

# Content Platforms Meet Data Storage, Retrieval Needs

## NASA Technology

Earth is under a constant barrage of information from space. Whether from satellites orbiting our planet, spacecraft circling Mars, or probes streaking toward the far reaches of the Solar System, NASA collects massive amounts of data from its spacefaring missions each day. NASA's Earth Observing System (EOS) satellites, for example, provide daily imagery and measurements of Earth's atmosphere, oceans, vegeta-

tion, and more. The Earth Observing System Data and Information System (EOSDIS) collects all of that science data and processes, archives, and distributes it to researchers around the globe; EOSDIS recently reached a total archive volume of 4.5 petabytes. Try to store that amount of information in your standard, four-drawer file cabinet, and you would need 90 million to get the job done.

To manage the flood of information, NASA has explored technologies to efficiently collect, archive, and provide access to EOS data for scientists today and for years to come. One such technology is now providing similar capabilities to businesses and organizations worldwide.

## Partnership

In 2004, Archivas Inc. of Waltham, Massachusetts partnered with NASA's Goddard Space Flight Center through the Small Business Innovation Research (SBIR) program. Founded by the former chief technology officer of *The New York Times*, who was seeking an effective means of digitally archiving the paper's 100-plus years of news, Archivas innovated new software technologies for preserving and providing useful access to vast digital repositories of data. The company began work with Goddard computer scientist Curt Tilmes to develop and test a beta form of its ArC technology for NASA to collect and store data from the Ozone Monitoring Instrument (OMI) onboard NASA's Aura EOS satellite.

This image from the Aura satellite's OMI instrument reveals tropospheric ozone over Indonesia. Aura monitors air pollution around the world on a daily basis.

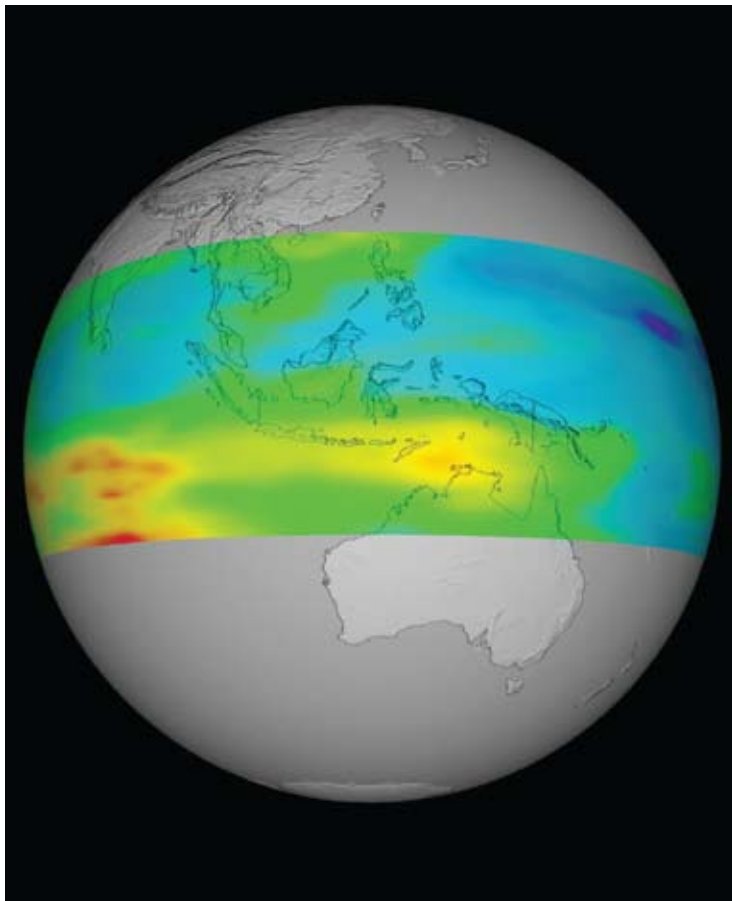
*“Working with NASA was instrumental in getting us the credibility we needed ... It was crucial to our early success.”*  
—Asim Zaheer, Hitachi Data Systems

Traditional methods of data storage at the time, such as tapes, had long proven slow to provide access to information and costly to maintain and scale up as an archive grew. NASA needed a solution for handling the large OMI data files and allowing them to be processed by different applications using different protocols. The result of the partnership with Archivas was “a single, consolidated repository of digital assets—in this case satellite images—that was accessible by multiple different applications used to process and store the images,” says Asim Zaheer, at the time vice president of marketing for the company.

The repository was capable of scaling up to extremely large sizes, a necessity considering the significant size of the individual satellite files. In addition, the technology enabled the quick retrieval of archived information, even years after collection, through a unique method of storing data as objects rather than files. Objects combine files with information about the file (its metadata) and a policy that provides some kind of instruction about how the information should be handled—whether it should be replicated or deleted after a certain number of years, for example.

The SBIR-derived technology became a long-term solution for NASA's OMI data collection and other Earth science missions. Archivas, in the meantime, translated its NASA work into a springboard for commercialization.

Archivas brought the ArC software to market in 2005 and found success across a range of industries. Hospitals needed to create repositories for medical imagery and patient records, and financial services firms required a solution for the long-term retention and preservation of



authenticity of financial records. Both industries have regulatory requirements that mandate record retention and the ability to recall these records as needed, and the Archivas software adapted to meet those needs.

“We evolved the technology from being simply a long-term record retention archival solution to a digital repository that can provide access to digital content anytime, anyplace, anywhere,” says Zaheer.

Recognizing the strength of the technology’s content management capabilities, Hitachi Data Systems Corporation, headquartered in Santa Clara, California, acquired Archivas in 2007.

## Benefits

Today, the technology Archivas advanced with NASA assistance is marketed as the Hitachi Content Platform, or HCP. Capable of scaling up to 40 petabytes in capacity in a single cluster, HCP allows users to securely store and preserve data for business, legal, compliance, and other purposes without the need for tape-based backup. HCP can be subdivided into separate tenants that can be uniquely configured with various data management policies and access rules; these tenants can be further divided into namespaces that can also be individually configured. The technology, which earned the 2009 “Information Management Innovation Award” from *Information Age* magazine, is adaptable to new data formats and applications, meaning users can easily maintain their repositories even as the information technology environment changes and evolves.

HCP provides secure content management for customers including Peak Web Consulting, Qualcomm, and Comdata. Payformance Corporation, a healthcare claim settlement solution provider, achieved an 80-percent increase in administrative efficiencies thanks to HCP technology.

Beyond data storage, HCP has proven to be an ideal technology for cloud computing applications—the use of computational resources based exclusively on a network rather than housed on a specific computer. HCP is now the cornerstone of Hitachi Data System’s cloud storage solutions. The technology also helps enable Hitachi Clinical Repository, a new information management

solution that offers healthcare providers a consolidated view of patient information, helping improve clinical decision making and patient care.

“In the retail sector, heavy manufacturing, technology, telecommunications, your cell phone provider, health care—you name it. This technology is broadly, horizontally leveraged, and it’s global,” says Zaheer, now Hitachi Data Systems’ vice president of corporate and product marketing. He attributes HCP’s commercial growth to the success of the early partnership with NASA.

“There are young startups developing technology left and right, and everyone feels they have the world’s best widget, but a lot of target customers don’t know how credible that story really is,” says Zaheer. “Working with NASA was instrumental in getting us the credibility we needed to engage with other organizations. It was crucial to our early success.”

Now HCP stands to play an increasingly significant role in the field of information management as data proliferates across industries in ever greater amounts. Google currently processes over 20 petabytes of information per day. Even individuals are producing large quantities of content, such as music, photos, and video, Zaheer says.

“All of that content has to live somewhere.” ❖

Hitachi Data Systems® is a registered trademark of Hitachi Ltd.

Hitachi Data System’s NASA-derived Hitachi Content Platform is an ideal tool for cloud computing applications that base computational resources on a network accessible by any number of devices.

