JUN
**20**
2023

**CLOUD**

# How Federal Agencies' Hybrid Ecosystems Have Transformed Fieldwork

The shift from on-premises resources to a dual-sided data structure has provided benefits for NSF, NIH, Census and others.

-------------------------------------------------------------------

**by Erin Brereton** 🐦

Erin Brereton has written about technology, business and other topics for more than 50 magazines, newspapers and online publications.

-------------------------------------------------------------------

⏸ **PAUSE**   08:24

Instead of housing resources solely in an onsite data center, many federal agencies now provide access to information and applications

## Latest Articles

**Samsung Creates Secure Digital Signage Solutions For Agencies**

**Data Center Solutions Rise To New Challenges**

**FEMA Drives Better Decision-Making For Disaster Planning And Relief With Interactive Dashboard**

**Agencies Moving To Digital Recordkeeping Shouldn't Neglect Backup And Recovery**

notable benefits, including enhanced fieldwork capabilities that allow employees to engage with and submit information from a variety of locations.

"There are a lot of federally funded resources that span the spectrum from very high-performance, large-scale computer systems to very specific, AI-centric systems that the National Science Foundation has been funding for a long time — and other agencies have been, as well," says Manish Parashar, director of NSF's Office of Advanced Cyberinfrastructure.

"These can be knitted together to provide a cloudlike environment, and they can be complemented with commercial cloud services that provide a computational ecosystem that can support research," he adds.

## Learn a unified approach to hybrid cloud:

*Related Content:*

*Learn how (and why) to establish a cloud center of excellence.*

*Discover why agencies should pay for servers like they do for cloud.*

*Explore how federal agencies can leverage the cloud for application modernization.*

Agencies often work with more than one cloud service provider, Parashar says, to take advantage of the various capabilities each offers.

"Tensor processing units, which are good for certain type of applications, are unique to Google Cloud, as an example," he says. "That's part of the reason why, depending on what research you're trying to do, you need to have access to a range of resources. It

**HARDWARE**

## Samsung Creates Secure Digital Signage Solutions for Agencies

**DATA CENTER**

## Data Center Solutions Rise to New Challenges

## How Agencies Can Seamlessly Manage Data

Agencies may be able to use a commercial cloud- and on-premises architecture to internally streamline data management.

Before the National Institutes of Health began disseminating data through cloud services, potential research collaborators in separate locations shared information via FTP or as an email attachment.

If the file was too large to email or transfer over the internet, some researchers shared data on thumb drives or CDs, which resulted in a lot of data duplication, says Nick Weber, program manager of NIH's Science and Technology Research Infrastructure for Discovery, Experimentation and Sustainability (STRIDES) initiative.

*EXPLORE: Here are three steps to optimize your hybrid cloud environment.*

Rather than having collaborators shuffle items between research centers — or submit NIH-affiliated data that relates to more than one discipline to several applicable repositories — contributors can now send the information to a general repository, allowing other researchers to access it through commercial cloud services, along with the computational tools the providers offer.

"Cloud has really flipped data sharing on its head," Weber says. "Major data sets are located in cloud environments to allow people to use the data and collaborate with others there.

"That was a major driver for the STRIDES initiative and our partnership with Amazon, Google and Microsoft — to be able to say, how can we make this even simpler for researchers? How can we bring some additional ways to use the technologies the cloud offers to accelerate their research?"

Prior to implementing the Open Data Dissemination Program, the National Oceanic and Atmospheric Administration used a number of segmented paths to distribute information, according to CTO Frank Indiviglio.

## Trending Now

**OMB Outlines Process For Federal Workers To Begin Returning To The Office**

**3 Key Considerations For Achieving The DOD's Zero-Trust Framework**

**DOD's Research Partnerships With Universities Prove Mutually Beneficial**

**Ransomware Attack Hits Federal Agencies**

you're trying to do, you need to have access to a range of resources. It doesn't make sense to go to only one vendor — or even only to cloud services.''

**Manish Parashar, Director, NSF Office of Advanced Cyberinfrastructure**

Certain climate-related metrics and computations, for instance, were published to the Climate Change Portal web interface by NOAA's Physical Sciences Laboratory. Weather forecasting models were available in another location.

Now, through an online NOAA repository list and commercial cloud offerings from Microsoft, Google and Amazon, the general public, academic institutions and other entities can access data culled from satellites, ships and other sources. These range from observational system readings to information that needs to be processed before being doled out, such as model data generated on a supercomputer.

''On the modeling side, there's more opportunity to share that data,'' Indiviglio says. ''We're talking about pretty big data sets; it's not a straightforward thing to just make them available. You have to package them up and get them out to folks.

''The value is, we can present it in a very consumable way, and people can get it without having to go through the hoops of 'I have to have the right network connection,' or 'I have to have this type of infrastructure.'''

## Successfully Scaling for a Hybrid System

Agencies may not have to proactively add many physical components to adopt a hybrid system; essentially, Parashar says, users need a basic access point, such as a laptop with network connectivity, to enter a portal that leads to data.

For instance, the Census Bureau implemented a hybrid cloud environment involving multiple vendors in preparation for the 2020 census, according to Barbara LoPresti, chief of the Decennial Information Technology Division. Information collected in the 2010 count was kept on-premises.

and personnel systems. The cloud services environment proved helpful, LoPresti says, as the agency's work related to data production escalated in the most recent census effort.

For the first time, the bureau's 300,000 field enumerators collected information door to door using Apple iPhones, transmitting it to operational control systems hosted by Amazon Web Services via AT&T (or, if needed in remote areas, through a local connectivity provider). The iPhones were returned to provider CDW·G when the census was complete.

*Click the banner below to learn about the benefits of hybrid cloud environments.*
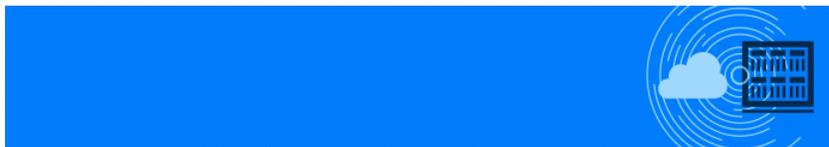
''We did not have to make a large capital outlay for computer equipment, servers and storage,'' LoPresti says. ''We could just expand our resources automatically by scaling in the cloud to handle those massive volumes, so it was a good fit for us. We moved from a large capital investment to more of an operating model.''

Based on the results and benefits — such as the flexibility the hybrid approach afforded the agency when it paused the count because of the COVID-19 pandemic — the Census Bureau's goal for the 2030 census is to put as much data as possible in the cloud, says LoPresti.

*READ MORE: What your agency should consider after selecting a hybrid cloud model.*

''In 2020, we used AWS GovCloud for our data collection applications,'' she says. ''The internet self-response application was running to collect data. That was very successful; we did not experience any downtime whatsoever with that application. If we had set up our own independent on-premises data center for that purpose, it would have been a lot more money and a lot more involved.''

Agencies may also achieve cost savings by negotiating an end-user discount with commercial cloud providers.

for researchers to use in studies.

To tap into patient and other associated data sets and computational items, an academic research institution or medical teaching hospital would only need to enroll in STRIDES, Weber says.

Agencies can also build parameters into cloud-based access to make findings available responsibly to a wider range of end users, potentially including citizen scientists or All of Us participants who want to explore patient data from the study, Weber says.

"That's brokered through the research portal in that program — to be able to say, 'Here are the rules in which individual users or research institutions can get access to this data, and here are some tools they can use to compute it,'" he says.

"Many research programs will stand up a central data resource. It may be open public-research data or controlled-access data from a participant who, at the time of the study, said, 'My data can be used in this way, but not that way.' We need to maintain that integrity."

# 55%

The percentage of government and public sector organizations that say they're proactively working to advance their cloud strategy

Source: KPMG, "2022 KPMG U.S. Technology Survey Report: Government and public sector industry insights," November 2022

## ENSURING ACCESSIBLE YET SECURE CODING

In 2016, the White House introduced the Federal Source Code Policy, which included a pilot program that urged federal agencies to share at least 20 percent of any new code they commissioned — noting that open-source software dissemination could help facilitate peer review and security testing.

In the years since, agencies such as the National Science Foundation have touted a similar approach to research-related data.

researchers who contribute to the open-source ecosystem. It builds a more vibrant research ecosystem."

Opening data to a larger audience via a commercial cloud service — which might help researchers obtain additional information for projects — could also facilitate agencies' open-source development efforts by serving as a type of crowd-sourced proofreading, according to Parashar.

"It gives us transparency into where the issues might exist with the technology," he says. "More people can look into it and see where there could be problems."

That's one reason the National Institutes of Health makes data available through commercial cloud service providers, according to Weber.

"If there's ever any question as to the conclusions that were drawn from the research, somebody else can use that data and try to validate or invalidate those conclusions, as opposed to just assuming that what people did was right and the conclusions we drew from that are actionable," Weber says.

AUGUST ALLEN/POLAR FIELD SERVICES

## More On    CLOUD COMPUTING   HYBRID CLOUDS

PRIVATE CLOUDS   PUBLIC CLOUDS

## Related Articles