

32位Linux设置超大OracleSGA的分析 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/645/2021_2022_32_E4_BD_8DLinux_c103_645615.htm 有不少用户认为在32位Linux，只能设置约1.7GB的Oracle SGA。也有不少用户在不同的Linux发行版中使用相同的配置过程，获得了不同的SGA最大值，便认为某些Linux发行版存在问题。这些想法都比较片面。实际上32位Linux上Oracle SGA 的容量取决于三个指标，即：Linux kernel 版本、Oracle Database 版本、Linux 内核参数shmmax。这里跟据一些网上文章做了些查正，做一些分析吧。一、内核版本的影响 在32位Linux平台，至少都有两套内核供用户使用。一个是smp核心，一个是hugemem核心。两个核心的区别在于直接映射的内核数据代码地址空间的区别：SMP 核心：在x86架构下，虚拟地址空间的大小为4G。在这4G空间中，用户空间占3G (0x00000000到0xbfffffff)，核心空间占1G(0xc0000000到0xffffffff)。这样的分配策略称为3G/1G分配。具体的分配方式如下：1. 0GB-1GB User space - Used for text/code and brk/sbrk allocations (malloc uses brk for small chunks) 2. 1GB-3GB User space - Used for shared libraries, shared memory, and stack. shared memory and malloc use mmap (malloc uses mmap for large chunks) 3. 3GB-4GB Kernel Space - Used for the kernel itself 这种分配方式对于拥有1G物理内存以下的系统是没有任何问题的，即使超过1G物理内存，3G/1G分配策略也没有什么问题，因为内核可以在高端内存区域(物理地址1G以上的内存)中存放一些内核数据结构(比如页缓冲等)。然而，随着物理内存的增多，3G/1G分配策略的问题也逐渐

会暴露出来。这是因为一些关键的内核数据结构(比如用于管理物理内存的mem_map[])是存放在1G核心空间之内的。对于32G内存的系统, mem_map[]会占用近0.5G的低端内存(物理地址896M以下的内存), 这样留给核心其他部分的内存就不到所有内存的1.5%。而对于64G内存的系统, mem_map[]本身就会耗尽所有的低端内存, 造成系统无法启动。但是把mem_map[]放到高端内存的做法也不太实际, 因为mem_map[]和内存管理, 体系结构相关底层实现, 文件系统以及驱动等几乎所有的核心的关键部分均有联系, 这时候就需要使用hugemem核心了。 hugemem核心: 与SMP的3G/1G策略不同, hugemem使用4G/4G分配方式。可以使核心空间由1G增加到4G, 而用户空间也由3G增加到4G。相比3G/1G分配策略, 对于4G物理内存系统, 使用4G/4G分配可以增加低端内存达3倍以上, 而对于32G物理内存系统, 则会有更多的提升, 达到原来的6倍。理论上, 4G/4G策略可以支持物理内存达200G的x86系统(如果硬件没有限制的话), 即使对于这样的系统, 4G/4G策略也能保证留有1G可用的低端内存。不论能否理解上面的解释, 只须要记住 smp 和 hugemem 两个核心, 一个是 3G/1G策略, 一个是4G/4G策略即可。

二、Oracle Database 版本: Oracle SGA 是挂载在内存用户空间中, 不同版本的 Oracle Database, 挂载SGA起始地址是不同的: Oracle 10g Release 1:挂载SGA的起始地址为0 × 50000000(1.25GB) Oracle 10g Release 2:挂载SGA的起始地址为0 × 20000000(0.5GB) (其它版本未查正, 有兴趣可以自己看看Oracle手册)从这个结果可以看出, 理论上若使用3G/1G的smp核心, Oracle 10g Release 1的SGA可设置到 3GB - 1.25 =

1.75GB。正是因此，也有人认为Oracle SGA 只能设置到1.75GB。而升级至 R2 版本，则可以设置到 $3\text{GB} - 0.5\text{GB} = 2.5\text{GB}$ 。相同的，使用 4G/4G 的hugemem 核心能够获得多大的SGA 就很清楚了。三、kernel shmmax 参数：shmmax定义单个共享内存段的最大值，它的取值范围区间是 $[0, 4294967295]$ ，单位为byte，4294967295 bytes即4294967296 bytes(4GB)减去1。一般来说，它应该足够大以容下整个SGA，避免SGA使用多个共享内存段造成Oracle性能下降。那么，将shmmax设置为最大值4294967295，使用 hugmemem 核心，Oracle 10g R2 版本，则理论上的SGA最大值为 3.5GB。尽管用户进程可用的虚拟地址空间为4GB以及shmmax的最大值为4294967295，仍然可以通过使用内存文件系统(in-memory filesystem，比如tmpfs、ramfs以及hugetlbfs)打开Oracle的Very Large Memory (VLM)特性来扩展SGA超过4GB，比如6GB。但是这种方法有个不方便的地方是，用户不能够再使用Oracle 10g中的Automatic Shared Memory Management了。总结一下：
SMP核心和hugemem核心的影响：
SMP核心：1GB 3GB，SGA最大值为1.75GB(3GB-1.25GB)
Hugemem核心：4GB 4GB，SGA最大值为2.75GB(4GB-1.25GB)
Oracle 不同版本的影响：
Oracle 10g Release 1:挂载SGA的起始地址为 0×50000000 (1.25GB)
Oracle 10g Release 2:挂载SGA的起始地址为 0×20000000 (0.5GB)
那么：Hugemem 核心 Oracle 10g Release 2，SGA最大值为3.5GB(4GB-0.5GB)
BTW：由于64位系统对应16EB寻址范围，而不是32位系统的4GB，所以想要获得更大SGA，性能更好的效果，应该优先使用 64 位系统，而不是通过32系统配合内存文件系统来配置了。更多优质资料尽在

百考试题论坛 百考试题在线题库 #0000ff>linux认证更多详细
资料 100Test 下载频道开通，各类考试题目直接下载。详细请
访问 www.100test.com