SEC (Simple Event Correlator) is a monitoring tool often used to create custom monitoring of HPC (High Performance Computing) systems. Without SEC , error tracking and automated hardware triage ticket creation is affected. When SEC is not operational on a scalable unit leader node , then its corresponding compute nodes are dropped. Therefore, I developed a program to establish a "heartbeat" for each scalable unit leader node per HPC system. This "heartbeat" notifies the admin node which then calls an action to store an epoch time for each leader node. The program then performs a threshold check for 60 minutes. When a leader node does not respond within the allotted time, a warning message is displayed onto the NCCS (National Center for Computational Sciences) Nagios dashboard so that engineers may be notified of issues. This has been applied to nearly every HPC system and has improved the efficiency HPC monitoring infrastructure.

Improving the High-Performance Computing System Monitoring Infrastructure to Reliably Detect Events of Interest

Daniel Moon[1], Jordan Webb[2]
University of Tennessee Knoxville[1], NCCS Division[2], Oak Ridge National Laboratory[2]

Introduction

- Simple Event Correlator (SEC) is a monitoring tool used to create custom monitoring of High- Performance Computing (HPC) systems.
- Assists with error tracking and automated hardware triage ticket creation.
- No tool to monitor its status and health on scalable unit leader (SU).
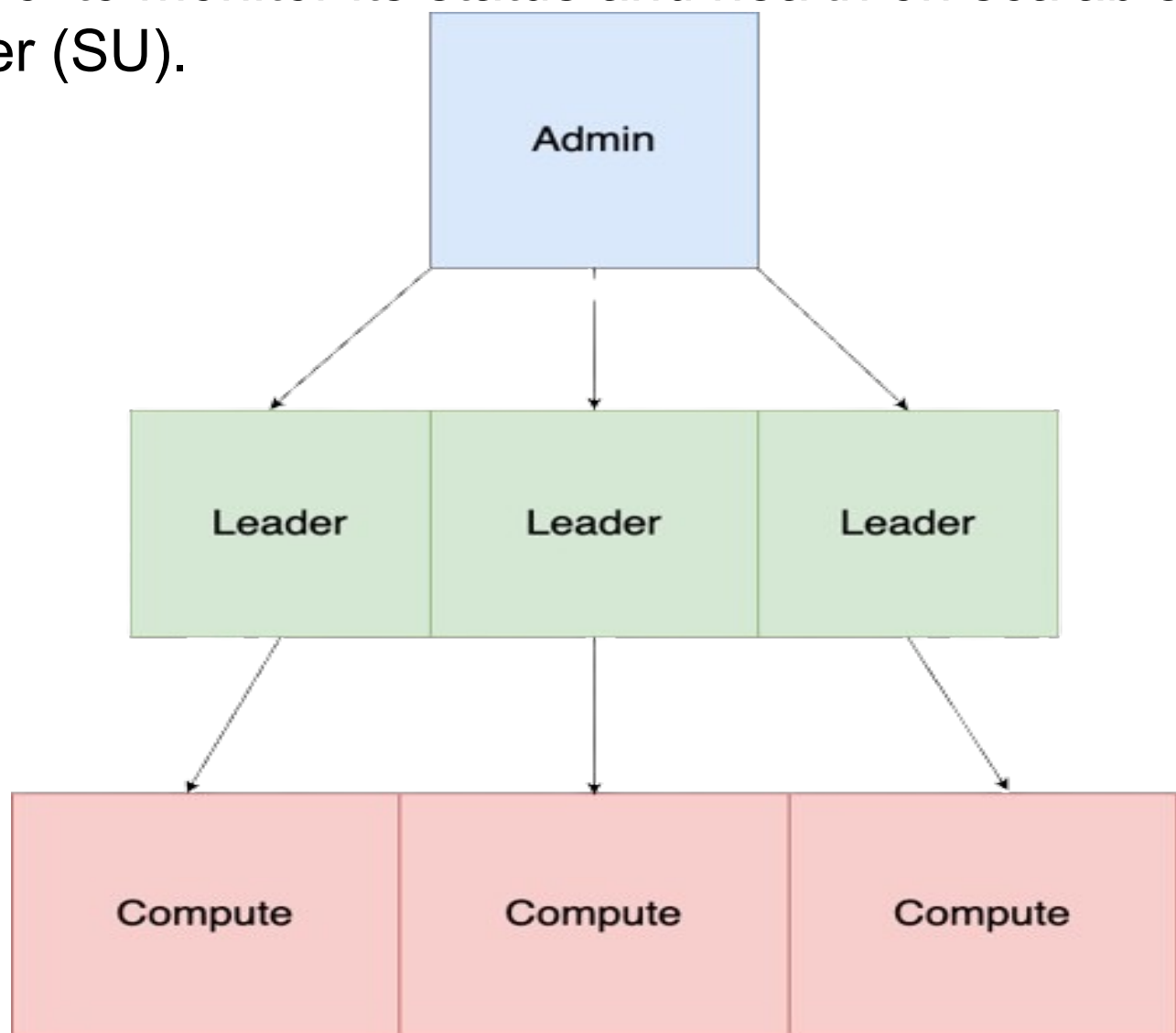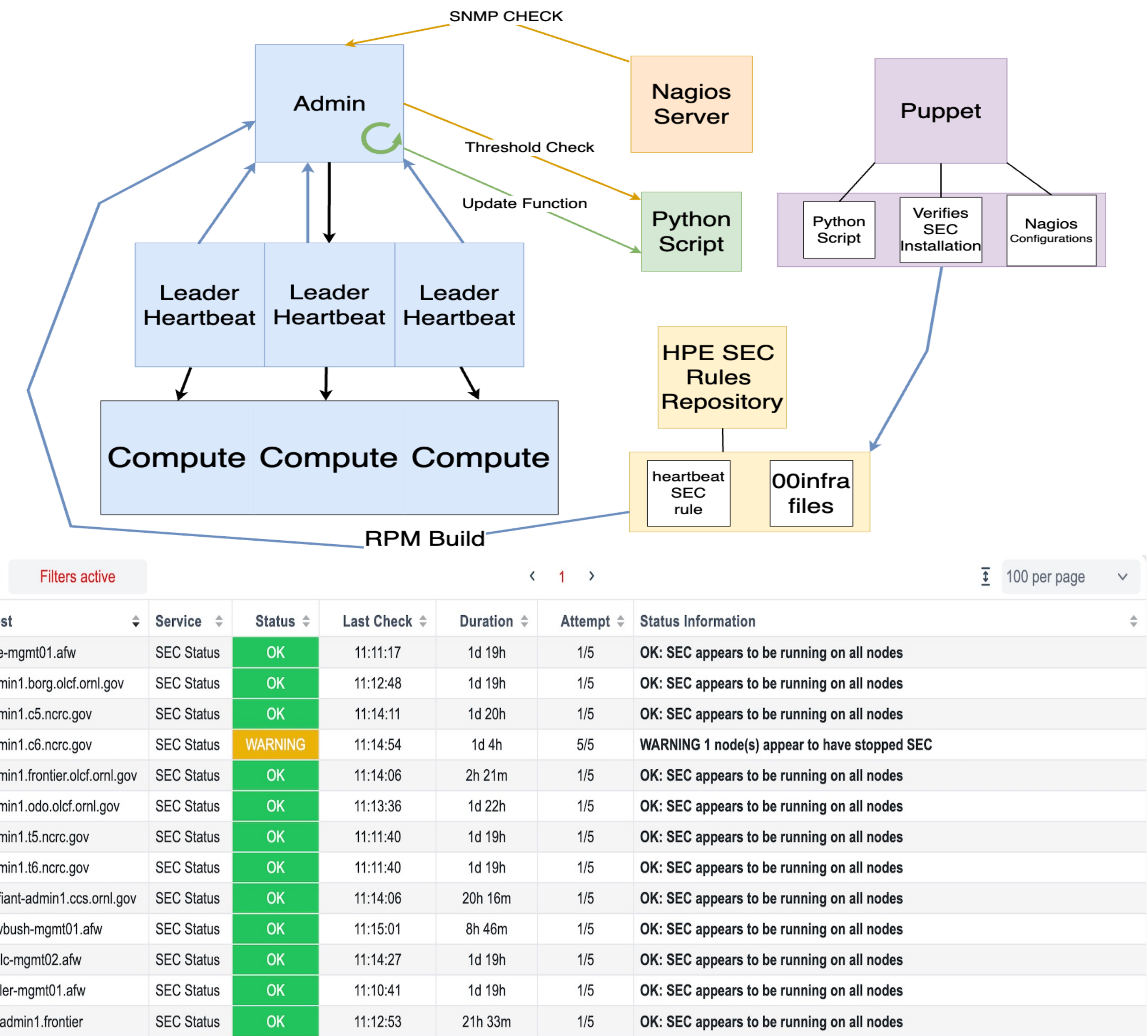

Methods

- Python Language
- Regular Expressions
  - Capable of tracking specific information from a string
- Threshold check
  - Tracks the status of the leader node and to ensure that SEC is operational.
  - If the login node reports within 60 minutes or less, then an OK message will display
  - Otherwise, a warning message will pop up
- A rule to establish a communication line between the leader node and the admin node.
- Visual display to show the SEC status on all HPC systems.
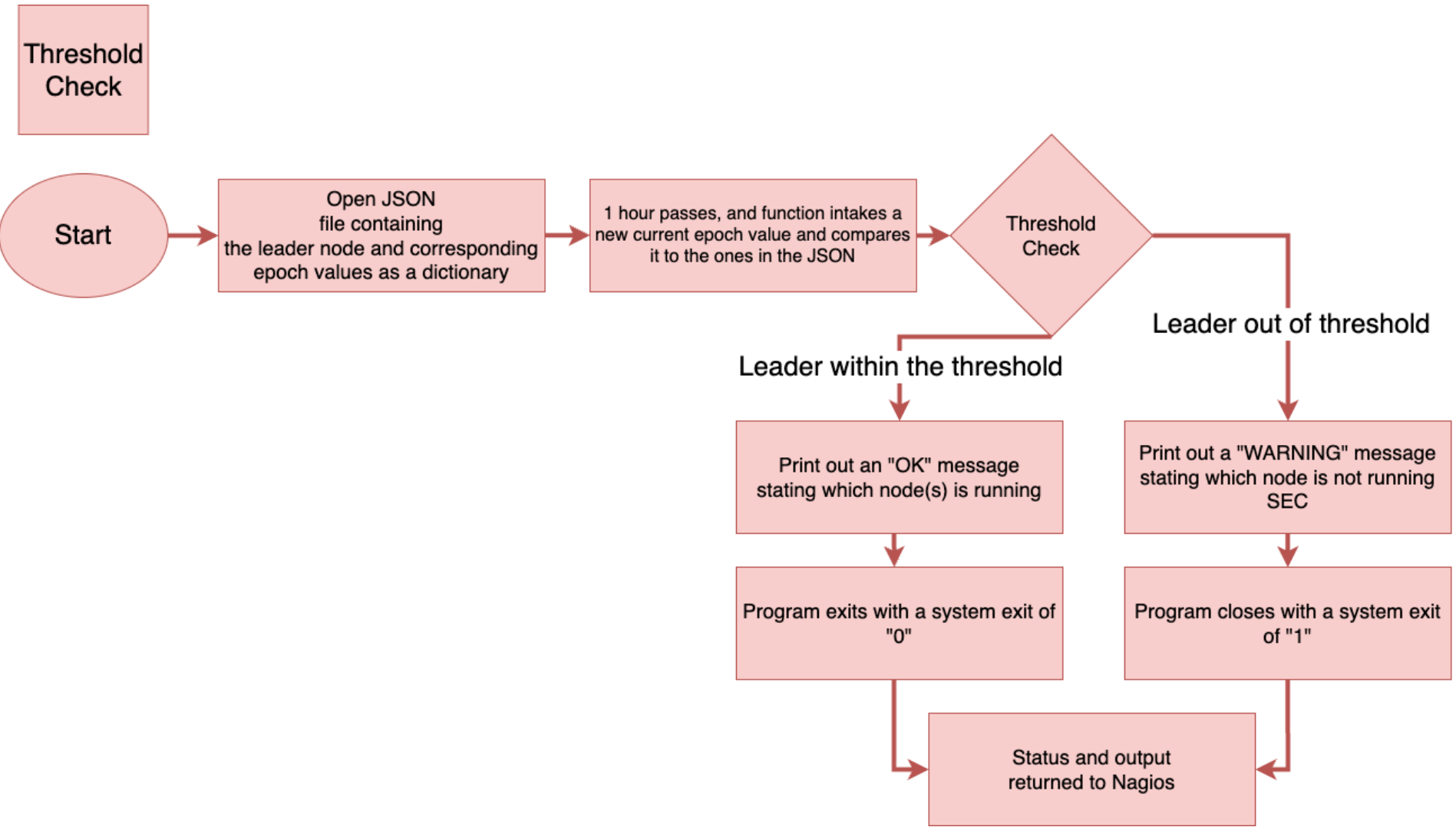
Results

- **Heartbeat Rule**
  - Calendar rule on the leader node to send a message to the admin node every hour.
  - Admin node receives the message and calls the "sec_health_check.py" script.
- **Health Check Script**
  - Obtain epoch value and associate with each leader node that sends a message to the admin node
  - JSON file is created to store this information.
  - Threshold check is run to ensure that the leader unit reports to the admin node within 60 minutes.
- **Visual Display on Nagios**
  - Nagios collects and organizes information about managed devices on IP networks.
  - Sends a Simple Network Management Protocol request to the admin node.
  - Displays the output of all systems that contain the "sec_health_check.py" script.
- **Automation and Production**
  - Enables python script, calendar rules, and Nagios configurations to be applied to every system.
  - Changes pushed to Puppet automatically update to the HPC systems the next time they run their puppet agent by the hour. An (RPM) package is installed onto the admin node.

Host	Service	Status	Last Check	Duration	Attempt	Status Information
ace-mgmt01.afw	SEC Status	OK	11:11:17	1d 19h	1/5	OK: SEC appears to be running on all nodes
admin1.borg.olcf.ornl.gov	SEC Status	OK	11:12:48	1d 19h	1/5	OK: SEC appears to be running on all nodes
admin1.c5.ncrc.gov	SEC Status	OK	11:14:11	1d 20h	1/5	OK: SEC appears to be running on all nodes
admin1.c6.ncrc.gov	SEC Status	WARNING	11:14:54	1d 4h	5/5	WARNING 1 node(s) appear to have stopped SEC
admin1.frontier.olcf.ornl.gov	SEC Status	OK	11:14:06	2h 21m	1/5	OK: SEC appears to be running on all nodes
admin1.odo.olcf.ornl.gov	SEC Status	OK	11:13:36	1d 22h	1/5	OK: SEC appears to be running on all nodes
admin1.t5.ncrc.gov	SEC Status	OK	11:11:40	1d 19h	1/5	OK: SEC appears to be running on all nodes
admin1.t6.ncrc.gov	SEC Status	OK	11:11:40	1d 19h	1/5	OK: SEC appears to be running on all nodes
defiant-admin1.ccs.ornl.gov	SEC Status	OK	11:14:06	20h 16m	1/5	OK: SEC appears to be running on all nodes
fawbush-mgmt01.afw	SEC Status	OK	11:15:01	8h 46m	1/5	OK: SEC appears to be running on all nodes
hallc-mgmt02.afw	SEC Status	OK	11:14:27	1d 19h	1/5	OK: SEC appears to be running on all nodes
miller-mgmt01.afw	SEC Status	OK	11:10:41	1d 19h	1/5	OK: SEC appears to be running on all nodes
spiadmin1.frontier	SEC Status	OK	11:12:53	21h 33m	1/5	OK: SEC appears to be running on all nodes

Conclusions

- Sustainable tool to monitor the health status of SEC on HPC systems.
- Without a tool to monitor SEC, the triage process is heavily impacted.
- System availability improves and helps ensure the integrity and continued operations of HPC resources.

Further Information

- Implemented a version number check utilizing Python Regular Expressions.
  - Called the RPM in the python script and used the Python Regular Expressions module to obtain the version number of SEC running on each system.
  - This ensures that a consistent version of SEC is running across every resource

Host	Service	Status	Last Check	Duration	Attempt	Status Information
ace-mgmt01.afw	SEC Status	OK	11:41:57	3d 34m	1/5	OK: SEC (1.80) appears to be running on all nodes
admin1.borg.olcf.ornl.gov	SEC Status	OK	11:43:50	1d 3h	1/5	OK: SEC (1.80) appears to be running on all nodes
admin1.c5.ncrc.gov	SEC Status	OK	11:42:50	6d 21h	1/5	OK: SEC (1.80) appears to be running on all nodes
admin1.c6.ncrc.gov	SEC Status	OK	11:44:45	1d 9h	1/5	OK: SEC (1.80) appears to be running on all nodes
admin1.frontier.olcf.ornl.gov	SEC Status	OK	11:44:51	7d 7h	1/5	OK: SEC (1.80) appears to be running on all nodes
admin1.odo.olcf.ornl.gov	SEC Status	OK	11:41:22	3d 11h	1/5	OK: SEC (1.80) appears to be running on all nodes
admin1.t5.ncrc.gov	SEC Status	OK	11:44:31	5d 8h	1/5	OK: SEC (1.81) appears to be running on all nodes
admin1.t6.ncrc.gov	SEC Status	OK	11:44:13	1d 13h	1/5	OK: SEC (1.80) appears to be running on all nodes
defiant-admin1.ccs.ornl.gov	SEC Status	OK	11:45:27	4d 21h	1/5	OK: SEC (1.81) appears to be running on all nodes

Literature Cited

1. R. Vaarandi, June 03, 2023, "Man page of sec," July 8, 2024
2. V. Bajrami, Nov. 27, 2020, "How to create a Linux RPM package," July 10,
3. 2024
4. Red Hat Documentation, "Introducing Configuration Management Using Puppet," July 10, 2024
5. A. S. Gillis, July 2023, "Nagios,", July 11, 2024

Acknowledgements