TTA 10

Efficient Tracking, Logging, and Blocking of Accesses to Digital Objects

Cyber Security Division
2013 Principal Investigators' Meeting

September 16th, 2013

Fabian Monrose
University of North Carolina at Chapel Hill
fabian@cs.unc.edu
919-962-1763

Michael Bailey, University of Michigan
Charles Schmitt, Renaissance Computing Institute

Team Profile

Homeland Security
Science and Technology

- Fabian Monrose: **Professor** of Computer Science at University of North Carolina at Chapel Hill. Prior to joining UNC, he was an Associate Professor at John Hopkins, and a founding member of their Information Security Institute.

- Michael Bailey: **Research Professor** at University of Michigan. He currently directs and contributes to research on the security and availability of complex distributed systems. Prior to working at the university, he was Director of Engineering at Arbor Networks, and a programmer at both Amoco Corporation and Andersen Consulting.

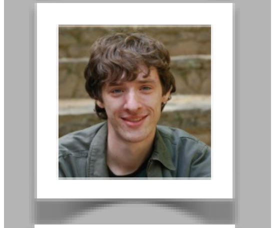
- Charles Schmitt: **Director of Informatics**. He provides technical leadership and management for biological and medical science related projects RENCI. Prior to joining RENCI, he was the senior computer scientist at BD Technologies, where he assisted in software development and bioinformatics support for programs in medical diagnostics and genomics.

F. Monrose

M. Bailey

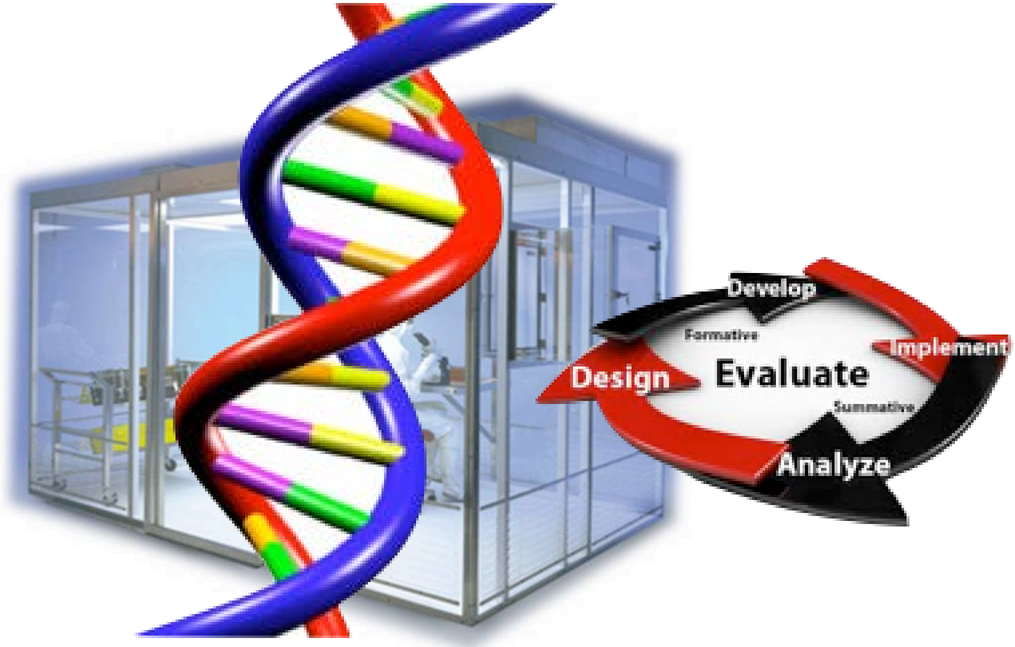
C. Schmitt

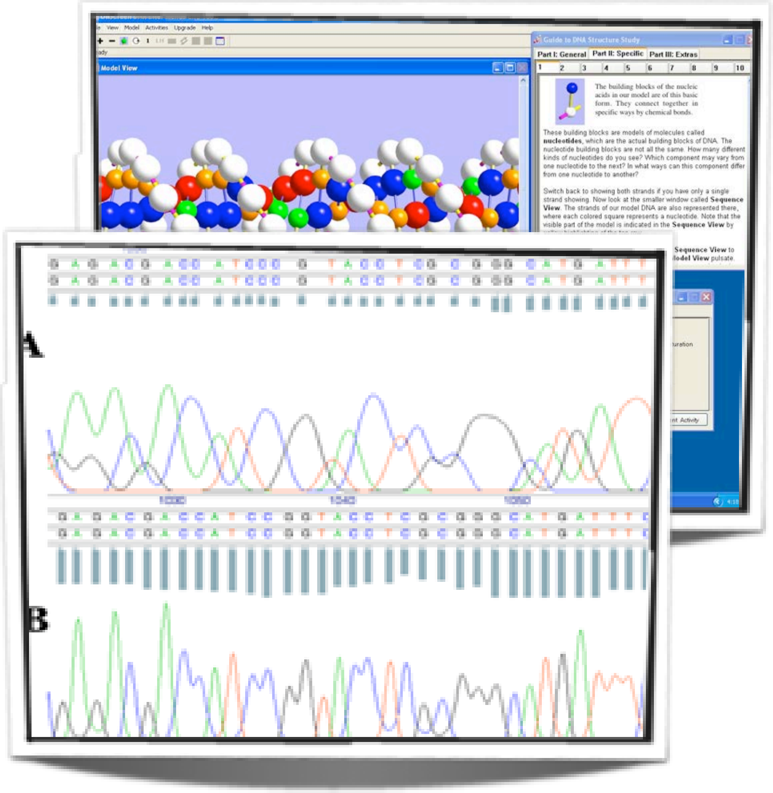
Engineers



Need

- Researchers and practitioners routinely require access to large corpora of **sensitive** data for a wide host of scientific activities





- Cleanrooms are secure, but cumbersome to use
- Do not meet today’s needs

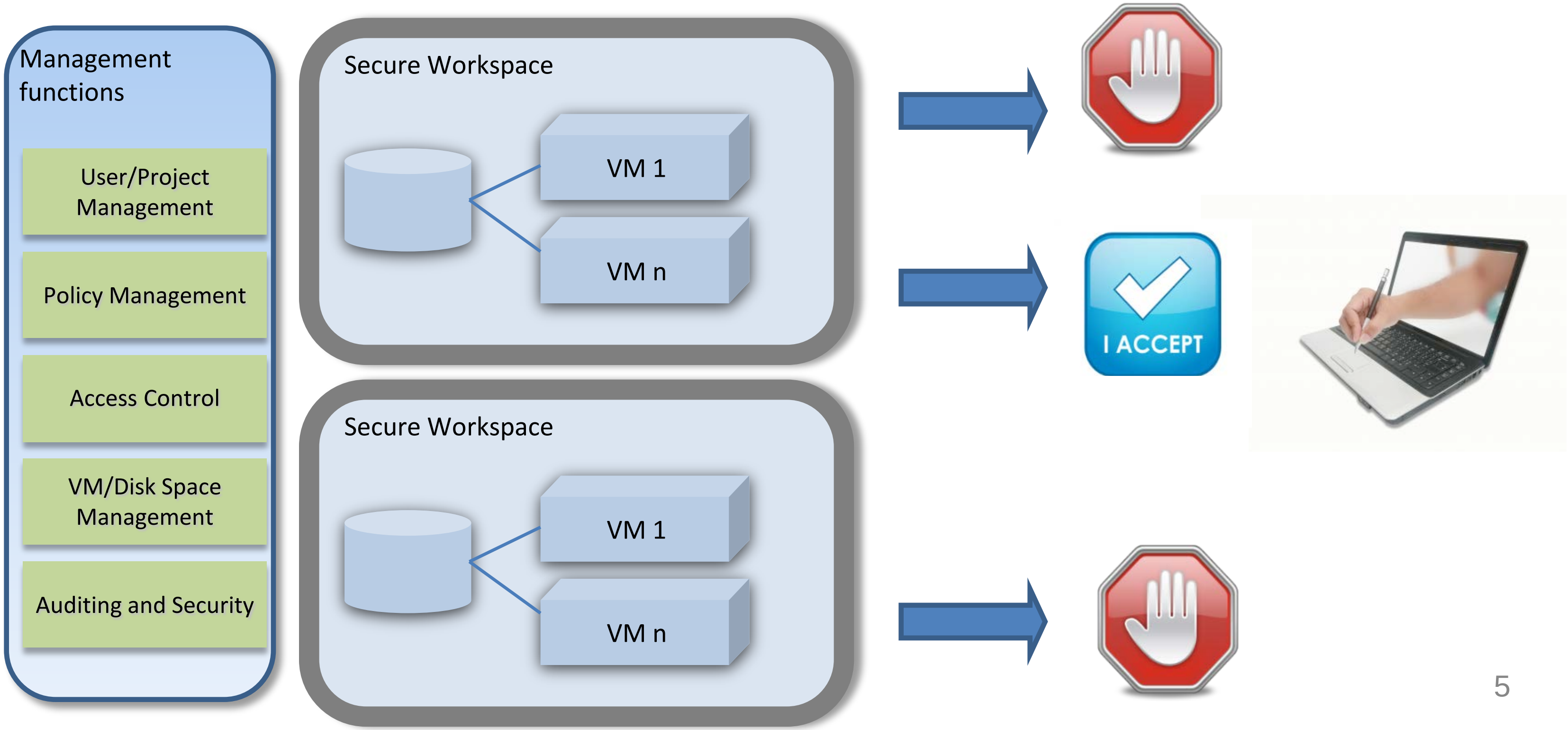
Need

- **Virtual data enclaves** have emerged as solutions for hosting sensitive data (e.g., medical records, call meta-data, network traffic, etc.)

- Yet, few solutions provide a secure environment for **reducing the risks** of **unauthorized** access to, and loss of, information hosted in these enclaves

Goals

- **To design and implement techniques for tracking the chain of custody of sensitive data hosted in enclaves**
- Particularly interested in an evaluation within the **Secure Research Workspace** at the Renaissance Computing Institute (RENCI)



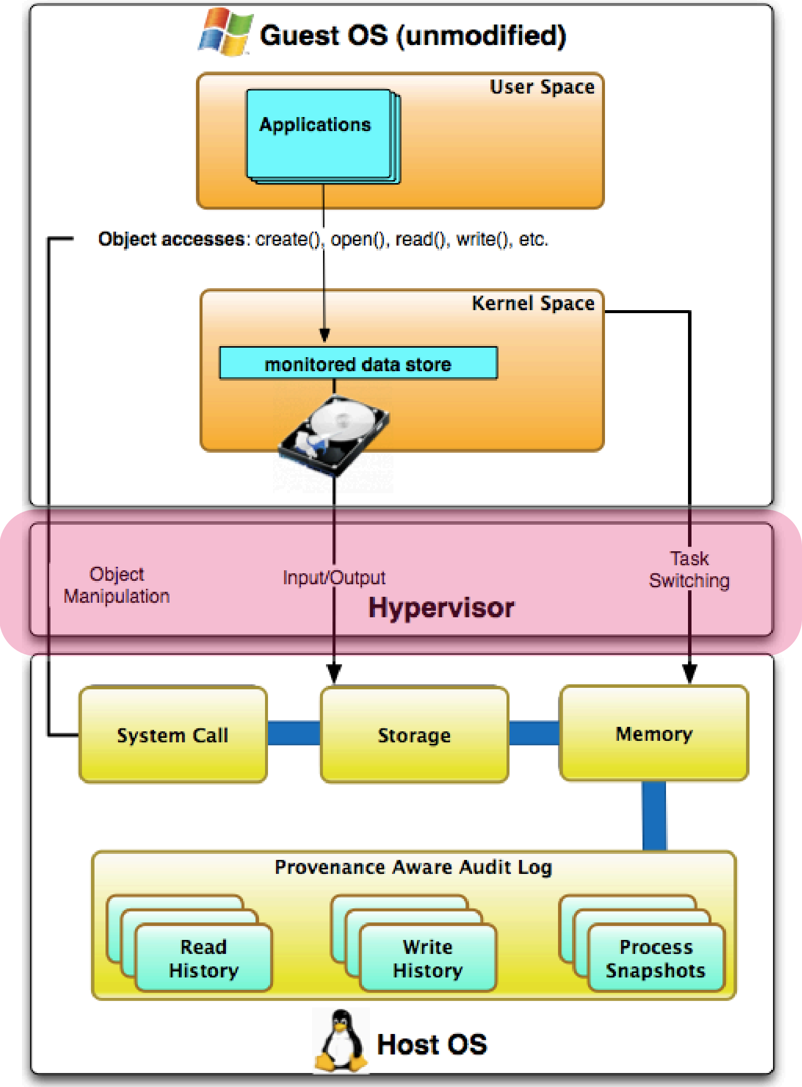
Goals

- **Deliverables:** An object tracking platform that will enable DHS and its customers to *(a)* **identify** and authenticate access to digital objects that originate from disk *(b)* **track accesses** to these objects on disks and in memory and *(c)* **track changes** to these objects via a provenance-aware audit trail

Approach

Homeland Security
Science and Technology

- Monitoring framework implemented within a Hypervsior
  - extends TrailOfBytes prototype
- Spans three layers: **storage, memory, and system-call** modules
  - *key idea is in monitoring access to physical memory when data is first loaded from a datastore*
- Semantic linkages captured in a **provenance-aware** filesystem



Benefits

Homeland Security
Science and Technology

- **Enhance state of the art in digital provenance in virtual data enclaves**

☑ **Improved data provenance representation**: A rich interface for managing and **mining the recorded information**, thereby providing deeper insights into **how** objects were manipulated

☑ **Limiting breaches**: Capabilities to not only **record**, but also to **deny**, unauthorized accesses or transfer of data from datastores for which provenance tracking has been enabled

Competition

- Several **research prototypes** or **government-led efforts** for hosting sensitive data

Statistisk sentralbyrå
Statistics Norway

UKDA
UK Data Archive

ICPSR
A PARTNER IN SOCIAL SCIENCE RESEARCH
INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH

cbs
Statistics Netherlands

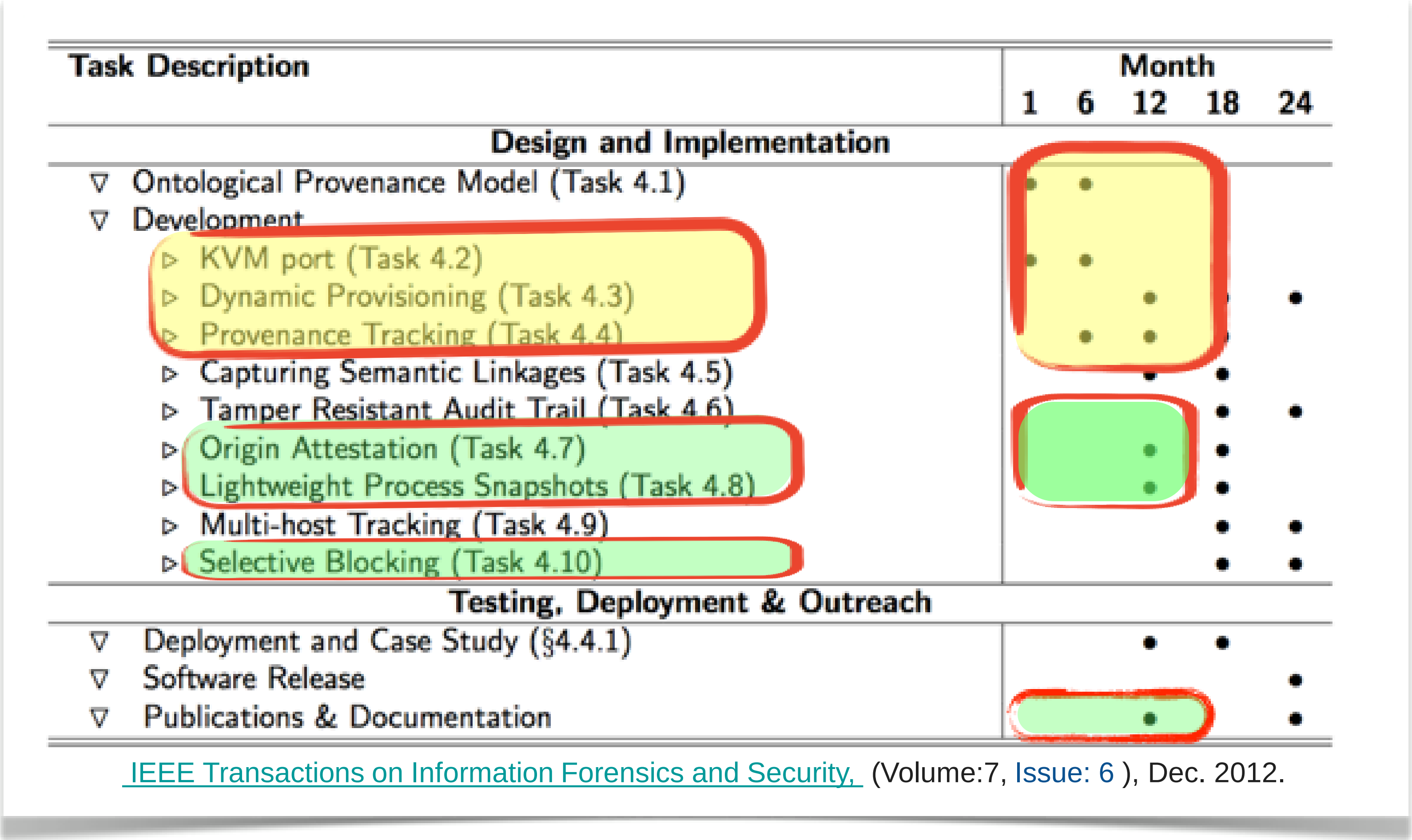
U.S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS

DHS PREDICT

NORC data enclave
the key to unlocking data securely

- But, no commercial ventures (that we ar
aware of).

Current Status



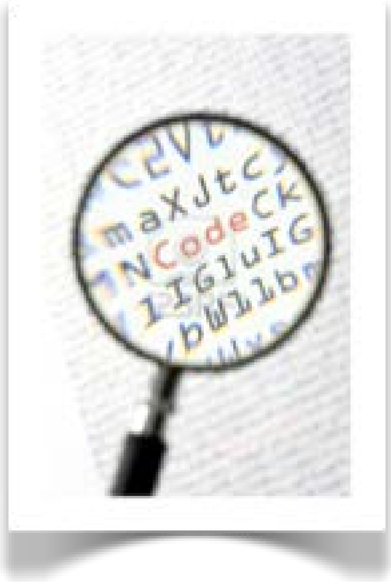
Task Description	Month				
	1	6	12	18	24
Design and Implementation					
▽ Ontological Provenance Model (Task 4.1)	•	•			
▽ Development					
▷ KVM port (Task 4.2)	•	•			
▷ Dynamic Provisioning (Task 4.3)			•	•	•
▷ Provenance Tracking (Task 4.4)		•	•	•	
▷ Capturing Semantic Linkages (Task 4.5)			•	•	
▷ Tamper Resistant Audit Trail (Task 4.6)				•	•
▷ Origin Attestation (Task 4.7)			•	•	
▷ Lightweight Process Snapshots (Task 4.8)			•	•	
▷ Multi-host Tracking (Task 4.9)				•	•
▷ Selective Blocking (Task 4.10)				•	•
Testing, Deployment & Outreach					
▽ Deployment and Case Study (§4.4.1)			•	•	
▽ Software Release					•
▽ Publications & Documentation			•		•

IEEE Transactions on Information Forensics and Security, (Volume:7, Issue: 6), Dec. 2012.

🟩 completed 🟨 in-progress / under revisions

Milestones: Origin Attestation and fast process snapshots

- **Attestation**: Implemented a technique for accurately determining what applications are accessing a monitored object
  - use information gathered from the guest VM and target process’ image (as available within the hypervisor)
- **Multi-host process snapshots**: Implemented a technique to support process snapshots and fast transfers (at 1255 MB/s)

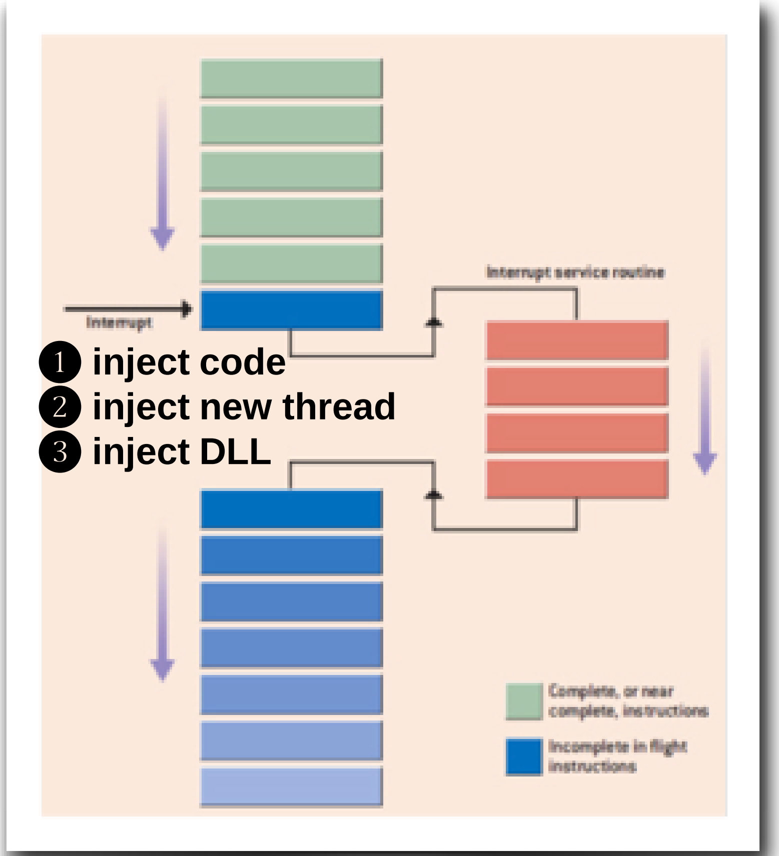


Fast Access™

Milestones:

(Provenance tracking and selective blocking)

- Explored several **within-guest OS** techniques for interrupting a running application:


Interrupt
Interrupt service routine
🟩 Complete, or near complete, instructions
🟦 Incomplete in flight instructions

1. inject code
2. inject new thread
3. inject DLL

Milestones:

(Provenance tracking and selective blocking)

- Adapted solution ❶ to a **within-hypervisor** mechanism that uses a system-call detection technique coupled with a code injection and process-redirection approach.

QuickTime™ and a
decompressor
are needed to see this picture.

Next steps
Deployment & Evaluation

- Technical assessment: several testing criteria including
  - ☑ accuracy (e.g., auditing and logging capabilities)
  - ☑ performance impact,
  - ☑ usability,

- System-level test with one of two existing **multi-institutional** studies using de-identified medical data:
  - medical visualization techniques to guide a medical decision support application for pediatric patients with epilepsy
  - medical visualization techniques to guide a medical decision support application for patients with a major depressive disorder

UNC SCHOOL OF MEDICINE

DukeMedicine

VE RI TAS
HARVARD MEDICAL SCHOOL