

€ 14,95 Deutschland € 14,95 Österreich € 16,45 Schweiz sfr 29,90

MySQL

- **Leistung: Performance-Engpässe vermeiden**
- **Verwaltung: Daten-Logger im Eigenbau**
- **Engines: Das richtige Tabellenformat wählen ab S. 60**

O/R-Mapper

Technik und Produkte für .NET S. 36

2 2009 DVD

database pro

Profi-Datenbanken

Tools & Specials

DVD: über 7,5 GB

DATENBANKEN

(Test- und Demoversionen, Auswahl)

- Advantage Database Server 9.1
- IBM DB2 Express-C 9.5
- IBM Data Studio 2.1 Developer
- IBM DB2 9.5
- Intersystems Caché
- MySQL 5.1 und 6.0

TOOLS & SPECIALS

(Test- und Demoversionen, Auswahl)

- Quest Toad DBA Suite for IBM DB2
- Quest LiteSpeed Engine for Oracle
- KeepTool 8.1.6.4
- HoraLight 8.1.6.4
- Altova XML 2009

DT-Control
geprüft:

Nicht jugend-
beeinträchtigend

Sharepoint

Effektives Formularsystem mit Infopath und SQL Server S. 22

Multithreading

Datenbank-Entwicklung für Multicore-Prozessoren S. 84

IBM

- Power Systems mit DB2 for i S. 44
- DB2 Express-C installieren S. 52
- Monitoring und Tuning S. 56

XML und XQuery verstehen

Praktisches Grundwissen für Datenbankprofis S. 28/S. 97

Tipps, Tests & Trends

Web-Intelligence mit Lixto Enterprise Mashup Framework S. 16
dazu: Oracle – Index tunen, Refactoring mit Liquibase, SSDS mit SOAP, VisualCockpit, MicroStrategy 9, Compliance im DB-Betrieb



Datenbank-Monitoring und -Tuning bei Hagebau mit DB2 und Speedgain for DB2

Der Weg ins Licht

Durch eine systematische Datenbanküberwachung lassen sich Performance-Probleme beim Einsatz eines DB2-Servers auch in komplexen Firmenumgebungen optimieren. **Martin Ortgies**

Auf einen Blick

Inhalt

Für den praktischen und effizienten Einsatz eines Datenbankservers ist es erforderlich, dessen Leistung permanent zu überwachen und bei Engpässen möglichst schnell die erforderlichen Leistungs-optimierungen vorzunehmen. Wie die Überwachungsergebnisse ermittelt, ausgewertet und zur Performance-Optimierung nutzbar sind, zeigt dieser Praxisartikel anhand der Datenbank DB2 und des Analysetools Speedgain for DB2 bei der Firma Hagebau.

Plattform

DB2

Technik/Anwendung

DB2-Datenbank-Monitoring, Überwachung, Auswertung, Optimierung

Voraussetzungen

DB2-Datenbankserver, Speedgain for DB2

Autor

Martin Ortgies ist als selbstständiger Kommunikationsberater tätig und leitet das Unternehmen Ortgies Marketing & Communications, das sich der Kommunikation und Pressearbeit widmet.

Die Kommunikation zwischen der Zentrale (Bild 1), den Lieferanten und den einzelnen Standorten erfolgt bei der Firma Hagebau vorwiegend über ein eigenes Kommunikationsportal, das Hagebau-Extranet. Hier finden die Standorte und die Zulieferer je nach Zugangsberechtigung alle erforderlichen Informationen und Funktionen von der Zentrale. Wo früher für die aktuellen Informationen viel Papier bewegt werden musste, erhalten die 12.000 registrierten Nutzer jetzt jede Nacht einen Newsletter mit den aktuellsten Dokumenten. Die Administration der Nutzer und Redakteure liegt in den Händen der Fachabteilung Extranet. Die Informationen für das Extranet werden von 300 Redakteuren in den Fachabteilungen der Hagebau-Zentrale sowie der Zentrale in Österreich und den Zentrallagern erstellt und veröffentlicht. Nach der Übernahme des Datenbank-servers in das eigene Rechenzentrum gab es mit der DB2-Datenbank zunächst massive Performance-Probleme in Form extrem langer Antwortzeiten, die durch ein systematisches Datenbank-Monitoring behoben wurden.

Extranet im eigenen Rechenzentrum

Das Extranet war in der Vergangenheit von einem externen Dienstleister betrieben worden. Wegen zu vieler Reibungsverluste hatte sich Hagebau entschieden, den Betrieb selbst zu übernehmen. Die notwendige Infrastruktur war schnell bereitgestellt; auch die Installation der zugrunde liegenden DB2-Datenbank verlief weitgehend unproblematisch. In der mehrwöchigen Umstellungs- und Migrationsphase traten nur kleinere Probleme und Fehler auf. „Ernsthafte Probleme gab es erst mit Aufnahme des Produktivbetriebs ab einer größeren Anzahl von Usern im Netz. Die Performance wurde sehr schlecht mit extrem langen Antwortzeiten“, beschreibt Torsten Wattering von der Zentral-IT Systemadministration das Problem. Die EDV-Mitarbeiter konnten beim Datenbankserver eine Überlastung der Prozessoren feststellen, ohne dass die Ursachen dafür ersichtlich waren. Das Team hatte zwar Erfahrungen mit den Datenbanken SAP, Oracle und Progress, DB2 war allerdings Neuland. Deshalb wurde zu den DB2-Spezialisten von ITGain aus Hamburg Kontakt



Die Zentrale der Firma Hagebau (Bild 1)

aufgenommen, die schnell vor Ort waren, um eine Analyse der Probleme durchzuführen.

DB2-Performance-Engpass

„Für einen optimalen Betrieb hätten auf dem Datenbankserver, ausgestattet mit vier CPUs, maximal sechs Anwendungen im ausführenden Status sein dürfen. Diese sechs Anwendungen wären im DB2-Applikationsstatus bei einer annehmbaren Performance optimal für eine maximale Auslastung. Tatsächlich haben wir 35 Verbindungen zu DB2 registriert“, berich- ▶

Die Hagebau-Unternehmensgruppe

Die Hagebau-Unternehmensgruppe in Deutschland und Österreich besteht aus circa 295 inhabergeführten mittelständischen Unternehmen aus dem Baustoff-, Holz- und Fliesenhandel mit 1.180 Betriebsstätten. Die angeschlossenen Unternehmen werden aus der Hagebau-Zentrale in Soltau beziehungsweise in Brunn am Gebirge mit Dienstleistungen aus den Bereichen Marketing, Versicherungen und EDV versorgt. Die EDV ist in Soltau mit 100 Mitarbeitern einer der größten Dienstleistungsbereiche und firmiert als eigene Gesellschaft unter dem Namen Hagebau Datendienst IT-Service GmbH.

Datenbank-Monitoring: Schritt für Schritt zum Kern des Problems

Thomas Kalb, Geschäftsführer von ITGain: „Datenbankprobleme sind oft komplex und schwer durchschaubar. Datenbank-Monitoring ist trotzdem kein Hexenwerk. Mit einem geeigneten Werkzeug wird das Performance-Tuning für jeden Datenbank-Administrator zu einem überschaubaren Handwerk. Die Voraussetzung: Zur Problemerkennung ist ein systematisches Vorgehen notwendig sowie ein geeignetes Tool, das dieses Vorgehen unterstützt.“

Schritt 1: Gibt es überhaupt ein DB2-Problem?

Liegt das Problem innerhalb oder außerhalb der Datenbank (zum Beispiel im Netzwerk, der Applikation und so weiter)? Welche Komponente hat den größten Anteil an der gesamten Laufzeit? Hat sich im Verhalten der Komponenten etwas verändert?

Schritt 2: Die Problemressource bestimmen

Wenn das Problem eindeutig in der Datenbank liegt, ist zu klären, welche Ressource den Engpass beziehungsweise den höchsten Zeitanteil darstellt (Bild 2). Die Monitoring-Analyse erfolgt nach dem Ausschlussprinzip: Ist es ein CPU- oder ein I/O-Problem? Welche Ressource wird überwiegend benötigt (CPU-Bound oder I/O-Bound)?

Schritt 3: Den Verursacher bestimmen

Im nächsten Schritt müssen die Abfragen identifiziert werden, die den Engpass beziehungsweise den höchsten Zeitanteil verursachen und folglich auch das größte Tuning-Potenzial bieten. Dafür wird der SQL-Workload systematisch untersucht. Als goldene Regel kann die Auswahl der SQL-Abfragen dienen, die für circa 80 Prozent des Gesamtverbrauchs verantwortlich sind.

Schritt 4: Die Tuning-Aktivitäten zur Problemlösung

Je nach ermittelter Problemressource müssen für die verursachenden SQL-Abfragen systematisch entsprechende Maßnahmen getroffen werden. Liegt das Problem etwa im I/O-Bereich, muss die Read-Effizienz gesteigert werden. Dies kann durch zusätzliche Indizes oder durch ein effizientes Caching erfolgen.

Schritt 5: Das Monitoring der Tuning-Maßnahmen

Um den Erfolg der Tuning-Maßnahmen bewerten zu können, wird ein geeignetes Datenbank-Monitoring-Programm benötigt. Es hilft, die Wechselwirkung zwischen der Tuning-Maßnahme und der Auswirkung zu prüfen. Um differenzierte Aussagen zu treffen, ist es wichtig, die einzelnen Tuning-Maßnahmen nur nacheinander durchzuführen.

tet Thomas Kalb, Datenbank- und Tuningexperte von ITGain. Um das Problem zu analysieren, nutzte ITGain den Datenbank-Monitor Speedgain for DB2. Ausgangspunkt waren die vom Datenbank-Monitor gespeicherten Datenbankwerte. Per Ausschlussmethode wurde das Problem Schritt für Schritt eingegrenzt und es wurden die Verursacher ermittelt (siehe Textkasten „Datenbank-Monitoring: Schritt für Schritt zum Kern des Problems“).

Ursachenanalyse und Diagnose

Am Anfang der Analyse stand die Frage, warum die Anwendung so lange brauchte. Dafür wurde anhand der Verbrauchsübersicht des Datenbank-Monitors zunächst analysiert, welche Res-

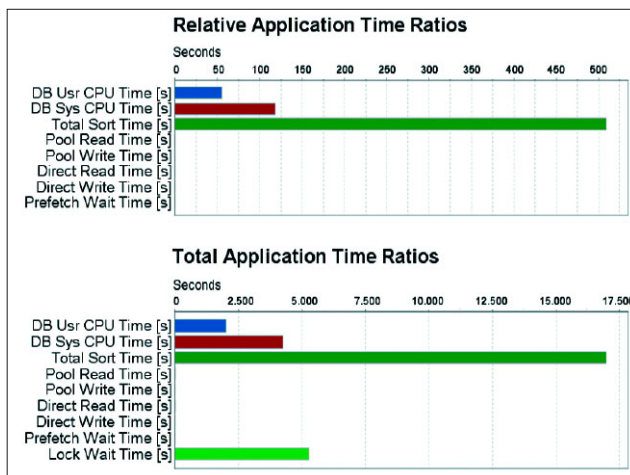
ourcen von der DB2-Datenbank am meisten verwendet wurden. Die Verbrauchsübersicht gibt einen Überblick, welche Ressourcen mit welchen Zeitanteilen beteiligt sind, und erlaubt eine direkte Diagnose (Bild 2). Auf den ersten Blick wurde sichtbar, wo sich der Engpass befand: Bei der Ergebnisausführung einer Abfrage benötigte die Datenbankanwendung viel zu viel Zeit für Sortierungen (Total Sort Time). Die Zeit für I/O-Zugriffe spielte dagegen keine Rolle.

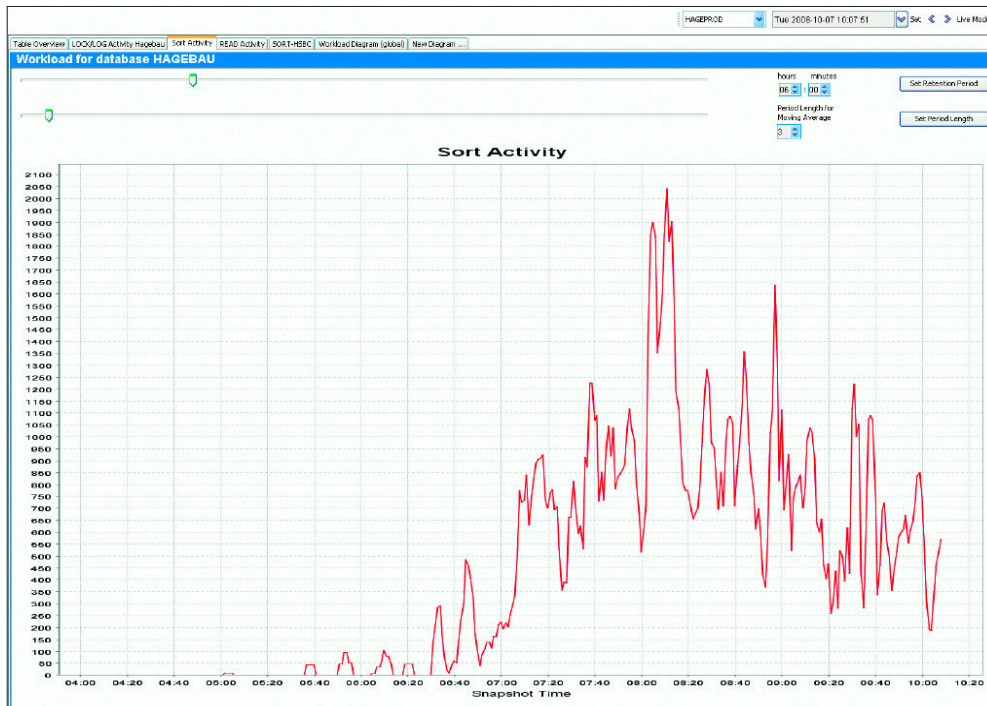
Im nächsten Schritt wurden die Sortierkennzahlen überprüft. Das Ergebnis: Die Anzahl der Sort-Operationen war enorm. Innerhalb einer Sekunde ergaben sich mehr als 2.000 Sortierungen (Bild 3). Gemäß der Vorgehenslogik einer problemorientierten Analyse wurde im nächsten Schritt systematisch nach dem Verursacher für das hohe Sort-Aufkommen gesucht. Dazu wurde der SQL-Workload analysiert, um diejenigen Abfragen zu identifizieren, die ursächlich für die hohe Zahl der Sortierungen waren. Nach den Erfahrungen von ITGain sind bei einer hohen Sortieranzahl in der Regel nur wenige SQL-Abfragen für die Masse der Sortierungen (rund 80 Prozent) verantwortlich.

Von der Analyse zur Problemlösung

Um den Ressourcenverbrauch zu reduzieren, gab es jetzt prinzipiell zwei Möglichkeiten: den Sort zu optimieren oder den Sort zu vermeiden. Bei nur wenigen verursachenden SQL-Abfragen lag es nahe, die Sortierungen durch die

Verbrauchsübersicht vor dem Tuning (Bild 2)





Anzahl der Sortierungen vor dem Tuning (Bild 3)

Definition zusätzlicher Indizes zu vermeiden. Durch die Analyse der sortierintensiven SQL-Anforderungen wurden diejenigen Zugriffsbestandteile sichtbar, die einen Sort benötigten (Bild 4). Die durchschnittliche Antwortzeit (rote Line, Baseline) betrug 3 Sekunden. Die Baseline entspricht der durchschnittlichen Antwortzeit aller bisher gemessenen Werte. Die blaue Linie gibt die aktuellen Verweilzeiten aller Anwendungen in DB2 wieder und die grüne Linie die Anzahl der Anwendungen. Auf Basis dieser Informationen wurden zur Problemlösung schließlich mehrere Indizes definiert. Diese Maßnahmen zur Sort-Vermeidung führten zum vollen Erfolg. Nach dem Tuning betrug die durchschnittliche Antwortzeit weniger als 0,25 Sekunden (Bild 5). Nach Nutzung der Indizes befanden sich auf dem Datenbankserver statt 35 nur noch 2 Anwendungen im Status *ausführend* (Durchschnittswerte). Die Zahl der Sortierungen reduzierte sich von 2.000 auf nur mehr 90 Sorts pro Sekunde (Bild 6). Die sogenannte Baseline zeigte einen Performancegewinn um den Faktor 15 von 3 Sekunden auf 0,2 Sekunden. Von der gesamten Verweilzeit entfällt nach dem Tuning nur noch ein geringer Anteil auf die Sortierungen (Bild 7).

Erfolgreiches Datenbank-Monitoring

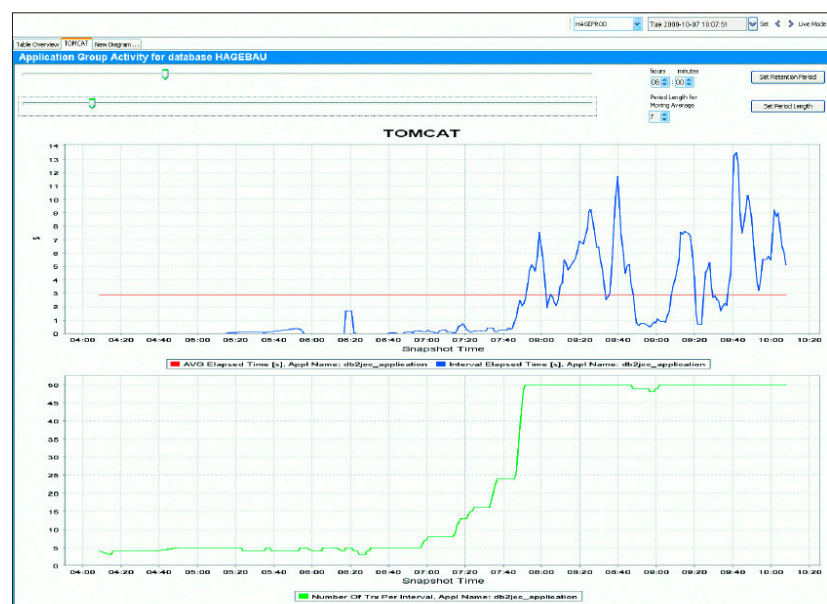
Mithilfe des Datenbank-Monitors Speedgain konnte ITGain die Performance-Probleme beheben. Das Hagebau-Extranet ist wieder voll verfügbar, und die Performance wurde gegenüber dem früheren Betrieb beim externen Dienstleister sogar noch verbessert. „Durch systematisches Datenbank-Monitoring wurden

komplexe Datenbankprobleme transparent und lösbar und zusätzlich weitere Optimierungspotenziale gefunden. Das hat uns überzeugt. Die Spezialisten von ITGain waren sehr engagiert und haben sich sehr flexibel gezeigt“, fasst Torsten Wettering die Ergebnisse zusammen. Hagebau hat sich nach den guten Erfahrungen entschieden, das Datenbank-Monitoring künftig weiter einzusetzen und als Dienstleistung von ITGain zu nutzen [1].

Anforderungen an ein Datenbank-Monitoring

Für das systematische Vorgehen ist ein geeignetes Datenbank-Monitoring-Programm wie Speedgain for DB2 sinnvoll. Ein Datenbank- ▶

Antwortzeiten vor dem Tuning (Bild 4)



Baseline-Monitoring

Gibt es wieder Probleme mit der Datenbank? Diese Frage eines Anwenders ist das tägliche Brot der Datenbank-Administratoren. Dass die Probleme nicht in der Datenbank liegen, akzeptieren viele Anwender allerdings nicht, weil es meistens an zuverlässigen und glaubwürdigen Nachweisen fehlt. Der neue Ansatz des Baseline-Monitorings verspricht Hilfe für Datenbankteams. Mit Einführung des Baseline-Monitorings kann der Administrator jetzt auf einen Blick dokumentieren, dass die Antwortzeiten innerhalb der Datenbank für eine Applikationsgruppe unverändert, besser oder schlechter sind als an den Tagen zuvor. DB2-Administratoren arbeiten mit vielen internen Kennzahlen, um ihre Datenbanken zu kontrollieren. Aber welche Aussagekraft hat die aktuelle Antwortzeit

oder der Wert für einen Datenbank-I/O? Sind die Werte gut oder schlecht? Das Problem: Ohne eine definierte Referenz ist keine nachvollziehbare Beurteilung möglich. „Wollen wir dem Datenbankanwender eine gute Dienstleistung liefern, müssen wir einen Nachweis erbringen, ob die aktuellen Werte besser oder schlechter sind als die zufriedenstellende Norm. Das Instrument dafür ist das Baseline Monitoring“, so Thomas Kalb. Baseline Monitoring soll dem Administrator die Mittel in die Hand geben, Performance- und Verfügbarkeitswerte zu überprüfen und zu bewerten, gezielt Schwachstellen zu erkennen, zu analysieren und auch proaktiv tätig zu werden. Baselines beschreiben Referenzwerte als „Normleistung“ der Datenbank, die der Nutzer erwarten kann.

Monitor liefert differenzierte Diagnosen zur Schwachstellenanalyse und zeigt außerdem Verbesserungspotenziale für die Optimierung der Datenbank auf.

- **Monitoring:** Der Datenbank-Monitor speichert Informationen zur Verfügbarkeit, Performance-Geschwindigkeit und Auslastung der Datenbanksysteme. Ein 24-Stunden-/7-Tage-Monitoring ermöglicht die laufende Überwachung und Benachrichtigung zum Zustand der Datenbank.
- **Alarmierung:** Die Alarmsituationen werden sofort angezeigt (je nach Programm durch optische Signale, akustische Meldungen oder über

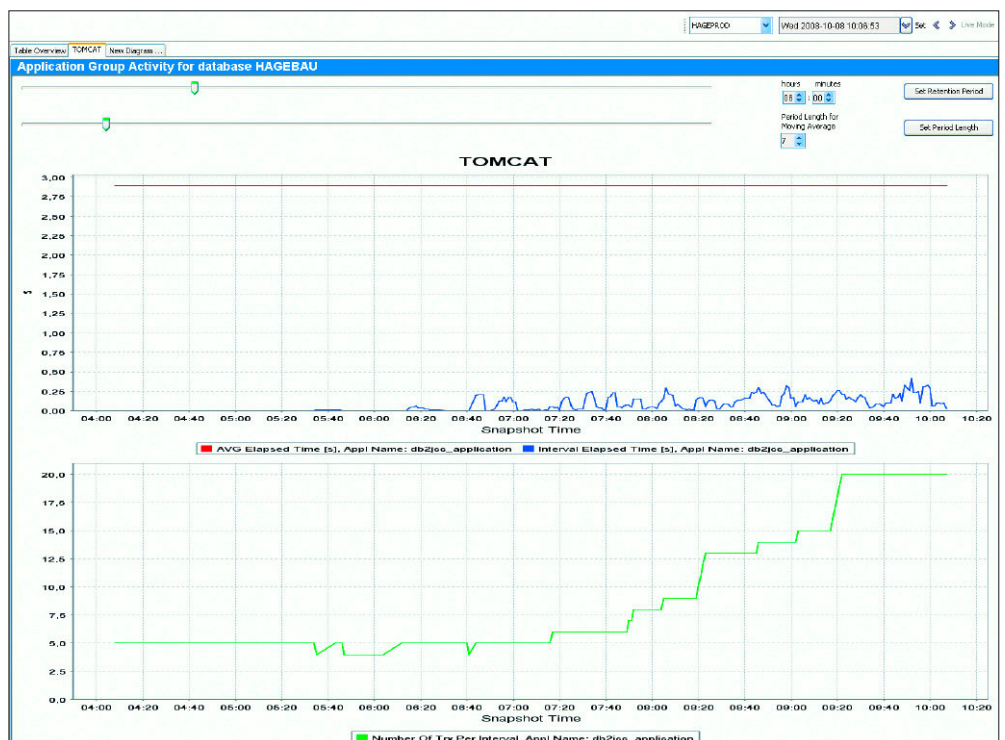
externe Schnittstellen an ein zentrales Enterprise-Monitoring).

- **Analyse:** Durch das Programm werden problematische Datenbankwerte angezeigt, die Ursachen analysiert und in Auswertungen dargestellt.
- **Historie:** Auf der Basis der gesammelten Daten können Langzeitanalysen sowie Kapazitätsplanungen für das Management zur Verfügbarkeit, Performance und Auslastung der Datenbanksysteme erstellt werden.
- **Workload-Analyse und Workload-Tuning:** Für ein effizientes SQL-Workload-Tuning ist es unabdingbar, dass der SQL-Workload in einer normierten Form vorliegt. Dafür werden gleichartige SQL-Statements mit dem gleichen Zugriffspfad zusammengeführt.
- **Baseline-Monitoring:** Ein neuer Ansatz ist hier das Baseline-Monitoring. Es liefert „Normwerte“ der Datenbank, die mit aktuellen Werten zu Verfügbarkeit, Performance und Auslastung vergleichbar sind, um sofort Abweichungen von der Norm erkennen zu können (vergleiche Textkasten „Baseline-Monitoring“).

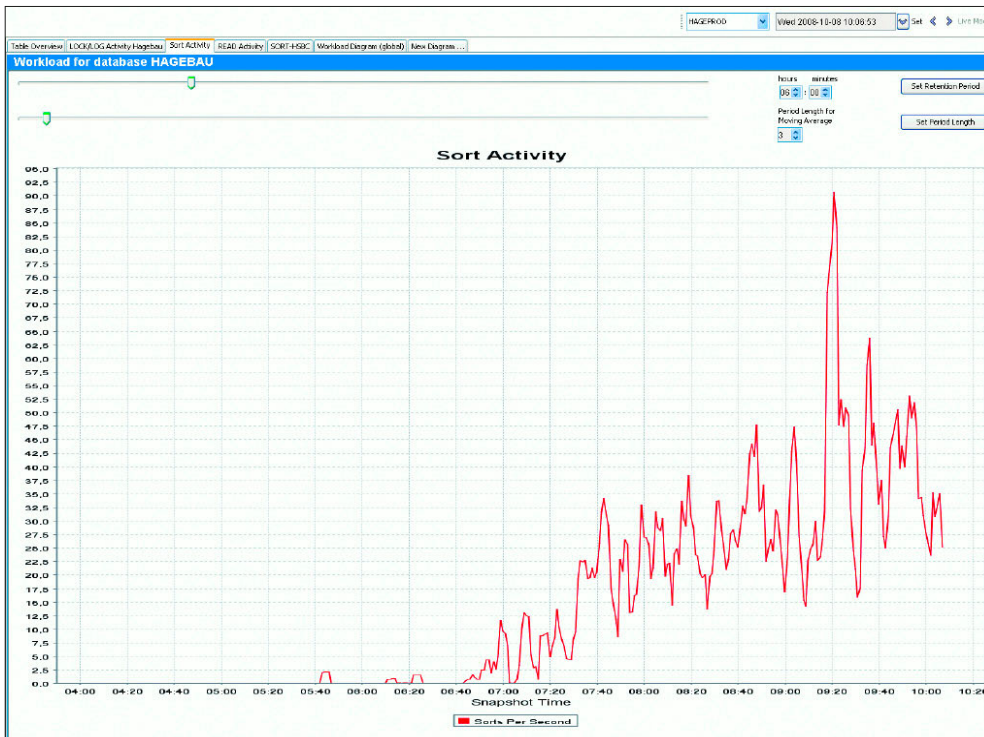
Der Ansatz für das Baseline Monitoring

Für jedes Anwendungssystem und für jede Abfrageebene werden über einen aussagekräftigen Zeitraum Durchschnittswerte ermittelt und als Normwert/Baseline definiert. Florian Boldt, bei ITGain Produktmanager für Speedgain for DB2, sagt dazu: „Das Ziel ist, differenzierte Durchschnittswerte als Basis zur Interpretation zu erhalten für typische Zeiten mit Leistungsspitzen / Leistungstälern / Wochentagen / Wo-

Antwortzeiten nach dem Tuning (Bild 5)



Anzahl der Sortierungen nach dem Tuning (Bild 6)



chenenden oder bekannten Lastzeiten/Abrechnungsterminen.“ Sind auf diese Weise Normwerte ermittelt, kann dem Anwender plausibel nachgewiesen werden, dass die Servicezeiten der Datenbank genauso gut (oder schlecht) sind wie in vergleichbaren Perioden. Datenbankteams können außerdem besser proaktiv werden, indem sie aussagekräftige dynamische Schwellenwerte festlegen. Weicht zum Beispiel die Antwortzeit signifikant über längere Zeit um zum Beispiel 75 Prozent von der Norm ab (gesetzter Schwellenwert), signalisiert die Warnampel Handlungsbedarf. Das Datenbankteam verfügt mit der Baseline auch über eine Messlatte, um seine Dienstleistung messbar und überprüfbar zu machen – zum Beispiel in Form von Service Level Requirements oder Service Level Agreements.

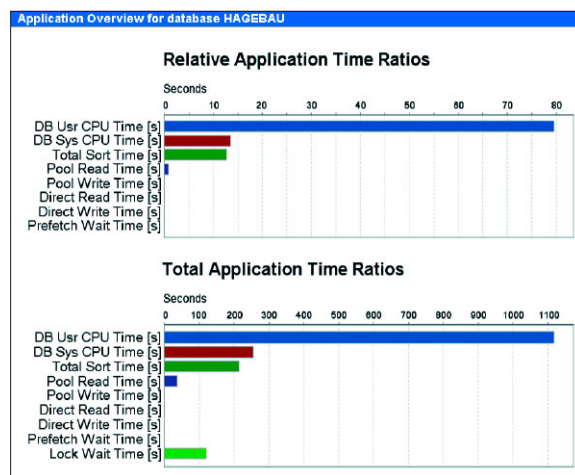
Externe und interne Baselines messen

Die externe Baseline beschreibt die „Normzeit“, die eine Anwendung in der Datenbank benötigt, und zwar von der Anfrage bis zur Antwort (Datenbank-Anwendungszeit). An der externen Baseline als Servicezeiten der Datenbank kann sich der Datenbank-Administrator exakt messen lassen. Sie kann auch für eine Service-Level-Vereinbarung genutzt werden. Die internen Baselines liefern dem Datenbank-Administrator ein Frühwarnsystem, um proaktiv tätig zu werden. Für die internen Baselines werden typische

interne Leistungsdaten genutzt, wie die durchschnittliche physische Getpage-Request-Zeit, Werte für das Schreiben eines Log-Buffers oder die Anzahl physischer Rows, die pro Sekunde gelesen werden. Nicht normale Abweichungen von diesen internen Baselines sind dann der Ausgangspunkt, um Performance-Schwachstellen frühzeitig zu erkennen und zu beseitigen. Diese Baselines für interne Leistungsdaten werden auch häufig genutzt, um aktuelle Schwachstellen des Operating-Systems zu erkennen.

Entsprechend dem in diesem Artikel vorgestellten Beispiel können Sie natürlich auch die Leistung anderer Datenbanken überwachen und die über gesonderte Zusatz-Tools abgefragten Leistungswerte zur Leistungsoptimierung heranziehen. [am]

[1] Hagebau-Portal im Internet; www.hagebau.de
 [2] ITGain-Portal im Internet; www.itgain.de



Verbrauchsübersicht nach dem Tuning (Bild 7)

IT im Hagebau-Extranet

Das Redaktionssystem für das Hagebau-Extranet wird als Java-Applikation unter Linux auf einer DB2-Datenbank betrieben. In zwei räumlich voneinander getrennten Rechenzentren läuft das Extranet unter VMware auf Intel-basierten Prozessoren. Ein EMC-Storage-System sorgt dabei für einen Disaster-Recovery-Schutz. Das gesamte Projekt wird mit insgesamt sieben Mitarbeitern betrieben, davon kommen drei aus der Fachabteilung (fachliche Betreuung der Anwender/Extranet-Hotline) und zwei aus der EDV (Technik/Administration).