ComConsult Technologie Information

Der Netzwerk Insider

September 1999

Technologie Information für die Ausbildung zum ComConsult Certified Network Engineer

Zum Schwerpunktthema

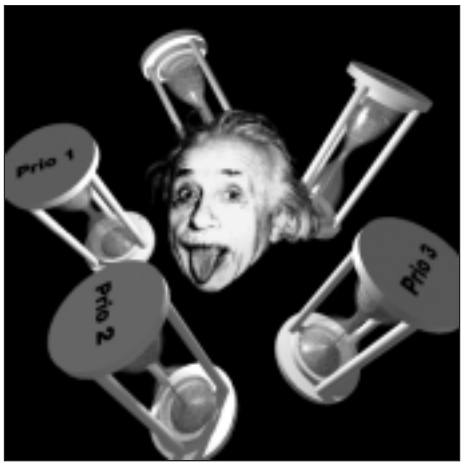
Prioritäten und Policies sind der falsche Weg

Von Dr. Jürgen Suppan

Die Hersteller von Netzwerk-Komponenten arbeiten an immer komplexeren Lösungen. Auch wenn die neuen Lösungen wie Prioritäten oder Policy Networking mit der Botschaft zentraler Steuerungs-Konsolen auf den ersten Blick attraktiv erscheinen, so zeigen sich doch bei näherer Analyse ernst zu nehmende Nachteile und Risiken:

- **sie führen zu einer sehr komplexen Fehlersuche**
- **es fehlt an Transparenz**
- **es entsteht eine starke Hersteller-Abhängigkeit durch inkompatible Ansätze**
- **es drohen höhere Personal-Kosten**

Um zu fundierten Aussagen zu kommen, wurde das Szenario des Einsatzes von Prioritäten von Dr. Kauffels mathematisch untersucht und einem einfacheren Ansatz der gezielten Überbuchung von Netzen (deutlich mehr Kapazität als im Mittel erforderlich) gegenüber gestellt.



Mathematische Analyse von Dr. Kauffels: Auch die Priorisierung unterliegt der Relativitätstheorie

Die mathematische Analyse bestätigt unabhängig von den Details der Modellannahmen, dass hinsichtlich der Qualität und der Effizienz der Prioritätsmechanismen noch erhebliche Unklarheiten und Risiken bestehen. Da mit dem Einsatz von Prioritäten und Policies speziell die notwendige Leistung für "mission critical" Anwendungen in Netzen gesichert werden soll, muss gerade für diesen Nutzungsbereich sichergestellt werden, dass keine Nachteile mit der Lösung verbunden sind. Die in dieser Ausgabe veröffentlichten Untersuchungen zeigen, dass primär folgende Nachteile bestehen:

- nur die höchsten beiden Prioritätsklassen kommen maximal in den Genuss einer leicht verbesserten Durchsatz- und Wartezeit-Situation
- allen anderen Prioritätsklassen drohen Nachteile bis zum Faktor 10, obwohl die erreichten Vorteile nur gering sind

weiter Seite 3

Schwerpunktthema

Leistungsanalyse: Gigabit Ethernet contra Policy Networks

von Dr. Franz-Joachim Kauffels

Schwerpunktthema dieser Ausgabe ist eine mathematische Untersuchung von Dr. Franz-Joachim Kauffels mit einem Vergleich der Effizienz des Einsatzes von Prioritäten und Policies gegenüber einem Gigabit-Ethernet-Ansatz.

Da die Untersuchung sehr mathematisch ist, werden vorweg die Ergebnisse in ihrer Konsequenz von Dr. Jürgen Suppan vereinfacht beschrieben. (s.o.)

ab Seite 3

Zum Geleit

Horrorszenario Token Ring? Wie groß ist das Risiko wirklich?

Nach unserer Berichterstattung über die veränderte Marktsituation im Token Ring Bereich kam die Frage auf, ob wir mit unserer Sorge um die zukünftige Entwicklung nicht übertreiben. Ich möchte dem erwidern, dass ich der Überzeugung bin, dass wir eher zurückhaltend berichtet haben. Es gibt eine Reihe von Gründen, die mich veranlassen, mit großem Ernst und echter Besorgnis auf die Risiken im Token Ring Markt hinzuweisen.

Der Hauptgrund meiner Sorge liegt in der Menge notwendiger Arbeiten und dem damit verbundenen Zeit-, Personal-, Know-how- und Kapitalaufwand, um aus dem Token Ring auszusteigen. Ein wesentlicher Teil dieser Arbeiten ist linear abhängig von der Größe der Installation, so dass große Installationen ein höheres Risiko tragen als kleinere. Der erhebliche Test- und Vorbereitungsaufwand betrifft aber jede Größenordnung von Installation.

weiter Seite 2

Horrorszenario Token Ring? Wie groß ist das Risiko wirklich?

Fortsetzung von Seite 1

Die Umstellung von Token Ring auf Ethernet wird in vielen Fällen die vorherigen Einführung von TCP/IP in einer ggf. umfangreicheren Form als bisher sowie die Installation neuer Kabelsysteme erfordern. Letzteres ist im laufenden Betrieb naturgemäß schwierig. Aus diesem Grund betragen die Umstellungszeiten unserer Kunden im Bereich von Netzwerken mit mehr als 10.000 Endgeräten (typischerweise auch noch in mehreren Standorten) zwischen 3 und 4 Jahren und erfordern Investitionen bis in den zweistelligen Millionenbereich.

Der notwendige Personalaufwand geht in den Bereich mehrerer 1000 Personentage (es ist zu beachten, dass die Netzwerk-Adapter in allen Geräten getauscht werden müssen und die gesamte Konfiguration für den Kommunikationsbereich geändert werden muss. Nimmt man einmal an, dass ein neuer Adapter 100 DM kostet, die Tauschzeit inklusive Test und Wegezeiten 30 bis 45 Minuten beträgt, dann entsteht daraus bei 10000 Endgeräten ein Gesamtaufwand von mindestens 1 Millionen DM Investition in Hardware und ein Schätzaufwand von 1000 Tagen Arbeit, ohne dass dabei die notwendigen Migrationsarbeiten im Bereich Switch

etc. berücksichtigt worden sind. Die Neuinstallation von Kabeln wird bei 10000 Geräten ca. 15000 Anschlüsse betreffen. Dies ergibt einen Schätzbetrag von ca. 4 bis 5 Mio DM Kosten und einen Arbeitsaufwand von mindestens 1500 Arbeitstagen. Letztere sind im laufenden Betrieb mit gleichzeitiger Störung für die Anwender zu erbringen).

Die Einführung von TCP/IP in einer Umgebung, die dies bisher noch nicht mit den Rahmenbedingungen einer Flächendeckung – TCP/IP liegt auf einigen Geräten vor, wurde aber nicht strategisch eingeführt – gemacht hat (kommt durchaus noch vor), kann neben dem Aufwand einer Neuplanung von TCP/IP (Adresskonzept, DHCP, DNS, Firewall, Sicherheitskonzept, Namenauflösung Netbios usw.) die Umstellung auf ein neues Betriebssystem-Release im Mainframe bedeuten, die Folgenotwendigkeit, die CPU-Leistung im Mainframe auszubauen usw. (mit den entsprechenden Folgekosten für alle Software-Lizenzen, die auf MIPS-Basis abgerechnet werden). In vielen Fällen kann dies zeitgleich mit der Klärung der Frage der Mainframe-Nutzung für E-Commerce erfolgen (siehe die aktuellen Ankündigungen der IBM zu diesem Thema auf der Mainframe-Seite). Dies ist einerseits zu begrüßen, wird aber andererseits zu Verzögerungen führen.

Gleichzeitig lässt sich leicht nachweisen, dass die 16 Mbit/s Version des Token Rings für alle Anwendungen außer der 3270-Terminal-Emulation zu wenig Leistung auf der Serverseite hat. Beispiel: eine 16 Mbit/s Anbindung für einen PC-Server entspricht der CPU-Leistung eines Intel 386-Prozessors. Dementsprechend gering ist der Multiuser-Grad, der sich auf der Basis von 16 Mbit/s erreichen lässt, wenn größere Datenmengen transportiert werden sollen. Daraus resultiert ein Bedarf auf 100 Mbit/s HSTR mindestens im Bereich der Server umzusteigen. Da dies mit ernstzunehmenden Investitionen verbunden sein kann, werden sich viele Anwender fragen, ob sie dieses Geld nicht gleich in den Umstieg auf Ethernet investieren. Wahrscheinlich liegt genau an dieser Stelle das größte wirtschaftliche Risiko für den Bestand der Token Ring Hersteller.

Ein Umstieg auf HSTR 100 Mbit/s ist auf absehbare Zeit nur mit Madge Connect machbar. Scheitert Madge, haben die betroffenen Anwender angesichts der zuvor beschriebenen Aufwände ein sehr ernstzunehmendes Versorgungs-Problem. Natürlich ist im Gegenteil zu hoffen, dass die Konzentration auf einen Anbieter eher zu einer Stabilisierung des Marktes führt, da sich das Geschäft für diesen Hersteller zumindest für eine Übergangszeit eher lohnen wird, als wenn sich mehrere einen zu kleinen Kuchen teilen müssen. Von daher hoffe ich, dass die Versorgungssicherheit tendenziell noch für mindestens 2 Jahre gegeben ist.

Betrachtet man die Summe der Argumente, so halte ich dies insgesamt nicht für ein Horror-Szenario, sondern ich bin überzeugt, dass viele Kunden Gefahr laufen, den Gesamtaufwand an Zeit und Investition für einen Ausstieg aus dem Token Ring zu unterschätzen. Hinzu kommt, dass die Migration von Token Ring zu Ethernet je nach Umgebung erhebliches Spezialwissen voraussetzen kann, über das in Deutschland nicht viele Personen verfügen. Diese Personen sind leider nicht clonebar (entsprechende Arbeiten der Bio- und Gentechnologie werden zumindest für diese Aufgabe zu spät kommen).

Naturgemäß besitzen zudem einige der bisherigen Token Ring Betreiber noch relativ wenig Wissen über den Einsatz von modernen Ethernet-Netzwerken. Auch wenn das Grundlagen-Know-how vorhanden ist, muss in einem ernstzunehmenden Umfang Ethernet-Wissen geschult werden, und es müssen Erfahrungen gesammelt werden. Die Erfahrung in den entsprechenden Seminaren hat deutlich gezeigt, dass seitens der Token Ring Betreiber das Ethernet-Thema zum Teil erheblich unterschätzt wird. Ethernet ist nicht mehr die Billig-Technologie wie sie es vor 5 Jahren vielleicht noch war. Die internen Technologie-Mängel von Ethernet haben zu Ergänzungstechnologien geführt, die in Summe ein erhebliches Leistungsspektrum abdecken. Durch diese Ergänzungstechnologien wird Ethernet allerdings bis an den Rand der Beherrschbarkeit komplexer, so dass hier ebenfalls ernstzunehmende Aufwände im Verstehen dieser Technologien zu erwarten sind.

Wer sich für dieses Thema interessiert, der wird auf unsere Sonderveranstaltung "Ausstieg aus dem Token Ring" vom 3. bis 5. November in Aachen verwiesen. Dort sind voraussichtlich auch alle relevanten Hersteller vertreten und stellen sich den Fragen der Teilnehmer.

Ihr
Dr. Jürgen Suppan

Impressum

Verlag:
ComConsult
Technologie Information GmbH
Pascalstraße 25
D-52076 Aachen
Telefon 02408/14907
Telefax 02408/149233

Herausgeber
und verantwortlich
im Sinne des Presserechts:
Dr. Jürgen Suppan

Gestaltung: Zweipfennig

Erscheinungsweise:
Monatlich,
12 Ausgaben im Jahr

Bezug:
Kostenlos als PDF-Datei
über den eMail-VIP-Service
der ComConsult Akademie

Für unverlangt
eingesandte Manuskripte
wird keine Haftung übernommen

Prioritäten und Policies sind der falsche Weg

Fortsetzung von Seite 1

Ein weiterer Risikopunkt liegt in den Schwächen des Prioritäten-Standards. Zwar wird dort festgelegt, wie grundsätzlich Pakete unterschiedlichen Klassen zugeordnet werden, es bleibt aber offen, nach welcher Strategie die Pakete einer höheren Priorität wirklich behandelt werden. Diese Entscheidung wird der individuellen Entwicklung der Hersteller überlassen. Aus der Entwicklung ähnlicher Mechanismen in Betriebssystemen ist bekannt (Stichwort: Scheduling), dass die Wahl einer falschen Bearbeitungsstrategie mit erheblichen Nachteilen für die Delay-Werte in niedrigeren Prioritätsklassen verbunden sein kann und im Extremfall sogar eine Blockade dieser Klassen möglich ist. Die mathematische Untersuchung von Dr. Kauffels für den Einsatz von Prioritäten in Switch-Systemen bestätigt genau diese Gefahr. Hinzu kommt, dass das jeweilige reale Verhalten der Switch-Systeme im Betrieb dem Betreiber Rätsel aufgeben kann, da genau der wichtige Weiterleitungs-Mechanismus herstellerspezifisch ist.

Für die Praxis kann daraus nur folgende Empfehlung abgegeben werden:

- Netzwerke und Switch-Systeme sollten konsequent unterbucht ausgelegt werden, d.h. sie werden mit wesentlich mehr Kapazität realisiert als eigentlich notwendig

- Gigabit-Technologien sollten im Backbone planerisch auf jeden Fall vorgesehen werden

Vorteil dieser Strategie ist u.a. ein stark vereinfachter Betrieb, da keine komplexe und aufwendige Parametrierung notwendig ist.

Der Nachteil liegt in den höheren Kosten von Switch-Systemen, die für den Einsatz mit Gigabit-Technik wirklich geeignet sind. Hier muß deutlich zwischen einem scheinbaren Angebot an Gigabit-Fähigkeit durch die Verfügbarkeit von Gigabit-Ports und einer wirklichen Gigabit-Eignung der Geräte in der Schaltmatrix und der Verbindung der Ports zur Schaltmatrix unterschieden werden.

Es ist auch zu berücksichtigen, dass neben der reinen Erfüllung von Leistungswerten wie Kapazität und Delay gerade für "mission critical" Anwendungen auch eine sehr hohe Verfügbarkeit sicher gestellt werden muss. Hier sollten grundsätzlich "load sharing" Technologien den Vorrang vor allen "hot stand by" Lösungen erhalten. Dies heisst konkret, dass eine Gesamtauslegung von Netzen durch

- sehr hohe Kapazitäten inklusive Gigabit-Technologie und

- "load sharing" Redundanz mit

- Link Aggregation im Layer 2 oder

- OSPF im Layer 3

vorrangig vor dem Einsatz von Policies und Prioritäten angestrebt werden sollte.

Prioritäten und Policies sollten erst zum Einsatz kommen, wenn diese Vorgehensweise wirklich konsequent ausgereizt wurde. Es mag sein, dass dies in 2 bis 3 Jahren der Fall ist. Bis dahin haben die Netzwerk-Hersteller Zeit Policy-Networking nicht nur in der Marketing-Abteilung sondern auch im Labor einer wirklichen Marktreife näher zu bringen.

Leistungsanalyse: Gigabit Ethernet contra Policy Networks

Vorspann

Die meisten der heutigen Netzwerk-Technologien wurden auf der Basis mathematischer Untersuchungen und Modellierungen entwickelt. Damit liessen sich denkbare Nachteile der Verfahren rechtzeitig ermitteln und wichtige Verbesserungen erreichen. In unserer schnellebigen Zeit ist diese Vorgehensweise aus der Mode gekommen. Wir machen unsere Erfahrungen am lebenden Testobjekt: dem Anwender. Die nachfolgende Untersuchung zeigt, dass sehr wohl durch geeignete mathematische Untersuchungen wichtige Aussagen über eine sinnvolle Weiterentwicklung unserer Netzwerke getroffen werden können. In diesem Fall wurde eine Netzwerk-Entwicklung unter dem gezielten Einsatz von Prioritäten durch einen Policy-Server einem einfachen unparametrierten Einsatz von Gigabit Ethernet gegenüber gestellt. Die Grundfrage der Untersuchung lautet: was ist besser, das parametrische Herauskitzeln der letzten Performance oder der bewußte Einsatz von Überkapazität.

Verfahren wie CSMA/CD haben sich in der Praxis besser gezeigt, als aufgrund der mathematischen Analysen angenommen werden konnte. Dies lag an den Vereinfachungen, die gemacht werden müssen, um die analytischen Modelle handhabbar zu machen. Trotzdem können aus Analysen wichtige Erkenntnisse gewonnen werden, wenn man sich auf wichtige Kernaussagen beschränkt. Beim Design moderner Netze geht es nicht nur um nackte Nominalleistung, sondern auch um das Minimieren von Verzögerungen (delay). Viele Hersteller von Netzkomponenten vertreten in den Marketing-Unterlagen die Ansicht, dass eine Kombination aus Fast Ethernet Switches und Priorisierungsverfahren fast ebenso gut sei oder besser als ein Gigabit Ethernet Switch ohne Priorisierung. Die Beschränkung auf genau diese Frage lässt eine Antwort durch ein mathematisches Modell zu, ohne dass zu grosse Abweichungen von der späteren Realität zu befürchten sind.

In der Diskussion um das Netzwerk-Redesign werden immer wieder Fragen nach Leistung, Verzögerung und Varianz der Verzögerung aufgeworfen.

Allgemein werden folgende maximale Verzögerungswerte für bestimmte Gruppen von Anwendungssystemen gefordert:

Application Sharing Cluster	< 0,1 ... 1 ms
Verteilte Applikationen	< 1 ... 10 ms
High End Video	< 10 ms
Low End Video	< 40 ms
Sprache	< 100 ms

Da Cluster eine eigene technologische Entwicklungsstufe darstellen und die Anforderungen für flächendeckende verteilte Applikationen auch nicht so häufig gegeben sind, wird für das Hauptmaß der Anwendungen der Delay-Wert 20 ms angestrebt.

Dabei haben sich drei Fraktionen herausgebildet: eine Fraktion möchte im Rahmen herkömmlicher Systeme, die eine Grundleistung von 100 Mbit/s. (Fast Ethernet) oder 155 Mbit/s. (ATM) mit zusätzlichen Verfahren wie RSVP, IEEE802.1p und q usf. Quality of Service durch Priorisierung sicherstellen, die zweite Fraktion setzt mit Gigabit Ethernet auf Leistung satt und die dritte Fraktion möchte am liebsten zu Gigabit Ethernet noch zusätzliche Verfahren packen.Der Autor hat schon früh die Ansicht vertreten, dass er vom "Aufmotzen" eines zu leistungschwachen Netzes mit zusätzlichen Verfahren wenig hält. Es ist so wie mit einem alten Mercedes Diesel: durch zusätzliche Spoiler und Schweller wird er nur noch schwerer und dadurch noch langsamer. Dabei ist es übrigens gleichgültig, ob man über Fast Ethernet oder ATM 155 redet: bei ATM trägt der "natürliche Overhead" in Form der LAN-Emulation oder des Verfahrens TCP/IP over ATM mindestens die Hälfte der Nominalleistung zu Grabe.

Hinsichtlich der Leistung von Gigabit Ethernet Switches hat der Autor beim letzten ComConsult/Dr.Kauffels Netzwerk Redesign Forum in den Raum gestellt, dass aufgrund der Warteschlangentheorie zu erwarten sei, dass bei einer Auslastung von bis zu 70% fast immer fast niemand warten muß und deshalb weitere Mechanismen zur Vorrangsteuerung und Priorisierung völlig überflüssig seien. Dieses wird im folgenden untersucht.

Da vielen die Modelle nicht mehr bekannt sein werden, soll hier in gewissen Grenzen versucht werden, eine Herleitung zu skizzieren. Wer sich wirklich dafür interessiert, sollte allerdings die Originalliteratur lesen.

Modellierung, Grundgedanken

Verschiedene wichtige Eigenschaften der Flüsse diskreter Einheiten im stabilen Zustand eines Systems können unabhängig von den genauen statistischen Verteilungen, der Parameter, die die Flüsse definieren, entwickelt werden. Dabei ist es vollkommen gleichgültig, ob eine Menschenmenge in einem Park, die Kassenschlange beim ALDI, der Verkehr auf einer Straße oder das Verhalten eines Netzwerk-Switches modelliert wird.

Wir betrachten ein begrenztes Gebilde, welches ein Netzwerk mit einem oder mehreren Ein- und Ausgängen darstellt. Wir nehmen an, dass das Netzwerk die Nachrichten bewahrt, es werden also keine Nachrichten hinzugefügt, weggenommen

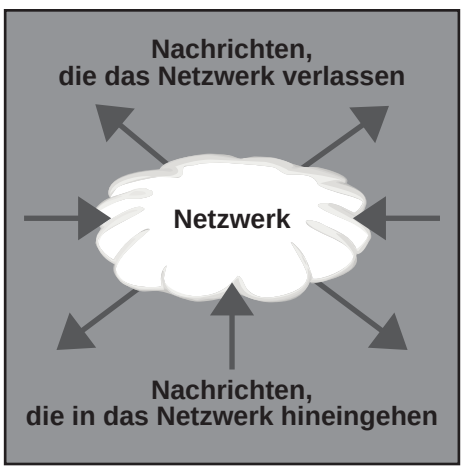
Der Autor

Dr. Franz-Joachim Kauffels ist einer der bekanntesten und erfahrensten Referenten der Netzwerkszene. Bekannt geworden durch vielfältige Publikationen, Seminare und Kongresse setzt sich der selbständige Unternehmensberater für eine fortschrittliche und wirtschaftliche Nutzung der neuen Technologien ein.

oder verändert. Deswegen können in das Netz über einen Eingang kommende Nachrichten entweder im Netz gespeichert, oder an einem Ausgang ausgegeben werden. Wenn die durchschnittliche Input-Rate die durchschnittliche Output-Rate übersteigt, muß das Netz die Nachrichten zwischenspeichern, wegwerfen gilt ja nicht. Wenn die durchschnittliche Output-Rate die durchschnittliche Input-Rate übersteigt, wird die Anzahl der zwischengespeicherten Nachrichten gegen Null fallen, und zwar bis zu dem Punkt, wo die Output-Rate die Input-Rate nicht mehr übersteigen kann, weil außer den über den Input kommenden Paketen keine mehr da sind.

Man sagt, dass ein solches Gebilde stabil ist, wenn über einen längeren Zeitraum betrachtet die Input- und die Output-Rate gleich sind. In dieser Hinsicht stabile Netze können typischerweise durch einen Zufallsprozeß modelliert werden, dessen Mittelwerte über die Zeit den Mittelwerten verschiedener statistischer Stichproben entsprechen.

Bild 1 ein Netzwerk innerhalb fester Grenzen



Sei $\alpha(t)$ die Anzahl der eingehenden Nachrichten und $\delta(t)$ die Anzahl der ausgehenden Nachrichten in einem Zeitintervall (0,t). Die Differenz dieser Werte, $N(t)$, bezeichnet das Wachstum der Anzahl der Nachrichten, die über das Zeitintervall zwischengespeichert werden:

$$N(t) = \alpha(t) - \delta(t) \quad (1)$$

Handhabbare deterministische Maße für den Fluss können durch die Mittelwertbildung von Variablen, wie sie gerade definiert wurden, erzeugt werden. Die mittlere Input-Rate λt über ein Intervall der Länge t wird definiert:

$$\lambda t = \alpha(t)/t \quad (2)$$

Durch Grenzwertbildung von t nach Unendlich über alle t kommt man zur allgemeinen Inputrate λ, die auch als mittlere Ankunftsrate bezeichnet wird. Ebenso kommt man mit dem Integral über die Werte von $N(t)$ und anschließender Grenzwertbildung zur mittleren Anzahl im System befindlicher Nachrichten N. Schließlich bezeichnen wir T_t als die mittlere Zeit während des Intervalls (0,t), die die Nachrichten während dieses Intervalls im System verbringen. Auch hier kommt man mittels Grenzwertbildung zu einem allgemeinen T. Diese Werte sind nicht unabhängig voneinander, sondern es gilt:

$$N = \lambda T \quad (3)$$

Diese Beziehung ist auch als "Littles Law" bekannt und besagt, dass die durchschnittliche Anzahl von Nachrichten, die in einem Netzwerk gespeichert werden, gleich dem Produkt aus mittlerer Ankunftsrate und durchschnittlicher Zeit, die diese Nachrichten im Netz verbringen, ist. Diese heuristische Herleitung kann auch stark bewiesen werden, allerdings braucht man dazu die Theorie stochastischer Prozesse, die wir hier nicht bemühen wollen. Littles Law stimmt für alle deterministischen oder stochastischen Input-Output-Systeme, in denen diskrete Einheiten das System betreten und zwischengespeichert werden, wenn alle verfügbaren Ausgänge gerade beschäftigt sind.

Eine weitere Vereinfachung des Modells kann man dadurch herbeiführen, dass man

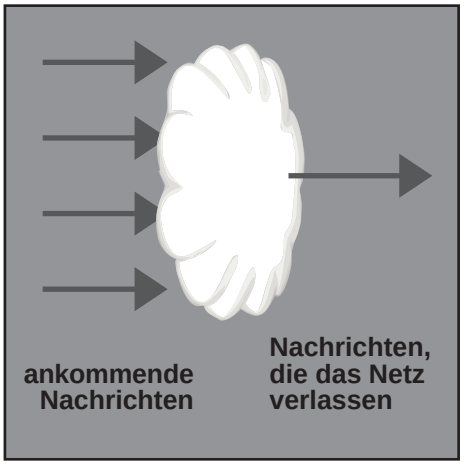
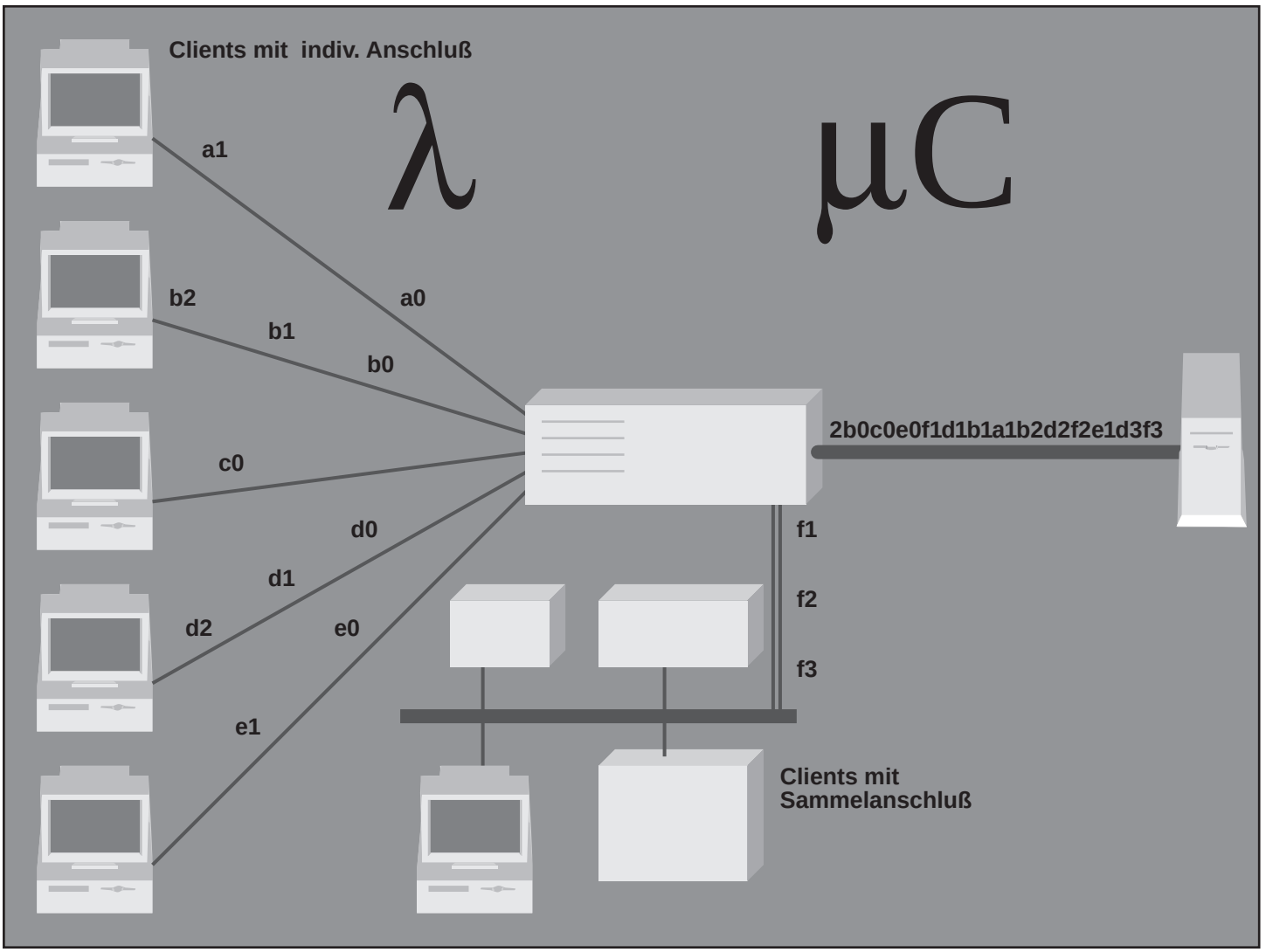


Bild 2 ein typisches Netz mit einem Output-Kanal

annimmt, dass zu einer Zeit nur eine Message das System verlassen kann und dass die Messages in der Reihenfolge verarbeitet werden, wie sie ankommen. Es ist allerdings erlaubt, dass zu einer Zeit mehrere Nachrichten ankommen.

Dies öffnet nicht nur mathematisch Wege geringerer Komplexität, sondern ist besonders geeignet für die Modellierung von Switch-Systemen. Man muss sich vor Augen halten, dass die Warteschlangenmodelle, die hier verwendet werden, bis zu Beginn der siebziger Jahre entwickelt und in den legendären Büchern von Kleinrock 1975 niedergelegt und bekannt wurden. Dann wurden sie u.a. für die Bewertung von Mehrfachzugriffsteuerungssystemen herangezogen, wie z.B. CSMA/CD. Allerdings hat hier oftmals eine gewisse Ungenauigkeit den

Bild 3 Modellierungsgrundlage: viele Clients, ein Server-Output



Spaß an den Ergebnissen verdorben, insbes. der CSMA/CD Algorithmus hat sich im täglichen Leben "besser" verhalten als prognostiziert, weil z.B. auch durch die Bearbeitungszeiten in den Adapterkarten zusätzliche Delays vor dem Beginn des Wettbewerbs um den Kanalzugang entstanden sind, die aber in die allgemeinen Exponentialverteilungen für die Zeiten zwischen den Ankünften von neuen Paketen nicht unkompliziert eingerechnet werden konnten. Ein gut gebauter Switch, der auch in Überlastsituationen souverän handelt und keine Pakete verliert, ist aber ein ideales Netzwerk im stabilen Zustand. Pakete werden von der Schaltmatrix in der Reihenfolge behandelt, wie sie ankommen. Kommen mehrere gleichzeitig an, werden sie durch ein Round Robin Verfahren etwas auseinandergezogen und dann der Reihe nach bedient. Für die Modellierung an sich spielt es keine Rolle, ob Pakete als Ganzes oder mit einem Cut Through Verfahren behandelt werden. Bei Cut Through sind eben nur die abgeschnittenen Enden die Pakete und der Rest spielt für die Analyse keine Rolle mehr. (Bild 3)

Durch diese Voraussetzungen kommen wir zu einem einfachen Warteschlangenmodell bestehend aus einem Speicher, in dem die ankommenden Nachrichten gesammelt werden (die "Warteschlange") und einem Bediener, der die Pakete weiterverarbeitet und somit die Warteschlange leert. Im Falle des Switches wäre der Bediener die Schaltmatrix, die die ankommenden Pakete an die Ausgangsports leitet. Die Leistung der Bedienung ist die Realisierung der entsprechenden Abbildung vom Eingangs- auf den Ausgangsport (Bild 4)

Wie bereits dargestellt, sammeln sich in einem stabilen System zwar manchmal Aufträge in der Warteschlange, aber insgesamt muss der Bediener mindestens so schnell sein, dass die im Mittel ankommenden Aufträge bearbeitet werden können und ggf. in der Warteschlange befindliche Aufträge nicht endlos aufgetürmt werden. Damit das System dauerhaft stabil bleiben kann, muss also der Bediener eigentlich etwas schneller sein als die mittlere Ankunftsrate von Aufträgen, sonst kann er ja die Warteschlange nicht abarbeiten. Das führt dazu, dass der Bediener auch einmal Leerzeiten hat. Der mittlere Bruchteil der Zeit, an der der Bediener beschäftigt ist (ohne Restriktion auf einen bestimmten Zeitpunkt t), wird als r bezeichnet. Nach unserer Definition gilt:

$0 \leq \rho \leq 1$

Normalerweise wird für die Modellierung lokaler Netze noch der Schritt vollzogen, dass man den Bediener mit einem Ausgangskanal der Bitrate C Bits pro Sekunde assoziiert.

Bezogen auf die Modellierung eines Switch Systems bedeutet dies letztlich, dass nicht die Datenrate an einem Ausgangsport kennzeichnend für das System ist, sondern die aggregate Gesamtleistung der Switching Engine. Dies ist allerdings nur dann ein korrektes Maß, wenn die Gesamtleistung bei oder über der Anzahl der Ports multipliziert mit der Port-Datenrate liegt. Dies ist völlig unabhängig davon, ob es sich um einen normalen, Fast- oder Gigabit-Switch handelt. Ist die aggregate Gesamtleistung zu schwach, kommt es zum gefürchteten Oversubscribing: zuerst laufen die Puffer voll und das Delay erhöht sich, dann kommt es zum Wegwerfen von Paketen:

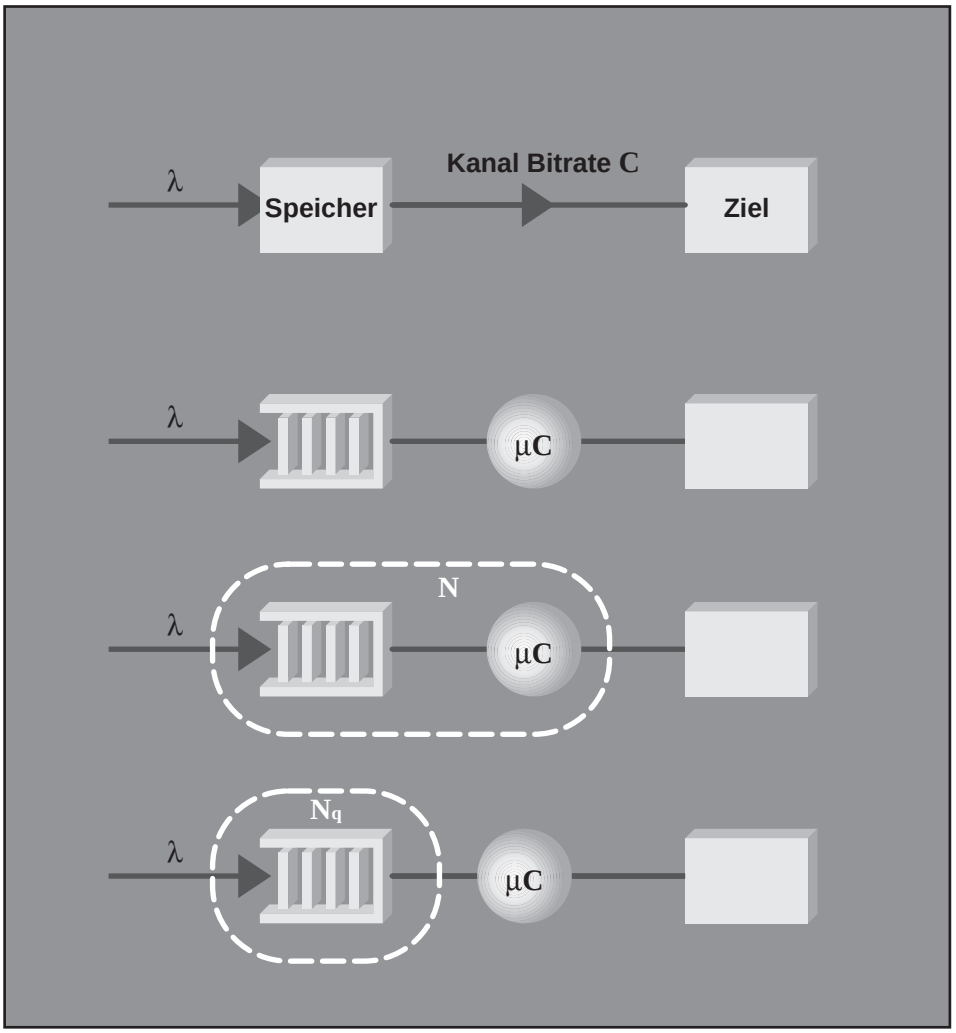
Ein Switch, bei dem aufgrund zu geringer aggregater Gesamtleistung ein Oversubscribing zu befürchten ist, ist kein stabiles System im Sinne unserer Modellierung.
(Bild 5 zum Oversubscribing)

Andererseits ist es so, dass vielfach vornehmlich das Verhalten eines Switches hinsichtlich einer Uplink-Konzentration vieler Clients auf einen Server zur Diskussion steht. Und hier ist das Modell eines Netzwerkes mit vielen Eingängen und einem Ausgang geradezu ideal. Wir setzen folgendes voraus:

- Die Leistung der Switch Matrix liegt weit jenseits der Leistung eines einzelnen Ausgangsports und bei oder über der Summe aller Portdatenraten. Dadurch ist das Oversubscribing ausgeschlossen.

- Die von unterschiedlichen Clients ankommenden Pakete werden als Aufträge für die Aussendung an einem Server Port entgegengenommen.

Bild 4 Das einfache Warteschlangenmodell mit Verdeutlichung der Werte **N** und N_q



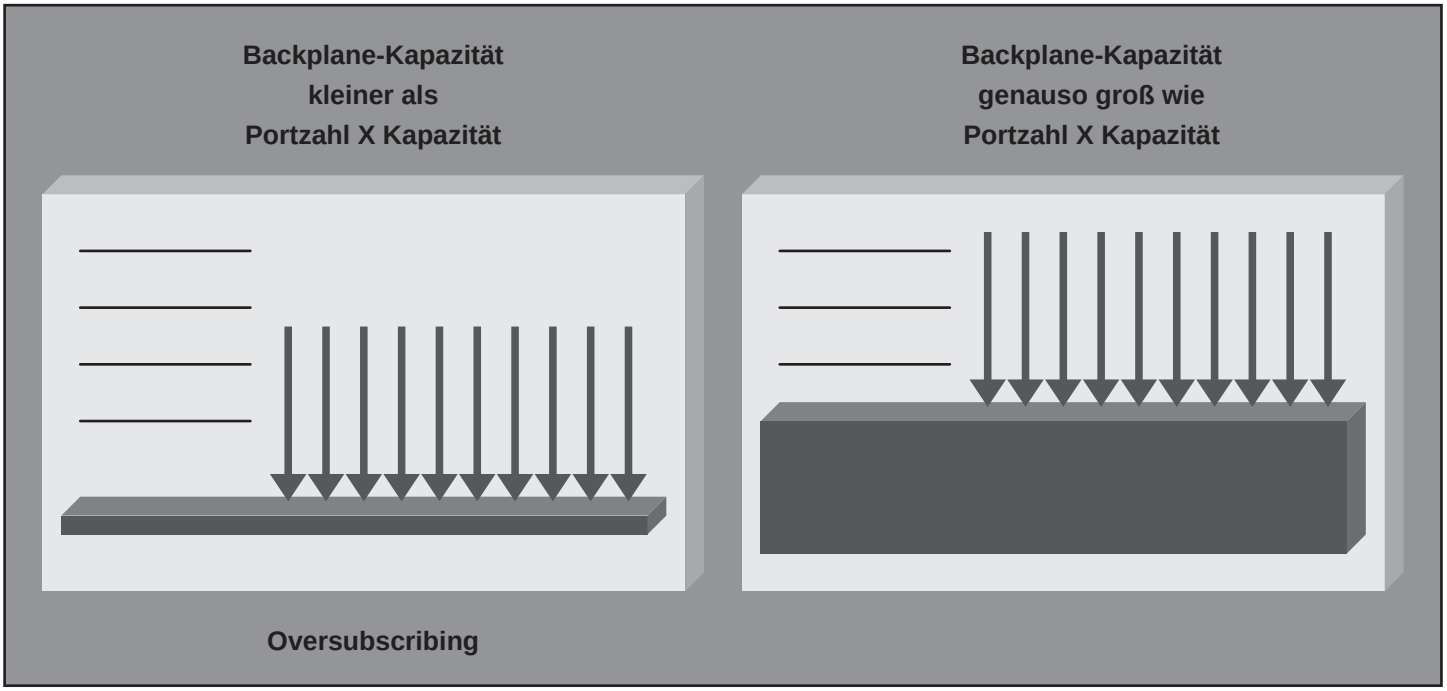


Bild 5 Merkmale von Switches: Backplane-Kapazität

- Der Server Port hat die Datenrate C nominell und real
- Das System ist in einem stabilen Zustand: die Ankunftsrate multipliziert mit der durchschnittlichen Paketgröße ist über ein längeres Zeitintervall betrachtet nicht größer als die mögliche Leistung des Server Ports

Mit diesen Voraussetzungen kommen wir in Riesenschritten zu einem Modell, dessen Verhalten bekannt ist. So können wir die interessierenden Beziehungen schnell ausrechnen.

Wir nehmen weiterhin an:

- die durchschnittliche Input-Rate ist λ Nachrichten pro Sekunde
- die Nachrichtenlänge unterlegt einer Zufallsverteilung
- die mittlere Nachrichtenlänge ist 1/m Bits pro Nachricht

Im Modell nehmen wir an, dass die Nachrichten mit allen Bits zu einem Zeitpunkt gleichzeitig ankommen. Das ist eigentlich etwas verwaschen, aber andererseits ist es ja so, dass eine Nachricht insgesamt als so lange im Netzwerk befindlich betrachtet wird, bis das letzte Bit "draußen" ist. Die miminale Verzögerung ergibt sich also durch die Zeit, die benötigt wird, um die komplette Nachricht bei ansonsten leerem System durch das System zu bringen. Bei einem Switch System kennt man diese Zeit durch Messungen. Bei Gigabit Ethernet Switches ist sie meist beeindruckend klein.

Da die mittlere Länge der Nachricht 1/m ist und die Kanal-Bitrate C bit/s. beträgt, ist die durchschnittliche Übertragungs- oder Verarbeitungszeit 1/mC Sekunden.

Im stabilen Zustand müssen durchschnittliche Ein- und Ausgaberate gleich sein, in diesem Falle λ Nachrichten pro Sekunde. Wenn Nachrichten im Kanal sind, werden sie mit einer durchschnittlichen Rate von mC Nachrichten pro Sekunde bedient. Der Unterschied zwischen λ und μC entsteht dadurch, dass der Kanal nicht dauernd beschäftigt ist und somit die mittlere Ausgangsrate aus Zeiten, wo der Kanal mit der Rate mC beschäftig ist und Zeiten, wo der Kanal nichts tut, zusammengesetzt ist. Da wir den mittleren Anteil der Zeit, an dem der Kanal etwas tut, mit ρ bezeichnet haben, gilt:

$$\lambda = \rho\mu C \quad (4)$$

oder

$$\rho = \lambda/\mu C \quad (5)$$

Die letzte Relation zeigt, dass die Verkehrsdichte der mittleren Ankunftsrate geteilt durch die mittlere Bedienrate entspricht, wenn man dies in verträglichen Einheiten ausdrückt. Damit haben wir auch theoretisch nochmal den gesunden Menschenverstand unterlegt, der eben besagt, dass man in ein System nicht dauerhaft mehr Pakete hineinstecken darf, als dieses auch bewältigen kann.

Littles Law kann dazu benutzt werden, um N, die mittlere Anzahl von Nachrichtenpaketen, die im System bestehend aus Warteschlange und Bediener zwischengespeichert werden und N_q, als mittlere Anzahl von Nachrichtenpaketen nur in der Warteschlange, zu korrelieren. Diese Analyse setzt auch die durchschnittliche Zeit im Puffer W mit der durchschnittlichen Gesamtzeit im System, T, miteinander in Beziehung. Betrachtet man das gesamte System, gilt offensichtlich:

$$N = \lambda T \quad (6)$$

Beschränkt man sich auf den Speicher, so gilt:

$$N_q = \lambda W \quad (7)$$

Die mittleren Verzögerungen T und W können dadurch korreliert werden, dass die mittlere Verzögerung für Warteschlange und Kanal zusammen gleich der Summe der mittleren Verzögerung in der Warteschlange und der durchschnittlichen Zeit $1/\mu C$ ist, die dazu benötigt wird, eine Bedienung durchzuführen, sobald das Paket am Kopf der Warteschlange angekommen ist:

$$T = W + 1/\mu C \tag{8}$$

Dabei wird allerdings die Zeit vernachlässigt, die das Signal vom Kopf der Warteschlange bis zum Bediener benötigt. Da bei den hier betrachteten Fällen die Signale mit ungefähr 200.000 km/s laufen, können wir das vernachlässigen. Hier unterscheiden sich die Umsetzungen in die Realität: modelliert man einen Switch, ist es gleichgültig, modelliert man die Kassenschlange beim ALDI, spielt diese Übergangsverzögerung ggf. eine große Rolle.

Multipliziert man beide Seiten von (8) mit λ und setzt aus (6) und (7) ein, ergibt sich

$$N = N_q + \rho \tag{9}$$

Das Ergebnis besagt, dass die Gesamtzahl der Nachrichten im System gleich der im Puffer zuzüglich ρ ist. Entweder bedient der Kanal ein Paket oder nicht. ρ zwischen Null und Eins gibt als Mittelwert sozusagen gerade den durchschnittlichen Bruchteil von Paketen in der Bedienung an.

Eine Gruppe geeigneter Warteschlangenmodelle

So schön die Herleitung auch bis jetzt gelaufen ist, werden wir jetzt zu allgemeinen Modellen kommen müssen, bei denen einfach eine Reihe von Ergebnissen ausgerechnet wurden und fertig vorliegen. Grundsätzlich gehen wir davon aus, dass der Input-Prozess als Poisson Prozess modelliert wird. Diese Annahme ermöglicht es, mit vernünftigem Aufwand Relationen für die durchschnittliche Anzahl von Nachrichten im Netzwerk, die durchschnittliche Verzögerung usf. aufzustellen. Für den Bedienprozeß gibt es allerdings verschiedene Modelle, darauf kommen wir gleich. In der Warteschlangentheorie hat es sich eingebürgert, ein Modell mit drei Symbolen zu bezeichnen: I/O/x, wobei I ein Symbol für den Inputprozeß, O ein Symbol für den Outputprozeß und x die Anzahl der Bediener ist.

M/G/1-Modell

Das "M" steht für "memoryless". Dies ist eine Charakterisierung des Poisson-Prozesses, die besagt, dass die Zeiten zwischen Ankünften von Paketen (sic: Zwischenankunftszeiten) unabhängig voneinander sind. "G" bedeutet genereller Bedienprozess, also ein Prozess, über dessen Verhalten man relativ wenig sagen kann, und die Anzahl der Bediener ist 1. Die Bearbeitungszeit für eine Nachricht, Y, ist definiert als die Nachrichtenlänge in Bits geteilt durch die Kanal-Bitrate. Von der Kanal-Bitrate nehmen wir an, dass sie konstant C Bits/s. ist und dass die Nachrichtenlänge zufällig verteilt ist. Obwohl diese Verteilung nicht näher spezifiziert ist, nehmen wir an, dass der Mittelwert dieser Verteilung, den wir als $1/\mu$ Bits/Nachricht notieren, bekannt ist. Somit ist die durchschnittliche resultierende Rate für die Bearbeitung von Nachrichten μC Nachrichten pro Sekunde.

Für den Ankunftsprozess gilt, dass die Wahrscheinlichkeit für k Ankünfte in einem Intervall der Länge T durch folgende Beziehung gegeben ist:

$$P\{k \text{ Ankünfte in T Sekunden}\} = \frac{(\lambda T)^k \, e^{-\lambda T}}{k!}$$

für $k = 0, 1, 2, \ldots$ (10)

Nach einer länglichen Rechnung, für die wirklich auf die Literatur verwiesen werden soll, gibt es eine Reihe von Ergebnissen für Delay und Wartezeit. Diese enthalten aber wegen des allgemeinen Bedienprozesses eine Reihe von Beziehungen, in die die statistischen Varianzen des Bedienprozesses einfließen. Es ist aber in unserem Fall gar nicht notwendig, diesen Aufwand zu treiben. Schon das CSMA/CD-System konnte wesentlich einfacher mit Hilfe des sog. M/M/1-Warteschlangenmodells behandelt werden, in dem auch die Verteilung der Bedienzeiten memoryless ist.

Übrigens: Bis zur sechsten Auflage meines Buches "Lokale Netze" 1993 habe ich CSMA/CD auf Basis dieser Ergebnisse analysiert. Erst ab der siebten Auflage habe ich eine präzisere Formulierung mit Erwartungswerten verwendet, in der 11. Aufl. zu lesen ab S. 326.

Jetzt betrachten wir allerdings einen Switch, genau genommen sogar nur den Teil eines Switches, der einen High Speed Uplink Ausgang bedient. Der Switch arbeitet aber mit Ausnahmen immer deterministisch. Im Cut Through Modus werden die ersten Bytes der Paketköpfe parallel gelesen und ausgewertet, danach wird die Switchmatrix umgestellt, das Paket durchgeschoben und die Bedienung ist fertig. Mit Varianzen von irgendwelchen statistischen Variablen machen wir uns hier nur unnütz das Leben schwer. Viel einfacher ist es, anzunehmen, dass der Switch deterministisch arbeitet und ggf. für etwaige Unwägbarkeiten am Ende eine Konstante zuzuschlagen. Dies führt auf das

M/D/1-Modell

Dieses Modell wurde in der Vergangenheit bei der Bewertung von Mehrfachzugriffschemen immer als Referenz genommen, es ist der "ideale" Bediener. Durch den Wegfall des CSMA/CD-Verfahrens liegt bei einem gut konstruierten Switch dieser Idealfall tatsächlich vor. Das M/D/1-Modell ist ein Spezialfall des M/G/1-Modells. Die bei letzerem als "Pollaczek-Khinchine"-Formeln eingegangene Formeln für die mittlere Anzahl N von Aufträgen im System oder der mittleren Zeit, die ein Auftrag im System verbringt, T können bei M/D/1 wesentlich vereinfacht werden. Es ist

$$N = \frac{\rho\,(2-\rho)}{2\,(1-\rho)} \tag{11}$$

$$T = \frac{2 - \rho}{2\mu C\,(1-\rho)} \tag{12}$$

Da wir ja annehmen, dass der Bediener konstant und zügig arbeitet, ist N_q noch interessanter als N.

$N = N_q - \rho$ also gilt:

$$N_q = \frac{\rho^2}{2\,(1-\rho)} \tag{13}$$

und

$$W = \frac{\rho}{2\mu C\,(1-\rho)} \tag{14}$$

Die mittlere Speicherverzögerung des M/D/1-Systems beträgt die Hälfte von der des M/M/1-Systems.

Die Kurven von N und μCT gegen ρ für die M/M/1 und M/D/1-Warteschlangenzeigen verschiedene generelle Eigenschaften.

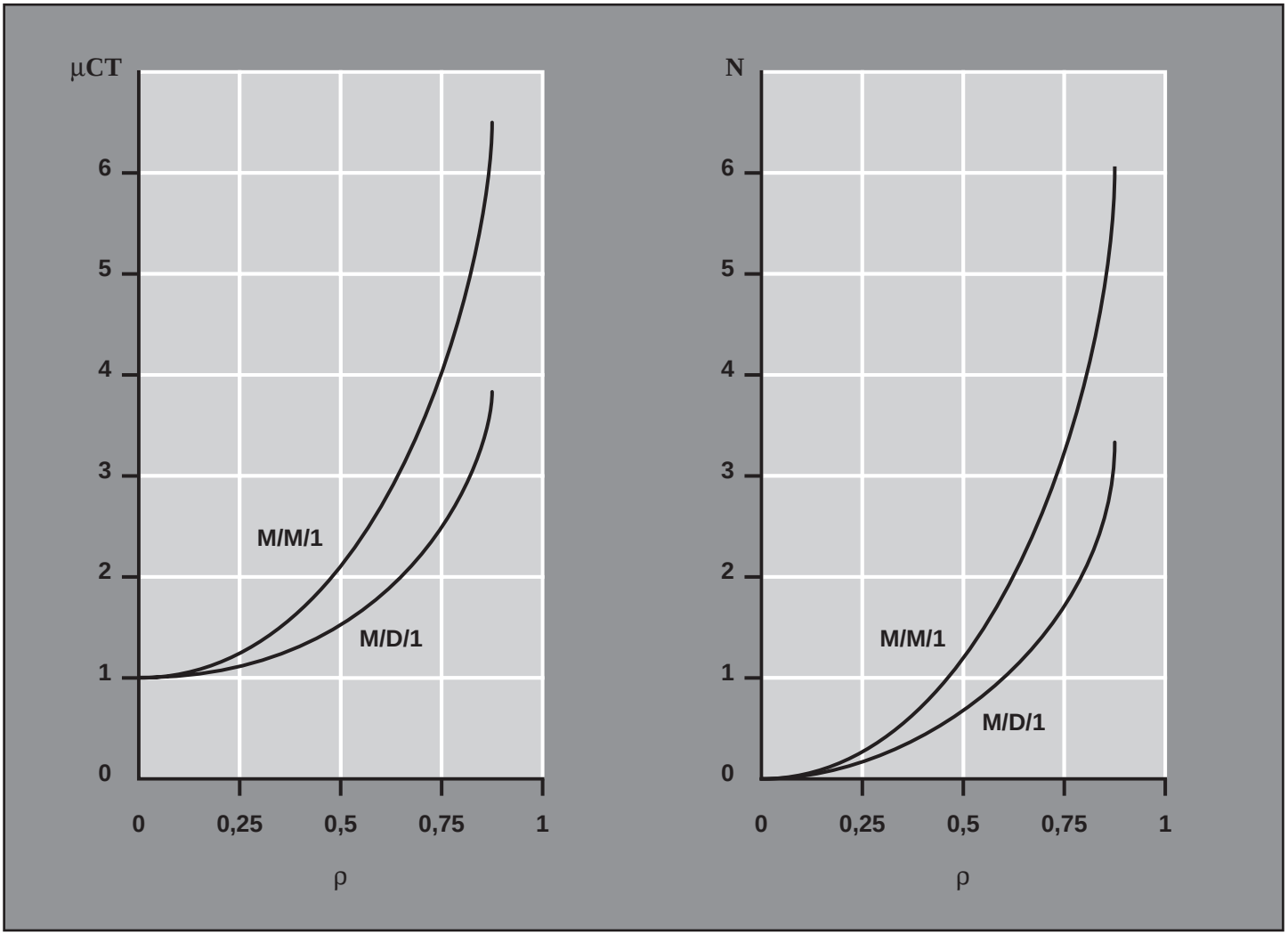


Bild 6 W M/M/1 und M/D/1

Wenn die Verkehrsintensität gegen 1 geht, wachsen N und µCT für beide Modelle ins Unendliche. Dies geschieht aber nicht ohne Vorwarnung, sondern wenn die mittlere Ankunftsrate λ in Richtung der Kanal(bearbeitungs)rate µC läuft. Hier entsprechen sich Modell und Wirklichkeit besonders gut: wenn λ dauerhaft durchschnittlich größer wird, als das, was das System leisten kann, läuft die Warteschlange (ins Unendliche) über und Aufträge werden in der Praxis weggeworfen. Es gibt auch verfeinerte Warteschlangenmodelle mit endlichen Warteschlangen, aber die brauchen wir hier gar nicht zu betrachten, weil schon eine dauerhafte Anhäufung von allzuvielen Paketen in der Warteschlange für unakzeptable durchschnittliche Delays sorgen würde.

Für ρ-Werte kleiner als 1 gibt es eine nichtlineare Beziehung zwischen N bzw. T und ρ. Die Kurven für die mittlere Nachrichtenverzögerung wurden angelegt, um die normalisierte Quantität µCT gegen ρ zu zeigen. Da 1/µC die durchschnittliche Nachrichtenübertragungszeit ist, ist T dividiert durch 1/µC , also µCT die durchschnittliche Verzögerung normalisiert zu (ausgedrückt in) Einheiten der durchschnittlichen Übertragungszeit.

1/µC ist übrigens auch das minimale durchschnittliche Delay, was erwartet werden kann, wenn eine Nachricht bei ansonsten leeren Puffer sofort drankommt.

Die Kurven für µCT sind ebenfalls nichtlinear und erreichen den Wert Unendlich für ρ gegen 1.

Die generelle Form der Kurven kann auch dadurch erklärt werden, dass bei geringen Werten für die Verkehrsintensität Nachichten selten im Puffer warten müssen und meist sofort bearbeitet werden. Sobald der Verkehr ansteigt, muss durchschnittlich länger gewartet werden. Dies kann man direkt an den Kurven von µCW und N_q sehen.

Ergebnisse

Wir wenden nun die Formeln an, um zu Ergebnissen hinsichtlich der Analyse der Leistungsfähigkeit von Gigabit Ethernet zu kommen.

Voraussetzungen:

• betrachtet wird die Bedienung vielfacher Eingänge (z.B. Clients) an einen Gigabit-Ausgang (z.B. Server, Uplink)

Leistungsanalyse: Gigabit Ethernet contra Policy Networks

- Der Switch hat eine Schaltmatrix, deren Leistung so hoch ist, dass sie in jedem Falle mehrere Gigabit-Ausgänge gleichzeitig behandeln kann ohne dass diese sich wesentlich wechselseitig beeinflussen. Es gibt solche Switches, z.B. den Corebuilder von 3Com mit 560 Gbit/s Backplane für max. 128 Gigabit Ports, den Cisco Catalyst 5500 mit 75 Gbit/s. Backplane für max. 32 Ports oder den Big Iron von Foundry Networks mit 256 Gbit/s. Backplane für max. 64 Ports. Auch verschiedene andere Hersteller haben Switches mit ähnlichen Verhältnissen zwischen Backplane/Matrix-Leistung und Switchports.

- Die Kanal-Bitrate C beträgt somit 1.000.000.000 Bit/s.

- Ein Ethernet-Paket kann minimal 64, maximal 1526 Bytes lang werden. Bei Gleichverteilung, aber auch bei Gaussscher Verteilung wäre der Mittelwert 731 Byte, also 5848 Bit. Um den Effekt der zwangsweisen Inter Frame Gap möglichst problemlos zu erfassen, runden wir dies auf 6.000 Bit auf. Die mittlere Nachrichtenlänge $1/\mu$ ist also 6000 Bit. Dadurch ergibt sich für μ der Wert $\mu = 1/6000 = 0{,}0001666$ 1/Bit

- Die mittlere Übertragungszeit ist

  $1/\mu C$ Sekunden.

  $\mu C = 166.600$ 1/s

  $1/\mu C = 0{,}000006$ s = 0,006 ms

Nach diesen Voraussetzungen können wir weitere Werte bestimmen. Zunächst hat mich einmal interessiert, was bei der geheimnisvollen Datenrate von 700 Mbit/s tatsächlich passiert.

Bei einer durchschnittlichen Paketlänge von 6000 Bits bedeuten 700.000.000 Bit/s. eine Ankunftsrate von

$\lambda_{0,5K,700M}$ = 116.666 Pakete pro Sekunde

Bei den Minimalpaketen von 64 Byte kämen wir OHNE Zuschlag für die Inter Frame Gap übrigens auf eine (in der Praxis völlig unrealistische) Rate von

$\lambda_{0,5K,700M}$ = 1.367.179 Pakete pro Sekunde, die uns aber für unsere Modellbetrachtung einen schönen Grenzwert liefert. In diesem Fall wären

$\mu = 0{,}002$,

$\mu C = 2.000.000$ und

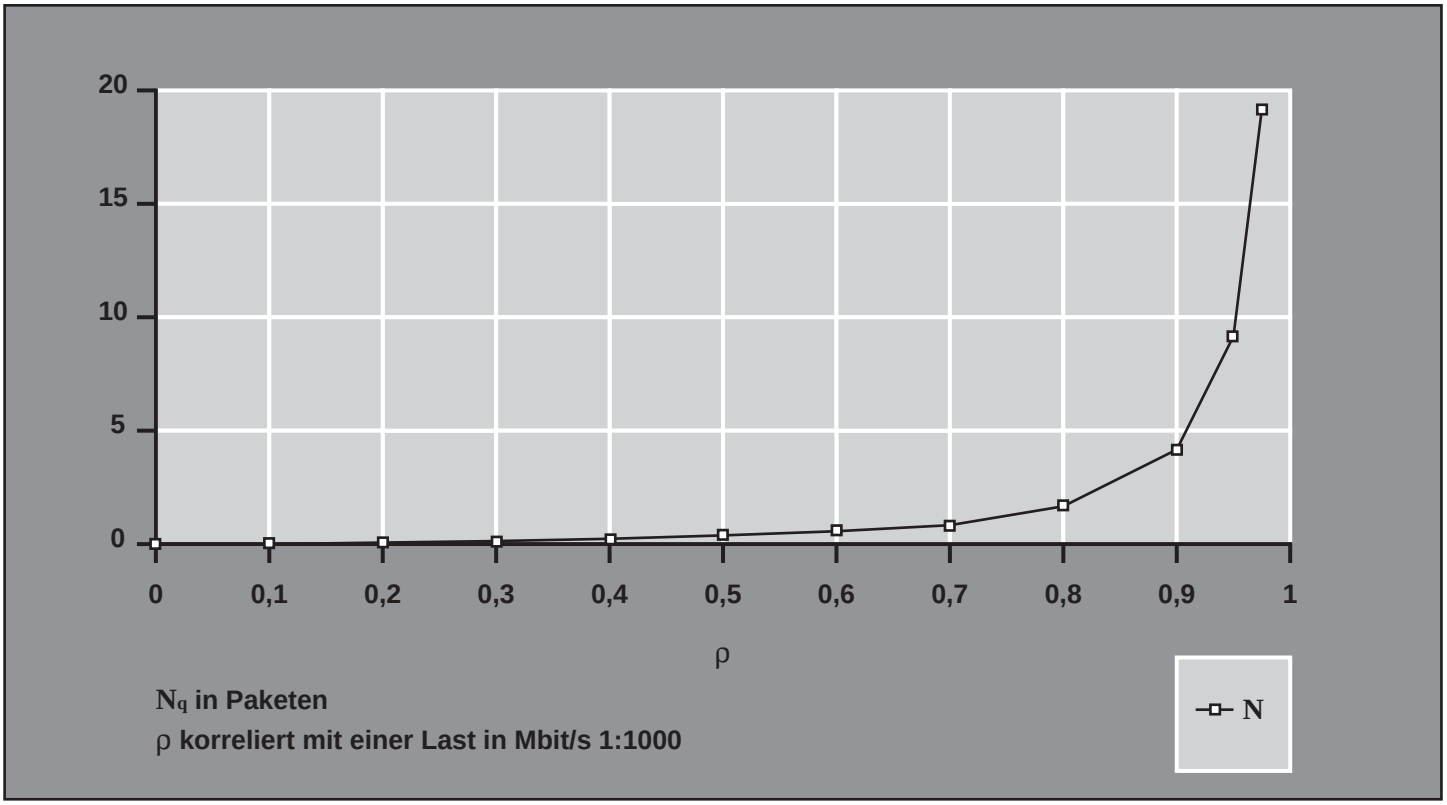
$1/\mu C = 0{,}000.000.5$ sec = 0.0005 msec.

Bleiben wir aber bei mittellangen Paketen und berechnen $\rho = \lambda / \mu C$

$\rho = 116.666 / 166.600 = 0{,}7$

Werden die Pakete kürzer, bleibt ρ i.w. gleich, wie Sie selbst leicht ausrechnen können. Das liegt einfach daran, dass der Bediener primär nach einer Datenrate bedient. Bei konstanter Datenrate kommen eben mehr kürzere oder weniger längere Pakete gleichzeitig an. Das ist ein ganz wesentliches Ergebnis, denn wenn wir jetzt Dinge ausrechnen, die nur auf ρ basieren, sind diese Ergebnisse ebenfalls unabhängig von der durchschnittlichen Paketlänge! Für die Praxis setzt das voraus, dass sich die Switches nicht an den kleinen Paketen verschlucken. Darum führt man Tests meist mit der kleinsten möglichen Paketlänge aus. Switches, die dann abschmieren, würde der Autor nicht empfehlen, weil sie nicht zu den stabilen Systemen gehören und mit ihnen alles Mögliche passieren kann.

Bild 7 Analyse Gigabit Ethernet: mittlere Warteschlangenlänge



Völlig unabhängig von der Nachrichtenlänge ist unter diesen Voraussetzungen die mittlere Anzahl von Nachrichten im Speicher (nach Formel 13):

$N_{q,700M} = 0,49 / 0,6 = 0,81$

Das Verhalten für andere Lastwerte zeigt sich in Bild 7. ρ korreliert mit der Last in Mbit/s bei Gigabit Ethernet im Verhältnis 1:1000.
Im stabilen Zustand befindet sich also bei einer Gesamtlast von 700 Mbit/s durchschnittlich noch nicht einmal ein einziges schlappes Paket in der Warteschlange, sondern nur ein Bruchteil. Die mittlere Wartezeit in der Schlange nach Formel 14 ist natürlich abhängig von der mittleren Paketlänge:

$W_{6K,700M} = 0,7 /99.960s = 0,000007s = 0,007ms$

Die mittlere Zeit T, die ein Paket im System verbringt, ist damit

$T_{6K,700M} = 1,3/99.960s = 0,000013s = 0,013ms$

Alle Systemzeiten über ρ zeigt Bild 8.

Viele Planer sind äußerst besorgt über das mögliche Delay. Es wird z.B. ein maximales Delay von 20 ms gefordert. Um dies zu erreichen, müssten sich aber schon über 1.500 mittellange Pakete in der Warteschlange stauen. Bei miminal langen Paketen fällt T auf ca. 0,001 ms und man müßte schon fast 40.000 von diesen Päckchen anstauen, um auf 20 ms Delay zu kommen. Und das schafft man nur, wenn man ρ längere Zeit ganz nahe gegen 1 treibt.

Gigabit Ethernet und Fast Ethernet, Priorisierung

Wendet man das gleiche Modell (mehrere Clients, ein Server-Uplink, stabiles System) auf einen Fast Ethernet statt auf einen Gigabit Ethernet Switch an, ändert sich lediglich die Bearbeitungsgeschwindigkeit C. Die Paketlänge bleibt bei durchschnittlich 6000 Bits,

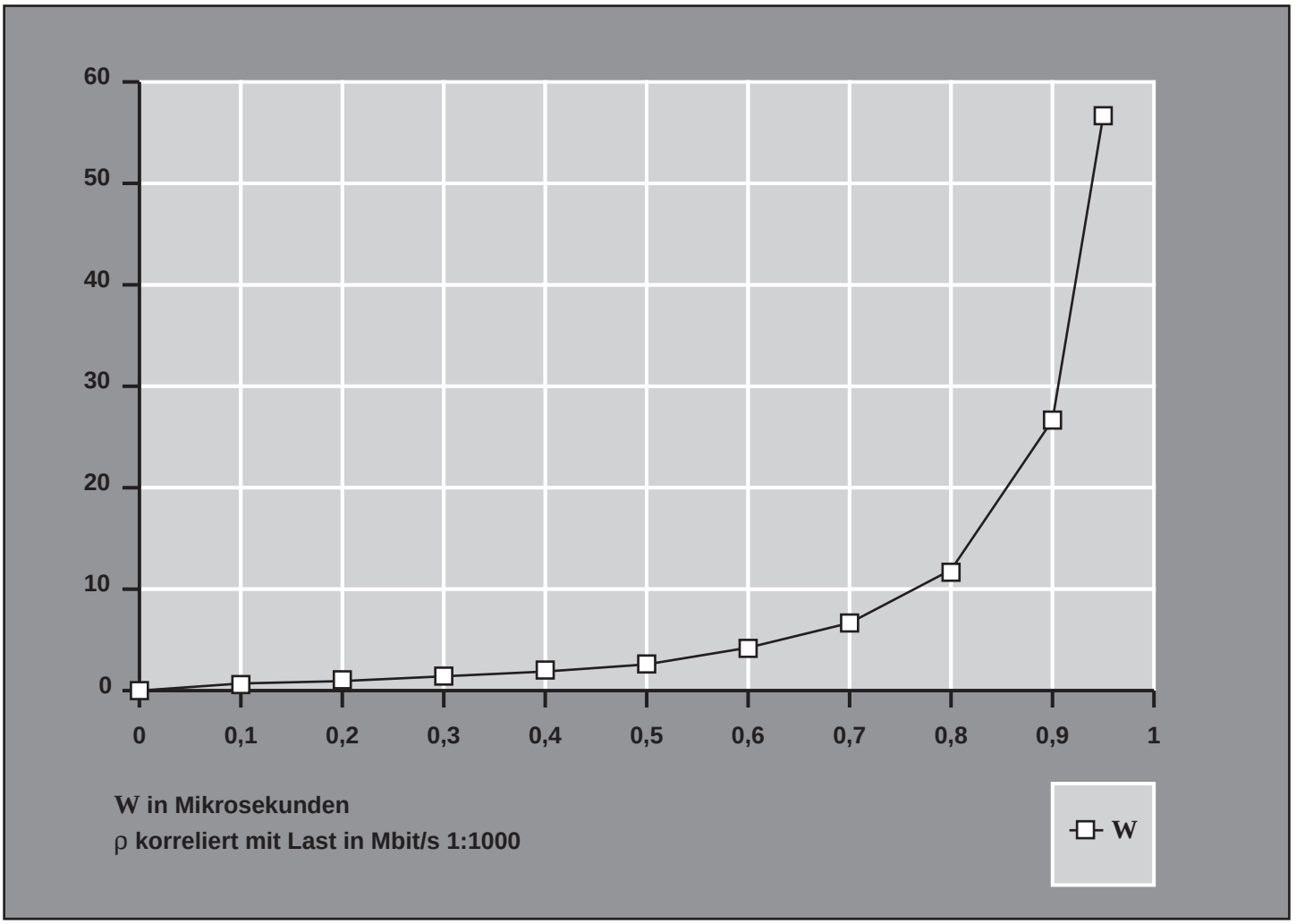
$1/\mu = 6000,$

$\mu C = 16.660,$

$1/\mu C = 0,00006$ und

$\lambda = 11.666$ Nachrichten pro Sekunde bei einer Gesamtlast von 70 Mbit/s. Da ist es auch kein Wunder, dass die Werte für W in µsek genau das Zehnfache ausmachen.

Bild 8 Analyse Gigabit Ethernet: mittlere Systemzeit



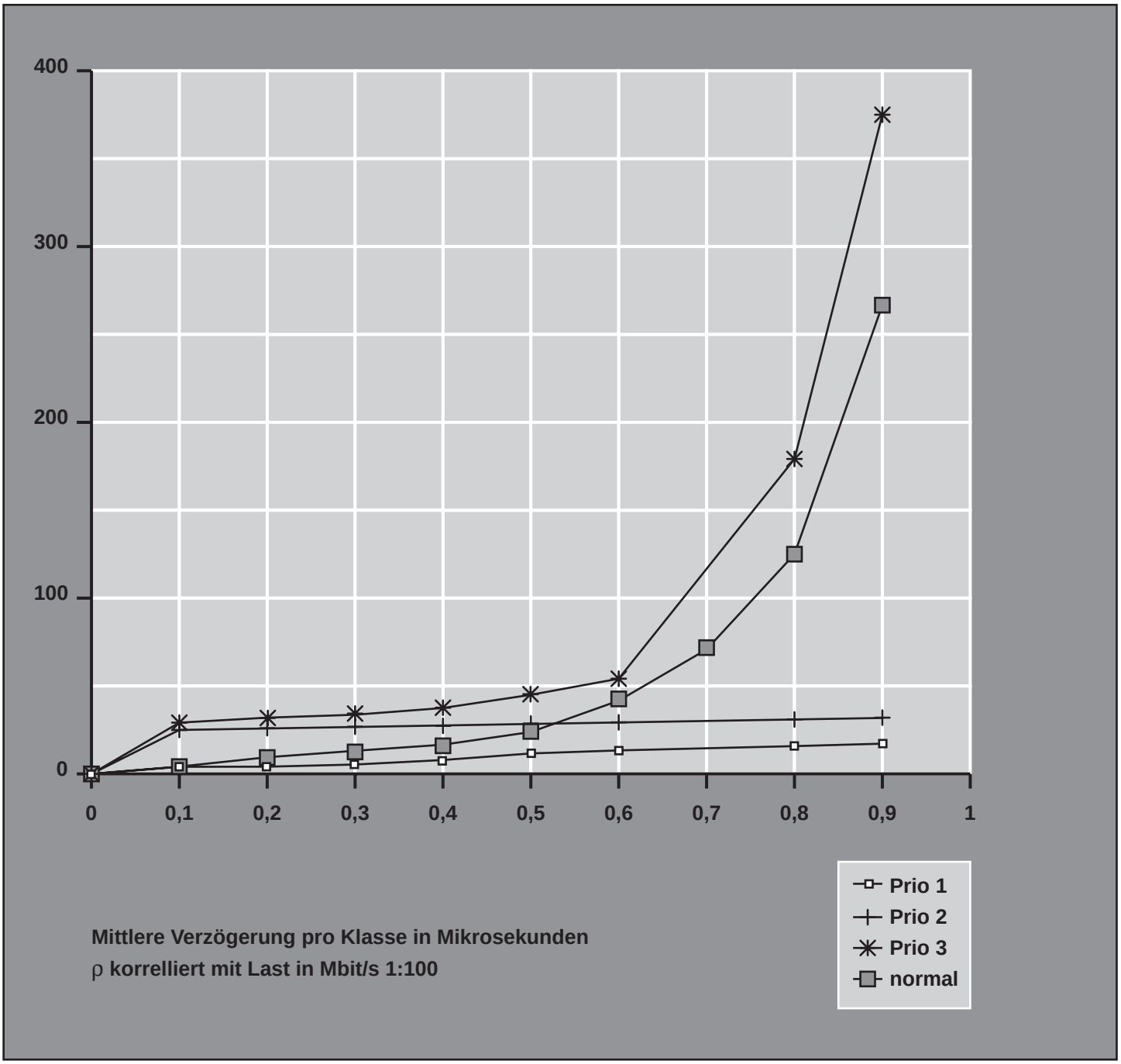
Erst bei einer Auslastung von ca. 96 % überschreitet W den Bereich zur Millisekunde. Schafft man es also, die Last in einem vernünftigen Bereich zu halten, sagen wir bis 70 Mbit/s., arbeitet auch ein Fast Ethernet Switch sehr gut.

Für die Bewertung der Priorisierung gibt es ebenfalls Modelle, die aber relativ komplex sind, so dass wir hier keine Formeln mehr zeigen können. Im Grunde genommen geht es bei diesen Modellen um die Frage, inwieweit die Pakete der einzelnen Priorisierungsklassen insgesamt verzögert werden. Es gibt auch z.B. bei Franta Simulationsmodelle.

In der Abbildung (Bild 9) sehen wir, wie sich ein M/G/1-Modell mit drei Klassen entwickelt. Zum Vergleich wurde der Verlauf der mittleren Wartezeit ohne Priorisierung gezeichnet.

Bis zu einer Auslastung von ca. 60% ist es eigentlich gleichgültig, in welcher Prioritätsklasse man sich befindet: die Klasse 1 wird etwas bevorzugt, die Klassen 2 und 3 liegen etwas schlechter als die Funktion ohne Priorisierung. Jenseits dieser Auslastung ändert sich das Bild aber dramatisch, da die Schlangen der Prios 1 und 2 nach wie vor sehr gut bedient werden, die Klasse 3 aber deutlich schlechter dran ist als ohne Priorisierung.

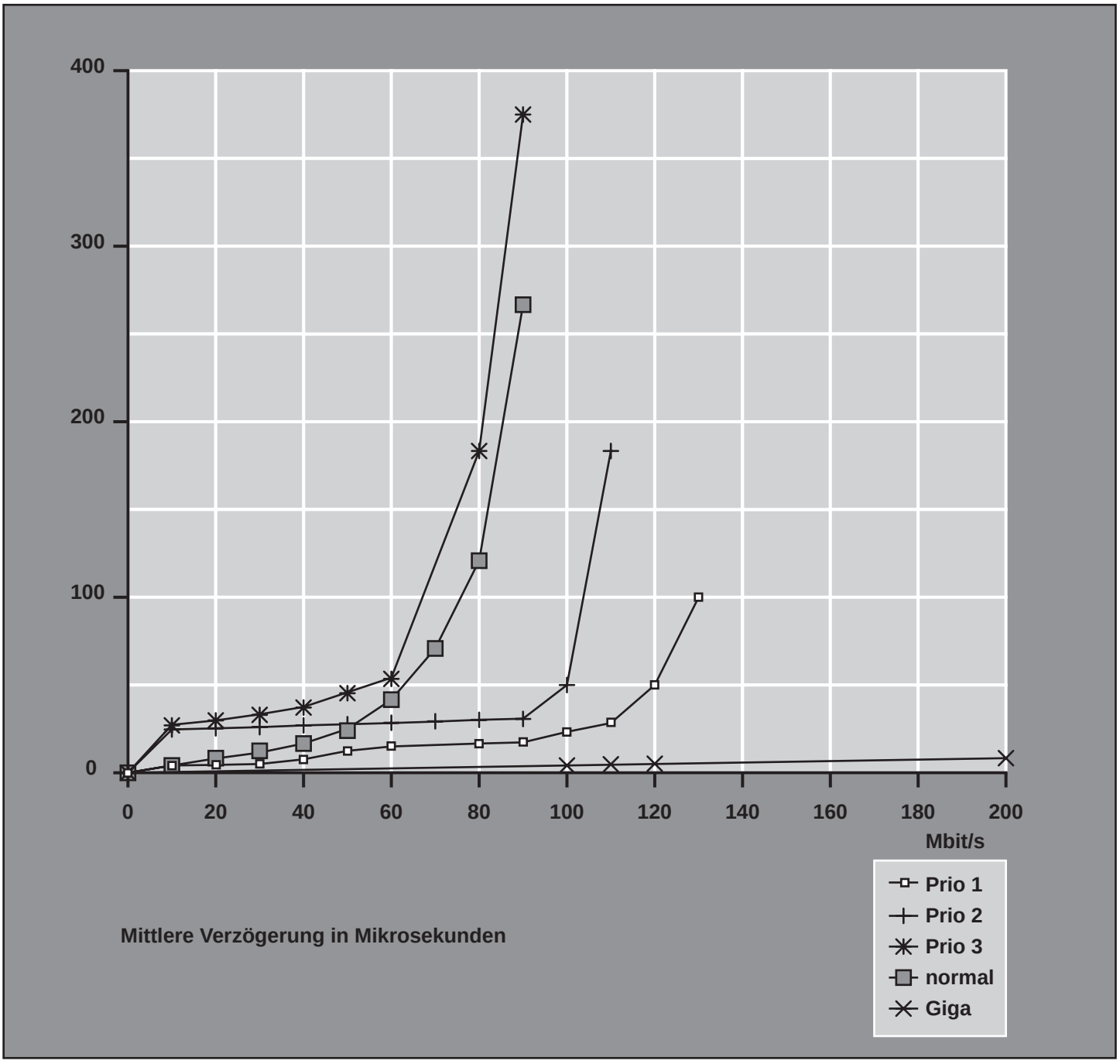
Bild 9 Priorisierung bei Fast Ethernet: M/G/1-Modell für drei Klassen



Dieses Modell zeigt den ganzen Wahnsinn der aktuellen Diskussion: in IEEE 802.1 wird über bis zu acht Prioritätsklassen geredet. Aber schon die dritte zeigt ein schlechtes Verhalten, so dass man auf alle, die noch geringere Priorität haben, wirklich verzichten kann.

Aber, damit noch nicht genug der falschen Annahmen. Viele Benutzer glauben, dass eine Priorisierung sie vor Engpässen in Überlastsituationen schützen könnte. Das stimmt auch, aber mit wesentlich deutlicheren Grenzen als die Werbung glauben macht. In der nächsten Abbildung (Bild 10) sehen wir, dass nur Leistung auf Dauer wirklich etwas bringen kann. Es wurde der Bereich von 0 bis 200 Mbit/s. aufgezeichnet. Bei 70 Mbit/s. verabschiedet sich die dritte Prioritätsklasse ins Nirvana, bei 85 Mbit/s werden die Werte für einen ungesteuerten Kanal wirklich schlecht. Die ersten beiden Prioritäten halten aber noch aus, bis in den Überlastfall hinein, die

Bild 10 Priorisierung bei Fast Ethernet vs. nacktem Gigabit: M/G/1-Modell für drei Klassen vs. Gigabit



zweite verabschiedet sich jenseits der 10% Überlast, die dritte bei ca. 30% Überlast. Diese Ergebnisse kommen aus der Simulation, weil ja ohnehin kein System im stabilen Zustand mehr vorliegt. Früher oder später wird der Switch einfach den Geist aufgeben, wenn man an der Überlast nichts ändert, da hilft auch die Priorisierung nichts mehr. Ganz unten diese flache Linie gehört zu einem Gigabit Ethernet Switch, der noch bei 10% Auslastung ist, wenn der Fast Ethernet Switch schon zufährt.

Zwischenfazit

Die Priorisierung hilft im Überlastfall nur bedingt, bis 70% Nominallast ist sie dagegen völlig überflüssig. Wer sein Fast Ethernet durchschnittlich pro Switch-Port nur bis 70% auslastet, kann dabei bleiben. Wer jetzt schon Befürchtungen hat, deutlich über 100 Mbit/s zu kommen, der sollte sich mit Gigabit Ethernet befassen, denn jenseits der 70 Mbit/s. wird es einfach heikel. Das Delay ist hingegen in diesen Fällen nicht so entscheidend, die gefürchtete Varianz des Delays kann man aber mit Priorisierung nur wenig eindämmen. Im Gegenteil: eine schlechte Prioritätsklasse kann zu einer wesentlich höheren Varianz führen als die Benutzung eines ungesteuerten Netzes.

Verkettete Switches

Normalerweise wird man ein Netz nicht nur aus einem einzelnen Switch bilden. Kleinrock hat netterweise gezeigt, dass M/M/1-Systeme unabhängig voneinander arbeiten, wenn man sie hintereinander hängt. Ausgehend von M/D/1 müssen wir das Ergebnis für die mittlere Systemzeit leider verdoppeln. Dann ist das Delay einer Kette gleich der Summe der Delays der einzelnen Komponenten. Bei Gigabit Ethernet befinden wir uns aber in einem Bereich, in dem wir noch auf etwas anderes achten müssen: die Signallaufzeit. Im Festkörper laufen die Signale etwa mit 200.000 km/s.

Das bedeutet aber, dass sie 1 µs auf 200 m benötigen. Verschiedene Hersteller bieten Extender an, mit denen die Grenzen des Standards für Gigabit Ethernet Glasfaserverbindungen wesentlich erweitert werden können. Gigabit Ethernet Switches arbeiten bei geringer Auslastung im Bereich weniger µs. Dehnt man die Entfernungen wesentlich aus, muss man damit rechnen, dass die Signallaufzeit eher zum limitierenden Faktor wird. Und auch daran wird Priorisierung nichts ändern können, denn das würde ja bedeuten, dass wir die Lichtgeschwindigkeit beeinflussen könnten. Und mit Einstein sollte man sich als einfacher Lokalnetzer wirklich nicht anlegen.

Literatur

Franta, W.R., Chlamtac, I.: Local Networks, Lexington Books, Lexington, Mass, 1981

Hammond, J.L, OReilly, P.J.P: Performance Analysis of Local Computer Networks, Addison Wesley Publ. Comp. , Reading, Mass. 1986

Suppan, Jürgen: Änderungen im Netzwerk-Design: Service-Level-Design contra Policy-based-Networking in Der Netzwerk Insider, August 1999, ComConsult Technologie Information

Kauffels, F.-J.: Lokale Netze, 11. Aufl., MITP Bonn 1999

Kleinrock, L: Queuing Systems, 2Bde, Wiley-Intersciences, New York 1976

Verzeichnis der verwendeten Parameter und Abkürzungen

$\alpha(t)$	Anzahl der ankommenden Nachrichten in einem Intervall (0,t)
$\delta(t)$	Anzahl der abgehenden Nachrichten in einem Intervall (0,t)
λt	Mittlere Ankunftsrate über ein Intervall (0,t)
λ	Mittlere Ankunftsrate
μ	Kehrwert der mittleren Nachrichtenlänge
ρ	Mittlerer Zeitanteil, an dem der Bediener beschäftigt ist
C	Kanal-Bitrate (brutto = Nominalleistung)
D	"deterministic" (bei Warteschlangenmodellen)
G	"general" (bei Warteschlangenmodellen)
M	"memoryless" (bei Warteschlangenmodellen)
N	mittlere Anzahl der im System befindlichen Nachrichten(paketen)
N_q	mittlere Anzahl der in der Warteschlange befindlichen Nachrichten(pakete)
N(t)	Wachstum der zwischengespeicherten Nachichten in einem Intervall (0,t)
T	durchschnittliche Gesamtzeit im System
W	durchschnittliche Gesamtzeit im Puffer

Anzeige

Ausbildung zum ComConsult Certified Network Engineer

Sonderaktion 1999: Palm IIIx

Drei 5-tägige Seminare und das 3-tägige Seminar*
zum Paketpreis von DM 11.900,-- oder EUR 6.084,37
(statt DM 13.970,--) zzgl. MwSt.
*** Bei Buchung erhalten Sie als Bestandteil des Seminarpaketes zur Planung Ihrer Akademie-Termine den 3Com »Palm IIIx«**

Vier 5-tägige Seminare*
zum Paketpreis von DM 12.800,-- oder EUR 6.544,54
(statt DM 14.560,--) zzgl. MwSt.
*** Bei Buchung erhalten Sie als Bestandteil des Seminarpaketes zur Planung Ihrer Akademie-Termine den 3Com »Palm IIIx«**

ComConsult Akademie

Nutzen Sie unsere Paketpreise!

Zwei 5-tägige Seminare und das 3-tägige Seminar zum Paketpreis von DM 9.250,-- oder EUR 4.729,45 (statt mind. DM 10.180,--) zzgl. MwSt.

Drei 5-tägige Seminare zum Paketpreis von DM 9.800,-- oder EUR 5.010,66 (statt DM 10.770,--) zzgl. MwSt.

Lokale Netze für Einsteiger

Einzelpreis: DM 3.500,-- (EUR 1.789,52) zzgl. MwSt.

- **11.10. - 15.10.99 Aachen**
- **22.11. - 26.11.99 Köln**
- **06.12. - 10.12.99 Aachen**
- **24.01. - 28.01.00 Aachen**
- **14.02. - 18.02.00 Aachen**
- **13.03. - 17.03.00 Aachen**
- **10.04. - 14.04.00 Aachen**
- **15.05. - 19.05.00 Aachen**
- **05.06. - 09.06.00 Aachen**
- **26.06. - 30.06.00 Aachen**

Internetworking

Einzelpreis: DM 3.790,-- (EUR 1.937,80) zzgl. MwSt.

- **14.02. - 18.02.00 Aachen**
- **27.03. - 31.03.00 Aachen**
- **05.06. - 09.06.00 Aachen**

Neue Ethernet Technologien

Einzelpreis: DM 3.690,-- (EUR 1.886,67) zzgl. MwSt.

- **29.11. - 03.12.99 Aachen**
- **27.03. - 31.03.00 Aachen**
- **22.05. - 26.05.00 Aachen**

TCP/IP und SNMP

Einzelpreis: DM 2.990,-- (EUR 1528,76) zzgl. MwSt.

- **25.10. - 27.10.99 Königswinter**
- **13.12. - 15.12.99 Hamburg**

Ab dem Jahr 2000 bieten wir »TCP/IP und SNMP« mit erweitertem Inhalt als 5-tägiges Intensiv-Seminar an.

Einzelpreis von DM 3.580,-- (EUR 1.830,42) zzgl. MwSt.

- **07.02. - 11.02.00 Bonn**
- **10.04. - 14.04.00 Berlin**
- **05.06. - 09.06.00 Hamburg**