Universität Heidelberg
Fachhochschule Heilbronn
Medizinische Informatik

Netzwerkmanagement in Nah- und Weitverkehrsnetzen

Henry Schäfer

Referent: Prof. Dr. Gerhard Peter
Korreferent: Prof. Lutz Müller

Vorwort

Motivation

Persönlicher Werdegang

Nach meinem Abitur begann ich im Frühjahr 1994 das Studium der Medizinischen Informatik, nachdem ich meinen Zivildienst in der Universitätsklinik Heidelberg absolviert hatte. Auf die Idee, Medizinischen Informatik zu studieren, hatte mich – wie viele andere – die Lektüre von *Studien- und Berufswahl* gebracht. Der Studiengang erschien mir als optimale Kombination für meine Interessen auf medizinischem und informationstechnischem Gebiet. Medizinischen Informatik

Meine ersten Berührungspunkte mit moderner Hardware hatte ich Mitte der achtziger Jahre, als ich einen gebrauchten VC-20 geschenkt bekam. Für erste Schritte, nicht zuletzt bei der Programmierung in BASIC, war dieser Homecomputer hilfreich. Allerdings vergingen von da an noch einige Jahre, bis ich einen „richtigen" Personal Computer besaß. Erst viel später, nämlich während meines Studiums, wurde ich mit Netzwerken konfrontiert.

Durch meine Arbeit im Linux-Pool der Hochschule konnte ich nicht nur Erfahrungen im UNIX-Umfeld sammeln, sondern ich bekam einen umfassenden Einblick in die Technik von Netzwerken. Durch meine Tätigkeiten neben meinem Studium im administrativen Bereich konnte ich später sowohl im Softwarelabor der Hochschule als auch bei dem Internetprovider und Systemhaus Steinle und Klöpfer die gesammelten Erfahrungen einsetzen und vertiefen. Softwarelabor Steinle und Klöpfer

Insgesamt verlief das Studium an der Fachhochschule Heilbronn und der Universität Heidelberg meinen Vorstellungen entsprechend und weites gehend planmäßig.

Die Entscheidung, die Diplomarbeit in einem Unternehmen zu schreiben, entstand aus mehreren Ideen. Zum einen bot es sich an, mit den bisherigen Erfahrungen auch inhaltlich neue Wege zu beschreiten. Zum anderen war es an der Zeit, einen neuen Schaffensplatz zu finden und diesen in einem befristeten Zeitrahmen kennenzulernen. Bei der Firma UUNET Deutschland habe ich diesen Schaffensplatz gefunden.

Danksagung

Viele Personen haben direkt oder indirekt dazu beigetragen, daß Ende September 1999 eine lesbare Version dieser Arbeit abgabefertig war.

So wie es im gesamten Network Engineering Team Standard ist, standen mir viele Personen der Firma UUNET bei Fragen und Problemen zur Seite. Besonders das Network Engineering Team ist immer wieder ein Quell der reinen Lebensfreude.

Prof. Dr. Gerhard Peter als meinem Referenten gebührt Dank für die problemlose Abwicklung und sein Verständnis bei der Unternehmung „Diplomarbeit“.

Für das kurzfristig anberaumte Korrekturlesen möchte ich mich ganz herzlich bei Martina Rotar, Daniel Seiler, Susanne Schurg und Henner Zeller bedanken. Mit diesen Freunden ist jeder Schritt ein leichter, ohne sie enthielte diese Arbeit weitaus mehr Fehler und Spaghettisätze, als sowieso darin sind.

Ebenso danke Tobias Häcker als Richtungsweisende und Freund, ohne ihn wäre mancher Schritt anders verlaufen.

Diese Arbeit wurde mit LaTeX unter Verwendung der teTeX-Distribution von Thomas Esser auf Linux- und Sun-Microsystems-Workstations erstellt.

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Kapitel 1

Einleitung

Eine kurze allgemeine Einführung in die Thematik. Persönlicher Hintergrund, Studium der Medizinischen Informatik und beruflicher Werdegang. Ein Überblick über die weitere Dokumentation und ihre Kapitelinhalte.

1.1 Fragestellung

Ein Titel, in dem das Wort „Management“ enthalten ist, erinnert in erster Linie an Begriffe wie *Leiten* und *Führen* im betriebswirtschaftlichen Sinn. Dies ist von der Wahrheit im technischen Umfeld gar nicht so weit entfernt. Im Sprachgebrauch gibt zwar viele Worte die mit Management zusammenhängen – Manager, Finanzmanagement, Managementstrategie, Managementsystem – jedoch auf den ersten Blick aus sehr unterschiedlichen Bereichen stammen.

Managen bedeutet jedoch so viel wie *leiten*, *unternehmen*, ja sogar *zustande bringen* [Dude96]. Das läßt sich sehr gut in das Management von Netzwerken einbringen. Die Sicht auf ein Netzwerk hängt sehr stark von den Möglichkeiten ab, ob und wie ein Netz dargestellt wird. Nur eine vollständige Sicht läßt alle Aspekte des Netzes – zum Beispiel Fehlermöglichkeiten, Konfigurationen, Monitoring – offen und macht es überschaubar. Die Frage nach vollständiger Überschaubarkeit stellt sich immer dann, wenn ein hohes Maß an Qualität erreicht werden soll. Doch die Gradwanderung von niedrigem Aufwand und Kosten auf der einen Seite und hoher Qualität auf der anderen, ist nicht einfach.

Netzwerkmanagement ist nur ein Oberbegriff und hat viele Teilaspekte, die sowohl in großen als auch kleinen Netzen Einsatz finden können. Einige Stichpunkte seien hier stellvertretend aufgeführt:

- *Statistiken*: Sie werden an vielen Stellen benötigt, sei es zur Darstellung von Auslastungen oder zur Berechnung von zukünftigen Bedürfnissen.

- *Sicherheitsmanagement*: An Orten, an denen viele Personen unterschiedlichste Rechte auf bestimmte Ressourcen haben sollen, stellt ein zentrales Management nicht nur eine Arbeitserleichterung dar, sondern auch die Möglichkeit einer zentralen Überwachungsinstanz.
- *Konfigurationsmanagement*: In ständig wachsenden Netzwerken ist es oft wichtig, Updates und globale Änderungen von Konfigurationen, an vielen Geräten gleichzeitig vorzunehmen.
- *Accountingmanagement*: Abrechnungen sind für jedes wirtschaftlich arbeitende Unternehmen von essentieller Bedeutung. Ohne Accountingdaten sind diese nicht möglich.

Welche Teile, nicht nur aus den genannten Beispielen, für ein jeweiliges Netzwerkmanagement-Gesamtkonzept notwenig sind, ist von Fall zu Fall verschieden. Die Anforderungen müssen selbst definiert werden. Wie in späteren Kapiteln beschrieben wird, hat OSI den umfassendsten Standard definiert, an dem sich bis heute jeder orientiert.

Auf Grund der wachsenden Zahl an Ressourcen, nicht zuletzt im Internet, steigt auch der Bedarf an Netzwerküberwachung und -management. Aber auch die Qualitätssicherung macht ein Managementsystem unabdingbar.

1.2 Vorgehensweise

Es zeigt sich, daß die im OSI-Management beschriebene Definition die umfangreichste ist. Die International Standard Organisation hat damit ein grundlegendes Paket geschaffen, dessen Beschreibung zum Verständnis der vielfältigen Implementierungen beiträgt.

Es ist notwendig einen Überblick über die Funktionsweise der im Internet verwendeten Protokolle zu bekommen, da die beschriebenen Lösungen sehr eng mit den zur Datenübertragung genutzten Techniken verknüpft sind.

Es werden die beiden Themen Netzwerkmanagement und Internet zunächst getrennt behandelt, um die Strategien und Funktionen im Folgenden verbinden zu können.

1.3 Übersicht

In Kapitel 1 *Übersicht* wird das Thema eingeführt und der Leser mit der Problematik vertraut gemacht.

Das Kapitel 2 *Grundlagen* gibt einen Abriß über Netzwerkmanagement, dessen Modelle und Sichtweisen. Anhand der Internetprotokollwelt wird das Prinzip und die Zusammenhänge erklärt und verdeutlicht.

Kapitel 3 *Netzwerkmanagement* beschreibt allgemeine Anforderungen, die wichtigsten Funktionen im Internet-Management und stellt diese der generischen OSI-Management Definition gegenüber.

Im Kapitel 4 *Auslastungsanalyse* werden Verfahren der Analyse von Zeitreihen beschrieben. Bestehende Lösungsansätze bei Vorhersageverfahren werden vorgestellt und verglichen. Die Umsetzung bei der Firma UUNET Deutschland wird anhand der theoritischen Erkenntnisse im letzten Abschnitt erläutert.

Das Kapitel 5 *Zusammenfassung, Perspektiven* rundet die Arbeit ab.

Kapitel 2

Grundlagen

Überblick über die Entwicklung des Internets, technische Grundlagen, Protokolle und Dienste, SNMP.

Das Kapitel 2 gibt einen Einblick in die Grundlagen des Netzwerkmanagements. Dabei werden Begriffe wie ISO/OSI, TCP/IP, SNMP und andere sowohl erwähnt als auch beschrieben, um im weiteren Verlauf darauf aufbauend das Hauptthema Netzwerkmanagement aufzuarbeiten.
In Kapitel 2.2 wird der Zusammenhang zwischen ISO/OSI-Referenzmodell und den praktizierten Schichten des TCP/IP-Protokoll erklärt. Das folgende Kapitel 2.1 setzt sich allgemein mit dem Netzwerkmanagement auseinander und stellt seinen Bezug zum Netz der Netze dar.

2.1 Netzwerkmanagement und das Netz

In den letzten Jahrzehnten wurden nicht nur immer mehr Rechnersysteme eingesetzt, sondern es stieg auch der Bedarf an Vernetzung (siehe Abbildung 2.1). Allein der Anstieg der am Internet angeschlossenen Rechner stellt eine Herausforderung dar. Dabei ist das Problem der Integration der heterogenen Systeme zu bewältigen, dessen Lösung nur eine Standardisierung sein konnte. Standardisierung steht allerdings auch für Einschränkung.

Eine der großen Leistungen der *International Organisation of Standardization* war die Beschreibung einer prinzipiellen Architektur von Telekommunikationssystemen. Beginnend im Jahre 1977 wurden 1984 die Ergebnisse als das heute bekannte ISO/OSI-Basisreferenzmodell veröffentlicht.

Die dem Modell zugrundeliegende Idee, die Aufteilung der komplexen Gesamtaufgabe in überschaubare Teilaufgaben, führte zu einer hierarchischen Gliederung in sieben Schichten. Die Zahl *sieben* ergab sich aus dem Kompromiß, einerseits möglichst wenige Schichten einzusetzen, um die Übersichtlichkeit zu bewahren, andererseits mit nicht zu wenigen Schichten die Aufgaben überschaubar zu spezifizieren. Weitere Ausführungen dazu folgen im Kapitel 2.2.

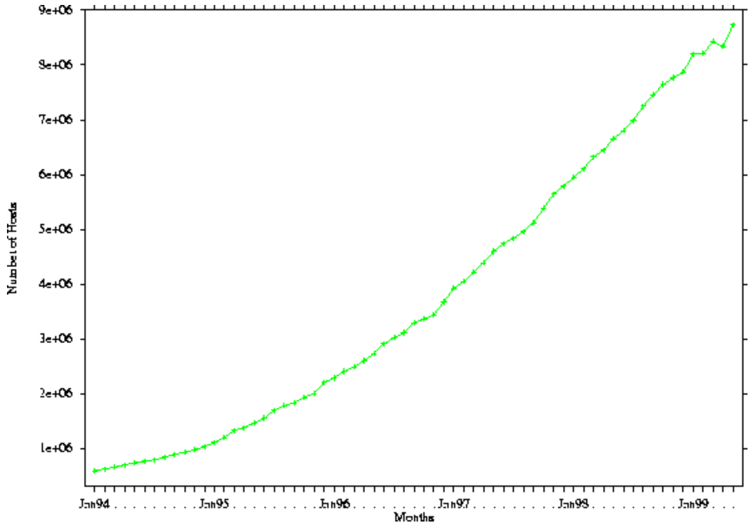


Abbildung 2.1: Übersicht gezählter Hosts in Europa [Ripe99]

Um den Inhalt des Netzwerkmanagements im allgemeinen zu verstehen, sollen OSI Systemmanagementkonzepte und OSI Management Information Base vorgestellt werden. Dazu stand hauptsächlich [Stal93] zur Verfügung.

2.1.1 Systemmanagementkonzepte

2.1.1.1 OSI Systems-Management Standard

Unter allen Standards die von OSI entworfen wurden, sind die OSI Systems-Management Standards die am umfangreichsten und komplexesten. Der erste Standard, der sich auf Netzwerkmanagement bezog, beschreibt das Managementarchitekturmodell und wurde von ISO unter dem Namen *OSI Management Framework and Overview* herausgegeben. Die folgenden Standardisierungen bezüglich des Managements wurden zusammen von ISO und dem internationalen Organ aller Postgesellschaften ITU[1] herausgegeben. Die Standards gliedern sich in fünf Kategorien:

- **OSI Management Framework and Overview**: ISO 7498-0 gibt eine allgemeine Einführung in Managementkonzepte, ISO 10040 gibt einen Überblick über die restlichen Dokumente.
- **CMIS/CMIP**: ISO 9595/9596 definiert den Common Management Information Service (CMIS), welcher Dienste für Managementanwendungen anbietet, sowie das zugehörige Common Management Information Protokoll, zuständig für Informationsaustausch des CMIS.

[1]International Telecommunication Union, früher CCITT

- **Systems Management Functions**: ISO 10164-1/-14 definiert spezielle Funktionen, die vom OSI Systemmanagement ausgeführt werden können.
- **Management Information Modell**: ISO 10165-1/-6 beschreibt die Management Information Base MIB, die alle OSI Objekte bezüglich des Netzwerkmanagements repräsentiert.
- **Layer Management**: definiert Managementinformationen, Dienste und Funktionen bezogen auf bestimmte OSI Schichten.

Die Zusammenhänge zwischen den fünf Kategorien stellt die Abbildung 2.2 dar, sie bietet einen Überblick über das Zusammenwirken der Standards.

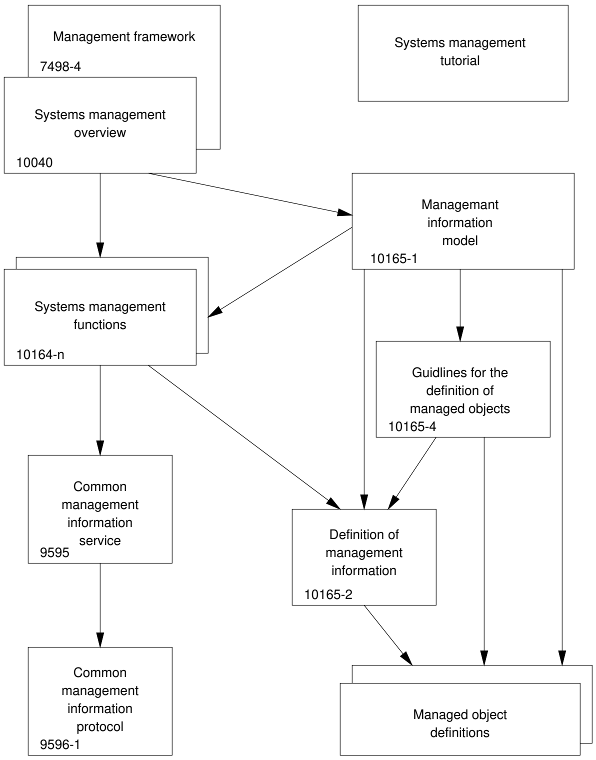


Abbildung 2.2: Beziehungen zwischen OSI Managementstandards [Stal93]

Das wichtigste und vielleicht auch schwierigste am OSI Modell ist, die Terminologie zu verstehen, also zu erkennen, wie die logische mit der physischen Ebene einhergeht. Zum besseren Verständnis der Standards sind in der Abbildung 2.3 die wichtigsten Begriffe enthalten. Der Begriff *systems management* wird im OSI-Kontext verwendet und stellt das Äquivalent zum Begriff *Netzwerkmanagement* in nicht-OSI-Literatur dar. Weiterhin basiert *OSI systems management* auf dem "hauseigenen" Schichtenmodell, die Darstellung der Standards ist also davon abhängig. Und *OSI systems*

systems management

Netzwerkmanagement

management ist beschränkt auf das Management von OSI-Komponenten in einer OSI-Umgebung[2].

2.1.1.2 Architekturmodell

Das OSI Architekturmodell ist in Abbildung 2.3 dargestellt. Es besteht aus 4 Schlüsselelementen:

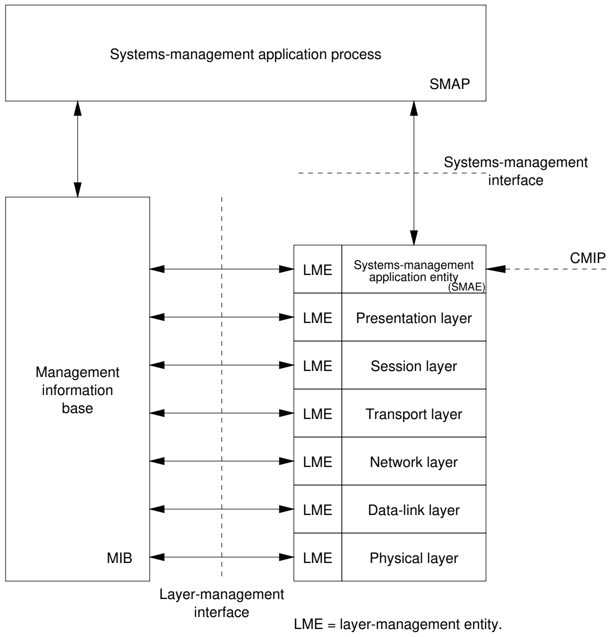


Abbildung 2.3: OSI Architektur Modell [Stal93]

- **SMAP – Systems-Management Application Process** ist die Software innerhalb des Systems, welche die Systemmanagementfunktionen ausführt. SMAP hat Zugriff auf verschiedene Systemparameter und kann deshalb alle Funktionen steuern und koordinieren, ebenso beim Austausch mit SMAPs anderer Systeme. Eine detailliertere Beschreibung des OSI *Application layer* kann [Stal93] entnommen werden.
- **SMAE – Systems-Management Application Entity** ist ein Dienst der Anwendungsschicht und ist verantwortlich für den Austausch von Managementinformationen, besonders mit Nachbarinstanzen wie auch mit zentralen Instanzen. Das standardisierte Anwendungsprotokoll *CMIP* wird zu diesem Zweck verwendet.

[2][Stal93]: not application programs, CD players, toasters, etc.

- **LME – Layer-management entity** stellt als Schlüsselelement, das in jeder Schicht der OSI-Architektur vorhanden ist, für die jeweilige Schicht spezifische Netzwerkfunktionen bereit.
- **MIB – Management Information Base** enthält eine Sammlung von Strukturbeschreibungen der Informationen eines jeden Knotens des Netzwerks.

Parallelen zu diesen Punkten sind im TCP/IP-Protokoll zu erkennen. Beispielsweise repräsentiert SMAE die Applikationsebene. Bei OSI wird beschrieben, daß SMAE aus mehreren *Application Service Elements* (ASE) besteht. Neben *Association Control Service Elements* (ACSE) und *Remote Operation Service Elements* (ROSE) sind noch zwei weitere Elemente enthalten, die besonders wichtig für das Netzwerkmanagement sind: *Common management information service elements* (CMISE) und *System Management Application Service Elements* (SMASE). — Zum Verständnis notwendige Details sind weiter unten beschrieben.

SMAE kann sowohl als Manager oder als Agent agieren. In der Rolle als Manager gibt der *Systems Management Application Process* (SMAP) Anweisungen, Operationen und Anforderungen an die darunter liegenden Agenten aus. Diese leiten wiederum Managementinformationen weiter und sind für die Rückantworten sowie Meldungen verantwortlich. Das Speichern der Daten ist im Gegensatz zum nicht standardisiert, lediglich der Datenaustausch nicht standardisiert.

Die Hauptaufgabe von SMASE besteht also in der Bereitstellung von Diensten, welche wiederum Netzwerk(management)funktionen repräsentieren wie man sie bereits kennt. Es werden fünf Bereiche unterschieden:

- Fehlermanagement (*Faultmanagement*)
- Leistungsmanagement (*Performance Management*)
- Konfigurationsmanagement (*Configuration Management*)
- Abrechnungsmanagement (*Accounting Management*)
- Sicherheitsmanagement (*Security Management*)

Die Funktionen dieser fünf Bereiche werden in Kapitel 3.1.2 und 3.3 beschrieben.

2.1.1.3 System Managementfunktionen

Die zu den System Managementfunktionen gehörenden Bereiche, die im OSI Framework definiert sind, wurden nicht als solche standardisiert. Sie nehmen vielmehr Bezug auf die im vorigen Kapitel erwähnten fünf Bereiche. Jeder dieser Funktionsbereiche der System Managementfunktionen nutzen einen oder mehrere dieser fünf Bereiche. Dazu gehören:

- Object Management
- State Management
- Relationship Management
- Alarm Reporting
- Event-report Management
- Log Control
- Security-alarm Reporting
- Security-audit Trail
- Access Control
- Accounting Meter
- Workload Monitoring
- Test Management
- Summarization

2.1.2 Management Information Base

Die Informationen über Ressourcen und zu verwaltende Elemente sind in einer Datenbank abgelegt, welche die Grundlage eines Managementsystems darstellt. In OSI wird die *Management Information Base* (MIB) verwendet und ist objektorientiert aufgebaut.

Die Datenbankelemente werden als *Managed Objects* (MO) bezeichnet. Damit stellt die MIB eine Sammlung von Managementobjekten dar. Diese können Software- wie auch Hardwareressourcen repräsentieren. Eine Systemressource kann durch ein oder mehrere Managementobjekte dargestellt werden; aber auch ein Managementobjekt kann mehrere Ressourcen umfassen. Ressourcen, die nicht in der MIB enthalten sind, können vom OSI-Management nicht verwaltet werden, da sie nicht zugänglich sind.

2.1.2.1 Management Informationsmodell

Der Standard ISO 10165-1, siehe Abbildung 2.2, liefert in vier Punkten einen Beschreibungsrahmen für Managementinformationen, deren erster Punkt näher beschrieben werden soll:

- Modell für Managementobjekte (MO) und deren Attribute
- Prinzipien der Benennung der Managementobjekte

- Definition der Struktur von Managementinformationen
- Konzepte vom Managementobjektklassen und deren Beziehung zueinander

Ein Managementobjekt beinhaltet Attribute, Operationen, die darauf angewendet werden können, sowie Verhaltensweisen und Meldungen (Attribute, Management Operation, Behaviour, Notification)[3]. Ein Managementobjekt ist Teil einer Klasse, die durch gleiche Eigenschaften charakterisiert ist[4].

Attribute spiegeln die Eigenschaften von Ressourcen wieder, sie geben als Datenelement einen aktuellen Status an. **Management Operationen** werden in attribut- und objektorientierte Operationen eingeteilt. Erstere wirken sich nur auf einzelne Attribute aus, während letztere gesamte Managementobjekte beeinflussen. Ein Managementobjekt beinhaltet gewisse **Verhaltensweisen**, nach denen es reagiert. Der Anstoß kann sowohl interner als auch externer Natur sein. So kann eine Aktion durch einen Zähler ausgelöst werden. Aber auch das Einwirken von Systemmanagementoperationen über CMIP auf das Managementobjekt, eben durch andere externe Systeme, kann eine definierte Verhaltensweise hervorrufen.

Die drei wichtigsten, in der MIB benutzten objektorientierten Methoden sind *Kapselung*, *Vererbung*, *Allomorphismus*.

2.1.2.2 Benennung der MIB Objekte

Mittels der objektorientierten Methode der Vererbung werden Klassenhierarchien aufgebaut. Jede Klasse ist durch einen eindeutigen Identifikator gekennzeichnet. Dieser Bezeichner ist eine Kette von Integerzahlen, die der Klasse einen eindeutigen Platz in einem der drei voneinander verschiedenen und unabhängigen Bäumen, dem *ISO-Registration-Tree* zuordnet. In diesem Baum sind die Definitionen für

- Managementobjektklassen,
- Attributdefinitionen,
- Aktionen,
- Meldungen und
- Pakete

abgelegt.

Die anderen beiden sind der *Inheritance*- und *Containment-Tree*.

[3]Das beharrliche Übersetzen im folgenden Absatz wird inhaltlich keine Verbesserung darstellen, jedoch wieder ein Beispiel dafür sein, das nicht alles *eingedeutschte* gut sein muß.

[4]...wie bereits vom OO-Ansatz bekannt...

2.1.2.3 Templates - Beschreibungsstruktur für Managementinformationen

Der Standard ISO 10165-1 schlägt Richtlinien für das Format von Definitionen von Managementobjekten, Attributen, Paketen und Meldungen vor. Diese Richtlinien liegen in Form von Templates vor, die in ASN.1 (Abstract Syntax Notation One) verfaßt sind. ASN.1 ist ein von der Zielumgebung unabhängiges Beschreibungsmittel, mit dem ausgedrückt werden kann, wie aus einzelnen Daten(-elementen) in korrekter Art und Weise komplexere Datenstrukturen und Informationsgebilde aufgebaut werden können. Deshalb ist ASN.1 gut geeignet um Managementobjekte zu beschreiben.

2.2 Internet-Protokollwelt und OSI-Modell

Mit TCP/IP wird allgemein eine Sammlung von Kommunikationsprotokollen bezeichnet. Als im Jahre 1983 UNIX 4.2 BSD durch Berkeley Software Distribution eingeführt wurde, erlangte TCP/IP einen erheblich größeren Bekanntheitsgrad als je zuvor. Dies war auch nicht weiter verwunderlich, zumal die US-Regierung diesen Teil von UNIX 4.2 BSD finanzierte. Die TCP/IP-Protokolle basieren auf Standards, die ursprünglich für die US-Regierung und die Forschungszentren der USA entwickelt wurden [Wash94]. TCP/IP

Daß das Protokoll TCP/IP die Grundlage des weltweit größten Netzes wurde, hat unter anderem folgende Gründe [Haec97]:

- Offene Protokollstandards, die herstellerunabhängig entwickelt wurden und die frei verwendbar sind. Es gibt praktisch für jede Art Netzwerkgerät TCP/IP-Unterstützung.
- Unabhängig von der verwendeten Netzwerktechnik: TCP/IP kann auf Ethernet, Token Ring, FDDI, ATM, X.25 usw. und ebenso auf Wählleitungen aufgesetzt werden.
- Ein gemeinsames Adressierungsschema für alle angeschlossenen Hosts, das eine eindeutige Identifikation jeder angeschlossenen Maschine ermöglicht.
- Einheitliche und offengelegte Protokolle in höheren Ebenen, die konsistente und verbreitete Services implementieren.

2.2.1 Die Protokollschichten

Ausgehend vom ISO/OSI[5]-Referenzmodell müßte man TCP/IP in 7 Schichten einteilen. Oft genug wurde dies auch versucht, denn das von ISO entwickelte Referenzmodell sollte einen Standard für die Entwicklung einer ganzen Protokollsuite bieten. ISO/OSI

[5] International Standard Organisation/Open Systems Interconnection

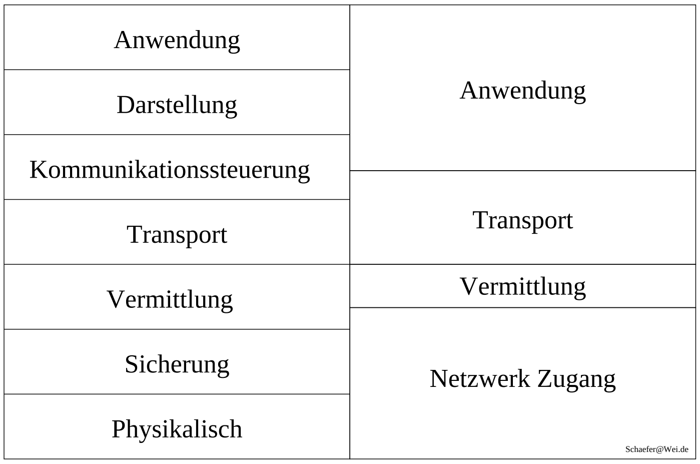


Abbildung 2.4: ISO/OSI- und TCP/IP-Schichtenmodell

TCP/IP Modell

Praktikabler ist es allerdings, das TCP/IP-Protokoll in nur vier Schichten einzuteilen [Haec97]:

- Anwendungsschicht
- Transportschicht
- Internetschicht
- Netzwerkschicht

Header

Die Schichten im Modell kommunizieren nur an ihren Kanten miteinander. Dabei werden zu sendende Daten von oben nach unten durchgereicht, wobei jede Schicht noch ein paar Informationen hinzufügt. Auf der Netzwerkschicht werden sie transportiert und beim Empfänger werden die Daten wieder nach oben durchgereicht. Beim Nachobenreichen werden die Informationen, sogenannte Header, wieder entfernt. Die Header sind dazu da, die Daten einzupacken. Mit Hilfe der Headerinformationen kann jede Schicht die Daten richtig interpretieren.

Eine kleine Übersicht vervollständigt das Bild 2.5 des Protokolls.

2.2.1.1 Netzwerkschicht

Die unterste Schicht dient der Übertragung der Daten. Normale Benutzer müssen sich mit dieser Schicht nicht befassen, da IP, UDP und TCP[6] diese Ebene transparent machen.

[6] unter Punkt Transportschicht erklärt

```
 +------+ +---+ +----+ +----+   +---+  +----+ +----------+ +-----+   +---+
 |Telnet| |FTP| |SMTP| |NAME|...|...|  |TFTP| |netbios-ns| |route|...|...|
 +------+ +---+ +----+ +----+   +---+  +----+ +----------+ +-----+   +---+
     |      |     |      |               |         |         |
     |      |     |      |               |         |         |
   +---------------------------+       +------------------+     +-----+
   |             TCP           |       |        UDP       | ... | ... |
   +---------------------------+       +------------------+     +-----+
                  |                             |
                  |                             |
          +---------------------------------------+
          |                                | ICMP |
          |-----+------+  Internet Protocol +------|
          | ARP | RARP |                           |
          +---------------------------------------+
                            |
                            |
               +-----------------------------+
               |     Local Network Protocol  |
               +-----------------------------+
```

Abbildung 2.5: Protokollrelationen

Auf der Netzwerkebene geschieht die Umsetzung vom IP-Datagrammen in Netzwerkframes und die Umsetzung von IP-Adressen in Hardwareadressen, sogenannte MAC-Adressen. Das zugehörige Protokoll ist das ARP-Protokoll[7]. Die MAC-Adresse[8] eines Netzwerkadapters wird normalerweise bei der Herstellung festgelegt und sollte wie ein Primärschlüssel in einer Datenbank einmalig (unique) sein. Das Format der MAC-Adresse ist xx:xx:xx:xx:xx:xx, wobei xx eine hexadezimale Zahl im Bereich 00..FF darstellt.

MAC und ARP

2.2.1.2 Internetschicht

Die Internetschicht von TCP/IP wird in RFC 791 [Post81a] beschrieben. Das Internet-Protokoll IP bildet den grundlegenden Paketdienst, den alle Protokolle darüber und darunter zur Zustellung von Daten nutzen. Dieses Protokoll ist für folgende Dinge zuständig [Haec97]:

- Festlegung des *Datagramms* als Grundeinheit der Übertragung im Internet
- Definieren des Internet-Adressierungsschemas
- Datenübertragung zwischen Transport- und Netzwerkschicht

[7] Address Resolution Protocol
[8] Media Access Control

- Routen von Datagrammen zu anderen Hosts
- Zerteilen und Wiederzusammensetzen von Datagrammen

Zwei wichtige Eigenschaften charakterisieren IP darüberhinaus [Haec97]:

1. IP ist ein *verbindungsunabhängiges* Protokoll. Vor der Datenübertragung wird, *nicht* wie bei verbindungsorientierten Protokollen, versucht, den Empfänger zu erreichen oder eine *Handshake*-Phase zur Verhandlung durchzuführen. Es ist die Aufgabe der darüberliegenden Protokolle, bei Bedarf solch eine Kommunikation zu implementieren.
2. IP ist ein *nicht verläßliches* Protokoll. Das gilt insofern, als es keine Fehlerprüfung oder Korrektur vornimmt. Dies bedeutet aber nicht, daß IP *unzuverlässig* ist. Es garantiert die korrekte Lieferung der Daten in das angeschlossene Netzwerk, nicht aber das korrekte Empfangen von Daten aus dem Netzwerk. Auch hierfür sind die Protokolle der darüberliegenden Schichten zuständig.

Die IP-Schicht enthält folgende Funktionalität:

- Datagramme
- ICMP - Internet Control Message Protocol

Das Internet-Protokoll ist ein paketvermittelndes Protokoll. Es werden alle Daten in Programmpaketen, sogenannten *Datagrammen*, versendet. Die Datagramme bestehen aus einem Header und den Daten. Im Header steht die Quell- und Zieladresse. Mit diesen und weiteren Informationen geht das jeweilige Paket seinen eigenen Weg zum Zielrechner. Je nach Routing können zusammengehörige Daten, in verschiedenen Paketen verpackt, unterschiedliche Wege gehen. Auf Grund der Zieladresse wird erkannt, ob der Empfänger im eigenen Netzwerk ist.

Datagramm
Header und Daten

Die MSS (Maximum Segment Size) ist die maximale Größe der Daten des Datenpaketes. Die MTU (Maximum Transmission Unit)[9] gibt an, welche Größe ein Datenpaket, inclusive Header, haben darf [Post83]. Wenn nun ein Datenpaket auf dem Weg zum Empfänger durch unterschiedliche Netzwerktechnologien übertragen wird, kann es sein, daß es fragmentiert wird, da unterschiedlich große Rahmen benutzt werden.

MSS
MTU

Um in der Internetschicht Kontroll-, Fehler- und Informationsfunktionen zu implementieren, ist in RFC 792 das *Internet Control Message Protocol* [Post81b] ICMP beschrieben. Dabei gibt es folgende Messagetypes:

ICMP

- 0 Echo Reply
- 3 Destination Unreachable

[9] die MTU bei Ethernet (10Mb) beträgt 1500 bytes

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|Version|  IHL  |Type of Service|          Total Length         |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|         Identification        |Flags|      Fragment Offset    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Time to Live |    Protocol   |         Header Checksum       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Source Address                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Destination Address                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Options                    |    Padding    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Begin of Data...                                            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Abbildung 2.6: Internet Datagram Header

- 4 Source Quench
- 5 Redirect
- 8 Echo
- 11 Time Exceeded
- 12 Parameter Problem
- 13 Timestamp
- 14 Timestamp Reply
- 15 Information Request
- 16 Information Reply

Daraus ergeben sich die vier grundlegenden Mechanismen [Haec97]:

1. Flußkontrolle: Wenn Datagramme zu schnell zur Verarbeitung ankommen, schickt der Empfänger eine *ICMP Source Quench Message* zum Sender, um diesen zeitweilig vom Senden abzuhalten.

2. Erkennung unerreichbarer Ziele: Ist ein Ziel nicht erreichbar, schickt das System, das den Fehler erkennt, eine *ICMP Destination Unreachable Message* an den Sender. Ist das gewünschte Ziel ein Netzwerk oder Host, sendet das vermittelnde Gateway die Nachricht. Ist das Ziel ein Port[10] auf dem Zielrechner, sendet dieser die Nachricht.

[10]Ports werden in einem der nächsten Kapitel behandelt

3. Umleitungsregeln: Eine *ICMP Redirect Message* enthält den Hinweis eines Gateways an einen Host, ein anderes — besseres — Gateway zu verwenden. Diese Nachricht kann nur verwendet werden, wenn Sender und beide Gateways im selben Netzwerk liegen.

4. Test eines entfernten Systems: Mittels einer *ICMP Echo Message* kann geprüft werden, ob ein anderer Netzteilnehmer ein funktionierendes IP in Betrieb hat. Das gängigste Kommando dafür ist `ping`.

5. Loop-detection: Bei jedem Schritt, wenn ein Paket einen Router durchläuft, wird der Wert des TTL Feldes (*Time to Live*) um eins erniedrigt. Bei null wird das Paket verworfen. Das verhindert unnötige Netzlast bei Fehlkonfigurationen aber auch daß Pakete terminiert werden, die zu viele Router passiert haben.

Darüberhinaus kann ICMP noch Zeitstempel und Informationsnachrichten verschicken.

2.2.1.3 Transportschicht

Die Transportschicht stellt eine konsistente Schnittstelle zu den Anwendungen und einen Mechanismus bereit, mit dem die Anwendungen ihre Kommunikation spezifizieren können. Abbildung 2.7 verdeutlicht, daß die Transportschicht für beide Endpunkte zuständig ist, unabhängig davon, ob die Kommunikations-Route durch ein einzelnes (möglicherweise brückenverbundenes) Netz oder durch mehrere routerverbundene Netze führt [Wash94].

UDP und TCP

Die meisten TCP/IP-Anwendungen verwenden derzeit entweder das Transportprotokoll UDP (*User Datagram Protocol*) oder TCP (*Transmission Control Protocol*). TCP ist dabei das bessere Protokoll für Übertragungen von Daten über weite Strecken, da die Performance beeinflußt werden kann. UDP hat keine solchen Parameter.

Einsatz von UDP

UDP ist im Vergleich zu TCP ein unsicherer Dienst. Unsicher deshalb, weil keine Garantie für die zuverlässige Übertragung der Daten gegeben werden kann. Wenn Anwendungen per UDP kommunizieren und die Sicherheit der korrekten Datenübertragung benötigen, so müssen die Anwendungen selbst mittels Fehlerkontrolle und -korrektur dafür sorgen. Andererseits sind Programmierschnittstellen zur Kommunikation per UDP wesentlich einfacher zu programmieren und deshalb auch sehr beliebt.

broadcast

Damit ist UDP ein Protokoll, welches sich hervorragend eignet, Informationen zu verteilen, welche sich ständig ändern und keine dauerhafte hundertprozentige Richtigkeit bedingen, da innerhalb kürzester Zeit erneut (die nächsten) Informationen anliegen. Unzuverlässige Datagramm-Dienste können ihre Daten an viele Empfänger rundsenden (broadcast), ohne eine Antwort zu erwarten.

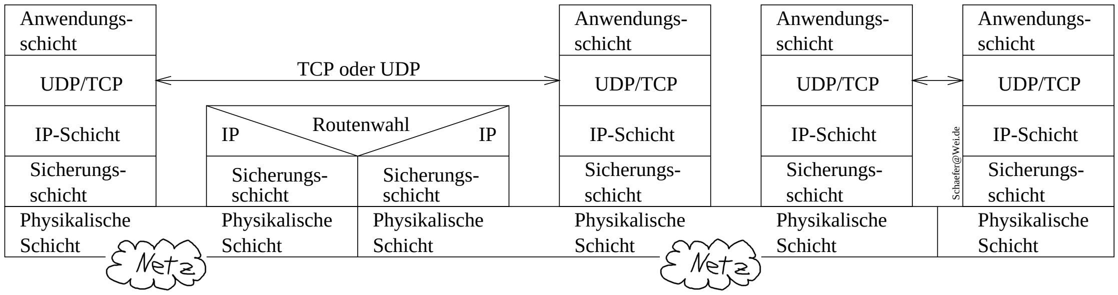


Abbildung 2.7: TCP/IP-Transportschichten

TCP bietet im Gegensatz zu UDP einen zuverlässigen Dienst mit End-zu-End-Signifikanz. Damit bietet TCP eine verläßliche, verbindungsorientierte Netzverbindung, deren Anwendung klar in Bereichen liegt, für die verläßliche Datenübertragung notwendig ist. Dazu gehören E-Mail- und Datei-Transfer sowie Zugriffe auf Terminal-Systeme etc..

Einsatz von TCP

2.2.1.4 Anwendungsschicht

Zur Anwendungsschicht gehören alle Programme, welche die Transportschicht verwenden, um Daten auszutauschen. Die Protokolle dieser Schicht sind im allgemeinen direkt für die Anwenderprogramme gedacht.

Anwenderprogramme

Die wichtigsten Protokolle sind:

1. TELNET: Das Netzwerkterminal-Protokoll, welches die Einwahl auf verschiedenen Rechnern erlaubt.
2. FTP: Das *File Transfer Protocol* dient der Übertragung von Dateien von entfernten und auf entfernte Rechner.
3. DNS: Der *Domain Name Service* bildet IP-Adressen des weltweiten, hierachischen Namensraums ab.
4. SMTP: Mittels des *Simple Mail Transfer Protocol* erfolgt die Übertragung elektronischer Post.
5. HTTP: Das *HyperText Transfer Protocol* überträgt Seiten und multimediale Elemente über das World Wide Web (*WWW*).

2.2.2 Ports, Sockets, Protokolle

2.2.2.1 Protokollnummer

Die Datei `/etc/protocols` beinhaltet die gängigsten Protokolle und die dazugehörigen Protokollnummern und, um die Stuktur der Datei komplett zu beschreiben, einen Aliasnamen. Die Protokollnummer ist die offizielle jeweilige Nummer, welche im Byte *Protocol* des IP-Headers steht. Sie bestimmt von welchem Protokoll das jeweilige Paket verarbeitet werden darf.

```
#
# /etc/protocols
#
# format: protocol number aliases ...
#
ip      0       IP      # internet protocol, pseudo protocol number
icmp    1       ICMP    # internet control message protocol
igmp    2       IGMP    # internet group management protocol
ggp     3       GGP     # gateway-gateway protocol
tcp     6       TCP     # transmission control protocol
egp     8       EGP     # exterior gateway protocol
pup     12      PUP     # PARC universal packet protocol
udp     17      UDP     # user datagram protocol
hmp     20      HMP     # host monitoring protocol
idp     22      IDP     #
rdp     27      RDP     # "reliable datagram" protocol
raw     255     RAW     #
```

Abbildung 2.8: Die Datei `/etc/protocols`

Diese Datei `/etc/protocols` (siehe Abbildung 2.8) enthält nicht alle Protokolle die es gibt, jedoch jene, die das System benutzt. Dabei sind, wie bereits erwähnt, ICMP, TCP, UDP und IP die wichtigsten. Wenn also ein Paket die IP-Ebene mit einem Wert von 6 im Byte *Protocol* erreicht, so wird das Paket vom jeweiligen Protkoll per TCP verarbeitet.

2.2.2.2 Ports

Die Datei `/etc/services` beinhaltet die gängigsten Dienste (*services*), die ein Unixsystem anbieten kann. Das sind jedoch weit mehr als wirklich benutzt werden. Dabei ist jeder Dienst einer Portnummer zugeordnet. Die Portnummern liegen im Bereich von 1 bis 65535. Ports unter 1024 (sogenannte *privileged ports*[11]) sind Diensten zugeordnet, die unter dem Superuseraccount root[12] laufen. Ports bis 256 gehören den reinen Internetdiensten wie z.B. HTTP (80), FTP (21), TELNET (23) bzw. SMTP (25) an. Ports ab 1024 sind eigentlich frei verfügbar und lokal einsetzbar, jedoch haben sich verschiedene Dienste dort fest angesiedelt. Dazu gehören zum Beispiel X-Window (ab 6000 aufwärts), Internet Relay Chat *IRC* (6667) oder *socks* (1080). Häufig findet man heute HTTP auch auf Port 8080.

Wenn sich ein entfernter Rechner an einem System anmelden will, so tut er das auf der jeweiligen Portnummer am Zielrechner, wie in der Datei

[11] auch *low numbered ports*

[12] dieser User hat auf einem normalen Unixsystem, von AFS und Kerberos (auf AFS und Kerberos wird in dieser Arbeit nicht eingegangen) abgesehen, alle Rechte

```
#
# /etc/services
#
# format:   service-name   port/protocol   [aliases ...]
#
tcpmux          1/tcp         # TCP Port Service Multiplexer
rje             5/tc          # remote job entry
echo            7/tcp
echo            7/udp
discard         9/tcp         sink null
discard         9/udp         sink null
systat          11/udp        users
systat          11/tcp        users
daytime         13/tcp
daytime         13/udp
netstat         15/udp
netstat         15/tcp
[...]
ftp-data        20/tcp
ftp             21/tcp
telnet          23/tcp
smtp            25/tcp        mail # Simple Mail Transfer
time            37/udp        timserver
time            37/tcp        timserver
name            42/tcp        nameserver
name            42/udp        nameserver
whois           43/udp        nicname
whois           43/tcp        nicname
nameserver      53/tcp        domain
nameserver      53/udp        domain
bootps          67/udp        bootp
bootps          67/tcp        # bootstrap protocol server
bootpc          68/udp
bootpc          68/tcp        # bootstrap client
tftp            69/udp
tftp            69/tcp        # trivial file transfer
snmp            161/udp
snmp-trap       162/udp
[...]
biff            512/udp       # used by mail system to notify users
exec            512/tcp
who             513/udp       # maintains data bases showing who's
login           513/tcp
syslog          514/udp
shell           514/tcp       cmd
printer         515/tcp       # spooler
talk            517/udp
ntalk           518/udp
utime           519/udp       unixtime
route           520/udp       router routed # local routing process (on site);
[...]
```

Abbildung 2.9: Datei `/etc/services`

`/etc/services` (siehe Abbildung 2.9) beschrieben ist.

2.2.2.3 Sockets

Neben den statischen Portnummern, welche in der Datei `/etc/services` festgelegt sind, gibt es außerdem *dynamisch zugeordnete* Ports. Diese sind vor der Anforderung nicht bekannt. Das System ist für die ordnungsgemäße Vergabe zuständig[13]. Wenn beispielsweise Kommunikationspartner denselben Port benutzen würden, wäre nur eine Netzwerkverbindung für den jeweiligen Dienst möglich, was der Multiuser-Fähigkeit entgegenstünde.

Das Wort Socket beschreibt die Kombination aus IP-Addresse und Portnummer, ein Socketpaar die Kombination zweier Sockets. Damit ist jede Netzwerkverbindung netzweit eindeutig definiert. IP-Addresse und Portnummer von Sender und Empfänger entsprechen einem Quadrupel. Socket netzweit eindeutig

2.2.3 Zusammenfassung

Abschließend soll eine Protokollverbindung die Grundlagen des TCP/IP Protokoll abrunden. In Abbildung 2.10 ist dies exemplarisch am TELNET Protokoll dargestellt. Wie in Kapitel 2.2.2 zu lesen war, erfolgt die eindeutige und mehrmals mögliche Zuordnung einer Session über die Portnummern. Auf der Clientseite wird eine zufällige Portnummer größer 1024 gewählt, über welche die Verbindung abgewickelt wird. Auf der Serverseite wird ein für den jeweiligen Dienst fest definierter Port angesprochen. Im Bild 2.10 ist das Port 23. Zusammen mit den beiden IP-Adressen wird eine internetweite Eindeutigkeit gewährleistet, die benötigt wird, um genau eine Session zu spezifizieren.

2.2.3.1 Nachtrag, RFCs

Die RFCs sind die Definitionsbasis des Netzes der Netze, dem Internet. Alle Protokolle, Kommunikationsarten und Beschreibung von Komponenten, die einen Status als Standard erhalten sollen, sind in solchen RFCs dokumentiert. Sie enthalten Implementierungsvorschriften bzw. -vorschläge. Die Einteilung erfolgt in fünf Kategorien: *Full Standard*, *Draft Standard*, *Proposed Standard*, *Experimental*, *Informational*. RFC

2.2.4 SNMP - Simple Network Management Protocol

Das *Simple Network Management Protocol* wurde aus dem *Simple Gateway Monitoring Protocol* (RFC 1028, [Davi87]) entwickelt. Die erste Version von SNMP SGMP

[13] zum besseren Verständnis folgt im Kapitel 2.2.3 das vollständige Beispiel zum Verbindungsaufbau

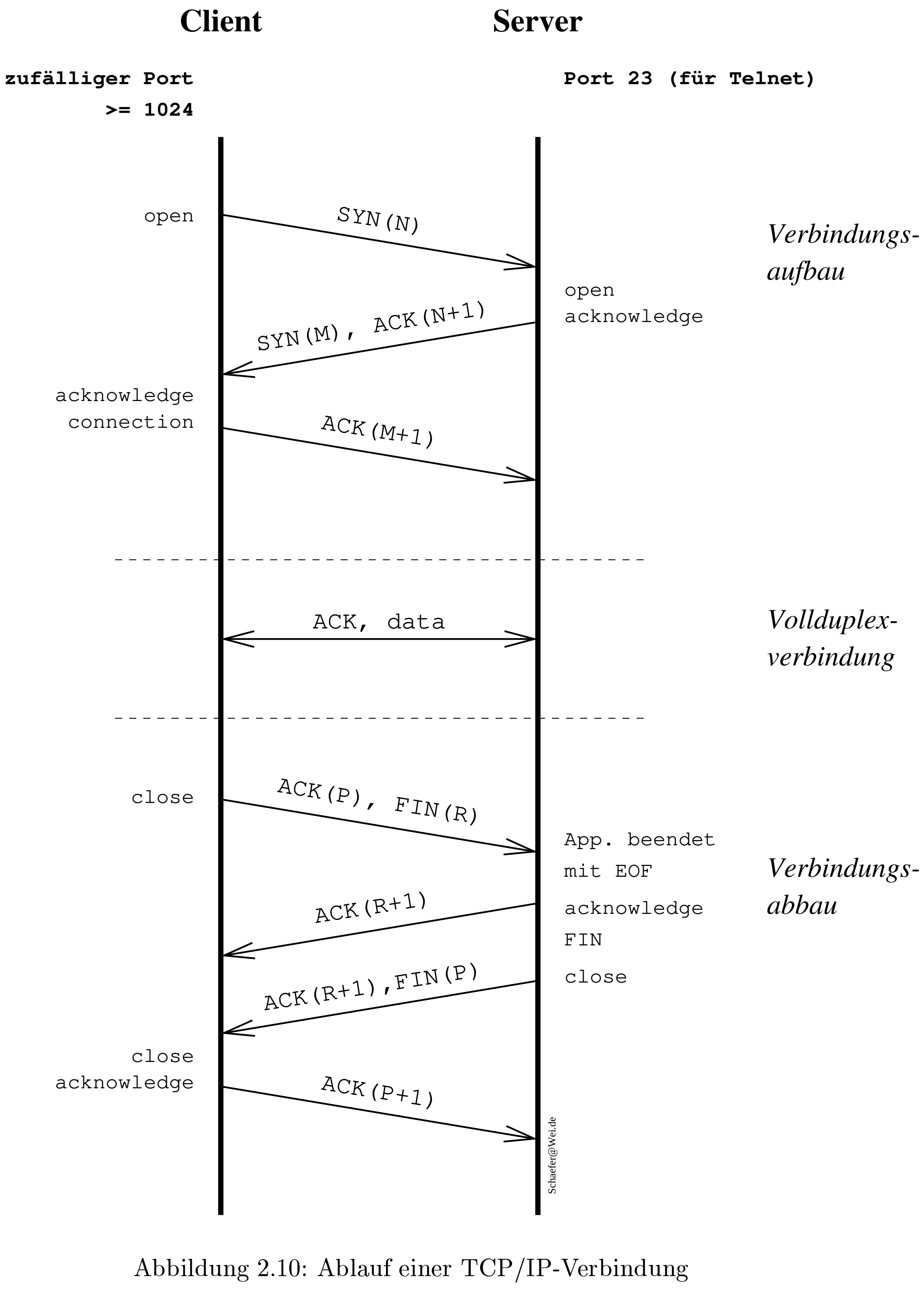


Abbildung 2.10: Ablauf einer TCP/IP-Verbindung

SNMP entstand 1988, wurde 1989 überarbeitet und erlangte 1990 (RFC 1157, [Case90]) Internet Standard [Hunt94].

Mittels SNMP können Netzwerkmanager von ihrem Arbeitsplatz aus entfernte Geräte wie Router und andere Komponenten beobachten und kontrollieren. Der Status eines jeden solchen Gerätes wird durch Daten in Form von Tabellen und Variablen dargestellt. Die Struktur der Daten ist durch die *Management Information Base* (MIB) definiert. Das eigentliche Monitoring, also das Beobachten, erfolgt durch Auslesen dieser Variablen mittels SNMP-Kommandos. Ein Verändern der Arbeitsweise der Netzwerkkomponenten erfolgt durch ein Setzen neuer Werte in den Variablen, ebenfalls durch SNMP-Kommandos. Der Umfang der definierten MIBs [McCl92] ist, in Anbetracht der Zahl der managebaren Objekte, sehr groß. Monitoring

SNMP ist ein sehr einfaches Protokoll, welches auf einem Client-Server Modell basiert und mit UDP realisiert wurde. Es wurde für lowbudget Systeme mit wenig Speicher und Ressourcen entwickelt. Auf Grund der Einfachheit gibt es heute auf allen Systemen Implementierungen, um eigene Systeme managen zu können.

Eine neue Version von SNMP - SNMPv2 - wurde 1993 von IETF, dem „Internet Engineering Task Force“, vorgestellt. Der große Unterschied zu SNMP ist der Einbau effizienterer Sicherheitsmechanismen, speziell kryptografischer Techniken, die Vereinfachung und der Einbau eines MIB-Formats, welches einfach kompiliert werden kann [Hunt94]. SNMPv2

Detaillierter wird SNMP im Kapitel Internet-Management (3.2) behandelt.

Kapitel 3

Netzwerkmanagement

Einführung in Netzwerkmanagement, Funktion und Aufgaben, Bereiche, Modelle, typische Verfahren

In Kapitel 2, speziell in 2.1 wurden die Grundlagen für den nun folgenden Teil beschrieben. Zugrunde lag hauptsächlich die Definition nach OSI. Wichtige Begriffe wurden erwähnt und erklärt. In diesem Teil soll Netzwerkmanagement in der aktuellen und gebräuchlichen Form, so wie es auch in der Industrie zum Einsatz kommt, beschrieben werden. Aber auch weiterführende Aspekte werden eine Rolle spielen. Sich ständig ändernde Transportmedien für Datennetze spielen ebenso eine Rolle, wie bereits etablierte Standards.

3.1 Anforderungen, Definition, funktionelle Bereiche

3.1.1 Anforderungen und Definition

Netzwerkmanagement umfaßt in einem Oberbegriff eine Vielzahl von Funktionen, Aufgaben und Problemen. Die Aufgaben und Funktionen sollen in diesem Kapitel ebenso beschrieben werden, wie dabei auftretende Probleme.

Kommunikation und Information

Einem Endbenutzer stellt das Netzwerk zwei grundlegende Funktionen zur Verfügung, Kommunikation und Information. Ein Benutzer erwartet die Verfügbarkeit eines Kommunikationspfades, einen Datenübertragungsmechanismus zum Ziel der Wahl. Möglicherweise werden darüber hinaus erweiterte Informationsangebote erwartet, dazu zählen Adreßrecherchemöglichkeiten genauso wie Informationen über die Inanspruchnahme des Netzes. Diese Forderungen werden zum Beispiel von der Television[1] nicht erfüllt.

[1]das generische Fernsehen

Zwar ist die Forderung nach Kommunikation und Information erfüllt, jedoch muß ein beschränktes, zur Verfügung gestelltes Angebot hingenommen werden. Weder das Ziel und dessen Inhalt kann unmittelbar bestimmt werden, noch der Kommunikationspfad[2]. Weitere Informationsangebote, wie Informationen über die Inanspruchnahme des Netzes, sind nicht abrufbar. Im Gegensatz zum Beispiel der Television können die Erwartungen an erweiterte Informationsangebote am Beispiel der Nutzung des Telefonnetzwerkes erfüllt werden. Hier sind nicht nur die Wahl des Ziels möglich (Kommunikation und Information als Dienstleistung eingeschlossen), sondern Adreßrecherchemöglichkeiten oder neuere, "intelligente" Nutzungsangebote (ISDN) stehen zur Verfügung. Das Netzwerkmanagement soll natürlich für die effektive und reibungslose (und billige) Realisierung dieser Funktionen sorgen.

Mit immer schnellerem Wachstum und immer größer werdender Heterogenität heutiger Netzwerke, wächst auch das Bedürfnis an einem effizienteren und leistungsfähigeren Ansatz für das Netzwerkmanagement. Dieses soll eine bessere Sicht auf die im Netz verfügbaren Ressourcen schaffen und deren Kontrolle ermöglichen. Fehler und Aktionen sollen erkannt und durch ein Managementsystem gesteuert werden. Rechtzeitige Maßnahmen können Ausfallzeiten verhindern, mögliche Netzlastenverteilungen erreichen, also eine durch Steuerung optimale Nutzung des Netzes und dessen Ressourcen erzielen. Damit ist nicht nur die Forderung des Netzbetreibers erfüllt, sondern auch eine bestmögliche Dienstqualität für den Endbenutzer erreicht [Chim98].

Übersicht und Kontrolle

Steuerung

Die Internationale Standard Organisation (ISO) hat einen großen Beitrag zum Netzmanagement geleistet, indem sie alle Aktivitäten zur Steuerung, Koordinierung und Überwachung im Netz festgelegt hat. Es wurde versucht, alle Anforderungen, die an ein Managementsystem gestellt werden können, zu formulieren. Daraus resultieren die in folgenden fünf Punkten zusammengefaßten Bereiche, die die Aufgaben bzw. Funktionen widerspiegeln.

3.1.2 Bereiche

Die von der International Standard Organisation definierten und zusammengefaßten Bereiche des Netzmanagementmodells werden (inzwischen) allgemein auch außerhalb des OSI-Kontextes verwendet. Die Bereiche sind [Chim98]:

- Fehlermanagement (*Faultmanagement*)
- Leistungsmanagement (*Performance Management*)
- Konfigurationsmanagement (*Configuration Management*)

[2]Es handelt sich um ein *Broadcast*medium.

- Abrechnungsmanagement (*Accounting Management*)
- Sicherheitsmanagement (*Security Management*)

Das Fehlermanagement umfaßt Erkennen und Lokalisieren von Fehlern aber auch das Eingrenzen derer. Es sollte ein Lösungsmodell für die Fehlerbehebung anbieten. Fehler sind nicht nur der Ausfall von ganzen Komponenten oder Diensten (*services*), bereits eine Verschlechterung der Dienstqualität muß erkannt werden. Das bedeutet, daß bereits die Symptome der Fehler erkannt werden müssen.

Aufgabe des Leistungsmanagements ist die Messung der Netzleistung, um gegebenenfalls qualitätsverbessernd einzugreifen. Die Leistung, beschrieben durch verschiedene Parameter wie Durchsatz, Antwortzeit und Auslastung, soll durch das Finden und Beheben von Engpässen optimiert werden. Die Vorgehensweise dabei erstreckt sich vom Sammeln der Daten bis hin zu ihrer Analyse. Dabei spielt die Ermittlung von Schwellwerten (*Baseline*) eine ebenso große Rolle wie die Überwachung des Regelverhaltens.

Das Konfigurationsmanagement ist zuständig für die Vergabe von Namen der Ressourcen wie auch für die Initialisierung, das Starten und Beenden. Im Rahmen des Konfigurationsmanagements müssen auch Betriebsparameter der Ressourcen gesetzt und geändert werden können. Eine Aufgabe ist die Verwaltung von Hardware- und Softwareinformationen, wie zum Beispiel Betriebssystem- und Schnittstellenkartenversionen o.ä..

Das Abrechnungsmanagement soll die Nutzung des Netzes durch Personen und Nutzergruppen erfassen, um eine Aufteilung bzw. Abrechnung der entstandenen Kosten zu verteilen. Dadurch werden auch Informationen über Nutzungsgewohnheiten und den Bedarf gesammelt, die bei späterer Netzplanung von Bedeutung sein können.

Schließlich hat das Sicherheitsmanagement die Aufgabe der Überwachung des Zugriffs auf Ressourcen des Netzes nach festgelegten Regeln. Dazu gehört die Authentifizierungen von Nutzern ebenso wie der Schutz sensibler Daten vor unbefugten Zugriffen.

Die Implementierung der genannten Managementbereiche in kommerziellen Produkten ist sehr unterschiedlich, jedoch ist das Fehlermanagement das in allen Produkten der am häufigsten integrierte Bereich [Chim98].

An der Stelle, an der in vergangenen Zeiten Netzwerkmanagement als Synonym für das Management von Kommunikationskomponenten, also Hardware wie Router und Hubs, stand, hat sich heute eine neue Erkenntnis durchgesetzt. Es wird von *Integriertem Management*, also der Verwaltung von Hardware- und Softwareressourcen, gesprochen und folgende Disziplinen werden unterschieden [Hege95]:

- Netzmanagement: Management von Kommunikationskomponenten
- Systemmanagement: Management der Endsysteme und deren Bestandteile

- Anwendungsmanagement: Management von verteilten Anwendungen, d.h. Anwendungen, die auf mehreren kooperierenden Systemen laufen.

Da die Trennung zwischen den einzelnen Disziplinen nicht immer eindeutig ist, dies aber auch nicht immer erforderlich ist, soll hier nur von Netzmanagement gesprochen werden.

3.2 Internet-Management

3.2.1 SNMP, SNMPv2, SNMPv3 bzw. SNMPng

Das *Simple Network Management Protocol (SNMP)* wurde zwischen 1987 und 1989 entwickelt und wird heute auch als SNMPv1 bezeichnet. Es ist der Nachfolger des *Simple Gateway Management Protocol*, welches dazu diente Router im Internet zu managen. Das schnelle Wachstum des Internets und der mangelnde Komfort des alten Protokolls machten die Entwicklung von SNMP notwendig.

Das Ziel der Entwicklung war es, ein einfaches, leicht zu implementierendes Protokoll zu entwickeln, das dennoch offen für Erweiterungen war. Das in der Schicht sieben des OSI Modells angesiedelte Protokoll kommt in der ersten Version mit nur vier Befehlen aus und ist damit tatsächlich einfach. Der Austausch der Daten zwischen Agenten und Managern erfolgt über das verbindungslose Transportprotokoll UDP. Vor- und Nachteile von UDP sind schon in den Grundlagen (Kapitel 2.2.1.1) beschrieben worden. Mit der Nutzung von UDP wird gewährleistet, daß in stark ausgelasteten Netzen keine Verbindungen aufgebaut werden müssen, die ein ohnehin schon belastetes Netz weiter einschränken würden.

Die Entwicklung von SNMP im allgemeinen wird durch die *Internet Engineering Task Force* (IETF) [Ietf99] im Auftrag der *Internet Activities Board* (IAB) [Iab99] weiter vorangetrieben [Chim98].

3.2.1.1 SNMP

Die vier Befehle von SNMPv1 und deren Bedeutung können Tabelle 3.1 entnommen werden. Wie bereits erwähnt ist es zwar gelungen ein einfaches Umfeld für den Umgang mit SNMP zu schaffen, doch wurden damit zugleich Schwächen integriert.

Wo die konzeptionellen Schwächen von SNMPv1 lagen, ist schnell zusammengefaßt [Chim98]:

- Das **Fehlen von Sicherheitsmechanismen** zur Übertragung von Managementdaten ist ebenso ein Problem wie die Authentifizierung des Ursprungs einer Nachricht mangels implementierter Möglichkeiten. Damit ist der Einsatz von SNMPv1 in WANs problematisch.

Befehl	*Funktion*
Get	Lesen des Inhalts einer Variablen aus der MIB eines Agenten.
GetNext	Lesen des Inhalts der unmittelbar folgenden MIB-Variablen.
Set	Wertzuweisung zu einer Variablen in der MIB.
Trap	Melden eines (kritischen) Ereignisses an den Manager durch den Agenten.

Tabelle 3.1: SNMPv1-Befehle [Chim98]

- Die **umständliche sequentielle Abfrage der MIB-Variablen** resultiert aus der Entscheidung für das verbindungslose Protokoll UDP. In einem Schritt kann nur so viel Information übertragen werden, wie in einem Datagramm verpackt werden kann. Dadurch gibt es keine Möglichkeit des Auslesens der gesamten MIB in einem Schritt bzw. durch eine Operation. Im Kontext von High-Speed Networking ist die sequentielle Arbeitsweise ungeeignet.

- Die **verbindungslose Übertragung der Managementdaten** sichert nicht die Empfangsgarantie. Sowohl Abfragen können verloren gehen als auch Ereignismeldungen und Antworten der Agenten.

Die Standards zu SNMP sind in RFCs verankert. Die grundlegendsten, sogenannte *Full Standards* sind RFC 1155, 1157 und 1213. RFC 1155 heißt *Structure and identification of management information for TCP/IP-based internets* und beschreibt allgemeine Strukturen und Identifikationsschemata für Managementinformationen auf der Basis von TCP/IP [Rose90]. RFC 1157, *A Simple Network Management Protocol (SNMP)*, beinhaltet die SNMP Architekturbeschreibung und Protokoll Spezifikation. Und RFC 1213, *Management Information Base for Network Management of TCP/IP-based internets: MIB-II*, beschreibt die zweite Version der Management Information Base (MIB-II) für Netzwerkmanagementprotokolle in TCP/IP-basierten Netzwerken. MIB-II

Die **Netzwerkmanagement Architektur** basiert auf einem Client-Server Prinzip und teilt sich in folgende Komponenten auf: Architektur

- Manager Station (Server)
- Management Agent (Client)
- Management Information Base (MIB)
- Netzwerkmanagementprotokoll

Der Manager ist ein einzelnes Gerät oder auch ein verteiltes System, welches eine für Menschen lesbare Sicht auf die Daten der Agenten gibt. So ist ein

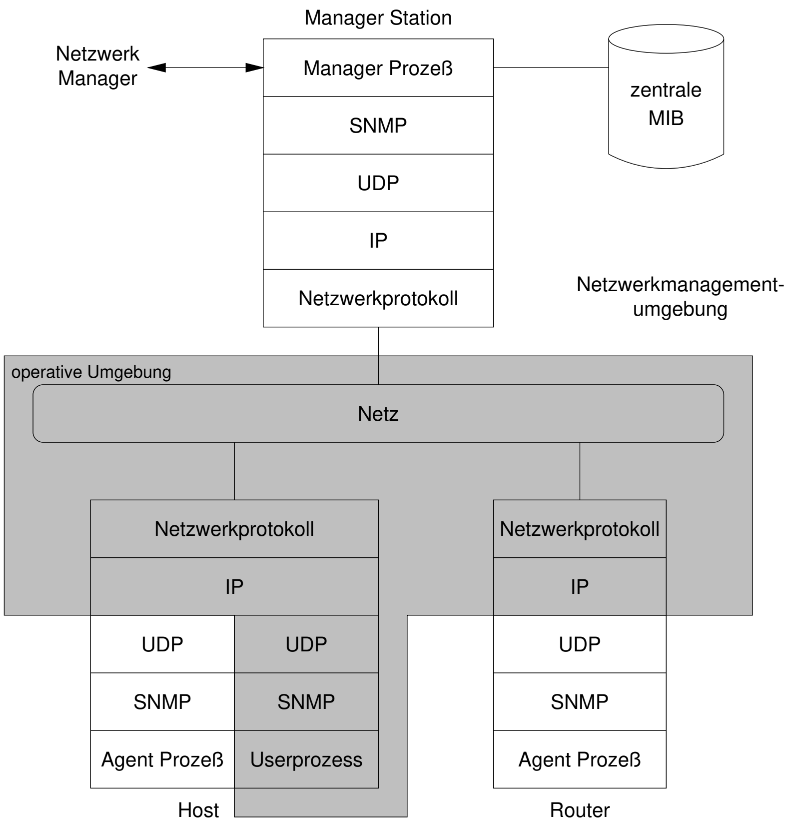


Abbildung 3.1: Arbeitsweise und Umgebung von SNMP [Stal93]

Satz von Datenanalysetools, Fehleranzeigen, Monitoren usw. vorhanden. Ebenso ist die MIB (Management Information Base) enthalten, um die Agenten abfragen und steuern zu können.

3.2.1.2 SNMPv2

Mit der Einführung von SNMP Version 2 (SNMPv2) als Erweiterung von SNMPv1 wurden Lösungen zu den Problemen und daraus resultierenden Schwächen adoptiert. So wurde der *Data Encryption Standard* (DES) eingeführt, um Managmentdaten verschlüsseln zu können. Nach [Chim98] kann mit Hilfe des *Digest Authentification Protocol* (DAS) die Prüfsumme der Protokolldateneinheiten berechnet werden. Weiterhin wurde der Befehl *GetBulk* eingeführt, so daß in SNMPv2 ein Auslesen der gesamten MIB am Stück möglich ist.

DES

Die funktionelle Beziehung zwischen Agenten und Managern wurde insofern verändert, daß Agenten ebenfalls Managerfunktionen ausführen können. Damit wurde eine Manager-Manager-Kommunikation eingeführt, welche bereits aus damaliger Sicht für die zukünftige Entwicklung wichtig erschien.

Die Rückwärtskompatibilität zu SNMPv1 ist mit der Version 2 nicht gewährleistet.

Es lassen sich die Schlüsselerweiterungen und -veränderungen im Vergleich zu SNMPv1 wie folgt zusammenfassen:

- Struktur der Managementinformationen
- Protokolloperationen
- Manager-Manager-Kommunikation
- Sicherheit

Die **Struktur der Managementinformationen** (*Struktur of Management Mnformation (SMI)* basiert auf der SMI von SNMPv1. Sie läßt sich unterteilen in Objektdefinitionen (*Object Definitions*), zusammenfassende Tabellen (*Conceptual Tables*), Nachrichtendefinitionen (*Notification Definitions*), Informationsmodule (*Information Modules*).

Wie in SNMPv1 SMI, werden die Objektdefintionen auch in ASN.1 Notation beschrieben, um Objekte zu verwalten. Das Makro `OBJECT-TYPE` soll als Beispiel für Syntax und Semantik dienen (Abbildung 3.2).

ASN.1

Da sowohl in SNMPv1 als auch in SNMPv2 Managementoperationen als skalare Objekte angewandt werden, können komplexere Informationen in Tabellen zusammengefaßt werden (*Conceptual Tables*).

Nachrichtendefinitionen (*Notification Definitions*) dienen, wie der Name es bereits vermittelt, der strukturierten Übermittlung von Nachrichten, welche einfache Informationen oder sogenannte *Traps* sein können.

```
OBJECT-TYPE MACRO ::=
          BEGIN
              TYPE NOTATION ::=
                                          -- must conform to
                                          -- RFC1155's ObjectSyntax
                                "SYNTAX" type(ObjectSyntax)
                                "ACCESS" Access
                                "STATUS" Status
                                DescrPart
                                ReferPart
                                IndexPart
                                DefValPart
              VALUE NOTATION ::= value (VALUE ObjectName)
              Access ::= "read-only"
                              | "read-write"
                              | "write-only"
                              | "not-accessible"
              Status ::= "mandatory"
                              | "optional"
                              | "obsolete"
                              | "deprecated"
              DescrPart ::=
                         "DESCRIPTION" value (description DisplayString)
                              | empty
              ReferPart ::=
                         "REFERENCE" value (reference DisplayString)
                              | empty
              IndexPart ::=
                         "INDEX" "{" IndexTypes "}"
                              | empty
              IndexTypes ::=
                         IndexType | IndexTypes "," IndexType
              IndexType ::=
                                  -- if indexobject, use the SYNTAX
                                  -- value of the correspondent
                                  -- OBJECT-TYPE invocation
                         value (indexobject ObjectName)
                                  -- otherwise use named SMI type
                                  -- must conform to IndexSyntax below
                              | type (indextype)
              DefValPart ::=
                         "DEFVAL" "{" value (defvalue ObjectSyntax) "}"
                              | empty
          END
```

Abbildung 3.2: Das Makro `OBJECT-TYPE`, aus [Rose90] und [Rose91]

```
ObjectName ::= OBJECT IDENTIFIER

               -- syntax of objects in the MIB

ObjectSyntax ::= CHOICE { simple            SimpleSyntax,
                          -- note that simple SEQUENCEs are not directly
                          -- mentioned here to keep things simple (i.e.,
                          -- prevent mis-use).  However, application-wide
                          -- types which are IMPLICITly encoded simple
                          -- SEQUENCEs may appear in the following CHOICE
                          application-wide  ApplicationSyntax }

SimpleSyntax ::= CHOICE { number  INTEGER,
                          string  OCTET STRING,
                          object  OBJECT IDENTIFIER,
                          empty   NULL }

ApplicationSyntax ::= CHOICE { address    NetworkAddress,
                               counter    Counter,
                               gauge      Gauge,
                               ticks      TimeTicks,
                               arbitrary  Opaque
                               -- other application-wide types, as they are
                               -- defined, will be added here
                             }

                  -- application-wide types
NetworkAddress ::= CHOICE { internet IpAddress  }

IpAddress ::= [APPLICATION 0] -- in network-byte order
                   IMPLICIT OCTET STRING (SIZE (4))

Counter ::= [APPLICATION 1] IMPLICIT INTEGER (0..4294967295)

Gauge ::= [APPLICATION 2] IMPLICIT INTEGER (0..4294967295)

TimeTicks ::= [APPLICATION 3] IMPLICIT INTEGER (0..4294967295)

Opaque ::= [APPLICATION 4]   -- arbitrary ASN.1 value,
                IMPLICIT OCTET STRING   -- "double-wrapped"
```

Abbildung 3.3: Das Makro `OBJECT-TYPE`, aus [Rose90] und [Rose91] (Forts.)

In SNMPv2 wurde das Konzept von Informationsmodulen (*Information Modules*) eingeführt. Damit lassen sich Informationen über die Module in der MIB definieren, respektive abfragen.

Nicht nur die **Protokolloperation** (*Protocol Data Unit* (PDU)) *GetBulkRequest* wurde eingeführt, um ganze Blöcke von Daten zu übertragen, sondern auch *InformRequest*, um Informationstraps zu anderen Agenten bzw. Managern zu schicken.

Die **Manager-Manager-Kommunikation** (auch M2M MIB) wurde in SNMPv2 eingeführt, um eine verteilte Managementarchitektur aufbauen zu können. Wie bereits erwähnt, kann so jeder Agent auch als Manager fungieren. So kann zum Beispiel ein Agent als *remote monitor* arbeiten, um Übertragungsdaten zu sammeln.

Allein das entwickele **Sicherheitskonzept** bedarf ein wenig mehr Erklärung. Abbildung 3.4 zeigt den Weg, welche Stationen nötig waren, um zu SNMPv2 zu gelangen. Dabei wird ersichtlich, daß erstmalig Sicherheitskonzepte integriert wurden. Die Sicherheitsmechanismen basieren auf denen der *Secure SNMP* (S-SNMP) Implementierung und wurden teilweise verändert.

Drei grundsätzliche Sicherheitsmechanismen sind vorgesehen: *Privacy*, *Authentication*, *Access Control*. Privacy basiert auf einem symmetrischen kryptographischen Protokoll, so daß nur Sender und Empfänger die übermittelte Nachricht entschlüsseln können. Dadurch ist für den Empfänger nicht nur die Authentizität der Nachricht gewährleistet (unter der Voraussetzung, daß der Schlüssel nur den beiden Parteien bekannt ist), sondern die Daten werden tatsächlich verschlüsselt übertragen. Der Mechanismus basiert auf einem konventionellen Verschlüsselungsverfahren, dem Data Encryption Standard (DES).

symmetrisches Protkoll

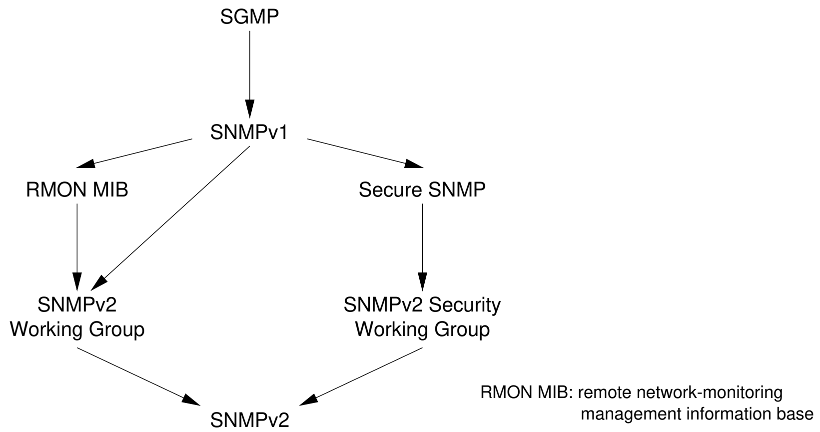


Abbildung 3.4: Die Entwicklung von SNMPv2 [Stal93]

Die eigentliche Authentizität, daß Daten nicht verändert wurden und auch von der erwarteten Quelle stammen, wird über ein asymmetrisches Verfahren gewährleistet. Über denselben Mechanismus kann auch zeitsynchroner Datenfluß garantiert werden, so daß Daten nicht verzögerbar bzw. in ihrer originalen Reihenfolge verarbeitbar sind. Der asymmetrische Authentifizierungsmechanismus von Secure SNMP (S-SNMP) und somit auch von SNMPv2 ist der *MD5 Message-Digest Algorithm*. Mit MD5 kann die Integrität der empfangenen Daten, die Ursprungsherkunft und Zeitlinearität verifiziert werden.

asymmetrisches Protkoll

3.2.1.3 SNMPv3

Seit 1996 läuft die Entwicklung von SNMPv3, auch als SNMPng bekannt. SNMPv3 bietet sicheren Zugriff zu Geräten durch eine Kombination aus einerseits Authentifizierung und andererseits Verschlüsselung vom Datenverkehr über das Netzwerk. Dadurch wird die Integrität der versendeten Informationen gesichert.

Durch die Authentifizierung wird die Quelle der Daten als solche identifiziert. Die Verschlüsselung sichert die Daten vor Veränderung aber auch vor dem Lesen Dritter.

Ein Sicherheitsmodell, die Authentifikationsstrategie der Aufteilung von Benutzer in Gruppen, wurde in SNMPv3 integriert. Diese Gruppen haben in Sicherheitsebenen verschiedene Möglichkeiten, je nach Rechten. Die Kombination aus Gruppenmodell und Sicherheitsebenen erlaubt ein sicheres Management.

Die Firma Cisco implementierte in der Routersoftware die in Tabelle 3.2 beschriebenen Sicherheitsmechanismen [Cisc99].

3.2.2 RMON, RMON2

Remote Network Monitoring (RMON) ist eine standardisierte Spezifikation zum Monitoring bestimmter Netzwerkkomponenten und dem Austausch der gesammelten Daten.

RMON wurde mit Unterstützung der Internet Engineering Task Force (IETF) entwickelt und letztlich in RFC 1757 als Draft Standard festgehalten.

In RMON sind Gruppen definiert, welche gesammelte Informationen zusammenfassen und einen spezifischen Satz an Daten darstellen. So beinhaltet die RMON Gruppe *Statistics* folgende Elemente eines Interfaces eines Routers: *packets dropped*, *packets sent*, *bytes sent (octets)*, *broadcast packets*, *multicast packets*, *CRC errors*, *runts*, *giants*, *fragments*, *jabbers*, *collissions*, ... [Cisc99a]. Anderer Gruppen sind *History*, *Alarm* usw.

Jede Gruppe ist insofern optional, daß es Herstellern überlassen wird, welche der Gruppen in Remote Network Monitoring Management Information Base (RMON MIB) unterstützt bzw. implementiert werden. Allerdings hängen manche Gruppen voneinander ab.

Modell	Ebene	Authentifizierung	Verschlüsselung	wie
v1	noAuthNoPriv	Community String	Nein	Community String als Authentifizierung
v2c	noAuthNoPriv	Community String	Nein	Community String als Authentifizierung
v3	noAuthNoPriv	Username	Nein	Username als Authentifizierung
v3	AuthNoPriv	MD5 oder SHA	Nein	Authentifizierung basiert auf dem HMAC-MD5 oder HMAC-SHA Algorithmus.
v3	AuthPriv	MD5 oder SHA	ja	Authentifizierung basiert auf dem HMAC-MD5 oder HMAC-SHA Algorithmus. DES 56-bit Verschlüsselung in Verbindung mit CBC-DES Authenifizierungs-standard (DES-56).

Tabelle 3.2: SNMP Sicherheitsmodelle und Ebenen

RMON verbindet folgende Vorteile:

- Vorhersage von Wachstum und Änderung
- Erhöhung der Netzzuverlässigkeit und -verfügbarkeit
- Schnelle Fehlersuche

3.3 OSI-Management

Da die in Kapitel 3.1.2 beschriebenen Punkte im Kontext des Netzwerkmanagements häufig auftreten, sollen sie detailliert beschrieben werden. Der Blickwinkel liegt hier jedoch nicht in der Beschreibung spezieller Bereiche, sondern vielmehr sollen durch Beispiele aus allen Bereichen der Einsatz und die Verfahren näher gebracht werden.

3.3.1 Fehlermanagement

Unter dem Fehlermanagement werden die Aufgaben verstanden, Fehler zu erkennen, zu lokalisieren und zu beheben. Dies ist besonders in großen Systemen nicht einfach, da viele Komponenten zusammen arbeiten. Die Überwachung der Funktion, also ob Ausfälle aufgetreten sind oder bevorstehen, ist von zentralem Interesse. Im Fehlerfall ist die Klärung der folgenden Punkte von Interesse:

- An welcher Stelle ist ein Defekt aufgetreten. Es geht hierbei um die physikalische wie auch um die logische Lokation[3].
- Der eingegrenzte Fehlerpunkt muß vom Rest des Systems abgeschottet bzw. ausgegliedert werden, um nicht weitere Fehler zu verursachen.
- Das Netzwerk muß rekonfiguriert bzw. angepaßt werden und dabei sollten die ursprünglichen Funktionen möglichst (temporär) ersetzt werden.
- Fehlerhafte Teile müssen ausgetauscht oder repariert werden.

Natürlich spielt es eine große Rolle, welche Komponenten ausgefallen sind. Dies kann im einfachsten Fall der benötigte Strom sein, aber auch Netzwerkinterfacekarten sind denkbar, wie auch physikalische Störungen an einem Kabel bis hin zu fehlerhaften Konfigurationen. Bei der Behebung dieser Probleme kann das Netzwerkmanagementsystem durch zum Beispiel Lokalisation nur helfend eingreifen und nicht alle Probleme automatisch lösen. So sind in der Praxis weitere Hilfsmittel notwendig. Manche von ihnen sind bereits auf jeder Workstation oder Router vorhanden[4], andere sind zum Beispiel *Cable Scanner* oder Netzanalysatoren.

Schließlich soll ein Netzwerkmanagementsystem nicht nur Reaktionen bewirken, sondern auch einen Administrator agieren lassen. Es soll Unterstützung bieten und mögliche Ausfälle bereits frühzeitig erkennen lassen. Das heißt aktuelle Zustände sollten ständig überwacht, aufgezeichnet und kontrolliert werden. In entsprechenden Alarmen und Statistiken sind die Zustände zu visualisieren. Nach erfolgreicher Behebung von einem oder mehreren Fehlern muß überprüft werden, ob sich das Netz in einem konsistenten Zustand befindet oder neue Fehler zu erwarten sind. Dies wird auch als *Problem-Tracking* und *-Control* [Zenk97] bezeichnet.

3.3.2 Leistungsmanagement

Wie die Überschrift verrät, handelt es sich um das Handling von Performance und Durchsatz. Grundsätzlich reicht es nicht aus, ein Netzwerk in irgendeiner Form aufzurüsten, sondern es ist vielmehr von Interesse an welcher Stelle eine sehr hohe Netzwerklast entsteht und welche Ressourcen dafür zur Verfügung stehen. Punkte großer Last sind natürlich an Stellen unterdimensionierter Leitungen zu vermeiden. So muß sich ein Netzwerkadministrator überlegen an welcher Stelle zentrale Server zu plazieren sind, da dort gewöhnlich mit sehr hohem Datenaufkommen zu rechnen ist. Die richtige Anordnung von Server und Workstation kann den Gesamtdurchsatz des Netzwerkes optimieren und Last von den dazwischengekoppelten

[3] Es ist ein Leitungsausfall denkbar, dessen Ursachen in einem fehlerhaften Interface zu suchen sind.

[4] Gemeint sind *traceroute* und *ping*.

Routern, Switches bzw. Bridges (falls überhaupt noch vorhanden) nehmen. Provider haben allerdings nicht diese Möglichkeit. Sie müssen sich dagegen überlegen ob ihre Leitungen entsprechend den Bedürfnissen ausgelastet sind oder ob Kapazitäten geschaffen werden müssen.

Grundsätzlich müssen die Fragen nach der Größe der Auslastung und damit verbundene Engpässe und Antwortzeiten gestellt werden. Da Performance von Benutzern immer subjektiv empfunden wird, stellt sich auch die Frage ob der Gesamtdurchsatz sich noch immer in einem akzeptablen Rahmen befindet. Zur Beantwortung sind Mechanismen notwendig, die eine aussagefähige Sicht auf Durchsatz und Auslastung bietet. Eine notwendige Maßnahme ist das Festsetzen von Schwellwerten.

Ein wiederholtes Auftreten von Übertragungen bereits gesendeter Datenpakete[5] deutet natürlich auf einen Fehlerfall hin und sollte von dem in Kapitel 3.3.1 beschrieben Fehlermanagement erkannt werden. Um so bedenklicher ist es, wenn dadurch eine Einschränkung der Performance erkennbar sein sollte. Ein Netzwerkmanagementsystem sollte also eine Vielzahl von Komponeten überwachen können, um Veränderungen im Netz, verursacht durch eine wachsende Zahl von Benutzern, zu erkennen. Auf diesem Weg ist eine genaue Einschätzung der aktuellen und möglicherweise auch zukünftigen Auslastung möglich. Allerdings sollte eine Sicht auf das gesamte Netz zur Verfügung stehen, nur so können sinnvolle Entscheidungen über Defizite und notwendige Kapazitäten getroffen werden. Jedoch sollten subjektive Eindrücke nicht der Entscheidungsträger für die Bewertung der Performance sein. Eine Bewertung dessen und auch von Antwortzeiten sind durch Schwellwerte am besten charakterisiert.

3.3.3 Konfigurationsmanagement

Mit dem Konfigurationsmanagement wird von einem Managementsystem die Möglichkeit einer (oft zentralen) Konfiguration gefordert. Dies ist insofern von großer Bedeutung, daß sich die Kommunikations- und Datentransporttechnologien in rasanter Geschwindigkeit entwickeln. Hier ist nicht nur ein automatisiertes Einspielen neuer Konfigurationen von Interesse, auch ein Update von Softwarereleases von Treibern und Routerbetriebssystemen ist gefordert.

Speziell kann es sich beispielsweise um folgende Fälle handeln. Beim Eingliedern neuer Subnetzwerke in ein bereits bestehendes Netz sind neue Routen zu konfigurieren, dies bedeutet ein Update der Routing-Tables. Auch Übertragungsraten auf bestimmten Netzteilstrecken können verändert werden.

Es handelt sich also um ein Hinzufügen, Verändern, Pflegen oder Entfernen von Netzwerkteilen bzw. Komponenten. Hierfür muß gewährleistet sein, daß das Konfigurationsmanagement vollen Zugriff auf die Komponenten hat,

[5]Retransmission

um diese falls notwendig zurückzusetzen aber auch abzuschalten, denn nur so können oftmals die neuen Funktionen verfügbar gemacht werden.

Templates

Oft vorkommende Konfigurationen sollten in Form von *Templates* gespeichert sein, um diese für weitere Geräte zu nutzen. In diesem Fall sind nur bestimmte Attribute zu ändern und die Konfiguration kann geladen werden. Ebenso wichtig ist das Speichern aktueller Konfigurationen um eine Historie der Veränderungen zu kennen. Das kann zum Beispiel mittels *CVS*[6] zentral geschehen.

Ein weiterer Aspekt ist natürlich die Zugriffskontrolle. Nur authorisierte Personen sollten Veränderungen vornehmen können. Aspekte der Sicherheit sollen allerdings in Kapitel 3.3.5 besprochen werden.

3.3.4 Abrechnungsmanagement

Es ist denkbar, daß sich große, flächendeckende Netze mehrere Parteien und damit auch die Kosten teilen. Dies ist insbesondere dann der Fall, wenn es sich nicht um öffentliche Netze handelt. Wer beispielsweise Transfervolumen bei einem Provider oder Carrier einkauft, kann heute häufig mit pauschalen oder grob gestaffelten Preisen rechnen. Eine 100%ige, auf das letzte Byte genaue Abrechnung macht kein Anbieter solcher Leistung. Dennoch ist es wichtig zu wissen, wem (also welcher Partei) welches Datenvolumen zu berechnen ist.

Man[7] kann sogar noch weiter gehen. Ein gutes Abrechnungsmanagement kann die effiziente Nutzung des Netzwerkes aufzeigen, indem es Nutzungsgewohnheiten vermerkt. Damit können zukünftige Ausbaustufen anwenderfreundlich geplant werden.

Die Ermittlung der Nutzung, durch geeignete Parameter ausgedrückt, stellt also eine Bedarfsermittlung dar und verzeichnet Nutzungsgewohnheiten, respektive Datenaufkommen einzelner Nutzer und Nutzergruppen. Die erzeugten Werte über die Nutzer sind insofern vertrauliche Daten, die entsprechende Autoritäten erfordern und entspechend zu behandeln sind.

3.3.5 Sicherheitsmanagement

Last but not least, das Sicherheitsmanagement umfaßt allgemein die Überwachung des Zugriffs auf Ressourcen des Netzes nach festgelegten Regeln. Das beinhaltet die Authentisierung von Nutzern zu bestimmten Diensten[8] wie auch die Zugriffskontrolle von Mitarbeitern zu bestimmten Komponenten und Ressourcen. Die Vorstellung über die Reichweite des Sicher-

[6] CVS: Concurrent Versions System

[7] Insider: Ein Rhetorikseminar würde mich an dieser Stelle eines besseren belehren, es müßte nicht mal auf Korsika sein...

[8] Gemeint sind Services im Sinne von Mail als Dienst etc..

heitsmanagements kann noch umfangreicher sein, denkbar wäre auch die Verwaltung von Zugangsdaten.

Grundsätzlich soll jedoch durch ein Sicherheitssystem Sabotage ebenso verhindert wie kontrollierter Zugang gewährt werden. Dies muß sich nicht nur im Bereich von Paßwörtern, Schlüsseln und anderen Zugangskontrollen abspielen, auch der Einsatz von Access-Listen gehört dazu und der restriktive Zugriff auf das Managementsystem selbst.

Zusammenfassend läßt sich sagen, daß das OSI-Management eine sehr allgemeine Beschreibung liefert, um ein Netzwerkmanagement zu definieren, und so nie implementiert wurde[9]. Allerdings hat dies bis heute in diesem Bereich einen Standard geschaffen, der in kommerziellen Systemen partiell adaptiert und implementiert wurde, was in Beispielen später beschrieben werden soll.

Eine andere Definition führte [Wagn97] ein, die hier zwar Erwähnung findet aber nicht mehr beschrieben werden soll, da sie inhaltlich dem OSI-Ansatz gleicht:

- traffic management (route planning)
- repair & maintenance
- (capacity) planning
- design & engineering / construction
- administration (billing, marketing etc.)

[9] Im Gegensatz zum ISO/OSI-7-Schichtenmodell. Leider, in diesem Fall, da der Ansatz ein sehr vollständiger und nützlicher ist.

Kapitel 4

Auslastungsanalyse

Einführung in die Analyse von Zeitreihen, Beschreibung des Programmes, Zweck, Funktion, Programmierung

4.1 Analyse von Zeitreihen

Zeitreihen wurden schon sein Jahrhunderten beobachtet und registriert. Dabei spielten eher praktische Bedürfnisse eine Rolle als die allgemeine wissenschaftliche Betrachtung. Die Entwicklung der statistischen Theorie begann erst in unserem Jahrhundert. Zuerst wurde in einzelnen Fachgebieten daran gearbeitet, so daß auch hier noch praktische Bedürfnisse eine große Rolle spielten. So wurden zum Beispiel in der Physik, Meterologie, Klimatologie wie auch in der Ökonomie solche Betrachtungen zuerst angestellt. Die Frage nach Trend und Saisonschwankung waren ebenso bedeutsam wie nach Periodizitäten.

Die richtige Interpretation der Ergebnisse solcher Betrachtungen hängt stark vom Verständnis der theoretischen Aspekte ab. So ist es oft nicht einfach, Theorie und Praxis zu verbinden.

Ebenso sind verschiedene Randbedingungen zu beachten. So kann es sein, daß nicht ausreichend viele Ergebnisse für die Betrachtung eines Ereignisses vorliegen. Auch das gewählte Modell zur Betrachtung der Daten spielt eine große Rolle. So sind natürlich Meßungenauigkeit und systematische Fehler zu vermeiden. Besondere Schwierigkeiten kann die Wahl der Zeitintervalle und die Größe des Stichprobenumfanges bereiten.

4.1.1 Beispiele und Behandlung von Zeitreihen

4.1.1.1 Beispiele von Zeitreihen

In diesem Kapitel soll nicht die gesamte Menge der statistischen Mittel und Verfahren beschrieben werden. Vielmehr soll ein Überblick über die

Schwierigkeiten aber auch Möglichkeiten der Betrachtung der bei einem Internetprovider anfallenden Daten gegeben werden. Dazu sollen einführende Beispiele aus anderen Fachgebieten Analogien zu im Netzwerk anfallenden Daten aufzeigen.

Ein Beispiel für physikalische Zeitreihen tritt besonders in der Meteorologie auf. Untersucht werden zum Beispiel Regenfall an aufeinanderfolgenden Tagen, Lufttemperaturen, Luftdrücke oder anderen Klimadaten. Bild 4.1 zeigt die mittleren Lufttemperaturen in München Riem an aufeinanderfolgenden Monaten.

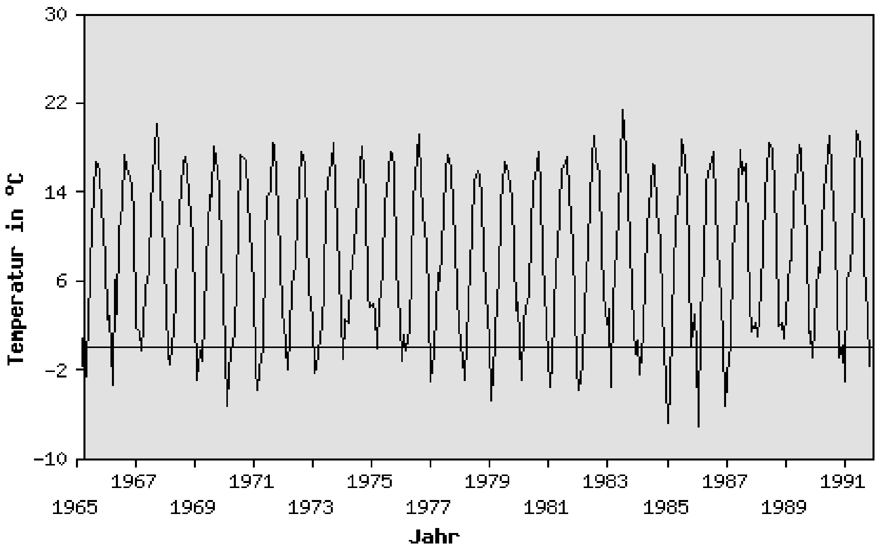


Abbildung 4.1: Mittlere Lufttemperaturen in München Riem an aufeinanderfolgenden Monaten [Wett99]

Ein Beispiel für eine andere Zeitreihe entstammt der Marktforschung. Um eine Produktion zu planen und die Entwicklung der Verkaufszahlen zu erkennen, werden Daten gesammelt anhand derer Tendenzen erkennbar sind. Im Bild 4.2 ist zum einen die steigende Tendenz zu erkennen aber auch die immer größer werdenden Schwankungen der verkauften Stückzahlen. Ähnliche Erscheinungen sind auch in anderen Bereichen erkennbar.

Zeitreihen spielen aber auch in der Demographie bzw. Epidemiologie eine wichtige Rolle. Heute sind Aussagen über das Bevölkerungswachstum oder Mortalitätsraten in der industrialisierten Welt sehr einfach, da der Staat jeden Menschen registriert.

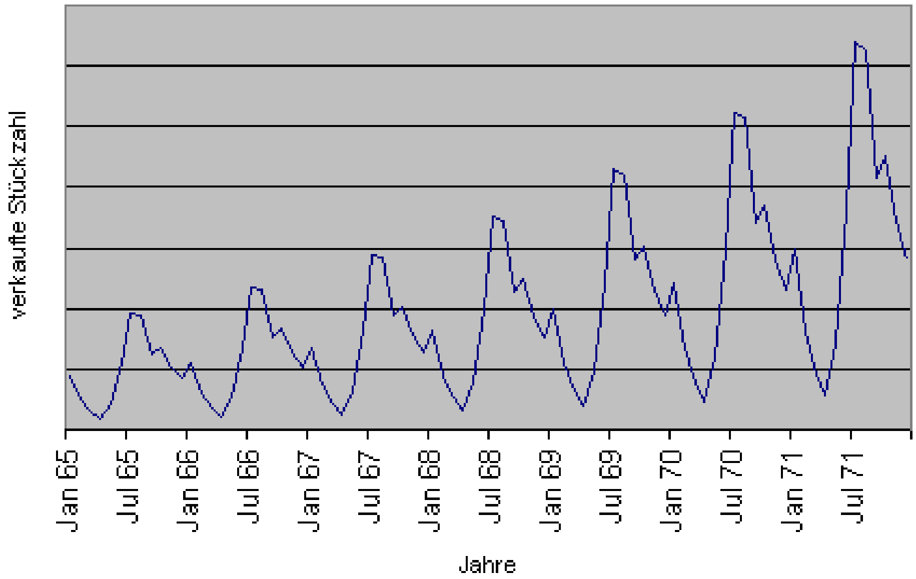


Abbildung 4.2: Beispiel für die Stückzahlentwicklung einer Produktion

4.1.1.2 Behandlung von Zeitreihen

Bei der Betrachtung von Zeitreihen spielen die Gesichtspunkte zur Analyse eine große Rolle für die Beurteilung der Daten. Der erste Schritt ist die zeichnerische Darstellung, anhand derer erste Betrachtung vorgenommen werden können. So kann ein grober Überblick über den Verlauf und die Eigenschaften einer Zeitreihe gewonnen werden. Wesentliche Beschreibungen sind Trends, Saisonschwankungen aber auch Ausreißer, die zu ermitteln sind. Besonders Ausreißer lassen sich oft mit statistischen Mitteln nur schwer erkennen und handhaben. Beschreibung

Werden beispielsweise mehrere Merkmale gleichzeitig beobachtet, kann möglicherweise die Variation einer Zeitreihe durch die Variation einer anderen erklärt werden. Ein Beispiel ist ein lineares System, dessen Eingangs- und Ausgangsgröße voneinander abhängen, also das System die Ausgangsgröße in Abhängigkeit von dem Eingangswerten verändert. Ein praktisches Bespiel ist die Verbindung des Traffic-Loads eines Routers und dessen CPU-Auslastung. Je mehr Verkehr ein Router transportiert, um so höher wird die CPU Auslastung sein. Erklärung

Aus einer Reihe von Daten sollen künftig Werte vorhergesagt werden. In der Literatur werden dabei die englischen Begriffe *prediction* und *forecasting* in gleicher Bedeutung benutzt. Andererseits wird *forecast* auch als ein Blick in die Zukunft bezeichnet und *prediction* hingegen für das systematische Vorgehen verwendet. Vorhersage

Steuern und Regeln ist eng verwandt mit Vorhersagen. Bei statistischen

Qualitätskontrollen wird nur in einen Produktionsprozess eingegriffen, wenn Abweichungen bei Beobachtungen auftreten. Ein mögliches Modell kann zum Beispiel wie folgt realisiert werden. Eine gemessene Zeitreihe wird an ein stochastisches Modell angepaßt, welches ab diesem Zeitpunkt zukünftige Werte vorhersagt. Dies kann als Maß für Abweichungen benutzt werden aber auch die Eingangsgrößen können so verändert werden, daß die Zielvorstellungen erfüllt werden.

4.1.2 Terminologie

Eine Zeitreihe wird kontinuierlich genannt, wenn die Beobachtung kontinuierlich in der Zeit gemacht wird. Ein Beispiel dafür ist die kontinuierliche Erfassung eines digitalen Signals, dessen 0/1 Werte zu jeder Zeit klar definiert sind. Im Gegensatz dazu wird eine Zeitreihe als diskret bezeichnet, wenn die Beobachtung nur zu bestimmten Zeitpunkten vorgenommen werden. Dies ist zum Beispiel in Netzwerken fast immer der Fall. So behandelt dieses Kapitel ausschließlich diskrete Zeitreihen.

kontinuierlich

diskret

Die Entstehung der Zeitreihen ist unterschiedlich. Aus einer kontinuierlichen Zeitreihe können in gleichen Zeitabständen Werte abgelesen werden (sampled series). Eine andere Form ergibt sich durch über gleiche Zeitintervalle aufsummierte Werte. Dies ist zum Beispiel der Fall, wenn Accountingdaten verschiedener Interfaces von Routern ausgelesen werden. Der Zähler eines solchen Interfaces zählt kontinuierlich jedes Paket, allerdings wird der Stand nur in bestimmten Intervallen abgefragt.

Ein großer Teil der statistischen Theorie beschäftigt sich mit Zufallsstichproben aus unabhängigen Beobachtungen. Die Eigenart der Zeitreihen ist aber, daß aufeinanderfolgende Beobachtungen nicht voneinander unabhängig sind. Bei der Betrachtung muß also die zeitliche Reihenfolge Beachtung finden. Würde man solche Beobachtungen als unabhängige betrachten, kann aus den vorangegangenen keine Erkenntnis für künftige Ereignisse gewonnen werden.

Abhängigkeit aufeinanderfolgender Beobachtungen

Kann eine Zeitreihe exakt vorhergesagt werden, wird sie als deterministisch bezeichnet. Jedoch sind die meisten Zeitreihen stochastisch. Eine Aussage über die Zukunft kann nur bedingt aus den Vergangenheitswerten getroffen werden. Für stochastische Zeitreihen ist eine exakte Vorhersage unmöglich. Zukünftige Werte haben eine Wahrscheinlichkeitsverteilung, welche durch die Kenntnis der vorangegangenen Werte bedingt ist.

deterministisch

stochastisch

4.1.3 Zeitreihenanalyse, Verfahren

Die Verfahren der Analayse der Zeitreihen ist sehr umfangreich. Sie beinhalten einfache beschreibende Verfahren, die Ausschau nach Trends und Saisonschwankungen. Ebenso stehen Wahrscheinlichkeitsmodelle für die Analyse von Zeitreihen zur Verfügung. Auch Verfahren zur Anpassung der Modelle an die Zeitreihen sind von Bedeutung. Das wesentliche diagnostische

Hilfsmittel ist die Autokorrelationsfunktion. Sie beschreibt die Entwicklung eines Prozesses in der Zeit, was oft als Analyse im Zeitbereich bezeichnet wird. Die Analyse im Frequenzbereich, die Spektralanalyse, also die Schätzung der Spektraldichtefunktion, welche die Gesamtvariation einer Zeitreihe durch die zyklischen Komponenten bei verschiedenen Frequenzen beschreibt, ist für diese Arbeit weniger von Interesse. Vielmehr sind Vorhersagemethoden von Interesse, die den möglichen weiteren Verlauf der Zeitreihen extrapolieren.

4.2 Vorhersageverfahren

Die Entscheidung, welche Verfahren einzusetzen sind, ist nicht immer einfach und erfordert einen Vergleich verschiedener Verfahren. Im folgenden sollen anhand der Menge der zur Verfügung stehenden Daten bezüglich der Zeitachse verschiedene Verfahren erläutert werden.

4.2.1 Kurzfristige Vorhersageverfahren

Zu den kurzfristigen Vorhersageverfahren [Gahs71] gehören die einfache und gleitende Mittelwertberechnung sowie die exponentielle Glättung erster Ordnung. Die letzten beiden liefern ähnlich gute Ergebnisse im Vergleich zu einfachen Mittelwerten.

4.2.1.1 Mittelwertbildung

Das typische Vorgehen bei der Berechnung des Mittelwertes ist einfach. Eine begrenzte Anzahl von Ausgangsdaten mit untereinander gleich großer Gewichtung wird zur Berechnung herangezogen und der Wert, der allen am nächsten ist, ermittelt.

4.2.1.2 Einfacher Mittelwert

Der einfache Mittelwert ist der Durchschnitt aller Zahlen, in dem diese aufaddiert und anschließend durch die Anzahl der herangezogenen Werte dividiert wird. Jedem Element wird die gleiche Bedeutung bzw. Gewichtung beigemessen. Bei konstantem Verlauf der Daten kann dieser Wert als Vorhersagewert dienen, wobei m der Anzahl der Vergangenheitswerte $A_1, A_2, \ldots, A_n$ entspricht.

$$A = \frac{1}{n} \sum_{i=1}^{n} A_i \tag{4.1}$$

Sollen nun $A_{n+1}, A_{n+2}, \ldots, A_{n+m}$ vorhergesagt werden, haben alle den Wert A:

$$A_{n+1} = A_{n+2} = \ldots = A_{n+m} = A \tag{4.2}$$

4.2.1.3 Gleitender Mittelwert

Die dem gleitenden Mittelwert zugrundeliegende Idee ist ebenfalls die einfache Mittelwertbildung, allerdings wird das Rechenverfahren mehrfach angewandt. Die Bildung des Mittelwertes wird für eine bestimmte Anzahl k von Perioden angewand, wobei das wesentliche Merkmal der Mittelwertbildung erhalten bleibt. Allerdings geht der statische Charakter in einen dynamischen über. Das bedeutet, daß eine Zeitreihe eine Glättung erfährt, indem für jeden Zeitpunkt i eine Anzahl von Werten zur Berechnung herangezogen werden. Auch wenn beim gleitenden Mittelwert, im Gegensatz zur exponentiellen Glättung, die Ausgangselemente gleichmäßig gewichtet werden, liegt die wesentliche Eigenschaft in der Glättung von Zeitreihen. Allerdings können damit ähnliche Eigenschaften erziehlt werden.

Es seien m Vergangenheitswerte $A_1, A_2, \ldots, A_m$ bekannt. Für k kleiner als m soll das Mittel über k Werte gebildet werden, so daß die gleitenden Mittelwerte wie folgt lauten:

$$M_1 = \frac{1}{k} \sum_{j=1}^{k} A_j$$

$$M_2 = \frac{1}{k} \sum_{j=2}^{k+1} A_j$$

$$\vdots$$

$$M_i = \frac{1}{k} \sum_{j=i}^{k+i-1} A_j$$

Damit ist die Anzahl der Mittelwerte auf $m-k+1$ begrenzt. Daraus ergibt sich folgende Formel:

$$M_i = \frac{1}{k} \sum_{j=i}^{k+i-1} A_j \quad \text{mit} \quad i = 1, 2, \ldots, m-k+1 \tag{4.3}$$

Damit zeigt sich im folgenden konstruierten Beispiel, das die Streuungsbreite für größere k verbessert, also die Glättung besser ist.

	i, j	1	2	3	4	5	6	7	8
	A_j	57	56	55	58	56	57	59	57
M_i	3	56,00	56,33	56,33	57,00	57,33	57,67	-	-
mit	4	56,50	56,25	56,50	57,50	57,25	-	-	-
k =	5	56,40	56,40	57,00	57,40	-	-	-	-

Tabelle 4.1: Beispiel für die einfachste Berechnung des gleitenden Mittelwertes

Die Streuungsbreite beträgt für

$$k = 3 \quad 57,67 - 56,00 = 1,67$$

$$k = 4 \quad 57,50 - 56,25 = 1,25$$

$$k = 5 \quad 57,40 - 56,40 = 1,00$$

Dies zeigt die bessere Glättung bei größerem k.

Für eine echte Glättung sollte allerdings der Mittelwert M_i nicht auf $i = 1$ bezogen werden, sondern auf die Mitte von $1, 2, \ldots, k$ der Werte, also auf $\frac{k+1}{2}$. Dafür ergibt sich (ungerade k vorausgesetzt) die folgende Formel:

$$M_i = \frac{1}{k} \sum_{j=i-l}^{k+i-l-1} A_j \quad \text{mit} \quad i = k-l, \ldots, m-l \quad \text{und} \quad l = \frac{k-1}{2} \tag{4.4}$$

Für eine Vorhersage aus dem gleitendem Mittelwert kann nun unter den folgenden Voraussetzungen gemacht werden:

- Die Vorhersage für die laufende Periode soll aus einer ungeraden Anzahl von Werten aufgebaut werden. Liegen diese selbst schon in der Zukunft, sind die entsprechenden Prognosewerte heranzuziehen.
- Alle Vergangenheitswerte sollen die gleiche Gewichtung haben.

Dies bedeutet eine weitere Rechtsverschiebung, so daß gilt:

$$M_i = \frac{1}{k} \sum_{j=i-1}^{i-k} A_j \quad \text{mit} \quad i = k, \ldots, m \tag{4.5}$$

Sollen n Werte $A_{m+1}, \ldots, A_{m+n}$ vorhergesagt werden, dann sind sie aus

$$A_i = \frac{1}{k} \sum_{j=i-1}^{i-k} A_j \quad \text{mit} \quad i = m+1, \ldots, m+n \tag{4.6}$$

zu ermitteln. Für $m = 10$, $n = 7$ und $k = 3$ zeichnet sich folgendes Bild ab. Je weiter i fortschreitet, desto mehr nähert sich der Vorhersagewert einem Grenzwert. Solche Verfahren sind bei der Verwendung von Computern zur

i, j	8	9	10	11	12	13	14	15	16	17
A_i, A_j	57	57	56	56,67	56,56	56,41	56,55	56,51	56,49	56,52

Tabelle 4.2: Beispiel für die Vorhersage mit gleitendem Mittelwert

Berechnung sehr gut zu handhaben, da sie in einfachen Schleifen realisierbar sind. So ist, wie in Tabelle 4.2 zu sehen, eine Realisierung durch die laufende Speicherung von k Vergangenheitswerten und einem Updating bei Periodenwechsel möglich.

4.2.1.4 Exponentielle Glättung erster Ordnung

Das Verfahren zur exponentiellen Glättung [Gahs71] hat verschiedene Vorteile. Ein wichtiger Punkt ist, daß die Vergangenheitswerte nicht mit gleicher Gewichtung in die Rechnung einfließen. Das heißt, daß ältere Werte eine geringere Gewichtung erhalten als neuere. Die Gewichtung der Werte nimmt in der Vergangenheit exponentiell ab. Weiterhin ist der Speicherbedarf und die Rechenzeit des Verfahrens relativ gering und das Verfahren arbeitet mit einem sich selbst regulierendem Lerneffekt.

Der jüngste Vergangenheitswert hat die Gewichtung α mit $0 \leq \alpha \leq 1$. Alle anderen Vergangenheitswerte werden durch den exponentiellen Mittelwert und der Gewichtung $1 - \alpha$ vertreten.

Wenn
E_i = exponentieller Mittelwert zum Zeitpunkt i,
α = Glättungsfaktor und
I_i = Istwert zum Zeitpunkt i
ist, dann ergibt sich die Vorhersageformel zu

$$\text{Neuer Mittelwert} = \alpha * \text{alter Istwert} + (1 - \alpha) * \text{alter Mittelwert}$$

Durch mehrfaches Einsetzen, die sich durch die Zeitverschiebung ergibt, gelten folgende Beziehungen:

$$E_i = \alpha * I_{i-1} * (1 - \alpha) * E_{i-1} \tag{4.7}$$

$$E_{i-1} = \alpha * I_{i-2} * (1 - \alpha) * E_{i-2}$$

$$E_{i-2} = \alpha * I_{i-3} * (1 - \alpha) * E_{i-3}$$

$$\vdots$$

$$E_i = \alpha * I_{i-1} + \alpha(1 - \alpha) * I_{i-2} + \alpha(1 - \alpha)^2 * I_{i-3} \ldots \tag{4.8}$$

oder als Summe

$$E_i = \alpha \sum_{j=1}^{i-1} (1 - \alpha)^{j-1} * I_{i-j} \tag{4.9}$$

Der erste Wert des exponentiellen Mittelwertes muß vorgegeben bzw. berechnet werden. Er kann auch empirisch angenommen werden als der erste Vergangenheitswert oder der arithmetische Mittelwert aller Vergangenheitswerte. Außerdem ist die Summe der Gewichtungsfaktoren 1. Um die Vergangenheit weitestgehend unberücksichtigt zu lassen, muß das α recht hoch gewählt werden. Für $\alpha = 0$ spielt der Vergangenheitswert keine Rolle, es gilt: $E_i = E_{i-1}$. Für $\alpha = 1$ ist der letzte Vergangenheitswert maßgebend, es gilt hier: $E_i = I_{i-1}$.

Die Glättungskurve mit $\alpha = 0,5$ reagiert viel schneller auf eine Störung bzw. Ausreißer, schwingt sich dafür auch schneller wieder ein als mit $\alpha = 0,1$. Ähnlich ist das Verhalten bei einem einmaligen Sprung der Wertereihe. Auch hier nähert sich mit $\alpha = 0,5$ die Kurve schneller dem angestrebten

Vergangenheit	Gewichtung		
j	$\alpha = 0,1$	$\alpha = 0,2$	$\alpha = 0,5$
1	0,100	0,200	0,500
2	0,090	0,160	0,250
3	0,081	0,128	0,125
4	0,073	0,102	0,063
5	0,066	0,082	0,031
6	0,059	0,066	0,016
7	0,053	0,052	0,008
8	0,048	0,042	0,004
9	0,043	0,034	0,002
10	0,043	0,034	0,002

Tabelle 4.3: Gewichtung bei verschiedenen Glättungsfaktoren

Wert. Nur bei kontinuierlichen gleichmäßigen Sprüngen versagt die einfache exponentielle Glättung gänzlich, da hier kein Trend berücksichtigt wird. In der Praxis wird immer eine Kombination aus Störungen und Sprüngen auftreten, wobei der Schwerpunkt auf den Störungen liegt.

4.2.1.5 Exponentielle Glättung erster Ordnung mit Trend

Das Verfahren zur exponentiellen Glättung mit Trend [Gahs71] berücksichtigt einen linearen Trend bei der Berechnung. Unterliegen die Istzahlen einem Trend, so laufen die Glättungswerte bei der exponentielle Glättung ohne Trend immer der Trendentwicklung hinterher. Der Grund liegt in der Gewichtung der Vergangenheitswerte. Je weiter die Werte in der Vergangenheit liegen, um so mehr weichen sie vom Vorhersagewert ab. Hierfür wird eine Trendkomponente eingeführt.

$$T_i = i * T$$

Der Trend T ist ein zeitabhängiger und periodenweise konstanter Betrag, so daß gilt:

$$I_{i-j} = I_i - j * T \tag{4.10}$$

Wenn nun 4.9 in 4.10 eingesetzt wird, ergibt sich

$$E_i = \alpha \sum_{j=1}^{i-1} (1-\alpha)^j * (I_i - j * T) \tag{4.11}$$

oder

$$E_i = \alpha * I_{i-1} \sum_{j=0}^{i-1} (1-\alpha)^j - \alpha * T \sum_{j=0}^{i-1} j * (1-\alpha)^j \tag{4.12}$$

Unter der Voraussetzung, daß der Index bis ∞ geht, gilt

$$\alpha * \sum_{j=0}^{\infty} (1-\alpha)^j = \frac{\alpha}{1-(1-\alpha)} = 1$$

und

$$\alpha * \sum_{j=0}^{\infty} j * (1-\alpha)^j = \frac{1-\alpha}{\alpha}$$

Durch Einsetzen in 4.12 folgt

$$E_i = I_{i-1} - \frac{1-\alpha}{\alpha} * T \tag{4.13}$$

Der Wert $\frac{1-\alpha}{\alpha} * T$ charakterisiert die Größe, um welche die Vorhersage hinterherläuft. Anders ausgedrückt, zur Vorhersage der nächsten Periode muß der letzte Trendbetrag addiert werden, d.h.

$$E_i = I_{i-1} + \frac{1-\alpha}{\alpha} * T + T = I_{i-1} + \frac{1}{\alpha} * T \tag{4.14}$$

$\frac{1}{\alpha} * T$ ist der Wert, linearen Trend vorausgesetzt, um den der Istwert bei Verwendung der exponentiellen Glättung erster Ordnung der Vorhersage hinterherläuft.

Damit kann eine Vorhersage mit Trend berechnet werden, wenn diese Korrektur berücksichtigt wird:

$$V_{i+1} = E_{i+1} + \frac{1}{\alpha} * T_i \tag{4.15}$$

Dabei sind
V = vorhergesagte Werte,
E = mit exponentieller Glättung erster Ordnung gemittelte Istwerte (mit Faktor α),
T = mit exponentieller Glättung erster Ordnung gemittelte Trendkomponenten (mit Faktor β).

Die Werte E_{i+1} und T_i sind die mit exponentieller Glättung erster Ordnung ermittelten Ist- und Trendwerte. Die Istwerte werden mit dem Glättungsfaktor α berechnet, der dem gleichen Faktor in 4.15 entspricht.

$$E_{i+1} = \alpha * I_i * (1-\alpha) * E_i \tag{4.16}$$

Dies entspricht der Gleichung 4.7, sie ist allerdings um eine Zeiteinheit verschoben.

Unter einem Trend T wird die Differenz zwischen zwei Periodenwerten verstanden, d.h. auf zufällige Schwankungen wird keine Rücksicht genommen.

$$T_i = \beta * (I_i - I_{i-1}) + (1-\beta) * T_{i-1} \tag{4.17}$$

β ist der Glättungsfaktor für die Trendkomponente.

Damit ergeben die Gleichungen 4.15, 4.16 und 4.17 einen Algorithmus:

$$\boxed{\begin{gathered} T_i = \beta * (I_i - I_{i-1}) + (1-\beta) * T_{i-1} \\ E_{i+1} = \alpha * I_i * (1-\alpha) * E_i \\ V_{i+1} = E_{i+1} + \frac{1}{\alpha} * T_i \end{gathered}}$$

Die Rechnung beginnt bei $i = 2$, da wenigstens zwei Perioden zur Erstaufnahme von Daten vergangen sein sollten. Diese Erstwerte sind I_1 und I_2. Ebenso gelten die Anfangswerte $T_1 = 0$ und $E_2 = I_1$.

Auch hier ist das Verfahren besonders gut zur Datenverarbeitung geeignet, da für eine Vorhersage nur jeweils der letzte Istwert I_{i-1}, die letzte gemittelte Trendkomponente T_{i-1} und der letzte gemittelte Wert E_i gespeichert sein müssen. Der Wert I_i ist der neue laufende Istwert, der die Mittelwerte T_i und E_{i+1} verändert.

Zur Vorhersage weiterer Perioden gilt dann

$$V_{i+1+j} = V_{i+1} + j * T_i$$

Ein ausführliches Beispiel kann [Gahs71] entnommen werden.

4.2.2 Mittelfristige Vorhersageverfahren

Die mittelfristigen Vorhersagen sollen Aussagen über Zeiträume von einem halben Jahr bis drei Jahren geben. Dazu gehören

- die exponentielle Glättung höherer Ordnung und
- die Regressionsanalyse.

Bei mittelfristigen Vorhersagen hat die Regressionsanalyse über die Minimierung der Abstandsquadrate Vorteile gegenüber der exponentiellen Glättung höherer Ordnung.

4.2.2.1 Exponentielle Glättung höherer Ordnung

Die exponentielle Glättung zweiter Ordnung enthält einen linearen Trendansatz und entspricht damit der exponentiellen Glättung erster Ordnung mit Trend. Zur Erinnerung, die exponentielle Glättung erster Ordnung glättet einen im allgemeinen konstanten Wert mit zufälligen Schwankungen. Die exponentielle Glättung dritter Ordnung beschreibt einen Parabelsatz zweiter Ordnung. Allgemein heißt das, daß die exponentielle Glättung n-ter Ordnung einen Parabel (n-1)-ter Ordnung beschreibt und damit der einfachste Fall, die exponentielle Glättung erster Ordnung eine Gerade darstellt.

Die Gleichung 4.7 beschrieb dieses Verhalten, wobei
M_i = Mittelwert zum Zeitpunkt i,
α = Glättungsfaktor und
I_i = Istwert zum Zeitpunkt i
ist, so ergibt sich die Formel zu

$$M_i = \alpha * I_{i-1} + (1 - \alpha) * M_{i-1} \tag{4.18}$$

Die Symbolisierung der ersten Ordnung erfolgt durch Einführung eines Index, so daß 4.18 sich zu

$$M_i^{(1)} = \alpha * I_{i-1} + (1-\alpha) * M_{i-1}{}^{(1)} \tag{4.19}$$

ergibt. Da die exponentielle Glättung zweiter Ordnung die Mittelwerte erster Ordnung glätten soll, werden die Mittelwerte zweiter Ordnung nach der Gleichung

$$M_i^{(2)} = \alpha * M_{i-1}{}^{(1)} + (1-\alpha) * M_{i-1}{}^{(2)} \tag{4.20}$$

berechnet. Das bedeutet allgemein für die exponentielle Glättung n-ter Ordnung folgendes:

$$M_i^{(n)} = \alpha * M_{i-1}{}^{(n-1)} + (1-\alpha) * M_{i-1}{}^{(n)} \tag{4.21}$$

Die Summe aus dem Mittelwert erster Ordnung und der Differenz zwischen den Glättungswerten erster und zweiter Ordnung entspricht der Vorhersage der exponentiellen Glättung zweiter Ordnung.

$$V_i^{(2)} = M_i^{(1)} + (M_i^{(1)} - M_i^{(2)}) = 2M_i^{(1)} - M_i^{(2)} \tag{4.22}$$

In Abbildung 4.3 wird ein linearer Trend vorrausgesetzt, der sich wie folgt

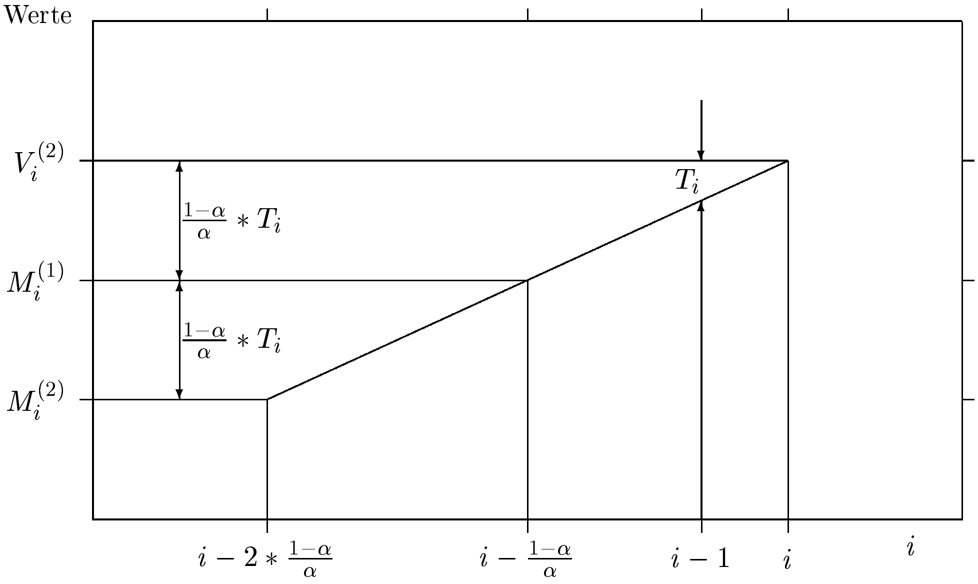


Abbildung 4.3: exponentielle Glättung zweiter Ordnung [Wett99]

berechnet:

$$\frac{1-\alpha}{\alpha} * T_i = M_i^{(1)} - M_i^{(2)}$$

also gilt

$$T_i = \frac{\alpha}{1-\alpha} * (M_i^{(1)} - M_i^{(2)})$$

So kann der periodische Trend T_i als Vorhersage verwendet werden; nach 4.22 ergibt sich:

$$V_{i+k}{}^{(2)} = 2M_i^{(1)} - M_i^{(2)} + k * T_i \tag{4.23}$$

So kann folgender Algorithmus zusammengefaßt werden:

$$M_{i+1}{}^{(1)} = \alpha * I_i + (1-\alpha) * M_i{}^{(1)} \quad \text{mit} \quad 1 \leq i \leq m$$
$$M_{i+1}{}^{(2)} = \alpha * M_i^{(1)} + (1-\alpha) * M_i{}^{(2)} \quad \text{mit} \quad 1 \leq i \leq m$$
$$T_{m+1} = \frac{\alpha}{1-\alpha} * (M_{m+1}{}^{(1)} - M_{m+1}{}^{(2)})$$
$$V_{m+k}{}^{(2)} = 2*M_{m+1}{}^{(1)} - M_{m+1}{}^{(2)} + k*T_{m+1} \quad \text{mit} \quad 1 \leq k \leq n$$

Mittels der ersten beiden Formeln werden m Vergangenheitswerte berechnet. Nachdem die Trendgröße T_i berechnet wurde, können durch die letzte Formel n Vorhersagewerte ermittelt werden.

Die Anzahl der Vorhersagewerte m hängt von den vorliegenden Bedingungen ab. Die Anzahl der Vorhersagewerte n kann beliebig gewählt werden. Dabei ist $M_1^{(1)} = M_1^{(2)} = I_1$ als Start zu wählen, die Rechnung beginnt bei $i = 2$.

Nach dem Fundamentaltheorem von Brown [Brow61] ist die Vorhersage bei einer exponentiellen Glättung j-ter Ordnung nach der Periode m

$$V_{m+1}{}^{(j)} = \sum_{i=1}^{j} \binom{j}{i} * (-1)^{i+1} * M_m^{(i)} \tag{4.24}$$

Das bedeutet für eine Vorhersage mit exponentieller Glättung dritter Ordnung:

$$V_{m+1}{}^{(3)} = 3 * M_m^{(1)} - 3 * M_m^{(2)} + M_m^{(3)} \tag{4.25}$$

4.2.2.2 Regressionsanalyse

Zur mittelfristigen Vorhersage ist die Regressionsanalyse besser geeignet als die exponentielle Glättung höherer Ordnung. Die Regressionsanalyse dient dazu, Schlüsse von einer Größe auf die andere zu ziehen, das heißt zwei Merkmale mit sogenannten bivariaten Verteilungen gleichzeitig zu beobachten. Probleme bei der Betrachtung entstehen durch zufallsbedingte Streuung. Zusammenhänge können durch diese einerseits vorgetäuscht werden, obwohl sie nicht existieren, andererseits kann ein vorhandener Zusammenhang überdeckt und somit nicht erkannt werden.

bivariaten Verteilungen

Der Korrelationskoeffizient macht Aussagen darüber, wie eng der lineare statistische Zusammenhang zwischen zwei Größen ist.

Korrelationskoeffizient

i / k	α	I_i	$M_i^{(1)}$	$M_i^{(2)}$	T_{m+1}	$V_{m+k}^{(2)}$
$i = 1$	0.2	121	121	121	-	-
2	0.2	123	121	121	-	-
3	0.2	124	121.4	121	-	-
4	0.2	127	121.9	121.1	-	-
5	0.2	130	122.9	121.3	-	-
6	0.2	129	124.3	121.6	-	-
7	0.2	135	125.2	122.1	-	-
8	0.2	137	127.2	122.7	-	-
9	0.2	136	129.2	123.6	-	-
10	0.2	142	130.6	124.8	-	-
11	0.2	142	132.9	126.0	-	-
$i = 12$	0.2	146	134.7	127.4	-	145.1
$k = 1$	0.2	-	137.0	128.9	2.0	147.1
2	0.2	-	-	-	-	149.1
3	0.2	-	-	-	-	151.1
4	0.2	-	-	-	-	153.1
5	0.2	-	-	-	-	155.1
6	0.2	-	-	-	-	157.1

Tabelle 4.4: Beispiel für die exponentielle Glättung zweiter Ordnung als mittelfristige Vorhersagemethode [Gahs71]

Die Regressionsgerade [Harm98] läßt sich als Schätzung von $y-$ aus $x-$Werten beschreiben. Mit der Annahme einer Punktwolke soll der lineare Zusammenhang in Form einer Geraden dargestellt werden. Für die Schätzung des $y-$Wertes sollte die Gerade so liegen, daß die Streuung in der $y-$ oder in vertikaler Richtung möglichst klein ist. Unter den vielen möglichen Geraden, die durch die Punktwolke skizzierbar ist, gibt es eine Gerade, die bei der Schätzung der $y-$Werte aus den $x-$Werten besonders günstig ist: die Regressionsgerade von y auf x. Es ist allerdings auch eine Darstellung der Regressionsgerade von x auf y denkbar, bei der die Streuung in horizontaler Richtung möglichst klein ist. Beide Geraden schneiden sich im Schwerpunkt der Punktwolke. Die Koordinaten des Schwerpunktes ergeben sich aus den Mittelwerten aller $x-$ und $y-$Werten: $\overline{x}$, $\overline{y}$.

Regressionsgerade

Die Regressionsgerade von y auf x stellt sich wie folgt dar:

$$\widehat{y} = b_{yx}x + a_{yx} \tag{4.26}$$

Die Steigung b_{yx} und der Achsenabschnitt a_{yx} werden so berechnet, daß die Gewichte y möglichst wenig von den Werten $\widehat{y}_i$ abweichen, die von der Regressionsgeraden als Durchschittsgewicht x_i angegeben werden. Dies bedeutet allgemein, die Summe aller Abstandsquadrate soll möglichst gering sein:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \quad \text{Minimum} \tag{4.27}$$

Genauer:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}(y_i - (b_{yx}x_i + a_{yx}))^2 \tag{4.28}$$

Wie sich mittels der Diffentialgleichung zeigen läßt, ergibt die Summe der Abweichungsquadrate ein Minimum, wenn für die Steigung b gilt:

$$b_{yx} = \frac{\sum(x_i - \overline{x})(y_i - \overline{y})}{\sum(x_i - \overline{x})^2} = \frac{\text{Summe der Abweichungsprodukte}}{\text{Summe der Abweichungsquadrate}} \tag{4.29}$$

Auf Grund des Wissens, daß die Regressionsgerade durch den Schwerpunkt $\overline{x}$, $\overline{y}$ verläuft, läßt sich der Achsenabschnitt a_{yx} errechnen, indem die Werte für x und b_{yx} in die Formel 4.26 eingesetzt werden. Aus

$$\overline{y} = b_{yx}\overline{x} + a_{yx}$$

folgt

$$a_{yx} = \overline{y} - b_{yx}\overline{x} \tag{4.30}$$

Wie bereits erwähnt, läßt sich auf dem selben Weg die Regressionsgerade von x auf y berechnen. Allerdings ist der Prädiktor, also die unabhängige Variable y, x ist in diesem Fall das Kriterium, die abhängige Variable.

Die genauere von beiden möglichen Geraden ist immer diejenige, welche die kleineren Abstandsquadrate aufweist, was wiederum vom Anstieg der Geraden abhängt.

Bedeutung der Regressionsgerade

Die Bedeutung der Regressionsgerade läßt sich sehr einfach darstellen. Die gefundene Gerade gilt zunächst nur für den berechneten, also untersuchten Definitionsbereich. Sofern die untersuchten Werte eine repräsentative Stichprobe aus allen möglichen Werten der Grundgesamtheit darstellen, gilt die Regressionsgerade für alle Werte, allerdings nur im Rahmen der zufallsbedingten Ungenauigkeit.

Die Annahme, daß die Abhängigkeit zwischen den Variablen x und y linear ist, ist häufig nur hypothetisch, so daß es nicht zulässig ist, den Wertebereich von x und y über den untersuchten Bereich hinaus auszudehnen bzw. zu extrapolieren. Wenn allerdings eine Vorhersage getroffen werden soll, kann dieser, unter der genannten linearen Einschränkung, durch Verlängerung der Geraden geschehen. Die Vorgehensweise ist wie bei den beschrieben Mittelwertbetrachtungen.

Bedeutung des Korrelationskoeffizient

Der Korrelationskoeffizient macht eine Aussage darüber, wie eng der lineare Zusammenhang zwischen $x-$ und $y-$Werten ist. Er kann Werte zwischen

+1 und −1 annehmen. Ein Korrelationskoeffizient von 0 bedeutet mathematisch keine lineare Korrelation, auch wenn mit dem Auge eine anderweitige Korrelation sichtbar ist.

Der Produktmoment-Korrelationskoeffizient nach Pearson [Gahs71], auch Maßkorrelationskoeffizient, ist immer dann gemeint, wenn von Korrelationskoeffizienten im allgemeinen gesprochen wird. Bei der Berechnung muß allerdings vorausgesetzt werden, daß

Produktmoment-Korrelationskoeffizient nach Pearson

- die Daten intervallskaliert sind,
- beide Variablen aus normalverteilten Grundgesamtheiten stammen und
- die Variablen in einer linearen Abhängigkeit zueinander stehen.

Das Ergebnis ist außerdem sehr stark von Ausreißern beeinflußt.

Der Produktmoment-Korrelationskoeffizient nach Pearson berechnet sich wie folgt:

$$r = \frac{\sum(x-\overline{x})(y-\overline{y})}{\sqrt{\sum(x-\overline{x})^2 * \sum(y-\overline{y})^2}} \tag{4.31}$$

Sind die oben genannten Voraussetzungen nicht erfüllt oder die Punktwolke ist von starken Ausreißern beeinflußt, sollte auf den Spearmansche Rankkorrelationskoeffizient zurückgegriffen werden [Gahs71]. Sind die Voraussetzungen für den Maßkorrelationskoeffizienten r erfüllt, sollten r und r_s fast identische Werte ergeben.

Der Spearmansche Rangkorrelationskoeffizient bietet bei nicht linearem monotonen Zusammenhang deutliche Vorteile. Er berechnet sich als

Spearmansche Rangkorrelationskoeffizient

$$r_s = 1 - \frac{6\sum D_i^2}{n^3 - n} \tag{4.32}$$

wobei n die Anzahl der Wertepaare und D_i die Rangdifferenz des einzelnen Wertepaares ist. Fallen ein oder mehrere Werte auf denselben Rangplatz, werden mittlere Ränge zugeordnet.

4.3 Umsetzung bei UUNET

Die zum bisherigen theoretischen Teil gehörige Analyse bei der Firma UUNET Deutschland[1] soll die gewonnenen Erkenntnisse anhand von Bandbreitenparametern verschiedener Router verdeutlichen. Welche Möglichkeiten existieren, wie eine Vorhersage bezüglich zukünftiger Auslastungen getroffen werden kann, sollen im folgenden beschrieben werden.

[1] Im folgenden nur UUNET genannt.

4.3.1 Programm

Da bei UUNET Wert auf Plattformunabhängigkeit gelegt wird und fast alle Applikationen webbasiert sind, war die Entscheidung ein Analysetool ebenfalls webbasiert zu entwickeln leicht. Es basiert auf den üblichen verfügbaren Mitteln (CGI[2]), die eine schnelle Implementierung zulassen und dennoch, bezüglich zusätzlicher Anforderungen, uneingeschränkt umfangreiche Manipulationen von Auswahl und Einstellung zuläßt. Die Sprache der Wahl ist Perl.

Bei den Bandbreitendaten handelt es sich um Angaben über Volumenflüsse auf bestimmten Interfaces von Routern. Das sind sowohl Bytes in einem mehr oder weniger festen Intervall, als auch die Bandbreitenangabe *kbps*. Auch die CPU- und Speicher-Auslastung der Router wurde in die Betrachtung mit einbezogen, da von dieser Größe die Funktion des Routings ebenso abhängt, wie die Auslastung[3] von Leitungen.

4.3.2 Zugriff auf Daten

Auf Grund der Fülle an Daten werden diese in einer Datenbank gespeichert. Der Zugriff erfolgt über eine quasi standardisierte Schnittstelle, DBI – Database Independent Interface – für die für alle gängigen Datenbanken Treiber zur Verfügung stehen. So ist gewährleistet, daß auch bei einer Umstellung auf eine andere Datenbank die Funktionalität der Programme erhalten bleibt.

4.3.3 Daten der Datenbank

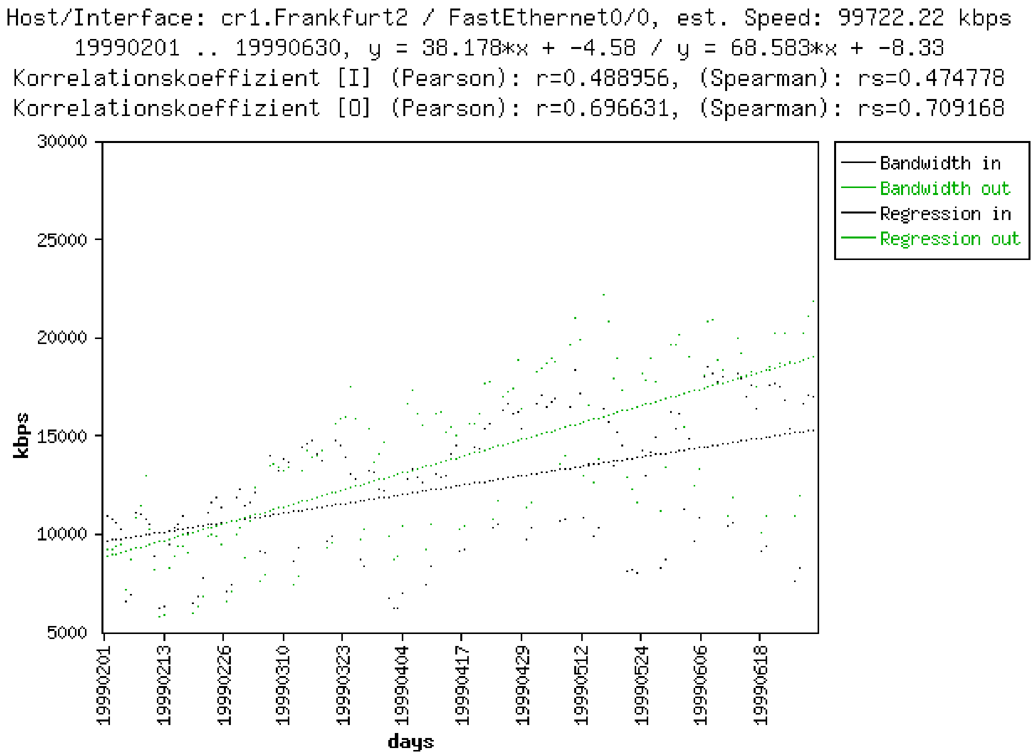
Da bei einer ungefähren Zahl von 500 Routern und 10000 Interfaces beim Speichern von Daten in einem fünf minütigen Intervall große Mengen an Daten anfallen, war es wichtig, die Daten zusammenzufassen, zu aggregieren. Der Vorteil ist, daß sehr viel Platz gespart wird, andererseits durch die Aggregation der Daten extreme Peaks der Mittelwertbildung zum Opfer fallen. Dies stellt allerdings insofern kein Problem dar, da für mittelfristige Beobachtungen die Entwicklung, also der Trend, eine Rolle spielt und kein Echtzeitmonitoring gefordert ist.

Die Daten werden in Stunden und Tagesintervallen gespeichert. Eine Auswertung über lange Zeiträume ist so einfach möglich, da zur Zeit der Abfrage keine Aggregierung mehr notwendig ist. Um den Nachteil, daß Peaks der Mittelwertbildung zum Opfer fallen, ausgleichen zu können, werden diese Werte (Minimum und Maximum) für die aggregierten Intervalle ebenfalls in der Datenbank gespeichert.

---

[2]Common Gateway Interface: Eine Schnittstelle zur Kopplung von Webservern mit externen Programmen.

[3]bei überlasteten Leitungen



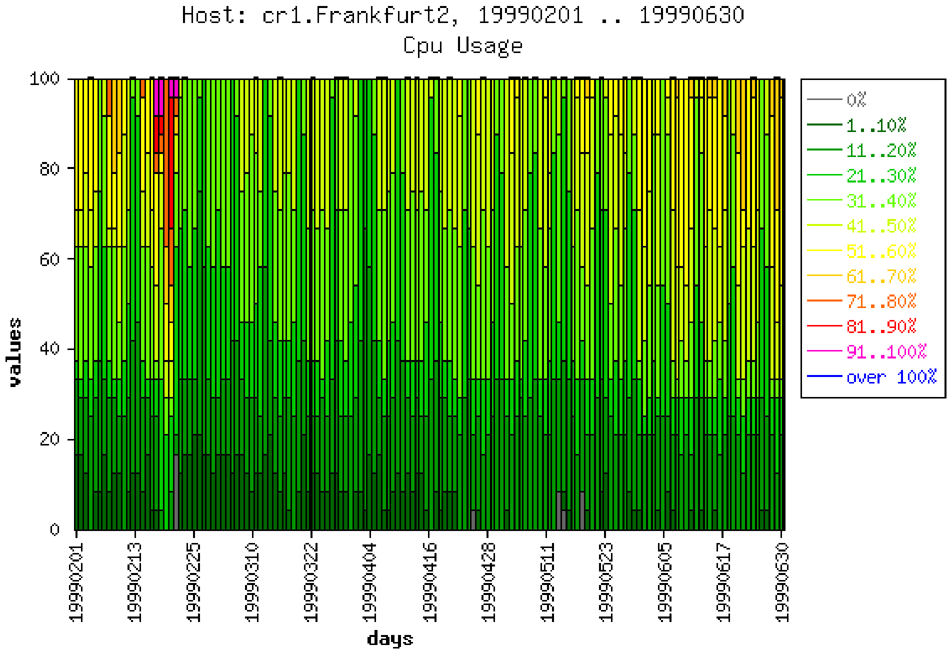


Abbildung 4.4: Beispiel für die Auslastung eines Interfaces und CPU-Auslastung des Routers

4.3.4 Funktionen

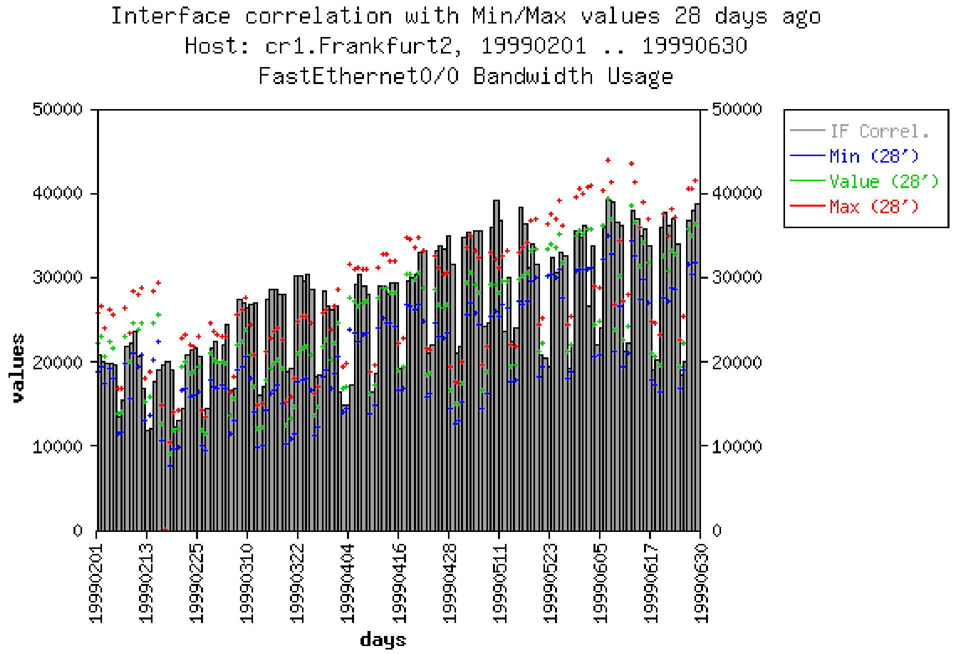
Zur Analyse der Daten wurde die Regressionsanalyse herangezogen, da sie für längerfristige Aussagen gut geeignet ist. Die Umsetzung erfolgte wie im vorangegangenen Kapitel beschrieben. Der Vorteil der Regressionsanalyse liegt darin, daß es ein Kontrollkriterium gibt, welches eine Aussage über die Güte der Ergebnisse macht. Eben dieses Kontrollkriterium, die Korrelationskoeffizienten, beschreibt, einen linearen Zusammenhang vorausgesetzt, ob die errechnete Regressionsgerade einen guten Wert für weitere Vorhersagen liefern kann.

Das Bild 4.4 stellt die Bandbreitenwerte für eingehenden und ausgehenden Traffic dar. Weiterhin sind die Regressionsgeraden, wie auch die Geradengleichungen zu sehen. Zur Kontrolle können die Korrelationskoeffizienten ausgewertet werden. In diesem speziellen Beispiel sind diese für den ausgehenden deutlich besser als für den eingehenden Verkehr. Das bedeutet, daß die Werte für den eingehenden Verkehr deutlich schlechter korrellieren, was offensichtlich an einer weniger harmonischen Entwicklung liegt. Anders ausgedrückt, die Werte des eingehenden Verkehrs sind deutlich mehr mit Ausreißern behaftet.

Ebenfalls ist in Bild 4.4 die CPU-Auslastung zu erkennen. Das Schaubild stellt die Auslastung in 100% dar. Alle Werte des jeweiligen Tages werden auf die entsprechenden Prozentualbereiche aufgeteilt.

Wie für die Interfaceauslastung eine Regressionsanalyse durchgeführt wurde, ist dies ebenfalls für den Verlauf der CPU-Auslastung möglich. Deutlich interessanter ist jedoch eine Darstellung wie in Bild 4.5 zu sehen ist. Diese Darstellung stellt den aktuellen Zustand und die jeweilig zugehörigen aktuellen Werte von vor vier Wochen dar (incl. der Minima und Maxima).

Ebenso interessant für einen Serviceprovider ist die Gegenüberstellung zueinandergehöriger Werte für ein- und ausgehenden Verkehr. Auch wenn dies keinen Forecastcharakter hat, ist zum Beispiel an Peeringpunkten bzw. Uplinks gut erkennbar, welchen Charakter ein Provider hat, in welche Richtung hauptsächlich der Verkehr fließt.



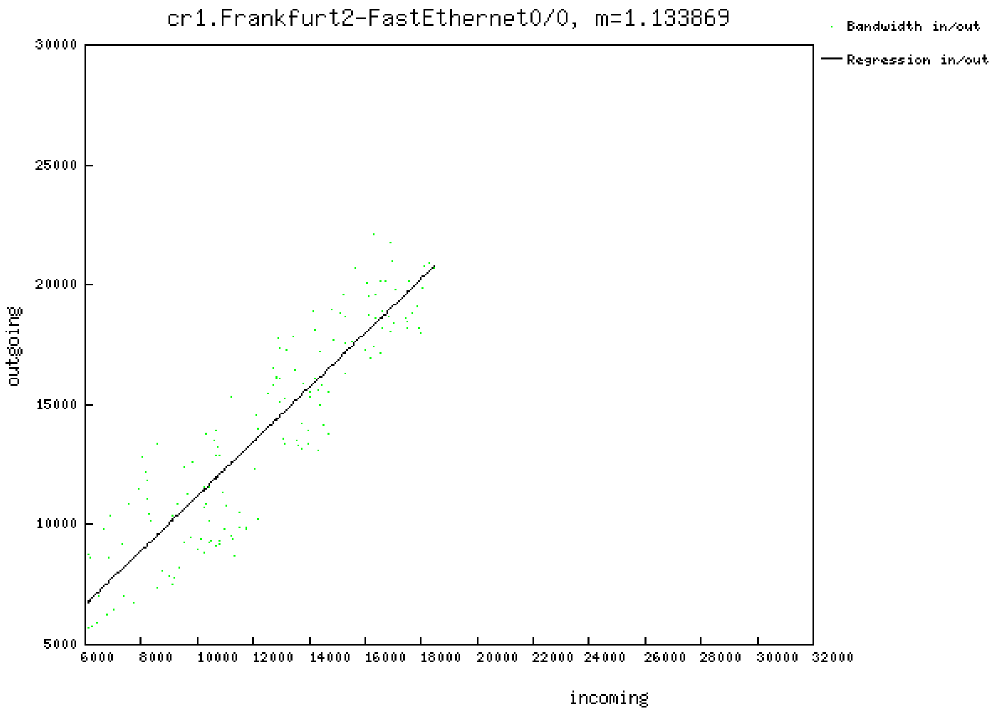


Abbildung 4.5: Beispiel für die Auslastung eines Interfaces mit Vergangenheitswerten, Gegenüberstellung von incoming und outgoing

Kapitel 5

Zusammenfassung, Perspektiven

Netzwerkmanagement beschreibt eine Vielzahl von Funktionen, Aufgaben und Problemen, und stellt einen Oberbegriff für das Management von Netzwerken im allgemeinen dar. Ein Netzwerk, welches sich für einen Anwender als transparentes Medium darstellt, ist für diejenigen, die die Funktionalität zur Verfügung stellen, nicht immer einfach zu handhaben. Je größer ein solches zu betreuendes Netzwerk ist, um so wichtiger ist es alle Funktionen steuern und überwachen zu können. Hier greift der Ansatz von Netzwerkmanagementsystemen. Sie sollen sowohl unterstützend als auch steuernd in solche komplexen Netzwerke eingreifen.

Ein Netzwerk kann als regelbares System betrachtet werden. Es gibt Kontrollgrößen, wie zum Beispiel der Bandbreitenverbrauch, dem ein Bedarf gegenüber steht. Die Kontrolle solcher Größen ist insofern notwendig, da sie eine Sicht auf das Netzwerk vermitteln, die mittels subjektiver Einschätzungen – das Netz ist zu langsam – nicht treffbar sind. Ein Anwender fordert die Übertragung von Informationen, nutzt ein Netzwerk zur Kommunikation. Auftretende Engpässe werden zwar von Anwendern schnell erkannt, jedoch hat eine solche Information wenig Gehalt, da von außen nicht erkennbar ist, an welcher Stelle Schwächen vorhanden sind.

An dieser Stelle kann ein entsprechendes Netzwerkmanagementsystem eingreifen. Es verschafft eine Übersicht über verfügbare Ressourcen und damit die Kontrolle. Fehler und Aktionen können erkannt, gesteuert aber auch verhindert werden, wenn frühzeitig Tendenzen erkannt werden. Nur so kann eine bestmögliche Dienstqualität für den Endbenutzer gewährleistet werden.

Wie bereits in Kapitel 3.1.2 beschrieben wurde, teilt sich das von der International Standard Organisation (ISO) definierte Netzmanagementmodell in folgende Bereiche:

- Fehlermanagement (*Faultmanagement*)
- Leistungsmanagement (*Performance Management*)
- Konfigurationsmanagement (*Configuration Management*)
- Abrechnungsmanagement (*Accounting Management*)
- Sicherheitsmanagement (*Security Management*)

Die zusammengefaßten Bereiche des Netzmanagementmodells werden allgemein auch außerhalb des OSI-Kontextes verwendet. Hersteller diverser Produkte implementieren diese Punkte oft nur partiell und in unterschiedlichen Protokollen, da die Vorgaben der International Standard Organisation sehr umfaßend formuliert wurden. Dennoch hat die Definition wegen ihres vollständigen Charakters ihre Berechtigung.

Netzwerkmanagement muß sich nicht nur auf IP-basierte Netzwerke beschränken. Bei den ständig wachsenden Anforderungen an Bandbreite im Internet gewinnt das Management von neuen aber auch älteren Technologien immer mehr an Bedeutung. So sind Managementsysteme nicht nur im IP-basierten Internet anzutreffen, sondern auch in ATM- und Telekommunikationsnetzen, also überall dort, wo die Größe des Netzes nicht auf einen Blick überschaubar ist und sich die Frage nach Verfügbarkeit stellt. Beim Einsatz von ATM ist sowohl das Verkehrsmanagement als auch das Leistungs- und Fehlermanagement von Bedeutung. In der Telekommunikationsindustrie ist das Management von Netzen und Diensten unter dem Begriff *OAM&P (Operations, Administration, Maintenance and Provisioning)* – Betrieb, Administration, Instanthaltung und Bereitstellung – bekannt.

Das Netzwerkmanagement in IP-basierten Netzen baut im Prinzip auf bekannte Komponenten auf. SNMP wird eingesetzt, um die Ressourcen abzufragen. Mittels der Management Information Base (MIB), die von den einzelnen Geräten unterstützt werden muß, können Accounting relevante Daten für das Abrechnungsmanagement ebenso gepollt werden, wie zum Beispiel Statusabfragen für das Fehlermanagment. Es können im Netz integrierte Geräte via SNMP komplett neu konfiguriert oder zum Beispiel nur einzelne Interfaces integriert werden.
Dies ist ein Beispiel für die Bedeutung des von der International Standard Organisation definierten Netzwerkmanagementmodells. Im Prinzip sind alle grundlegenden Punkte erfüllt, die Implementierung im Detail variiert allerdings von Hersteller zu Hersteller.

Der Schwerpunkt des praktischen Teils dieser Diplomarbeit lag in der Vorhersage der Entwicklung von Bandbreiten. Die Evaluierung, welche Verfahren sinnvoll und wie anzuwenden sind, nahm einen großen Teil der Zeit in Anspruch. Es wurden die Randbedingungen ermittelt um eine adäquate Aussage über notwendige Daten zu treffen.

Da bereits entsprechende Abrechnungsdaten zur Verfügung standen, konnte auf diesen aufgebaut werden. Allerdings mußten diese aggregiert werden, da sonst das Volumen vorhandene Kapazitäten gesprengt hätte. Dies stellt jedoch kein Problem dar. Für mittelfristige Vorhersagen werden keine Daten in kurzen Zeitintervallen benötigt, so daß die Daten durch eine Mittelung über Stunden und Tage noch ausreichende Informationen enthalten.

Die Speicherung der aggregierten Werte erfolgt in einer Datenbank. Abfragen können so einfacher realisiert werden als ein Zugriff über Flatfiles.

Zur Berechnung der Entwicklung von Bandbreitenbedürfnissen wurde die Regressionsanalyse herangezogen. Sie hat den Vorteil, daß sie für mittelfristige Entwicklungen hervorragende Aussagen macht und vor allem der exponentiellen Glättung höherer Ordnung vorzuziehen ist. Wie in Kapitel 4.2.2.1 beschrieben wurde, stellt die exponentiellen Glättung höherer Ordnung ein komplexeres Verfahren dar. Die Stärken der Regressionsanalyse liegt ebenfalls darin, daß eine Aussage über die Güte der berechneten Geraden getroffen werden kann. Dafür stehen zwei Korrelationskoeffizienten zur Verfügung.

Die Visualisierung der Ergebnisse erfolgt über ein webbasiertes Interface (CGI), zur Darstellung wird nur ein Webbrowser benötigt. Die plattformunabhängige Darstellung der Ergebnisse ist ein allgemeine Anforderung bei der Firma UUNET.

Warum wird in die Entwicklung von Netzwerkmanagementsystemen so viel Manpower gesteckt?

Auf dem Markt sind einige Produkte verfügbar, die den notwendigen Anforderungen an Netzwerkmanagementsystemen genügen. Die Funktion dieser ist generell so, daß die Daten von den Geräten, allgemein gesagt die Ressourcen, via SNMP gelesen werden. Die Speicherung erfolgt entweder in einer mitgelieferten oder auch generischen, bereits verfügbaren Datenbank. In der Auswertung der Daten unterscheiden sich die Produkte allerdings deutlich. Es gibt Lösungen, die auf proprietären, nicht standardisierten Applikationen aufsetzen. Andere Anbieter sind den Schritt zu plattformunabhängigen Ausgabeformaten bereits gegangen, damit sind webbasierte Lösungen gemeint so wie auch Implementierungen in Java.

Die Schwächen fertiger Produkte sind vielseitig. Wenn sich eine Firma im Netzwerk-, spezieller im Providergeschäft, zu etablieren versucht, ist das der richtige Zeitpunkt für die Integration eines Netzwerkmanagementsystems. Zu diesem Zeitpunkt stehen aber nicht genügend Mittel zur Verfügung, um eine der ausschließlich teuren Produkte zu installieren. Zu einem späteren Zeitpunkt sind solche Lösungen sicher bezahlbar. Allerdings sind dann die Anforderung an ein solches Produkt oft sehr speziell. Zum Beispiel kann der Wunsch auf bereits bestehende Accountingdaten aufzusetzen nicht realisiert werden, da die benötigten Tabellen der Managementsysteme andere Strukturen haben, anders organisiert sind.

Eine Umstellung auf ein solches Produkt würde einen großen Einsatz an Ressourcen bedeuten. Bei größeren Netzwerken können nicht alle Bedürfnisse abgedeckt werden. Der Zugriff auf die gesammelten Daten ist nicht transparent, da die Implementierung der Produkte nicht offengelegt wird. Das bedeutet, sollten parallel Inhouse-Lösungen installiert sein, müßten alle Daten doppelt gehalten werden. Unter anderem schreckte dies UUNET bis heute ab eine fertige Lösung zu installieren.

Eine mögliche Lösung für die Zukunft könnte die Ausgabe von statistischen Informationen im XML-Format sein. Dies würde bestehende grundlegende Prinzipien nicht brechen und dennoch einen offenen Ansatz darstellen.

Der Bedarf an Netzwerkmanagement ist gegeben. Verschiedene Lösungen sind im Einsatz. Welche die richtige Lösung ist, muß immer von Fall zu Fall entschieden werden.

Durch den Einblick in die praktische Arbeit bei der Firma UUNET Deutschland habe ich meinen Bildungshorizont erweitern können und die Erfahrung, an der Evaluierung und Implementierung einer eigenen Lösung arbeiten zu können, möchte ich nicht missen.

Index

Literaturverzeichnis

[Brow61] Brown, R. G., Meyer, R. F.: The fundamental theorem of exponential smoothing. In Operations Research, Vol. 9, No. 5, 1961

[Case90] Case, J.,Fedor, M., Schoffstall, M., Davin, J.: „A Simple Network Management Protocol (SNMP)“. RFC-1157, May 1990.

[Case90a] Case, J., Fedor, M., Schoffstall, M., Davin, J.: „A Simple Network Management Protocol (SNMP)“. RFC-1157, May 1990.

[Chim98] Chimi, Emmanuel: High-speed networking: Konzepte, Technologien, Standards. München; Wien: Hanser, 1998

[Cisc99] http://www.cisco.com/univercd/cc/td/doc/product/software/ios120/120newft/120t/120t3/snmp3.htm

[Cisc99a] http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito_doc/rmon.htm

[Davi87] Davin, J., Case, J., Fedor, M., Schoffstall, M.: „A Simple Gateway Monitoring Protocol“. RFC-1028, November 1987.

[Dude96] Duden, Der: Die deutsche Rechtschreibung. Dudenverlag, 21. Auflage, 1996.

[Gahs71] Gahse, S.: Mathematische Vorhersageverfahren und ihre Anwendung. Verlag Moderne Industrie, 1971.

[Haec97] Häcker, Tobias: *Sicherheit im World Wide Web*, Universität Heidelberg, Fachhochschule Heilbronn, März 1997.

[Hege95] Hegering, H.-G., Neumair, B., Wies, R.: Integriertes Management verteilter Systeme: Ein Überblick über den State-of-the-Art. GI/ITG-Fachtagung "Kommunikation in Verteilten Systemen", Tutorium, Technische Universität Chemnitz-Zwickau: Februar 1995

[Hunt94] Hunt, Craig: *TCP/IP Network Administration.* Fivth Edition, O'Reilly & Associates, Inc., May 1994.

[Iab99] http://www.iab.org/

[Ietf99] http://www.ietf.org/

[Harm98] Harms, V.: Biomathematik, Statistik und Dokumentation, 7. überarbeitete Auflage, Harms-Verlag, 1999.

[Kopk94] Helmut Kopka, LaTeX: Einführung, Band 1, First Edition, Addison-Wesley, 1994.

[Lamp94] Leslie Lamport, LaTeX: A Document Preparation System, Second Edition, Addison-Wesley, 1994.

[McCl91] McCloghrie, K., Rose, M.: „Management Information Base for Network Management of TCP/IP-based internets: MIB-II". RFC-1213, March 1991.

[McCl92] McCloghrie, K., Davin, J., Galvin, J.: „Definitions of Managed Objects for Administration of SNMP Parties". RFC-1353, July 1992.

[Post81a] Postel, Jon(Ed.): „Internet Protocol: Protocol Specification". RFC-791, September 1981.

[Post81b] Postel, Jon: „Internet Control Message Protocol - ICMP". RFC-792, September 1981.

[Post83] Postel, Jon: „The TCP Maximum Segment Size and Related Topics". RFC-879, November 1983.

[Rose90] Rose, M., McCloghrie, K.: „Structure and Identification of Management Information for TCP/IP-based Internets". RFC-1155, May 1990.

[Rose91] Rose, M., McCloghrie, K.: „Concise MIB Definitions". RFC-1212, March 1991.

[Ripe99] Source: RIPE Network Coordination Centre
http://www.ripe.net/,
(http://www.ripe.net/statistics/hostcount.html), Juni 1999.

[Stal93] Stallings, William: SNMP, SNMPv2, and CMIP: the practical guide to network management statndards / William Stallings; Addison-Wesley Publishing Company; Third Printing; October 1993.

[Wagn97] Wagner, A.: Künstliche Intelligenz im Netzwerkmanagement; Vortrag am Fachbereich Informatik (20) der Johann Wolfgang Goethe-Universität Frankfurt am Main im Rahmen des Seminars "Neue Architekturen für das Management verteilter Systeme" der Professur Verteilte Systeme und Betriebssysteme im Sommersemester 1997. (Der Vortrag fand statt am 2.7.1997 im Seminarraum 307 des Fachbereichs.); URL: http://www.vsb.cs.uni-frankfurt.de; E-Mail: A.Wagner@stud.uni-frankfurt.de; Frankfurt/Main (Germany): 1997

[Wash94] Washborn, Kevin und Jim Evans: *TCP/IP: Aufbau und Betrieb eines TCP/IP-Netzes.* Addison-Wesley, 1994.

[Wett99] http://www.wetterzentrale.de/klima/pmuenchr.html

[Zenk97] Zenk, Andreas: Lokale Netze - Kommunikation der 90er Jahre: LAN-Betriebssysteme, Internetworking, Groupware; Bonn(?): Addison-Wesley-Longman, 1997