Using Content Distribution Networks for Astronomy Outreach

Mathias Jäger
ESA/Hubble Public Information Officer
mjaeger@partner.eso.org

Lars Lindberg Christensen
Head of ESO education and Public Outreach Department (ePOD)
lars@eso.org

Mathias André
Web & Advanced Projects Coordinator at ESO
mandre@eso.org

Keywords
Content distribution network, data exchange, sharing outreach materials, video streaming speeds

Thousands of people from all over the world search the internet on a daily basis for the newest discoveries in astronomy: be it in the form of press releases, high resolution images, videos or even planetarium fulldome content. The growing amount of data available, combined with the increasing number of media files and users distributed across the globe, leads to a significant decrease in speed for those users located furthest from the server delivering the content. One solution for bringing astronomical content to users faster is to use a content delivery network.

A growing problem

The basic architecture of the internet and its organic pattern of growth result in a number of bottlenecks that impact its overall performance and mean that its end-to-end reliability cannot be guaranteed (Nygren, 2010).

During an exchange of data both the server and the user have to continuously exchange acknowledgements of the data packets being sent. So the distance between the server and the end user can become the overriding bottleneck in download speed — or throughput — and have a large impact on video viewing quality. Therefore, if the server hosting the files is not close to the end user then the streaming of HD videos, or even the download of large files, becomes difficult, or even impossible.

The magnitude of this effect can be seen in Table 1 from Nygren et al. (2010). Their

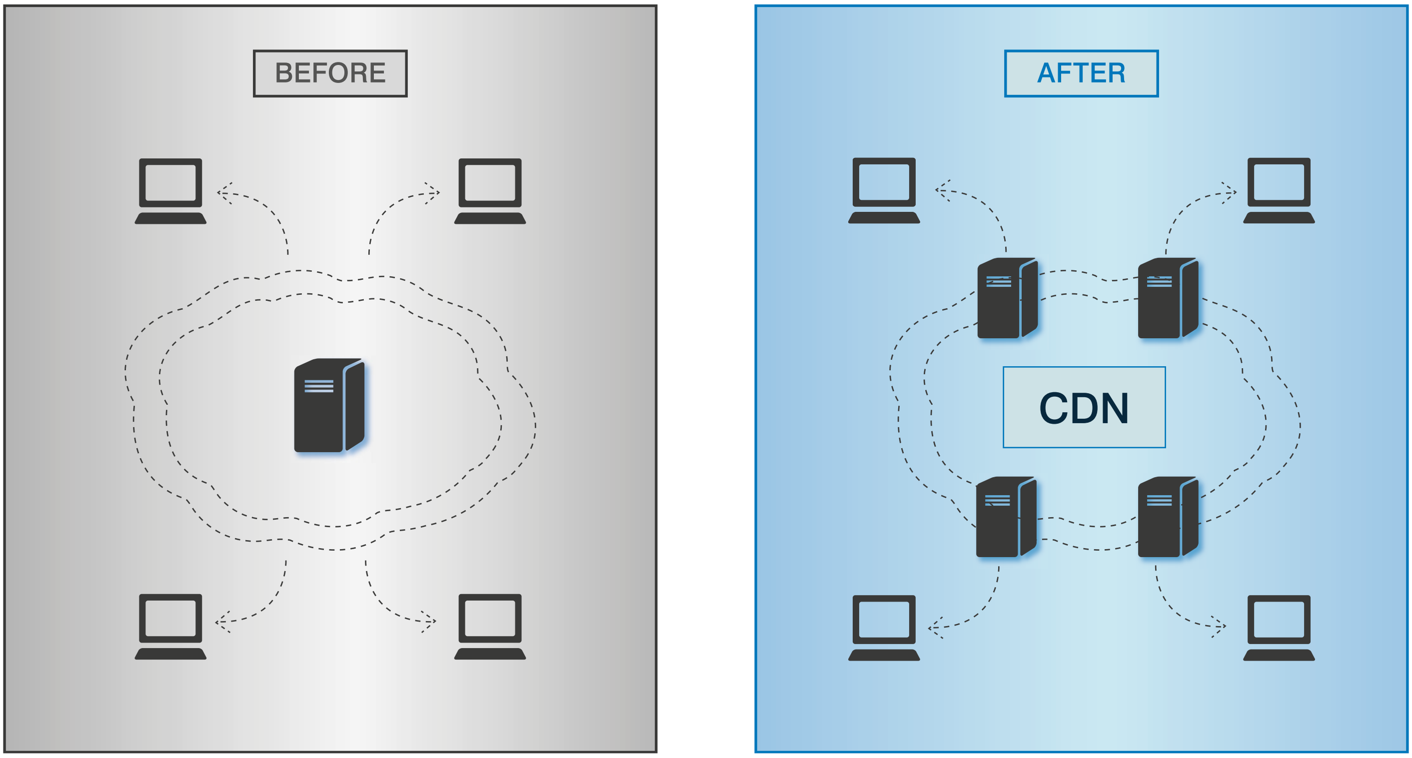

Figure 1. *Comparison between content distribution from a single server (left) versus a content delivery network (right).*

Distance from server to user (km)	Network round-trip time (ms)	Typical packet loss (%)	Typical throughput (Mbit/s)	Typical download time for a 4-gigabyte DVD movie
Local: < 150	1.6	0.6 %	44 (high quality HDTV)	12 minutes
Regional: 800–1500	16	0.7 %	4 (basic HDTV)	2.2 hours
Cross-continent: ~ 5000	48	1.0 %	1 (SDTV)	8.2 hours
Multi-continent: ~ 8000	96	1.4 %	0.4 (poor)	20 hours

Table 1. *Effect of distance on typical throughput and download time for a fast consumer connection. Credit: Nygren et al. (2010).*

measurements show that the network round-trip time, which measures the time required for a data packet to travel from a server to an end-user and back again, can reach up to about 100 milliseconds (ms) for a distant server with more than 1% of all packets lost on the way. The resulting throughput in megabits per second decreases dramatically the further away the server is.

Until February 2015 the European Southern Observatory (ESO), which also hosts the European Space Agency's Hubble webpage spacetelescope.org, only had one server location, at ESO Headquarters in Garching bei München, Germany. All content from this leading ground-based observatory, including almost 15 000 breathtaking high-resolution images and more than 3500 spectacular videos in many formats up to, and including, Ultra High Definition[1], was distributed across the globe from this single location. This posed a problem, as analysis showed that more than 40% of the public traffic to both sites originated from outside Europe. Since the content was available only from the server in Garching, this led to a significant decrease in speed for those users accessing it from outside Europe. This issue is accentuated further in the case of videos that are designed to be viewed on screen in real-time, as high latency and slow transfer speeds result in stuttering video playback.

This issue is not unique to ESO, but is relevant to anyone distributing content via the internet to international users. So far, few scientific organisations have addressed this issue, meaning that significant parts of the content fail to reach audiences on the other side of the globe.

Recently, this problem has been exacerbated as High Definition (HD)[2] and Ultra HD videos have become a much larger part of worldwide internet traffic. Many scientific and astronomical organisations, including ESO, now distribute video podcasts in HD or Ultra HD format. Some of the recent video files available from ESO and ESA/Hubble can easily reach several gigabytes, especially those in formats produced for broadcasters. ESO has also found that a

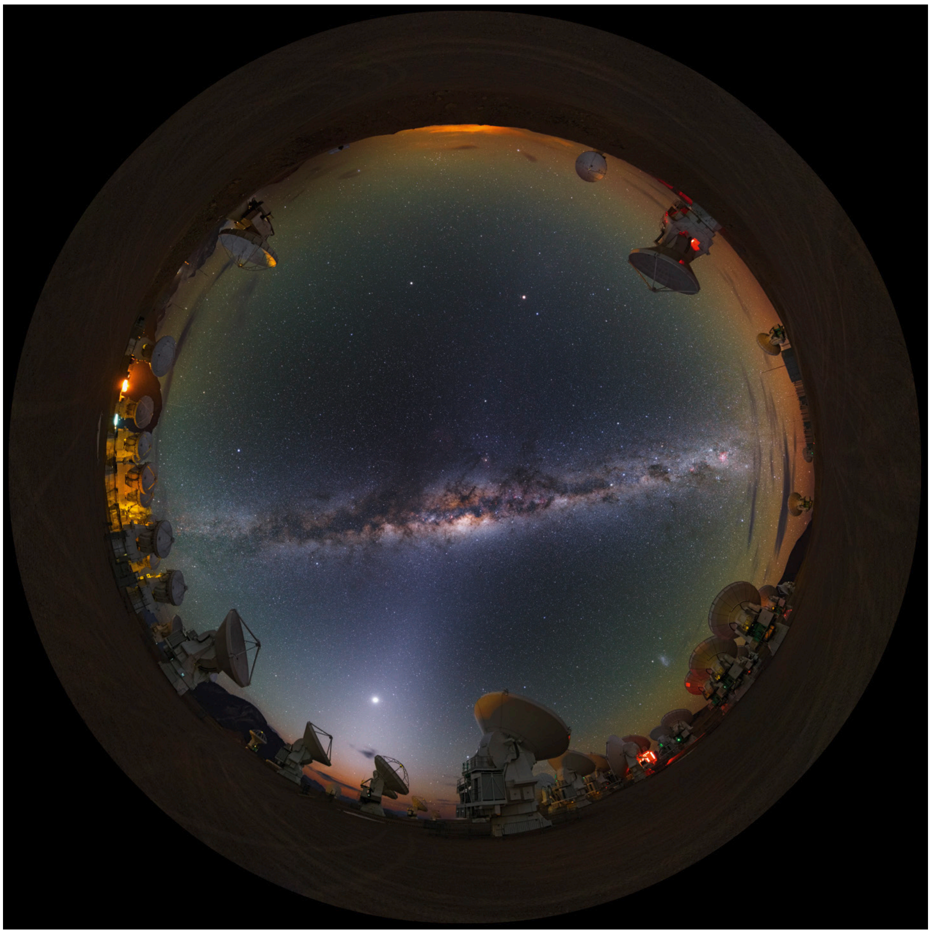

Figure 2. *Fulldome frame for use in a planetarium. Credit: Y. Beletsky (LCO)/ESO*

good fraction of the more technologically avid of our audience want the images and videos in the highest resolution possible. To fulfil this need, we have had to find a way to deliver these products effectively.

The biggest challenge: Distributing planetarium materials online

Another recent trigger that has spurred ESO and ESA/Hubble to look into this problem has been a need to pursue the production and distribution of fulldome planetarium shows for the upcoming ESO Supernova Planetarium & Visitor Centre[3] at ESO's Headquarters in Garching, Germany. The planetarium community is expressing a deep and urgent wish to receive more fulldome productions[4] from scientific organisations, as well as more direct and timely access to visualisations of new discoveries and data. These fulldome movies can reach 400 gigabytes for a full 30-minute show and make the need to solve the above-mentioned geographical problem even more pressing.

To demonstrate the feasibility of distributing files of this gargantuan size online — provided the problem can be solved — Table 2 shows the download times of planetarium clips and shows for different typical internet connections. The throughputs of the connections quoted are ideal ones with little associated latency, packet loss, network outages, and inter-network friction.

Table 2 demonstrates that distributing planetarium material — even full shows — is entirely feasible, although there are several requirements:

- A decent internet connection is needed, similar to a good private 16 Mbit/s internet connection.
- The user needs a bit of patience and a full night or weekend at their disposal.
- The use of a download manager like Free Download Manager[5], or a command line tool like aria2[6] in case there is an internet outage, is required on either side.
- Sufficient free disk space is required at the location where the file is to be saved.
- A nearby server is needed to deliver the material and solve the geographical problem.

None of these expectations are unreasonable for many facilities, although there are naturally local challenges that may make the download times longer.

Content delivery networks

Content delivery networks (CDNs) can solve all of the above-mentioned problems. A company provides a distributed system of servers deployed in multiple data centres worldwide. This network ensures that users access data from a geographically nearby location, which gives a better guarantee of availability and high performance. As part of the service the CDN provider ensures that all content is synchronised and available from all of the servers.

CDNs are not a new concept: they have been around for more than ten years, but they have only recently become an affordable option for outreach and communication departments, such as that of ESO, which delivers around 0.3 petabytes of outreach internet traffic per year. Due to changes in the market and the arrival of new, cheaper and more flexible CDN providers, prices now start in the range of 50–100 euros per terabyte of data transferred, depending on how many data centres are offered, and their locations.

Originally, CDNs improved website performance by caching static site content in order to avoid intervening bottlenecks as much as possible. But the range of applications has now expanded, and CDNs are extremely important for transferring large files, such as high-resolution images and videos, and for live streaming of media. They are also used to serve normal web content — including JavaScript or CSS (Cascading Style Sheets) files and graphics — e-commerce applications and social networks.

CDNs do not only provide better performance and availability of the customer's content, they also offload the traffic served directly from the content provider's servers which can then effectively accommodate large-scale and short-term variations in demand for the content. This reduces the pressure on content providers to accurately predict capacity needs and enables them to gracefully absorb spikes in website demand — sometimes colloquially known as the Slashdot effect. Managing these spikes in demand can lead to a possible cost saving for the content provider, as sites no longer need to support a substantial local server infrastructure as a contingency against large sudden demands, that may sit underutilised except during popular events.

How hard is it to implement a CDN solution?

With limited time and manpower available to make information technology (IT) infrastructure changes in a scientific

File size	Small DSL connection (4 Mbit/s)	Mobile 3G (16 Mbit/s)	Large DSL connection (16 Mbit/s)	Mobile LTE (100 Mbit/s)	Fibre cable connection (100 Mbit/s)	Large institutional cable connection (1 Gbit/s)
Fulldome clip (10 Gigabytes)	5.5 hours	1.4 hours	1.4 hours	800 seconds	800 seconds	80 seconds
Fulldome movie(400 Gigabytes)	9.26 days	2.3 days	2.3 days	8.9 hours	8.9 hours	0.88 hours

Table 2. *Theoretical shortest download times of planetarium materials for different typical internet connections.*

Figure 3. *Location of CDN77 servers worldwide. Credit: CDN77*

	Local		European			Non-European			
	Munich, Germany	Augsburg, Germany	London, UK	Helsinki, Finland	Batalha, Portugal	ESO, Santiago, Chile	New Jersey, USA	Salt Lake City, USA	Cartagena, Colombia
Pre-CDN average through-put (Mbyte/s)	0.02	0.11	0.05	0.05	0.04	11.94	0.04	0.12	0.04
Post-CDN average through-put (Mbyte/s)	0.01	0.10	0.06	0.08	0.08	20.42	0.05	0.61	0.11
Improve-ment	-53%	-7%	39%	47%	73%	71%	30%	428%	188%

Table 3. *Summary of a simple test of download speeds from nine geographical locations before and after ESO's and ESA/Hubble's CDN installation.*

organisation it is important to ask how hard it will be to implement a CDN solution. In our experience it is fairly simple to make the changes needed. To get a simple system running it is only a matter of updating the links to static files on the web pages to point to the CDN system instead of the local server. The CDN system will then automatically fetch the files from the nearest server on demand.

However, especially if dealing with large files, the files should be uploaded to the CDN system in advance and then pre-loaded onto the many servers on the network. This is again simple to implement, but it is important to ensure that a tight workflow is followed when dealing with the data locally. Any files that change on the local master server also need to be re-synchronised with the CDN system to ensure that changes percolate through the network. It is much easier to implement the more advanced features of a CDN system if good data workflows are already in place, and followed.

ESO's partnership with CDN77

Since February 2015 ESO and ESA/Hubble have been using a CDN. As part of a collaboration, the CDN provider CDN77[8] has now placed all of ESO and ESA/Hubble's images and videos on 28 servers worldwide. They are distributed in 23 different countries and cover all continents besides Africa. This gives users — particularly those outside Europe — significantly faster access to content and allows them to play back videos live onscreen and download even the largest videos and images in a reasonable time. This is especially advantageous for web visitors from ESO's 16th Member State Brazil, which is currently undergoing the last stages of ratification and Chile, the location of ESO's telescopes.

As a sanity check on the usefulness of the CDN solution ESO performed a basic and independent test before and after the installation of the CDN network (Table 3). This test confirmed that the throughput of ESO and ESA/Hubble's images and videos outside Europe was improved with an average factor 179% (between 30% and 428%) using a variety of small, medium and large files. This is consistent with the values of up to 130–400% faster downloads quoted by Nygren et al. (2010). The ability to stream live video was also significantly improved as expected, but quantitative data were not measured.

It was not possible to control the local test conditions when collecting the results other than by normalising for the available internet speed at the time of testing[9]. Despite being a fairly crude test there were also indications that, as expected, there was a small drop in throughput within the local radius of the master server, in our case those users within 15 kilometres. This is because the CDN-routed traffic will need to go to the nearest CDN server, which is usually further away than the local server used. In our case the nearest CDN server is in Frankfurt, Germany, which is 300 kilometres as the crow flies. For the rest of the test locations in Europe a small improvement in download speed could be seen.

In conclusion it can be stated that a CDN system provides quicker and more convenient access to content worldwide for users who are far from the location of the master server. This was measured both for downloading files and streaming video, with a measured average improvement outside Europe of 179%, and up to more than 400%, depending on the end-user's location and the proximity of the nearest CDN server.

Acknowledgements

Many thanks to those working with us on this project from CDN77, including Veronica Minovska and Tomáš Sýkora. Thanks also go to our colleagues spread over large parts of the world for helping with the testing: Javier Enciso, Adam Hadhazy, Georgia Bladon, Max Rößner, Bruno Rino, Simo Pirinen, Francisco Rodriguez, Kevin Scott, Michael Naumann, Klaus Herzig and Daniel Schweizer.

Notes

1. Ultra High Definition is 3840 x 2160 pixels per frame.
2. High Definition is 1920 x 1080 pixels per frame.
3. More information on the ESO Supernova Planetarium & Visitor Centre: http://supernova.eso.org/
4. The most common fulldome standard format is 4096 x 4096 pixels per frame.
5. Free Download Manager is available at: http://www.freedownloadmanager.org/download.htm
6. Aria2 is a command line tool that downloads a file from multiple sources/protocols and tries to utilise your maximum download bandwidth to speed up your download. It is available here: http://aria2.sourceforge.net/
7. More information on content delivery networks: http://en.wikipedia.org/wiki/Content_delivery_network
8. More information on CDN77: http://www.cdn77.com/
9. Internet speeds at the time of testing were established using speedtest.net.

References

Nygren, E. et al. 2010, ACM SIGOPS Operating Systems Review archive, vol. 44, 3, 2–19, (ACM: New York, USA) http://www.akamai.com/dl/technical_publications/network_overview_osr.pdf

Biographies

Mathias Jäger is an astronomy communicator from Austria. He obtained a PhD in astronomy from the University of Heidelberg, then worked for the Haus der Astronomie before being an intern at the European Southern Observatory (ESO) in the education and Public Outreach Department (ePOD). Currently he is a science communication freelancer for organisations including ESO and ESA/Hubble.

Lars Lindberg Christensen is a science communication specialist and Head of the ESO education and Public Outreach Department (ePOD) in Munich, Germany. He leads public outreach and education for the La Silla Paranal Observatory, for ESO's part of ALMA and APEX, for the European Extremely Large Telescope, for ESA's part of the Hubble Space Telescope and for the IAU Press Office.

Mathias André obtained an MSc in Computer Science in England and worked for several years as a Unix system administrator and IT Operations Manager before joining the ESO outreach group to tackle new challenges as Web and Advanced Projects Coordinator.