# Projects and Software

## AttackTagger

The cyber-infrastructure that supports science research faces the daunting challenge of defending against cyber attacks. Modest to medium research project teams have little cyber security expertise to defend against the increasingly diverse, advanced and constantly evolving attacks. Even larger facilities that have with security expertise are often overwhelmed with the amount of security log data they need to analyze in order to identify attackers and attacks, which is the first step to defending against them. AttackTagger can scale to be able to address the dramatic increase in security log data, and detect emerging threat patterns in today's constantly evolving security landscape. AttackTagger is a sophisticated log analysis tool designed to find potentially malicious activity, such as credential theft, by utilizing a Factor Graph model. AttackTagger integrates with existing security software so as to be easily deployable within existing security ecosystems and consumes a wide variety of system and network security logs.

## Blue Waters Supercomputer

Blue Waters has served thousands of engineers, scientists and educators across the nation, including those who scientists with unique jobs that can't be served by other systems. With tens of thousands of nodes and 450 Gbps of external WAN connectivity on an open network, Blue Waters presents a unique set of challenges to secure.

NCSA CyberSecurity team worked to redesign security from the ground up with this new system and the National Petascale Computing Facility during the 5 years going up to full operations. Using risk based methods to develop a new security program and architecture, we have achieved the gold standard of security in the community of NSF cyberinfrastructure being the first system to require two-factor authentication, designing and deploying one of the first 100Gbps network monitoring infrastructures, and testing many other security technologies at unprecedented scale.

## Center Cyber-protection

Cybersecurity at NCSA provides for the protection of the center's digital assets and those of key partners through the many services we provide. We have a 24/7 incident response team that performs full digital forensics and coordinates with law enforcement and other institutions. Preventative security is provided by our vulnerability management program, risk assessments & security architecting, automatic blocking, and more. We run over 60 servers to provide the monitoring, logging and other security services, including one of the largest production Zeek clusters in the world. The CyberSecurity division is also responsible for training staff, notifying reliant parties of new vulnerabilities, creating policies and much more security awareness work. Finally, we participate in several organizations and maintain collaborations with XSEDE, CERN, and others in the community.

## Custos

The Custos project provides security middleware for science gateways. The project is a collaboration between Indiana University, Johns Hopkins University, and NCSA.

## CILogon

The CILogon project enables use of federated identities by science projects. The project develops open source software that implements security standards including OAuth, SAML, and X.509. CILogon is an InCommon federation research and scholarship service provider that enables federated access to Globus, OSG, LIGO, XSEDE, and other cyberinfrastructure. NCSA offers subscriptions to organizations that use CILogon.

## Duo at NCSA

Duo is a Multi-factor authentication system provided to NCSA account holders by an agreement with the University of Illinois. Through this agreement NCSA is provided with a Duo instance that is separate from the University's to accommodate the unique challenges at NCSA.

## Large Synoptic Survey Telescope (LSST)

The Cyber Security Division at the NCSA provides cyber security policy and security operations services to the Large Synoptic Survey Telescope.
CSD helps secure and ensure that the LSST fulfills its scientific mission that includes delivering 200 petabyte set of images and data products.

## Science DMZ Actionable Intelligence Appliance (SDAIA)

SDAIA aims to secure Science DMZs and cyber-infrastructure, and provide the cybersecurity research community with a rich, real-world intelligence source upon which to test their theories, tools, and techniques. Science DMZs support big data and access to high-performance computation through very high bandwidth networks in an open environment that presents new challenges to the traditional university security stance. SDAIA provides a holistic approach that will address the special Science DMZ architecture through a virtual security appliance that benefits from external, shared intelligence to protect the site, and further provide intelligence to the wider community of both DMZ operators and cybersecurity researchers.

## SciTokens and SciAuth

The management of security credentials such as passwords and secret keys for computational science workflows is a burden for scientists and information security officers. Problems with security credentials (e.g., expiration, privilege mismatch) cause the workflows to fail to fetch needed input data or store valuable scientific results, distracting scientists from their research by requiring them to diagnose the problems, re-run their computations, and wait longer for their results. In an effort to avoid these problems, scientists often use long-lived, highly-privileged credentials (e.g., enabling the workflow to fully impersonate their identity), increasing risks to their accounts and to the underlying computational infrastructure and resulting in complexity for information security officers managing the infrastructure. The SciTokens project delivers open source software to help scientists manage their security credentials more reliably and securely. The project includes participants from the Laser Interferometer Gravitational-Wave Observatory (LIGO) Scientific Collaboration and the Large Synoptic Survey Telescope (LSST) project to ensure relevance and facilitate adoption. Integration with the widely-used HTCondor software and collaboration with Open Science Grid and the Extreme Science and Engineering Discovery Environment (XSEDE) facilitates adoption by the wider scientific community.

To address the challenges and risks of managing security credentials for scientific workflows, the SciTokens project delivers an open source software infrastructure that uses IETF-standard Open Authorization (OAuth) tokens for capability-based secure access to remote scientific data. SciTokens uses OAuth refresh tokens, maintained securely on the submission node, to delegate short-lived, least-privilege OAuth access tokens to scientific workflows, to enable their remote data access. The access tokens convey the specific authorizations needed by the workflows, rather than general-purpose authentication impersonation credentials. These least-privilege authorization tokens help to address the risks of scientific workflows running on distributed infrastructure including NSF resources (e.g., LIGO Data Grid, Open Science Grid, XSEDE) and public clouds (e.g., Amazon Web Services, Google Cloud, Microsoft Azure). By improving the interoperability and security of scientific workflows, the SciTokens project 1) enables use of distributed computing for scientific domains that require greater data protection and 2) enables use of more widely distributed computing resources by reducing the risk of credential abuse on remote systems.

The SciAuth project supports adoption of SciTokens by NSF cyberinfrastructure. It provides community engagement, support for coordinated adoption of community standards, assistance with software integration, security analysis and threat modeling, training, and workforce development to enable improved interoperability and usability for security credentials across NSF cyberinfrastructure. SciAuth aims to help the community realize the benefits of an interoperable, capability-based ecosystem when transitioning between credential technologies.

## Trusted CI

Trusted CI is the NSF Cybersecurity Center of Excellence. Its activities include one-on-one engagements with NSF projects to address their cybersecurity challenges; education, outreach, and training to raise the state of security practice across the scientific enterprise; and leadership on bringing the best and most relevant cybersecurity research to bear on the NSF cyberinfrastructure research community.

## XSEDE Federation

XSEDE is a federation of service providers and virtual organizations that have come together to bring high-performance computing to scientists at research institutions across the U.S. The mission of XSEDE is to enhance the productivity of scientists and engineers by providing them with new and innovative capabilities and thus facilitate scientific discovery while enabling transformational science/engineering and innovative educational programs.

The XSEDE project is led out of NCSA, and the security operations team in particular is co-led by NCSA CyberSecurity director Alex Withers and Derek Simmel at PSC. Jim Basney of NCSA's CyberSecurity division is also the security lead for XSEDE's Requirements Analysis & Capability Delivery (RACD) team, driving many of the IdM and security projects like the single-sign-on hub and Duo two-factor authentication integration.

## Advanced Computational Health Enclave (ACHE)

The Advanced Computational Health Enclave (ACHE) is a special environment with restricted physical and electronic access at NCSA. The ACHE is a physically isolated 1000-square-foot data center that is strictly managed to support electronic Protected Health Information (ePHI). It follows HIPAA standards and University of Illinois policies and procedures to ensure conformance and protection of ePHI. The ACHE undergoes annual security assessments and is currently undergoing SOC2 Type2 certification.

### Nightingale

The Nightingale system is a tenant of the ACHE and provides researchers with an environment to process and store ePHI.  Nightingale provides both compute and storage technologies and facilities for the ingest of data and processes for the export of de-identified data.  Nightingale provides the capability to meet a wide variety of researcher's computing needs ranging from small systems to large multi-node systems supported GPUs and low latency inter-connects.  Nightingale leverages NCSA's experience and expertise in meeting these scientific, high through put and high performance computing needs.