Implementing Oracle Products on Linux for System z

SHARE Session #9295

Denver, CO     August 26, 2009

Gaylan Braselton, IBM
gbrasel@us.ibm.com

Tom Kennelly
kennelly@us.ibm.com

Marc Connolly
Marc.connolly@oracle.com

Oracle Products on Linux on z

Tom Kennelly
- IBM Certified Technical Sales Specialist
- Oracle for Linux on System z
- kennelly@us.ibm.com

© 2009 IBM Corporation
Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

- AIX®
- DB2®
- System z®
- z/OS®
- zSeries®

The following are trademarks or registered trademarks of other companies.

- Intel, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation in the United States, other countries, or both.
- Java and all Java-based trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States, other countries or both.
- Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.
- UNIX is a registered trademark of The Open Group in the United States and other countries.
- Microsoft, Windows, Windows NT and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
- Red Hat, the Red Hat Shadow Man logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.
- Oracle is a registered trademark of Oracle Corporation and/or its affiliates. More information on Oracle trademarks can be found at www.oracle.com/html/copyright.html.

Notes:

The information contained in this document has not been submitted to any formal IBM test and is provided "AS IS" with no warranties or guarantees either expressed or implied. All examples cited or described in this document are presented as illustrations of the manner in which some IBM products can be used and the results that may be achieved. Actual environmental costs and performance characteristics will vary depending on individual client configurations and conditions. IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply. All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have obtained. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions. This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the products or services available in your area. All statements regarding IBM’s future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

Agenda

- Objectives
- Definitions
- Oracle Products as seen through High Availability options
- Oracle HA Solution Overviews
  - Foundation for Oracle Maximum Availability Architecture and Oracle Grid
- Disaster Recovery
- Summary
High availability is critical in today’s environment. The direction is always towards continuous availability.

This presentation looks at the availability options from an Oracle standpoint and is not meant to preclude IBM alternatives.

The key technologies of Oracle’s Maximum Availability Architecture are discussed.

Objectives

- High Availability (HA) – Provide service during defined periods, at acceptable or agreed upon levels, and masks unplanned outages from end-users. It employs Fault Tolerance; Automated Failure Detection, Recovery, Bypass Reconfiguration, Testing, Problem and Change Management.

- Continuous Operations (CO) – Continuously operate and mask planned outages from end-users. It employs Non-disruptive hardware and software changes, non-disruptive configuration, software coexistence.

- Continuous Availability (CA) – Deliver non-disruptive service to the end user 7 days a week, 24 hours a day (there are no planned or unplanned outages).

The goal is to strive to provide Continuous Availability.

Definitions

- High Availability (HA) – Provide service during defined periods, at acceptable or agreed upon levels, and masks unplanned outages from end-users. It employs Fault Tolerance; Automated Failure Detection, Recovery, Bypass Reconfiguration, Testing, Problem and Change Management.

- Continuous Operations (CO) – Continuously operate and mask planned outages from end-users. It employs Non-disruptive hardware and software changes, non-disruptive configuration, software coexistence.

- Continuous Availability (CA) – Deliver non-disruptive service to the end user 7 days a week, 24 hours a day (there are no planned or unplanned outages).

The goal is to strive to provide Continuous Availability.

Definitions provided by the HA Center of Competence in Poughkeepsie.
Fundamentals of High Availability

- Redundancy, Redundancy, Redundancy – Duplicate everything to eliminate single points of failure.

- Protect Data Consistency – Provide ability for data and file systems to return to a point of consistency after an unplanned outage.
  - Journaling databases
  - Journaling file systems
  - Mirroring
  - Routine database backups

- Automate Detection and Failover -- Let the system do the work in order to minimize outage windows.
  - Multipathing
  - VIPA
  - Monitoring and heart beating
  - Clustered middleware
  - Clustered operating systems

Building Blocks of HA for Oracle on Linux for System z

Server Provided HA
Oracle

Operating System HA
Linux
z/VM

Hardware Provided HA
System z

RAC
Data Guard
Flashback
CRS
Grid Control

Linux Clustering
Oracle ASM

Mature Hypervisor
Hardware assist for performance and recovery

Spare CPUs
N+1 power supplies
Chip sparing
Concurrent Maintenance
50 years MTBF (system failures)
Oracle Database without Oracle MAA

Guards against:
- Hardware failure – z10

Comments:
- Sufficient for many databases
Oracle Database - building Oracle MAA

Guards against:
• Hardware failure – z10

Comments:
• Added Oracle’s Cluster Ready Services (CRS)
• Now ASM is a cluster ready file system and is ready to be shared.
Oracle Database - building Oracle MAA - Hot standby

- Guards against:
  - Hardware failure – z10
  - Linux OS or Oracle DB failure
  - Allows for maintenance to Linux and possibly Oracle in the Prod guest

- Comments:
  - Added a hot stand by Linux to same Oracle DB.
  - Through the use of CRS, with its heartbeat, and ASM shared disk storage do a failover to the stand by guest.

Oracle Hot Stand By Approach

- Comments
  - Can also be accomplished across LPARs using HiperSockets connections.
  - Can be accomplished across different System z platforms using appropriate network connectivity.
  - Only allowed between Oracle databases using the same binaries (i.e. Linux on z in this case)
  - An outage that can affect users occurs but can be of a short duration
Oracle Database building Oracle MAA – RAC

Guards against:
• Hardware failure – z10
• Linux OS or Oracle DB failure
• Allows for maintenance to Linux and possibly Oracle in the Prod guest

Comments:
• Unlike hot standby there will be little impact to the end users of Linux node failure.
• Load balancing is occurring between the RAC nodes.
Oracle Database - building Oracle MAA - RAC

Guards against:
- Hardware failure – z10
- Linux OS or Oracle DB failure
- Allows for maintenance to either z/VM, Linux and possibly Oracle DB in either Prod guest

Comments:
- Physically separate z10s

Overview of Major RAC Components
Oracle RAC as an HA Solution

- RAC implies a HA Solution
  - RAC provides high availability for database instances

- Have you taken into account single points of failure for:
  - Disk failures?
  - IPC Interconnect failures?
  - Are the servers on the same electrical circuit?
  - Are the servers under the same sprinkler?
  - If the nodes are in a different building, is it a single cable run?
  - Did you do appropriate capacity planning for a node or multiple node failures?

- Your availability is as solid as your planning for any platform on which you implement a RAC solution
  - If you plan well, it is a very Highly Available software solution

Deploying RAC for High Availability

- RAC – Real Application Clusters
  - Active/Passive configuration
    - One node processes work
    - The other node waits for the first node to fail
  - Active/Active configuration
    - All nodes process work
    - If any node fails the cluster is re-mastered.
  - Besides availability, RAC can be used for workload distribution
    - All work does not have to go through all nodes
  - Deploy
    - In the same LPAR for test/dev applications
    - Across LPARs for LPAR maintenance or software failures (most common implementation)
    - Across CECs when taking entire systems down is a “common” occurrence
### Oracle Standby and Replication Solutions for Disaster Recovery

- **Standby – replication to standby database**
  - **Oracle Data Guard**
    - Uses redo log shipping for log apply or SQL Apply
    - Less data transmitted than replication
    - Sync or async
    - Various configurations of logical and physical standby databases
  - **Data Broker** monitors database and affects transition
  - Both production and standby databases must be installed from same CD/DVD
    - Support for heterogeneous systems not supported yet
    - Both systems must match for endian, chip set and headers
  - Data Guard generally deployed between CECs

#### Standby Database - Data Guard

- **Primary LPAR or CEC**
- **Application Server(s)**
- **DR LPAR or CEC**
- **Linux**
- **Oracle Database**

**Comments:**
- Think of Data Guard for disaster recovery as well as RMAN backups
**High Availability with Oracle on Linux for System z**

- System z – reputation for the most highly available platform on the planet
  - Attention to detail over decades of engineering
    - Fault Tolerant (HA) design
    - Elimination of single points of failure
  - Driving to 100 years MTBF
- **Oracle Maximum Availability Architecture**
  - Best Practices based on Oracle technology
    - Best HA/DR in distributed database technology (Forrester, Oct 2006)
    - Spans all Oracle products
    - Constantly evolves with new releases
- **Synergistic**
  - Continue on your path with Grid using System z
  - Develop a Grid strategy for Oracle on Linux for System z
  - Take advantages of the HA/DR features of IBM and Oracle technologies

**How do I get started for existing workloads**

- Choose servers to consider consolidating
- Engage IBM for a SCON study
- Determine memory requirements
- Understand z/VM and virtualization
- Undertake a Proof of Concept
- Realize the benefits of Oracle DB/AS on Linux on z
Additional Information Sources

  - SG24-6482-00 Experiences with Oracle Database 10g on Linux for zSeries
  - SG24-7191-00 Experiences with Oracle 10gR2 Solutions on Linux for System z
  - SG24-7573-00 Using Oracle Solutions on Linux on System z
  - SG24-7634-00 Experiences with Oracle Solutions on Linux for IBM System z

- http://www.oracle.com/ibm
  - IBM platform information

- http://otn.oracle.com
  - (Select “Downloads”)

  - General z/VM Tuning Tips
    - Lot’s of information on Linux for zSeries
    - Hints and Tips for tuning Linux on System z
  - http://www.zseriesoracle.sig.org
    - Special Interest Group of Oracle users on the mainframe (z/OS and Linux)
  - http://www.mail-archive.com/linux-390%40vm.marist.edu/
    - Marist List Server
  - http://www.oracleinsight.net/2008/02/06/the-mainframe-renaissance/
    - The Mainframe Renaissance

Any Questions?

Thanks!!!
Oracle Database Single Instance Failover using Oracle Clusterware and ASM

- Implementation Components
  - Oracle Clusterware (CRS)
  - Oracle Automatic Storage Management (ASM)
  - 1 or More non-RAC Oracle Databases Sharing Clustered ASM Instance
- Component Illustrations
- Installation, Configuration and Testing
  - Component Installation Sequence and Sample Functional Tests
- Resources
**Oracle Clusterware**

- Formerly, Cluster Ready Services (CRS)
- Initially Developed for RAC
- Eliminated need for 3rd Party Products
- Reduces Customer and Internal Oracle Costs
- Supports Non Oracle Database Requirements

**Oracle Clusterware Components**

- Virtual IP's
  - IP Address which can be “failed over”
  - Provides Mechanism to Automate Reconnections
- Voting Disks
  - Shared file or disk device for quorum management
- Cluster Registry (OCR)
  - Cluster Configuration and Status
- Daemons
  - crsd – Cluster Resource Services
  - ocssd – Cluster Synchronization Services
  - evmd – Event Management Logger (Event Generation)
  - oprocd - I/O Fencing
- Oracle Supplied & Custom Applications
**Oracle Clusterware Illustrated**

Clusterware "monitors" file system

```
/example
```

Clusterware "detects" file system failure

```
/example
```

**Server A**

**Server B**

Shared Disk
Oracle Clusterware Illustrated

Clusterware fails over VIPs
Enables and mounts files system

Server A

Server B

Shared Disk

Oracle Automatic Storage Management

- Storage Management for Oracle Databases
- Volume Manager
  - 1MB/128KB Striping, Flexible Mirroring
  - Online Disk Reconfig & Auto Rebalancing
- File System
  - Even Data Distribution for Optimal Performance
  - Automatic File Management via Oracle Managed Files (OMF)
- Clustered Configuration Support
  - Using Oracle Clusterware
- Oracle Enterprise Management Integration
  - In addition to Command Line Utility Management

ASM Disk Group

ASM Disk
ASM Disk
ASM Disk
ASM Disk
ASM Disk Group

+DATA
datafiles
file1... filen
controlfile
redolog1
Oracle Automatic Storage Management

**One DB Instance to One ASM Instance**

Database Instance

- SGA
- ASM file extents
- DB Processes
- I/O Service Requests

Database Kernel Performs I/O

ASM Instance

- Manages Volume and File Layout
- Captures & Uses I/O Statistics
**Oracle Automatic Storage Management**

**Multiple DB Instances to One ASM Instance**

- Multiple Database Instances
  - DB Processes
  - SGA
  - ASM file extents

- Single ASM Instance

**Single ASM Instance**

- ASM

**I/O Service Requests**

---

**Oracle Automatic Storage Management**

**Multiple DB Instances to One Clustered ASM Instance**

- Multiple Database Instances
  - DB Processes
  - SGA
  - ASM file extents

- Clustered ASM Instance
  - ASM Node 1
  - ASM Node 2
Oracle Database Single Instance Failover
Installation & Configuration

1 LPAR

Linux A

z/VM

Linux B

Gather Media & Patch Sets
Install CRS
Prep for Reg/Voting Files
Verify Cluster Config
Run OUI
Verify Cluster Operational

Install Clustered ASM
Prep Data or LUNS
Run OUI
Create Disk Groups
Verify ASM on Each Node Operational

Perform Software Only DB Install on Each Node
Run OUI on Each Node
Shut down DB on 1st Node (via sqlplus)

Perform DB Create on 1st Node
Run DBCR on 1st Node
Create crapped on Other Nodes
Create Admin Directory on Other Nodes
Test DB Start/Stop on in Sequence Other Nodes

Config Other Nodes for DB Failover
Copy 1st Node initSid to Other Nodes
Create orapwd on Other Nodes
Create Admin Directory on Other Nodes
Test DB Start/Stop on in Sequence Other Nodes

Config CRS with Actions
On 1st Node Create & Install Failover Scripts
Copy Failover Scripts to Other Nodes
Create CRS App Profile
Register CRS App Profile
Perform Functional Tests

OK?
**Oracle Database Single Instance Failover**

*Installation & Configuration*

1. Start/Stop DB via `crs`
2. Database Instance Relocation via `crs`
3. Manual Shutdown via `sqlplus`
4. Manual Shutdown of Linux Kernel with DB Running
5. Test 1-4 from Other Nodes

**Sample Functional Tests**
Resources

- Redbooks:
  - Experiences with Oracle Solutions on Linux for IBM System z (SG24-7634-00)

- Oracle
  - Oracle Database Installation Guide 10g Release 2 (10.2) for IBM zSeries Based Linux (B25400-01)
  - Oracle® Database Oracle Clusterware and Oracle Real Application Clusters Installation Guide 10g Release 2 (10.2) for Linux (B14203-09)
  - VIPCA / SRVCTL / OUI Issues See Metalink Note 414163.1

Questions

Thank you.
Oracle Database
Advanced Security SSL
and
System z Crypto Support
(PKCS11)
**Cryptographic Functional Areas & Algorithms**

**Data Confidentiality**
- Shared Secret Key
- Data Transfer Encryption
  - VPN, SSL/TLS...
- Data Storage
  - Databases, Archives
- Short Key Lengths
- Algorithms
  - DES
  - T-DES
  - AES

**Data Integrity**
- One Way Hash Algorithms
- Data Transfer Verification
  - VPN, SSL/TLS...
- Data Storage
  - Databases, Archives
- Short Key Lengths
- Algorithms
  - MD5
  - SHA
  - MAC, MDC

**Key Confidentiality**
- Public Key Cryptography
- Hand Shake Authentication*
  - VPN, SSL/TLS...
- Secure Key Distribution
  - HSM Key Export, e.g.
- Long Key Lengths
- Algorithms
  - RSA
  - DSA

**Symmetric**

**Checksum**

**Asymmetric**

---

**SSL Handshake**

1. (1) Client "hello" Msg
   - Crypto capabilities & prefs, i.e., ver of SSL, cipher suites, & compression methods

2. (2) Server "hello" Msg
   - Selected cipher suite and compression method & its cert & may request client cert (optional)

3. (3) Verify Support for Cipher Suites, Check Server's Cert, Send Client cert (optional)

4. (4) Client Key Exchange Msg
   - Contains client pre-master secret key and msg auth codes (MAC) encrypted with Server's Public Key

5. (5) Sends Client Cert (optional)

6. (6) Sends Cert Certify Msg (optional)
   - Encrypted with Client's Private Key

7. (7) Verify Client Cert (optional) using client's public key

8. (8) Convert pre-master to master secret key

9. (9) Sends Change Cipher Msg
   - Requesting Server to Use Negotiated Cipher Suite

10. (10) Sends "finished" Msg
     - Message encrypted with cipher method & new-master key

11. (11) Server Sends its own Change Cipher & "finished" Msg

12. (12) Exchange Messages
     - Encrypted with shared secret key
Sender Encryption Sequence

Unsigned Message

Hash Algorithm

Hash

Concenate

Hash

Data

Encryption using Sender’s Private Key

Signed Message

Secret Symmetric Key Encryption

Clear Secret Key

Encryption using Recvr’s Public Key

Recvr

Signed & Encrypted Message

Encrypted Clear Secret Key

Messages

Receiver Decryption Sequence

Messages

Signed & Encrypted Message

Secret Key Decryption

Signed Message

Encrypted Clear Secret Key

Clear Secret Key

Private Key Decryption using Receiver’s Private Key

Public Key Decryption using Sender’s Key

Data

Hash

Hash

Equa?
Oracle Database Advanced Security SSL

- Database to Database to Client Authentication
  - Database can Request Client Certs
  - Clients can Request DB Certs
- File System Wallet (Credential) Support
- PKCS #11 openCryptoki Support for Linux
  - HSM Support for Secure Key Storage

```
Linux
DB Instance
Listener
PKCS#11 -
openCryptoki
libica
z90crypt
PKCS#11 -
openCryptoki
Job
Oracle
ASO
Java
IBMPKCS11Impl
Provider
.java
JCE
Oracle
SQL
Oracle
SSL
z/VM
PCI Crypto
Adapter
OPNCF
刻画administration/DB01/wallet/Wallet_server
Oracle PKCS#11 Wallet
/home/admin/DB01/wallet/Wallet_client
Oracle PKCS#12 Wallet
/usr/lib/pkcs11/PKCS11_API.so64
Accelerated Asymmetric Key Operations
Secure Key Operations
```
Resources

• Redbooks:
  • Experiences with Oracle Solutions on Linux for IBM System z (SG24-7634-00)
  • Security on z/VM (SG24-7471-00)
  • System z Cryptographic Services and z/OS PKI Services (SG24-7470-00)

• Oracle
  • Metalink Note 453523 (How to use HSM's with Oracle DB)
  • Oracle® Database Advanced Security Administrator's Guide (B14268-02)

• Other
  • IBM 4764 PCI-X Cryptographic Coprocessor FAQ
    • http://www-03.ibm.com/security/cryptocards/pcixcc/4764FAQ.shtml

Questions

Thank you.