

AARNet's Cloudstor+ Cloud storage initiative: Cloud-based value-added services

Alex Reid
AARNet

Summary: This paper describes the rationale behind AARNet's involvement in providing a range of value-added services over its Research and Education high-speed network. Special emphasis is given to services offered 'in the Cloud', and in particular concentrating on its latest service offering, Cloudstor+, which provides Cloud storage to researchers across Australia.

Introduction

AARNet, the Australian national higher education and research network organisation, has launched a number of initiatives for its members in the 'cloud computing' space. This paper is concerned primarily with a new service called "Cloudstor+", but it first sets out the context and rationale for providing this service.

AARNet is owned and operated by 38 Australian universities together with CSIRO. It was founded in 1989, and so is celebrating 25 years this year, its 20th anniversary being celebrated by publication of a book (Korporaal 2009). It provides high-speed and high-availability network connections between its members and the Internet at large. It has its own fibre on redundant paths across Australia and it leases dark fibre across the Pacific, which are now lit variously at 10Gbps, 40Gbps and 100Gbps; all its member organisations are connected at 10Gbps or better. While enhancing and upgrading its network (now moving to its fourth generation, implementing MPLS) remains AARNet's first priority, it has for many years now sought to offer value-added services layered over the network.

National Collaborative Research Infrastructure Strategy

AARNet works collaboratively with other Higher Education (HE) and research-sector agencies around Australia to implement the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS). This strategy, which has seen many hundreds of millions of dollars of government and sector investment between about 2005 and the present day, seeks to provide a comprehensive suite of capabilities to support research. It includes high-performance computing, data storage, tools, authentication and authorisation, and of course networks. AARNet is the operator of the network component of this suite, known as the Australian Research & Education Network (AREN). AARNet also works collaboratively with equivalent organisations overseas, generally called National Research & Education Networks

(NRENs), mainly so that seamless connectivity can be achieved between Australian and international researchers across the globe.

Within the context of these two sets of collaborators, AARNet looks to see which needs of its communities (researchers, academics, students, administrators) are not being (adequately) met either by suitable commercial offerings, or by other NCRIS providers. It also looks to see what services are being offered by its sister NREN organisations in other countries, and whether they would be equally applicable here.

AARNet Service Candidates

Services which become candidates for further consideration are those that leverage the network effectively, that are ‘natural fits’ within the suite of services already being offered, and for which widespread need can be observed across the AARNet client base. One of the key characteristics of candidate services is that they enhance the interactions and collaborations among the academic community. Normally, AARNet prefers to ‘bundle’ its value-add services, which is one reason it looks for widespread appeal.

This need for widespread appeal has been particularly noted by the eResearch support group within AARNet, which has recognised that much of its attention to date has been directed towards a small group of researchers with specialised, and usually demanding, needs. For instance, effort has been put into meeting the needs of radio-astronomers, who need dedicated 1Gbps and 10Gbps circuits across Australia and around the world for their Very Long-Base Interferometry (VLBI) projects ([VLBI 2013](#)). Likewise, particle physicists analysing data from the Large Hadron Collider at CERN in Switzerland, which has led to the discovery of the Higgs Boson, require tuned networks to achieve consistent >1Gbps downloads from CERN to Melbourne, with occasional peaks to as much as 9Gbps sustained over several hours ([LHC 2012](#)).

While not wishing to undervalue the needs of these researchers, nor the support effort they require, we have been keen to find ways of meeting the needs of the bulk of researchers, going beyond straightforward email, Web browsing and general Internet access, which of course play a large role in researchers’ activities these days, as they move to embrace the eResearch paradigm. The eResearch paradigm represents a fourth generation of scientific endeavour, starting with empirical, then theoretical, more recently computational, and now eResearch or eScience ([Hey, A et al 2009](#)), also called ‘data-intensive scientific discovery’.

The task of finding ways to move the bulk of researchers more fully into the eResearch paradigm has been identified as the ‘eResearch Chasm’ by Wolski & Richardson ([Wolski & Richardson 2010](#)). They base their analysis on the ‘technology adoption chasm’ of Geoffrey

Moore ([Moore 1991](#)), arguing that it applies equally to researchers. This is illustrated in Figure 1 below.

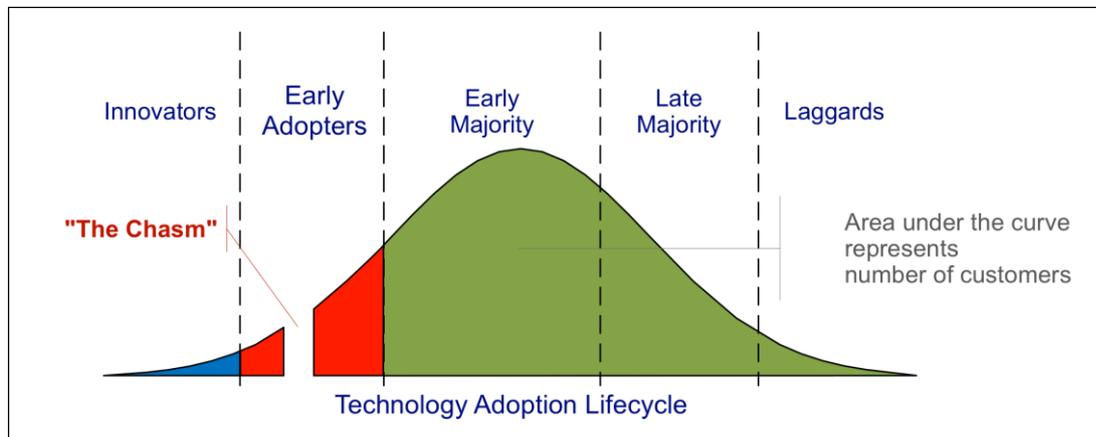


Figure 1: The Technology Chasm

They postulate that all but a few innovators and early adopters do not want the hassle of learning new technologies and having to adapt their ways of doing things, unless a simple, intuitive, reliable service is provided.

Cloudstor/FileSender

It was just this scenario that gave birth to the idea for Cloudstor+'s predecessor, Cloudstor, a file sending system (especially designed to overcome size limits on email attachments) developed jointly by AARNet and the NRENs of Ireland and Norway, and more recently the Netherlands. Cloudstor has an Australian user base of 11,500 users, and is also implemented by AARNet's partners in 26 other countries, where it is known as FileSender. It filled a gap in the market for tools and services not met by any other offering.

This need for simple systems that have broad appeal is also the reason why Dropbox has had widespread adoption, and also why researchers (and other members of the AARNet community) have taken so much to Skype rather than other more elaborate videoconferencing systems.

It should be noted that AARNet does not have a strong software development capability, so it normally aims to leverage products and services which require little adaptation and/or which it can acquire from other NRENs.

AARNet Cloud Services Catalogue

A very good example of a service which was acquired from overseas NRENs is *eduroam* ([eduroam 2012](#)). This is a WiFi service that enables a member of one participating institution to gain access, seamlessly, to the WiFi network of all other participating institutions, by

authenticating their connection via their home institution. It started in Europe over 10 years ago. To join, AARNet was required to install some equipment and undertake some configuring (it depends on Radius servers), but, with almost no adaptation, this allowed Australia to join a world-wide network of *eduroam*-enabled institutions. Most Australian universities and CSIRO have now deployed *eduroam*, typically employing existing equipment, usually across their entire WiFi network, as have many other nations. So a travelling academic will find that they can open up their laptop in any of hundreds if not thousands of universities worldwide and be instantly and transparently, and securely, connected to the Internet.

AARNet's strategies of collaborating with other NRENs and borrowing systems and services from them, and of seeking services that add value to the underlying network, have their current manifestation in a collection of services termed NET+ Cloud Services (NET+ 2014). This collection includes *eduroam* and Cloudstor, as well as new services which are brokered from third parties, such as Box.

AARNet adds the following benefits to the interaction between customers and Cloud Service Providers:

- Standardised contracting and service provisioning that can be re-used and shared across the sector;
- Services customised for the Research and Education sector;
- Reliable high performance networks, with large bandwidth and low latency;
- Private point to point and multipoint network connections;
- Integration with middleware such as federated authentication systems.

Where the service is provided by a third party, there will generally be a Partnering Agreement between AARNet and the Service Provider, which governs the manner in which the service is provided over the AARNet network. The Partnering Agreement also contains standard pricing and terms which flow through to the Customer Agreement entered into by the Customer and Service Provider.

Appropriateness of Cloud Services to a Network Provider

It is significant that in many discussions of cloud services, there is much talk of the provision of servers, the need for software to implement cloud services running on those servers, and the need for comfortable user interfaces and tools to make access simple. But there is rarely a discussion of the need for fast, secure, ubiquitous, reliable networks. This is because the network is now taken for granted; it has become largely invisible. It is, of course, entirely appropriate that this should be the case – that has been our goal as builders and operators of networks, and should be a matter of great pride that the network is now taken for granted. Of

course, this does not diminish the effort that is required to maintain and enhance the network. But it does mean that network providers should start to look at providing value-added services running over the network layer, as AARNet has done.

The existence of the network actually invites consideration of ‘cloud services’, which are defined as any computer-based services which are provided across the network, potentially at multiple locations, usually by third parties (‘virtualised third party providers’), and where the actual location and nature of the service provision is hidden from the end-user.

Rationale for Cloudstor+

When AARNet considered that there was a need for a service like Cloudstor, many of the factors discussed above came into focus. These included recognition of the need for a particular value-added service across a broad range of the client base; that it seemed to be a natural fit among other services; and that it enhanced the interactions and collaborations among the academic community.

In addition to these general prerequisites, we had been encouraged for some time by users of Cloudstor (which has been in operation for over four years) to consider adding a permanent storage capability to it (typically, Cloudstor files are held in ‘the cloud’ only for a short period, till they are picked up by the intended recipient). Users clearly saw a need for their files to be held in the cloud on a more permanent basis. However, there was already such a service provided to the academic research community, called Data Fabric, which had been developed and offered as part of the NCRIS services through an earlier government initiative, called the Australian Research Collaboration Service, or ARCS ([ARCS Data Fabric 2011](#)). It was proving to be quite popular with the community, with some 5,000 users registered. Of course, there was also Dropbox, and other commercial offerings, which were also much used within our community.

However, AARNet became aware that there were some misgivings about the future of the Data Fabric. In part this was because it was built on iRODS ([iRODS 2014](#)), but heavily modified, which was becoming a maintenance headache. This in turn led to an issue with meeting its running costs. The funding provided through the NCRIS initiative was intended primarily for construction and development, not for ongoing operating costs. When ARCS reached the end of its contract in 2011, and was not continued, the responsibility for operating some of its services, including Data Fabric, were taken over by the State-based eResearch support organisations, whose continuity was rather more certain. Nevertheless, the ongoing burden of supporting the Data Fabric proved too onerous, and the decision was made mid-way through 2012 to discontinue that service.

In the light of these developments, and against the backdrop of encouragement by Cloudstor users, AARNet had been quietly investigating possible alternatives. When the announcement was made that Data Fabric would be discontinued, with no indication of any possible replacement, AARNet ramped up its effort to fill this gap. As indicated above, the new service would need to be a natural fit with its other value-add services, and a natural complement to Cloudstor. AARNet announced its intention to launch the new service, to be called Cloudstor+, in October 2012. As it happens, at the same time, and unbeknown to us, RDSI had been looking at a possible replacement, known as share.edu ([share.edu 2013](#)). RDSI (Research Data Storage Infrastructure) is the agency funded under NCRIS to establish substantial storage facilities for significant data collections.

Development and trial by a collection of beta testers during the course of 2013 resulted in a formal release of Cloudstor+ on 30 September 2013. By now it has a 1000-strong user base.

Cloudstor+ Capabilities

The primary purpose of Cloudstor+ is to provide data storage ‘in the cloud’ for researchers as a means to augment and manage their local storage facilities and to encourage easy data sharing. Researchers at Australian universities and the CSIRO (members of the AARNet community) are each given a free allocation of 100GB of storage, and can acquire more at an approximate cost of \$30/TB per month. Users can access this storage in a variety of ways: a Web address enables them to log in to their storage area, where they can upload or download files, make them accessible to others and generally manage their files through simple Web interfaces.

Users can also access their files from mobile devices (clients are available for Apple and Android smartphones and tablets), and also arrange for automatic synchronisation of files on their various devices (desktops, laptops, mobile devices). Data is replicated across data stores in several locations across Australia, which are directly connected to the AARNet backbone at 10Gbps. Data is stored and accessed from the nearest physical storage node, ensuring extremely fast uploads and downloads, and is replicated in the background to provide high reliability and accessibility. Preliminary tests indicate data transfer rates several orders of magnitude greater those experienced for commercial offerings: this is one good reason why AARNet considers offering such services. Data are all stored in Australia, so there are no issues of data sovereignty.

Another key bonus is that authentication and authorisation are provided via the Australian Access Federation ([AAF 2014](#)), which means that researchers use their credentials from their home institution. Although three copies of the data are retained at any one time, data is not backed up in a formal sense. The system draws on AARNet’s expertise and capability in high-reliability service provision, aimed at ensuring the highest levels of reliability and availability.

It builds on the infrastructure used for AARNet's Mirror and CloudStor services, which have very high reliability and availability – AARNet monitors all its systems on a 24x7 basis. Because it exploits existing infrastructure, it has been inexpensive and quick to implement; it can be expanded rapidly in response to demand; and it also enables the use of a range of storage types.

The system is based on the community-developed open source ownCloud software on the user-facing interface ([ownCloud 2014](#)). This gives AARNet considerable control over any interfaces and add-in modules it may wish to adapt, as well as being able to draw on the resources of a large world-wide development community. It also avoids any undesirable commercial lock-in. At the same time, AARNet and its other NREN collaborators have found the support provided by ownCloud to be very responsive. Several other NRENs are very keen to deploy the Cloudstor+ architecture themselves and to assist in further development. This community gives great comfort to an organisation like AARNet which, as indicated above, has not been set up to develop software and systems itself.

Conclusions

While AARNet is primarily a network service provider, it recognises the importance of providing a range of services that run over the network, especially where the services meet a number of important criteria. These include services that leverage the network effectively, that are 'natural fits' within the suite of services already being offered, and for which widespread need can be observed across the AARNet client base. AARNet does not have a significant software development capability, so it looks to exploit its relationships with overseas NRENs and to draw on community development wherever possible. This has been the case with Cloudstor+.

It is recognized that researchers, based on their past experiences, may well be understandably reluctant to commit their valuable data to yet another sector cloud storage service provider. For this reason, AARNet has stressed that it is committed to sustaining the Cloudstor+ service indefinitely, and can point to its 25-year history of reliable and trustworthy service to the sector. AARNet is committed to the provision of value-added services running over its primary service, the network, for the long haul, and particularly to a range of Cloud services as being especially pertinent to a Network provider. It remains to be seen if this latest service is as successful in meeting the needs of researchers as Cloudstor has been to date.

References

- ARCS Data Fabric. 2011. <http://ands.org.au/guides/storage.html> (background information only).
- Australian Access Federation (AAF). 2014. <http://aaf.edu.au/>.
- Cloudstor. 2013. <http://www.aarnet.edu.au/services/netplus/cloudstor>.

- Cloudstor+. 2014. <http://www.aarnet.edu.au/services/netplus/cloudstorplus>.
- eduroam. 2012. <https://www.eduroam.org/> (2012) and <http://www.eduroam.edu.au/> (2010).
- Hey, A et al. 2009. The Fourth Paradigm: Data-Intensive Scientific Discovery, Microsoft Research, 2009.
- iRODS. 2014. Integrated Rule-Oriented Data System, <http://irods.org/>.
- Korporaal, Glenda. 2009. AARNet: 20 Years of the Internet in Australia, 1989-2009.
- LHC. 2012. <https://www.aarnet.edu.au/CaseStudy/Probable-Higgs-Boson-Discovery.aspx>, 4-Jul-12.
- Moore, Geoffrey A. 1991. Crossing The Chasm. New York: HarperBooks. Figure 1 is available, under Creative Commons Attribution 3.0 Unported License, at <http://commons.wikimedia.org/wiki/File:Technology-Adoption-Lifecycle.png>.
- NCRIS. 2013. <http://education.gov.au/national-collaborative-research-infrastructure-strategy-ncris>.
- NET+. 2014. <http://www.aarnet.edu.au/services/netplus>.
- ownCloud. 2014. <http://en.wikipedia.org/wiki/OwnCloud>.
- RDSI. 2013. <https://www.rdsi.edu.au/>.
- share.edu. 2013. <http://conference.eresearch.edu.au/eres2013/news-2013/february-2013/>.
- VLBI. 2013. <https://www.aarnet.edu.au/CaseStudy/Radio-Astronomy-Exploring-the-Universe.aspx>.
- Wolski, Malcolm & Richardson, Joanna. 2010. Moving Researchers Across the eResearch Chasm, Ariadne, issue 65, 29-Oct-10 <http://www.ariadne.ac.uk/print/issue65/wolski-richardson>.

Cite this article as:

Alex Reid. 2014. *AARNet's Cloudstor+ Cloud storage initiative*. Australian Journal of Telecommunications and the Digital Economy, Vol 2, No 2, Article 35. <http://doi.org/10.7790/ajtde.v2n2.35>. Published by Telecommunications Association Inc. ABN 34 732 327 053. <http://telsoc.org>