

## 保持 RAC 群集环境稳定当前必须要做的 11 件事 (Doc ID 1525819.1)

### 文档内容

[用途](#)

[适用范围](#)

[详细信息](#)

- [1. 在您的环境中应用最新的 Patchset Update \(PSU\)](#)
- [2. 确保 UDP 缓冲区大小合适](#)
- [3. 在所有版本 10.2 和 11.1 集群上将 DIAGWAIT 的值设置为 13](#)
- [4. 在 Linux 环境中实施 HugePage](#)
- [5. 实施 OS Watcher 和 \(或\) Cluster Health Monitor](#)
- [6. 按照最佳实践配置 OS 设置](#)
- [7. 确保在 AIX 平台上应用合适的 APARS 以避免出现过量分页/交换问题](#)
- [8. 应用 NUMA 补丁](#)
- [9. 增加 Windows 非交互式 Desktop Heap](#)
- [10. 运行 RACcheck 实用程序](#)
- [11. 使用 slewing 选项实施 NTP](#)

[参考](#)

### 适用于:

Oracle Cloud Infrastructure - Database Service - 版本 N/A 和更高版本  
Oracle Database Backup Service - 版本 N/A 和更高版本  
Oracle Database Cloud Exadata Service - 版本 N/A 和更高版本  
Oracle Database Cloud Service - 版本 N/A 和更高版本  
Oracle Database - Enterprise Edition - 版本 10.2.0.1 到 11.2.0.2 [发行版 10.2 到 11.2]  
本文档所含信息适用于所有平台

### 用途

许多 RAC 不稳定的问题皆可归因于没有实施一个其实很短的最佳实践和 (或) 配置列表。本文档旨在提供一种简单的方法, 找出这些常被遗漏的最佳实践和 (或) 配置问题的列表, 以期可以防止这些问题引起的系统不稳定。

### 适用范围

本文章适用于所有 RAC 实施过程。

### 详细信息

#### 1. 在您的环境中应用最新的 Patchset Update (PSU)

**适用平台:** 所有平台

**原因:** 10.2.0.4 及更高版本中引入了 Patchset Updates (也称为 PSU), 主要是为了改进 CPU 修补策略。PSU 按季度推出, 其中包括最新的 CPU, 另外它们也包含其它对保持您的环境稳定性非常重要的修正。如果要进行全新安装, 应始终应用最新的 PSU 作为您的基线。对于现有安装, 一个必须实施的策略就是定期和持续的应用最新的 PSU。许多提交给 Oracle Support 并被确认为属于 bug 的问题都是已知 bug, 其中许多 bug 已在最新的 PSU 得到修正。请注意, 在 Windows 上, 会更经常推出累积型的补丁包, 但是在季度 PSU 版本期间发行的 Windows bundle patch 中包含了最新的 PSU 修正程序。

**更多信息:** 有关 PSU 的更多信息, 请参阅以下文档:

[Document 854428.1](#) Intro to Patch Set Updates (PSU)

[Document 1082394.1](#) 11.2.0.X Grid Infrastructure PSU Known Issues

[Document 756671.1](#) Oracle Recommended Patches -- Oracle Database

[Document 161549.1](#) Oracle Database, Networking and Grid Agent Patches for Microsoft Platforms

#### 2. 确保 UDP 缓冲区大小合适

**适用平台:** Windows 除外的所有平台

**原因:** 私网可以说是 RAC 数据库的命脉。但是, 如果未向 UDP 分配合适的缓冲空间用以发送和接收缓冲, 则私网的性能将大幅降低。这将会导致您的集群出现稳定性问题。

**更多信息:** 有关正确调整 UDP 缓冲区的更多信息, 请参考以下文档:

[Document 181489.1](#) Tuning Inter-Instance Performance in RAC and OPS

[Document 563566.1](#) gc lost blocks diagnostics

**注意：** Windows 集群对 Cache fusion 通信使用 TCP，因此，UDP 缓冲区设置不适用于 Windows。

### 3. 在所有版本 10.2 和 11.1 集群上将 DIAGWAIT 的值设置为 13

**适用平台：** Windows 以外的所有平台

**原因：** 在 10gR2 (10.2.x) 和 11gR1 (11.1.x) 中，OPROCD 守护进程的默认容差仅设置为 500 毫秒 (0.5 秒)。对于非常繁忙的系统，此容差可能过小，因此负载繁重的系统可能会出现错误重启的情况。将 diagwait 设置更改为 13 后，OPROCD 的容差变成 10,000 毫秒 (10 秒)，为繁忙的系统提供了更长容差，可避免出现错误重启的情况。另外，如果出现节点重启的情况，设置diagwait 能够提供更多的时间将诊断信息刷新到跟踪文件中，以供进一步诊断使用。此更改不能包含在补丁集中，因为必须关闭整个集群才能实施。但是，我们仍然强烈建议在所有 10gR2 和 11gR1 群集上将此值更改为 13。对于新实施的集群，应在安装后立即进行此更改。对于现有的安装，应当安排停机时间，以尽快进行此项更改。可通过以下命令确认当前设置：

```
# $CLUSTERWARE_HOME\bin\crsctl get css diagwait
```

**注意：** 此设置不适用于 Windows 环境，也不适用于 11gR2 版本 (11.2.0.1 和更高版本)。对于 10gR2 来说，该设置需要 patchset 10.2.0.4 (或更高版本)。

**更多信息：** 有关 DIAGWAIT 的更多信息，请参考以下文档中的内容：

[Document 559365.1](#) Using Diagwait as a diagnostic to get more information for diagnosing Oracle Clusterware Node evictions  
[Document 567730.1](#) Changes in Oracle Clusterware on Linux with the 10.2.0.4 Patchset

### 4. 在 Linux 环境中实施 HugePage

**适用平台：** 所有 LINUX 64 位平台

**原因：** 在 Linux 环境中实施 HugePage 能够极大地提高内核性能。对于内存较大的系统，效果尤其明显。一般而言，所有 RAM 大于 12GB 的系统都适合使用 Hugepage。系统中的 RAM 越大，系统启用 Hugepage 后获得的好处也越大。这是因为内核为映射和维护内存页表所做的工作量会随着系统内存的增大而增加。启用 Hugepage 能够显著地降低内核要管理的页面数，而且能提高系统的效率。经验表明，如果未启用 Hugepage，内核挤占关键的 Oracle Clusterware 或 Real Application Clusters 守护进程的情况会很常见，而这会导致实例或节点驱逐出现。

**注意：** 在 Linux 平台上，11g Automatic Memory Management (AMM) 与 HugePage 不兼容。最佳实践是禁用 AMM，以支持 HugePage。有关 Linux 上的 AMM 和 HugePage 的更多信息，请参阅[Document 749851.1](#)

**更多信息：**

[Document 361323.1](#) HugePages on Linux: What It Is... and What It Is Not...  
[Document 401749.1](#) Shell Script to Calculate Values Recommended Linux HugePages / HugeTLB Configuration

### 5. 实施 OS Watcher 和 (或) Cluster Health Monitor

**适用平台：** 所有平台

**原因：** 虽然 OS Watcher 和 Cluster Health Monitor 与稳定性并不直接相关，但是，对于确定 OS 状态和分析导致节点或实例驱逐的许多问题的潜在根本原因方面，它们却是非常好用的工具。如果在第一次发生某个问题后就有合适的数据可用于诊断这个问题，则可缩短确定根本原因的时间，而且能防止以后出现停机。大部分类似的第三方数据收集工具的收集间隔时间都比较长 (如 5 分钟或更长)，而且 (或者) 它们很难被解释，或收集到的数据不正确。OS Watcher 是一款非常简单的小型工具，每 30 秒钟 (默认) 收集一次 OS 的基本信息。Cluster Health Monitor 虽然不适用于所有平台，但它能够更精细地实时收集数据，可以补充 OS Watcher 的不足之处。应当在所有集群节点上全天候运行这两个实用程序或其中一个，有助于更快地诊断和调试问题。

**更多信息：**

[Document 301137.1](#) OS Watcher User Guide  
[Document 1328466.1](#) Cluster Health Monitor (CHM) FAQ  
[Document 580513.1](#) How To Start OSWatcher Black Box Every System Boot (Linux specific)

-

### 6. 按照最佳实践配置 OS 设置

(请参阅 Oracle / IBM 针对系统稳定性进行内存优化而联合编写的白皮书)

**适用平台：** 所有 AIX 版本

**原因：** [Oracle Real Application Clusters on IBM AIX Best practices in memory tuning and configuring for system stability](#) (Oracle Real Application Clusters on IBM AIX 针对系统稳定性进行内存优化和配置的最佳实践) 白皮书是这两大供应商根据共同的经验进行联合测试与合并最佳实践后的精华。经验表明，如果遵照本白皮书的建议，可以解决 RAC/AIX 集群中的大多数稳定性问题。AIX 版本 6.1 已将其中的多数建议作为默认值包含在内，但仍应在所有 AIX RAC 群集上确认这些设置，不论哪个 OS 或 Oracle 版本。

**更多信息：**

白皮书下载地址：<http://www.oracle.com/technetwork/database/clusterware/overview/rac-aiX-system-stability-131022.pdf>  
[Document 811293.1](#) RAC Assurance Support Team: RAC Starter Kit and Best Practices (AIX)

## 7. 确保在 AIX 平台上应用合适的 APARS 以避免出现过量分页/交换问题

**适用平台:** 所有 AIX 版本

**原因:** 经验表明, 这是很常见的会影响 AIX 环境的问题。鉴于此问题的性质, 任何易受此问题影响的用户应该都有过系统完全挂起的经历。在非 RAC 环境中, 此问题会导致系统挂起, 直到进行手动干预为止。而在 RAC 环境中, 此问题会由于节点无法响应导致出现节点驱逐的情况。

**更多信息:** 有关此问题的更多信息, 请参考 [Document 1088076.1](#) Paging Space Growth May Occur Unexpectedly on AIX Systems With 64K (medium) Pages Enabled

**注意:** 该文章中列出的 APAR 版本和编号特定于给定的 Technology Level (TL)。您需要应用的实际 APAR 或修正程序编号将取决于您使用的 AIX (Technology Level, TL)。请与 IBM 联系, 确认是否已进行此项修正, 如果尚未进行, 确认需要哪种 TL 或 APAR 才能进行此项特定的修正。

## 8. 应用 NUMA 补丁

**适用平台:** 所有平台

**原因:** 从 10.2.0.4 和 11.1.0.7 RDBMS 补丁集开始, 可在支持 NUMA 的平台 (取决于 OS 和硬件) 上进行 NUMA 优化。在 (支持 NUMA 的系统中的) RDBMS 代码中应用 NUMA 会触发一些导致数据库性能降低和不稳定的 bug。与 10.2.0.4 和 11.1.0.7 中的 NUMA 优化相关的症状/问题相关的完整列表, 请见 [Document 759565.1](#)。如果要运行 10.2.0.4 或 11.1.0.7 补丁集, Oracle 强烈建议将 [Patch 8199533](#) 应用于系统, 以预先解决这些 NUMA 相关的问题。

## 9. 增加 Windows 非交互式 Desktop Heap

**适用平台:** Windows 平台

**原因:** 现已发现, Windows 集群上的非交互式 Desktop Heap 的默认大小不够。这会导致出现应用程序连接问题和集群总体不稳定 (挂起和/或崩溃) 的问题。要有效地解决此问题, 建议将非交互式 Desktop Heap 增加到 1MB。如果没有 Microsoft 参与, 不应超过建议的 1MB。

**更多信息:** 关于如何对非交互式 Desktop Heap 进行上述调整的说明, 请见 [Document 744125.1](#)。

## 10. 运行 RACcheck 实用程序

**适用平台:** Linux (x86 和 x86\_64)、Solaris SPARC 和 AIX (使用 bash shell)

**原因:** RACcheck 是一款 RAC 配置审核工具, 主要用于审核 Real Application Clusters (RAC)、Oracle Clusterware (CRS)、Automatic Storage Management (ASM) 和 Grid Infrastructure (GI) 环境中各个重要的配置设置。此实用程序用于验证由 RAC Assurance 开发和支持团队维护的 RAC 和 Oracle Clusterware 最佳实践和初学者指南文章 (请参阅 [Document 810394.1](#)) 系列定义的最佳实践和成功因素。我们强烈建议在 RACcheck 支持的平台上运行 RAC 的客户使用此工具识别会影响集群稳定性的潜在配置问题。

**更多信息:** 有关 RACcheck 的更多信息和下载此实用程序的链接请参见 [Document 1268927.1](#)。

## 11. 使用 slewing 选项实施 NTP

**适用平台:** 所有 Linux 和 Unix 平台。

**原因:** 如果没有 slewing 选项, 在时间差异超过特定 (取决于平台) 阈值时, NTP 将向前或向后调整系统时钟。大幅度向后调整时间会导致 Clusterware 以为错过了签到, 从而发生节点驱逐的情况。出于此原因, 我们强烈建议将 NTP 配置调整为 slewing time (加快或减慢) 时钟时间以同步时间, 以防止此类驱逐情况的发生。有关如何在您的平台上实施 NTP 时间调整的更多信息, 请参考平台特定的 RAC 与 Oracle Clusterware 最佳实践和初学者指南文档 (见下文)。

**更多信息:**

[Document 811306.1](#) RAC and Oracle Clusterware Best Practices and Starter Kit (Linux)  
[Document 811280.1](#) RAC and Oracle Clusterware Best Practices and Starter Kit (Solaris)  
[Document 811271.1](#) RAC and Oracle Clusterware Best Practices and Starter Kit (Windows)  
[Document 811293.1](#) RAC and Oracle Clusterware Best Practices and Starter Kit (AIX)  
[Document 811303.1](#) RAC and Oracle Clusterware Best Practices and Starter Kit (HP-UX)

---

## 参考

[NOTE:401749.1](#) - Oracle Linux: Shell Script to Calculate Values Recommended Linux HugePages / HugeTLB Configuration

[NOTE:756671.1](#) - Primary Note for Database Proactive Patch Program

[NOTE:1054902.1](#) - How to Validate Network and Name Resolution Setup for the Clusterware and RAC

[NOTE:811271.1](#) - RAC and Oracle Clusterware Best Practices and Starter Kit (Windows)

[NOTE:810394.1](#) - RAC and Oracle Clusterware Best Practices and Starter Kit (Platform Independent)

[NOTE:181489.1](#) - Tuning Inter-Instance Performance in RAC and OPS

[NOTE:301137.1](#) - OSWatcher (Includes: [Video])

[NOTE:361323.1](#) - HugePages on Linux: What It Is... and What It Is Not...

[NOTE:1082394.1](#) - 11.2.0.1.X Grid Infrastructure PSU Known Issues

[NOTE:1088076.1](#) - AIX: Paging Space Growth May Occur Unexpectedly With 64K (medium) Pages Enabled  
[NOTE:759565.1](#) - Oracle NUMA Usage Recommendation  
[NOTE:559365.1](#) - Pre-11.2: Using Diagwait as a diagnostic to get more information for diagnosing Oracle Clusterware Node evictions  
[NOTE:563566.1](#) - Troubleshooting gc block lost and Poor Network Performance in a RAC Environment  
[NOTE:567730.1](#) - Changes in Oracle Clusterware on Linux with the 10.2.0.4 Patchset  
[NOTE:744125.1](#) - Windows: Connections Fail with ORA-12640 or ORA-21561  
[NOTE:811306.1](#) - RAC and Oracle Clusterware Best Practices and Starter Kit (Linux)  
[NOTE:749851.1](#) - HugePages and Oracle Database Automatic Memory Management (AMM) on Linux  
[NOTE:811280.1](#) - RAC and Oracle Clusterware Best Practices and Starter Kit (Solaris)  
[NOTE:1328466.1](#) - Cluster Health Monitor (CHM) FAQ  
[NOTE:161549.1](#) - Oracle Database, CRS, ASM, Networking and EM Agent Patches for Microsoft Platforms  
[NOTE:811293.1](#) - RAC and Oracle Clusterware Best Practices and Starter Kit (AIX)  
[NOTE:854428.1](#) - Patch Set Updates for Oracle Products  
[NOTE:1268927.1](#) - ORAchk - Health Checks for the Oracle Stack  
[NOTE:1427855.1](#) - AIX: Top Things to DO NOW to Stabilize 11gR2 GI/RAC Cluster  
[BUG:13623902](#) - NODE EVICTIONS ON RAC CLUSTER AFTER EXCESSIVE PAGING

Didn't find what you are looking for?