Motivation für Datenreplikation im Netzwerkmanagement*

Steffen Geschke

Otto-von-Guericke-Universität Magdeburg
Institut für Technische Informationssysteme
Postfach 4120, D–39016 Magdeburg
E-mail: geschke@iti.cs.uni-magdeburg.de
Tel.: ++49-391-67-12816 Fax: ++49-391-67-12020

Es werden Voraussetzungen und Lösungsansätze vorgestellt, um Daten redundant im Netzwerkmanagement zu verwalten. Die Bewertung der Replikationsstrategien erfolgt mit Hilfe der Datencharakteristik und -verteilung, der Konsistenzsicherung sowie anhand der Auswirkungen auf die Managementumgebung. Ein Ausblick auf alternative Anwendungsbereiche der Verteilungskonzepte rundet die Betrachtungen ab.

1 Einleitung

Das Bedürfnis, jederzeit und überall erreichbar zu sein und trotz Mobilität nicht auf vertrauten Netzdienste verzichten zu müssen, hat zu einem starken Anwachsen mobiler Kommunikationssysteme geführt. Die Bedeutung des Netzwerkmanagements steigt mit der Größe der zu verwaltenden Komunikationssysteme.

Die charakteristischen Merkmale eines mobilen Kommunikationssystems erfordern eine Neubewertung des Kommunikationsaufwandes und des Ausfallverhaltens von Kommunikationssystemen. Die Replikation von Daten ist unter diesen neuen Bedingungen bereits als Lösung für primäre Netzdienste in mobilen Systemen akzeptiert [CFLS91]. In globalen Informationsnetzen ist der Einsatz von Cache-Strategien ebenfalls erfolgsversprechend [ABM90], [Dil96]. Diese Untersuchungen belegen, daß Replikation nicht an mobile Kommunikationssysteme gebunden ist.

Die Replikation von Managementdaten ist ein geeignetes Mittel, ein fehlerrobustes Netzwerkmanagement mit minimalen Verzögerungszeiten zu definieren. Zugleich unterstützt die Replikation die Forderung nach Skalierbarkeit der Managementsysteme.

2 Daten- und Netzwerkmanagement

Netzwerkmanagement ist ein sekundärer Netzdienst, welcher die Planung, Organisation und den Betrieb des Rechnerverbundes sowie der primären Netzdienste umfaßt. Drei Standards haben sich auf dem Gebiet des Netzwerkmanagements durchgesetzt [Pra95]:

OSI Systems Management: Das OSI Systems Management definiert Informations-, Organisations-, Funktions- und Kommunikationsaspekte des Netzwerkmanagements auf der Grundlage eines konsequent objektorientierten Ansatzes. Bedingt durch die hohe Komplexität des Modells verfügt das OSI Systems Management jedoch kaum über praktische Bedeutung.

Telecommunication Management Network (TMN): Das TMN ist konzeptionell ein selbständiges Netzwerk, welches seine Managementfunktionen über Schnittstellen zum eigentlichen Kommunikationsnetz wahrnimmt. Analogien zum OSI Systems Management sind besonders anhand des Manager-Agent-Prinzips, des objektorientierten Paradigmas und der Management-Domains nachweisbar. Das TMN wird als aussichtsreicher Anwärter für den Managementstandard öffentlicher Kommunikationsnetze angesehen.

Internet Management: Das Managementmodell im Internet wird durch die dominierende Stellung des Simple Management Network Protocol (SNMP) geprägt. SNMPv2 behebt Probleme des ursprünglichen Standards bezüglich Sicherheit, Leistungsreserven sowie Managementhierarchie und gilt als Quasi-Standard der TCP/IP-Welt.

*Diese Arbeit wird von der DFG im Rahmen des Schwerpunktprogramms "Mobilkommunikation" unter dem Aktenzeichen Sa 465/1 gefördert.

Alle Managementmodelle beruhen auf einer einheitlichen Organisationsform für Managementdaten, die als Management Information Base (MIB) bezeichnet wird. Die MIB stellt ein verteiltes DB-System der Managementdaten dar. Entsprechend ihrer Implementation sind zusätzlich Aspekte objektorienterter bzw. relationaler Datenbanken wirksam.

Das Verteilungskriterium der Managementdaten ist implizit durch den Ort des Datenanfalls vorgegeben. In Einzelfällen wie den Registerdaten mobiler Kommunikationssysteme werden die Daten zusätzlich an ihrem Einsatzort verwaltet. Dieses Verfahren sichert einen schnellen Zugriff auf die Teilnehmerdaten und minimiert die Zeiten für den Verbindungsaufbau.

Die Management Information Base als Echtzeitdatenbank wird im Rahmen Integrierter Management Systeme durch das Manager Repository ergänzt [Rab92]. Das Repository stellt ein klassisches Datenbanksystem dar, welches Managementapplikationen mit historischen Managementdaten für Organisation und Planung unterstützt. Im Bestand des Repositories sind zudem ergänzende MIB-Daten und Duplikate der MIB zu finden.

3 Replikationsstrategien

Die Replikation von Managementdaten unterstützt ein fehlerrobustes Netzwerkmanagement durch eine Datenhaltungsstrategie, welche

1. eine hohe Verfügbarkeit der Daten sicherstellt,
2. die Zugriffszeiten auf die Managementdaten reduziert.

Redundanzen sind verbunden mit Overhead und Konsistenzproblemen. Aus diesem Grund ist eine Bewertung des Aufwandes für den Erfolg der Replikationsmechanismen unerläßlich.

Welche Daten können oder müssen dupliziert werden? Eine Einschränkung des duplizierten Datenvolumens ist notwendig, da für die Konsistenzsicherung Ressourcen verbraucht werden, die den primären Netzdiensten zustehen. Eine konkrete Auswahl der Replikationsdaten wird anhand des Replikationsziels getroffen. Sich ändernde Echtzeitdaten und abgeleitete Managementdaten sind generell von der Replikation auszuschließen.

In einer mobilen Kommunikationsumgebung erfolgt die Suche nach geeigneten Knoten zur Aufnahme der Duplikate innerhalb des Festnetzes, wo Speicherressourcen und Bandbreite im notwendigen Umfang bereitstehen. Generell lassen sich drei Lokalitätskategorien als Verteilungsorte für die Duplikate charakterisieren:

Unmittelbare Nachbarschaft: Die Knoten, welche für die Aufnahme von Duplikaten ausgewählt wurden, sind direkt mit dem Originalknoten verbunden.

Mittlere Distanz: Die lokale Ausweitung ist durch eine physische oder logische Abgrenzung sichergestellt. Als mittlere Distanz sind beispielsweise Zelle, Cluster oder Domain geeignet.

Global: Die Verteilung der Datenduplikate erfolgt ohne Einschränkungen innerhalb der gesamten Infrastruktur des Kommunikationssystems.

Sicherheit in verteilten Systemen bedeutet auch eine Verteilung der Sicherheitsverantwortung. Diese muß im Rahmen des Datenbestandes durch Kontrolle der Zugriffsrechte und durch ein robustes Update-Protokoll gewährleistet sein.

3.1 Replikation für hohe Verfügbarkeit

Liegt in einem Informationssystem eine Anfrage für einen nicht verfügbaren Informationsträger vor, wird ein Stellvertreter gesucht, welcher diese Anfrage im vollen Umfang befriedigen kann. Die Fähigkeit als Stellverteter zu fungieren wird erworben, indem der Stellvertreter das Wissen des angefragten Informationsträgers teilt. Dieses Wissen wird ihm durch Replikation zugänglich.

Datenreplikation kann somit intuitiv als ein natürliches Instrument zum Zweck einer hohen Verfügbarkeit von Managementdaten angesehen werden.

Die Fehlerrobustheit fester Kommunikationsnetze ist zwischenzeitlich stark gewachsen. Im Gegensatz zu den konventionellen Netzen müssen mobile Kommunikationssysteme weiterhin als sehr fehleranfällig betrachtet werden [IB94]. Zusätzlich erfordert die Mobilkommunikation eine erweiterte Interpretation des Verfügbarkeitsbegriffs:

Uneingeschränkte Verfügbarkeit: Die Ressource eines Kommunikationssystems gilt als verfügbar, wenn der Zugriff auf diese Ressource ohne Einschränkungen möglich ist.

Eingeschränkte Verfügbarkeit: Die Verfügbarkeit einer Ressource kann durch äußere Einflüsse beeinträchtigt sein. Diese eingeschränkte Verfügbarkeit wird beispielsweise durch eine herabgesetzte Bandbreite des Kommunikationskanals deutlich.

Unkontrollierte Nichtverfügbarkeit: Fällt die Ressource aus, handelt es sich um Nichtverfügbarkeit. Die Nichtverfügbarkeit im Fehlerfall erfolgt unplanmäßig.

Kontrollierte Nichtverfügbarkeit: Die Mobilkommunikation ergänzt den Status der Nichtverfügbarkeit durch eine koordinierbare Form. Sie wird insbesondere durch eng bemessene Energiereserven und durch lokale Besonderheiten hervorgerufen. Die kontrollierte Nichtverfügbarkeit kann durch geeignete Maßnahmen beim Verlassen des Kommunikationsverbundes begleitet werden. Systemrelevante Daten werden auf Nachbarressourcen ausgelagert.

Bei unkontrollierter Nichtverfügbarkeit sind vorbereitende Maßnahmen zu treffen, die einem potentiellen Ausfall der Ressource gewachsen sind. Das Auslagern wichtiger Managementdaten muß vom System geplant und periodisch wiederholt werden. Der Kommunikationsaufwand für das zyklische Aktualisieren der Duplikate ist hoch und erfordert eine starke Einschränkung jener Daten, die mit Hilfe der Replikation dauerhaft zur Verfügung gestellt werden sollen. Eng bemessene Speicherreserven erfordern eine weitere Einschränkung des Datenvolumens. In die Replikation sind die Daten einzubeziehen, welche Einfluß auf die globale Funktionsfähigkeit des Managementsystems besitzen, zum Beispiel Leistungswerte für Lastlenkungsmechanismen, Teilnehmer- und Gesprächsabrechnungsdaten.

Die Duplikate der Managementdaten sind in unmittelbarer Nachbarschaft vom Originalknoten zu verwalten. Wird ein globaler Fehler wie der Ausfall einer lokalen Routingfunktion erwartet, muß auch die Replikation in mittlerer Distanz in Erwägung gezogen werden.

Ein einzelner Nachbarknoten als Träger der Duplikate bietet sich aufgrund einer einfachen Handhabung der Datenauslagerung an. Nachteilig wirkt sich die hohe Belastung einer einzelnen Ressource bezüglich der Speicherkapazität und des stark genutzten Kommunikationskanals aus. Durch eine Verteilung auf mehrere Duplikatsknoten können diese Nachteile auf Kosten einer komplizierten Cluster-Bildung reduziert werden.

Das Aktualisieren der Duplikate kann vom Original oder vom Duplikat angefordert werden. Dabei werden die klassischen Reportmechanismen des Netzwerkmanagements verwandt:

Polling: Das Duplikat fordert die Werte des Originals durch periodisches Abfragen an. Ändert sich die Werteausprägung des Originals unmittelbar nach einem Polling, treten zeitlich Inkonsistenzen auf.

Trap: Das Original überwacht Werteänderungen und meldet diese an das Duplikat. Für die Kontrollfunktion werden zusätzliche Ressourcen benötigt. Zudem kann bei frequentierten Werteänderungen das Kommunikationssystem durch Aktualisierungsnachrichten überlastet werden.

Trap-gesteuertes Polling: Das Duplikat erfragt zyklisch die Werte des Originals. Das Original hat die Möglichkeit, die Abfrageperiode zu variieren. Die Vorteile von Polling und Trap werden vereint.

Ist ein Netzwerkknoten nach zwischenzeitlicher Nichtverfügbarkeit in den Rechnerverbund zurückgekehrt, werden seine ausgelagerten Daten wiederhergestellt. Zuvor wird geprüft, ob die Duplikate dem Systemstatus des Knotens entsprechen.

Durch Terminalmobilität, welche es einem mobilen Teilnehmer erlaubt, die Position innerhalb des Kommunikationssystems zu wechseln, kann der Nachbarschaftsverbund zwischen Original und Duplikat verlorengehen. Zur Minimierung des Kommunikationsaufwandes sind in diesem Fall die Duplikatsknoten nach dem Wiederherstellen der Originaldaten freizugegeben und Replikationsknoten in der neuen Nachbarschaft zu ermitteln.

3.2 Replikation zur Minimierung der Zugriffszeiten

Mit zunehmendem Wachstum von Kommunikationssystemen steigen auch die Zugriffszeiten auf die Netzdienste und die zugehörigen Daten. Mobile Teilnehmer eines Kommunikationsnetzes sind zusätzlich durch eine minimale Bandbreite und eine relativ hohe Ausfallwahrscheinlichkeit eingeschränkt.

Die Verbindungswege und somit die Übertragungszeiten lassen sich durch das gezielte Einfügen von Redundanzen minimieren. Das Cache-Prinzip hat sich bereits in der Mikroprozessortechnik bewährt. Im

Netzwerkmanagement ist diese Strategie geeignet, die Zugriffszeiten auf Managementdaten zu minimieren und die Daten zusätzlich am Ort ihrer Verarbeitung zu verwalten. Dieser Ansatz wird bereits bei Registerdaten mobiler Kommunikationssysteme realisiert. Jedes Organisations- und Verwaltungszentrum verfügt über eine Kopie der Teilnehmerdaten. Der Zugriff auf den lokalen Datenbestand verringert die Verbindungsverzögerung, welche durch die notwendige Authentifizierung entsteht.

Der Ansatz der Datenreplikation kann einheitlich gestaltet und auf andere Managementdaten ausgedehnt werden. Das Spektrum der Daten, welche in die Replikation einbezogen werden, ist freizügig zu bemessen. Dedizierte Replikationsknoten können mit ausreichend Speicherreserven versehen werden. Der Kommunikationsaufwand für den Zugriff und die gleichzeitige Aktualisierung der Duplikate beträgt maximal $2K$ (K = Kosten der direkten Kommunikation). Dem ensteheden Mehraufwand sind jedoch die minimalen Kosten für erfolgreiche Sekundärzugriffe auf die Duplikate gegenüberzustellen.

Die Konsistenzsicherung erfordert die Einführung einer Latenzzeit. Die Duplikate werden zum Zeitpunkt T mit einem Zeitstempel versehen, der Auskunft über die letzte Aktualisierung der Daten gibt. Innerhalb des Zeitintervalls $< T, T + Latenzzeit >$ werden die Duplikate als vertrauenswürdig betrachtet. Außerhalb dieses Gültigkeitsintervalls müssen die Duplikate als überholt eingestuft werden. Bei Nichtverfügbarkeit des Originals kann in Einzelfällen auch mit Hilfe von überholten Daten operiert werden. Die Entscheidung darüber wird an die Managementapplikation delegiert, da sie über die Semantik der Daten informiert ist.

Die Einführung von Latenzzeiten ruft zusätzlichen Overhead in der duplizierte MIB hervor. Die direkte Zuordnung von Latenzzeiten zu Einzeldaten erscheint uneffektiv. Daten lassen sich zu Datenkategorien gruppieren, welchen eine repräsentative Latenzzeit zugewiesen wird. Im einfachsten Szenario verfügen alle Daten eines Knotens über die gleiche Latenzzeit.

Das Ersetzen konkreter Zeitintervalle durch Aufzählungstypen führt zu einer weiteren Minimierung des Overheads für die Speicherung der Latenzzeit.

Die Aktualisierung der Duplikate erfolgt auf Anfrage, sofern die Duplikate den Status “überholt” aufweisen. Für zeitkritische Anfragen kann zusätzlich ein Update-Intervall eingeführt werden, welches das Ende des Gültigkeitsintervalls überdeckt. Anfragen innerhalb des Update-Intervalls können durch die Duplikate befriedigt werden. Zugleich wird eine Aktualisierung der Duplikate angefordert.

Das Duplikat selbst kann eine Aktualisierung bei abklingendem Gültigkeitsintervall vornehmen. Dazu werden jedoch zusätzliche Ressourcen für die Überwachung der Duplikate benötigt und eine Aktualisierung erfolgt unabhängig von der bestätigten Annahme, daß diese Daten vom Managementprozeß frequentiert benötigt werden.

Die Verteilung der Replikationsknoten erfolgt global im Kommunikationssystem. Die erforderlichen Strategien stellen ein komplexes algorithmisches Problem dar, welches einer weitergehenden Analyse bedarf.

3.3 Globale Verteilungsstrategien

Eine abstrakte Problembeschreibung lautet: Zu einem gegebenen Graphen X ist ein Subgraph Y zu ermitteln, so daß $\forall x \in V(X)$ gilt: $d(x, \{y|x_i\}) \leq Kostenlimit$.

Die Kostenfunktion $d(x, y)$ liefert im einfachsten Fall die Anzahl der Hops des Verbindungspfades von x nach y. Leistungsorientierte Funktionen wie Durchsatz oder Ausfallwahrscheinlichkeit spiegeln praktische Erfordernisse auf Kosten eines höheren Verarbeitungsaufwandes gezielter wider.

Eine Analyse aller kombinatorischen Verteilungsmöglichkeiten scheitert am Verarbeitungsaufwand von $\Theta(n^n)$. Aus diesem Grund wird ein heuristischer Algorithmus vorgeschlagen, welcher eine akzeptable Lösungsgenauigkeit bei minimalen Verarbeitungsaufwand und Speicherbedarf liefert. Die Verteilung der Replikationsknoten erfolgt iterativ:

1. Erstellen einer Kostenmatrix
2. Für alle Replikationsknoten x_i:
   (a) Positionieren des Replikationsknotens x_i
   (b) Validieren der Position aller Replikationsknoten $\leq x_i$

Ausgangspunkt dieses Verfahrens ist eine Kostenmatrix, welche für alle Verbindungen im Kommunikationsverbund die Werte entsprechend der genutzten Kostenfunktion bereitstellt. Zum Einsatz kommen Algorithmen der Klasse “All Pairs Shortest Path”. Das Erstellen der Kostenmatrix geht bereits mit einem Verarbeitungsaufwand von mindestens $\Theta(V(X) + log(X))$ in die Berechnungskomplexität ein.

Das Positionieren des Replikationsknotens in Schritt 2 verringert die Kosten für die ausgewählte Verbindung, welche mit einem maximalen Wert in der Kostenfunktion vertreten ist. Durch Zeilen- und Spaltenvergleich wird eine optimale Verkürzung entlang dieses Verbindungspfades ausgewählt.

Eine Validation der bisher positionierten Replikationsknoten in Schritt 3 ist notwendig, da das Einfügen eines weiteren Replikationsknotens Auswirkungen auf jene Verbindungen hat, welche zum Einsatz voriger Replikationsknoten führte. Die Positionsvalidation erfolgt mit Hilfe von filterorientierten Zeilen- bzw. Spaltenvergleich.

Um ein starkes Anwachsen des Verarbeitungsaufwandes zu unterbinden, kann das Überprüfen der Positionen nur in einem abgegrenzten, lokalen Rahmen erfolgen.

Eine Minimierung des Speicherplatzbedarfes kann durch ein Zusammenfassen der Knoten einer Organisationseinheit, zum Beispiel einer Domain, erzielt werden. Ein repräsentativer Knoten dieses Verbundes wird mit den maximalen Verbindungskosten der Gruppe bewertet.

Die vorgeschlagene Verteilungsstrategie operiert zentral auf der Gesamtheit aller Kostenfaktoren. Das Berechnungsergebnis kann als Ausgangspunkt für dynamische und dezentrale Mechanismen dienen.

4 Zusammenfassung und Ausblick

Datenreplikation ist ein einfaches intuitives Instrument, um die Verfügbarkeit von Daten im Netzwerkmanagement zu erhöhen. Die erforderlichen Mechanismen zur Konsistenzsicherung erfordern einen Kommunikationsaufwand, welcher dem Netzwerkmanagement in konventionellen Systemen nicht zugebilligt wird. Durch Ausnutzung von Phasen unterschiedlicher Systemauslastung kann die erforderliche Bandbreite zur Aktualisierung der Duplikate bereitgestellt werden.

Die Anwendung der vorgeschlagenen Replikationen und Verteilungsstrategien ist für feste und mobile Netze zu untersuchen. Die engen Randbedingungen bezüglich Kommunikations- und Speicherressourcen ermöglichen eine Ausweitung der Ergebnisse auf primäre Netzdienste.

Werden konkrete Zugriffsmechanismen vernachlässigt, ist der Einsatz der Verteilungsstrategien in globalen Informationsnetzen möglich. Im Internet kann ein solches Informationsnetz aus Archie- und FTP-Servern sowie Proxy-Servern für World Wide Web und andere Netzdienste bestehen. Durch eine gezielte Verteilung der Server lassen sich die Zugriffszeiten auf die Netzdienste für alle Teilnehmer minimieren. Die Planung einer solchen Informationsstruktur kann durch Anycast-Adressen vervollständigt werden. Dieser in IPv6 neu eingeführte Adreßtyp, wählt den Server mit den kürzesten Verbindungskosten aus.

Literatur

[ABM90] R. Alonos, D. Barbara, and H. G. Molina. Data caching issues in an information retrieval system. In *ACM TODS*, pages pp. 359–384, Sept. 1990.

[CFLS91] M. J. Carey, M. J. Franklin, M. Livny, and E.J. Shekita. Data Caching tradeoffs in client-server DBMS architectures. In *ACM SIGMOD '91*, May 1991.

[Dil96] J. Dilley. Building Large, Widely Distributed Systems. In *ICDP'96: IIndustrial Stream*, pages pp. 13–19, 1996.

[IB94] T. Imielinski and B. R. Badrinath. Wireless Mobile Computing, Challanges in Data Management. Technical Report DCS-TR-296, Department of Computer Sciences, Rutgers University and WINLAB TR-49, 1994.

[Pra95] A. Pras. *Network Management Architecture*. PhD thesis, University of Twente, Entschede, 1995.

[Rab92] S. Rabi. Integrated Network Management: Technologies and Implementation Experiences. In *Infocom '92*, 1992.