HYPERION RESEARCH

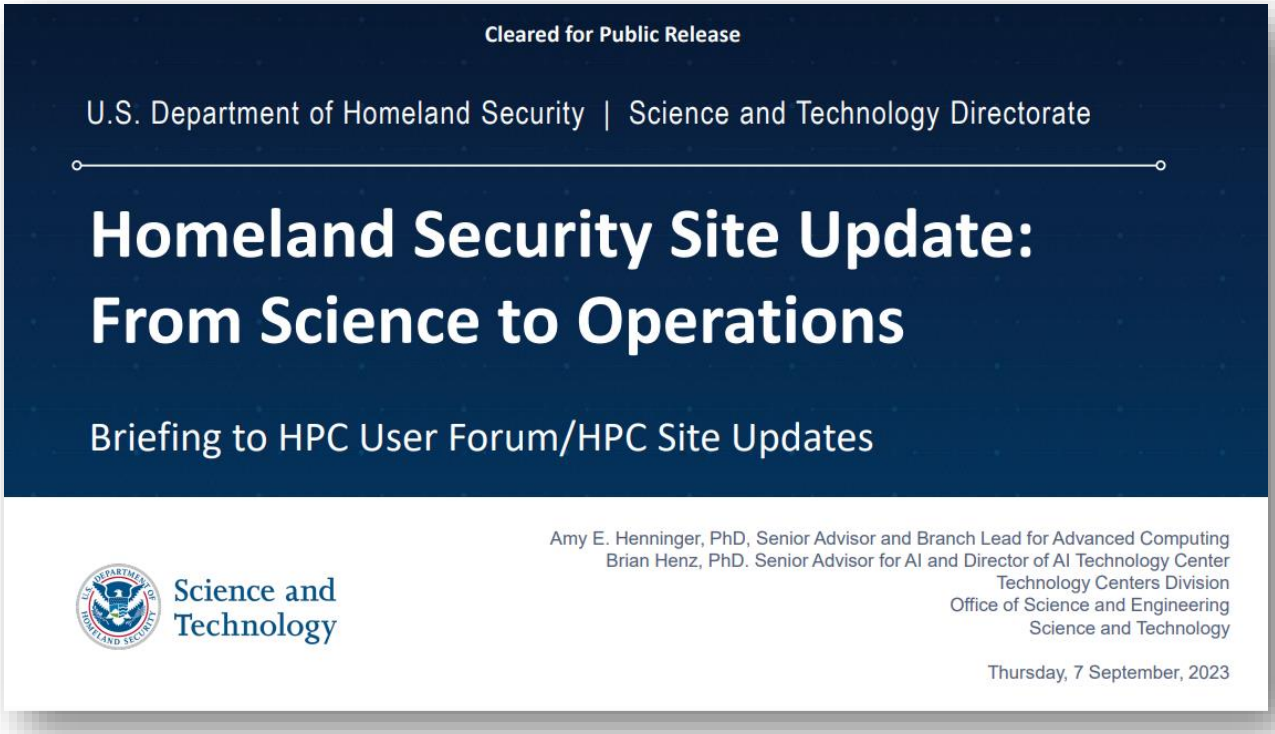
HPC User Forum Update

AI Foundation Models on the Horizon for DHS S&T

Thomas Sorensen
February 2024

IN THIS UPDATE

The HPC User Forum was established in 1999 to promote the health of the global HPC industry and address issues of common concern to users. In September 2023, the 83rd HPC User Forum took place in Tucson, Arizona. This update summarizes a presentation from that conference given by Amy Henninger, Senior Advisor and Branch Chief of Advanced Computing at the DHS Science and Technology Directorate/Technology Centers. In addition to providing updates on the various clusters and labs under the DHS umbrella and their application spaces, she also gave an update on the department's goals for the future as well as their relationships with new technologies like quantum and AI.



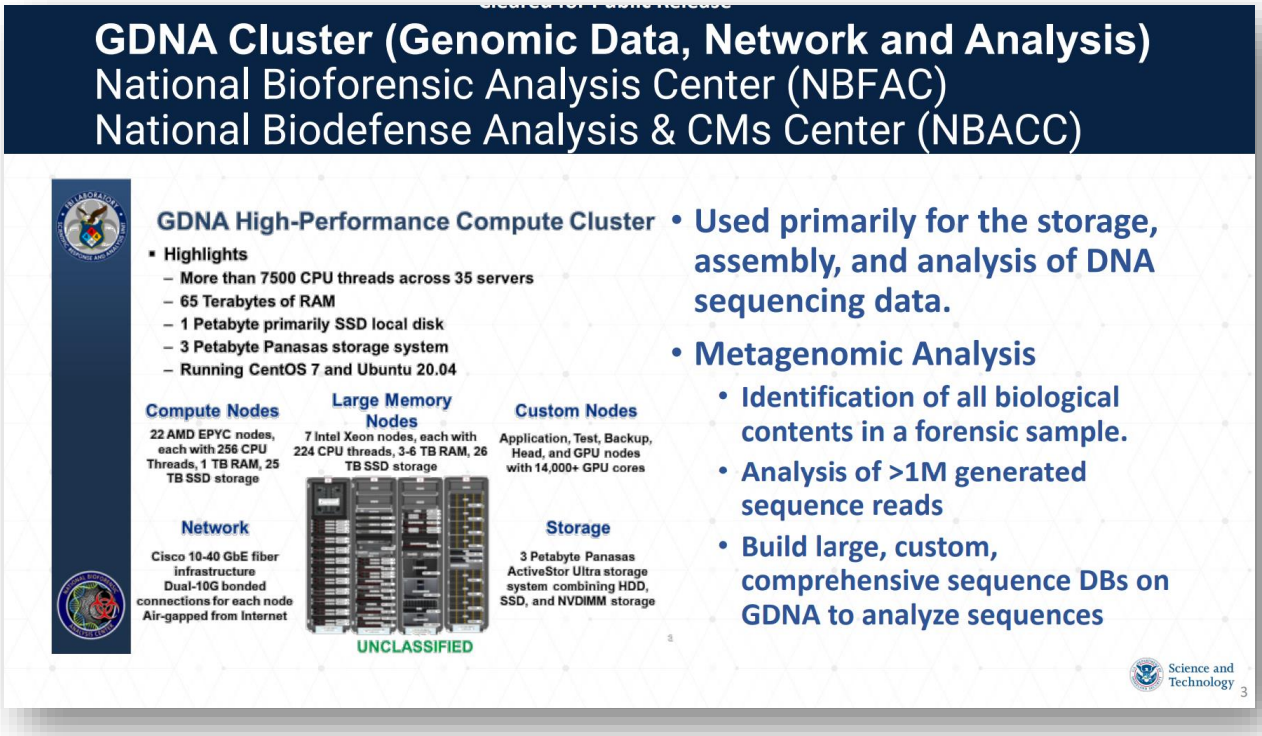
Source: DHS S&T, 2023

PRESENTATION: AMY E. HENNINGER, PHD, OFFICE OF SCIENCE AND ENGINEERING, DEPARTMENT OF HOMELAND SECURITY

Henninger began by highlighting the massive scope of activity within the DHS and the agencies they support: over 260,000 employees doing jobs ranging from aviation, border security and emergency response to cybersecurity analysis and chemical facility inspection. In all DHS is composed of 22 sub-agencies with 8 major players:

- Citizen and Immigration Services (USCIS)
- Coast Guard
- Customs and Border Protection (CBP)
- Cybersecurity and Infrastructure Security Agency (CISA)
- Federal Emergency Management Agency (FEMA)
- US Immigration Customs and Enforcement (ICE)
- US Secret Service
- Transportation Security Administration (TSA)

FIGURE 1



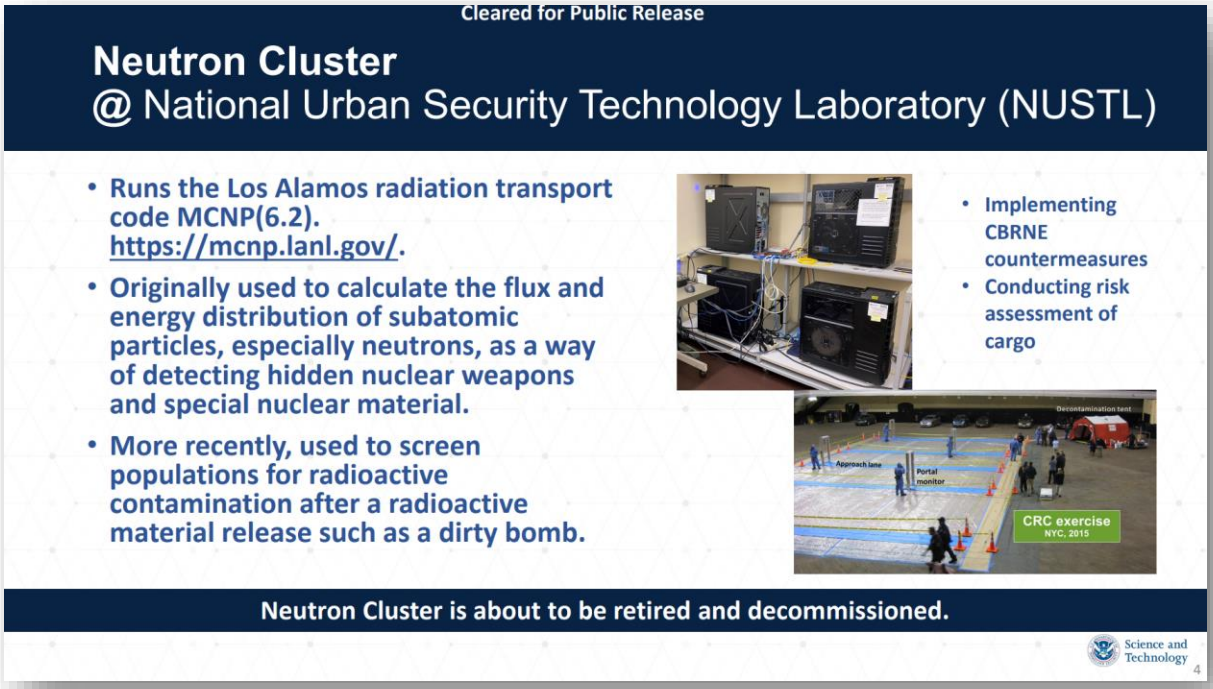
Source: DHS S&T, 2023

DHS S&T supports all the wide-ranging science within these groups. Currently, and increasingly so, HPC and related advanced data center technology is being used to accelerate or make possible these wide-ranging mission goals. In addition to providing this support through the use of traditional HPC methods, Henninger noted that DHS S&T is actively researching and deploying cutting-edge technologies in their agencies and labs.

According to Henninger, when she joined DHS a little over a year ago, there were only two HPC user groups. The first cluster was the GDNA Cluster (Genomic Data, Networking and Analysis), which supports the National Bioforensic Analysis Center (NBFAC) and is an office of the Federal Bureau of Investigation (FBI). A common use case for this cluster would involve manipulation and storage, and analysis of DNA sequencing data usually linked to the activities of a wet lab.

The second system was the Neutron Cluster from the National Urban Security Technology Laboratory (NUSTL). This system essentially runs Monte Carlo simulations using Los Alamos National Laboratory (LANL) radiation transport code to estimate particle interactions. An exercise that this cluster might support would be devising methods of screening for radioactive contamination in individuals in emergency situations. “By modern standards,” Henninger admitted, “this is not really HPC. But at DHS in the CIO office, it is still designated as HPC.”

FIGURE 2



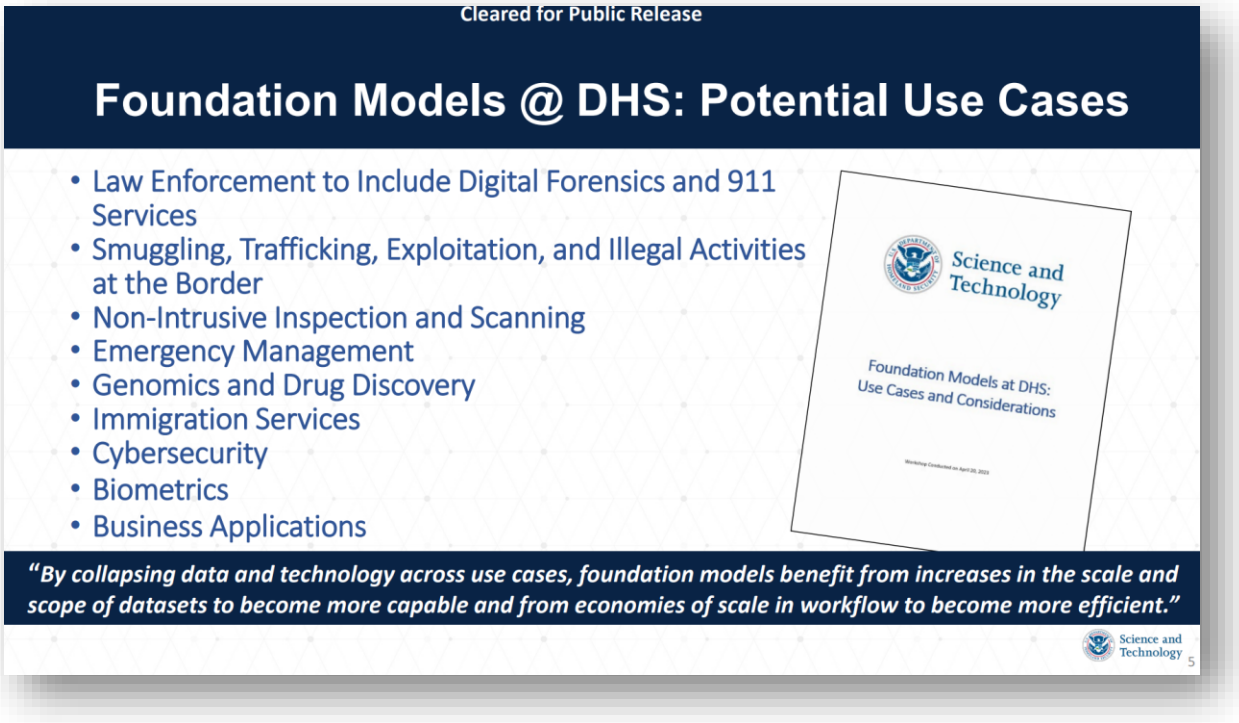
Source: DHS S&T, 2023

According to Henninger, while the portfolio may currently be modest, there is a big vision for DHS in the compute space. There was a recent DHS workshop conducted in partnership with IBM R&D to assess how many and to what extent DHS projects could benefit from foundation models. The report, *The Role of Science and Technology in Preparing for Future Change*, has recently been published on the DHS S&T website, with the list of general areas set to receive HPC integration being:

- **Law Enforcement to Include Digital Forensics and 911**
    - Digital Forensic tools are extremely helpful to DHS case agents in solving cases like child exploitation, counter-fentanyl and counterfeiting. Henninger believes foundation models could be used to support this. Locations, dates, names, activities, can all be examined and retrieved from unstructured data. Next generation 911 capabilities will help to allow texts, pictures, and videos to be sent to emergency dispatchers.
- **Smuggling, Trafficking, Exploitation, and Illegal Activities at the Border**
    - The CBP covers over 7,000 miles of border and they have a multitude of data resources including multi-spectral data, text reports, imagery, sensor data, even social and news media. According to Henninger, foundation model forecasting could help them in allocating resources in a more optimized and timely way. The information captured in this way could also stand to also serve multiple DHS sub-agencies if leveraged properly.
- **Non-Intrusive Inspection and Scanning**
    - The TSA generates 5.5 million images a day. When they scan luggage and individuals, they are all captured as images and stored. That does not even include those 7,000 miles of land ports and maritime ports covered by customs and border control where they are also scanning vehicles and cargo containers. These two collection efforts create a massive data source for DHS S&T that Henninger believes could be better leveraged in terms of an image and video-based foundation model.
- **Emergency Management**
    - FEMA and other emergency response organizations need accurate real-time information to support the optimal allocation of resources during disasters. Weather reports, geospatial images, social media, news reports -- all of these are data sources that could be used as part of a foundation model to help FEMA quickly coordinate their actions and optimize their response. Henninger also sees this technology aiding in the administrative part of relief provision.
- **Genomics and Drug Discovery**
    - Genomic data from centers like NBFAC and NBACC could be used to train foundation models in identifying new zoonotic diseases and creating treatments for those as well. Furthermore, there is data contained within multiple sources such as papers, publications, and disparate digital forms that could be aggregated and leveraged for scientists with the help of HPC acceleration.
- **Immigration Services**
    - Immigration services are a highly administratively demanding practice. Not only are applicants often in need of a translator to complete the over 100 application documents, but some of these forms also still require handwritten data to be entered. This process could benefit greatly from digitization and automation, going hand-in-hand

with the introduction of a foundation model. Automating tasks allows more resources to be allocated to those cases most in need of active human attention.

FIGURE 3



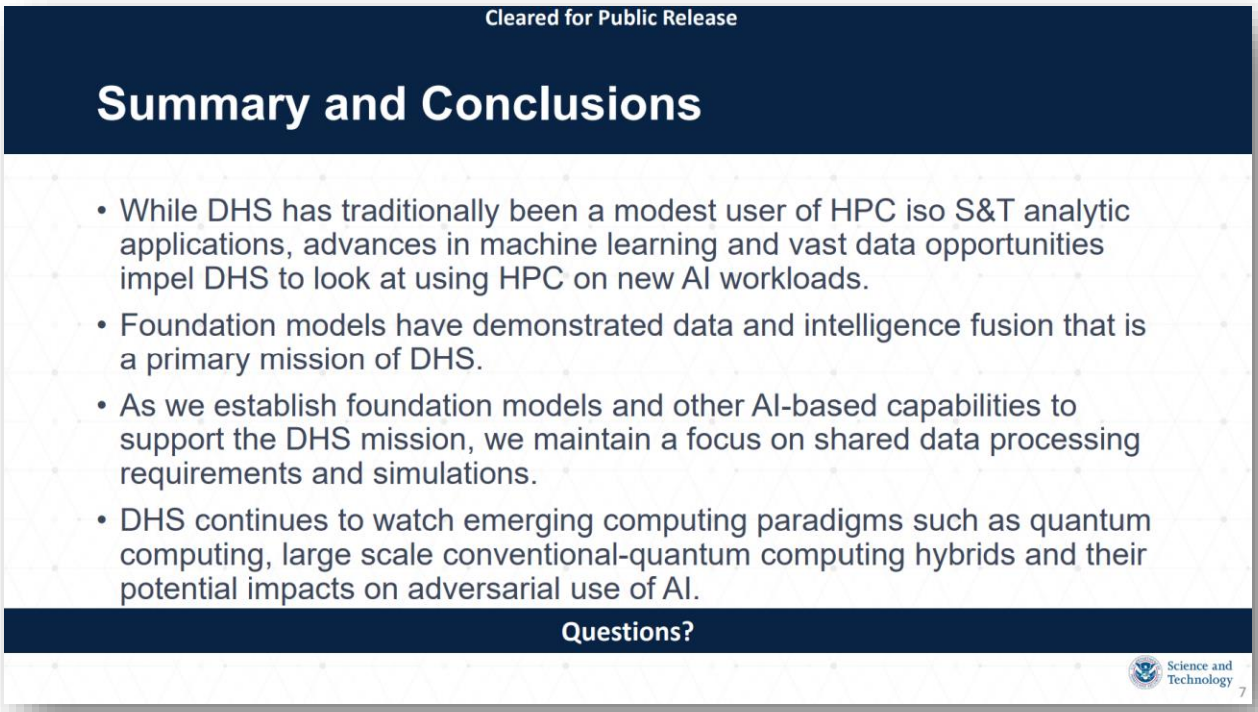
Source: DHS S&T, 2023

- **Cybersecurity**
  - There is a clearer path in cybersecurity for foundation models as the data is already formed and is easier to analyze. DHS supports many agencies with cybersecurity activities, and CISA alone receives 10TB+ of cybersecurity data a day from many different sources, mostly for threat hunting. Tasks that would make use of this data include behavior analytics, threat monitoring, automated incident response, process automation, and more.
- **Biometrics**
  - The biometrics work at DHS is focused on assessing vendor-supplied models. Models from industry, academia, and institutes go through an assessment process at NIST and then to the test facility at DHS. Despite this, DHS has one of the largest biometrics databases in the world and all that data is 'untapped' in these biometric models that are being assessed. Henninger admits that while the privacy regulations that keep that data unused serve a good purpose, their prospects as the basis for a powerful foundation model are extremely promising. Perhaps, she maintains, transfer learning or a process

of fine tuning an existing industry model using in-house DHS data would allow for a powerful yet still secure foundation model.

- **Business Application**
  - Software development, IT analysis, help desk, data science, and visualization are just a subset of DHS activities that Henninger views as eligible candidates for AI foundation models.

FIGURE 4



Source: DHS S&T, 2023

DHS S&T sees the adoption of foundation models as a way of improving on or even leapfrogging beyond their previously light HPC and ML portfolios. But while foundation modelling is at the front and center of DHS S&T, Henninger notes, "There is a horizon that we need to be aware of. There is goodness in integrating first principles models and modelling capabilities with AI-based workloads."

Despite not traditionally relying heavily on HPC, DHS expects, like many other organizations, their HPC activity to considerably uptick with the expected adoption of AI technology, specifically in the foundation modelling space. "This provides us the opportunity to do the kind of data and intelligence fusion that DHS was really stood up to do … we need to be doing better." Henninger described a future DHS ecosystem in which, "the whole is better than the sum of the parts." The ability to share and analyze data of different forms from disparate sources with a combination of traditional and cutting-edge technology could unlock new capabilities that enable DHS to provide critical services more optimally.

For more information or to view this and other presentations given at HPC User Forums dating back to 2008, visit www.hpcuserforum.com.

About Hyperion Research, LLC

Hyperion Research provides data driven research, analysis and recommendations for technologies, applications, and markets in high performance computing and emerging technology areas to help organizations worldwide make effective decisions and seize growth opportunities. Research includes market sizing and forecasting, share tracking, segmentation, technology, and related trend analysis, and both user & vendor analysis for multi-user technical server technology used for HPC and HPDA (high performance data analysis). We provide thought leadership and practical guidance for users, vendors, and other members of the HPC community by focusing on key market and technology trends across government, industry, commerce, and academia.

Headquarters

365 Summit Avenue
St. Paul, MN 55102
USA
612.812.5798
www.HyperionResearch.com and www.hpcuserforum.com
