Minimalinvasive Überwachung von Datenbanken für optimale Verfügbarkeit

Ralf Appelbaum
TEAM GmbH
Paderborn

Schlüsselworte
Enterprise Manager, Nagios, proaktive Überwachung, minimale Metriken, Monitoring

Einleitung
Überwachung von Datenbanksystemen ist ein wichtiges Thema, um deren Verfügbarkeit zu gewährleisten. Über Werkzeuge zur Überwachung, wie z. B. Enterprise Manager oder Nagios, wurde schon viel berichtet. 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 will ich bzw. was muss ich mindestens überwachen, um die für mein Unternehmen optimale Verfügbarkeit zu gewährleisten?
Sicher kann man alle möglichen Messwerte erfassen lassen. Aber eine Überwachung soll möglichst minimalinvasiv, d.h. mit kleinstmöglichem Aufwand eingreifend, erfolgen und damit das Datenbanksystemen wenig zusätzlich belasten.

Im Rahmen der proaktiven Überwachung unserer „Oracle Administration Services“ haben wir uns bei verschiedensten Kundenumgebungen Gedanken dazu gemacht, welche Überwachung unter dem Aufwand-/Nutzen-Aspekt sinnvoll ist. Aufwand meint hier sowohl Kosten als auch Systembelastung und Nutzen ist die sichergestellte Verfügbarkeit.
Mit diesem Vortrag zeige ich auf, welche Messwerte in welchen Konfigurationen aus unserer Sicht minimal überwacht werden sollten und welche Ausfälle damit vermieden werden können. Aus unseren Erfahrungen berichte ich auch, welche Ausfälle sich nicht gänzlich vermeiden lassen.

Objekte der Überwachung
Im Rahmen unserer Dienstleistung, den Betrieb der Datenbanksysteme bei unseren Kunden zuverlässig zu gewährleisten, treffen wir auf die verschiedensten Konfigurationen und Komponenten von Datenbanksystemen. Daraus lässt sich eine Liste von Aspekten ableiten, die zwingend berücksichtigt werden muss:

- 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

Als Beispiel kann man die komplexe Konfiguration aus Kombination von RAC und Data Guard betrachten, die den Hauptbestandteil der so genannten Maximum Availability Architecture (MAA) bei Oracle bildet (hier eine vereinfachte Installation mit nur einem RAC, siehe Abb. 1).

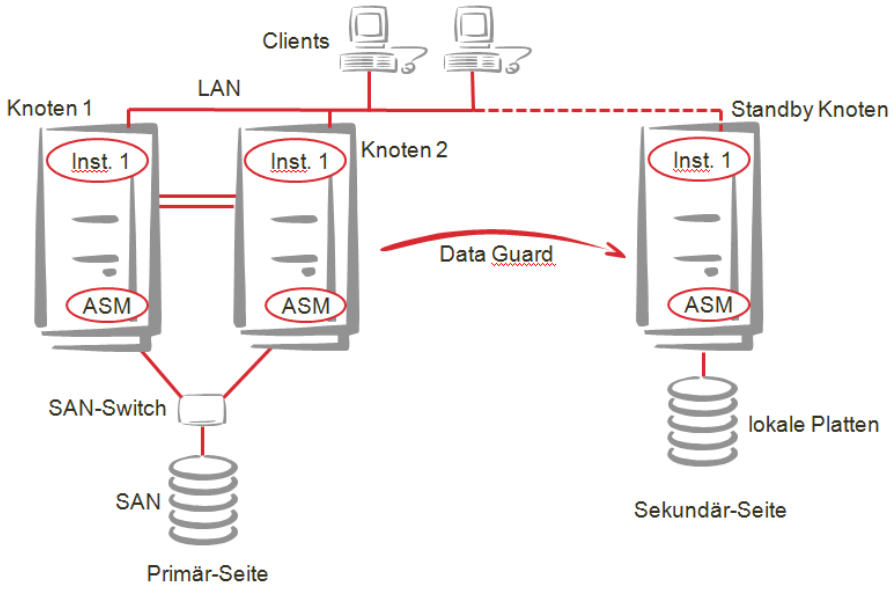

Abb. 1: vereinfachte Maximum Availability Architecture (MAA)

In dieser Konfiguration trifft man zahlreiche der oben aufgeführten Aspekte wieder. Dort finden sich Single Instance, RAC, Data Guard, lokales Dateisystem, SAN, ASM, Netzwerk mit virtuellen IPs und bei installiertem Database Control auch ein WEB Server. Diese Konfiguration lässt sich sowohl unter Unix/Linux realisieren, als auch unter Windows.
Und 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 Prüfungen ermitteln, ob eine Komponente des Datenbanksystems einwandfrei funktioniert oder ob Fehler gemeldet werden. In der Regel ist es bei einer Fehlermeldung aber schon zu spät (z.B. I/O nicht möglich da Dateisystem voll), weil durch den Ausfall einer einzelnen Komponente das gesamte Datenbanksystem ausfällt.
Daher werden überwiegend Prüfungen durchgeführt, die einen numerischen Wert, einen Messwert, liefern (z.B. Füllgrad des Dateisystems). 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 Datenbanksystems entgegenwirken.

Ziel der Überwachung ist also, den Betrieb eines Datenbanksystems ohne Unterbrechung sicher zu stellen. Sollte sich ein Ausfall nicht vermeiden lassen, so ist das Ziel zumindest das Datenbanksystem schnellstmöglich wieder in Betrieb zu setzen.

Dazu sind nur Prüfungen erforderlich, welche Aussagen über eine Gefährdung des Betriebs des Datenbanksystems oder dessen Ausfall machen, keine Prüfungen bzw. Messungen, welche ausschließlich zukünftiger Ressourcenplanung oder dem Performancetuning dienen.
Selbstverständlich können fast alle Messungen, wenn man die Werte im zeitlichen Verlauf betrachtet, auch zukünftige Ressourcenplanungen und Performancetuning unterstützen.

Fragestellungen zur Überwachung
Wie bereits in der Einleitung erwähnt, stellen sich Fragen, wenn man eine neue Überwachung einrichtet.
Nicht nur die eine:

„Was will ich bzw. was muss ich mindestens überwachen?“

sondern auch:

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

Auch wenn dieser Vortrag die erste Frage, „Was will ich bzw. was muss ich mindestens überwachen?“, zum Kernthema hat, möchte ich doch einige Hinweise zur Beantwortung der anderen Fragen geben.

Ebenen der Überwachung
Die Tabelle 1 listet mögliche Ebenen auf, an denen die Überwachung mit Prüfungen ansetzen kann.

Ebene	**beispielhafter Messwert**
Betriebssystem	Auslastung
Netzwerk	Erreichbarkeit
Speicherplatz	Verbrauch bzw. 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

Optimal ist sicher eine Überwachung von der obersten bis zur untersten Ebene eines Anwendungsstacks, d.h. von der Benutzeroberfläche (Frontend) einer Anwendung über den Applikationsserver, das Netzwerk, die Datenbank bis zum darunter liegenden Betriebssystem und schlussendlich bis zur Hardware. Und das soll möglichst in Echtzeit erfolgen.

Vollumfassende Prüfungen auf allen Ebenen sind aber unrealistisch und unmöglich. Sie sind auch nicht nötig, denn mit einer Prüfung auf einer höheren Ebene (z.B. ASM) können auch Aussagen über die Funktionsfähigkeit darunter liegender Ebenen (z.B. SAN) getroffen werden. Wichtig sind aber 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 aber sagen, je detaillierter das Monitoring ist, desto einfacher wird später die Fehlersuche, 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. Darin findet sich 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, IBM mit Tivoli und einige andere bieten Werkzeuge zur Überwachung von Datenbanksystemen und mehr. Ein kostengünstiges, weil Lizenzkosten freies Werkzeug ist Nagios, welches Thema mehrerer DOAG Vorträge war und ist. Es gibt noch weitere Lizenzkosten-freie Alternativen.

All diese Werkzeuge bieten integrierte Prüfungen oder vorkonfigurierte Plug-ins für Prüfungen. Für weitere Prüfungen können eigene Plug-ins erstellt werden. In selbst erstellten Plug-ins können Prüfungen/Messungen mit auf dem Server vorhandenen Befehlen ausgeführt werden (siehe Tabelle 2).

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

Bei der Auswahl der Programme mit denen eine Prüfung bzw. Messung erfolgt ist das Ziel, möglichst keine oder nur wenige zusätzliche Installationen durchzuführen.

Zeitintervalle für Prüfungen
Eine wichtige Entscheidung ist auch, in welchem Intervall wird eine Prüfung ausgeführt. Erfolgt die Prüfung zu oft, wird der Datenbankserver unnötig stark belastet. Erfolgt eine Prüfung zu selten, dann stellt man das Überschreiten einer kritischen Schwelle eventuell erst fest, wenn es für Korrekturmaßnahmen zu spät ist um einen Ausfall zu verhindern.
Ein optimales Intervall ist für jede Prüfung individuell festzulegen. Dabei kann man eine begrenzte Anzahl Intervalle ansetzen, z.B. alle 5 Min., alle 15 Min., jede Stunde, alle 6 bzw. 12 Std., einmal am Tag.
Schlägt eine Prüfung fehl bzw. überschreitet ein Messwert die kritische Schwelle, ist es nicht immer sinnvoll sofort Alarm zu schlagen (z.B. kurzfristige CPU Last über Faktor X). Hier sollten mehrere aufeinander folgende Prüfungen / Messungen abgewartet werden 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, z.B. den Oracle Enterprise Manager. Diese realisieren aber teils sehr umfassende Prüfungen, um auch andere Aspekte, z.B. Tuning, zu unterstützen. Die Erfahrung bringt dann mit der Zeit die Erkenntnis, welche Prüfungen Alarm liefern und tatsächlich vor einem Ausfall gewarnt haben.
Die hier dargestellte Auswahl spielt unsere Erfahrungen wieder. Für jede Ebene der Überwachung werden die relevanten Prüf-/Messpunkte aufgelistet.

Betriebssystem

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.	

Die rsh-/sshVerbindung wird von einem fernen Rechner aus geprüft, die Übrigen über lokale Prozesse.

Netzwerk

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

Diese Prüfungen erfolgen von einem fernen Rechner. Ein Netzwerkausfall von 5 Min. ist schon sehr lang und wird voraussichtlich vorher schon von den Anwendern bemerkt bzw. gemeldet. Alternativ können hier auch die Antwortzeiten gemessen und mit einem Schwellwert versehen werden.

Speicherplatz

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

Für Platzverbrauch und Logdateigröße sind auch längere Intervalle zur Überwachung denkbar, z.B. 1 Tag.

Storage

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

Oracle Datenbank

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

Eine Prüfung auf „Tablespace User-Quota“ z.B. ist sinnvoll für die Sicherstellung des Betriebs einer Anwendung, nicht aber für die Sicherstellung des Betriebs der Datenbank.

Oracle DB Instanz

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.	

Eine Prüfung auf „blockierter Session“ z.B. erkennt eher ein Problem der Anwendung, ist aber nicht für die Sicherstellung des Betriebs der Instanz relevant.

Oracle Clusterware / Grid Infrastructure

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

Die Prüfungen der Clusterware/Grid Infrastruktur erfolgt in kurzen Intervallen, da deren Einsatz ja bedeutet, dass Hochverfügbarkeit gefordert ist. Im Umfeld einer Konfiguration für Hochverfügbarkeit sind auch die Intervalle für Prüfungen der anderen Bereiche eventuell zu reduzieren.

Oracle ASM

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

Oracle ASM Instanz

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.	
ORA-Fehler in Alertlog-Datei	15 Min.	im Laufe der Einschwingphase werden einige Meldungen von der Prüfung ausgenommen

Oracle Listener

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 Listenerlog-Datei	1 Std.	
Fehler in Listenerlog-Datei	15 Min.	im Laufe der Einschwingphase werden einige Meldungen von der Prüfung ausgenommen

Bei einem tnsping von einem fernen Rechner aus kann man das normale ping auf die IPs der Net-Aliasnamen auch einsparen.

Data Guard

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

Grid Control / Cloud Control

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

Auch ein Grid Control, das mit Diagnostic Packs zur Überwachung von Datenbanksystemen dient, sollte überwacht werden.

Web-Server / Infrastruktur

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

Der Ausfall einer URL wird voraussichtlich vorher schon von den Anwendern bemerkt bzw. gemeldet.

Bei einigen unserer Kunden erfolgen die Prüfungen nicht durch ein Überwachungssystem beim Kunden, sondern durch ein zentrales System bei TEAM. In diesen Fällen sind die Prüfintervalle meist erheblich länger, z.B. stündlich oder alle 6 Stunden. Für diese Kundeninstallationen gab es zuvor keine regelmäßige Überwachung, diese wurde erst im Rahmen der TEAM „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 (z.B. Plattenfüllgrad).

Eine genauere Diskussion, in welchen Konstellationen welche Prüfungen erforderlich sind und warum andere Prüfungen ggf. nicht erforderlich sind, soll auch in der Fragerunde am Ende des Vortrages stattfinden.

Resümee

Es gibt nicht die eine Antwort auf die Frage, was minimal überwacht werden muss. Denn jede Installation ist anders. Und trotz Überwachung, kann es in Extremsituationen zu unerwarteten Ausfällen des Datenbanksystems 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 initialen 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 aber auch beachten und die Ursachen gewissenhaft beheben.

Kontaktadresse:
Ralf Appelbaum
TEAM
Partner für Technologie und
angewandte Methoden der
Informationsverarbeitung GmbH
Hermann-Löns-Str. 88
D-33104 Paderborn

Telefon:	+49 (0)5254 / 8008-37
Fax:	+49 (0)5254 / 8008-19
E-Mail:	ra@team-pb.de
Internet:	http://www.team-pb.de