Eine Überwachung von Datenbank-Systemen ist wichtig, um deren Verfügbarkeit zu gewährleisten. Sicher kann man dazu alle möglichen Messwerte erfassen. Aber eine Überwachung soll möglichst minimalinvasiv, also mit kleinstmöglichem Aufwand eingreifend, erfolgen und damit das Datenbanksystem wenig zusätzlich belasten.

Minimalinvasive Überwachung von Datenbanken für optimale Verfügbarkeit

Ralf Appelbaum, TEAM GmbH

Der Artikel zeigt, welche Messwerte in welchen Konfigurationen minimal überwacht werden sollten und welche Ausfälle damit vermieden werden können. Der Autor berichtet aus seinen Erfahrungen auch, welche Ausfälle sich nicht gänzlich vermeiden lassen.

Über Werkzeuge zur Überwachung wurde schon viel geredet. Die Werkzeuge ermöglichen es, viele, teilweise vordefinierte Messwerte (Metriken) regelmäßig zu erfassen, mit Schwellwerten zu versehen und bei deren Überschreiten Alarme zu senden. Bei jedem dieser Werkzeuge stellt sich aber die Frage, was man möchte beziehungsweise was man mindestens überwachen muss, um die für das Unternehmen optimale Verfügbarkeit zu gewährleisten?

Im Rahmen der proaktiven Überwachung der „Oracle Administration Services" hat sich der Autor bei verschiedensten Kundenumgebungen Gedanken darüber gemacht, welche Überwachung unter dem Aufwand-/Nutzen-Aspekt sinnvoll ist. „Aufwand" meint hier sowohl Kosten als auch Systembelastung und „Nutzen" ist die sichergestellte Verfügbarkeit.

Im Rahmen der Hochverfügbarkeit trifft man auf die verschiedensten Konfigurationen und Komponenten von Datenbank-Systemen:

- *Betriebssysteme*
  Unix/Linux
  Microsoft Windows
- *Einfache Konfigurationen*
  Single Instanz
- *Hochverfügbare Konfigurationen*
  Oracle Real Application Clusters (RAC)
  Oracle Data Guard
  Oracle Fail Safe
- *Datenbankspeicher*
  Dateisystem
  Storage-Area-Netzwerke (SAN)
  Oracle Automatic Storage Management (ASM)
- *Netzwerk und Netzwerkkomponenten*
- *Oracle-Editionen*
  Enterprise Edition
  Standard Edition
  Standard Edition One
  Express Edition
- *Produkte*
  Oracle-Datenbank
  Oracle-Grid-Infrastruktur
  Oracle Grid Control
  Oracle Application Server

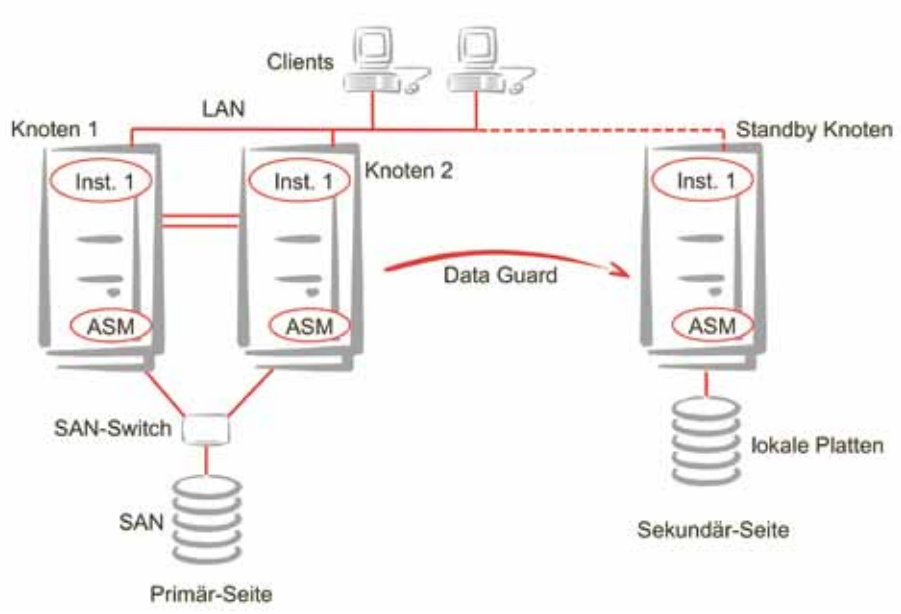

Abbildung 1: Vereinfachte Maximum Availability Architecture (MAA)

Als Beispiel kann man die komplexe Konfiguration aus Kombination von RAC und Data Guard betrachten, die den Hauptbestandteil der sogenannten „Maximum Availability Architecture" (MAA) bei Oracle bildet (siehe Abbildung 1).

In dieser Konfiguration trifft man viele der oben aufgeführten Aspekte wieder: Single Instance, RAC, Data - Guard, lokales Dateisystem, SAN, ASM, Netzwerk mit virtuellen IPs und bei installiertem Database Control auch einen Web-Server. Diese Konfiguration

lässt sich sowohl unter Unix/Linux realisieren, als auch unter Windows. In dieser Umgebung ist eine Überwachung auf jeden Fall erforderlich, um die mit der Konfiguration implizierte höhere Verfügbarkeit auch tatsächlich zu gewährleisten.

Ziel der Überwachung

Im Rahmen einer Überwachung werden in regelmäßigen Intervallen Prüfungen ausgeführt. Einige ermitteln, ob eine Komponente des Datenbank-Systems einwandfrei funktioniert oder ob Fehler gemeldet werden. In der Regel ist es bei einer Fehlermeldung aber schon zu spät (zum Beispiel, falls ein I/O nicht möglich ist, wenn das Dateisystem voll ist), weil durch den Ausfall einer einzelnen Komponente das gesamte Datenbank-System ausfällt.

Daher werden überwiegend Prüfungen durchgeführt, die einen numerischen Wert, einen Messwert, liefern (wie Füllgrad des Datei-Systems). Erreicht oder überschreitet der Wert an einem Messpunkt eine vorher festgelegte kritische Schwelle, dann droht die Komponente auszufallen. Erfährt man frühzeitig, dass ein Messwert seine kritische Schwelle überschritten hat, kann man dem Ausfall der Komponente und damit dem Ausfall des gesamten Datenbank-Systems entgegenwirken.

Ziel der Überwachung ist, den Betrieb eines Datenbanksystems ohne Unterbrechung sicherzustellen. Sollte sich ein Ausfall nicht vermeiden lassen, so ist das Ziel zumindest, das Datenbank-Systems schnellstmöglich wieder in Betrieb zu setzen. Dazu sind nur Prüfungen erforderlich, die Aussagen über eine Gefährdung des Betriebs des Datenbank-Systems oder dessen Ausfall machen, jedoch keine Prüfungen beziehungsweise Messungen, die ausschließlich zukünftiger Ressourcen-Planung oder dem Performance-Tuning dienen. Selbstverständlich können fast alle Messungen, wenn man die Werte im zeitlichen Verlauf betrachtet, auch zukünftige Ressourcen-Planungen und Performance-Tuning unterstützen.

Ebene	beispielhafter Messwert
Betriebssystem	Auslastung
Netzwerk	Erreichbarkeit
Speicherplatz	Verbrauch beziehungsweise freier Platz
Storage	Erreichbarkeit
Web-Server	URL erreichbar
Oracle-Datenbank	Tablespace-Auslastung
Oracle-DB-Instanz	Verfügbarkeit
Oracle Clusterware	Status der Ressourcen
Oracle ASM	Status Diskgruppen/Failuregruppen
Oracle-ASM-Instanz	Verfügbarkeit
Oracle Listener	Alle Instanzen/Services registriert
Grid Control	Agent Upload OK
Data Guard	Status der Standby-Datenbank

Tabelle 1: Übersicht möglicher Überwachungsebenen

Programmtyp	Programmbeispiele
Betriebssystembefehle	top ping du / df ls / dir swap –l
spezielle Programme/ Programmkomponenten	http Request (URL) prüfen
Oracle-Programme	lsnrctl srvctl crsctl asmcmd dgmgrl emctl dcmctl opmnctl
SQL/SQL*Plus Statements	select status, logins, blocked from v$instance; connect dbsnmp/<passwort>@<TNS-Alias>

Tabelle 2: Beispiele von Prüfprogrammen

Fragestellungen zur Überwachung

Wie bereits in der Einleitung erwähnt, stellen sich Fragen, wenn man eine neue Überwachung einrichtet. Nicht nur die: „Was will ich beziehungsweise was muss ich mindestens überwachen?“, sondern auch folgende:

- An welcher Stelle beziehungsweise auf welcher Ebene setze ich Prüfungen an?
- Wie beziehungsweise womit führe ich die Prüfungen durch?
- Was genau, also welche Messwerte, lassen sich prüfen beziehungsweise ermitteln?
- Wie häufig erfolgen die Prüfungen?
- Wo liegen die kritischen Schwellwerte für Messwerte?

Ebenen der Überwachung

Tabelle 1 listet mögliche Ebenen auf, an denen die Überwachung mit Prüfungen ansetzen kann. Optimal ist sicher eine Überwachung von der obersten bis zur untersten Ebene eines Anwendungs-Stacks, also von der

Prüf-/Messpunkt	Intervall	Bemerkung
CPU-Auslastung	5 Min.	
Anzahl Prozesse	5 Min.	
Hauptspeicher-Auslastung	5 Min.	
Swap-/Paging-Auslastung	5 Min.	
rsh-/ssh-Verbindung möglich	15 Min.	
Länge der Mailqueue	15 Min.	relevant, wenn auf dem Server Mails versendet werden
Anzahl Zombie-Prozesse	15 Min.	

Tabelle 3: Vorgeschlagene Prüf-Intervalle für das Betriebssystem

Prüf-/Messpunkt	Intervall	Bemerkung
Ping auf primäre Server IP	5 Min.	
Ping auf weitere Server IPs wie virtuelle IPs	5 Min.	bei Oracle Restart, RAC, Fail Safe und anderen Konfigurationen mit mehreren IPs je Server

Tabelle 4: Vorgeschlagene Prüf-Intervalle für das Netzwerk

Prüf-/Messpunkt	Intervall	Bemerkung
Platzverbrauch bzw. freier Platz auf Devices, Dateisystem, Datenträger	30 Min.	
Temp-Auslastung	30 Min.	
Größe ausgewählter Logdateien	1 Std.	Wichtig, wenn Logrotation nicht konfiguriert ist oder nicht konfiguriert werden kann

Tabelle 5: Vorgeschlagene Prüf-Intervalle für den Speicherplatz

Prüf-/Messpunkt	Intervall	Bemerkung
Ping auf Storage IP	5 Min.	Prüfung nur bei iSCSI erforderlich

Tabelle 6: Vorgeschlagene Prüf-Intervalle für den Storage

Prüf-/Messpunkt	Intervall	Bemerkung
Port erreichbar	15 Min.	
URL abrufbar	15 Min.	
Gültigkeit von HTTPS-Zertifikaten	1 Tag	

Tabelle 7: Vorgeschlagene Prüf-Intervalle für Web-Server und Infrastruktur

Benutzeroberfläche (Frontend) einer Anwendung über den Applikationsserver, das Netzwerk, die Datenbank bis zum darunterliegenden Betriebssystem und schlussendlich bis zur Hardware – möglichst in Echtzeit.

Vollumfassende Prüfungen auf allen Ebenen sind allerdings unrealistisch und unmöglich. Sie sind auch nicht nötig, denn mit einer Prüfung auf einer höheren Ebene (wie ASM) können auch Aussagen über die Funktionsfähigkeit darunterliegender Ebenen (wie SAN) getroffen werden. Wichtig sind jedoch Prüfungen, die vorab ein Indiz für einen drohenden Ausfall darstellen. Diese Abhängigkeiten werden wir später bei den einzelnen Prüfungen noch näher betrachten. Losgelöst vom Aspekt „minimalinvasive Überwachung" lässt sich sagen, dass, je detaillierter das Monitoring ist, desto einfacher später die Fehlersuche wird, wenn es doch zu einem Ausfall kommt.

Werkzeuge zur Überwachung

Als zentrales Werkzeug zur Überwachung bietet sich zunächst der Oracle Enterprise Manager an. Für diesen ist zumindest das Diagnostic Pack als zusätzliche Option erforderlich. Es findet sich darin aber eine Vielzahl von Prüfungen für die meisten genannten Ebenen bereits vorkonfiguriert. Leider ist das Diagnostic Pack nur in der Enterprise Edition der Oracle-Datenbank lizenzierbar.

Auch andere Anbieter wie BMC, Quest mit Foglight for Oracle, HP mit OpenView und IBM mit Tivoli bieten Werkzeuge zur Überwachung von Datenbank-Systemen und mehr. Ein kostengünstiges, weil lizenzkostenfreies Werkzeug ist Nagios, das Thema mehrerer Vorträge bei der DOAG war und ist. Es gibt darüber hinaus noch weitere Lizenzkosten-freie Alternativen.

Alle Werkzeuge bieten integrierte Prüfungen oder vorkonfigurierte Plug-ins dafür. Für weitere Prüfungen können eigene Plug-ins erstellt werden. Darin können Prüfungen/Messungen mit auf dem Server vorhandenen Befehlen ausgeführt werden (siehe Tabelle 2). Bei der Auswahl der Programme, mit denen eine Prüfung beziehungs-

weise Messung erfolgt, ist das Ziel, möglichst keine oder nur wenige zusätzliche Installationen durchzuführen.

Zeit-Intervalle für Prüfungen

Eine wichtige Entscheidung ist auch diejenige darüber, in welchem Intervall man eine Prüfung ausführt. Erfolgt diese zu oft, belastet man den Datenbank-Server unnötig stark. Erfolgt eine Prüfung zu selten, dann stellt man das Überschreiten einer kritischen Schwelle eventuell erst fest, wenn es für Korrektur-Maßnahmen zu spät ist, um einen Ausfall zu verhindern. Für jede Prüfung lässt sich ein optimales Intervall festlegen. Dabei kann man eine begrenzte Anzahl von Intervallen ansetzen, etwa alle fünf Minuten, alle fünfzehn Minuten, jede Stunde, alle sechs oder zwölf Stunden, einmal am Tag.

Schlägt eine Prüfung fehl beziehungsweise überschreitet ein Messwert die kritische Schwelle, ist es nicht immer sinnvoll, sofort Alarm zu schlagen (etwa eine kurzfristige CPU-Last über Faktor X). Hier sollten mehrere aufeinander folgende Prüfungen/Messungen abgewartet und nur Alarm ausgelöst werden, wenn alle das gleiche kritische Ergebnis liefern. Dabei ist es günstig, wenn sich das Intervall bei den Wiederholungsprüfungen verkürzt, insbesondere wenn das ursprüngliche Intervall lang ist.

Minimal erforderliche Prüf-/Messpunkte

Die Auswahl geeigneter Prüf- und Messpunkte ist nicht einfach. Dabei orientiert man sich häufig zunächst an bekannten Überwachungswerkzeugen wie Oracle Enterprise Manager. Diese realisieren aber teils sehr umfassende Prüfungen, um auch andere Aspekte wie das Tuning zu unterstützen. Mit der Zeit liefert die Erfahrung die Erkenntnis, welche Prüfungen einen Alarm liefern und tatsächlich vor einem Ausfall gewarnt haben. Die hier in den Tabellen 3 bis 15 vorgestellte Auswahl spiegelt die Erfahrungen des Autors wider. Für jede Ebene der Überwachung werden die relevanten Prüf-/Messpunkte aufgelistet.

Prüf-/Messpunkt	Intervall	Bemerkung
Tablespace-Füllgrad bzw. freier Platz im Tablespace	30 Min.	Prüfung nicht erforderlich, wenn Bigfile-Tablespace mit unbegrenzter Größe
Füllgrad der Fast-Recovery-Area (FRA)	15 Min. 6 Std.	wenn Ziel der Archivelogs wenn nur Ziel von Backups
Füllgrad Archivelog-Ziel	15 Min.	wenn FRA nicht verwendet wird

Tabelle 8: Vorgeschlagene Prüf-Intervalle für die Datenbank

Prüf-/Messpunkt	Intervall	Bemerkung
Instanz läuft im richtigen Modus (open/mount)	15 Min.	
Anmeldung an Instanz als nicht SYSDBA möglich	5 Min.	
Größe Alertlog-Datei	1 Std.	
ORA-Fehler in Alertlog-Datei	15 Min.	Im Laufe der Einschwingphase werden einige Meldungen von der Prüfung ausgenommen
Anzahl Session	5 Min.	
Intervall Redo-Log-Switches	15 Min.	
DB-Links erreichbar	5 Min.	

Tabelle 9: Vorgeschlagene Prüf-Intervalle für die Datenbank-Instanz

Prüf-/Messpunkt	Intervall	Bemerkung
Init-Prozesse aktiv	5 Min.	
alle Ressourcen aktiv	5 Min.	
alle Ressourcen auf den richtigen Knoten	5 Min.	Nur sinnvoll, wenn DB-Services auf einem Knoten spezifisch zugeordnet sind
CRS-Fehler in Alertlog-Datei	5 Min.	

Tabelle 10: Vorgeschlagene Prüf-Intervalle für die Clusterware

Prüf-/Messpunkt	Intervall	Bemerkung
Füllgrad der Diskgruppen	30 Min.	Wenn Daten-Dateien automatisch wachsen können oder Diskgruppe Ziel von Archivelogs bzw. Backup ist

Tabelle 11: Vorgeschlagene Prüf-Intervalle für Oracle ASM

Prüf-/Messpunkt	Intervall	Bemerkung
Instanz läuft im richtigen Modus (open)	5 Min.	
Anmeldung an Instanz als SYSDBA möglich	5 Min.	
alle Diskgruppen gemountet	15 Min.	
Größe Alertlog-Datei	1 Std.	Im Laufe der Einschwingphase werden einige Meldungen von der Prüfung ausgenommen
ORA-Fehler in Alertlog-Datei	15 Min.	

Tabelle 12: Vorgeschlagene Prüf-Intervalle für die ASM-Instanz

Prüf-/Messpunkt	Intervall	Bemerkung
Listener läuft	5 Min.	
Listener erreichbar: tnsping geht	5 Min.	
alle Instanzen registriert	5 Min.	
zusätzliche Services registriert	5 Min.	Im RAC werden die Services als Ressourcen geprüft
Größe Listener-Log-Datei	1 Std.	
Fehler in Listener-Log-Datei	15 Min.	Im Laufe der Einschwingphase werden einige Meldungen von der Prüfung ausgenommen

Tabelle 13: Vorgeschlagene Prüf-Intervalle für den Oracle Listener

Prüf-/Messpunkt	Intervall	Bemerkung
Agent läuft	15 Min.	
Agent Upload OK	15 Min.	
OMS läuft	15 Min.	

Tabelle 14: Vorgeschlagene Prüf-Intervalle für Grid Control

Prüf-/Messpunkt	Intervall	Bemerkung
Status Standby-DB	5 Min.	
SCN-Differenz	15 Min.	Abhängig von der gewünschten Verzögerung
Log-Transport-Verzögerung	15 Min.	
Log-Apply-Verzögerung	15 Min.	Abhängig von der gewünschten Verzögerung

Tabelle 15: Vorgeschlagene Prüf-Intervalle für Data Guard

Dazu noch folgende Anmerkungen: Die „rsh-/ssh“-Verbindung beim Netzwerk wird von einem fernen Rechner aus geprüft, die übrigen über lokale Prozesse. Ein Netzwerk-Ausfall von fünf Minuten ist schon sehr lang und wird voraussichtlich vorher schon von den Anwendern bemerkt beziehungsweise gemeldet. Alternativ können hier auch die Antwortzeiten gemessen und mit einem Schwellwert versehen werden.

Für den Platzverbrauch und die Logdatei-Größe im Speicher sind auch längere Intervalle zur Überwachung denkbar, beispielsweise ein Tag. Der Ausfall einer URL beim Web-Server wird voraussichtlich vorher schon von den Anwendern bemerkt beziehungsweise gemeldet. Bei der Datenbank ist eine Prüfung auf „Tablespace User Quota“ sinnvoll für die Sicherstellung des Betriebs einer Anwendung, nicht aber für die Sicherstellung des Betriebs der Datenbank. Eine Prüfung auf „blockierte Session“ in der Datenbank-Instanz erkennt eher ein Problem der Anwendung, ist jedoch nicht für die Sicherstellung des Betriebs der Instanz relevant. Die Prüfungen der Clusterware/Grid-Infrastruktur erfolgt in kurzen Intervallen, da deren Einsatz ja bedeutet, dass Hochverfügbarkeit gefordert ist. Eventuell sind die Intervalle auch auf eine Minute zu reduzieren. Im Umfeld einer Konfiguration für Hochverfügbarkeit sind auch die Intervalle für Prüfungen der anderen Bereiche eventuell zu reduzieren. Beim Oracle Listener kann man bei einem „tnsping“ von einem fernen Rechner aus das normale „ping“ auf die IPs der Net-Alias-Namen auch einsparen.

Auch ein Grid Control, das mit Diagnostic Packs zur Überwachung von Datenbank-Systemen dient, sollte überwacht werden. Bei Data Guard unterscheiden sich die Mittel der Prüfung einer Standby-Konfiguration, je nachdem, ob Standard oder Enterprise Edition im Einsatz sind.

Bei einigen Installationen erfolgen die Prüfungen nicht durch ein Überwachungssystem vor Ort, sondern durch ein zentrales System im Unternehmen des Autors. In diesen Fällen sind die Prüf-Intervalle meist erheblich länger, beispielsweise stündlich oder alle sechs Stunden. Für diese Kunden-Installationen gab es zuvor keine regelmäßige Überwachung, sie wurde erst im Rahmen der „Oracle Administration Services“ eingerichtet. Dort sind an sich nur solche Messpunkte sinnvoll, bei denen der Wert kontinuierlich anwächst und nach Überschreiten eines Schwellwertes die Auswirkungen auf den Betrieb erst erheblich später erfolgen (wie Platten-Füllgrad).

Fazit

Es gibt nicht die eine Antwort auf die Frage, was minimal überwacht werden muss. Denn jede Installation ist anders. Trotz Überwachung kann es in Extrem-Situationen zu unerwarteten Ausfällen des Datenbank-Systems kommen.

Aber auch eine hoch verfügbare Konfiguration bietet nur Sicherheit, wenn ihre Funktion überwacht wird. Daher sollte eine permanente, proaktive Überwachung in jeder Konfiguration erfolgen. Aus den zuvor aufgezeigten Prüf- und Messpunkten kann man sehr gut einen Initial-Satz passend zur eigenen Konfiguration auswählen. Die Schwellwerte werden zunächst mit viel Spielraum festgelegt. Dabei nimmt man Fehlalarme in einer Einschwingphase in Kauf. Sind die Schwellwerte später restriktiver gesetzt, sollte man Alarme hingegen beachten und die Ursachen gewissenhaft beheben.

Ralf Appelbaum
ra@team-pb.de