Securing NASA's Most Powerful Supercomputer

Thomas H. Hinke, Ph.D., CISSP
NASA Advanced Supercomputing (NAS) Division
NASA Ames Research Center
Located in Silicon Valley, California

Outline

- **Background**: NASA Advanced Supercomputing (NAS) Division's high performance computing systems
- **Protection**: Security approaches for protecting the NAS systems
- **Detection**: Security approaches for detecting possible attacks on NAS systems
- **Research**: Improving methods to extract actionable information from the mountain of data that inundates the NAS systems

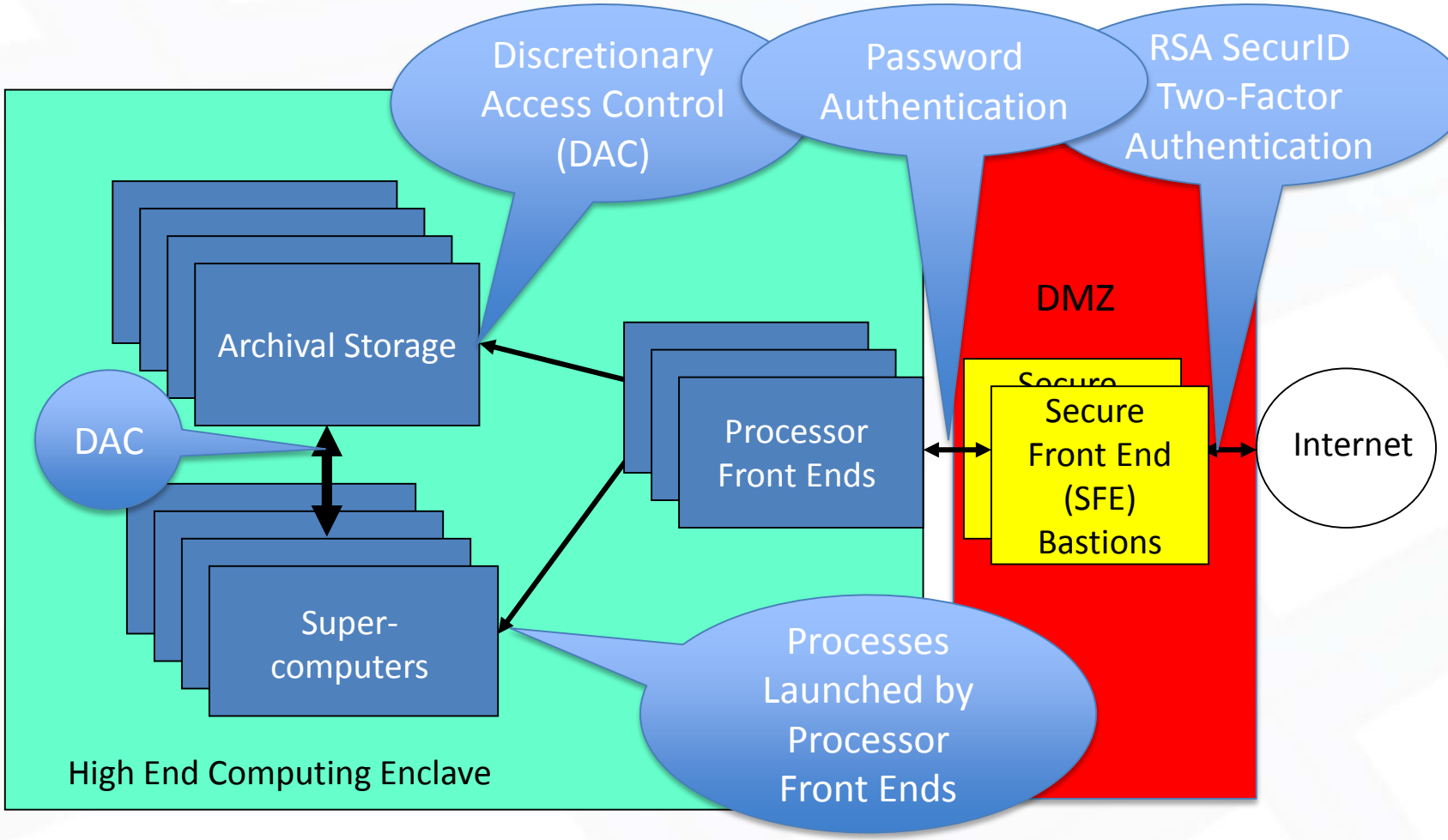
Background: Computer Systems at the NASA Advanced Supercomputing (NAS) Division

- Pleiades: NASA's most powerful supercomputer
  - No. 15 on TOP500 list of the world's most powerful supercomputers
  - 7.25 petaflops (PF) theoretical peak
  - 246,048 cores and 938 terabytes (TB) of memory
- Merope: Uses repurposed processors from Pleiades.
  - 162 Tflops/s theoretical peak
  - 13,824 Cores and 26 TB of memory
- Endeavour: Shared memory system
  - 32 teraflops (TF) theoretical peak
  - 2 nodes - 1–24 cores with 4 TB and 512 cores with 2 TB memory
- All three systems:
  - Manufactured by: SGI with Intel processors
  - Run SUSE® Linux

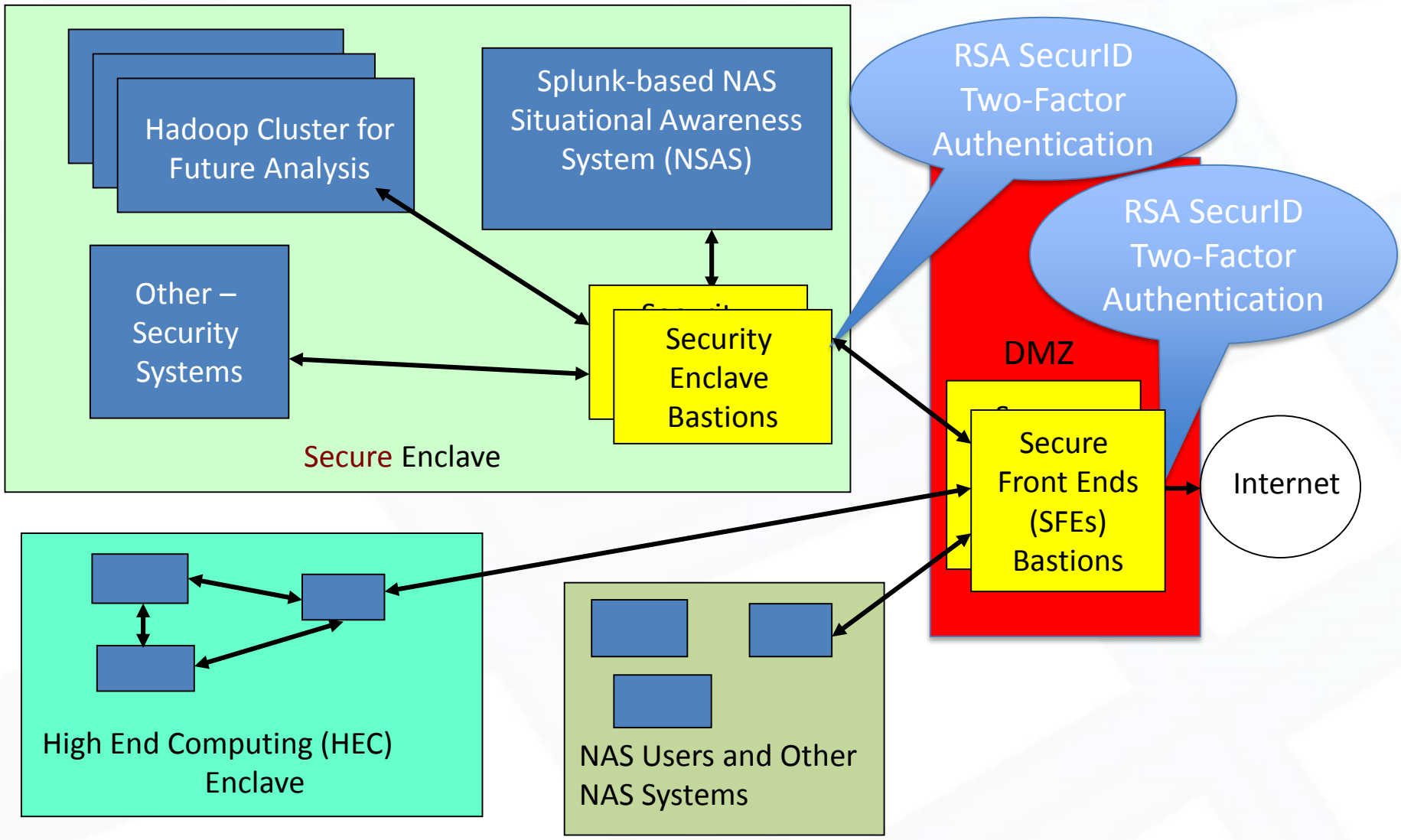
Background: Use and Data Sensitivity of Data on NAS Systems

- NAS systems provide supercomputing resources and user support services for science and engineering projects at:
  - NASA Centers
  - Universities performing NASA research
  - Companies performing NASA research
- NAS system are rated as moderate for data under FIPS 199
  - Loss of confidentiality, integrity, or availability will have serious adverse effect on the organization
- Number of NAS users
  - 1798 active accounts
    - Includes several hundred accounts for NAS staff and non-supercomputer users who are using other systems hosted at NAS
  - 440 accounts are for users who are not U.S. citizens
    - Includes graduate students and faculty
- Users can
  - Access the system over the Internet
  - Import and run their own codes

Protection: Supercomputers Protected Inside a High End Computing (HEC) Enclave



Both Local and Remote Users' Access to HEC Enclave Must be Authenticated by SFE



Security of Access to High End Computing Enclave

- Interactive access to the HEC Enclave is through an SFE using SSH to encrypt the connection
- File transfers into the HEC Enclave use several mechanisms:
  - SFEs support two-stage file transfers into the HEC Enclave
    - From source system to SFE
    - From SFE to Processor Front End or Storage
  - DMZ secure file server also supports two-stage file transfers
    - Requires only password authentication, so can be used for unattended file transfers
  - Secure Unattended Proxy supports one-Stage File Transfers
    - Also supports unattended file transfers using limited lifetime public-private keys to authenticate actual transfer
    - Files can be transferred to internal HEC filesystem
- File transfers out of the HEC Enclave are not restricted

Secure Front End Is An Attack Resistant Security Reference Monitor

- **Always Invoked Requirement**
  - Network Access Control Lists (ACLs) ensure that SFEs authenticate interactive user's access to Enclave-resident systems
- **Tamper Proof Requirement**
  - Design of the SFE minimizes the opportunity to attack the SFE
  - SFE implemented as a separate device so that it is isolated from tampering by users of other NAS systems
  - SFE implemented with a jailed (chrooted) environment for all users, which:
    - Limits user access to system directories
    - Permits users to access only those functions required to log in and perform file transfers
  - Capabilities of the SFE's Linux operating system
    - Has only those capabilities that are absolutely needed
    - Minimizes the possibility of including unneeded capabilities with potential security vulnerabilities
- **Correctly Enforces the Desired Security Policy Requirement**
  - Authenticates users using two-factor authentication

Use of SSH Rather than VPN For Secure Interactive Access

- Advantages of SSH
  - NAS uses SSH for all remote, interactive access to its systems
    - Users must first SSH into the SFE
  - SSH provides a high level of encryption protection
  - SSH is widely available
  - SSH can be used from multi-user systems and single-user systems
- Disadvantages of VPN
  - It places the remote system on a NASA network with a NASA IP address
  - The entire remote system must be given this NASA IP address when a remote user is using VPN
    - Thus, this is not appropriate for users on multi-user systems such as might be found at a university
  - Any user's action on their remote system while connected to a NASA VPN actually comes from a NASA IP address, which might place NASA in a bad light, if the user does something bad

Ongoing Actions to Protect NAS Systems

- NAS performs weekly credentialed scans using Nessus during which the scanner:
  - Logs onto each of the NAS systems
  - Identifies missing patches or existing vulnerabilities that make the systems susceptible to attack
  - Provides weekly reports to NAS system administrators and NAS management
- NAS sinkholes (null routes) foreign scans and all brute-force attacks, as well as all known malware sites, so that the bad actor gets no response

Detection Provided By NAS Situational Awareness System (NSAS)

- Our in-house NAS Situational Awareness System (NSAS) identifies security events that require human or automated mitigation
- Data sources include:
  - Bi-directional network flow data
  - Intrusion detection system (IDS) data
  - Log data
  - Nessus vulnerability scanner data
  - Domain Name Server (DNS) requests
  - Other external data sources (e.g., list of hostile sites, National Vulnerability Database)
- A preprocessor normalizes and enriches data into a common information model including geographic and “whois” information before sending data to Splunk
- Splunk is used for analysis and display

NAS Research is Underway to Develop Enhancements to NSAS

- NASA organizations are inundated with a mountain of network data with nuggets of valuable security information buried in the mountain
- Goal of Research: To reduce in near real-time the mountain of data that has to be analyzed *to discover actionable malicious security events*
- Approach is to categorize:
  - Entities (IP addresses, ports, organizations, and users)
  - Flows (stream of network packets)
  - Into the following risk categories, using advanced data analysis tools:
    - **Acceptable risks** are network flows and their associated IP addresses associated with authorized, authenticated users
      - In general, these can be eliminated from further consideration
    - **Hostile risks** are network flows and associated IP addresses that have been identified as associated with a hostile actor
    - **Unknown risks** are network flows and associated IP addresses that will be the focus of analysis to move them into either the acceptable or hostile risk categories

Research is Underway to Develop Enhancements to NSAS (cont.)

- Want to improve detection of actionable security events, such as:
  - Attack flows
  - Recurrent flows, which may indicate attempts by implanted advanced persistent threat software to contact a command and control center
  - Unauthorized exfiltration of NASA data
- Want to develop techniques to visualize associations between security events in order to identify unknown patterns of hostile attack

Questions?