boldonjames

White paper

Date: March 2006

Bandwidth management for

Defence Messaging

Software and Security Solutions

boldonjames

EXECUTIVE OVERVIEW

As a networks data carrying capacity or Bandwidth rises so does the amount of data it is required to carry. It is generally accepted that the richness and volume of data available will increase until it outstrips the ability of a given network's capacity to deliver in a timely manner.

Never truer than in the tactical battlespace where "Information Superiority" is all about delivering rich content to the "Warfighter" in a timely manner.

So how does an organisation ensure "Information Superiority" when the richness and volume of data is growing faster than capacity?

Existing approaches to bandwidth management, including using operational procedures such as banning large data movements during peak times and network layer compression techniques go some way to achieving the required goal but don't fully satisfy timely "Information Superiority" especially across low capacity network connections.

This paper presents a new approach to bandwidth management called "Application Level Ultra Compression" that enables the timely delivery of "Information Superiority" with compression ratios 20 to 50 times higher than existing methods.

The paper assumes that an organisation has already applied some degree of 'good practice' to the distribution and management of its information but that Bandwidth cost and throughput remain an issue to the organisation.

The paper focuses on the "Deployed Tactical" space within the Defence Sector, but the issues and implications apply equally to a commercial environment.

BANDWIDTH - AN ETERNAL PROBLEM?

Throughout the history of computing, network bandwidth has caused difficulties. As technological advances have repeatedly doubled throughput; the applications and related systems requirements utilising that bandwidth have expanded to make use of (and quickly exceed) every increase. Stepwise network bandwidth improvements are usually expensive and whilst it is certainly true that there have been several orders of magnitude of reduction in the time taken to deliver any individual piece of information, the ever increasing volume and diversity of priority and content mean that optimising the use of limited network bandwidth remains a significant problem.

Some organisational information flows can be designed to take account of network bandwidth limitations by strategically placing replicas of often accessed information at key points in the network and having these periodically maintained. However, some classes of information are unpredictable both in when they occur and their content. The growth in importance of e-mail (both formal and informal) to the business of organisations coupled with its relatively 'uncontrolled' nature means that the impact of Messaging traffic on network bandwidth utilisation is an increasing problem.

In the Defence Sector strategies such as "Information Superiority" demand an increase in the amount of information (maps, photographs, reports, blueprints etc) delivered to the war fighter. In order to achieve Information Superiority we must establish more effective means of delivering information to the right place at the right time. Slow or busy communications links conspire against this strategy.

The available network bandwidth may have to carry a wide range of traffic including specific application information, file transfer and voice. Each class of traffic may have its own characteristics and issues.

Because messaging is so easy to use and the transmission of very large amounts of information may occur as a result of a simple 'send', it is very difficult to predict the load generated from a user driven tool like email. The rest of this paper focuses on messaging information flows.

SOME APPROACHES TO THE PROBLEMS

Whilst assuming that technology will continue to deliver improved bandwidth, the problems outlined above can be viewed as two separate questions:

- How to reduce the actual number of bytes transmitted whilst preserving information integrity and security?
- How to prioritise the information that is sent especially when a single link is used for a broad range of information flows?

Various techniques have evolved to attack the problem from each of these standpoints.

Operational Procedures

Organisations may lay down good practice operational guidelines that can be very effective in reducing messaging traffic, such as sending links to information that is already available locally rather than including actual attachments. Whilst this has clear merit it can be difficult to enforce, and does not cater for circumstances where an updated version of some document is being passed for feedback prior to general publication.

Blanket limitations

Some systems may inhibit messages or attachments over a certain size from crossing low bandwidth links. The Boldon James SAFEmail.mil defence messaging client, for example, offers a “Can-Send?” check that can be configured to prohibit messages greater than a certain size from being sent to certain individuals/roles or organisations. Whilst this clearly reduces traffic, without careful management it can lead to situations where key personnel do not have access to critical information, or specific administrator action is required to override system limits.

Network prioritisation

Network prioritisation tools are mandatory elements of a modern tactical network, but they are limited in their effectiveness because they can only attempt to prioritise packets that have already been sent. They do not provide the ability to control prioritisation at its source, namely at the point the packets are sent. The key is to prioritise at the higher network protocol level, namely the application level, as opposed to the lower levels such as the IP level.

Within the network protocols it is possible to define ‘data’ priorities that can be used by the network to ensure that higher priority traffic takes precedence. Attempts to set these at an IP level on an application basis can be thwarted by overrides applied by routers, or may operate only on an IP address basis making it impossible to apply different priorities to different data destined for the same address.

Network compression

Hardware products are available that will compress the data streams between end points. Some will even use caches of previous transmissions allowing (small) indexes to be transmitted in place of the actual data. These compression mechanisms are known as ‘ultra compression’ and will typically achieve much higher compression levels than simple byte level compression using similar algorithms to WINZIP.

Such hardware products can deliver worthwhile benefits, but are dependent upon additional hardware being accredited and installed with the potential integration and maintenance issues that may bring.

These products cannot compress encrypted data or will be of limited value if higher layer network protocols have already compressed the data. In addition, because such solutions work at a low network protocol level they have to cache all the network traffic to be able to deliver benefits. This results in needing to cache more data, requiring more disk space, and the time taken to compare new data against cached data will put an upper limit on the data throughput.

Application level Ultra Compression

As with network prioritisation a better solution is to implement ‘ultra compression’ at a higher protocol level and in software, as opposed to at a low level in hardware. A high level software solution can enable ultra compression to act on data before it gets encrypted or compressed resulting in better compression. In addition, application level ultra compression can be more selective about what data to cache so as to avoid having to compare all data against all previous data sent by any application. This will also result in better compression ratios relative to the size of disk cache used.

Finally, because application level ultra compression can be performed before the data is transmitted the compression does not need to be performed in real time on the transmitted data. The effect of this is that using ultra compression does not place an upper limit on the maximum data throughput. The net effect of all this is that hardware ultra compression will typically yield traffic reductions of between 50% and 80% (2X to 5X compression ratios), while software ultra compression can provide traffic reductions of between 90% and 99% (10X to 100X compression ratios) and in some circumstances better than 99% reduction.

Email acceleration

E-mail content has some specific characteristics that need to be addressed in any solution that accelerates the delivery of email, specifically:

- Every message is different to every other message – even if it is only the time stamp information.
- Messages are unpredictable – they arrive randomly, and may vary in size from small (subject only) to enormous (50Mb attachment).
- Messages may have levels of service requirements that mean that a large message may have to be sent 'immediately' at the expense of smaller messages that have been waiting for a gap in the bandwidth.
- The information flow is bi-directional with MTA (Message Transfer Agent) to MTA handshakes and message specific reports (Delivery Reports and Receipt Notifications) included in the information flow.

With the use of technology that allows the detection of sequences of information that have been transmitted previously (regardless of the actual position of the sequence), then the following qualities of messages become significant:

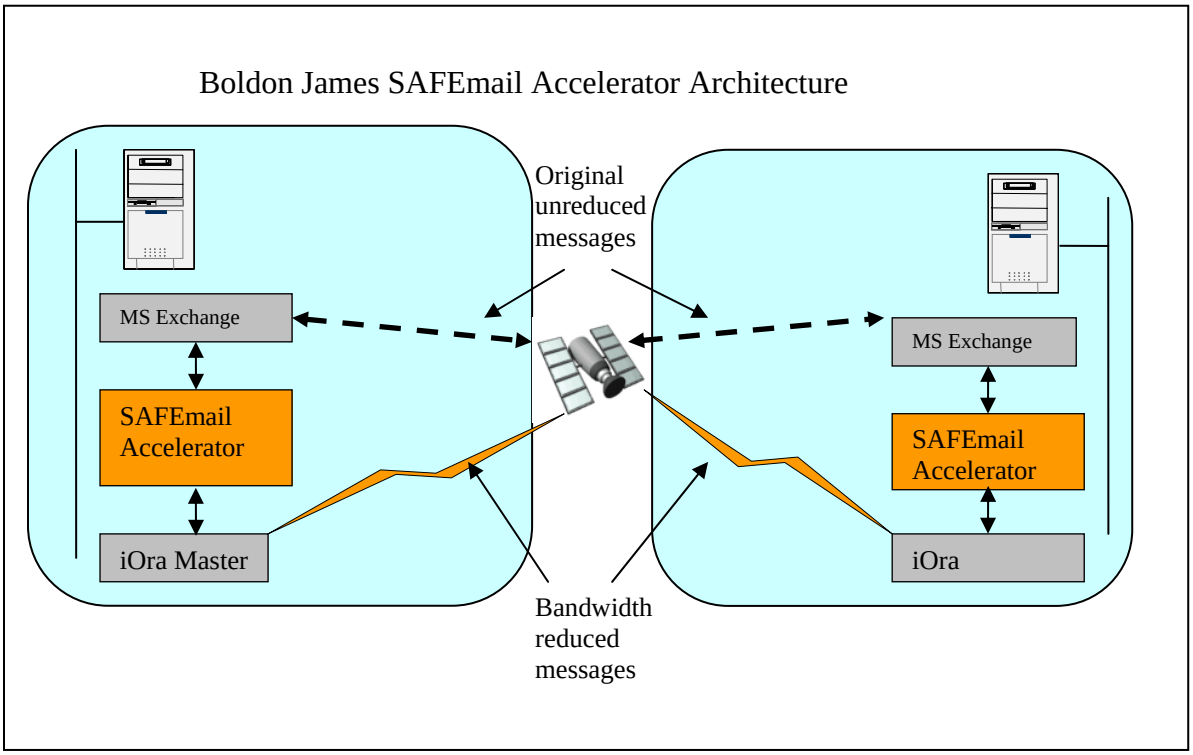
- Large messages usually contain (large) attachments which have been sent before with little (even zero) change between attachment content.
- Even the smallest message has a level of overhead when in transit that makes it worth compressing.
- Even the smallest message may contain sequences that have been seen before (e.g. recipient lists)
- Messages have priority information built in which can be used to directly prioritise them.

The next section explores one method of achieving ultra compression for e-mail by combining Boldon James messaging expertise with iOra's best of breed WAN acceleration technology to produce Boldon James SAFEmail Accelerator.

SAFEMAIL ACCELERATOR - WHAT IS IT AND HOW DOES IT WORK?

Boldon James SAFEmail Accelerator provides bandwidth reduction and bandwidth management for e-mail messages between two Exchange Servers. All message traffic is handled irrespective of the message content including Basic email, IPM.NOTE, P772, PCT, SMIME, DMS, Outlook custom forms and Calendaring messages, etc.).

The diagram below shows an overview of the components and the flows with and without acceleration.



In the un-accelerated system, the two Exchange servers will send and receive message traffic via the Wan link and apply standard priority management to the traffic (highest priority first). The complete message is always transmitted and hardware based network level compression may be installed. Messages are directed across the link using the defined Exchange address routing mechanisms.

When SAFEmail Accelerator is installed at both ends, iOra publications are created for each single directional link at their sending end. The following summarises how messages are transmitted:

- All messages that would have been sent to the other Exchange system will be processed by SAFEmail Accelerator.
- Each message is transferred by SAFEmail Accelerator into the local iOra Publication store associated with the target Exchange server (see below for further details).
- iOra is invoked to publish the information to the other end.
- iOra processes the changes to its store (the new messages).

  - iOra detects sequences that have been transmitted previously (either as part of a previous message, or even as part of another application's data that is published through the same Publication store).
  - Sequences previously seen are replaced by an 'index'.
  - The resulting message(s) (new sequences + indexes) are then compressed the resulting 'Publication amendment' is transmitted to the other end.

- The receiving end of iOra applies the 'Publication amendment' to its local publication copy
- SAFEmail Accelerator is invoked which submits the new received messages to its associated Exchange server which proceeds as if the message had arrived direct from the sending Exchange system.

Messages in the reverse direction (including Delivery Reports etc.) follow a mirror image operation.

The iOra replication activity uses a standard TCP/IP link typically using HTTP or HTTPS.

How effective is it?

Clearly each system will have its own message profile in terms of the average size, typical number of recipients and the type, frequency and volatility of attachments, and therefore actual performance will vary from system to system.

The underlying iOra technology often delivers compression effects in excess of 90% (see http://www.iora.com/technology/epsilon.asp for further details) for non-messaging information flows. Whilst message traffic has a significantly different profile to other applications (e.g. SharePoint) very significant savings can be still be achieved.

The results below were obtained using an arbitrary set of actual user messages ('Size' indicates total uncompressed size or size range, 'Saving' is a percentage reduction in transmitted message size):

Content	Size (Kb)	Saving
New Messages with Word or EXCEL attachments	130➔180	79%
Resend of messages with some modifications to the attachments and message body	136➔183	88➔96%
Arbitrary set of messages including Forward, Reply, DR and RN	5011	84%

A previous study of Naval e-mail traffic indicated that attachments constitute around 80% of message traffic, and that the application of iOra technology would yield an 80 to 90% reduction for these attachments. Benefits generally increased with time in line with the likelihood that a version of an attachment had been sent before. These figures are in line with the results shown in the table above.

It should be noted that maximum benefit is achieved where a message contains either or both of:

- Information that contains a high percentage of sequences seen before (e.g. resend of an attachment with a small amount of modification).
- Information that is inherently compressible (e.g. a WORD document).

Similarly there is minimum benefit where message content matches neither of these criteria. For example:

- An encrypted message which by its nature contains random sequences (even if the underlying content between two messages is exactly the same). Note that signed only messages do NOT present this characteristic, and are essentially as compressible as unsigned messages.

How is message Priority handled

A message will have a Message Priority associated with it (High, Medium, Low), but other factors such as message size, message content and Message Precedence (for Military messages) may be used to determine an overall (derived) message priority.

SAFEmail Accelerator maps each message into one of 6 folders based on its derived priority with the mapping being configurable by the administrator. Folder mapping is defined in descending 'priority' order set along the following example lines:

Folder	Criteria
Level 1	Override Precedence
Level 2	Flash Precedence
Level 3	Immediate Precedence
Level 4	Priority Precedence with size < 100Kb
Level 5	All other Priority Precedence, & Routine Precedence
Level 6	All other messages

Messages are deposited in the appropriate folders, and then are selected in descending 'priority' order to be included in the next publication update. As it is included in an update it is removed from the Level folder.

Publication updates may be limited in size (for example to fit within a fixed time slot), and it is even within the architecture to publish a message in fragments, where message size is too large for a single slot. Any such fragmentation will not affect message integrity as the whole message will ultimately be 'delivered' as a unit to the remote Exchange server.

Service level requirements are critical, and it is essential that if an urgent message arrives for transmission it does not get stuck behind a (large) lower priority message that preceded it. This is handled by the SAFEmail acceleration architecture by putting an upper limit on the size of the 'pipeline' of submitted amendments for a target Exchange server. This then allows any newly arrived urgent message to be published and delivered in a guaranteed timeframe subject to the availability of network bandwidth and the size of urgent messages.

NETWORK BANDWIDTH – A SOLUTION?

It is likely that network bandwidth will continue to be problematic as the ever increasing demands for information keep pace with technology improvements, and that each of the approaches outlined in section 3 may have a part to play within an organisation's network.
It should be noted that the benefits listed below offered by SAFEmail Accelerator technology can be achieved independently of, or in conjunction with, other bandwidth management techniques.

- **Compression:** iOra's Epsilon compression engine eliminates the overhead of re-sending the same blocks of data twice. By configuring iOra to support application acceleration between a central and remote location the total bandwidth associated with this content is typically reduced by over 80%.
- **Prioritisation:** Individual WAN connections can be configured to use a maximum bandwidth per application. Within SAFEmail Accelerator, prioritization of email messages can be controlled based on selected message attributes.
- **Compliancy:** Boldon James SAFEmail supports all NATO and CCEB specifications for the format of military messages including ACP120, ACP123, NATO STANAG 4406, P772, P22 and S/MIME signatures. SAFEmail Accelerator also works in Medium Grade Messaging Domains with conventional COTS message formats.
- **Compatibility:** Boldon James SAFEmail solution is fully compatible with Microsoft Exchange and all its client options. iOra Accelerator is compatibly with any IP network infrastructure.
- **Management:** Centrally managed configuration of bandwidth utilization for specific WAN links as well as specific users, messages or document types.
- **Security:** Boldon James SAFEmail is a proven solution for handling military messages at any security level.
- **Scalability:** Supports hundreds of thousands of users and terabytes of content. The underlying Boldon James and iOra architectures have no inherent upper limits on number of users or size of content.
- **Synergy:** Can be deployed along with other enterprise applications using iOra thereby maximising efficient use of bandwidth.
- **No new hardware:** Can be run on existing Windows servers, thus avoiding the expense of deploying new hardware. By reducing bandwidth avoids the need to upgrade WANs or add new networking equipment.

About Boldon James

We are a successful organisation focused on providing messaging and connectivity software solutions tailored to key vertical markets including, Defence, Aviation and Government. In the UK our high grade consultants extend this offering by assisting organisations to manage IT and operational risks through a portfolio of services.

The utilization of commercial off-the-shelf (COTS) software at the core of our messaging and directory solutions enables us to provide products which meet the requirements of high grade formal messaging environments such as Defence, Intelligence and Aviation as well as other environments in Government and Commerce.

The Consultancy Services division is focused on the creation of Security Policy, BS7799 implementation (through to certification if required), Risk Assessment, Training and Awareness campaigns, Business Continuity and Penetration Testing. These services are tailored for Central and Local Government, the NHS, Utilities, Police Services and large Corporations and led by consultants who are CESG Listed Advisor Scheme members (CLAS) and CHECK accredited team leaders.

Our robust data communications software provides core functionality in many OSI networks around the world as well as including support for the majority of users of ICL VME (Fujitsu Services) Mainframe Computers both in the UK and around the world.

Boldon James is proud of its reputation which is built on the competence of our staff, many of whom can be categorised as world class performers in their field. Our staff hold an impressive array of relevant professional qualifications and accreditations. We believe that customer service is the key to success.

For further information please visit: http://www.boldonjames.com