Best Practices for Oracle on Linux for System z
Getting Started

Denny Dutcavich
Technical Sales Specialist
dutch@us.ibm.com

Trademarks

Trademarks

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries. For a complete list of IBM Trademarks, see www.ibm.com/legal/copytrade.shtml: AS/400, DBE, e-business logo, ESCO, eServer, FICON, IBM, IBM Logo, iSeries, MVS, OS/390, pSeries, RS/6000, S/30, VM/ESA, VSE/ESA, Websphere, xSeries, z/OS, zSeries, z/VM

The following are trademarks or registered trademarks of other companies

Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation
Java and all Java-related trademarks and logos are trademarks of Sun Microsystems, Inc., in the United States and other countries
LINUX is a registered trademark of Linux Torvalds
UNIX is a registered trademark of The Open Group in the United States and other countries.
Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.
SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.
Intel is a registered trademark of Intel Corporation
* All other products may be trademarks or registered trademarks of their respective companies.

NOTES:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

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 achieved. 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 product 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.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

Any proposed use of claims in this presentation outside of the United States must be reviewed by local IBM country counsel prior to such use.

The information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

Memory

- Sizing memory is still critical (post z/VM 5.2)
  - A Linux guest running Oracle should have memory for
    - Oracle SGA
    - Oracle PGA (generally very small)
    - Linux (about 256 Mbytes)
    - DBControl needs memory if always running
- Not enough memory
  - Excessive paging will occur affecting performance
  - Linux will randomly shut down processes in an effort to stay alive
  - Monitor paging to avoid poor performance
    - Add memory as necessary
- Real Memory Requirements
  - R:V = 3:1 is a good place to start
    - Can be altered depending on activity of virtual machines
    - More aggressive for guests with very low utilization
- Paging happens
  - Paging to memory helps

Paging

- VM Paging Space is critical
  - Configure expanded storage as the first paging device for z/VM
    - Configure at about 20% (2 Gig Max) of physical memory for xStore for z/VM 5.2
  - Paging Space
    - Provide enough disk to keep 50% free
      - About twice the amount of virtual memory allocated to all guests
    - Separate swap space from
      - guest data
      - sysres
- Linux Swap
  - Linux swaps (pages) to manage memory
    - Insure swap space configured
    - How much for R1 and R2?
      - 512 Mbytes
    - Ignore Swap Space error message from OUI during installation
  - Configure VDISKS for Linux Swap Space
    - Swaps are done to memory with VDISK
    - No memory utilized until swap occurs
      - Use judiciously as it uses main memory
    - Can be used to determine correct memory size by monitoring swapping on devices
      - Use two smaller VDISKS (or 1 VDISK and 1 Minidusk)
      - If swapping to second VDISK, add memory

Storage (DASD, Disk) Considerations

- Distribute data across arrays (IBM Storage)
  - Not doing this can cause BIG performance issues
- LVM Striping
  - ESCON one stripe per channel
  - FICON one stripe per volume
  - FCP one stripe per volume or LUN
  - Optimal stripe sizes
    - 16K or 32K (but may vary depending on your workload)
    - IBM used 8K for some internal verification testing on OLTP workload

Storage (DASD, Disk) Considerations

- Minidisk vs Dedicated
  - Minidisks with the standard ECKD DASD drivers have slightly higher throughput with little difference in overall CPU costs
- Minidisks can use Minidisk Caching (MDC)
  - Can yield a large benefit in throughput and significant CPU savings depending on the workload
    - Benefits more heavy R/O workloads
    - Benefits can be more substantial when sharing r/o minidisks between multiple Linux guests (i.e. /usr)
- FCP
  - Dedicated disks have significantly higher throughput and lower CPU costs
- SCSI vs ECKD
  - In general
    - SCSI:
      - Somewhat higher throughput, lower CPU cost per I/O
      - Larger volumes possible
      - Disk space (LUNS) easier to manage
    - ECKD:
      - Better “multi-pathing”
    - See Report
      - http://www.vm.ibm.com/perf/reports/zvm/html/520lxd.html

Changes for Oracle

- No specific changes recommended
  - Oracle 10g Databases moved to System z Linux should not require changes
    - Monitor performance after moving database (i.e. tune the database)
    - Asyncio is built into the kernel
- Things to be aware of
  - Optimizer
    - 10g requires CBO (Cost Based Optimizer)
      - Tuning may be necessary
  - Parallel Query
    - In 10g one CPU (IFL) is still beneficial
    - Will performance be the same if moved from 4 CPUs, 8 CPUs, etc
      - Depends on real capacity needs and differences in CPU speeds and architecture
- Otherwise normal tuning

What Do I Run in a Linux Virtual Machine with Oracle?

- Production
  - Single database is easiest to manage
  - Application Server can be in same virtual machine
  - However, one server in one virtual machine provides
    - Best performance characteristics and tuning
    - Best for capacity planning and accounting from z/VM for applications
- Test/Dev/etc
  - Multiples of anything are OK
  - You may experience performance issues with multiple DBs
    - If there are not enough resources such as memory, the guest may experience high page rates
    - Ensure you allocate enough resources
- Why
  - Linux does not differentiate between Oracle database instances
  - If Linux needs memory (i.e. import or load)
    - It may page other SGAs
  - You can lock the SGA with Oracle Database 10g from the init.ora
  - Or you can use
    - CP SET RESERVED <userid> <nn> (where nn is a number of page frames) for the userid
    - This could affect other guests if memory constrained
  - Probably best to do neither
    - Provide appropriate resources

Monitoring Performance

- z/VM has the only accurate view of what resources each guest is consuming. Any investigation should start in z/VM.
- The goal of z/VM is convince each guest that it is running alone on the machine and has all of its own resources.
  - Therefore, no guest can possible tell what is really happening on the system.
  - Don't under estimate the value of sar or vmstat
- Consider either
  - z/VM Performance Toolkit
  - ESAMON from Velocity Software
    - http://www.velocity-software.com/
  - Omegamon for Linux virtual machine performance
  - You can also use CP commands when you suspect problems
- Use Oracle Statspack to monitor your instances
  - Or DBControl
    - DBControl can use a bunch of memory
  - You can use Grid Control on another platform
    - Grid Agent available for Oracle on System z Linux

Key Success Factors

- Run z/VM 5.2!!!!
  - z/VM 5.2 alleviates the contention for storage below 2 Gbyte on previous releases
  - Always apply the most current Recommend Service Updates (RSUs)
  - Insure Fixedbuffer Feature is not enabled for z/VM 5.2
- If you are not at z/VM 5.2 yet apply fixed buffer feature
  - Use for ECKD DASD (does not apply to FCP)
  - Apply most current RSUs
- Memory sizing is critical
  - Small virtual guest size – less is better
  - Let z/VM manage memory not Linux
    - Use VDISK for Linux swap

Key Success Factors

- Monitor resource
  - Understand limits
  - Make changes (i.e. tuning) that may be necessary
- Paging and swap space necessary
  - Both should use memory devices
- Avoid I/O bottlenecks
  - Distribute data in the ESS across arrays
  - Stripe the LVM or use Oracle ASM
  - Faster is better, use FICON or FCP

Information Sources

- http://www.ibm.com/redbooks
  - Oracle, Linux and System z Redbooks
- http://www.oracle.com/ibm
  - IBM platform information
- http://otn.oracle.com
  - (Select “download code”)
- http://www.vm.ibm.com/perf/tips
  - General z/VM Tuning Tips
- http://www.vm.ibm.com/perf/reports/zvm/html/520lxd.html
  - Linux Disk I/O Alternatives
- http://www-124.ibm.com/developerworks/oss/linux390/index.shtml
  - Lot’s of information on open source and IBM products
- http://www-128.ibm.com/developerworks/linux/linux390/perf/index.html
  - Hints and Tips for Selecting and Tuning I/O options
- http://www.zseriesoraclesig.org
  - Special Interest Group of Oracle users on the mainframe (z/OS and Linux)