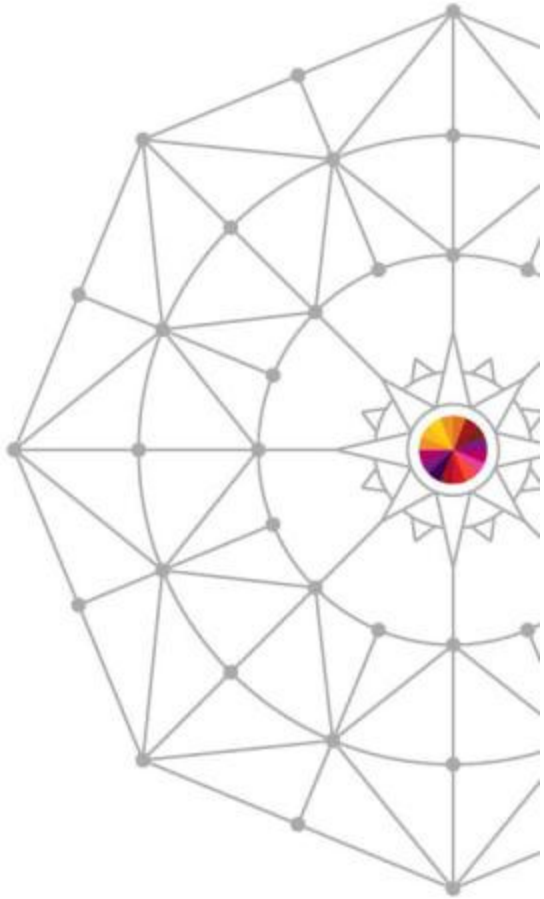
#SHAREorg

SHARE

How an Oracle Database on Linux on System z works well

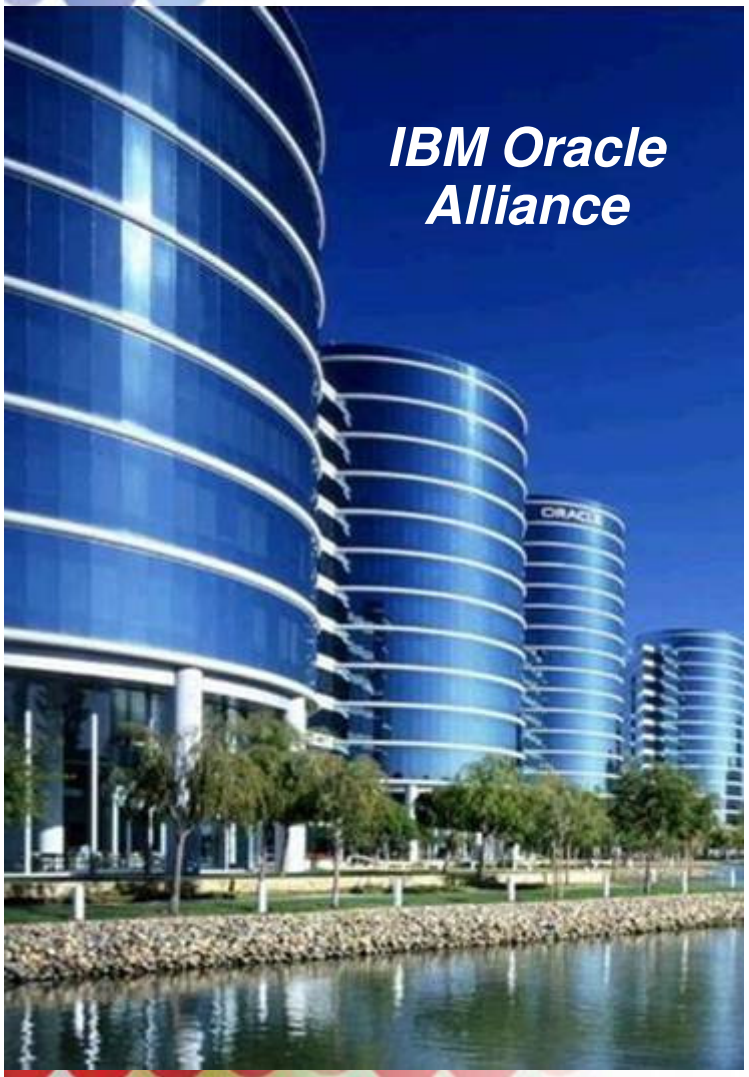
Wilhelm Mild
IT Architect
IBM Germany

2014/3/11
Session 15393



SHARE in Anaheim

IBM and Oracle Have a Long-Standing Relationship



Sustaining relationship of 150K + clients

- Oracle 25 years, PeopleSoft 23 years, JD Edwards 35 years, Siebel 13 years

Mutual executive commitment

- Dedicated, Executive-led Alliance teams, Regular Senior executive reviews

Vibrant technology relationship (Diamond Partner)

- Sustained investment in skills and resources including dedicated international competency centers

Market-leading services practice

- IBM GBS is Oracle's #1 SI partner (7,500 joint projects) with 5,000 people dedicated to Oracle

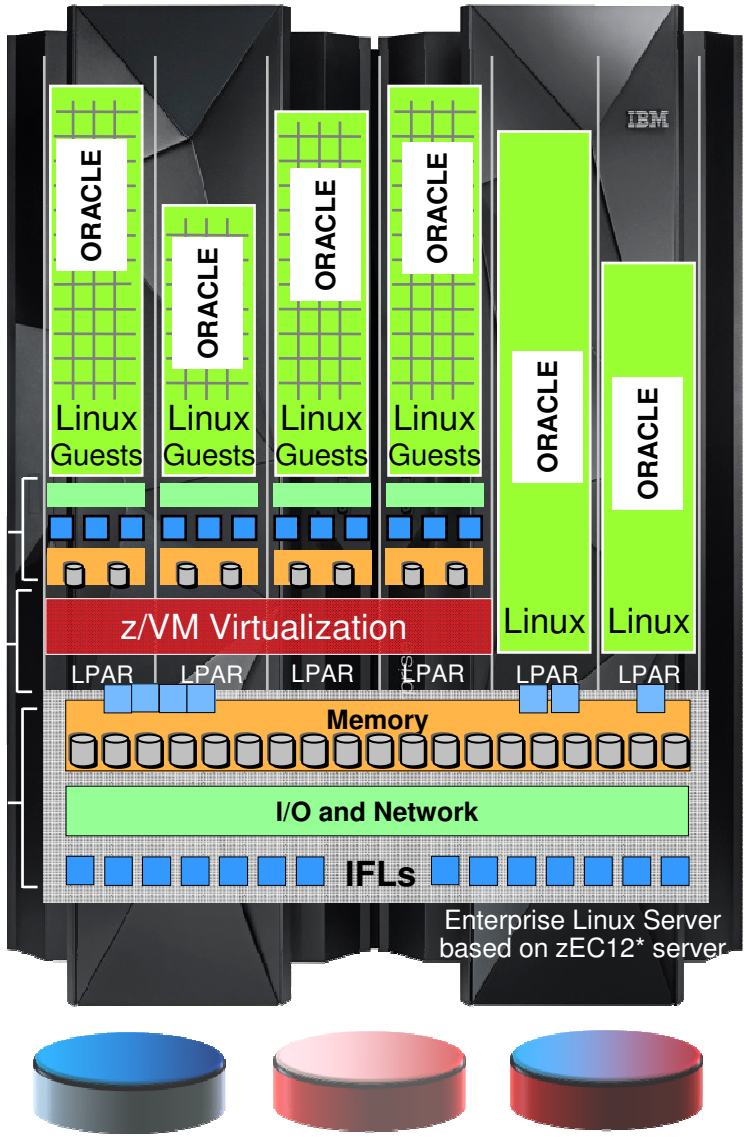
Unrivalled client support process

- Dedicated on-site resources and significant program investments ($77M on 1000+ assets)

Oracle 12c on Linux on System z

- Oracle on Linux on System z is the same Oracle as anywhere else - the code is ported to the new environment
- Oracle 12c (12.1) on SUSE x86-64 is the same as Oracle 12.1 on IBM Linux on System z.
- There is no difference between Linux distributions – it's the same Oracle image for any Linux on System z
- Supported is SUSE and Red Hat in exactly the same way

Enterprise Linux Server and Oracle

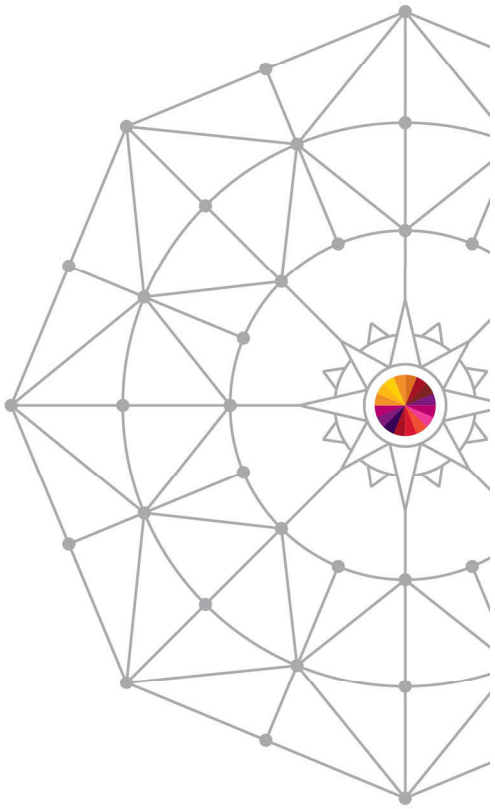


LPAR	Logical Partition = subset of hardware resources, virtualized as a separate computer; up to 60 LPARs can be configured
IFL	Integrated Facility for Linux = core; Enterprise Linux Server with zEC12* server: 5.5 GHz per core, up to 101 cores
Linux Guest	virtual Linux Guests running the workload such as database server, etc.; hundreds of virtual Linux Guests can be hosted on one Enterprise Linux Server (ELS)

* zEC12 = IBM zEnterprise EC12

Workload is dependend on all layers

- Storage
  - type, ECKD, FCP
  - attachment perfomance
- Virtualization
  - Type, LPAR, z/VM, Oracle
- Oracle 12c database implementation
  - single
  - container
  - HA / clustered

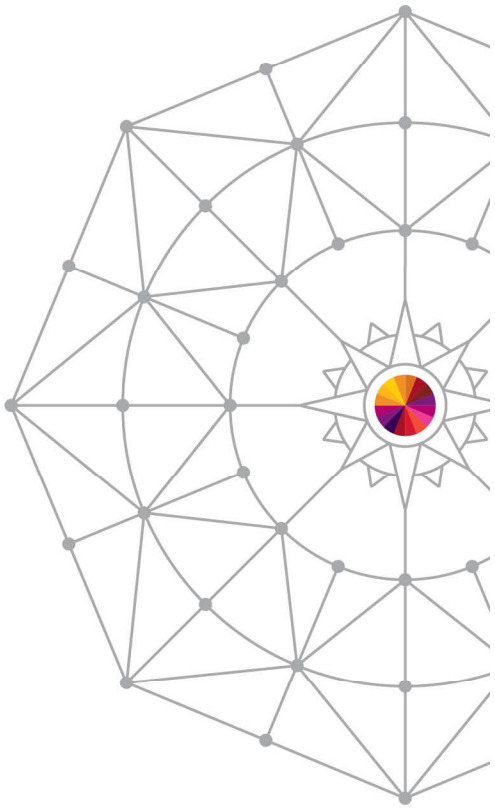


Agenda

1. Infrastructure for Oracle 12c database workloads
   - Disk options
   - FICON devices with HyperPAV
   - FCP devices and Multipathing
   - Comparing FICON and FCP disk devices

2. **Virtualization layers**
   - System z LPAR Virtualization
   - z/VM Virtualization
   - Oracle 12c virtualization

3. **Oracle Database layer**
   - Optimizing Oracle 12c setup
   - Linux on z and Oracle 12c work well

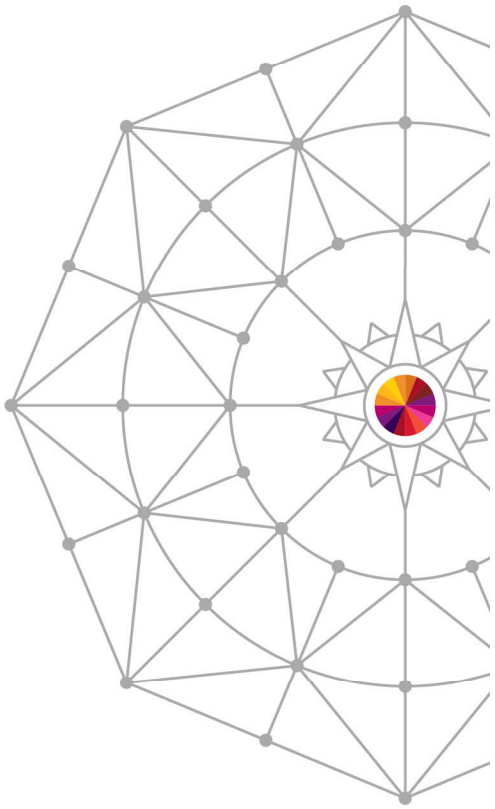


Agenda

1. **Infrastructure for Oracle 12c database workloads**
   - **Disk options**
   - **FICON devices with HyperPAV**
   - **FCP devices and Multipathing**
   - **Comparing FICON and FCP disk devices**

2. **Virtualization layers**
   - System z LPAR Virtualization
   - z/VM Virtualization
   - Oracle 12c virtualization

3. **Oracle Database layer**
   - Optimizing Oracle 12c setup
   - Linux on z and Oracle 12c work well



Linux on System z – Disk storage connectivity options

- z/OS
- Linux native
- Linux guest

Linux CDL on CKD VM Minidisk *

FICON

Linux data on ECKD disks

DS8000 only

- Linux native
- Linux guest

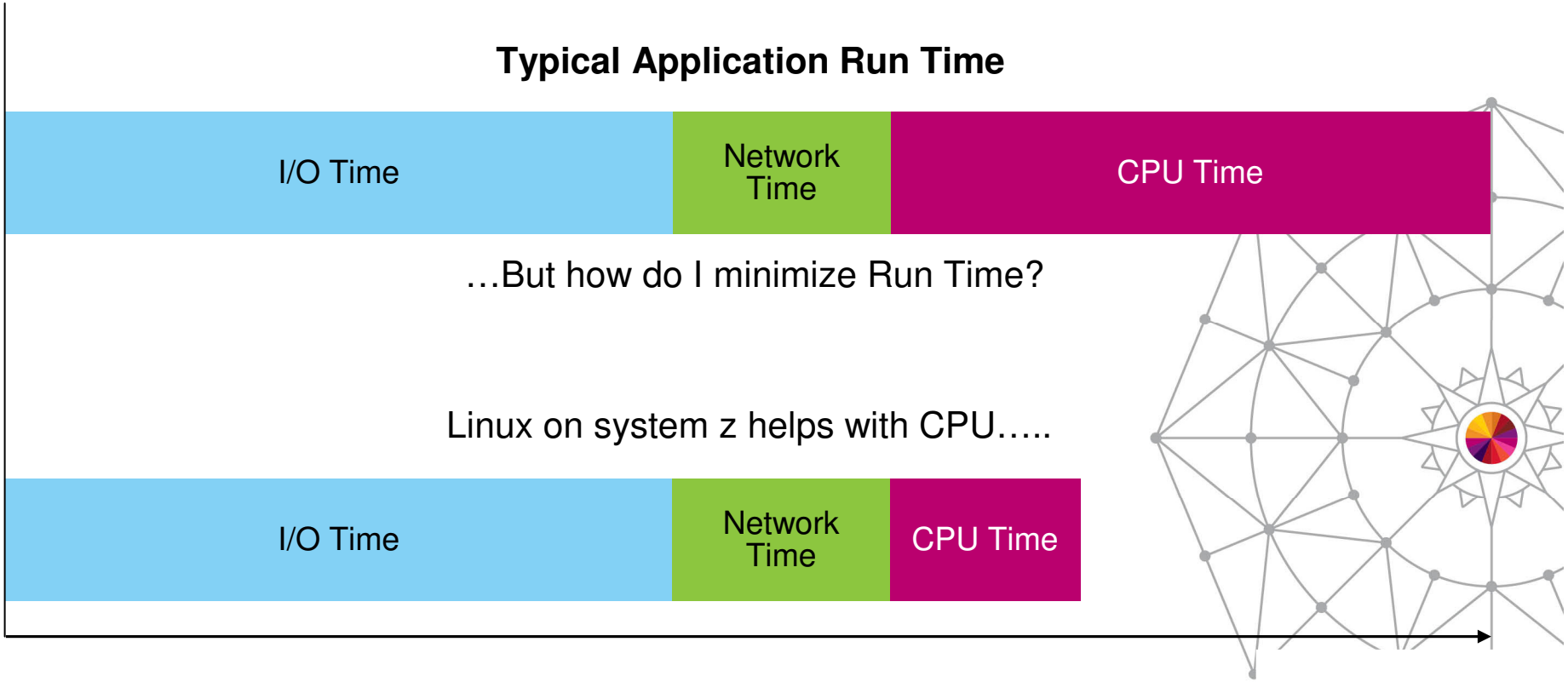
FCP

Linux on FBA VM Minidisk (EDEV)

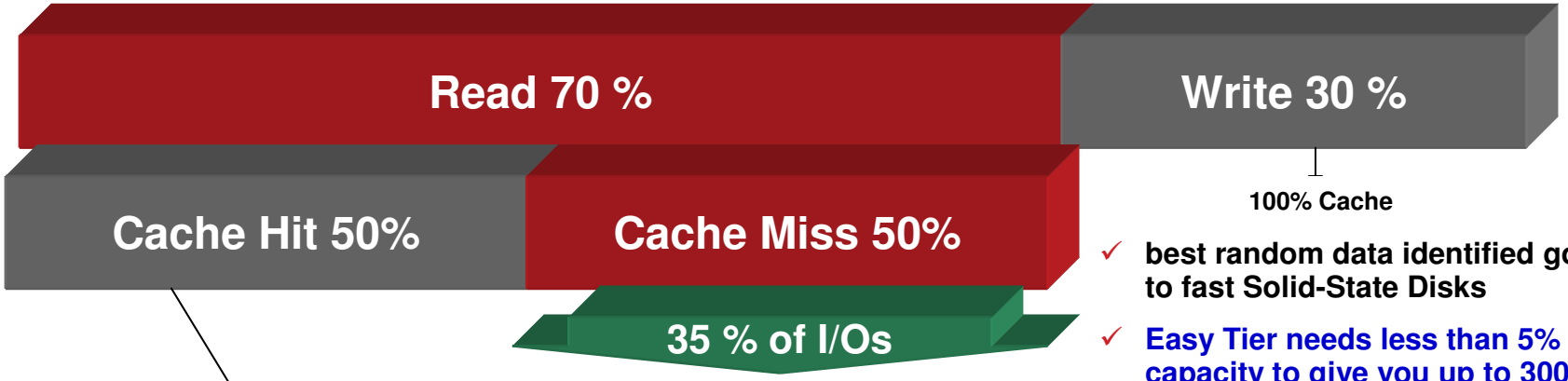
Linux data on FBA disk

DS8000, XIV, Storwise, FlashSystem

Addressing I/O pressures together

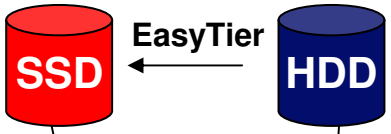


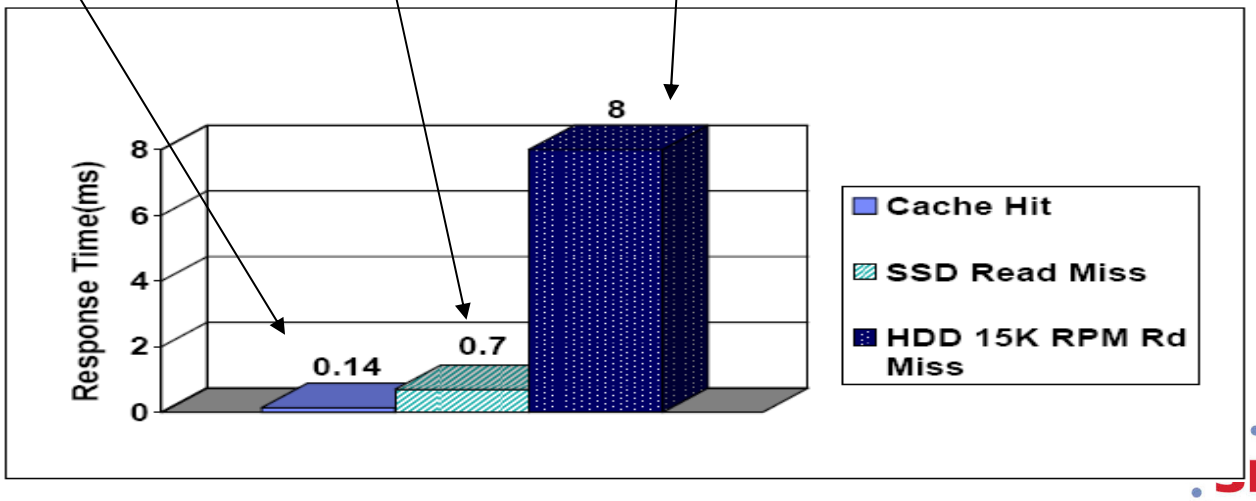
Easy Tiering => Performance made easy



100% Cache

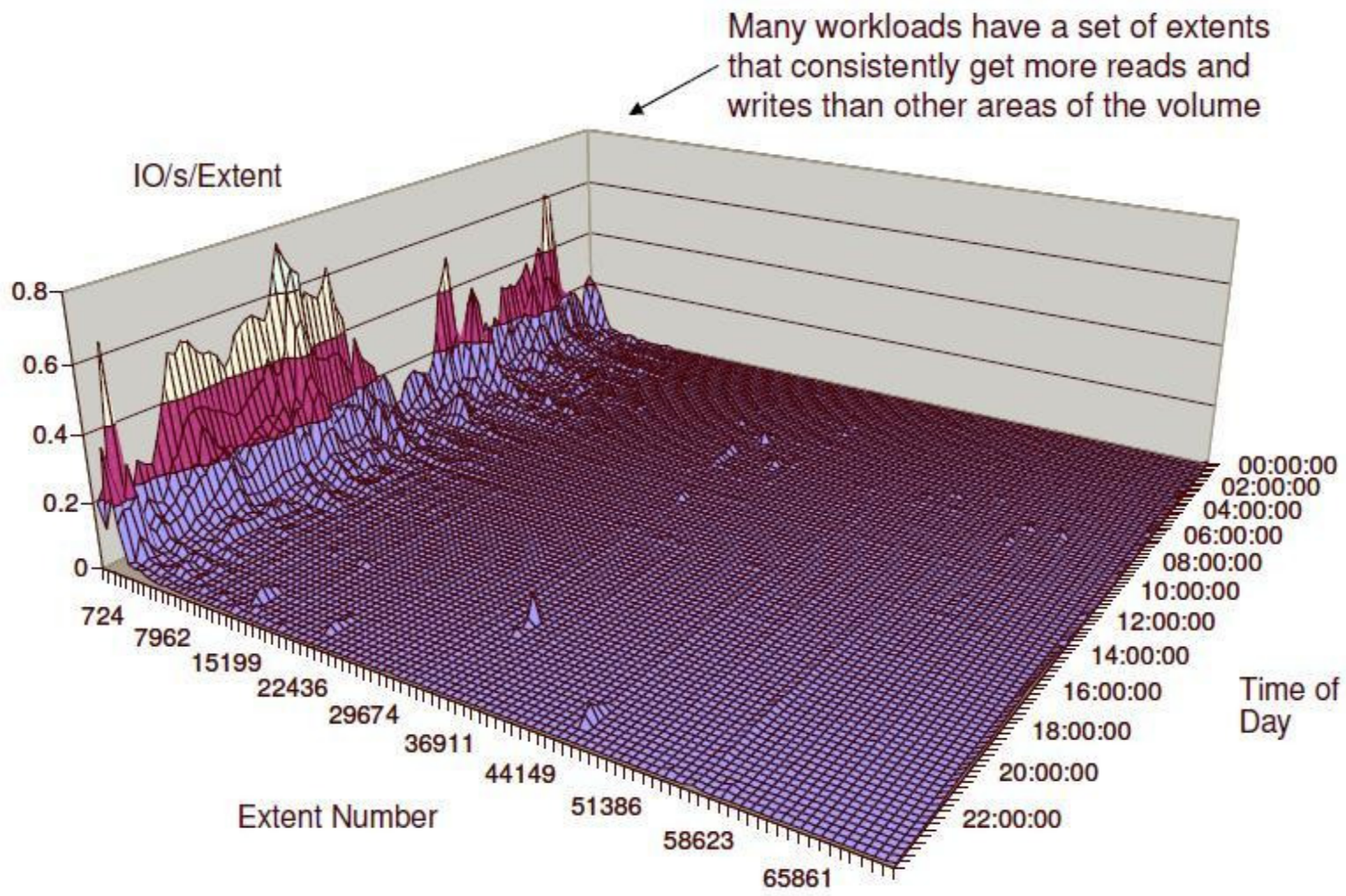
- ✓ best random data identified goes to fast Solid-State Disks
- ✓ Easy Tier needs less than 5% capacity to give you up to 300% performance



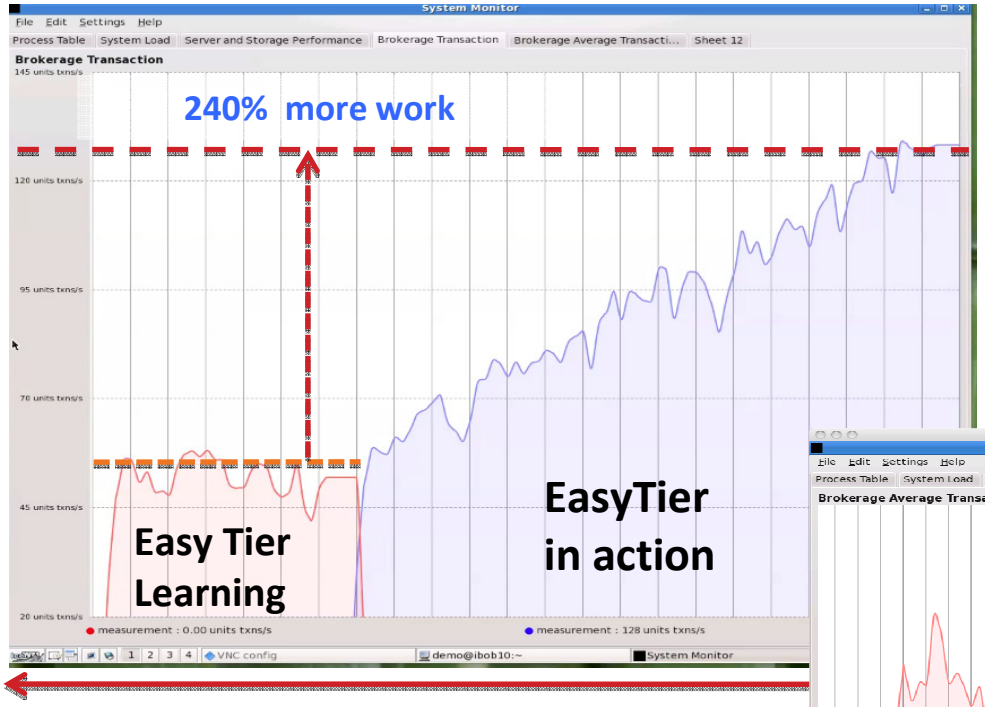


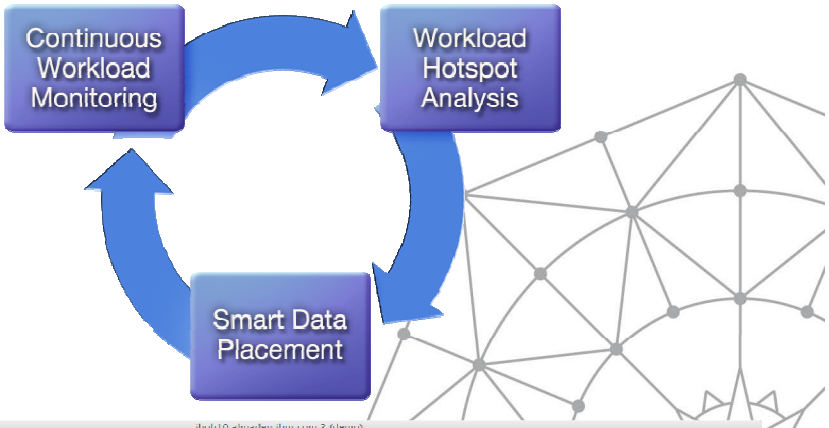
EasyTier for
- **DS8000**
- **Storwize V7000**
- **SVC**

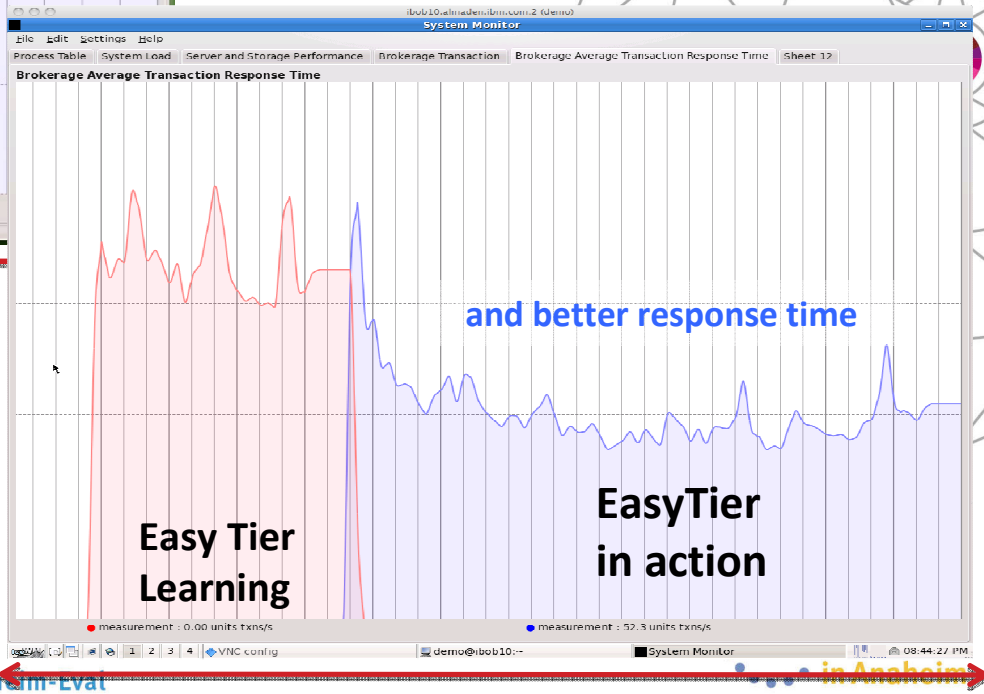
Easy Tiering – the I/O statistic



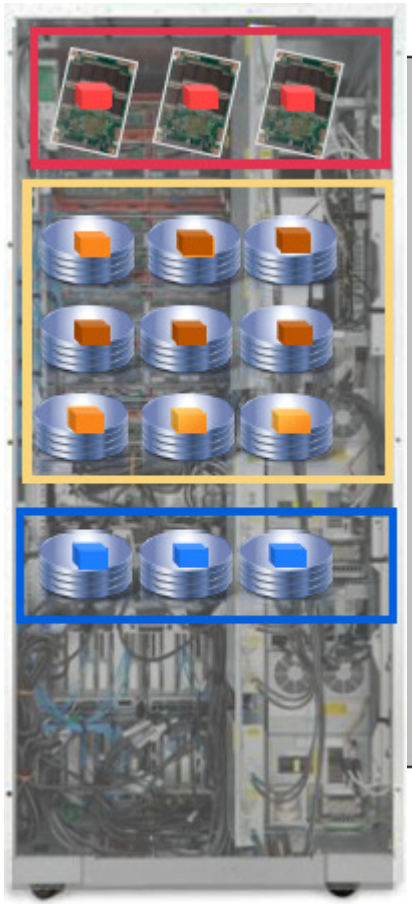
EasyTier – more transactions + better response time

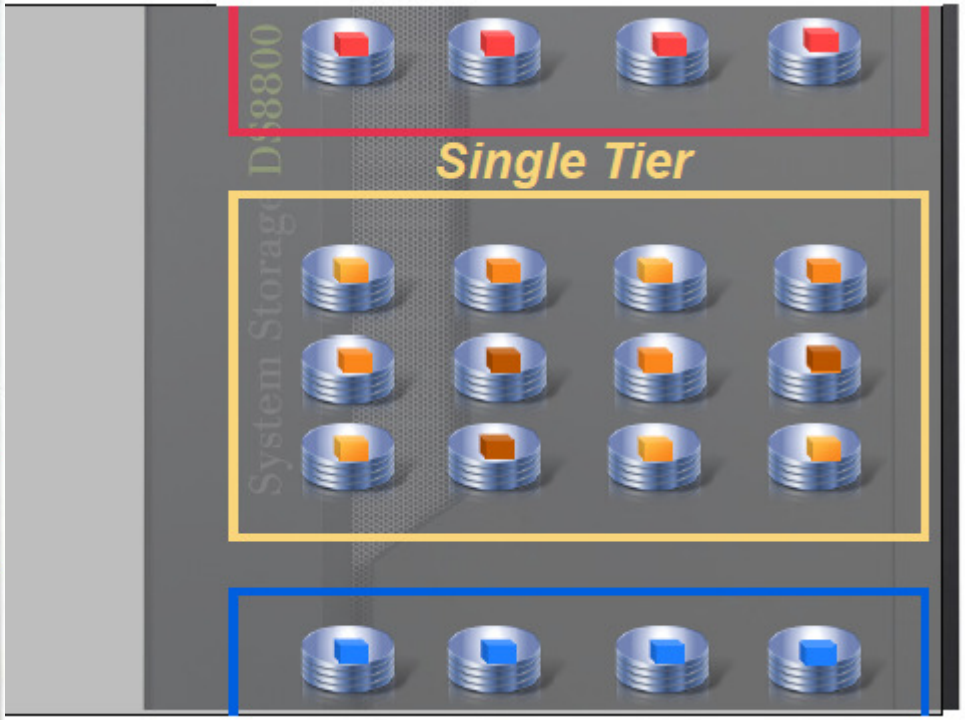


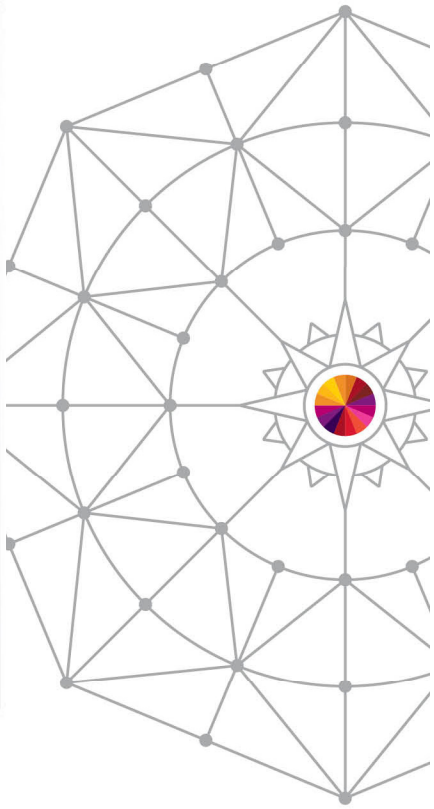




Storage tiering - for increased performance



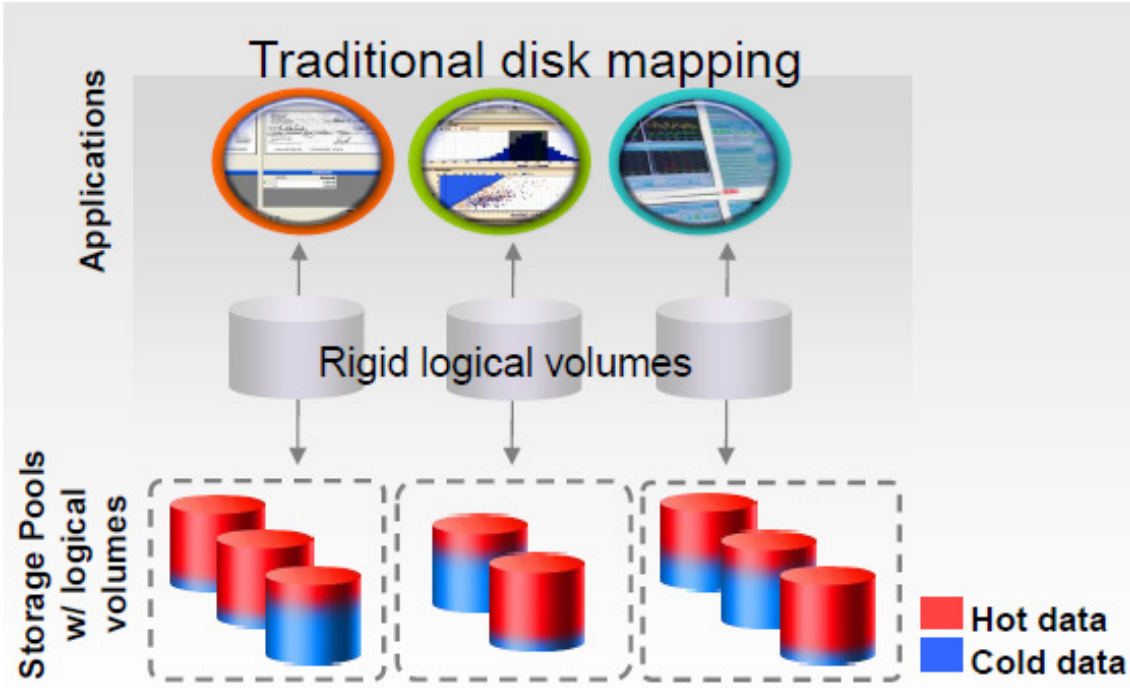




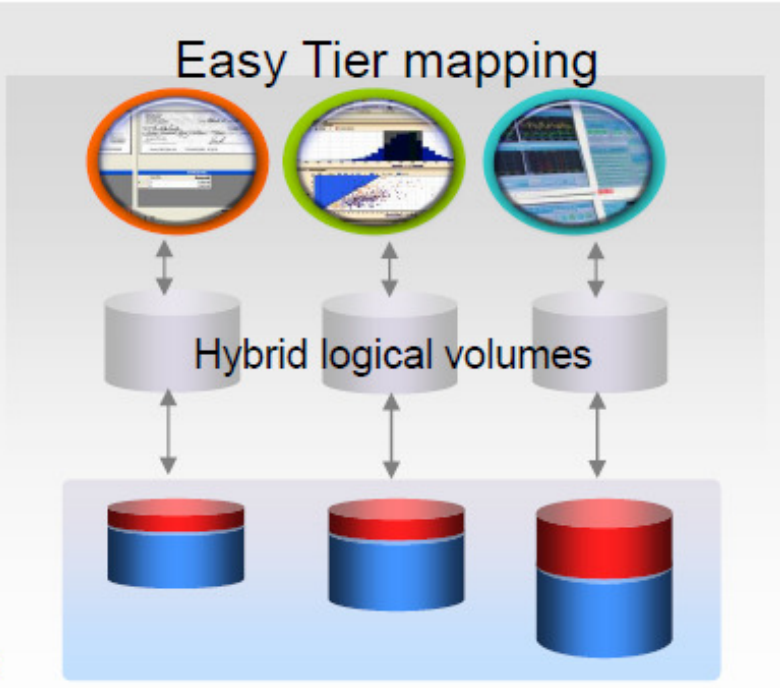
Continuous rebalancing within each tier maintains peak performance across all drives

Set it and forget it...

Storage tiering - for increased performance

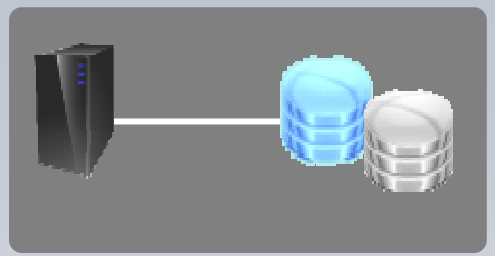


Volumes have different characteristics. Applications need to place them on correct tiers of storage based on usage.



All volumes appear "logically" homogenous to apps. But data is placed at the right tier of storage based on its usage through smart data placement and migration.

IBM Flash in Storage



Flash in Storage

FlashSystem 820 & 840
FlashSystem Solutions

All Flash Array
Fibre Channel, InfiniBand, FCoE
4TB to 48TB per 3U
IBM MicroLatency™

DS8870

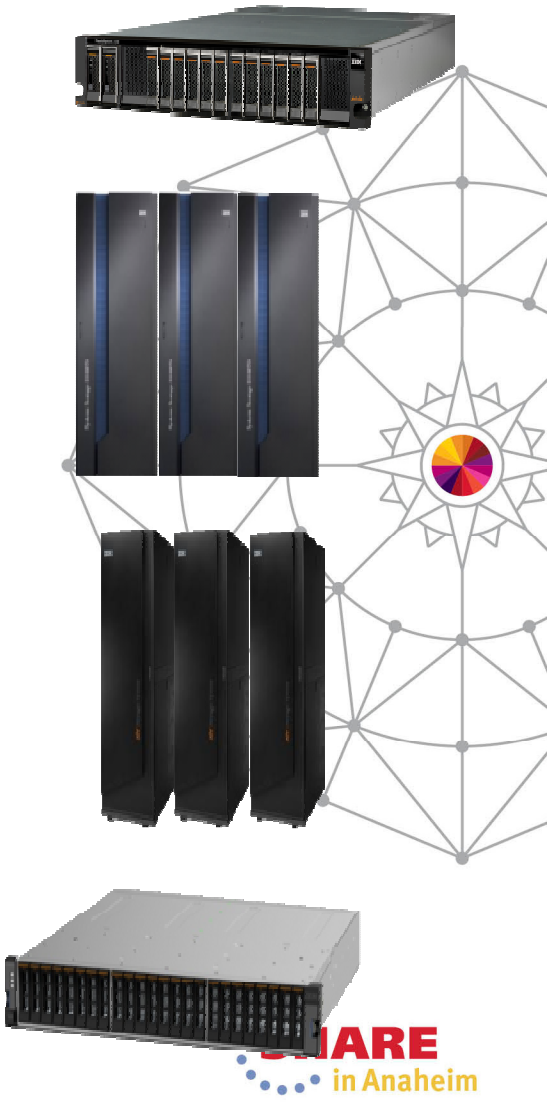
Enterprise All Flash or Hybrid
Fibre Channel, FICON
Easy Tier

XIV

Cloud-optimized scale-out
Fibre Channel, iSCSI
Up to 12TB Flash cache

Storwize

All Flash or Hybrid
Software defined storage



Business Critical Data Economics on *DS8870*
Flash Optimized - delivers immediate ROI…

Today

- ✓ ***Superior Data Economics***
- ✓ ***Leading Performance*** *Across Varying & Dynamic Workloads*
- ✓ ***Proven Security & High Availability*** *Resiliency Architecture & Multi-site Disaster Recovery*

New

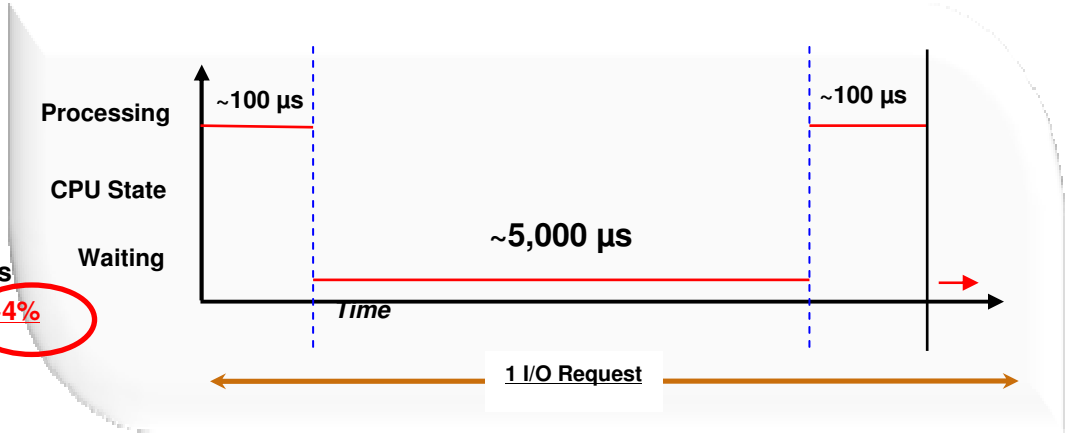
- ***Flash-at-scale performance***
  - *Consistent Microsecond Latency*
  - *Up to 20% IOPS Improvement*
- ***Footprint reduction up to 30%*** *vs. HDD*
- ***Power reduction up to 60%*** *vs. HDD*
- ***New Entry System*** *Scalable to 1,056 drives*

Microlatency effects storage run time

I/O Serviced by Disk

1. Issue I/O request ~ 100 µs
2. Wait for I/O to be serviced ~ 5,000 µs
3. Process I/O ~ 100 µs

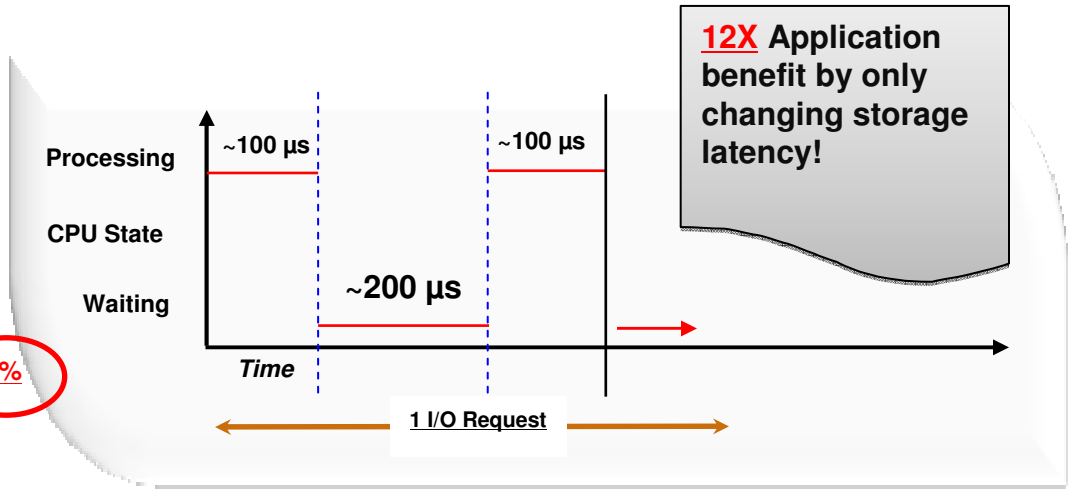
- Time to process 1 I/O request = 200 µs + 5,000 µs = 5,200 µs
- CPU Utilization = Wait time / Processing time = 200 / 5,200 = ~4%



I/O Serviced by IBM FlashSystem

1. Issue I/O request ~ 100 µs
2. Wait for I/O to be serviced ~ 200 µs
3. Process I/O ~ 100 µs

- Time to process 1 I/O request = 200 µs + 200 µs = 400 µs
- CPU Utilization = Wait time / Processing time = 200 / 400 = 50%



FlashSystem 820 Oracle performance results with Linux on System z

Performance of Linux on System z with FlashSystem and Oracle:

I/O bound Oracle databases can benefit from IBM FlashSystem over spinning disks.

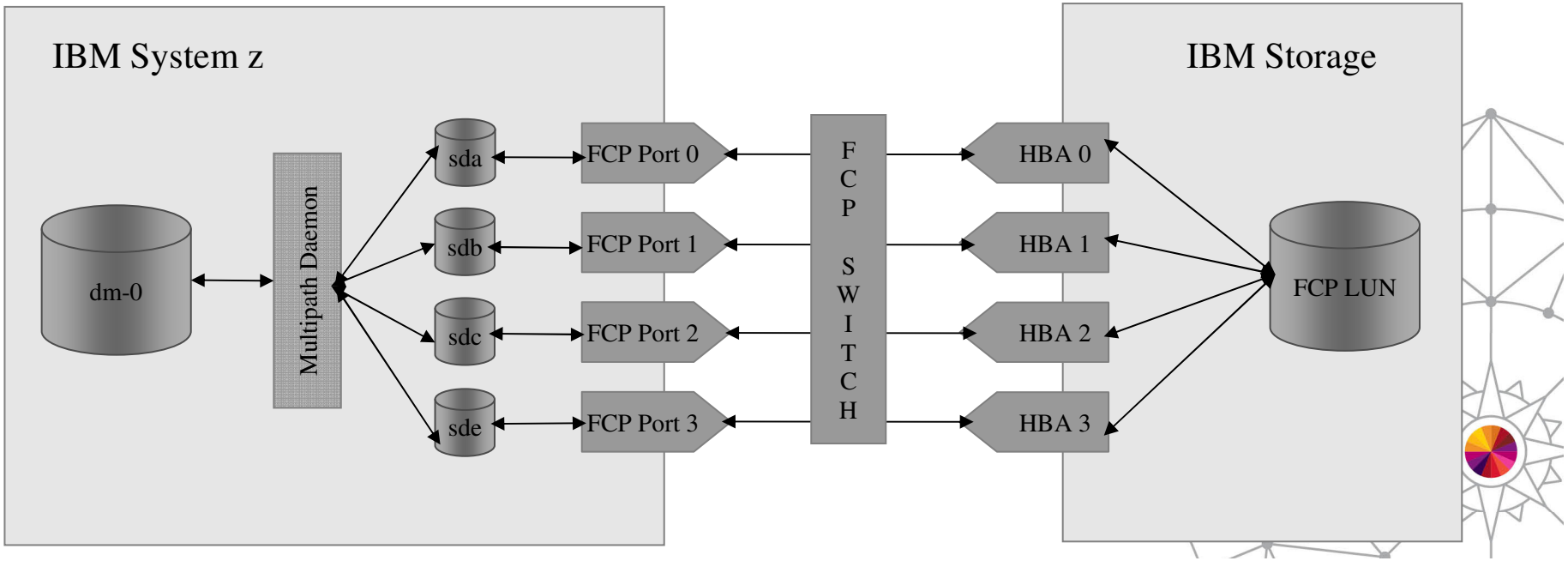
- **21x** reduction in response times
- **272%** improvement in CPU utilization
- **957%** improvement in IO wait times

*System z FiconExpress **8s** I/O cards can provide an additional 10% throughput running with FCP*

Now certified to attach to System z for Linux on System z, with or without an SVC

Source: IBM Actual Results and Estimates

FCP devices and efficient multipathing



- Multipath setup (device mapper)
  - use a 1:1 relation ship between FCP ports on System z and on the HBA on storage server,
    - except the bandwidth is different, (i.e one side 8GBits/sec to the switch and the other side 4GBits/sec)
- Performance relevant multipath parameters are path_grouping_policy and rr_min_io
  - **Note:** parameter rr_min_io is named **rr_min_io_rq** in SLES11 SP2, RHEL 6.2
  - Be aware that there are defaults predefined for certain disk devices (check with *multipath -t*) which overrules the settings in the defaults section

Multipath options change

/etc/multipath.conf “path_grouping_policy multibus”

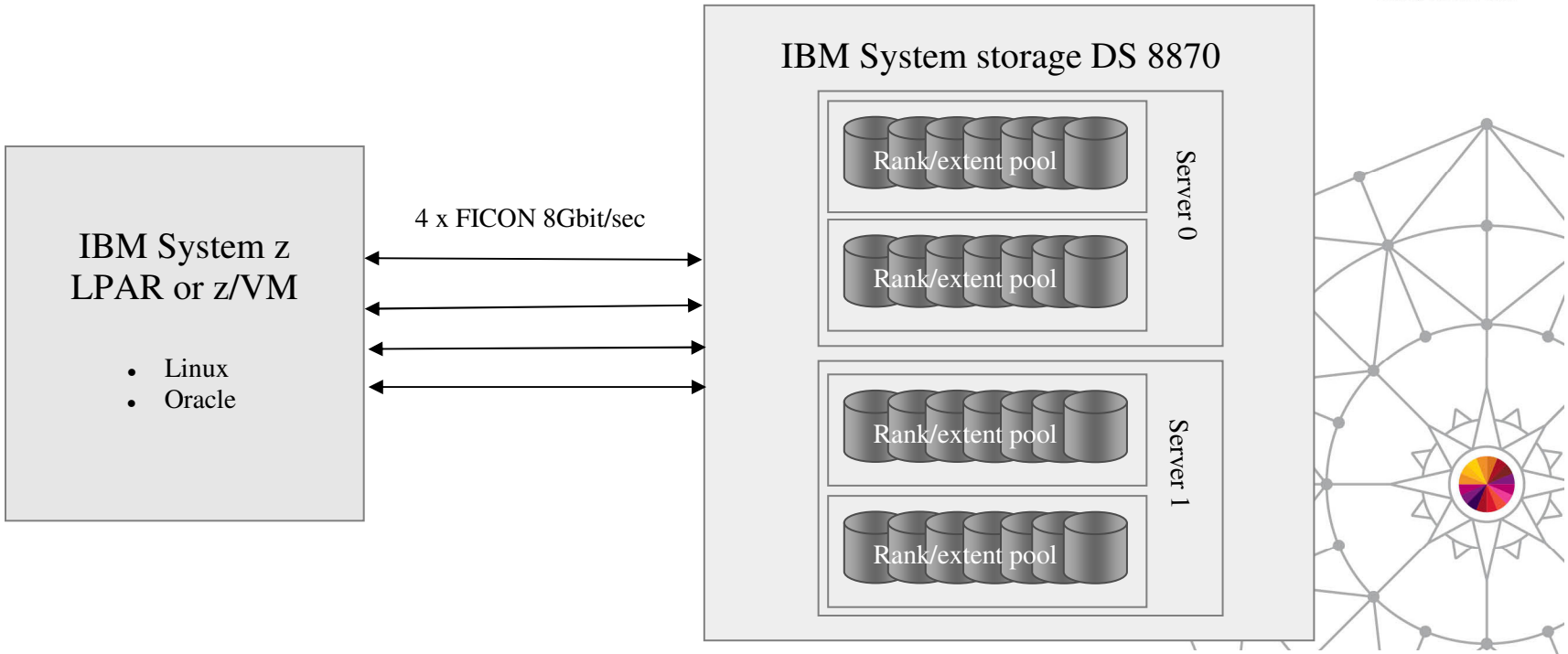
- rr_min_io defines the number of I/O operations that are sent to path before switching to the next (round robin)

- In RHEL6.2+ / SLES11 SP2+ now called **rr_min_io_rq**

- **Device mapper does accept old rr_min_io value**
  - **but does nothing with it**
- *Watch for upgrades!*
- Adapt **rr_min_io_rq** value – *it is storage dependent*
  - DS8K with rr_min_io=100 provided good results
  - XIV recommends rr_min_io=15

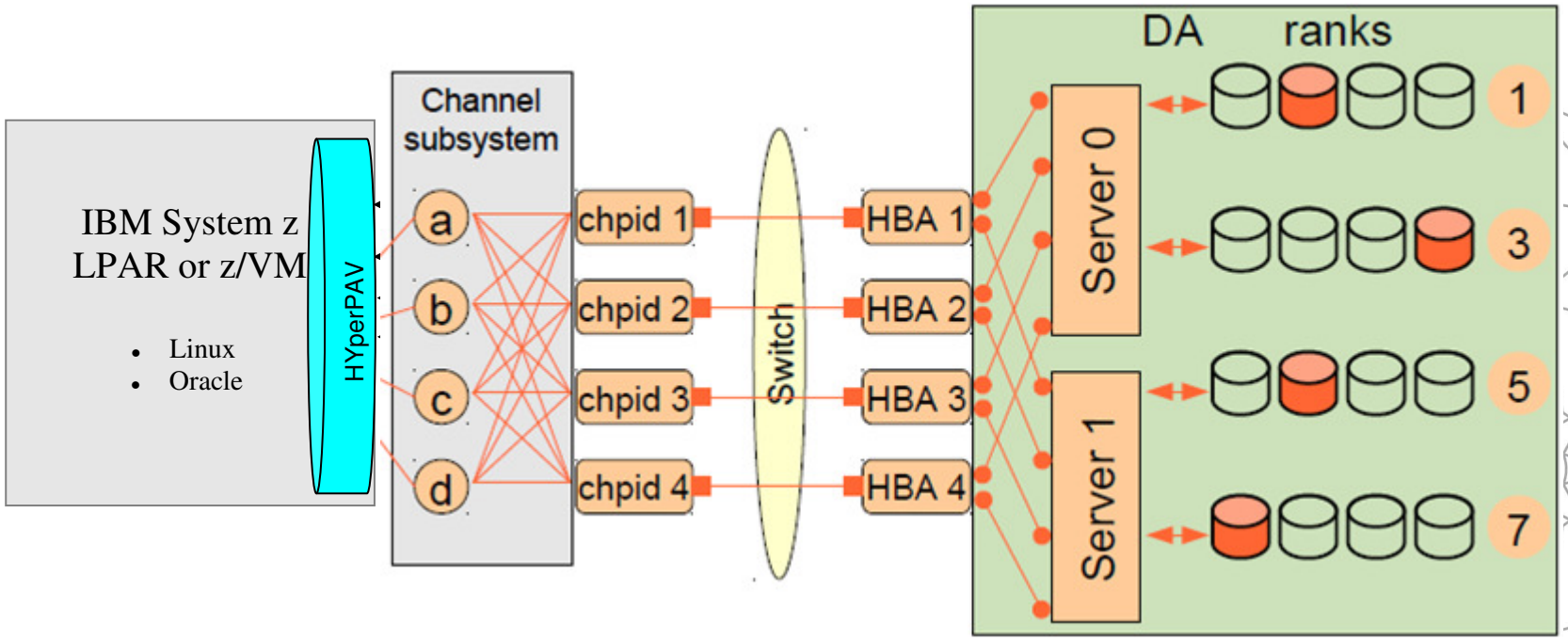
Thre default value of rr_min_io should be checked / changed !

FICON Environment setup



- LPAR or z/VM based environment using Oracle
  - ECKD disks attachment type
- FICON attachment
  - Possibility for HyperPAV devices

FICON Environment setup with HyperPAV

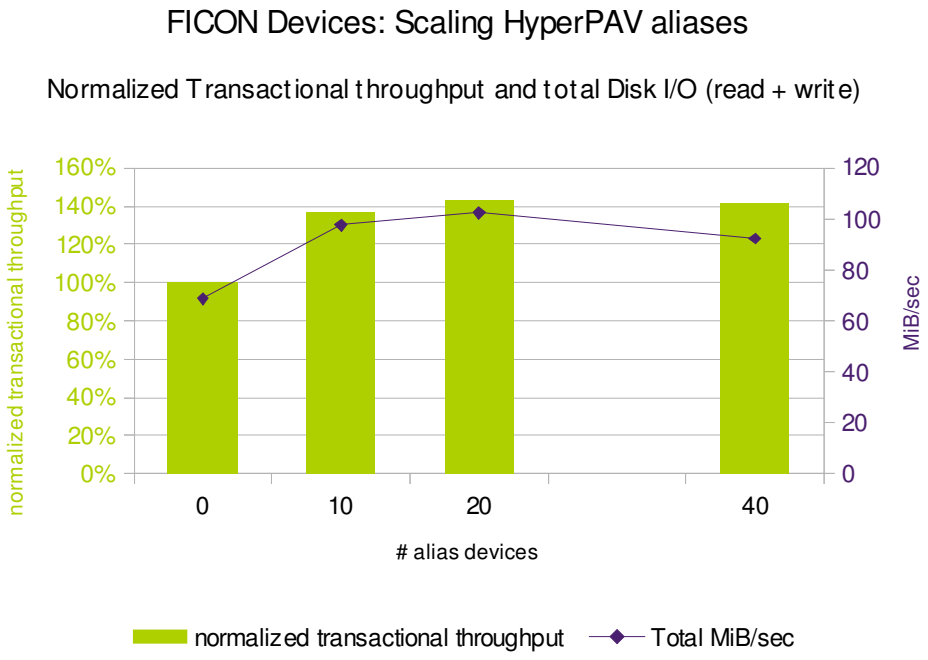


- LPAR or z/VM based environment using Oracle
  - ECKD disks on different ranks attached
- FICON attachment
  - HyperPAV devices

ECKD Disc I/O – think about it

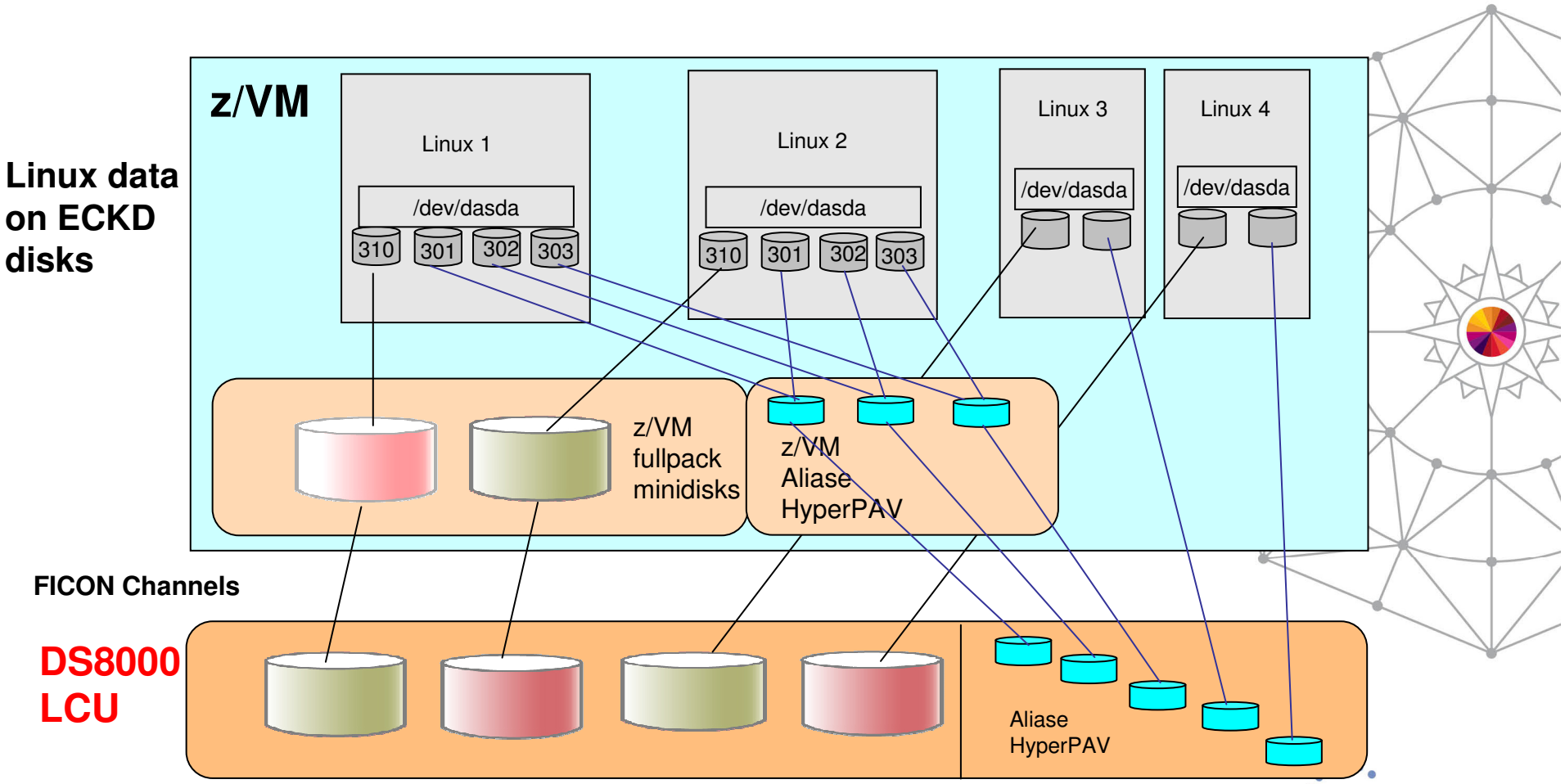
- FICON Express dedicated storage connection to multiple Host Bus Adapters (HBA)
- Multiple channels, more CHPIDs, cache and more paths
- ECKD: storage pool striping / spread data across multiple arrays
- PAV / Hyper PAV avoids subchannel busy
- LVM with striping, mount with option 'noatime‘
  (Inodes statistics will not be updated for each operation)

Oracle – DASD disk devices and HyperPAV



- ECKD protocol: when one I/O request is sent to a subchannel (e.g a disk), the subchannel is busy and blocked.
  - Impact scales with the level of parallelism, typical database disk I/O is highly parallelized
  - Large disks like mod 27 or mod 54 can become a serious bottle neck
- HyperPAV devices are a suitable solution to overcome that issue
  - 40% improvement with 20 alias devices
  - Slight degradation in regard to the maximum with 40 alias devices

HyperPAV in z/VM and Linux on z



Summary disk connections for Oracle

- Storage server:
  - Storage pool striping with many ranks
  - Use disks from both internal servers
- Infrastructure:
  - Use the highest possible link speed of the FICON/FCP channel, but ensure it is supported from all elements in the path
- FICON Disks
  - Use an appropriate amount of HyperPAV aliases (e.g. 10 per LCU), the larger the more are needed
  - FICON devices are easy to administer, they save CPU due to the use of the SAPs
- FCP Disks
  - Use multipath policy multibus and rr_min_io_rq/rr_min_io in the area of 100
  - FCP devices are driven by the CPU/IFLs, but this allows higher throughput values
- Further tuning options: FCP queue depth (default 32)
  - Indicator : iostat/sar reports larger values for avgqu-sz (for example > 10)

Evaluate your Storage

- Test your storage subsystem, (without Oracle database installed !) with disk utility **Orion** provided by Oracle to help decide various storage configurations for your Oracle Databases.
  http://www.oracle.com/technetwork/topics/index-089595.html

- You can also use the Oracle I/O calibrate routine from the script provided from the Oracle Database PL/SQL Packages and Types Reference 12c Guide:
  http://docs.oracle.com/cd/E16655_01/appdev.121/e17602/d_resmgr.htm#ARPLS67598

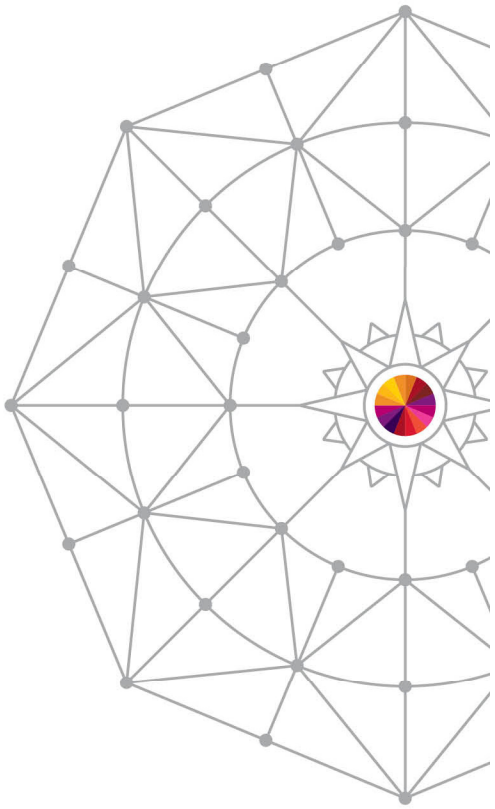
- Oracle's /IO calibrate routine does not harm the database or the underlying data files. I/O calibrate does require a database to be created.

Agenda

1. Infrastructure for Oracle Database workloads
   - Disk options
   - FICON devices with HyperPAV
   - FCP devices and Multipathing
   - Comparing FICON and FCP disk devices

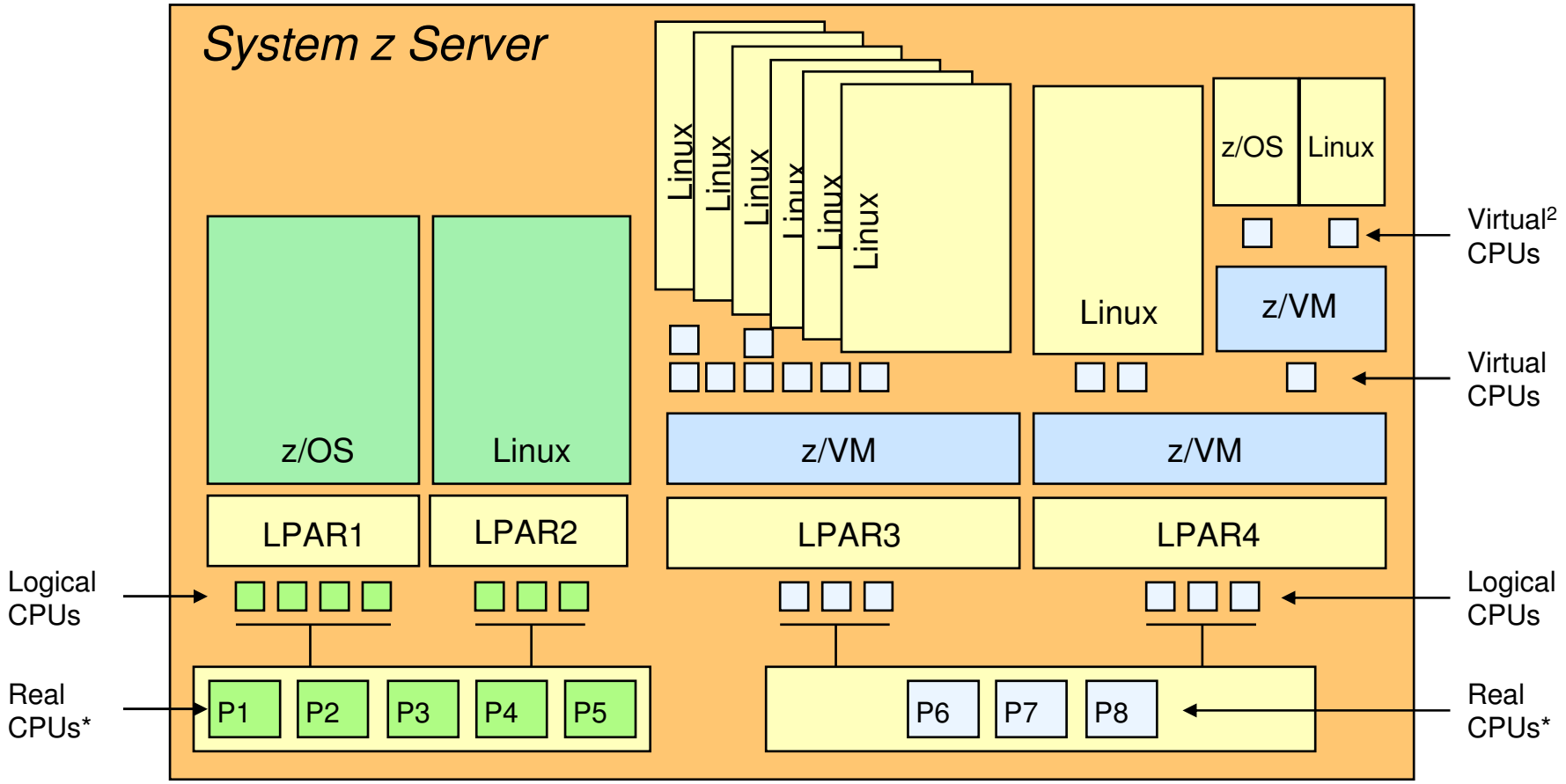
2. **Virtualization layers**
   - **System z LPAR Virtualization**
   - **z/VM Virtualization**
   - **Oracle 12c virtualization**

3. Oracle Database layer
   - Optimizing Oracle 12c setup
   - Linux on z and Oracle 12



Effective System z Virtualization

Note: There are typically dozens or hundreds of Linux servers in a z/VM LPAR.



P1 – P8 are Central Processors (CP) or Integrated Facility for Linux (IFL) Processors

* - One shared Pool of CPUs per machine (CEC) only

Virtualization Recommendations

Use Virtualization method based on workload

LPAR suited Oracle environments

- Non-dynamic Production servers run in LPAR
  - Sharing resources is limited (i.e Memory, CPU)
  - Better behavior for high number of users
- Active-standby HA with 2 LPARs sharing CPU
- Servers with 24X7 and > 100GB in LPAR
- Combination of LPAR and z/VM workload possible

LPAR Virtualization

- When processor utilization becomes high, LPAR weights must be used to prioritize the LPARs.
- From best practices: an LPAR should not have excessive logical CPUs defined, since the weight is spread between logical CPUs.
- The ratio of logical CPUs defined to physical CPUs is proportional to LPAR overhead.
- Be aware in z/VM 6.3, if too many virtual CPUs are defined in the LPAR for the current workload, z/VM will “park” one or more logical CPUs and have the weight distributed across fewer virtual CPUs (Vertical CPU management)

Virtualization Recommendations – z/VM

z/VM benefits for Oracle database servers

- Dynamic workload runs best in z/VM
- Variation of resources & servers
- Servers with small peaks over the day
- Large number of servers with staged workload peaks
  - Test, QA, pre-production server groups
- z/VM share everything has big advantages
  - Same network for many guests
  - Shared memory over many Linux guests

z/VM can compete with new Oracle container concepts

z/VM Recommendations for improved I/O

- Use EDEV with FCP disks for paging
  - can parallelize I/O (not possible with ECKD)
  - Avoid subchannel busy contention
  - increases throughput by factors
- Use Spool with FCP
- Use DCSS segments as EDEV
  - decreases initial load time

- Avoid to use EDEVs for data disks
  - no HyperPAV possible
  - subchannel devices (ECKD) w/o HyperPAV suffer due to subchannel busy contention
  - emulation FCP -> EDEV is CPU intensive, slower than DASD w/o HyperPAV

z/VM 6.3 virtualization

- IBM z/VM V6.3 extends the mainframe virtualization limits
- Improved economies of scale with z/VM support for 1 TB of real memory
- Better performance for larger virtual machines (page reorder is obsolete)
- Reduced LPAR sprawl for additional horizontal scalability
- Considerably more virtual machines can be consolidated into a single LPAR, depending on workload characteristics.
- Expanded memory obsolete
- Improved performance with HiperDispatch

Oracle 12c pluggable concept

- Oracle Database 12c Release 1 introduces a new multi-tenant architecture that makes it easy to deploy and manage databases
  - Oracle multitenant pluggable databases
  - For consolidating multiple databases
  - Automatic Data Optimization with Heat Map
- Virtualization based on container concept
  - Common ASM
  - Common SGA for pluggable databases

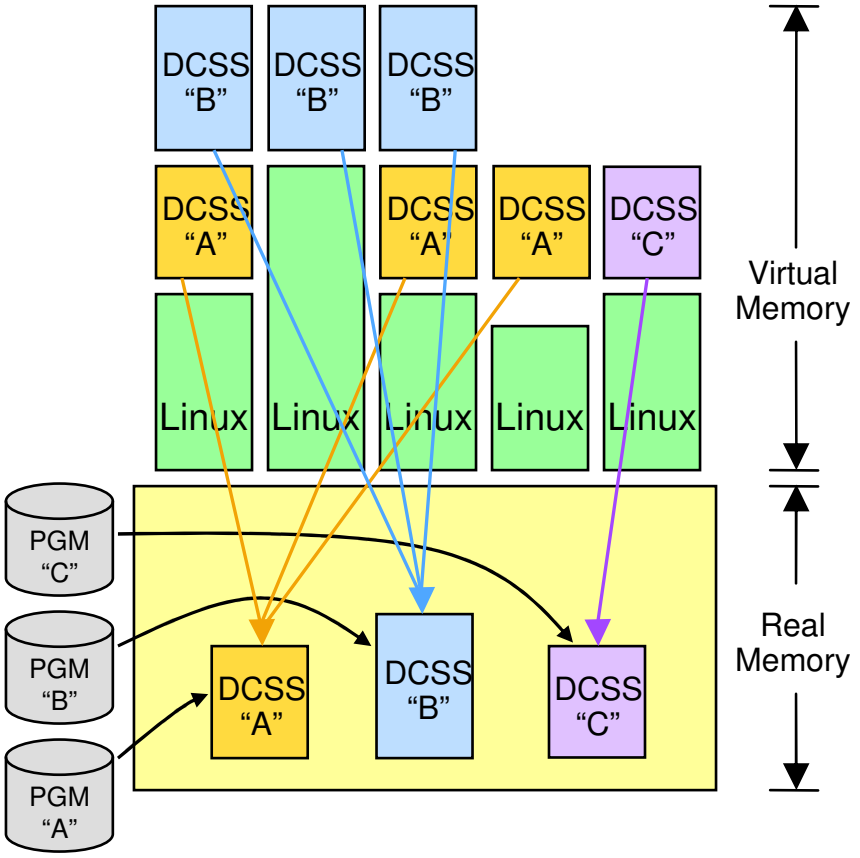
- A pluggable database may be moved from one container to another simply by unplugging it from one and plugging it into another
- Software patches and upgrades may be applied at the container level. Upgrading the one container database also upgrades all pluggable databases at once.

Effective Virtualization with Linux on z and z/VM shared memory

Linux Shared Memory Exploitation for many Virtual machines

z/VM Discontiguous Saved Segments (DCSS)

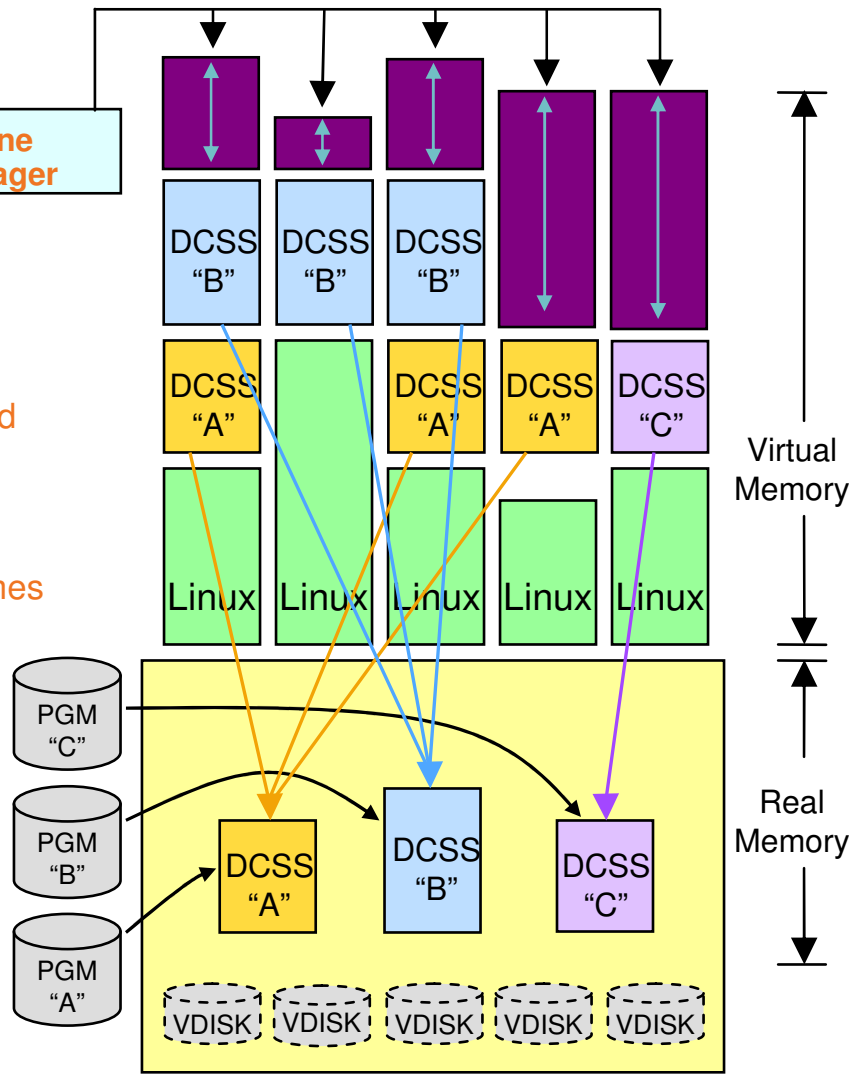
- **DCSS support is Data-in-Memory technology**
  - Share a single, real memory location among multiple virtual machines
  - Can reduce real memory utilization
- **Use Cases:**
  - **As fast Swap device**
  - **For sharing read only data**
  - **For sharing code (e.g. program executables/libraries)**
- **The large DCSS allows the installation of a full middleware stack in the DCSS (e.g. WebSphere, Databases, etc)**
- **The DCSS becomes a consistent unit of one software level**
- **NSS – Named Saved System – for a bootable Linux image**



http://public.dhe.ibm.com/software/dw/linux390/perf/ZSW03186USEN.PDF

Effective Virtualization with Linux and z/VM SRM

- **Real memory constraint corrected by z/VM**
  Virtual Machine Resource Manager (SRM)
  - Linux images signaled to reduce virtual memory consumption
  - Demand on real memory and z/VM paging subsystem reduced
- **z/VM Virtual Disks in Storage (VDISK)**
  - Simulate a disk device using real memory
  - Use VDISKs for Linux swap devices instead of real disk volumes
  - Reduces demand on I/O subsystem
- **Linux guest: shared program executables**
  - Execute-in-place (xip2) file system
  - Access to file system is at memory speeds; executables are invoked directly out of the file system (no data movement required)
- **Data-in-Memory technology**
  - Share a single real memory location among multiple VMs
  - Reduce real memory utilization



http://public.dhe.ibm.com/software/dw/linux390/perf/ZSW03186USEN.PDF

inactive virtual memory

Optimize in z/VM with DCSS

- A DCSS can include pages up to 512 GB in size, which exceeds the previous limitation of 2 GB.
- A large DCSS can reside almost anywhere in addressable storage.
- The Linux dcssblk device driver concatenates DCSSs of 2GB, multiple DCSSs appear to Linux as a single device

General Recommendations – z/VM

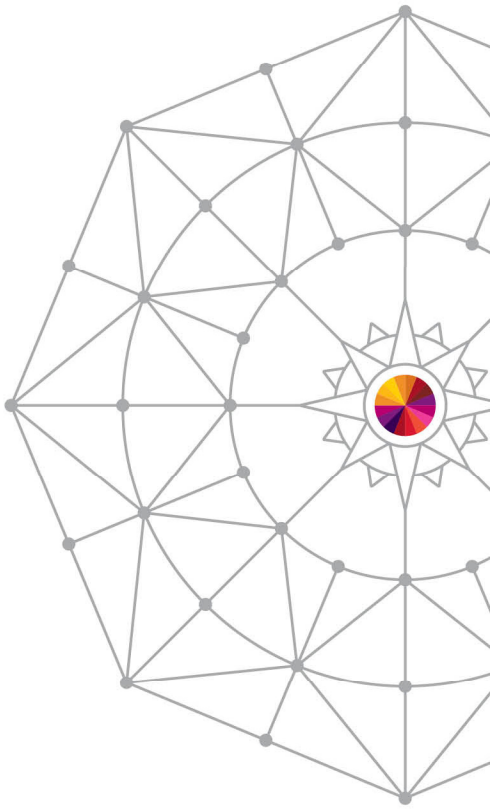
- **z/VM Performance Toolkit**
- Ensure the virtual to real memory ratio stays in an appropriate range for the workloads
  - Indicators of impact:
    - z/VM Paging activity
      - *Report 'User Paging Activity and Storage Utilization' (UPAGE, FCX113)*
      - *Columns: 'X>DS' paging to DASD, critical: Reads paging from DASD*
    - z/VM Guest Waits
      - *Report 'Wait State Analysis by User' (USTAT,FCX114)*
      - *Especially columns %PGW, %PGA, and %CFW*
    - z/VM CPU load
      - *Report 'System Performance Summary by Time' (SYSSUMLG, FCX225)*
      - *Report 'General CPU Load and User Transactions' (CPU, FCX100)*
- Disable Page reorder for guests larger than 8 GB (prior z/VM 6.3)
  - Solved with z/VM 6.3
  - Find more information at http://www.vm.ibm.com/perf/tips/reorder.html

Agenda

1. Infrastructure for Oracle Database workloads
   - Disk options
   - FICON devices with HyperPAV
   - FCP devices and Multipathing
   - Comparing FICON and FCP disk devices

2. Virtualization layers
   - System z LPAR Virtualization
   - z/VM Virtualization
   - Oracle 12c virtualization

3. **Oracle Database layer**
   - **Optimizing Oracle 12c setup**
   - **Linux on z and Oracle 12**



Oracle Storage access options

- ECKD DASD with HyperPAV
  - DASD uses less CPU but lower throughput
- FCP/SCSI provides greater throughput – heavier CPU
  - FCP/SCSI LUN with multipathing

- LVM for Disk storage access
  - Total control of disk placement
  - Self control of striping
  - Self control of failover

- Oracle Automatic Storage Manager (ASM) for Disk storage access
  - Oracle controls multipathing
  - ASM controls disk striping
  - ASM controls failover

ASM and LVM should be used exclusively – not combined

Automatic Storage Management (ASM)

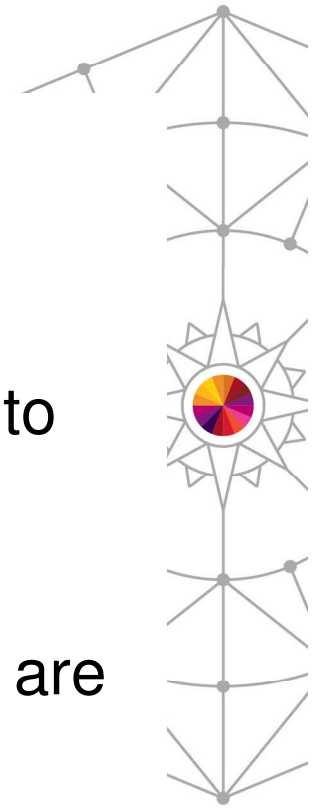
- Automatic Storage Management (ASM) instance
  - Instance that manages the disk group metadata
  - RAC requires shared disks and ASM
  - ASM is both a volume manager and a file system for database files
- Disk Groups
- Logical grouping of disks
- Determines file mirroring options

Oracle 12c

- Self-contained PDB for each application
- Applications run unchanged
- New virtualization concept with containers
  - Multitenant Container Database (CDB)
- New portability concept
  - Pluggable Database (PDB)
- Shared memory and common background processes
- More applications per server
- Common operations performed at CDB level
- Manage on container level (upgrade, HA, backup)
- Granular control when appropriate

Pluggable database flexibility

- New model supports many Pluggable PDBs
  - Up to 252 databases
- PDBs share common SGA
- Foreground sessions see only the PDB they connect to
- A PDB is fully backwards compatible with a pre-12.1 database
- Only small increments in memory as additional PDB's are added



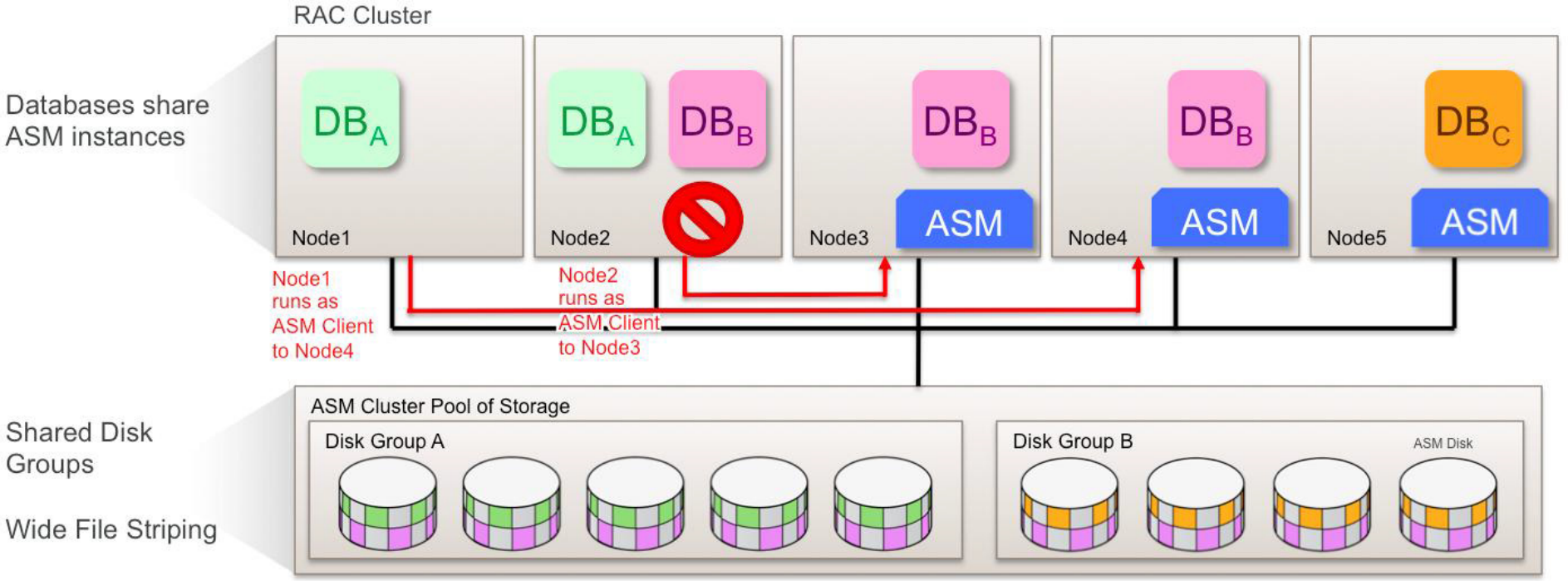
New Flex ASM in Oracle 12c

- New Flexible Architecture
- ASM instance on separate Server to Database Instance
- Smaller clusters of Oracle ASM instances can support more database clients while reducing the Oracle ASM footprint for the overall system.
- With Oracle Flex ASM, you can consolidate all the storage requirements into a single set of disk groups. All disk groups are mounted and managed by a set of Oracle ASM instances running in a single cluster.
- Private ASM network configurable

- Higher Redundancy:
- DB Servers are automatically relocated to another instance if an Oracle ASM instance fails. If necessary, clients can be manually relocated eg:
- SQL> ALTER SYSTEM RELOCATE CLIENT 'client-id';

Note: Releases before Oracle Database 12c require local Oracle ASM

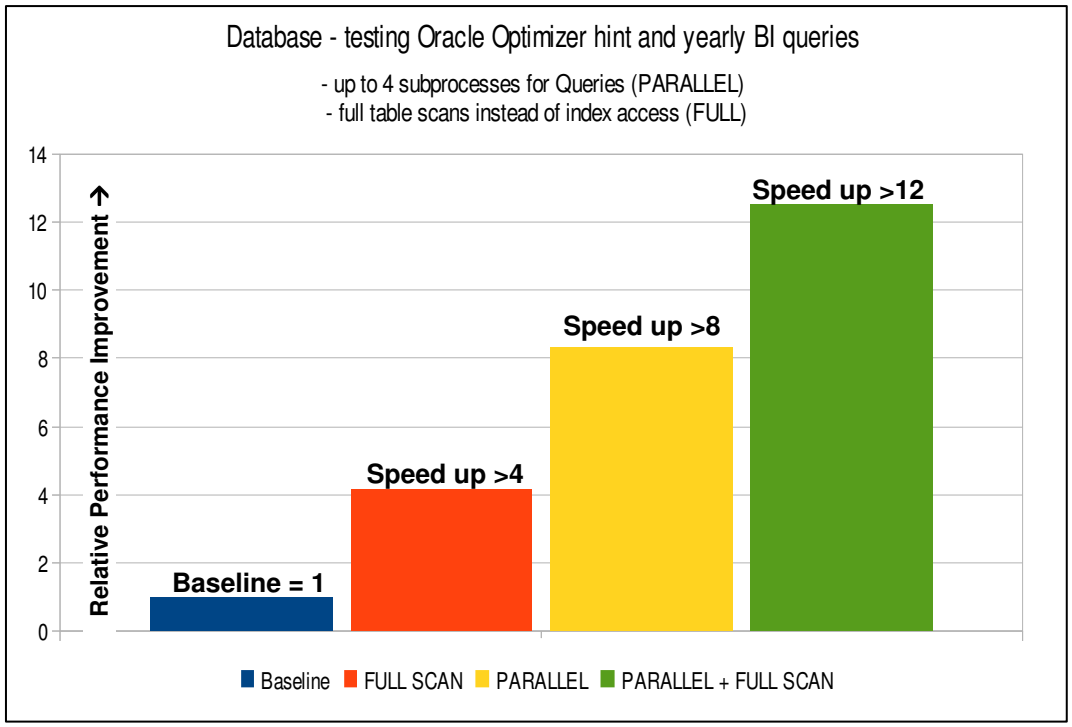
Flex ASM



http://www.oracle.com/technetwork/products/cloud-storage/oracle-12c-asm-overview-1965430.pdf

Oracle DB Tuning Activities

- **Oracle optimizer hints** are specific for the SQL statement where specified
  - 'FULL' force table scans vs index access
  - 'PARALLEL' forces breaking up the statement into parts which can be executed in parallel in the same time
  - 'PARALLEL' and 'FULL'



- Risks
  - Forcing a table scan can result in a severe performance degradation, when index access is the appropriate access method
  - There might be reasons that a certain statements can not be executed parallel, then the behavior will not change

Oracle parameters - Recommendations

- **Review existing optimizer hints!**
- Customer workload specific experience with Oracle optimizer hints:
  - Got very good improvements with the hints FULL(<table name>) and PARALLEL(<table name>, <number of CPUs>) for BI queries
  - Suggest to review existing optimizer hints. Examples:
    - Combination of full(t) and parallel_index(t, 12) seems to be contradictory because usage of full table scan or index are mutually exclusive
    - Degree of parallelism specified with 12 seems to be far too high for a system with 4 vCPUs. A typical level for parallelism is <amount of vCPUs> or <amount of vCPUs + 1>, the upper limit is no more than 2X the number of cpus/virtual cpu
  - For Oracle 11g consider to specify parallel_degree_policy=AUTO instead of explicit optimizer hints to let Oracle decide about parallelism
- **Log Setup**
  - Place redo logs on separate disks
    - Single disks are sufficient, striped LVM not needed
    - Ensure to have no other activity on these disks
  - Recommendation: Usage of larger log files
    - e.g. 4x 1 – 1.5 GB to reduce the frequency of log switches

Oracle 12c adaptive execution plans

- **Adaptive Execution Plans**
- Defer final plan for statement until execution time
- Contains pre-determined subplans
  - e.g.Switch nested loop joins to hash joins
- Dynamic Statistics Collection
  - In-flight Statistics collection
- Optimizer picks final plan based on cardinality observed during execution
- Requires following parameters:
- **OPTIMIZER_FEATURES_ENABLE=12.1.0.1 (OR LATER)**
- **OPTIMIZER_ADAPTIVE_REPORTING_ONLY=FALSE**
  - reporting-only mode is off, and the adaptive optimizations are enabled as usual.

General Recommendations - Linux

Memory requirements:

- Don't over-configure Linux memory
  - Excess memory allocated to the Linux guest is used by Linux for I/O buffer and File system cache
  - In a virtualized environment under z/VM, oversized guests place unnecessary stress on the VM paging subsystem
  - Real memory is a shared resource, caching pages in a Linux guest reduces memory available to other Linux guests.
  - Larger virtual memory requires more kernel memory for address space management.

- Consider setting *vm.swapiness* to 0 (sysctl.conf) for all systems which are running primarily databases using page cache I/O
  - Defines a preference to reuse page cache pages instead of swap application pages

Example of memory sizing for Oracle

- Standard Memory estimation = sum of:
  - Memory required for Linux Kernel: 512 MB
  - Memory required for Oracle SGA: As per DBA estimation
  - Memory required for Oracle PGA: As per DBA estimation
  - Memory required for Oracle ASM: 256 MB to 512 MB (If ASM is used)

  - Memory required for additional agents like OEM, Tivoli etc., as needed by the application
  - Linux Overhead requirements: 5 % of the total memory

Starting size = SGA + PGA + 0.5GB for Linux + ASM (if used)

- Memory over-commitment (relationship of virtual to real memory)
  - Limit/avoid memory over-commitment for critical production databases
  - Test/development guests can benefit from z/VM memory over-commitment capability

Linux Huge pages - recommendation

- With HugePages each OS pagetable mapping (virtual to physical) maps memory points to a 1MB page (as opposed to 4KB)
- Decreases page table overhead
- Pages are locked in memory and never swapped out, which provides RAM for shared memory structures such as SGA
- For Huge Pages you must use ASMM (Automatic Shared Memory Management)
- Huge Pages incompatible with AMM (Automatic Memory Management)
- Less operating System overhead, fewer cpu resources
- Recommended when SGA > 8GB

Linux Huge pages - recommendation

- If huge pages are configured, this amount of memory is no longer available for applications using 4K pages
  - Starting with Oracle 11g, the use of huge pages is done automatically
    - If the SGA can not be allocated as a whole in huge pages, the fall back is to allocate the whole SGA in 4KB pages, which can produce a heavy memory pressure.
  - Ensure to have enough huge pages defined that the full SGA from **all** Oracle databases in that system server fits into

- To verify usage of Hugepages
  - Monitor value of *HugePages_Free*
    - When starting Oracle the amount value of HugePages_Free must be lower (reduced by the SGA size)

Optimization for small environments

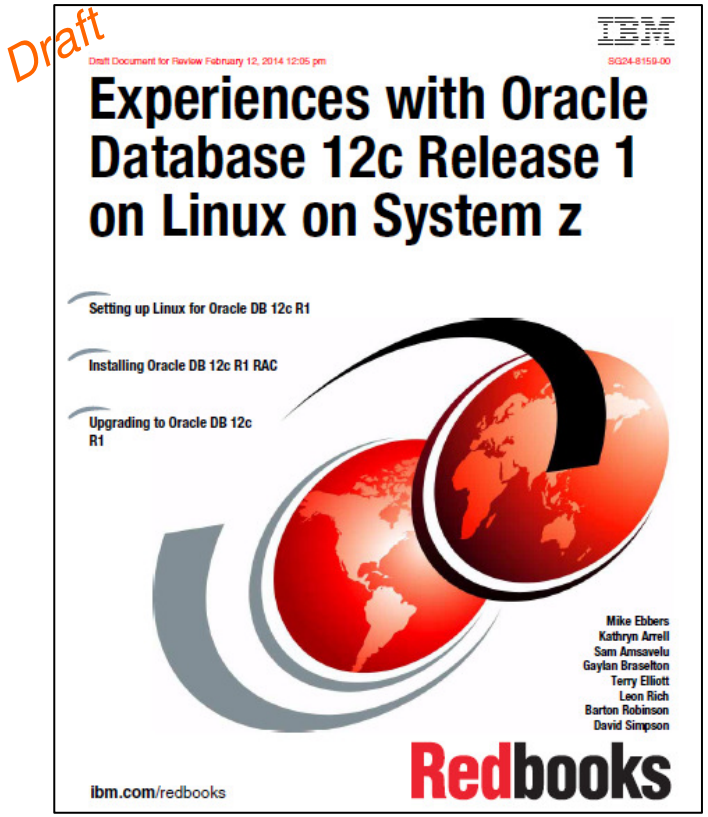
- Define more than one virtual CPU in Linux
  - This influences the execution plan for more parallel workload
- Over commitment of memory for SGA and PGA not recommended
- More CPU is used with fast I/O subsystem or FCP attached devices

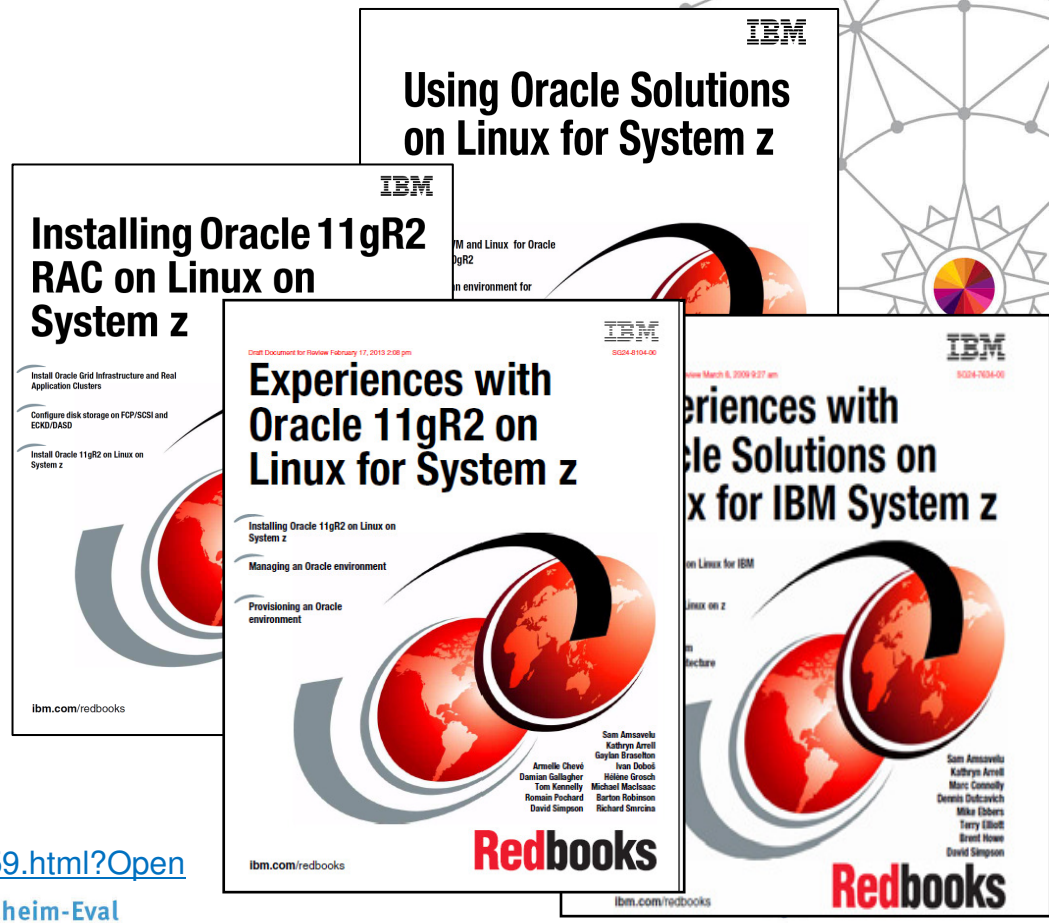
Summary Best practices – Oracle and Linux on System z

- Big database servers (SGA >100 GB) should be run in LPAR
- Busy Linux database servers as z/VM guest should be given enough guest memory so that paging for this guest can be minimized
- Page Reorder process in z/VM 6.3 not existent anymore
- No Expanded Storage for z/VM 6.3
- Size a Linux database server as z/VM guest that it just does not swap
- Use direct I/O for database files
  - Right-sizing the buffer pool is more beneficial than having additional Linux page cache
- Separate database disks and disks for logging/archive log
- Define sufficient I/O bandwidth for database disks
  - For SCSI discs, define multipathing and failover
  - For ECKD disks, use HyperPAV
- Use storage data striping
  - ASM is Oracle's proffered methodology for striping database files across as many disks
  - XIV disk storage system has its own internal striping
  - Tiered Storage solution helps in fast I/O
  - Flashsystem boosts I/O performance by factors

Oracle and Linux on System z – IBM & Oracle working together

- Oracle database 12cR12 available on Linux on System z (since 1Q2014)
- Database Performance at developerworks
  http://www.ibm.com/developerworks/linux/linux390/perf/tuning_database.html





http://www.redbooks.ibm.com/redpieces/abstracts/sg248159.html?Open

Some Reference Customers
Linux on System z with Oracle

Bank of New Zealand

Questions?

Wilhelm Mild

IBM IT Architect

IBM Deutschland Research & Development GmbH
Schönaicher Strasse 220
71032 Böblingen, Germany

Office: +49 (0)7031-16-3796
mildw@de.ibm.com

THE Open GROUP Master Certified IT Architect

Thanks to Dr. Juergen Doelle for his contribution to this content.

The IBM Oracle Centers

MISSION

Help IBM customers to deliver integrated solutions with Oracle Software Products on IBM Infrastructures

STRENGH

Cross platform team with strong knowledge on Oracle products and a wide network within IBM and Oracle ecosystem

ACTIVITIES

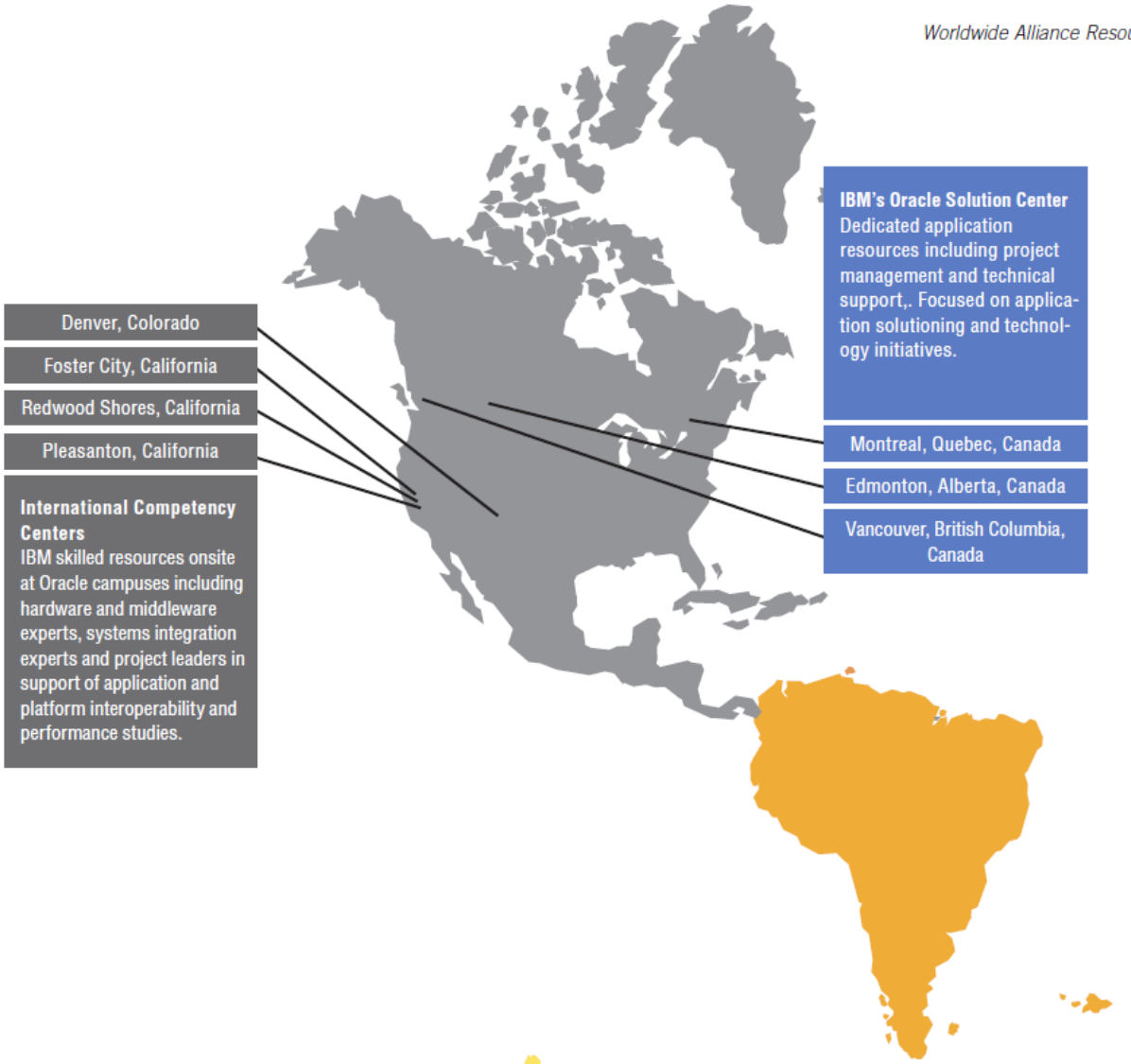
- Convince : Briefings & Conferences
- Build : Architecture, Design, Sizing
- Demonstrate : Proof-of-Concept, Benchmarks
- Deliver : Publications & Workshops

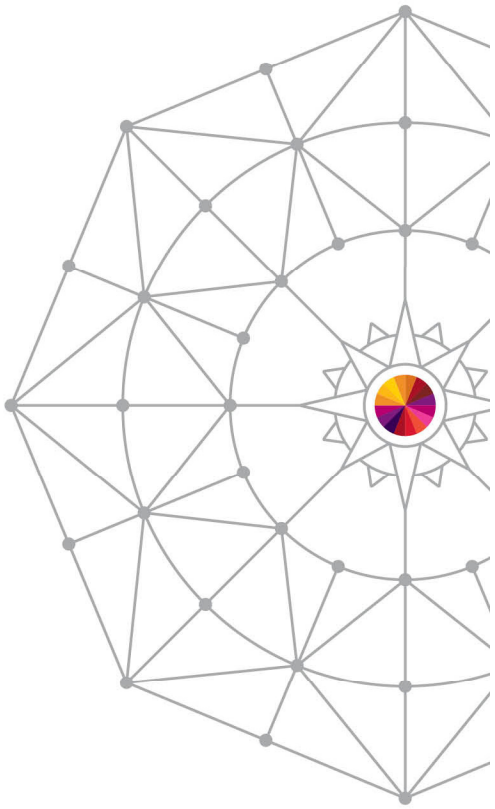
COVERED PRODUCTS

- IBM Platforms (System z,Power, System x, Total Storage)
- Oracle Technologies (Oracle DB, RAC, ASM, Dataguard)
- Oracle Applications (EBS, Siebel & OBI & OWI)
- Entry point to other on Industry Solutions (BRM, iFlex, RETEK, Weblogic…)

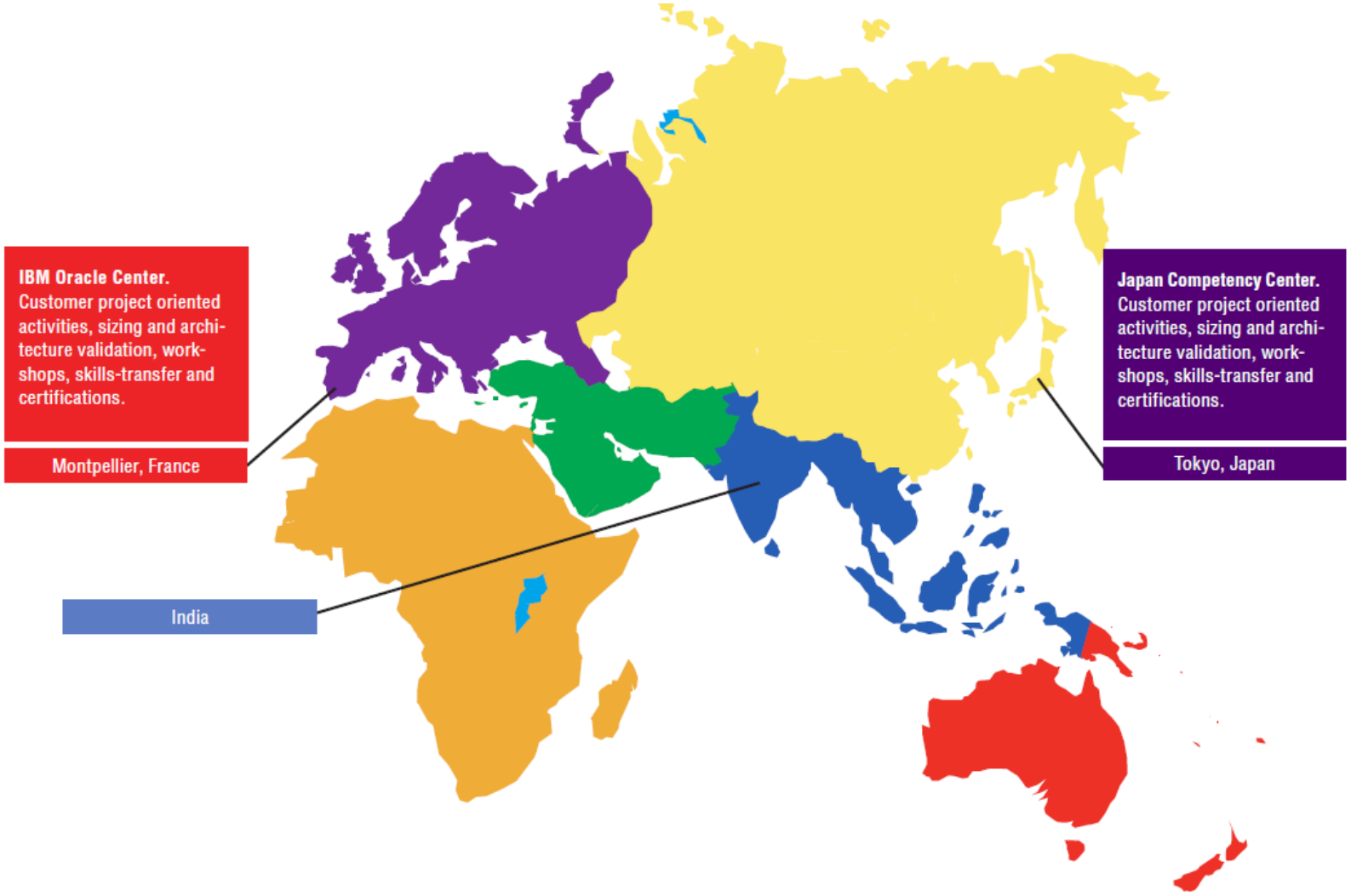
ibmoracl@us.ibm.com

IBM Oracle Competence Centers





IBM Oracle Competence Centers



Notices

This information was developed for products and services offered in the U.S.A.

Note to U.S. Government Users Restricted Rights — Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to: IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This 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.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of 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.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

Trademarks

- This presentation contains trade-marked IBM products and technologies. Refer to the following Web site:

  http://www.ibm.com/legal/copytrade.shtml