ntz Heft 1/2010

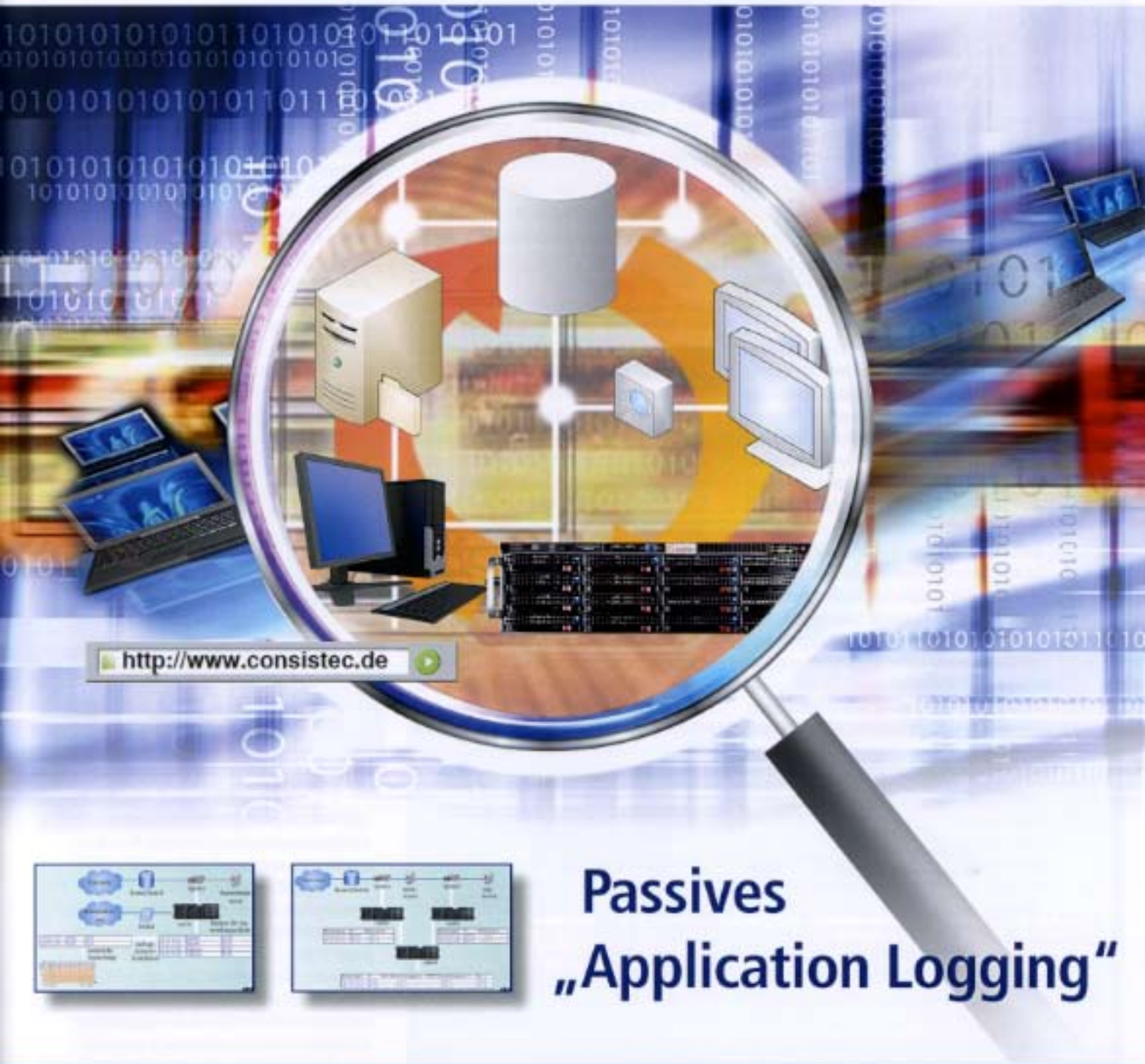
Sonderdruck

ntz

VDE VERLAG

Fachzeitschrift für

Informations- und Kommunikationstechnik



Passives „Application Logging"

Passives „Application Logging“ zum Betreiben und Überwachen komplexer Systemlandschaften

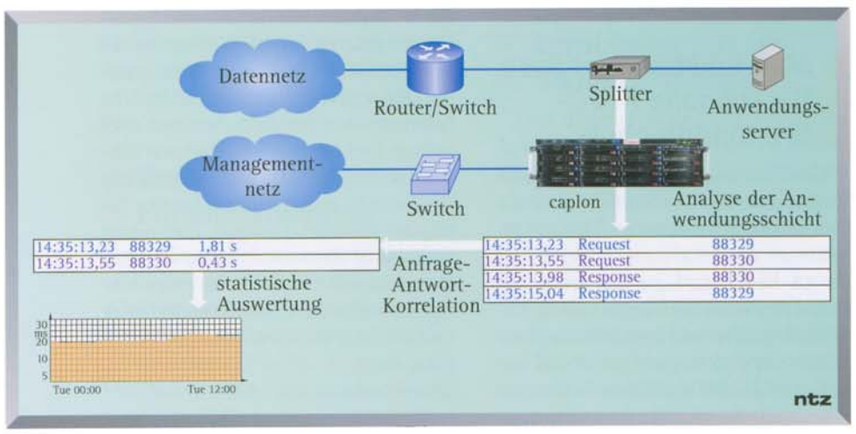
Martin Nicolay

Die Fokussierung auf Geschäftsprozesse in Unternehmen, ihre Standardisierung (ISO 9001) und Automatisierung ist untrennbar mit dem Siegeszug von serviceorientierten Architekturen (SOA) verbunden. Dabei entstehen in der IT verteilte Systeme, die Informationen unterschiedlicher Drittsysteme bündeln. Auf diese greifen wiederum verschiedene Dienste zu, die jedoch eigene Strukturen beinhalten. Solche Systeme benötigen angepasste Überwachungsstrukturen, um die Leistungsfähigkeit jederzeit zu prüfen und zu gewährleisten.

Die Verknüpfung einzelner Systeme mit den Services eines Unternehmens, z. B. die Kundendatenbank, ist sehr tief in die SOA-Strukturen integriert. So muss etwa die Buchhaltung auf Kundendaten ebenso zugreifen, wie der Vertrieb, der Support oder das Marketing. Die Anforderungen an solche Systeme hinsichtlich Verfügbarkeit und Leistungsfähigkeit wachsen daher kontinuierlich.

Mit den immer größer werdenden Bandbreiten steigen auch die Datenmengen, die die Systeme zu verarbeiten haben. Durch die Einbindung immer weiterer Prozesse nimmt auch die Anzahl der Anfragen zu. Dies kann zu Leistungseinbrüchen führen, die sich in verlängernden Verarbeitungszeiten und Antwortzeiten einzelner Systeme äußern. Probleme mit der Leistung eines einzelnen Systems im Systemverbund können dabei Auswirkungen auf die Leistung und das Funktionieren anderer Verbundsysteme und damit auch anderer Dienste haben – beispielsweise aufgrund von auftretenden Zeitfehlern (timeouts), wie folgendes Beispiel zeigt:

Die Adressdatenbank eines Unternehmens ist im Laufe der Zeit kontinuierlich angewachsen. Die Marketingabteilung plant nun eine Kampagne, die alle Kunden einbeziehen soll. In diesem Zusammenhang wird unter anderem ein Serienbrief mit einem personalisierten Anschreiben erstellt. Dazu wird zur Abfrage der Anreden und Adressen aller Kunden auf die Kundendatenbank zugegriffen. Das die Kundendatenbank beherbergende System kann durch eine derart umfangreiche Anfrage so stark ausgelastet sein, dass Anfragen von anderen Systemen erst mit einem merklichen Zeitverzug bedient werden – oder im ungünstigsten Fall sogar gar nicht.



Funktionsprinzip des passiven Application Logging

Auf einen Blick

Heutige Computersysteme zeichnen sich durch einen hohen Grad an Serviceorientierung aus. Monitoring-Systeme sollten diesem Umstand angepasst werden und Dienste direkt in der Anwendungsschicht protokollieren und nicht nur die Transportschicht betrachten, um damit die Betreibbarkeit von komplexen Systemlandschaften deutlich zu verbessern.

Die Folgen zeigen sich aber erst durch die Unzuverlässigkeit anderer Services, die die Kundendatenbank zum gleichen Zeitpunkt abfragen mussten. Unter Umständen braucht der entsprechende Prozess erheblich länger als sonst oder schlägt ganz fehl.

Häufig ist es sehr schwierig, in komplexen Systemstrukturen aus dem auftretenden Fehlerbild auf die eigentliche Fehlerursache („Root Cause“) zu schließen. Die Folge: Der Prozess der Fehlerbehebung wird sehr zeit- und kostenintensiv. Es ist also nötig, die Datennetze engmaschig zu überwachen.

Viele Probleme werden durch die Protokollierung von Kenndaten (KPI, Key Performance Indicators) der unteren Protokollschichten (Transportschicht) erkannt. Dazu zählen unter anderem die Laufzeiten (RTT, Round Trip Time), Wiederholungen (Retransmits) und die Fragmentierung.

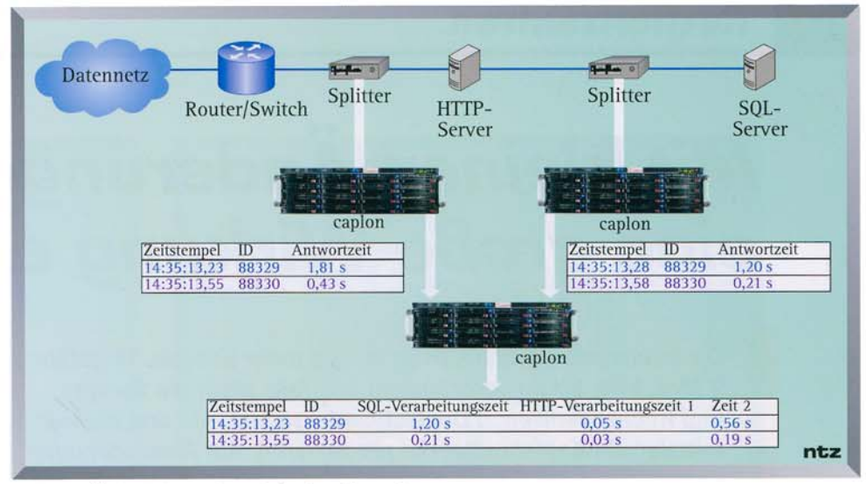
Diese etablierten Methoden stellen sicher, dass die vorhandene Infrastruktur verfügbar und eine Kommunikation zwischen den Komponenten des Systemverbunds grundsätzlich möglich ist. Informationen über den eigentlichen „Nutzverkehr“ werden jedoch nicht geliefert, da dieser auf den höheren Schichten (Anwendungsschicht, Application Layer) übertragen wird. Viele Fehlersituationen oder Leistungsprobleme lassen sich aber nur sinnvoll analysieren, wenn auch die Anwendungsschicht ausgewertet wird.

Grundsätzlich liefern die auf den Systemen laufenden Anwendungen in ihren Logdateien Informationen über den Nutzverkehr und aufgetretene Fehler. Die Auswertung von Logdateien stellt damit ein weiteres Instrument zur Überwachung von Diensten dar. Hierbei gibt es aber grundsätzlich zwei Problemfelder:

- Zum einen werden nicht alle notwendigen Informationen von den Anwendungen vollständig geloggt bzw. nicht alle Fehlersituationen können durch Logmeldungen komplett abgebildet werden.
- Zum anderen wirkt sich ein Erhöhen des Loglevels negativ auf die Leistungsfähigkeit der jeweiligen Systeme aus. Aus Leistungsgründen muss daher oft auf ein ausführliches Mitloggen verzichtet werden. In diesem Fall haben Betreiber solcher Systeme oft gar keine richtige Möglichkeit, an Informationen über den „Nutzverkehr“ zu gelangen, um ggf. eine Fehleranalyse vorzunehmen.

Die Lösung für die oben genannten Probleme heißt: Passives Logging.

Passives Logging bedeutet, dass Logdateien auf Basis des Netzverkehrs ohne Beeinflussung der Systeme und Dienste erstellt werden. Dabei macht man sich zunutze, dass Netzdienste eine genau spezifizierte Schnittstelle benutzen und die Kommunikation über das Netz genau diesem Protokoll folgen muss. Das Mitschneiden des Netzverkehrs eines zentralen Dienstes eines Unternehmens (z. B. der Kundendatenbank) versetzt das Trace-System in die Lage, die komplette Kommunikation zwischen den Systemen auf der Anwendungsschicht zu erfassen. Hieraus können dann alle relevanten in den Anfragen (Requests) bzw. Antworten (Responses) enthaltenen Informationen oder auch Fehlerfälle geloggt werden. Es können aber auch nahezu beliebige Leistungsdaten abgeleitet werden (Antwortzeiten, Ausführungszeiten, Anzahl der Anfragen pro Sekunde usw.), ohne dabei Einfluss auf die eigentliche Funktion und Leistungsfähigkeit des Netzes und des Dienstes zu nehmen. Auf diese Weise kann der Dienst unabhängig von seinem Hersteller lückenlos protokolliert werden. Die Logging-Lösung kann damit genauso mitwachsen, wie der Dienst – selbst wenn der Hersteller einmal gewechselt wird.



Schema: Überwachung einer einfachen Dienstekette

Auf Grundlage der Antwortzeit auf jede einzelne Anfrage kann die Leistungsfähigkeit genauso berechnet werden wie die Fehlerhäufigkeit. Bei unternehmenskritischen Anwendungen kann – wo es sinnvoll ist – auch das Auslösen eines Alarms implementiert werden; entweder als SMS-Benachrichtigung an einen Administrator oder indem ein zentrales System etwa per SNMP-Trap alarmiert wird.

Grundvoraussetzung für die korrekte Funktion eines solchen Trace-Systems ist das verlustfreie Mitschneiden des Netzverkehrs. Um Anwendungsdaten überhaupt nutzen zu können, müssen die je nach Protokoll über mehrere Pakete verteilten Daten wieder korrekt und lückenlos zusammengesetzt werden. Bereits ein einziges verlorenes Paket kann sich zu einem „false positive“ auswirken, also zu einem Alarm führen, der fälschlicherweise ausgelöst wird. In Hochgeschwindigkeitsnetzen stellt dies besondere Anforderungen an die Hardware des Trace-Systems, um den gesamten Netzverkehr komplett (ohne Paketverlust) und vollständig (komplette Nutzlast) erfassen und auf die Festplatte schreiben zu können.

Wird diese Lösung in mehrfacher Ausführung in einem Unternehmen an mehreren Stellen eines verteilten Systems eingesetzt, kann die Kombination der einzelnen Informationen aus den Loggern – etwa durch eine speziell für diese Aufgabenstellung vorgesehene Appliance – ein wertvolles Plus an Informationen bieten. Dies leistet unter anderem bei der Fehlersuche und der Leistungsoptimierung sehr nützliche Dienste, die eine isolierte Betrachtung von Einzelsystemen nicht liefern kann. So können die Durchlaufzeiten der einzelnen Dienste zueinander in Beziehung gesetzt werden. Dadurch lässt sich die aktuelle Leistung des gesamten verteilten Systems beurteilen, die sich nicht zwangsläufig aus der Summe der Einzelsysteme bildet. Zudem kann auf diese Weise vermieden werden, dass die Optimierung eines einzelnen Systems andere Systeme in Mitleidenschaft zieht und dadurch die Gesamtleistung des verteilten Systems sinkt. ■

Martin Nicolay ist Senior Technical Consultant bei consistec Engineering & Consulting in Saarbrücken.
E-Mail: mn_ntz@consistec.de