

Sparen mit Datenkompression

In unserer national und international vernetzten Welt spielen Mietleitungen eine wichtige Rolle. Nicht zuletzt fallen sie als Budgetposten immer mehr ins Gewicht. Die Mietleitungskosten sind direkt abhängig von der Leistung (Übertragungskapazität) und der Leitungslänge. Gelingt es nun, mittels geeigneter Technologien, die benötigte Übertragungskapazität zu reduzieren ohne den Datendurchsatz zu beeinträchtigen, so können billigere Leitungen eingesetzt oder die bestehenden Leitungen bei gleichen Kosten besser ausgenützt werden. Im folgenden wird nur die Kompression für direkte, transparente Verbindungen erläutert. Datenkompression für die Datenübertragung auf öffentlichen Netzen, wie X.25, Frame Relay, ISDN, etc. wurde nicht berücksichtigt. Diese können jedoch mit ähnlichen Methoden beschleunigt werden.

Effizienzsteigerung

Verschiedene Faktoren beeinflussen direkt die Effizienz einer Datenübertragung. Will man die Leitungskosten reduzieren, müssen alle Faktoren angemessen berücksichtigt werden.

Schlussendlich sollen über eine gegebene Leitung mit einer zusätzlichen Einrichtung netto mehr Daten übertragen werden können. Die daraus resultierenden Ersparnisse sollen eine rasche Amortisation der zusätzlichen Investitionen ermöglichen.

Oft wird auch festgestellt, dass der Engpass nicht die Übertragungsleitung ist. Faktoren ausserhalb der eigentlichen Übertragung, z.B. ein langsames Applikationsprogramm, können für die Verzögerung verantwortlich sein. Der grösste Teil der Probleme bei zu langsamen Reaktionszeiten könnten durch geeignetere Programmierung der Datenbanken behoben werden. Wie z.B. Proxy-Server oder Datenbank Doppel die gegenseitig abgeglichen werden.

Die Kompression allein ist wertlos, sofern nicht auch dazu die Verzögerungszeit, die Art der Komprimierung und der angewandte Algorithmus beachtet werden. Grosse Kompressionsfaktoren, die in Datenblättern angepriesen werden, sind mit äusserster Vorsicht zu betrachten. Messungen mit einer Ethernet Bridge eines bekannten Herstellers ergaben, dass die Kompressionsfaktoren über alles gemessen bei weitem nicht seinen Angaben entsprachen. Während die Bridge bei 9,6kbit/s noch einen Faktor 5 erbrachte, reduzierte sich der Faktor bei 64kbit/s auf 2. Bei 2Mbit/s benötigte das Gerät für die Kompression mehr Zeit als für die Übertragung ohne Komprimierung. Für die Messungen, die zu den Herstellerangaben führten, waren vermutlich ideale Bedingungen kreierte worden. D.h. Text-Daten lassen sich gut und gerne um den Faktor 10 - 20 komprimieren, jedoch Bild-Daten (die meist bereits in einem komprimierten Format vorliegen) erreichen nur eine sehr geringe Komprimierung von ca. 1.1!

Kompression

Die Datenkompression ist, wie der Begriff richtig ausdrückt, eine Verdichtung der Daten. Dabei sind Daten im heutigen Sinne als Bits und Bytes zu verstehen, deren Ursprung Daten sein können, aber nicht unbedingt sein müssen. Die Kompression basiert auf einem 8-Bit Byte. Übermittelt werden die Daten in Paketen, getrennt von Fullcharakteren (idle characters).

Je nach Häufigkeit der auftretenden Charakter oder Sequenzen von Charakteren, teilt man ihnen einen Kurzcode zu. Damit reduziert sich die Anzahl Bits für dieselbe Information. Schon Samuel Morse hatte 1844 genau dieses Konzept angewandt. Je nach der Häufigkeit der Buchstaben verwendete Morse in seinem Code 1,2,3 oder 4 Zeichen.

Die heute verwendeten Algorithmen sind meist eine Variante des in den 70er Jahren von Abraham Lempel und Jacob Ziv entwickelten Kompressions-Algorithmus. Der Lempel-Ziv Algorithmus erstellt eine Tabelle repetitiver Charakter und Sequenzen, denen ein kurzer Code mit fixer Länge, genannt Token, zugeteilt wird. Nur diese Token werden übertragen und auf der Empfängerseite, mit Hilfe derselben Tabelle, die ursprünglichen Daten wieder hergestellt.

Soll die Datenkomprimierung effizient sein, muss diese Tabelle dynamisch, d.h. in regelmässigen Abständen den Daten angepasst werden. Dies bedingt jedoch, dass jede Änderung in der Tabelle der Empfängerseite mitgeteilt und damit auch übertragen wird. Je häufiger die Tabelle angepasst wird, desto öfter muss sie übertragen werden und desto mehr unproduktive Daten sind auf der Leitung. Auch die Länge der reduzierbaren Sequenzen hat einen direkten Einfluss auf die Effizienz der Komprimierung. Lange Sequenzen ergeben eine hohe Komprimierung, jedoch auch grosse Verzögerungen. Je länger die Sequenz

desto grösser ist die Verarbeitungszeit, da die ganze Sequenz erst in den Speicher eingelesen werden muss bevor ein Token zugeteilt und gesendet werden kann.

Ein positiver Nebeneffekt der Komprimierung ist eine einigermaßen sichere Verschlüsselung der Daten. Da die Tabellen dynamisch verändert werden, benötigt man zum Entschlüsseln genau dasselbe Gerät in gleicher Konfiguration. Die fehlenden Standards helfen in diesem Fall der Sicherheit. Bereits sind schon Produkte auf dem Markt, die zusätzlich zur Komprimierung die Daten auch verschlüsseln.

Komprimierungsarten

Es gibt zwei verschiedene Komprimierungsarten, eine reine Software Lösung und eine mehr oder weniger auf Hardware basierende Technologie. Die reine Software Lösung ist genügend für Übertragungsgeschwindigkeiten bis max. 64kbit/s. Bei höheren Geschwindigkeiten wird die Verarbeitungszeit (latency) zu gross. Hardware Lösungen sind heute für Geschwindigkeiten von bis zu 8Mbit/s erhältlich. Es bleibt noch die Frage, ob ein spezielles Gerät für die Datenkompression oder eine in Bridge/Router integrierte Option eingesetzt werden soll. Bei der integrierten Variante ist noch zu unterscheiden, ob das ganze Paket oder auch nur die Daten ohne Rahmen (frame) komprimiert werden können. Wird das ganze Paket komprimiert, so ist der Einsatzbereich, wie für das spezielle Gerät, auf Punkt-zu-Punkt Anwendungen begrenzt. Werden nur Daten ohne Rahmen behandelt, hat das den Vorteil, dass die Routing-Informationen in Klartext vorhanden sind. Die Pakete können so über mehrere Knoten gesendet werden, ohne dass eine Dekomprimierung stattfinden muss. Dieser Vorteil ist durch den Mangel eines Standards aber sehr minim, da er nur in einem homogenen Netzwerk zum tragen kommt. In homogenen Netzwerken müssen alle Knoten kompatibel und vom gleichen Hersteller sein. In heterogenen Netzwerken ist die Datenkompression auf Punkt-zu-Punkt Anwendungen limitiert. Generell kann festgestellt werden, dass eine integrierte Datenkompression wesentlich mehr Verarbeitungszeit benötigt als ein spezialisiertes Gerät. Der Grund ist, dass Mikroprozessor und Speicher primär für andere Aufgaben eingesetzt sind und dass meistens eine reine oder aber mindestens partielle Softwarekomprimierung angewendet wird.

Die auf Datenkompression spezialisierten Geräte verwenden die äusserst schnelle Hardwarekomprimierung. Dies bringt dafür den Nachteil von zusätzlichen Elementen in der Übertragungsstrecke. Ein kleiner, dank der Zuverlässigkeit der heutigen Technologien, fast vernachlässigbarer Nachteil.

Ein weiterer zu beachtender Faktor ist die angewandte Methode der Fehlerkorrektur. Bei jeder Übertragung können Fehler entstehen. Eine spezielle Kontrolle erlaubt das Erkennen von Fehlübertragungen. Nach einem festgelegten Modus quittiert der Empfänger dem Sender den korrekten oder aber inkorrekten Erhalt der Daten. Bis zum Eingang dieser Quittung müssen die unkomprimierten Daten im Speicher bleiben um sie eventuell nochmals senden zu können. Je grösser die komprimierten Sequenzen sind und je mehr Datenpakete gesendet werden dürfen, bis eine Quittung abgewartet werden muss, desto mehr Zeit wird für die Korrektur verwendet und desto weniger Zeit steht für Nutzdaten zur Verfügung.

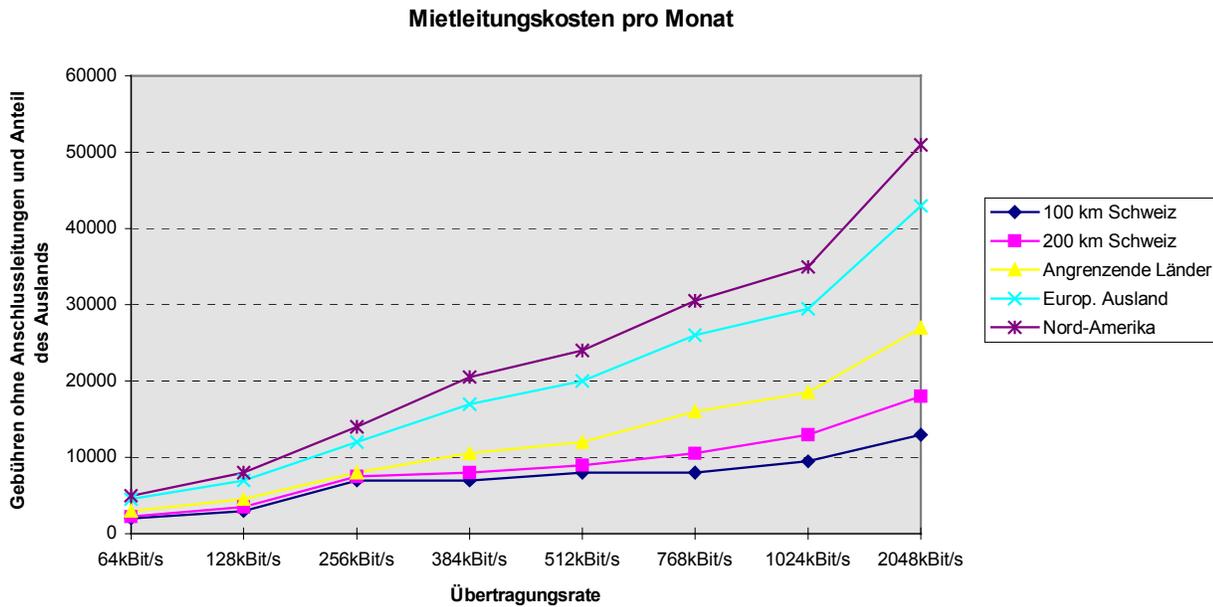
Standards

Bis heute bestehen keine Standards für die Datenkompression. Alle erhältlichen Systeme verwenden eine proprietäre Technologie. Unter dem Namen "Synchronous Data Compression Consortium" (SDCC) haben sich etwa zwei Dutzend, in der Datenkompression tätigen Firmen zusammengeschlossen. Das Ziel ist, sich erst auf einen standardisierten Algorithmus zu einigen. für später ist geplant, einen eigentlichen Standard zu erarbeiten der u.a. auch Fehlererkennung und Fehlerkorrektur enthalten soll. Bis zur Verabschiedung eines Datenkompressions-Standards können aber noch einige Jahre vergehen.

Sparpotential

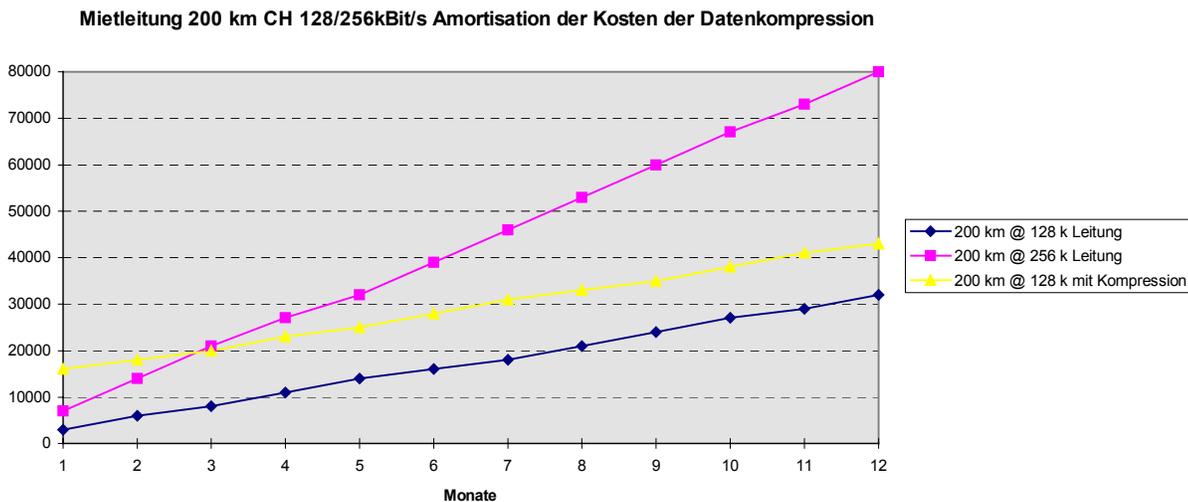
Die Mietleitungskosten sind ein enormer Budgetposten. Die Grafik 1 zeigt die monatlichen Gebühren für eine Mietleitung von 200 km in der Schweiz und je eine Leitung in das benachbarte Ausland (angrenzende Staaten und BENELUX), in ein entferntes europäisches Land und nach Nordamerika, in Abhängigkeit der Übertragungsrate. Auffallend ist dabei der sprunghafte Anstieg der Kosten zwischen 128 kbit/s und 256 kbit/s. In dieser Grafik wurden für alle Auslandsleitungen nur die Mietgebühren ohne lokalen Anschluss und ohne Anteile des Auslandes berücksichtigt. Die effektiven Kosten sind daher um einiges höher.

Grafik 1



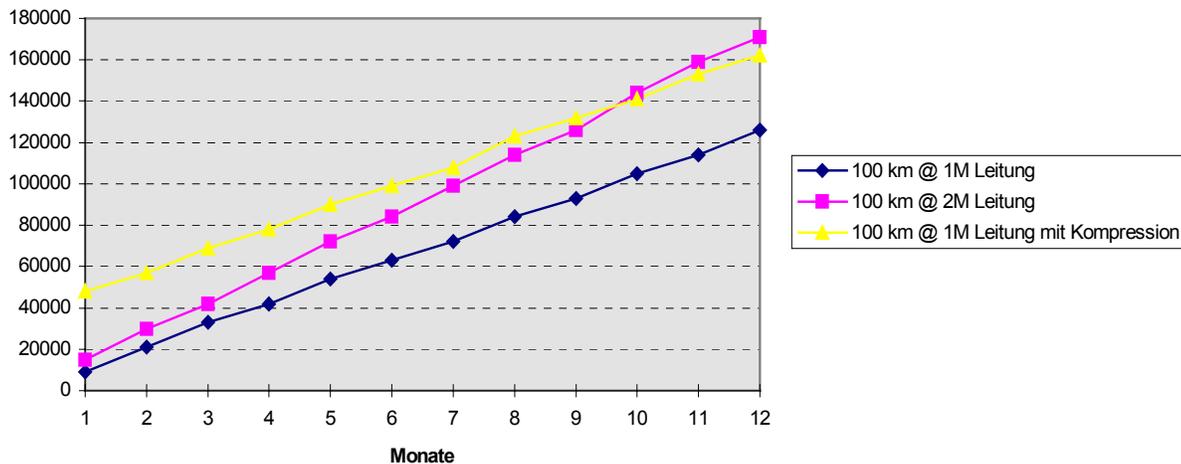
Um die Rentabilität einer Installation von Kompressionsgeräten berechnen zu können, muss von einem minimalen Kompressionsfaktor ausgegangen werden. Unter Berücksichtigung der vorhin erklärten Faktoren und unter der Annahme des Einsatzes spezialisierter Geräte sowie unter Beachtung eines Sicherheitsfaktors, kann mit einer minimalen Kompression von Faktor 2 gerechnet werden. D.h., dass die Übertragungsraten mindestens halbiert werden können.

Grafik 2



Grafik 3

**Mietleitung 100 km CH 1024/2048 kBit/s
Amortisation der Kosten der Datenkompression**



Die Auswirkung der reduzierten Übertragungsgeschwindigkeit auf die Mietleitungskosten zeigen für zwei Beispiele die Grafiken 2 und 3. Beide Beispiele betreffen Inlandleitungen. Einsparungen auf Auslandleitungen sind noch markanter. Die Grafik 2 hat zur Annahme eine 200 km Mietleitung, die von 256 kbit/s auf 128 kbit/s reduziert werden kann. Die Kosten für die Kompressionsgeräte, inklusive deren Installation von ca. Fr. 13'000.- sind innert 3 Monaten amortisiert. Danach bringt jeder Monat eine Ersparnis von ca. Fr. 3'900.-. Grafik 3 zeigt die Verhältnisse bei einer 100 km langen 2 Mbit/s Leitung die auf 1 Mbit/s reduziert wurde. Die Investitionen werden hier innerhalb von weniger als 10 Monaten amortisiert und nachher jeden Monat Einsparungen von ca. Fr. 3'500.- realisiert.

Schlussbemerkung

Das Sparpotential in der Datenkommunikation ist bei Mietleitungen sehr beachtenswert. Die Randbedingungen sind jedoch sorgfältig zu überprüfen. Unter Berücksichtigung aller Faktoren darf mit einer Halbierung der benötigten Übertragungsrates gerechnet werden. Sehr zu empfehlen ist ein Versuchsbetrieb bei dem die Übertragungszeiten definierter Daten mit und ohne Kompression gemessen werden. Bei der Verwendung von effizienten, externen Kompressionsgeräten beträgt die Amortisationszeit der Zusatzinvestitionen durch die Kostenreduktion in der Regel weniger als 12 Monate.