FINDING BIG FISH IN by Jenny Mangelsdorf BIG DJATA

Challenge:

- Capture and process satellite-generated data reliably
- Archive hundreds of millions of environmental observations
- Process increasing requests for environmental and climate data

Solution:

- Unique mirroring architecture that ensures big data reliability
- Infrastructure and network capabilities that support NOAA researchers and more than 500 external users
- Big data systems engineering, integration and testing, and operations

Results:

- Creation of the premier online facility for the distribution of NOAA and U.S. Department of Defense environmental satellite data
- Efficient management of a dynamic system that grows more than five terabytes a day
- Accessible global weather and climate data

Off Florida's Gulf Coast, scientists are diving into a sea of data. They're tracking tiny Atlantic bluefin tuna larvae, and their parents, to better protect and monitor the environmentally delicate and highly migratory fish.

To help locate the fish, which can grow to more than 1,500 pounds, researchers turn to Mitchell A. Roffer, president of Roffer's Ocean Fishing Forecasting Service. He taps the systems that capture the world's climate and weather data, including the U.S. National Oceanic and Atmospheric Administration's Comprehensive Large Array-data Stewardship System, or CLASS.

With CLASS, which has ocean and weather data ranging from temperatures and currents to sea color, Roffer can better forecast where bluefin are spawning, oil spills are spreading and ships should be moving.

"CLASS is a valuable system, providing data that helps us better understand the effects of environmental conditions," Roffer says. As a forecaster and researcher supporting NASA, NOAA, fishing, and oil-and-gas industry research projects, Roffer has accessed this data for decades.

Storing a world of data

CSC works with NOAA's data, too. First supporting the predecessor to CLASS, CSC helped develop CLASS and now helps ensure that the system, which acts as an extremely big, interactive electronic library, can process and store satellite data and respond to requests from users such as Roffer.

Some users, especially those focused on climate research, want to dig into the past. The CLASS team continually works to ensure that the system's archive, which includes hundreds of millions of environmental observations dating back to the mid-1970s, is available and will last into the next century.

"Just as today's scientists and researchers revisit historical weather data and run new models, future generations will want to run models on this era's weather, a great deal of which is on CLASS," says Kern Witcher, NOAA CLASS project manager.

As the CLASS team members work to ensure the data's longevity, they also strive to improve the system's performance. "A big challenge is throughput," Witcher says. "We're getting to a point where in the next decade we'll be ingesting 35 terabytes of data a day."



CLASS receives data generated by numerous geostationary an polar-orbiting satellites. One in particular, launched last Octobe required NOAA to supercharge the ability of CLASS to manage the increased amounts of data it would be receiving.

Now circling the Earth every 102 minutes, the Suomi National Polar-orbiting Partnership (NPP) satellite collects data at a resolution of 25 meters per pixel, providing a fourfold improveme in quality over older instruments. With those instruments, CLA was capturing about 300 gigabytes of data a day. When Suom NPP went live, CLASS added five terabytes to its daily diet.

"That first day was like a rollercoaster — it started very slowly and then went extremely fast," recalls Constantino Cremidis, CS CLASS project manager. "At the beginning, it was quite a ride."

Protecting data integrity

Besides helping CLASS store and manage its data, another goal is making the CLASS data easily accessible to the public. "Of course, when a user wants months of data and there are multi-terabytes of data a day, this is a difficult challenge to meet," Cremidis says.

About 500 users and organizations around the world tap CLASS for climate science and research, or for data for the financial, agricultural and oceanographic industries.

nd er, Ie	The requests are diverse, the data demands are substantial, and the data is irreplaceable. To meet these challenges, CSC helped NOAA create an innovative architecture that would better pro- tect and manage the CLASS data. As part of the architecture, they built two mirroring systems, located in two different U.S. states that simultaneously receive and store CLASS data.
ent SS ni	"Everyone copies their data, but not at this level," Cremidis says. "We keep the two systems online and available at all times."
SC	"The system has become almost like a living organism," Witcher adds. "We have so many things happening with it. It's always evolving, moving, shifting and dynamically changing."
	Change for every part of the CLASS system seems a constant. CSC supported the system's predecessor at a time when few used the Internet. Today the CLASS team is exploring how it can improve its use of social media to better meet user needs, while preparing for the next jump in data. ■

JENNY MANGELSDORF is a writer for CSC's digital marketing team.