed SLAs for payment processing as volume has grown.

Finally, while VocaLink has realized significant benefit from Exadata's ability to accelerate implementation, deliver high performance, achieve high availability and simplify support, as with any significant new system there were some 'teething' issues experienced along the way. These issues were resolved in large part by leveraging VocaLink's Exadata test environment and Oracle MAA Best Practices, resulting in Exadata proving itself as both a reliable and performant solution.

[27] www.oracle.com/goto/maa

Appendix A

- Oracle RAC = Real Application Clusters
- ASM = Automatic Storage Management
- BDP = Bandwidth Delay Product
- MAA = Maximum Availability Architecture
- SCAN = Single Client Access Name
- SLA = Service Level Agreement
- DBFS = Database File System
- HCC = Hybrid Columnar Compression
- ACO = Advanced Compression Option
- FAN = Fast Application Notification
- RMAN = Oracle Recovery Manager
- HA = High Availability
- DR = Disaster Recovery
- RPO = Recovery Point Objective
- RTO = Recovery Time Objective
- FRA = Fast Recovery Area
- CRS = Cluster Ready Services
- ASR = Automated Service Request

VocaLink – Oracl e Exadata Case Study
April 2013

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200

oracle.com