

About Storage Services Gateway

Since 2007, Penn State University has been researching ways of addressing the challenges associated with storage support for the tremendous growth of information and data within Penn State University Libraries and the university as a whole. Simple processes used in the past, such as copying data sets, finding a specific subset of data, or replacing storage arrays, are becoming extremely complex operations that are too large and time consuming to accomplish within acceptable levels of service. Software solutions built to manage large repositories of data also require application-specific knowledge, long-term migration, and integration challenges that often result in large and unwieldy silos of data.

To begin to address these challenges the Storage Networking Industry Association (SNIA) is endorsing the use of storage tiers and classes, data life-cycle strategies and new standards for managing structured data, such as the eXtensible Access Method (XAM). As Penn State began researching the XAM standard, implementation hurdles became apparent; hurdles such as integration with existing applications, working across heterogeneous data silos, and hitherto lack of application of metadata standards in the storage domain. In order to assess XAM's potential, Penn State initiated an Archival Storage Prototype project which entailed development of a Storage Services Gateway that resides between applications and storage. The prototype entailed the development of a Storage Services Gateway that consists of a policy engine and a metadata engine, and serves as a mechanism to access our storage pool via standards-based protocols including AJAX, SOAP, REST, HTTP, and FTP.

Our project team included metadata experts who led the definition of technical and administrative metadata used to search federated content, create retention policies, and route objects to tiered storage. Additionally, we wanted to demonstrate how the use of metadata is instrumental in the lifecycle management of data saved into our storage cloud. Another key component of the Archival Storage Prototype project was to test and demonstrate XAM capabilities for enhanced storage management. The project team focused on our tier 2 (fixed content) storage service using XAM as the framework to access private cloud storage targets to save, manage, and retrieve fixed content objects and associated metadata. Fundamental to the Storage Services Gateway, was the development of a REST API that is capable of performing XAM translations when utilizing our storage cloud. REST is a well-established software architecture for many web developers and is common in software repositories, therefore enhancing adoption of the XAM standard with little change to existing code or skill sets.

We've also integrated our Storage Services Gateway with Fedora 3.2.1 to demonstrate the flexibility of our architecture to incorporate existing capabilities and solutions in a diverse storage environment. We utilized Fedora's native support of the http protocol via REST APIs to access external control groups and reference content from within our Storage Services Cloud. In addition to pulling the actual saved objects into Fedora for management and viewing, we were also able to list XAM generated metadata within Fedora such as the XAM assigned unique ID (XUID), and ingest times and dates.