2. The need for bandwidth optimization

Part of INASP's work enables access to online academic and scientific publications. For these programmes to be effective, the institutions targeted need to have a usable Internet link.

Several universities in this study have an Internet connection of between 512 Kbps and 1 Mbps (as at May 2003).[1] This is about as much as a DSL connection (512 Kbps to 1.544 Mbps), which is typically used to connect a single household in the West to the Internet. Bristol University, by contrast, has a 2.5 Gbps link; which is 5120 times as much as the University of Dar es Salaam has.

Students and researchers in the West tend to take free, fast access to the Internet for granted. While it is not necessary to have very fast access to the Internet for it to be usable, there is a limit below which it becomes frustrating. Usability studies show that an average Web page should load within 10 seconds; if the text starts loading immediately, followed by the graphics, load times of up to 39 seconds can be acceptable.

Unfortunately, proxy servers (described in this document and in the glossary), cause a delay and then the page loads all at once. As will be seen in this report, it is essential to implement a proxy server. Therefore, a typical page-load time of around 10 seconds could be a target for IT departments.

Usability is important for researchers because of the nature of Web searches. A user might have to load many pages and scan through them before finding the right document. If each 'false lead' takes a long time to load, Web searches become a frustrating experience.

2.1 Participants in the study

The Universities of Addis Ababa (Ethiopia), Bristol (United Kingdom), Dar es Salaam (Tanzania), Makerere (Uganda) and Moratuwa (Sri Lanka) were asked to supply information about their networks and optimization efforts. Bristol was included in order to compare the others with a university that has cheap access to high amounts of bandwidth.

Some institutions could not supply all the required information for reasons such as the information requested not being known.

The MIMCOM network (a network for Malaria researchers across several African countries) and the Malawi College of Medicine were also included because they provided interesting examples or were well described and documented.

Information about the University of Zululand is also included because they did interesting work on a charging system, but Zululand was not part of the original case study.

The table below compares the bandwidth situation of the universities and other networks included in the case studies.

[1] These connection speeds are explained in the glossary entry under the entry 'data rate'.

	Users	Computers	Bandwidth	GB/month	Connection	ISP*
Addis Ababa**	4000	Unknown	512 Kbps	106	Leased line	VSAT
Bristol***	22000	16000	2.5 Gbps	4500	Leased line	Cable
Malawi College of Medicine	Unknown	250	128 Kbps	Unknown	Wireless	VSAT
Dar es Salaam	11000	2000	512 Kbps	Unknown	Leased line + own VSAT	VSAT
Makerere	25000	Unknown	1.5 Mbps	144	Leased line	VSAT
MIMCOM	1400	1200	800 Kbps	90	Own VSAT	N/A
Moratuwa	Unknown	Unknown	2 Mbps	Unknown	Leased line	Cable & VSAT
Zululand	6000+	750	1152 Kbps	120	Leased line	Cable

* How the ISP makes its connection to the Internet.

** Addis Ababa has over 20,000 students, but only academic staff, postgraduate students and some undergraduate students in IT-related departments have Internet access.

*** Bristol is connected to the Janet network, which is a network of educational institutions in the UK. Janet is a very large and well-funded network, more a part of the core Internet infrastructure than something that is *connected to* the Internet.

2.2 Why bandwidth is expensive

Bandwidth to developing countries is so expensive that most universities cannot afford more than 1.544 Mbps – equivalent to the average Western household with ADSL connection. The reasons for this situation include the following:

- In many cases, Internet access to the country is available only via satellite connections, which are much more expensive than cable. Reasons why many organizations opt for satellite connections are given below. A map showing the marine cable connections to Africa (and the fact that only a few countries are connected in this way) can be viewed at: <http://network.idrc.ca/ev.php?URL_ID=6568&URL_DO=DO_TOPIC&URL_SECTION=201&reload=1053101936>
- Where marine fibre cables do exist, they may not carry enough traffic to achieve the economies of scale that make transatlantic bandwidth to Europe, for example, so affordable. In some of the countries that are connected via a marine fibre cable, the telecommunications infrastructure for connecting it to most of the country does not exist.
- The wired telecommunications networks in many developing countries reach only a small part of the population, and many areas (even parts of cities) are not covered at all. The development of wired networks cannot follow the same course as it did in industrialized countries owing to small populations or low population densities in some areas, poverty, the rise of mobile and satellite communications.
- Some telephone companies that have telephone lines lack the capacity (owing to low demand) to create leased-line connections. Low demand exists mainly because many companies and institutions bypass the national telecommunications grid by using VSAT.
- Leased lines are sometimes analogue instead of digital. On an analogue line, a modem is used for digital transition (such as connection to the Internet), resulting in a maximum speed of 56 Kbps. Digital lines are capable of much higher speeds.
- Bandwidth is also expensive due to the comparative weakness of the currencies of developing countries that have to pay in US dollars or euros or other major currencies for most or all of their upstream international bandwidth.
- While the cost of the telecommunications link between two countries is generally shared, in the case of African countries (and possibly of many other developing countries) the cost of the international link is paid for entirely by the African country. This amounts to reverse subsidization of developed countries.
  (see <http://www.afrispa.org/HalfwayDocs/HalfwayProposition_Draft4.pdf>).

- Considerable congestion exists at ISPs where many users share a small amount of bandwidth; ISPs simply have too many customers for their capacity.
- Inter-country links do not exist between most developing countries. For example, most communications between African countries must be routed at great expense through Europe or the USA.
- Communications and computing equipment is expensive for African organizations as a result of weak currencies, high transport costs, small budgets and high tariffs. In many African countries computer equipment is classified as a luxury item and taxed accordingly, though this counter-productive policy is likely to change in the medium term.

The regulatory situation in some countries is another problem. Active bureaucratic resistance is often experienced when changes to the telecommunications environment are proposed. For example, it may take a very long time before a licence to operate a satellite system is granted or refused. Refusal may be designed to protect the state-owned telecommunications company, or to protect a foreign company that has bought the original national operator. Governments also consider security issues when deciding whether to grant permission for new communication links. This might concern their own security (by limiting their population's access to outside information and contacts) or in response to the terrorist threat that is affecting ever more countries.

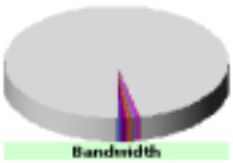
Lack of government investment in telecommunications is widespread because of more pressing priorities such as health and education.

Universities cannot afford a decent link, or in some cases do not see its value.

2.3 Bandwidth problems associated with university campuses

The graphic below shows how Web access through a proxy server at an academic institution can be analysed using inexpensive software (Sawmill in this case). Since the proxy server log keeps a record of every Web site visited, as well as the amount of bandwidth consumed, the top sites by bandwidth can be identified (sites can also be sorted by number of visits, but that is of less interest for bandwidth analysis).

The graphic shows no academic sites in the top ten sites (by bandwidth). The top three sites were not even visited by users: they are connection attempts by worms to URLs on the Internet that were programmed into the worm's code. Apart from that, there is MSN, which is not a particularly useful site, Yahoo's e-mail service, and Gator, a shopping service that installs a 'helper' on unsuspecting users' PCs. This example shows how the bandwidth usage of an academic institution can be monopolized by malicious worms and the commercial interests of large corporations. The academic purpose of the network is not reflected in the bandwidth usage, and it is clear that this network is out of control.



	Page or folder	Page views	Bandwidth	Visitors	Page views bar
1	http://www.opasoft.com/work/zcheduler.php?(parameters)	97,331	86.32M	4	
2	http://www.n3t.com.br/work/zzcheduler.php?(parameters)	97,272	86.25M	4	
3	http://www.gwnnet.com.br/marcos/gayer.php?(parameters)	96,639	48.34M	4	
4	http://shttp.msg.yahoo.com/notify/ (default page)	43,479	18.06M	16	
5	http://140.239.165.252/http_page.html?(parameters)	21,976	17.14M	1	
6	http://rad.msn.com/ADSAdClient31.dll?(parameters)	19,197	15.08M	73	
7	http://h.msn.com/c.gif?(parameters)	15,745	7.21M	70	
8	http://bannerserver.gator.com/bannerserver/bannerserver.dll?(parameters)	15,044	53.03M	12	
9	http://uk.adserver.yahoo.com/a?(parameters)	12,961	13.22M	38	
10	http://mail.opi.yahoo.com/online?(parameters)	9,460	4.46M	63	
	205587 other items	757,491	9.60G	-	
	Total	1,187,585	9.94G		

Other problems particularly associated with university campuses include:

- Students typically have more time, are less supervised, and are under less pressure from work targets than, for example, office workers. Therefore, a university network is one of the most challenging environments to manage.
- People use the Internet in many different ways, some of which are inappropriate or do not make the best use of the available bandwidth. For example, while it may not generally be a problem if a student downloads a music file, plays on-line games or experiments with the latest Microsoft service pack, it becomes a problem when the bandwidth consumed by this activity prevents a researcher from downloading or viewing a scientific article
- Even where high amounts of bandwidth are available, control, monitoring and optimization are necessary because users (and especially students) will always find a way to fill the available amount of bandwidth.
- Hacking: students experiment with their computing knowledge, and connect to exposed computer systems both on the campus and elsewhere in the world.
- Peer-to-peer (P2P) networking (using Kazaa and other programs) is very popular among students.
- Universities may need to police the amount of bandwidth they are getting from a shared system because they might be competing for bandwidth with other customers. For example, if a university gets its bandwidth from an ISP, it is likely that the ISP also sells bandwidth to other users, such as local companies. An organization should have a clear understanding of the nature of the shared system – how much minimum bandwidth they are paying for (the Committed Information Rate, CIR) and how much Burst Excess (BE) they can get. When there is a BE, what is the contention ratio? (These issues are explored in more detail in Section 6.1.)
- 'Long fat pipe network' factors affect TCP/IP performance in networks that have relatively large bandwidth but high latency (delays), as can be found in satellite networks. The high latency is due to the long distance that signals must travel from the VSAT dish to the satellite, and in some satellite networks this requires steps to enable TCP/IP to make full use of the available bandwidth. This issue is discussed in more detail in the glossary and Appendix A.
- Latency due to a connection via satellite also makes it important that functions such as DNS (see Section 4) are provided locally rather than across the Internet link. When the DNS server is on the local LAN, most DNS queries will take less than 10 ms (milliseconds). The satellite link adds at least another 550 ms because of the long distance between the earth and the satellite.
- Staff issues. Without proper management, IT staff may themselves become part of the problem (see Sections 10 and 11).
- Control and monitoring is necessary in order to make informed decisions about how much bandwidth is needed. If the graphs show that bandwidth is used mainly for recreational activities, or is consumed by virus activity and Windows updates, then control is more urgently required than additional bandwidth.
- The after-effects of a new installation. Soon after any organization connects its staff to the Internet and e-mail, it can expect to see its IT-support problems spiral. It takes a long time to get these problems under control, even for a rich private company. A university's IT environment can easily get even more chaotic. Management needs to plan for this transition.