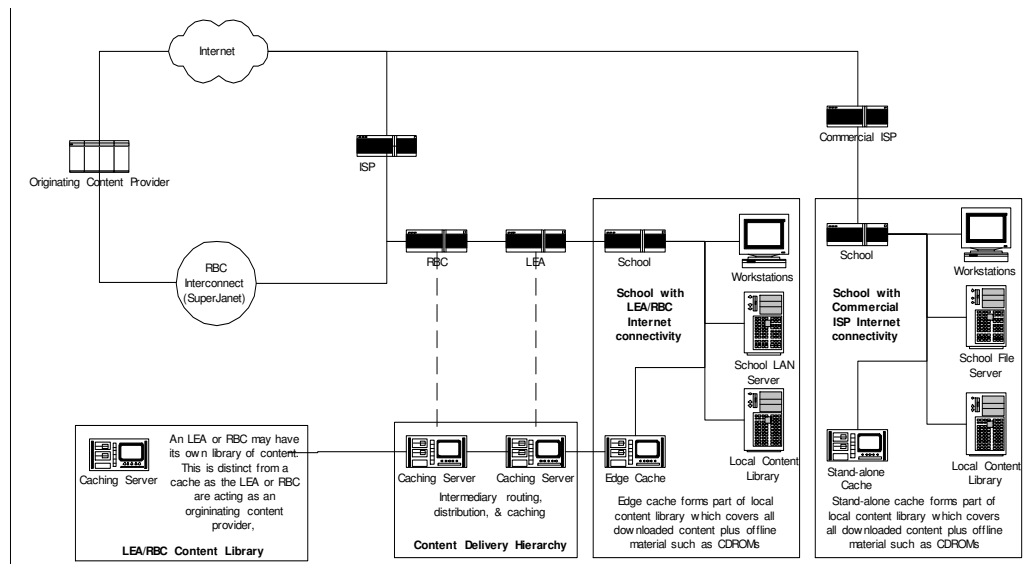


## Specification

### 1. Overview

The purpose of this paper is to provide a minimum specification for caching equipment and associated services. Additionally, this document is intended to increase understanding of issues surrounding delivery of media rich content to school networks over the Internet. As usage of the Internet and media-rich broadband content increases, there is an emerging market for systems known generically as 'Content Delivery Networks'. Content caches are one element of such systems, and in this paper we address two caching models (although there are others):

- **'Stand alone'** caches (sometimes known as 'proxy caches') – servers placed within a schools local network that hold copies of content from the Internet, and therefore lead to dramatically reduced access time for users of this content.
- **'Edge caching'** as part of a wider content delivery network – as well as providing the same functionality as a stand alone cache, an 'edge cache' would normally be part of a wider content delivery network that provides a more holistic solution across a number of organisations. Typically such systems would be implemented at LEA or RBC level, and in addition to content caching and delivery, might also provide learning content management functionality (managed access to a local library of content, sometimes provided through an LCMS – Learning Content Management System).



This paper aims to define the sort of device that might be used in either of these situations and is aimed at specifying the functions that a stand alone or edge cache should have (i.e. systems deployed into schools).

We do not prescribe the architecture that would be required above school level in LEAs or RBCs, although any infrastructure must not inhibit the model we have defined for caching.

Within this paper we describe caches, learning platforms and learning content management systems as independent but interoperable technologies. However, as these functions are interrelated there may be products that perform more than one of these tasks.

Schools should consult their LEA and RBC before making purchasing decisions, since LEAs and RBCs may already have, or be planning specific infrastructure for content delivery. It should be noted that using current technology there is no guarantee that two arbitrary systems can be integrated.

For the most part this specification covers outputs and functions rather than precise technical details. Where there is technical detail this will be the subject of regular revision to reflect market and technology developments

## 2. Introduction

The current cost and speed of broadband provisioning has lead to the setting of targets that will provide all schools with broadband Internet access by 2006. Until this is achieved the DfES has recognised the need to encourage the use of caching and content delivery services deployed to assist in the efficient use of existing bandwidth within both broadband and non-broadband connected schools (although non-broadband schools will derive greater immediate benefit).

This document is aimed primarily at LEA and RBC level purchasers, but equally at individual schools that are aiming to invest in caching services, where buying in accordance with this specification will offer the best possibility of future integration with LEA/RBC services.

- The specification is intended to be used for '**edge caches**' within institutions where Internet connection is part of a managed wide area network of an LEA or RBC. Where this is the case, the specification should be applied as a minimum set of requirements at every level of any hierarchy structure of caches, to enable the whole infrastructure to be compliant with the specification. It is clearly one element of the service and unlikely to be the sole criteria from the customer perspective.
- Additionally, institutions with direct access to the Internet through a commercial ISP may wish to make use of this specification for '**stand alone**' caching systems, not only for the best possibility of future integration with an LEA or RBC service, but because this specification will ensure that unrestricted access to the wealth of on-line material is possible. It is highly recommended that schools consult with their LEA or RBC before purchasing themselves, as there may be additional requirements for future integration with LEA/RBC services.

This document will form the basis of any procurement specification should DfES/Becta undertake such a central procurement exercise. The specification is based on seminars held in 2002 hosted by Becta. The specification is a minimum specification, and providers will be expected to provide scalable solutions upward from this specification.

## 3. Top level functionality

From an end users point of view, the caching system is intended for two purposes:

- Teachers and learners should be able to access content in the school as they normally would when accessing it on the Internet, but have this content delivered to them faster than would be possible without the caching system. I.e. the technology must be invisible to the end user, but provide a marked increase in performance where the bandwidth of a schools Internet connection and/or the number of potential concurrent users would otherwise make the use of such materials impractical.
- ICT co-ordinators need to monitor and manage how the caching system is working, including configuring the cache, managing what is stored, pre-loading new content.

To achieve this any caching service must as a minimum:

- Enable 'pre-loading' of specific content from the Internet, offline media such as CD/DVD, and other content sources such as LEA/RBC content libraries.
- Increase the performance of Internet connectivity for commonly used content using 'reactive caching' i.e. without the need for human intervention.
- Provide monitoring, management and reporting features.

These and other related requirements are described in detail below.

#### **4. Integration**

- At a minimum, the service must integrate with TCP/IP networks.
- Service providers must also ensure compliance with the protocols specified in e-Government Interoperability Framework – Section 6 Interconnection Specifications. These must be offered in addition to, rather than instead of, TCP/IP.
- LAN integration must be possible with minimal reconfiguration, and must not compromise the existing network design in relation to the Quality of Service, necessary for Video Conferencing, Voice Over IP and live streaming of data.
- The service must integrate with managed and personalised learning platforms and learning content management systems:
  - Must not affect settings applied to content such as user access controls.
  - Must not prevent applications from sending/retrieving user information to/from host(or other) sites required to deliver user specific information.
  - Must ensure sending/retrieval of learning objects required for learning platforms is seamless.
  - Must be compliant with learning environments conforming to SCORM version 1.2.
- In addition to integration with generic learning platforms defined above it must be possible to integrate the service with learning platforms as defined by the Curriculum Online Learning Platform Regime.

#### **5. Capacity**

Storage should be sized to offer best performance in line with the requirements of the institution.

- Hard disk storage must be a minimum of 70GB.
- Storage must be upgradeable and managed to ensure that large amounts of data do not need to be repopulated by the user.
- The end user must be able, if they wish, to define minimum and maximum portions of the disk for reactive caching and preloading of content. Ideally these levels should be dynamically managed.

#### **6. Reactive caching - delivery of on-demand content**

- The service must be capable of the reactive caching of content, and enhance the delivery of on demand content by minimising the time to display.
- The service must be configurable to make best use of out of hours bandwidth usage.
- The service must enable browser clients to access cached content transparently.
- The service must support the HTTP and FTP specifications listed in the current version of the e-Government Interoperability Framework - Section 6 Interconnection specifications. Specifically:
  - HTTP 1.0, 1.1 (including the pre-fetch command).

- FTP.
- Termination of HTTPS connections to the network.
- Internal HTTP connections.
- In addition the service must also support:
  - ICP version 1 and 2 (RFC2186).
- For interoperability with other systems, the service must be capable of handling the full list of formats and document types specified in the current version of the e-Government Interoperability Framework - Section 13 Information Access Specifications. The following subset of the e-GIF specifications are seen as being of particular relevance to caching:
  - HTML with associated resources and derivatives, e.g. CSS, JS, XML.
  - Office documents DOC, PPT, XLS, RTF, TXT, PDF.
  - Animation and Activities - current and latest versions of Flash and Shockwave.
  - Images – GIF, PNG and JPEG.
- The service must provide a mechanism to service requests for cached information, even if the origin server is not available, however this requirement must not allow authentication and authorisation systems to be compromised.
- The service must provide a mechanism to specify URLs that should not be cached.

## 7. Pre-loading of content

- The service must be capable of preloading user-specified content via provision of a simple control mechanism.
  - The service must be capable of preloading the following types of content:
    - Non-packaged content (i.e. individual web pages or digital assets).
    - Content packages that conform to the IMS Content Packaging specification.
    - Content packages that conform to the approved Curriculum Online Content Packaging standard (which follows the IMS format, but specifies the Curriculum Online metadata scheme for description of resources).
  - WAN-based preloading will be defined using a simple open standard for content packaging (referred to as a “manifest”). This, likely to be XML based format, is to be defined by DfES/Becta in consultation with content and service suppliers. Preloading should support this format irrespective of provider.
  - The device must check regularly (at least daily), with a remote source, that the manifest (XML or other) it holds is current, and make changes as a result, including the removal of files no longer required, and the addition of new cacheable components.
  - In the future these remote manifests might be integrated with portals and electronic content procurement, or held in a central directory.
  - Updating must be connection-agnostic and be able to use ISDN, Broadband, Satellite, etc.
  - The provision of local pre-loading of content must be possible, via either LAN or direct using CD/DVD.
- All pre-positioned content must be reliably retrieved without errors, and in pristine condition.
- A mechanism must be included to enable the management of content on the device, including the ability to define the expiry point of content.
- The service must provide the facility to control the amount of bandwidth used to update pre-positioned content during working hours.
- The service must allow schools to incorporate their own content.

- Unauthorised users or websites must be prevented from storing content on the device. e.g. to protect the system from unauthorised promotional content.
- **OPTIONAL:** The service provider should offer preloading of content before deployment to schools in conjunction with content suppliers.

## 8. Streaming Media

Support of streaming media and file formats must be provider-agnostic and comply with:

**OPTIONAL:** Support for streaming media formats – ASF, RAM, etc. Ideally this could be provided as a remote software update.

**MANDATORY:** Provision of non-hardware upgrade to enable caching, splitting and pre-positioning of streaming media.

- When operating as a streaming media cache, the device must:
  - Simultaneously deliver standard streaming media formats. MMS, RTP, RTSP, Quicktime and MPEG.
  - Simultaneously be able to split WMT, RTP/RTSP, Quicktime and MPEG into 4 live streams.
  - Be capable of handling multi-bitrate streams.
- The service must be able to block user specified streaming media (i.e. Internet radio).
- The service must be configurable to the amount of bandwidth allocated to streaming media.

Where the option to not provide streaming media is taken, the service should not prevent the network accessing streaming media.

## 9. Headless operation

- The device must operate as an appliance with no requirement for keyboard, mouse or display for effective operation, with exception for installation/configuration.

## 10. Monitoring and Management

Devices to be capable of remote management through a simple web-based interface by appropriately authenticated users and must be able to:

- View a list of all preloaded content, including administrative information such as:
  - Date added.
  - Size on disk.
  - Expiry date.
  - Source.
  - Last accessed time .
- Manage this content, for example:
  - Store, retrieve and version control.
  - Delete content ahead of its expiry.
  - Add new content by entering a target URL for a remote manifest list.
  - Pin content within the cache indefinitely.
- Change the relative amounts of storage allocated for preloaded and reactive content.
- Provide a secure web based interface for authenticated users to reconfigure any system parameter both locally or remotely.

- Provide a centralised mechanism to distribute and activate software upgrades (including OS) and patches for the device.
- The service must support SNMP to allow the use of existing SNMP monitoring tools.
- Where applicable, allow multiple caches to be monitored, updated, supported and maintained, by groups or by other structure.
- Notify administrators of serious or 'immediate action required' alerts.

## 11. Reporting

The service must provide:

- Subscription history details, and allow the management of subscriptions.
- Reporting functions that allow logs and summaries to be remotely interrogated (for example to show percentage of requests fulfilled from cache and most used sites).
- The ability to store log files in CSV or similar industry standard format for optional further processing.
- A mechanism to specify the parameters that are logged.
- Centralised log management for multiple caches.
- A configurable mechanism to manage the storage consumed by log files.

## 12. Access Control, Authentication and Digital Rights Management (DRM)

The service will need to ensure that the rights of content providers remain protected and also that the purchasing authority (school, LEA or RBC) stay within the terms of their licensing agreements. As DRM technologies develop, they might want to provide features such as management of subscriptions, or a unified single login to control access to a variety of content and services, and the cache should not prevent them from doing this.

Specifically:

- The original content owners' permissions must be enforced by the cache. I.e. in serving content to users from the cache, the system should not impair DRM features put in place by the content owner, nor allow them to be bypassed.
- The device should integrate with existing LEA or RBC authentication methods.
- Integrate with industry standard systems including LDAP, RADIUS, NTLM and Kerberos.
- Enforce permissions on served content in a way which is indistinguishable to content delivered from the origin server.
- The system should not inhibit the use of Digital Object Identifiers (DOIs) for referencing the location of resources, and controlling access to them.
- The system should not inhibit the use of Access Control systems or DRM technologies in general.

## 13. Content Filtering, virus checking & security

- The service must not prevent or impair the normal operation of filtering of Internet content, including filtering services supporting several age/role appropriate levels of filtering.
- The service must not prevent or impair the normal operation of virus scanning tools.
- The service must not prevent the normal operation of firewall security.
- Services should support the ICAP protocol.
- Content must only be downloaded to the cache at the request of the institution, i.e. the service must protect the cache from promotional material.

#### **14. Service specification**

- Content delivery must be provided as a service to schools rather than a hardware purchase. This managed service might be supplied, for example, by an LEA, RBC or a commercial supplier. The contract is for the service, and not tied to specific hardware so, for example, in the case of hardware failure the supplier would be expected to provide replacement equipment as is required to continue the service without additional costs.
- Support for all services must be provided via a single point of contact, at no more than national call rates, from 8am to 6pm every working day, excluding bank holidays.
- At a minimum, on site technical support must be provided within one working day of the reporting of a fault, and replacement equipment must be provided within 24hours of the reporting of a fault, should technical support not be able to rectify the fault.
- Service should include ongoing management, maintenance and technical support.
- Quality of service - to be defined as part of procurement exercise.
- Performance – to be defined as part of procurement exercise.
- The service provider must provide an installation service for equipment.
- Failure of the service, or part of the service, must not prevent access to the Internet or use of other network services, or require manual reconfiguration of other equipment, including workstations.
- The service should require minimal technical expertise for integration, system outage reinstallation, or recovery process.

#### **15. Advice and guidance**

- Service suppliers must be able to offer appropriate guidance for large, split site or clusters of schools.
- Guidance must be provided on sizing services and the number of clients and bandwidths supported by devices.
- Service must support and offer an integrated hierarchy of devices to best deliver content within a network – for example by locating devices at LEA and RBC as well as at the school edge.

#### **16. Training**

- Training on local content management functions must be offered at the time of installation and be available for the lifetime of the service at the request of the institution.

#### **17. Availability**

- The service should be offered to broadband and non-broadband schools, to facilitate easy migration without requiring device and contract replacement.
- Remote management of content should be possible through manifests. For example, a school could choose to preload content as suggested by its LEA. To implement this, the school would choose to load a manifest hosted by the LEA. The LEA would then update the manifest and the school would then automatically synchronize. This model might also be offered by commercial suppliers as a service.
- In the future these remote manifests might be integrated with portals and electronic content procurement, or held in a central directory.

## **18. Useful links**

Note – these documents are subject to review and update without notice. It is the responsibility of the service provider to keep themselves up-to date with latest developments.

E-Government Interoperability Framework v4.0

[http://www.e-envoy.gov.uk/oeo/oeo.nsf/sections/framework-egif4/\\$file/egif4.htm](http://www.e-envoy.gov.uk/oeo/oeo.nsf/sections/framework-egif4/$file/egif4.htm)

How to prepare a Local Area Network for Curriculum Online

<http://www.ictadvice.org.uk/index.php?section=te&cat=007000&rid=1740>

IMS Content Packaging

<http://www.imsproject.org/content/packaging/index.cfm>