

Bridging the Bandwidth Gap: Open Educational Resources and the Digital Divide

Björn Haßler and Alan McNeil Jackson

Abstract—This paper presents an overview of Digital Divide issues associated with the global disparity in availability and cost of bandwidth. We give examples and discuss the relevance of this to global use of Open Educational Resources. Strategies for mitigating low-bandwidth challenges are discussed.

Index Terms—Low-bandwidth operation, Transcoding, data communications aspects, network management.

1 THE DIGITAL DIVIDE

THE term “Digital Divide” has a range of meanings. In the context of the current paper, we consider the term “Digital Divide” to mean the global disparity regarding available Internet bandwidth. Broadband Internet connections are extremely expensive in the developing world. Many users cannot afford them, and so have great difficulty accessing the Internet effectively and satisfactorily. This is particularly true when accessing the many sites designed for high bandwidth connections. No matter how important its content, if a site takes too long to load it becomes unusable.

As broadband is becoming the norm in the developed world, website and OER designers are forgetting the lessons of optimization learned when the best you could hope was that your users would have a 56 kbps modem dial-up connection. By allowing websites and OER content to grow in size not only do we provide a poor user experience for all users, but we make our sites virtually unusable with a slow Internet connection.

1.1 The Cost of Bandwidth

Physical bandwidth in developing countries is low (cf. Fig. 1), because it is prohibitively expensive. In 2004, low-income countries were paying more than 10 times the price that high- and middle-income countries did for a broadband connection [1]. A more recent survey (2007) shows that some African universities are spending the equivalent of 20 full-time salaries for a 2,000 kbps¹ Internet connection shared between 500 and 600 computers.

The basic reason for the cost of Internet connectivity in developing countries is the lack of an extensive wired infrastructure. There are two aspects to this: One is international connectivity, the other is “last mile delivery.” International connectivity often has to rely on satellite, both for high bandwidth satellite connectivity (e.g., for an Internet Service Provider) or by end users directly. This is expensive: The cost of satellite connections is several 1,000 USD per Mb/s per month. The second issue is the national connectivity (“last mile delivery”). This is relevant where an international connection (say by an Internet

1. Typical domestic broadband connections in the UK in 2009 are described as “8,000 kbps” with the average achieved bandwidth over 3,000 kbps.

- B. Haßler is with the Centre for Commonwealth Education and the Centre for Applied Research in Educational Technologies, University of Cambridge, 1st floor, 16 Mill Lane, Cambridge CB2 1SB, UK. E-mail: bjoern.hassler@sciencemedianetwork.org.
- A.M. Jackson is with Aptivate, The Humanitarian Centre, Fenner’s, Gresham Road, Cambridge CB1 2ES, UK. E-mail: alan@aptivate.org.

Manuscript received 4 Mar. 2009; revised 29 Apr. 2009; accepted 21 Jan. 2010; published online 20 Apr. 2010.

For information on obtaining reprints of this article, please send e-mail to: lt@computer.org, and reference IEEECS Log Number TLTSI-2009-03-0033. Digital Object Identifier no. 10.1109/TLT.2010.8.

Service Provider) is distributed to end users, through a wired infrastructure or other means. The cost of this is also considerable, and typically similar to the cost of international connectivity.

It is difficult to make longer term predictions, but there seems to be broad agreement that, while some reduction in cost is expected over the next few years, the cost will become nowhere near as cheap as the cost of connectivity in the global North. (This does take new undersea fiber and new satellites into account.) One should be suspicious of the claim that “soon we’ll all have infinite bandwidth at no cost,” based on the experience of the broadband explosion in the North, which leverages our copper infrastructure via technologies like ADSL. There is no equivalent existing infrastructure in developing countries that could be leveraged for end users. Another disparity is the fact that the majority of the content is hosted within the North and there are high capacity undersea fiber interconnects connecting Europe and the US.

So, in absolute terms, connectivity in the South is much more expensive than connectivity in the “wired” North, and while there will be improvements, connectivity is nevertheless likely to remain much more expensive for the foreseeable future. However, there is an additional factor: Rather than looking at cost of connectivity in absolute terms, one should look at the cost of connectivity related to some economic measure, such as purchase power parity. When this is done, it only exacerbates the disparity, cf. Fig. 2. We should also bear in mind that while there are a growing number of regions that are connected now, there are many other regions that are not connected and will only gradually move into minimal provision over the next few years.

This view of connectivity is accurate, despite improvements in mobile Internet connectivity. Data transfer for mobile browsing still has to get in and out of the country, so ultimately (in many regions) has to rely on satellite. However, much of mobile browsing in the developing world is based on GPRS, and while it can be comparatively cheap, it is also comparatively slow. We will come back to this later.

1.2 Estimating Available Bandwidth

In this paper, we distinguish “physical bandwidth” and “available bandwidth.” The physical bandwidth is the size of the pipe: For instance, the size of the pipe to a country, to an institution, or to a home: An upper limit on connectivity. The “available bandwidth” is the actual bandwidth available to a typical user. Of course the “available bandwidth” is always lower than “physical bandwidth”: There are technical reasons, such as “contention,”² but also usage reasons: Typically, the “physical bandwidth” is shared between several users, each user having a fraction of the “physical bandwidth” as their available bandwidth.

Another important factor in utilizing bandwidth is the management of the local network. What is important to realize is that for a user on a poorly managed network, the available bandwidth may be a small fraction of the physical bandwidth: The remainder of the bandwidth might be wasted on network overheads, such as viruses or illegal file sharing. The African Tertiary Institutions Connectivity Survey [4] found that almost 2/3 of universities practice little or no

2. Typically, an Internet Service Provider shares bandwidth between several customers. The number of customers determines the “contention ratio,” so if bandwidth is shared between 10 customers the connection is said to have a contention ratio of 10. There are two main places where this contention can take place that are relevant to our discussion. First, between the customer and their in-country ISP. The second place is the contention over the international bandwidth. Even if an institution has a high bandwidth connection and low contention to their ISP, they may still have a high contention for the international bandwidth. Considering that most web resources are hosted in the US and Europe and that many developing countries do not have fully functioning Internet Exchange Points, which consequently means that even in-country traffic is often routed via the US and Europe, it is the contention over international bandwidth that is usually the most significant.

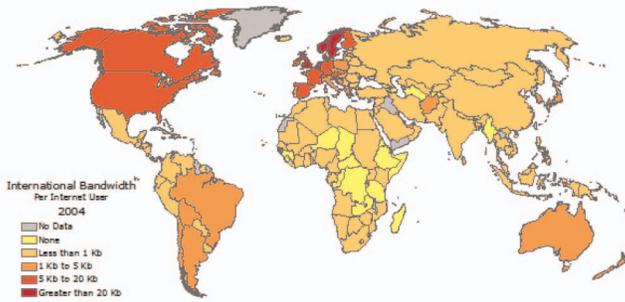


Fig. 1. International bandwidth per Internet user (2004).

management³ of their connections: Universities have a hard time retaining skilled staff, and there has been a lack of awareness among management and funders as to the need and means to build up good network administration, policy, and training.

What is the typical available bandwidth in the developing world? In many countries in the developing world, international connections are routed through high latency satellite links, and it is hard to make a good estimate of available bandwidth: The available bandwidth to the end user depends on how much bandwidth is brought in, but also on how well this bandwidth is managed, and to how many users it is distributed. We chose to base our estimates on African Universities because there are surveys of their available bandwidth. According to the aforementioned ATICS survey [4], the average bandwidth available to an African university is 1,254 kbps. On average, there are about 600 connected computers in each university. Assuming that 1 in 20 of these computers is using the Internet at any one time (based on ADSL contention ratios), this would give each user about 40 kbps on average. However, most universities do not have a Committed Information Rate so their true bandwidth is roughly half of the stated bandwidth they have bought. In addition, many developing world institutions are not implementing effective bandwidth management practices which further reduces available bandwidth [5].

Taking this into account, we arrive at a figure of 20 kbps as a sensible upper limit on bandwidth available to the user, which is several orders of magnitude slower than the available bandwidth at a European and American university.

Briefly returning to mobile Internet connections, we also note that this connectivity is similar to mobile connectivity under GPRS, which is the dominant mode of connectivity in the developing world: Although better connectivity (such as EDGE or 3G) is becoming available in urban areas, basic phones that are used by the majority of users only support GPRS, if they are Internet capable at all.

1.3 Estimating the Size of a Usable Web Resource

How does available bandwidth translate into “usable” web resource sizes? What is the webpage size for low-bandwidth usable webpages? The Web Guidelines for Low Bandwidth [2] suggest a page size of between 25 and 75 kB. This is based on usability research, which found that users will abandon webpages that take longer than 10 seconds to load. Although if useful data start to appear within 2 seconds, they are prepared to wait up to 30 seconds for the page to finish loading [3]. Hence, the guidelines [2] suggest that at 20 kbps, the page can be no larger than 25 kB (to load within 10 s), or up to 75 kB if the page loads progressively.

We need to recognize that these limits are for the total page including images, external JavaScript files and external CSS files.

3. This may mean no firewall protecting the local network from intrusion or viruses, no proxy cache to avoid repeated downloads of the same content, no antivirus software and no Acceptable Use Policy.

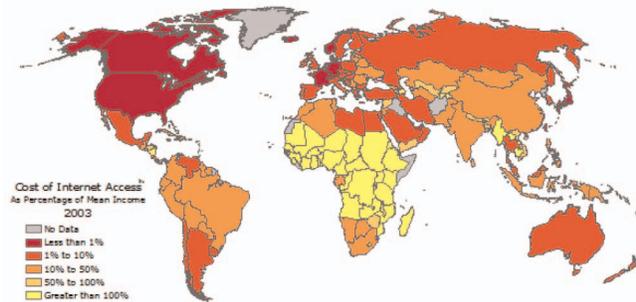


Fig. 2. Cost of Internet access as a percentage of mean income (2003).

Although sharing such files between pages allows them to be cached by the user’s browser and reduces the loading time of subsequent pages, they should still be counted in the page size because they will slow down each user’s first page view of your site, affecting their first impressions. If a user abandons a page because it takes too long to load they are not likely to revisit it. Although we have no evidence, it seems reasonable to assume that a user who is absolutely certain that they have found a useful resource will tolerate longer download times to wait for large documents, PDF files, or other media such as video or audio recordings. The tolerable time will depend on the perceived value, which in turn will relate to the ratio of the size of the information content to the size of its presentation. But this is only true when the user is certain they have found what they are looking for and not true for speculative downloads. The webpages that allow a user to find and access these resources and to preview their content must still conform to the above timing guidelines.

Overall, we are not advocating the use of large or unoptimized media objects, quite the contrary. However, we recognize that some information can only compress so far. Take, for instance, a large table of epidemiological data. To a health researcher, the value of this data may mean they would be willing to wait for potentially hours for it to download. However, if this table of textual data is presented in a PDF as an image for the sake of pretty presentation (as is sometimes the case) multiplying the size of the file by a factor of 10 or more, then this is not acceptable.

1.4 Web-Based Access to Open Educational Resources

The recommendation for “low-bandwidth usable” page sizes is to be contrasted with the size of the typical webpage [3], [6], that is now estimated to be about 300 kB (in 2008), having increased from about 100 kB in 2003. A page size of 300 kB is 4 to 12 times the recommended page size of 25-75 kB. This disparity of usability is widening, because the rate of average page size growth is much greater than the growth rate of available bandwidth in developing countries. Given these findings about average webpage size, it is no surprise that the average OER webpage would be of a similar size. As a use case, we consider navigating through MIT Open Courseware to a particular resource, cf. Table 1.

The timings are calculated from uncached page sizes, assuming an available bandwidth of 20 kbps. The page load times are given in minutes. For a typical user in the North, corresponding times would be seconds. We have used MIT OCW, because it is a leading exponent of OER. However, the MIT OCW website itself is typical, and many other OER related pages are similar. For instance, the OCWC consortium⁴ front page comes in at 243 kB, while the Open University OpenLearn⁵ front page comes in at 743 kB. From this use case, and comparison with other OER sites, it is quite clear that in terms of web accessibility, the vast majority of OER sites does not meet low-bandwidth accessibility requirements.

4. <http://www.ocwconsortium.org>.

5. <http://openlearn.open.ac.uk>.

TABLE 1
Navigating through MIT OCW [8]

Page	Size	Time
MIT OCW front page	240kB	1.6 min
Courses page	856kB	5.7 min
Search	120kB	0.8 min
8.01 Physics I, 1999	149kB	1 min
Syllabus	121kB	0.8 min
Lecture Notes	126kB	0.8 min
Download	118kB	0.8 min
<i>Total =</i>	<i>1.7MB</i>	<i>11 min</i>
Additional download of OCW zip file:		
8-01Fall-1999.zip	11MB	75 min

1.5 An OER Users View of Bandwidth

We have established that there are bandwidth related access issues affecting a large number of Internet users particularly in developing countries. We do not suggest that issues around bandwidth are the sole cause for poor use and slow uptake of OERs in developing countries (cf. for example Fig. 3), but compared to other barriers, we emphasize that bandwidth issues are an important contributing factor.

We may still ask to what extent the OER community subjectively experiences this issue. As part of a recent discussion on access to open educational resources conducted within the UNESCO OER Community [9], the issues around lack of bandwidth were echoed quite widely, and we include a number of quotes from that discussion to give a more individual and subjective perspective on the issues:

One of the barriers of using OERs in central and southern African universities is the issue of BANDWIDTH. They must pay very expensive for that and they have no money ... (Rwanda)

If OER projects want to be helpful for developing countries...there is a crucial need to develop resources accessible in low bandwidth...and by low..I mean almost dial-up! ... I hope OER developers keep this in mind.... (Mexico)

I would like to enthusiastically embrace the idea of working harder on bandwidth management. ... more effort should go into making resources usable in low-bandwidth environments (which is after all the target audience of this group). (Brazil)

When ever I send the websites of free available e-resources to our students, teachers and researchers they complain that they could not download the materials because of slow Internet or some times non accessibility. (Pakistan)

Last time we participated in the identification of OER materials but what we faced was the trouble of having access to Internet connectivity. ... became extremely expensive and at the moment some schools cannot even afford to continue to have this connectivity. (Teacher at a Zambian secondary school)

... not to forget the issue of the bandwidth, which is much exaggerated by the cost. It is quite often to loose connection in a University because of the high bill to be paid. (Sudan)

With reference to bandwidth, this is an ongoing issue for teachers in my project. They cannot download videos, or watch them, because the CTC where they go to use the Internet has measured service via satellite, and once the bytes are used in a month, service is shut down until the bill is paid. (Guatemala)

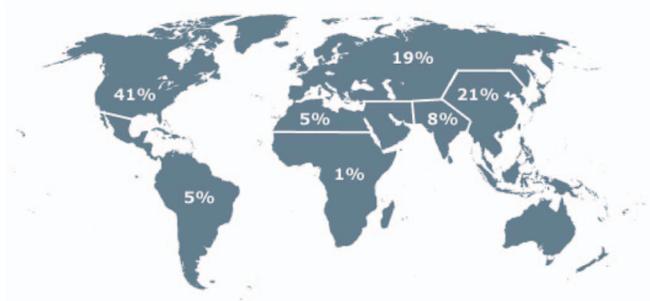


Fig. 3. World-wide visitors to MIT OCW. Source: MIT website [10].

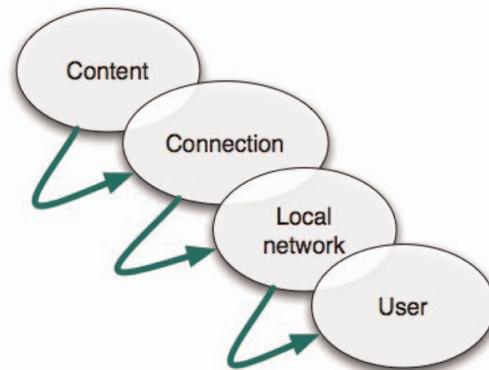


Fig. 4. The content delivery chain.

Arguably such connectivity issues were discussed more widely than other access related issues (cf. [9] for the full UNESCO discussion report). That is not to say that other access issues are not important, but we might say that some of the other issues are currently “Northern issues,” in the sense that they will only become relevant to the global South once some of the issues pertaining specifically to the South have been addressed. Access issues are generally important, but the South is suffering first and foremost from lack of bandwidth.

2 THE CONTENT DELIVERY CHAIN

We use a Content Delivery Chain model as a framework for thinking about the different aspects of bandwidth and connectivity [7]. This helps to clarify the factors contributing to available bandwidth.

The content delivery chain (cf. Fig. 4) represents the pathway of content: Starting with an item of content, it is delivered through an Internet connection to a local network and finally to the user. We refer to this model as a “chain” because successful content delivery depends on every link of the chain:

- Content providers have their role to play: They must ensure that their content is usable over existing connections. As we have seen, if the content provided does not meet low-bandwidth accessibility requirements, it is unlikely to reach recipients in the developing world.
- Of course some connectivity (physical bandwidth) must be present.
- Likewise, the local network must be well managed to fully utilize the bandwidth.
- Users are also critical. In many ways, user behavior is the largest factor determining the effectiveness of any Internet connection, and appropriate policies need to be in place.

The content delivery chain is limited by the weakest link, in the sense that total available bandwidth is limited by the weakest link in the chain.

We note that, rather than considering the whole content delivery chain, much of the discussion around bandwidth only focuses on the connection (that is the “physical bandwidth” of the connection). However, it is all elements of the chain that determine whether a piece of content is available to the user: All elements of the chain determine whether the available bandwidth is “effective,” i.e., whether the available bandwidth can be used effectively to achieve a certain task, such as the successful use of content.

2.1 Effective Use of Bandwidth

We may ask whether an available bandwidth of 20 kbps means that you can make effective use of that bandwidth? In practice, even just 20 kbps is a viable connection, and a lot is possible with 20 kbps. However, it does require simultaneously good content in appropriate formats, good local network management and design, good bandwidth management policy (in particular, an Acceptable Use Policy), and user education to understand their responsibilities to use the scarce and expensive resource effectively.

As we see through the Content Delivery Chain, physical bandwidth (“pipe size”) is only part of the issue: There are important considerations in terms of formatting the content, network management, user policies, caching, and smart use of connections that would make a huge difference (even with current pipe sizes). In the present context of OER, this has the advantage that it can be implemented by the OER community (providers and consumers) primarily with local effort, and without having to rely on external factors, such as improvements in physical bandwidth or sustained donor payment for satellite broadband.

Of course, one could argue that some of this is generic, and doesn’t pertain specifically to OER. However, there are issues (such as formatting the content) where it becomes a specific OER issue. Also, one could argue that an effective, global OER community needs to take current barriers into account (regardless of whether they are specifically to do with OER or not), and do as best as we can to address these issues.

Another way of looking at this is to distinguish between content size and content information. For instance, the complete works of Shakespeare can be compressed to 2 MB in text only form,⁶ equivalent to only six average webpages [6].

2.2 Content Delivery Chain: Content

Let us now look at some aspects of the content delivery chain in more detail. We start with the first element of the chain: Content.

Often content can be rendered in different formats, and in terms of enabling low-bandwidth access, it is important to provide alternatives. For text-based materials, a plain text version can be provided as an alternative. PDF carries a significant overhead, and where PDF is used, it should be optimized, cf. [2, section “PDF optimization”].

Multimedia objects (such as images, flash, and embedded sounds) should be used carefully, with consideration for the circumstance of the users. Where such objects are used, they should be linked to, rather than embedded in a page. Users should be warned of the size of the object when linking to it. Where possible, text alternatives should be provided for multimedia objects, such as a description, transcript, or plain text version of a document.

Files should not contain redundant information. For instance, quite often audio files are presented as stereo, when they are actually just mono files: Presenting the file as mono can halve the bandwidth without any loss of information or quality.

To give some sense of scale, Table 2 shows the time to download a 20-minute video, audio versions, and the script, over a connection that provides an effective bandwidth of 20 kbps.

This is assuming a continuous connection—in reality, a lengthy download is unlikely to run to completion and may have to be resumed several times.

It is quite clear that the provision of alternative formats enhances low-bandwidth accessibility. For instance, for video, we refer to the

TABLE 2
Time to Download Different Versions

File type	Size	Time
Video, 512kbps, mp4	77MB	9 hours
Video, 64kbps, 3gp	10MB	1 hour
Audio, 32kbps, mp3	5MB	33 min
Audio, 6kbps, AMR-NB	800kB	5 minutes
Script, plain text	20kB	8 sec

“video-audio-text” cascade [2], recommending that video should be provided in multiple formats, including low-bandwidth friendly mobile formats. However, video should also be provided as an audio-only version, as well as a text transcript if possible. The provision of multiple video formats, and the transition from video to audio is particularly important, because it can often be achieved in automated workflows without increasing the complexity of the overall publishing process. For automated video workflows see, for instance, the OpenCast project.⁷

All users are effectively making cost/benefit decisions when choosing whether or not to download materials. While it may be argued that an audio-only version of a video resource is not as “valuable,” this judgment is based solely on the resource itself and does not factor in the user’s experience of accessing the resource. An audio version that can be used because it is readily accessible is far more valuable to its user than a video resource that cannot be accessed. Presenting resources in multiple formats and taking a “graceful degradation” approach empowers the user to make their own value judgments. It also gives users the choice to evaluate the materials before committing to a longer download.

Although this is the case for many OERs already, it is critically important that users can download materials, as opposed to materials just being available online (or, in the case of video, just available as streaming video). Obviously objects (such as images, javascript, and css) should be reused where possible to benefit from caching. To ensure that objects can be cached, sensible cache headers need to be set.

OpenLearn provides an example from the OER community, where multiple formats are made available programmatically, including print formats, xml formats, and various package formats. While these formats aren’t necessarily provided for low-bandwidth use, and don’t necessarily contain low-bandwidth formats, it is nevertheless straight forward to see how low-bandwidth formats could be added programmatically, without adding significant overheads to an existing multiformat delivery process.

2.3 Content Delivery Chain: Connection

Often the Content Delivery chain is seen from this aspect alone: The total physical bandwidth (“pipe size”) of the connection. Where Digital Divide and bandwidth issues are discussed, often those issues are seen purely in terms of such infrastructural considerations. Of course, connectivity is an issue: There is a disparity in terms of connectivity, and particularly in terms of the cost of connectivity.

However, even with much improved connectivity, the other elements of the chain are still limiting factors on available bandwidth and how effectively this bandwidth can be used.

2.4 Content Delivery Chain: Local Network

It is important that local area networks (LANs) are well managed, including

6. <http://www.it.usyd.edu.au/~matty/Shakespeare/shakespeare.tar.gz>.

7. <http://www.opencastproject.org>.

- good design of the LAN, subnets, and routing,
- appropriate use of caching, and
- provision for virus protection.

Moreover, a well-managed LAN includes the ability to diagnose (and resolve) network problems, for which monitoring tools are essential. It is important to also take into account acceptable use policy [11]. The combination of the policy, monitoring, and tools are summarized in the so-called “BMO triangle,” see [7] for details. Much of these are standard procedures in the North. However, in many institutions in the developing world, it is hard to retain suitably trained staff, and consequently local networks are not managed as well as possible.

Local area networks are important also from an OER perspective: Educational content can be provided on local area networks, so that users do not need to access the content via international networks.

2.5 Content Delivery Chain: Users

User behavior is one of the biggest factors in terms of effective use of bandwidth. Users share the available bandwidth with each other, and if some users are making heavy use of illegal file sharing, it stops other users from downloading OERs. In other words, it is essential to have an Acceptable Use Policy, and to raise awareness among users. As an example, we refer to TENET’s Acceptable Use Policy.⁸ If developing world institutions want effective use, then this can be made possible, but they must make some hard choices—one of them is prioritizing what this expensive scarce resource can be used for and by whom, cf., for instance [11].

But users can also empower themselves to make the most of existing bandwidth by using bandwidth optimizing tools. For instance, Aptivate hosts a free web-based service called Loband⁹ that reformats any webpage into a text-only form that radically reduces its size.

2.6 Mobile Connectivity

Some Internet users may be familiar with the fact that mobile phones work quite well in some regions even in developing countries, and that even web browsing on the mobile phone can be affordable and reasonable. How does this experience compare to the Content Delivery Chain and scarcity of international bandwidth?

All mobile Internet connectivity still needs to go through national Internet gateways, and as such incurs the same cost as broadband traffic. However, one of the reasons why mobile connectivity is apparently cheaper to the end user is because it is slow.

In many parts of the developing world, mobile browsing (where available at all), will continue at GPRS connectivity. This is similarly slow as a typical desktop connection (20 kbps as discussed above). However, often this limited connectivity is usable on a mobile device, because the mobile device has optimized applications (such as Google Mail for Mobile, and the Opera Mini browser).

However, the same connectivity used on a netbook or desktop computer (with standard browsers and webmail) would be painstakingly slow indeed. To give an example: On Google Mail for Mobile, only the message itself is transferred, and would be of the order of a few kB. However, sending the same message using some webmail clients might require 100 times the size of the message (or more).

There are two important lessons from this: On the one hand, accessing the Internet through GPRS (as well as faster connections) typically means a slow and unreliable connection. At the same time, using smart strategies to work with that slow connectivity, such as content transcoding in Opera Mini, can make this slow connection usable.

In this way, mobile devices work with certain key elements of the Content Delivery Chain better than other devices, and thus

allow to make better use of connections. We will now apply similar ideas to the delivery of OER.

3 FUTURE STRATEGIES

Because of the bandwidths available and the current size of OER content, from a user’s perspective in many Southern universities, it is not correct to say that they are fully online—in the sense that users do not have a useable connection with OER content.

At the same time, there is a viable Internet connection that could be of tremendous value if used optimally, so it is also not right to say these universities are fully offline when considering access to OER.

In fact, there is a mixture of both modes, and this hybrid online/offline scenario has historically not been well provided for. In analogous situations, we have seen technology models swing from offline to online modes and back again as at first bandwidth increases and then the size of content. And yet there is some indication that hybrid modes may finally be addressed prompted by the use of mobile phones for Internet access. By their nature, mobile phones have intermittent connectivity and some effort is being made to address the issues of synchronization and fluid models that allow applications to run remotely or locally or in arbitrary combinations of those modes.

While we are just starting to see this approach used in mobile applications, OER would benefit from scaling this approach up to large databases of OER content and infrastructure at the institutional level.

What could a hybrid approach look like? We can imagine a scenario where a user obtains a content collection, which is put online within the local area network, and then occasionally synchronized. Because of the way bandwidth is utilized by users, the available bandwidth varies with the time of day, and throughout the week. Times where the network is less busy can be used for bandwidth managed synchronization.

There are some initiatives that have partially implemented this process, including the MIT OCW mirror site program,¹⁰ which makes MIT OCW available in bulk, together with a password protected “rsync” facility. There are similar programs, such as the eGranary,¹¹ that make a range of educational resources available offline.

Providing offline content is important, but there are important requirements:

- Any OER/OCW provider should be able to easily provide their content as hybrid online/offline content.
- Any OER/OCW consumer should be able to easily obtain such content.
- Content needs to be updatable. Such updates need to respect the available bandwidth, and need to be automatic in as much as possible.

Clearly what would be most appropriate is a hybrid, synchronizing system that is both online and offline: A system that supports both modes and everything in between optimally.

What are the requirements for setting up such a system? In the following sections, we identify a number of components.

3.1 Federated Content Catalogues and Syndication

To be able to successfully and systematically make content available in such hybrid models, a good “catalogue” of OER content is needed. In other words, it is not sufficient to publish OER content in human readable form, but machine readable metadata and syndication features need to be provided, so that content can be pulled into hybrid delivery systems. This of course also has added benefits, such as enabling federated searching.

8. http://www.tenet.ac.za/about-us/AUP_v3-1.pdf/view.

9. <http://www.loband.org>.

10. <http://ocw.mit.edu/OcwWeb/web/about/mirror/>.

11. <http://www.widernet.org/digitallibrary/>.

We do note that various “content catalogues” are available, that go some way toward this. This includes, for instance, the Open CourseWare consortium OPML feed,¹² that links to various institutional OCW index feeds. Unfortunately, none of the feeds terminate with a link to the actual resources, not to mention alternative versions of the same resource.¹³

A notable exception to this is the OpenLearn index feed, which does terminate in links to actual resources. However, this OPML feed is not the one that is included in the OCWC OPML feed. The implication of this is that anybody wanting to gain access to the resources would have to go through the provider’s webpages, which, as noted above, is substantial in size.

While certain tools (such as wget) could be used to automate the retrieval of these pages, this is fragile, and would have to be tailored for each content provider. A solution like the OpenLearn index feed (that links to resources) is a more robust solution and we think should be the standard for inclusion in OCWC’s index feed, rather than feeds linking back to webpages.

We may compare this with the use of RSS feeds in podcasting, where a link to the actual resource (the media file itself) is included. Moreover, the yahoo media standard,¹⁴ allows for providing all important content alternatives via the <media:group> element.

If such detailed feeds and federated catalogues were available (that linked through to the actual resources), it would be possible to ingest OER and OCW materials into hybrid content delivery systems easily and systematically. This would also help searching: At the moment, there are a number of initiatives that aggregate OCW content essentially manually, and these could be helped by providing more automatic means.

Moreover, where content is available in machine readable form, automated transcoding and content reformatting can be applied to make content available in a broader range of low-bandwidth friendly formats (similar to Lo-band for webpages). With those elements, a content delivery network could be built, that would allow users to share and reorganize content in a systematic way. Content could, thus, be easily cached offline, or shared on local area networks through peer-to-peer methods. At the moment providing content systematically is difficult, because there isn’t sufficient organization of the content, and insufficient standardization on metadata and syndication formats. For further discussion, the resources of the Steeple project¹⁵ can be consulted.

3.2 Content Caching Mechanisms on Local Networks

Once the content is available in suitable catalogues, it can then be cached. For this to work effectively, innovative caching systems are needed. Key elements of such systems would include the ability

- to prefetch resources when the network is less busy (and in this way gradually retrieve entire OCW/OER content catalogues),
- to be “primed” through offline data transfer (content from DVD or memory stick; once the cached is “primed” it can retrieve updates online), and
- to reshare data through offline means (e.g., through local writing to DVD, or through local peer-to-peer file sharing).

3.3 Contributing Content Back to the Community

A key component of a hybrid online/offline OER delivery system would be the ability for users in low-bandwidth circumstances to contribute resources back to the global community more easily: Users in low-bandwidth circumstances would use a local repository to deposit their OERs, whether this is newly written OERs, or other OERs that have been adapted in some way. These resources would then be available to other local users. However,

the resources would also be retrieved from the local repository (in a low-bandwidth compatible way), and could be deposited with an international mirror server in a location with better connectivity.

In this way, locally authored resources could be shared globally: South to North, as well as South to South. In the recent UNESCO discussion [9], improving South to South sharing was recognized as an important and highly desirable element.

4 CONCLUSION

The strategies outlined could lead to significant improvements in low-bandwidth accessibility of Open Educational Resources, and we urge the OER community to address these issues with high priority. The issues are of high importance, particularly for OER producers and consumers in the global South, and are often missed from access discussions.

We see the methods that we have outlined above as pragmatic solutions to the situation now. We, as an OER community, are empowered to implement these relatively simple guidelines, and thereby provide significantly greater access for those who suffer from low-bandwidth connections, while taking nothing away from users in high bandwidth environments. This is not about creating a lowest common denominator Internet, but we advocate creating multiple accessible pathways to address differing needs.

We take exception to the view that developing countries will have to wait until their infrastructure improves before they get meaningful access to these valuable open content resources. Of course, we do not argue against bandwidth improvements. However, we need to do what we can do. We (as the OER community) are empowered to increase accessibility of the resources we produce, but we are not directly empowered to create infrastructure in developing countries.

The possibility of obtaining and utilizing significant amounts of OER may be an important incentive for improving and managing local area networks, enabling other aspects of OER use, including reusing and contributing OERs back to the community. It will also improve non-OER related aspects, such as the ability to communicate in general, and to make effective use of the connection for other educational or academic purposes.

While ever increasing bandwidths in the North allow us to put ever larger forms of content online, we must at the same time ensure we address the issues around much slower bandwidths if we wish to build OER resources that are truly globally accessible. We cannot allow ourselves the luxury of waiting for globally universal broadband access. We can and we must make very significant progress right now with the approaches we have outlined above.

REFERENCES

- [1] *Information and Communications for Development—Global Trends and Policies*. World Bank, 2006.
- [2] Aptivate, *Web Design Guidelines for Low Bandwidth*, <http://www.aptivate.org/webguidelines/Home.html>, 2010.
- [3] A. King, *Speed Up Your Site: Web Site Optimization*. New Riders Publishing, 2003.
- [4] *African Tertiary Institutions Connectivity Survey*, <http://www.aau.org/renu/docs/ATICS2006.pdf>, 2006.
- [5] Aptivate, *Bandwidth Management Position Paper*, <http://www.aptivate.org/attach/Projects.BMOPositionPaper/AptivateBMOPositionPaper.pdf>, June 2007.
- [6] *Average Web Page Size Triples Since 2003*, <http://www.websiteoptimization.com/speed/tweak/average-web-page>, 2010.
- [7] *How to Accelerate Your Internet: A Practical Guide to Bandwidth Management and Optimisation Using Open Source Software*, <http://bwmo.net>, 2009.
- [8] MIT OCW, http://www.sciencemedianetwork.org/wiki/MIT_OCW_Page_Sizes, 2010.
- [9] United Nations Educational, Scientific, and Cultural Organization, *Access to Open Educational Resources*, <http://oerwiki.iiep-unesco.org/index.php?title=Access2OER/Contents>, 2009.
- [10] *Free Online MIT Course Materials—Site Statistics*, <http://ocw.mit.edu/OcwWeb/web/about/stats>, 2010.
- [11] D. Greaves, “Making Strategic Decisions About Bandwidth,” *Int’l Network for the Availability of Scientific Publications Newsletter*, no. 35, pp. 3-5, <http://www.inasp.info/uploaded/documents/No-35-Newsletter.pdf>, Spring 2008.

12. <http://www.ocwconsortium.org/feeds/ocwmemberfeeds.opml>.

13. See <http://www.sciencemedianetwork.org/wiki/Syndication>.

14. <http://search.yahoo.com/mrss/>.

15. <http://www.steeple.org.uk/wiki/Syndication/formats>.