

Taming big data intel



BY JOHN EDWARDS

Sophisticated big data reporting, analysis, visualization, integration and development tools are essential for turning ever-growing mountains of intelligence data into useful, collaborative information that can be efficiently distributed to parties across the military and intelligence communities.

Both communities aim to understand how and when to use the information culled from big data. “Big data means different things to different people, but most commonly refers to the volume, velocity and variety of data,” Mike Friedel, manager for defense at SAS Federal, said. “It is going to take time to learn how to leverage big data. There is a great deal of research and development that is tackling this very thing.”

The military and intelligence communities have adopted big data as a strategic technical solution to address cyber insights, said Air Force Brig Gen Brian Dravis, chief of the Defense Information Systems Agency (DISA) Joint Information Environment Technical Synchronization Office. “Big data is an ecosystem consisting of data collection, storage and computation architecture and analytics,” he said.

Military organizations must use analytics to avoid getting buried under the massive amount of information they generate, said Greg Gardner, chief architect for government and defense solutions at NetApp. “As intelligence, surveillance and reconnaissance data is gathered and integrated, for example, it must also be both mapped out and separated from insignificant or unnecessary ‘noise’ that is irrelevant to decision-making.”

Big data insights are now essential to the conduct of modern warfare, and DISA is working to provide big data capabilities to its mission partners. “Data normalization and standard application programming interfaces support sharing,” Dravis said. He noted there are now several defense and intelligence working groups focused on normalizing data specifications and content dictionaries for cyber operations and defense. “Additionally, the [Department of



Dravis

Defense] has accepted the Ozone Widget Framework as a common application and presentation layer to quickly share analytic applications between communities,” Dravis said.

SEARCHING FOR TECHNOLOGIES

As the volume of data grows and data sources expand, it becomes increasingly important to look ahead and determine which technologies — out of the many available options — are going to be most useful. “Storing data in information and technology silos and data warehouses is not sufficient,” Friedel said. “Both the military and the intelligence communities really need to be able to manage vast and varied data in the most effective and efficient ways possible.”

Cloud architecture and the open-source Apache Hadoop platform allow the distribution of large data sets within virtually any environment, making big data easier to manage. Such technologies, along with analytical and visualization tools, will continue to have a major impact on military and intelligence community adopters, Friedel said. “High-performance analytics and visualization technologies are essential to analyzing and understanding quickly what the data’s telling them,” he said.

The military and intelligence communities are working jointly to build a platform capable of handling big data information and analytics. “This provides a common environment that enables any organization, cyber mission team, cyber protection team, or anyone with a situational awareness need, to develop analytics on common data ... for respective mission needs,” Dravis said.

Finding and deploying the best big data technologies means tapping into both open-source and corporate resources. “Today, we are leveraging open-source capabilities in support of cyber defense, but recognize that the future will require us to leverage industry to ensure rapid deployment of analytics in a common framework, ensuring the war fighter’s ability to operate in all domains,” Dravis said.



RICK NAYSTATT/NAVY

Intelligence hidden deep inside servers, such as these systems located at the Navy Enterprise Data Center in San Diego, can be discovered and made useable and shareable by sophisticated big data tools.

Commercial solutions have leapt forward over the last two years in terms of capability and extensibility, Dravis said. “We would be remiss if we do not compare current government solutions against commercial counterparts to deliver the most optimized and capable solution to satisfy performance and functionality requirements,” he said. “We must maintain agility to adjust to operational requirements and evolving technology.”

DISA plans to investigate technologies in three basic areas: interoperable solutions, analytics-as-a-service and out-of-the-box analytics. “Our goal would be to ensure we do not end up with stovepiped analytic solutions, because that would put us back into the same position before we had big data,” Dravis said.

Several big data solutions, as well as pilot projects focused on specific data analytical tasks, have already been deployed. “Big data solutions promise to discover cyber ‘insights’ lurking among petabytes or exabytes of data generated annually by cyber systems,” Dravis said. “Efforts like [Joint Information Environment], which plan to improve DoD operational efficiency and effectiveness by streamlining cyber command-and-control under fewer management domains, increases the need for securable and scalable big data solutions.”

Big data technology is evolving rapidly, so efforts to reap the approach’s benefits must be both adaptive and aggressive.

“Currently, Map/Reduce [M/R] systems are proving to be the most useful technology,” said Dravis, who noted that the most widely adopted and piloted M/R in military and intelligence networks is based on the Hadoop platform. “Hadoop is attractive because of its wide adoption among commercial solution vendors and its extensible platform, which allows the government to keep research and development costs lower as we research new cyber analytics.”

The intelligence community has an advantage in evaluating and piloting big data analytics, since it is “so closed off and secure” that it can experiment on the best ways of using analytic technologies with relatively little risk of inadvertent data exposure, Friedel said. “As for the military, the technology must be proven and mature so

that it can stand up to the rigors and discipline of combat,” he said.

FACING AND ADDRESSING CHALLENGES

Governance is the most significant big data challenge DISA currently faces, according to Dravis. By its very nature, big data requires data sources fed in from multiple organizations, which means they must define data-sharing legal, policy, oversight and compliance standards. “Governance, therefore, must account for equities and concerns from a broad community. The challenges are not insurmountable, but require coordination,” Dravis said.

Finding people who are qualified to deploy, configure and manage big data systems is another challenge confronting military and intelligence leaders. Analysts capable of interpreting the meaning and significance of big data intelligence are also in short supply.

“As helpful as technology is in sorting through PC-PAD [planning and direction, collection, processing and exploitation, analysis and production, and dissemination], making sense of that data requires a major investment in the training and experience of both analysts and leaders,” Gardner said. “Commanders must build the right mix of people to help make sense of the various data streams that come into their organizations on a mission-by-mission basis, particularly in the ISR arena.”

The military is already hunting for big data experts. “We can’t speak for the intelligence community; however, DoD cyber workforce initiatives are well underway,” Dravis said. “The demand for capability has fostered partnerships where a lot of skilled minds come together to provide joint solutions.”

People are the key to big data success. “We must recruit, educate, and train our workforce to think with a big data mentality,” Dravis said. Yet finding big data experts won’t be easy, particularly since the military and intelligence communities must compete against an array of government and commercial organizations for scarce big data talent.

Unfortunately, the big data expert shortage is likely to continue for many more years. “The U.S. lags behind in technology, engineering and mathematics compared to other countries,” Friedel said.

Despite the current challenges, military and intelligence community leaders can’t afford to neglect big data planning. “Any and all organizations should be looking at and deploying sandboxes now; they should start experimenting in order to get ahead of the current situation,” Friedel said. “It’s not advisable to abandon legacy systems that have been in place 20 or 30 years, but it is time to experiment with figuring out if these systems can be utilized in the mix of big data.” □