

Herrmann & Lenz



Services GmbH

ORACLE AUF MS WINDOWS –
32 ODER 64 BIT? EIN ÜBERBLICK

ORACLE AUF WINDOWS

IST WINDOWS ÜBERHAUPT DIE RICHTIGE PLATTFORM FÜR ORACLE-DATENBANKEN?

- ⊙ „Windows ist doch gar kein richtiges Server-Betriebssystem!“
 - ⊙ Die Welt hat sich weiter gedreht – wir schreiben das Jahr 2009. NT 4.0 ist Geschichte...
 - ⊙ Spätestens der Windows Server 2003 und erst recht der Server 2008 sind mindestens ebenso als Plattform für produktive, große und ressourcenintensive Oracle-Systeme geeignet wie beliebige andere Plattformen!

WINDOWS 32 ODER 64 BIT?

ENTSCHEIDUNGSHILFEN

- ⊙ Hier müssen wir uns einige Fragen stellen:
 - ⊙ Ist 64bit-fähige Hardware vorhanden?
 - ⊙ Ist unsere Anwendung 64bit-fähig?
 - ⊙ Wie viel Hauptspeicher soll / kann eingesetzt werden?
 - ⊙ Welches Release des Windows Servers soll eingesetzt werden?

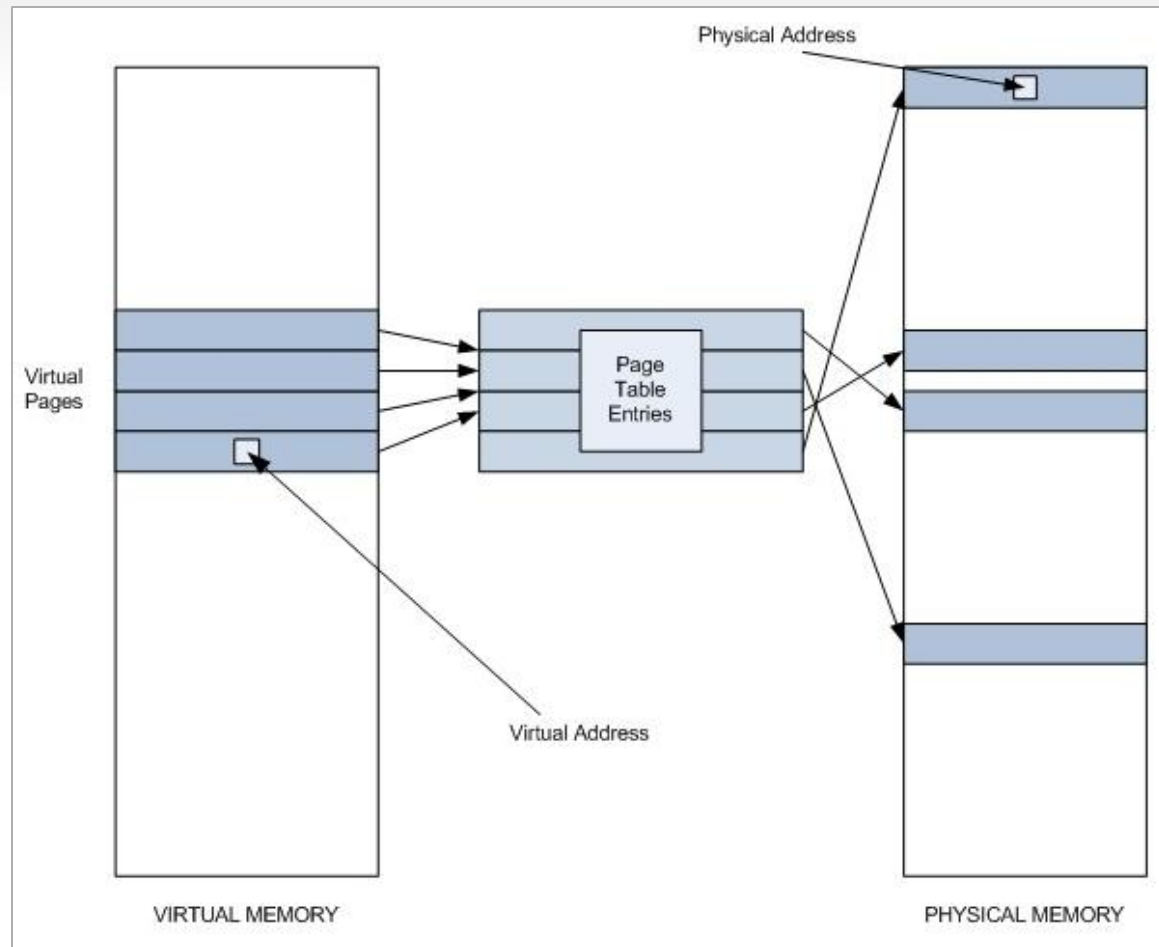
HARDWARE (x32 vs. x64)

- ⊙ 32bit-Windows läuft auf allen (auch älteren) handelsüblichen Maschinen mit intel-basierten CPUs
- ⊙ 64bit-Windows (wie auch 64bit-Anwendungen) erfordert geeignete Prozessoren
 - ⊙ Intel / HP Itanium
 - ⊙ Intel bzw. AMD x64
- ⊙ Moderne Prozessoren unterstützen sowohl 32- als auch 64bit-Code (Itanium nur 64bit)

- ⊙ 32bit-Systeme
 - ⊙ Adressregister
 - 32bit-Prozessoren können maximal 4GB RAM bzw. Adressraum direkt verwalten ($4.294.967.296 = 2^{32} =$ die größte mit 32bit adressierbare Zahl).
 - Für die Verwaltung von mehr RAM sind bestimmte Techniken erforderlich (PAE).
 - Von diesen 4 GB werden auf vielen Windows-Systemen lediglich ~3,2 GB im Task Manager angezeigt – die letzten 800MB werden oft durch sog. „memory-mapped Devices“ wie z.B. PCI-Karten besetzt. Um diese „hinter“ die 4GB-Grenze zu verschieben, werden passende Treiber sowie ein entsprechendes BIOS benötigt.

- ⊙ 32bit-Systeme
 - ⊙ Virtueller Adressraum
 - 2 GB RAM sind für den Windows-Kernel reserviert; jeder Prozess im User-Mode kann also maximal 2 GB privaten virtuellen Speicherplatz beanspruchen.
 - Sog. Pages werden in Stücken von 4K (mit Large Memory Page Support 2 MB oder mehr, s. Metalink Doc ID 422844.1) aus dem virtuellen in den physikalischen (oder ausgelagerten) Speicher gemappt – für Read-Only-Zugriffe kann die selbe Page aus dem physikalischen Speicher in den virtuellen Adressraum vieler verschiedener Prozesse gemappt werden.
 - Will ein Prozess die betreffende Page ändern, so erhält er eine beschreibbare Kopie, während alle anderen Prozesse das Read-Only-Original behalten.

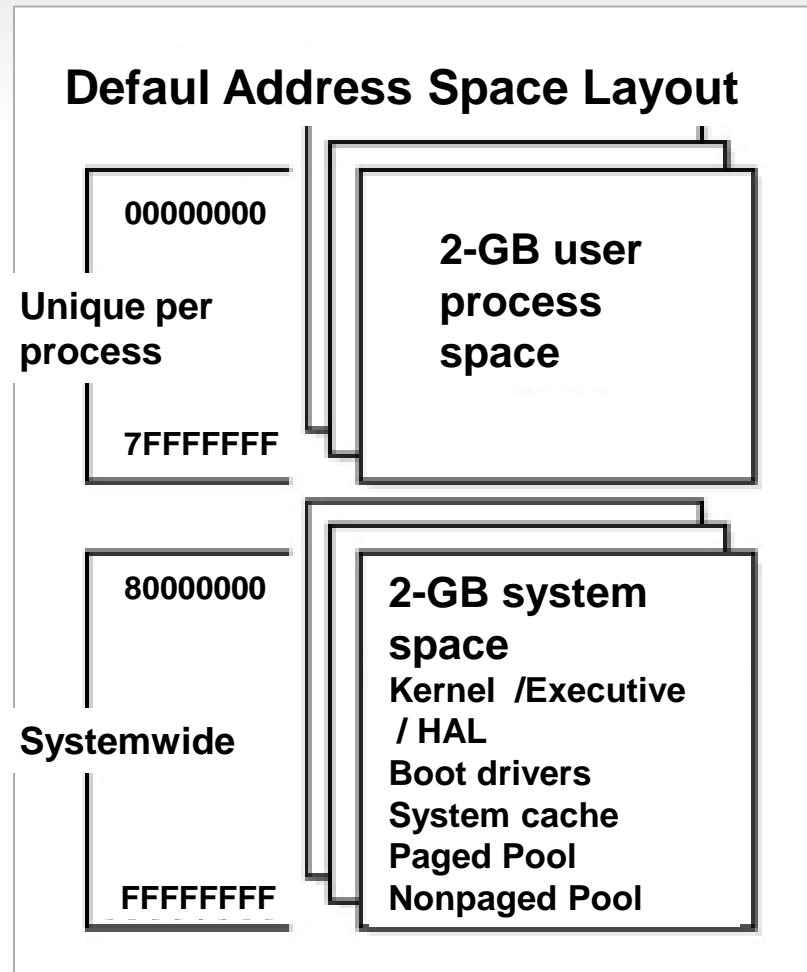
PAGE MAPPING



ARCHITEKTUR - UNTERSCHIEDE (3)

- ⊙ 32bit-Systeme
 - ⊙ Kernel-Memory
 - Nur Code im Kernel-Mode kann zugreifen.
 - Besteht aus verschiedenen Bereichen: Paged Pool (kann ausgelagert werden), Non-Paged Pool (kann nicht ausgelagert werden) und System PTEs (Page Table Entries).
 - Nur max. 2 GB für die gesamte Prozessverwaltung – egal wie viele User-Prozesse gleichzeitig aktiv sind.
 - Evtl. müssen Hunderte oder auch Tausende Prozesse, geöffnete Dateien und Datei-Handler verwaltet werden – bei einem Terminal-Server ggfs. ein Vielfaches davon.
 - Also: VORSICHT bei „Optimierung“ durch den /3GB-Schalter (insbesondere bei Terminal-Servern)!

KERNEL MEMORY



- ⊙ 64bit-Systeme
 - ⊙ Bei 64bit-Prozessoren sind die Adressregister auf 64 bit erweitert; tatsächlich werden von heutigen Prozessoren jedoch „nur“ bis zu 48bit genutzt (entspricht einer möglichen Adressierung von maximal 256 TB RAM). MS limitiert die Nutzung auf 44bit ~ 16 TB.
 - ⊙ Hiervon würden dann max. 8 TB auf den Windows-Kernel (Kernel Mode Memory) und max. 8 TB auf einen beliebigen User-Prozess (User Mode Memory) entfallen.
 - ⊙ Die Maximalgröße für jeden der Bereiche Paged Pool, Non Paged Pool und System PTE beträgt auf jedem x86-64-System 128 GB.
 - ⊙ Installation von Oracle Server-Software nur 64bit-Versionen auf 64bit Betriebssystem! Clients evtl. abweichend auch 32bit-Versionen.

SPEICHERUNTERSTÜTZUNG UNTER MS WINDOWS

Betriebssystem	32bit-Limit *	64bit-Limit
Win2003 SP2 Datacenter	128 GB (64 GB mit 4GT)	2 TB
Win2003 Storage Server EE	8 GB	Nicht verfügbar
Win2003 R2 Datacenter, Win2003 SP1 Datacenter	128 GB (16 GB mit 4GT)	1 TB
Win2003 SP2 EE	64 GB	2 TB
Win2008 SE, Webserver, Small Business	4 GB	32 GB
Win2008 Datacenter, EE, Itanium	64 GB (nicht für Itanium)	2 TB
Win2008 R2 EE, Datacenter, Itanium	Nicht verfügbar *	2 TB
Win2008 R2 SE, Web Server	Nicht verfügbar *	32 GB

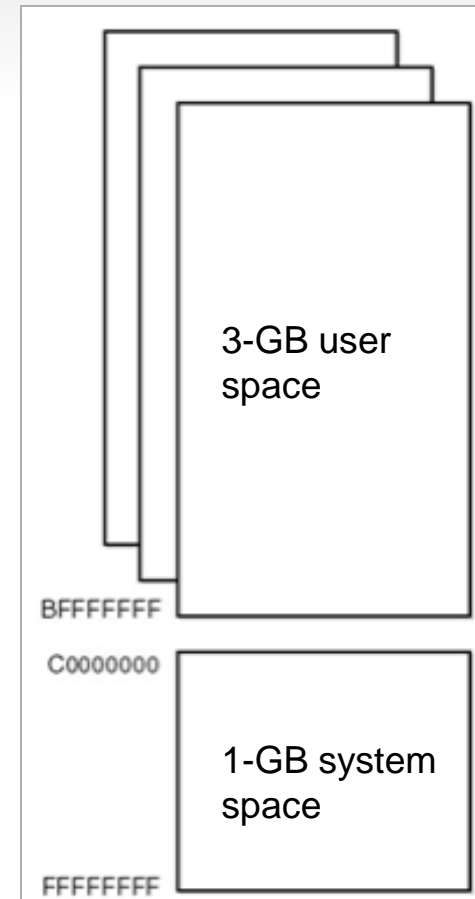
*** Win2008 R2 ist nur noch in der 64bit-Ausgabe verfügbar**

*** >4 GB bei 32bit-Systemen erfordern PAE**

OPTIMIERUNGSMÖGLICHKEITEN

4 GIGABYTE-TUNING (4GT)

- ⊙ Steht ab Windows 2000 Advanced Server und Data Center zur Verfügung
- ⊙ Wird durch den /3GB-Schalter in der boot.ini aktiviert
- ⊙ Verschiebt die Grenze zwischen User- und Kernel-Memory **innerhalb** des maximalen 4GB-Speicherplatzes



PAE / AWE

- ⊙ PAE (Physical Address Extension) = zusätzliche Fähigkeit des IA32-Prozessors, mehr als 4 GB Speicher zu adressieren. Auf 64bit-Systemen wird PAE nicht unterstützt.
- ⊙ AWE (Address Windowing Extension) = Sammlung von APIs zu den Funktionen des Windows-Speicher-Managers, mit denen Programme mehr als 4 GB Speicher adressieren können → Feature des Windows-Servers.
- ⊙ PAE bzw. AWE sind hardwareseitig bei Prozessoren ab Intel Pentium Pro oder vergleichbaren AMD-Prozessoren verfügbar.
- ⊙ Wird in der boot.ini durch den /PAE-Schalter aktiviert (automatisch aktiviert bei Server 2003 mit Hot Swappable Devices).
- ⊙ Ist allein oder zusammen mit dem /3GB-Schalter nutzbar.
- ⊙ Speicherbereiche > 16 GB können NUR mit dem /PAE, aber OHNE /3GB genutzt werden!

PAE UND ORACLE (1)

- ⊙ Oracle unterstützt PAE in folgenden Versionen:
 - ⊙ Oracle 8.1.6.x
 - ⊙ Oracle 8.1.7.x
 - ⊙ Oracle 9.2.x
 - ⊙ Oracle 10.1.x
 - ⊙ Oracle 10.2.x
 - ⊙ Oracle 11.1.x
 - Achtung: KEINE Unterstützung in 9.0.1.x!

PAE UND ORACLE (2)

- ⊙ Was bewirkt der Einsatz von PAE/AWE in der Datenbank?
 - ⊙ Verfügbarkeit von Speicherbereichen jenseits der 4 GB-Grenze für den **Buffer Cache**
 - ⊙ Zugriff auf Blöcke im hohen Speicherbereich durch Mapping aus einem „Fenster“ innerhalb der regulären „unteren“ 2 bzw. 3 GB – evtl. Mapping und Unmapping
 - ⊙ Größe des Fensters definiert durch Registry-Key `AWE_WINDOW_MEMORY` in `HKLM\Software\Oracle\[Key]Homexxx` – Default ist 1 GB
 - ⊙ `AWE_WINDOW_MEMORY` gilt für ALLE DB-Instanzen!

PAE – VORAUSSETZUNGEN FÜR DIE NUTZUNG IN DER ORACLE DATENBANK

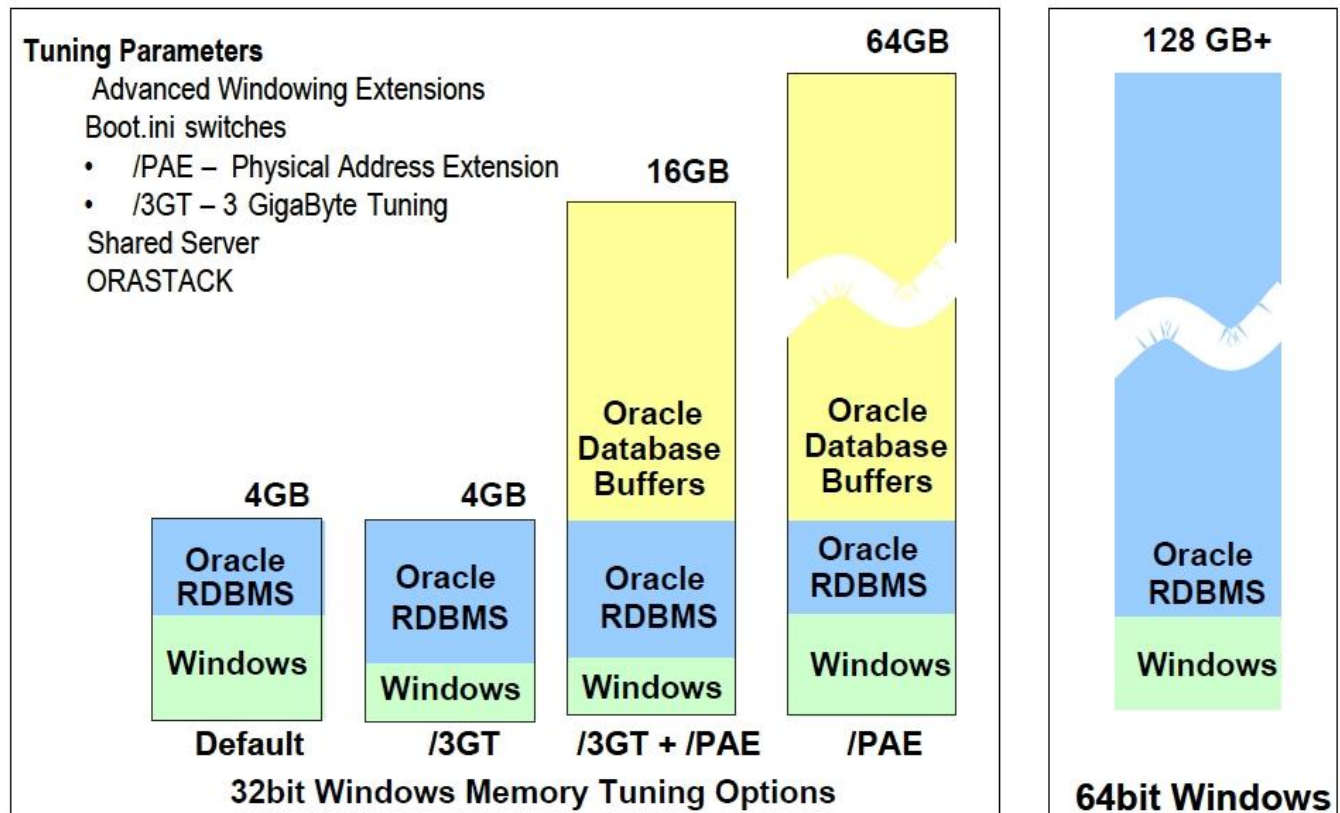
- ◉ OS-Account des Oracle-Prozesses muss das Recht „Lock Pages in Memory“ besitzen (Default für den LOCAL SYSTEM-Account)
- ◉ AWE_WINDOW_MEMORY-Key in der Registry gesetzt
- ◉ Init.ora-Parameter
USE_INDIRECT_DATA_BUFFERS=TRUE
- ◉ Nutzung des init.ora-Parameters
DB_BLOCK_BUFFERS → **nicht** DB_CACHE_SIZE;
keine abweichenden Blockgrößen!
- ◉ Rest der SGA (Shared Pool, Large Pool etc.), PGA, Overhead für oracle.exe und dlls, Stack Space für alle Threads **müssen** in den unteren 2-3GB liegen!

PAE - RECHENBEISPIEL

- ⊙ Der maximal mögliche Speicher für den Buffer Cache berechnet sich aus
 - ⊙ Physikalischer Speicher > 4 GB plus
 - ⊙ AWE_WINDOW_MEMORY
- ⊙ Ein Datenbank-Server mit 12 GB RAM und der Defaultgröße 1 GB für AWE_WINDOW_MEMORY kann also theoretisch einen Buffer Cache von 9 GB nutzen:
$$\text{RAM} - 4\text{GB} + \text{AWE_WINDOW_MEMORY} = 12 - 4 + 1 = 9$$

(reell jedoch etwas weniger → Overhead)
- ⊙ Bei 8k Blockgröße also DB_BLOCK_BUFFERS=1179648 für 9 GB Buffer Cache.

ÜBERSICHT TUNING-PARAMETER WINDOWS 32 UND 64 BIT



FAZIT: 64BIT RULES!

- ⊙ Es gibt diverse Möglichkeiten, ein 32bit-System zur Nutzung von mehr als 4 GB RAM zu „überreden“
- ⊙ Es gibt jedoch auch oftmals teilweise massive Probleme (Prozessabstürze, Verlangsamung des Systems bis zum vollständigen „Einfrieren“ etc.)
- ⊙ Die deutlich besseren Erfahrungen haben wir mit 64bit-Systemen gemacht!

DOKUMENTATION

- ⊙ Interessante Artikel zum Thema:
 - ⊙ MS Knowledge Base Doc ID 283037, 297812 u.a.
 - ⊙ Blog des MS Performance Teams ASKPERF: <http://blogs.technet.com/askperf>
 - ⊙ Oracle Database Platform Guide for Windows, hier: Oracle Database Scalability on Windows u.a.
 - ⊙ Oracle Metalink Doc ID 115349.1, 422844.1 u.a.

FRAGEN ?

Herrmann & Lenz 

Services GmbH

susanne.jahr@hl-services.de

<http://www.hl-services.de>