
Grand
Challenges
Online Research
The changing landscape
of Online Delivery
BBC R&D

Grand Challenges

Online Research

The Changing Landscape
of Online Delivery

Foreword

Welcome!

The past 30 years have been an exciting time in entertainment technologies, network research, and online delivery. With the stealth rise of ubiquitous computing through mobile phones; the shrinking of the planet by audio and video conferencing via instant messaging; democratisation of news, knowledge and opinion reporting via wikis and blogs, the future appears to be even more exciting.

In *Building Public Value*, the BBC has committed to the delivery of BBC content to the British Public over this multitude of systems, in a way that is enjoyable, accessible and repurposable by the British Public in the way they choose.

Whilst this is an exciting and intriguing future, this does require a new evolution in the way large media such as video, and audio are stored, and delivered. This document explores this vision, and suggests a number of key challenges that need to be addressed along the way. These are however only the first steps into the 21st century that we are taking. In order to take the next steps we must meet these challenges head on, resolve to solve them and move forward. Our view is that this is a challenge that needs to be met in participation with industry and the community.

Michael Sparks
Senior R&D Engineer,
Digital Media Group, BBC R&D
http://bbc.kamaelia.org/

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Choice is Challenge

The BBC will launch a Creative Archive – free access to BBC content for learning, for creativity, for pleasure.
The BBC will make its services available when and where people want them, with a new generation of BBC on-demand services.

Building Public Value, Chapter 2, http://www.bbc.co.uk/thefuture/bpv/media.shtml

This vision poses a number of challenges on the long term horizon simply due to the vision's grand scope. This document discusses a number of these, so that they can be met, resolved, and the vision delivered.

The British Public comprises 60 million people and is often approximated as 20 million homes. This belies a hidden assumption of one display per home. However, many homes feature more than one display already. A growing number of homes have more than one computer in the home. In such homes, laptops are rapidly becoming second screens.

Thus whilst many calculations are based on the possibility of 20 million homes, it is entirely possible that we may well in the future need to consider the potential audience in terms of 60 million people.

Internet Search://

> Long Tail Distribution Wired

The BBC archive comprises of over 1 million hours of programming. Even offering as much as 10% of this archive to 60 million people has consequences. Online services tend to follow a trend: offer a choice, and all options get chosen. This occurs whether 10, 100, or 100,000 options are provided This behaviour is often called a long tailed distribution, and appears to be innate human behaviour. Music charts, cities sizes, letter distributions, etc all demonstrate this.

Essentially this means that given a choice, somewhere, someone will take that choice. Wildly popular items turn out to be just a small subset of all choice made. This drives the economics of online music stores. The bulk of the business comes from the long tail - the variety of content available.

Choice itself forms unique challenges.

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Offering the entire audience such a wide choice will have the following result:

- They will access almost every piece of content offered
- They will do so often
- They will not access the same data at the same time

Bottom line is we're potentially talking about 60 million people watching 60 million different things. Scaled down to just 1 million people, the challenges remain. Even if they're watching the same thing many will be watching from different time points.

All of this will occur in the face of a changing landscape for technology as used by our audience. Platforms proliferate, storage increases and platforms change. Much of the online delivery world for content is based on proprietary systems rather than open standards.

Any future plan and discussion of long term challenges must be discussed against this background. Inside each challenge there are latent issues that will cause issues when scaling BBC online delivery. After identifying common themes we will come back to the grand challenges.

Challenges: Short Version

To give a flavour of these challenges :

- It will be simpler to deliver this choice to and ***from*** every home where possible. It could fit in your home, now.
- Our audience will be sharing their own creations, and will expect to share ours.
- Scaling the BBC's online delivery infrastructure using proprietary systems is not cost effective
- Current distribution methods cannot deliver the BBC's vision.
- Moore's law has ended in terms of raw CPU power(Mhz) - this will radically change the way systems we use are built.
- The BBC has always sought the best possible quality for a given delivery system. Growth in bandwidth at home and at the BBC's side will offset each other.

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Changing Technologies

Network delivery speeds to the home have increased 10 fold over the past 10 years. Home storage capabilities have increased 1000 fold over the same time period. Mobile phones **have** become the true converged ubiquitous computing platforms, with gigabytes of storage, 3D visual capabilities, and video/audio recording and playback.

Internet Search://

> media consumption devices

Computers **have** shrunk to laptops, intelligent displays, pen and touch screens. Home entertainment platforms have gone from multichannel TV with simple computer games to near interactive 3D movies, gesture based input, surround sound audio and entire collections of video fitting inside ring binder sized wallets.

There is currently reason to believe that these trends will continue over the next 10 years. Home network delivery speeds will continue to rise relatively slowly, but at least 10 fold. Storage may continue to grow at a rate of1000 fold per decade. Incredible as this seems today, this means that even a laptop may be capable of storing dozens of terabytes.

Devices will continue to become more personal, and current esoteric platforms, such as paper with e-ink may become more common. Connectivity to and from these devices will become faster, more convenient and movement of data between devices will be an **expectation**, not a privilege. If content providers do not enable this, the people will enable it for themselves, sharing their results.

3D interactive capabilities of "games" consoles will become film quality, leading to an escalation of format wars leading beyond 2D gesture capabilities to full 3D resolution and scene recognition. Camera integration is leading to video phone style systems becoming common place, very soon. Speech recognition will become common place. Audio/Video search will take on a brute force approach, supplementing and often replacing metadata approaches of the past.

It is into this environment the BBC will be delivering its content.

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Deliver The Archive?

The BBC audio archive comprise of 500,000 hours of audio. At bit rates for an iPod is a mere 27TB of storage. In terms of *current* technology this would fit on commodity disk stores which could sit under a desk. In terms of solid state storage – you need 27,000 MMC cards. In practice that means 1 cubic foot of storage. Using video bit rates used for the Olympics you're talking either 3 desks for disk based storage, or 1 shelf full for solid state.

Internet Search://

> gutenberg project

Clearly these bit rates are not suitable for archival and would be inappropriate for television production. The idea that even the entire BBC archive could fit in most people's homes today is a startling one.

If the amount of the archive the BBC chose to make available was as much as 10%, then it becomes realistic for the audience to download now, choose later – even on a laptop – if storage does continue to as projected.

In this world it makes more sense to deliver the content available en masse to the entire audience, than it does to pick and choose. Obviously there are technical hurdles to jump over in achieving this, but conversely having the content almost everywhere massively simplifies this distribution. We allow the network and users to populate the network.

It seems inefficient to force the audience to return to the BBC for content when they will be able to locally store all the content we make available.

At this point in time the question then becomes not "how do we serve millions of concurrent stream", but rather "how do we ensure that we populate the millions of storage devices out there?"

Challenge: How do we deliver a BBC repository, dozens of Terabytes in size, to every home in the UK, regularly refreshed, in a way that copes with disk failures and data loss.

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Content Sharing

One of the earliest rules small children are taught is to share. It is one of our greatest instincts and people are measured by their generosity and willingness to share. Our audience already has the means to create and share their creations. A digital camera for £100 today allows the budding video artist to create their own works trivially – given talent or determination.

Internet Search://

> Darknet

Our audience is choosing to share their work online. They will expect to be able to share works they create, derived from our content, online as well. Some will be willing to share their results with us as well as with each other.

Currently the mechanisms used for sharing are either slow or inappropriate – such as websites. They are driven underground due to legal fights over technology – an arms race simply evolving stronger, more capable P2P systems.

Despite being often used in clearly illegal ways, P2P is inherently a legal sharing system and is becoming more stable. It's usage by the general population for sharing is likely grow beyond its already unprecedented levels. Growth of legal founded P2P systems provides opportunities for the industry, and the public.

Lawful use of P2P will lead people to changed expectations. People will expect to use content they have purchased in much the same way as they use content they produce. Placing content in locked envelopes, may to lead to the audience find means of opening it, and then using it as they wanted to – or they may just go elsewhere.

The resulting scenario will be that people will expect high quality content to continue to be created and made available by the BBC supplementing BBC content. New forms of culture may spring up, perhaps back to simpler but richer story telling.

Challenge: Provide systems, acceptable to this mature audience. Systems that foster this culture, but ***ensuring*** that people are still paid for producing content, whilst recognising that the audience will create a lot of content themselves.

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Proprietary Past, Open Future

Online audio/video delivery has taken 3 forms to date. Traditional media has *largely* chosen to use systems from either Microsoft or Real, with a small subset choosing systems from Apple and others. More niche but highly profitable companies have explored sealed Java applets or simple media download over HTTP. All have found their content "leaks" onto private sharing systems, often based on P2P technologies.

Internet Search://

> IETF
> ISOC

By and large however, the proprietary ***systems*** – consisting of closed protocols, codecs, players and servers - are well known for producing systems that currently deliver the goods as far as online AV delivery is concerned. Even so these systems deal with thousands, not millions.

Proprietary server systems essentially use a mechanism for charging that boils down to audience size. A physical server can handle a fixed maximum number of listeners or viewers. Charging per server essentially penalises popularity.

Increasing the audience to millions from thousands is a 1000 fold increase. Whilst system providers would jump at such a financial opportunity. However, it is not reasonable to increase costs to the license fee payer 1000 fold.

At present, proprietary systems fail to provide the BBC any way of mitigating costs. Open systems based on open standards often replace the early adopted proprietary systems, simply in order to scale. This has been likened to proprietary medicines being replaced by generic drugs over time.

Aspects impacting any new standards would include: agility in adopting new video formats that come into existence every day; ability to switch from streaming to downloading; ability to switch from single source to multiple sources.

Challenge: Open *systems* for media delivery must made to scale to very large scale unidirectional streaming, or else many of the other visions in this document falter.

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

P2P and Multicast, Together

Multicast is similar to traditional broadcast in that the BBC sends a single stream of data out onto the Internet. The underlying networks copy it to the interested clients. This can allow for additional traditional style live services over the Internet.

Peer to peer (P2P) allows home users to query the network to find content that exists on the network, and retrieve that content from as many sources as necessary to get a good download speed. This can allow access to a huge range of content that is already in the network. Also since clients do not need to return to the BBC it minimises license fee payer costs.

Internet Search://

No documents match your search

When taken alone, neither multicast delivery nor peer to peer (P2P) are able to deliver the full vision in Building Public Value. Multicast lets us offer higher-quality streams of live content as long as the audience is all watching same thing. P2P is very well suited when the audience chooses to watch different things.

However when we look at the problems both approaches have we note that the weaknesses of one are the strengths of the other.

Issues with multicast: (potential strengths in P2P)

- It is limited by the size of your smallest consumer's connection
- It is not generally deployed on the wide area Internet for both technical and business reasons
- Losing content can result in either breakage or thousands (and potentially millions) of clients saying "please resend".

Issues with P2P: (potential strengths in multicast)

- The network cannot be trusted – viruses, spam, defacement
- Distributed search and navigation systems are difficult and are only in their nascent stage of development.
- P2P relies on delivery between homes. How can we deliver 8Tbit/s downlink from a 4Tbit/s uplink?

Despite the issues with both, a solution based on *both* appears to hold great possibilities. P2P can deal with the problems of Multicast and Multicast can deal with the problems of P2P.

We have the desire to make available this choice to every home in the UK. Clearly not everyone would choose to receive it, this is be feasible in terms of home storage. Multicast is however available now for a subset of our audience, thanks to work done for the 2004 Olympics.

In this scenario the entire P2P network becomes a large intelligent cache which is able to fill in missing blanks and missing chunks without needing to return to the BBC.

Seeding a P2P system is often a slow process. Initially one person adds content and on demand usage grows. If however we wish to seed a P2P system, as a broadcaster we have other options worth consideration:

- We can broadcast our content specifically flagged for caching into a P2P system with associated meta data
- We can multicast our content, as it is digitised from the archive, over a network connection for people to store and cache in their local P2P systems.
- Any requests for content that come back to the BBC can be sent out also on the multicast channel, so that the number of sources in the network becomes very large very fast for our audience that does not have multicast available.

This has a number of advantages – P2P can bridge the gap between multicast islands or multicast deserts by rebroadcasting locally. Clients receiving our content via multicast can ask for missing content via the P2P network, rather than all simultaneously clogging the central server.

Since the network becomes inherently untrusted, mechanisms for trust (by users) need to be added to this system, which deals with a number of security issues in both multicast and P2P systems.

Challenges: We need to build new Collaborative Client-Hub protocols, allowing integration between broadcast, multicast, P2P systems and our archives. P2P and multicast systems need to be adapted to allow and take advantage of thousands of P2P sources becoming simultaneously active. Security mechanisms such that the audience can trust content is a key issue in such a system.

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Who's Free Lunch is Over?

> *On Intel chips, we reached 2GHz a long time ago (August 2001), and according to CPU trends before 2003, now in early 2005 we should have the first 10GHz Pentium-family chips. A quick look around shows that, well, actually, we don't. Herb Sutter, Dr Dobb's Journal, February 2005.*

Moore's law was formulated by Intel's Gordon Moore in 1965 and suggested that the number of transistors in a given area would double every 18 months. He was right, and the observation prompted many analogies in other areas – including raw CPU speed. This held true and developers could gain speed as a "free lunch" simply by waiting.

Internet Search://

> The Free Lunch is Over

Moore's law has ended in terms of raw CPU power as measured in raw clock speed.

Developers can no longer expect performance gains by simply waiting. CPU manufacturers are utilising the old maxim of "many hands make light work", and moving to highly parallel architectures. This includes home systems like Sony's Playstation 3 which features 9 CPUs inside a single chip.

Software must be written with parallel systems in mind, changing dramatically the way it is produced. Whilst this is not a problem directly for the audience, this does affect the BBC's long term ability to deliver online.

Taking advantage of these CPUs requires specially designed software. Parallel computing is considered by many as difficult – which needs to change, since in a few years parallel systems will be common, and the BBC will need maximise the use these systems. Online delivery is a naturally highly parallel activity, and so is needs amongst the earliest to benefit after the change.

Challenge: Create a usable system for producing parallel software systems that promotes reuse, resilience, scales on new concurrent hardware platforms, and due to design naturally encourages the design of parallel software systems.

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Moore's Law at Home vs Desire for Quality

The growth in network upload capacity for servers will help mitigate BBC distribution costs from a central location. However growth in download capacity by our audience's tumbling costs will wipe out any savings.

Internet Search://

No documents match your search

Traditionally the BBC has aimed to send the best quality audio and video services that the user can receive. Over dial-up we provided video at download speeds that fit over a dial-up. For the Olympics – we assumed that broadband users have 10 times that capacity. As a result we provided video 10 times larger and much closer to the casual observer to digital TV quality (10 times more on both sides).

In years to come home download speeds will grow 10 fold. Our audience will be able and expect to, receive HDTV quality video over broadband. Likewise the BBC will also want to deliver this quality. Upload capacity would need to grow 10 fold again, *simply to service the current audience size – eliminating savings in upload costs for the older poorer quality.*

As every person buying broadband knows – greater capacity costs more. Simply changing from 2-4 million streams *per day* to 2-4 million users *per second* means our upload capacity means costs to the license fee payer grow dramatically. Grow that 10 fold again, and you get a very large number indeed!

Since Moore's law doesn't really help us, any work we can do to reduce the need for users to user their full download capacity saves us money. Reducing the bandwidth required for HDTV quality video will enable the BBC to provide the audience with higher quality video using a flexible, scalable approach.

Challenge: To find a way of capping the *need* for bandwidth usage on the client end. This may mean that research into higher quality video in smaller and smaller bit rates becomes ever more important to allow the BBC to remove Moore's law from the client side of the equation.

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Summary

The key aim is to pull people's awareness towards specific key problems which are thought to be very difficult, but achievable in a decade or two.

The sheer scale and scope of the problems caused by 20-60 million different items of content to be viewed *simultaneously* from the BBC, potentially online is a challenge that should be undertaken, but should not be undertaken lightly.

There *are* BBC projects underway looking at these challenges, but the scope of this document is larger. The challenges laid out in this document will exist, and need resolving – whether the BBC does it or asks industry to help. Clearly the ultimate solution will be a mix of both in areas of appropriate expertise.

Meeting Challenges

BBC R&D projects exist that seek to address some of the issues laid out here. The Kamaelia[1] & Dirac[2] projects are aimed specifically at the some of the challenges in this document. BBC and Industry collaborations such as *Share It!*[3] have been looking at issues related to these challenges. Some of these challenges naturally fit elsewhere with other groups.

However, the reality is that the BBC and industry cannot assume that the public will follow. P2P has shown that the Internet Community – that is anyone competent, willing and connected to the network – will go their own way if the public is not involved. There is almost always a better way.

The style of this document is inspired by the "Grand Challenges in computing" documents produced by UK Academia[4]. These grand challenges have a long history. One notable success in recent times has been the Human Genome Project. A project that was successful at driving forward state of the art, but that which ultimately failed in it's original goal was Hilbert's Grand Challenge to Mathematics around 1900.

[1] http://kamaelia.sourceforge.net/
[2] http://dirac.sourceforge.net/
[3] http://www.extra.research.philips.com/euprojects/share_it/
[4] http://www.bcs.org/BCS/Awards/Events/GrandChallenges/conferencereports.htm

Grand Challenges, Online Research

The Changing Landscape of Online Delivery

Deliver The Archive ?

Challenge: How do we deliver a BBC repository, dozens of Terabytes in size, to every home in the UK, regularly refreshed, in a way that copes with disk failures and data loss.

Content Sharing

Challenge: Provide systems, acceptable to this mature audience. Systems that foster this culture, but ***ensuring*** that people are still paid for producing content, whilst recognising that the audience will create a lot of content themselves.

Proprietary Past, Open Future

Challenge: Open *systems* for media delivery must made to scale to very large scale unidirectional streaming, or else many of the other visions in this document falter.

P2P and Multicast, Together

Challenges: We need to build new Collaborative Client-Hub protocols, allowing integration between broadcast, multicast, P2P systems and our archives. P2P and multicast systems need to be adapted to allow and take advantage of thousands of P2P sources becoming simultaneously active. Security mechanisms such that the audience can trust content is a key issue in such a system.

Who's Free Lunch is Over?

Challenge: Create a usable system for producing parallel software systems that promotes reuse, resilience, scales on new concurrent hardware platforms, and due to design naturally encourages the design of parallel software systems.

Moore's Law at Home vs Desire for Quality

Challenge: To find a way of capping the *need* for bandwidth usage on the client end. This may mean that research into higher quality video in smaller and smaller bit rates becomes ever more important to allow the BBC to remove Moore's law from the client side of the equation.