What is Compression?

This white paper is aimed at providing a high level overview of compression in relation to Opteq iQ. It is not intended to be an exhaustive technical explanation or definition of the solution but rather a business oriented overview and insight to the philosophy and design principles of Opteq iQ and the benefits of these. This document is best read in conjunction with the Opteq iQ module brochures and datasheets.

Increasing Real Throughput with Compression.

Can I really get more the 2Mb/s on a T1? The laws of physics say that 1.544Mb/s is 1.544Mb/s. However physics doesn't consider that the data may be compressed, then transported, and then decompressed, creating more than just an illusion of greater speed. Throughput is a function of time, and if you indeed do decrease the time to transport any given piece of data, then the gain is real. You really have shrunk the time of acquiring the data, and thus increased the throughput beyond what the physical medium might dictate. So yes, you really can get more than 2Mb/s on a T1.

What is compression?

Compression is an organized reduction in the volume of data, where "organized" implies that you can "decompress" it to get something similar to or exactly the same as the original. You use compressed data all of the time. You use "zip" files all of the time. Your system backups are compressed. You might even be unlucky enough to have used a compressed file system. But it's more than just that. Virtually every graphic file is compressed in one form or another. Just about all multimedia is compressed. Without compression there is too much information to transport over non-dedicated networks.

Newton's Telecom Dictionary defines compression as *"reducing the bandwidth or number of bits needed to encode information or a signal, typically by eliminating long strings of identical bits or bits that do not change in successive sampling intervals."* Compression works as a kind of shorthand to cram more traffic onto the same pipe. That traffic is then decompressed at the destination without changing the actual encoded data.

There are 2 general "kinds" of compression: "lossy" and "lossless". "Lossy" compression allows for decompression that yields something that is not exactly what was originally compressed. JPEG (.jpg) images are the most common example. With graphics, if a few dots here and there are missing or wrong, your eyes can't tell the difference generally. So it doesn't matter that much. With audio and video, a minced bit here and there won't be noticed, unless you are planning on saving it many times (as if you are mixing audio for a CD). For data, lossless compression is required. You can't have a bit in an IP packet lost, or the checksum will fail or the file probably won't work.

You might ask why anyone would ever use "lossy" compression as opposed to "lossless", since it seems a lot better to get exactly was originally compressed rather than something slightly different. The answer is that lossless compression has limits as to what compression ratios can be obtained and are much less than lossy compression. You can cheat more with lossy compression, so you can get much better compression ratios. But you don't get exactly what you started with, so it's no good for the purposes discussed here.

The Opteq iQ implementation.

Opteq iQ implements selective compression between ANY 2 POINTS in a data stream that have a unit. It doesn't matter what is in between, only that the data that is compressed also passes through the second unit so it can be decompressed back to its original content before it reaches the target device. If it seems pretty simple, it's because it is. If you have an Opteq "box" at your office and you have one at home, you can compress all of the traffic from your office network to your home PC transparently. So if you have a 256kbs connection, it will appear to be 2-3x faster, in most cases. The same holds true for inter-branch traffic.

About Compression Ratios

OK, so you read the "in most cases" above and now you're wondering what's going on. Why is there a conditional attached to this? Well, you have to know something about Internet content and compression to understand. When you read about compression products the first thing that most people check is the compression ratio. Vendors advertise "4:1 compression" or "8:1 compression", but you have to be very careful. "4:1" implies that the data is 1/4 the original size. In the context of an Internet data compression unit like ours, that means that you can get 4 or 8 times the throughput. But compression is highly dependent on what is being compressed and the size of what you are compressing. So there is a dependence on the traffic.

First and foremost, a general rule is that you can't compress things that are already compressed. Think of it like a trash compactor. Once you run the compactor, the trash is compressed to maybe half of what it was, but running it again isn't going to do anything more. It's already as small as its getting. The same goes for data. So once it's compressed, that's it. As mentioned before, graphics are already compressed. So are videos. And if you're downloading a .tgz file, it's already compressed, so you won't see any "gain" for such files.

Secondly, there is the matter of size. A product like ours compresses 'on the fly", which means that it doesn't try to figure out what the data is, it just grabs packets and tries to compress them. Internet packets are generally not larger than 1500 bytes, so you're not going to get as good compression on a small chunk as you might get on an entire file if you were compressing it on your hard drive.

Compression Ratios and the Internet

Compression ratios obtained in the context of a network appliance require some understanding of internet traffic and the way the compression is utilized. Since the packet needs to continue to be "IP functional", that is, it needs to be a routable packet, the IP header cannot be compressed nor can the Ethernet header. Also its worthy of note that most 60 byte packets only have 20 bytes of IP "data", so compressing them won't do much good, as the "filler" to fulfil the minimum Ethernet packet size isn't going to be transported over your internet connection anyway.

You also must consider that compression is not a freebie in terms of overhead. It takes time to compress and decompress a packet, and it also takes time to "try" to compress a packet. The theory is that it's worth the CPU cycles if compression can be obtained, because the benefit in terms of gain in throughput outweighs the cost of the reduction in overall performance of the compression device, particularly at today's CPU speeds. But when you try to compress a packet and fail, either because the packet is uncompressible or because there is no "gain" in the compression (i.e. the resulting packet would not be smaller than the original), then you have lost CPU cycles with no benefit.

Almost half of the packets that pass through an Internet router or bandwidth manager are small packets. Most of them are TCP ACKs or handshake packets for starting and terminating connections. None of these packets can be compressed. They are more than 50% overhead, and the data content is minimal.

Compression Tuning

An important strategy when using the compression feature on the Opteq is to minimize wasted overhead. We allow you to set general strategies for packet size and realized compression ratios, and also allows you to override the settings for particular "connections" or for specific protocols. If system capacity is a concern, you can elect to only try to compress large packets, or only HTTP traffic, for example. You can select a minimum compression ratio so that you don't waste cpu cycles compressing data that is only marginally compressible.

Realistic Compression Ratios Obtainable

For general http browsing, our empirical tests show about a 70% increase in real throughput with compression enabled. This implies that on a T1 line you'd get about 2.6Mb/s. Web pages are a highly differentiated mix of text and graphics, and we've already mentioned that graphics generally can't be compressed further. With text files (say if you were editing files via telnet or ssh, or downloading pure text like with Google groups), ratios of about 250% can be obtained. Database files, say backing up a mySQL database, the ratios are closer to 400%. For binaries (such as .exe files), the gains are in the 40-50% range. Email is largely text, although attachments are generally already compressed. So what you'll get is far too dependent on the content to try to predict.

The basic rule is, the more recurring data elements in the application or file, the higher the compression ratio-and hence, the bandwidth savings.

Real Life Use of Compression

First, each location must have 1 or more "Bandwidth Manager" systems. Then any of the locations can selectively choose to compress data sent to any of the other locations. In a perfect world everyone would use the same device, but realistically such scenarios would have to be planned.

Scenario 1: ISPs offering Dedicated Services over WAN

Perhaps the most valuable use of compression would be for ISP's, as they could specifically benefit by offering an "enhanced" T1 service to any of their customers by simply offering a bandwidth manager solution to the client. Since each bandwidth manager can handle compression to any number of locations, they would only need a single unit at their location to handle any number of subscribers. So their variable cost would be nothing after the initial purchase. Clients who opted for the service would have to buy a bandwidth manager box and then pay a monthly fee for the service, but the cost of a 50-70% likely increase in throughput would be a small fraction of the cost of getting a second T1 and load balancing the lines. The solution is particularly useful on a shared connection like DSL, because the reduction in congestion on the shared access line would give everyone better throughput as well as more reliable service.

Scenario 2: Increased Bandwidth between Locations of a Multi-POP ISP or Corporation

It's likely that multiple POP's or locations share more that a bit of information, whether it is a centralized database or backups. Enabling compression between them will increase the speed of the transactions, as well as reduce the amount of bandwidth that internal activities use, making more available to sell or use.

Scenario 3: Inter-Corporation Compression

The Allied and Acme corporations may do a lot of business, and may move large documents back and forth. Cutting the download time of those documents increases productivity. As long as both companies have an Opteq bandwidth management device, they can compress traffic between them. If both companies have a T1 isn't 3Mb/s better than 1.5Mb/s?

Scenario 4: Increased Throughput for Home Access by Company Personnel

Suppose your company is on a cable modem network, and you are on the same or a different one at your home? Cable companies have lower bandwidth upstream, so the traffic coming out of your office will be 256k or 1Mb/s or something in between, depending on what system you are on. By putting a bandwidth manager at the office and the homes of any executives that regularly access the office network, you can significantly increase access speeds with compression.

About Opteq iQ

Opteq International is acknowledged as an emerging leader in the rapidly evolving world of true network management. By designing, developing and manufacturing our own proprietary applications on unique hardware server platforms, we retain total control over the quality, throughput and reliability of our products. The Opteq flagship product, Opteq iQ, is the unique customised vehicle through which we measurably impact the businesses of our customers by fully capitalising on the investment made in their network. The driving idea behind Opteq IQ is to get the most out of our customer's network …to thrive in today's fast-paced, data-intensive economy where our customers depend on the corporate network to be truly responsive and secure. Opteq iQ is designed as a single solution to facilitate the three converging key deliverables of the network management word – Performance, Security and Management. To facilitate this Opteq iQ has a comprehensive suite of fully integrated application modules such as Bandwidth iQ, Web iQ, Compress iQ, Intrusion iQ, Mail iQ, Firewall iQ and Real Time iQ all residing on one common solution platform Singular iQ. All of the Opteq iQ application modules reside on, and exploit the rich functionality provided by Singular iQ. While each module delivers industry leading functionality and performance in its own right these can be combined to deliver the most comprehensive Network Management solution available today. This modular architecture fundamentally differentiates the Opteq iQ product set from its competitors and underpins Opteq's approach and philosophy towards managing network infrastructure - end-to-end management, customised for each unique installation, from one transparent solution.
