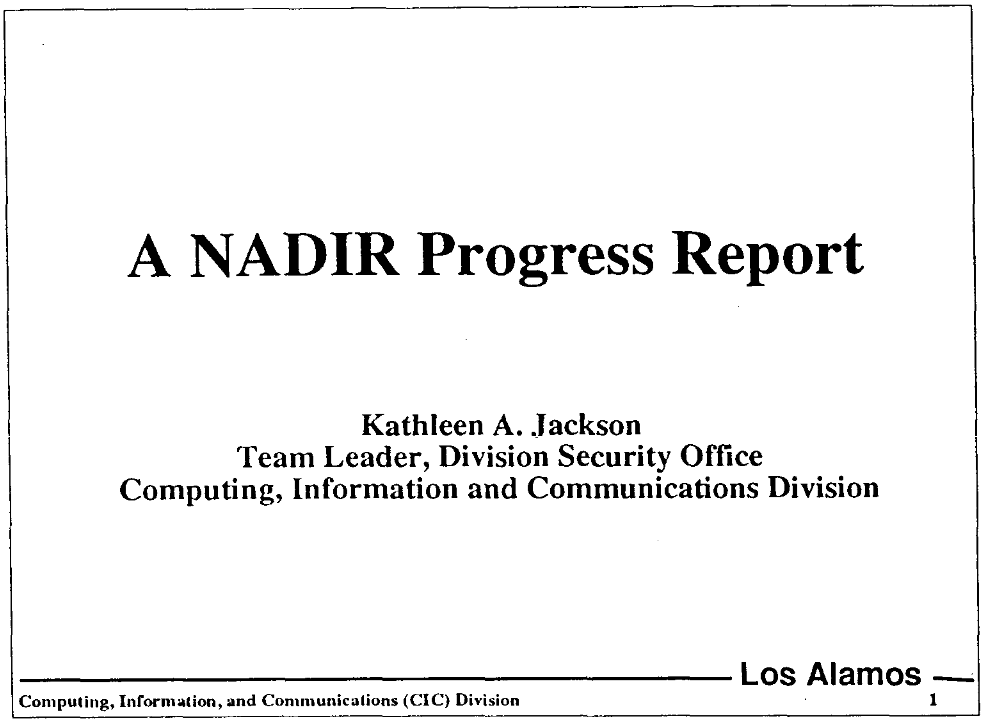
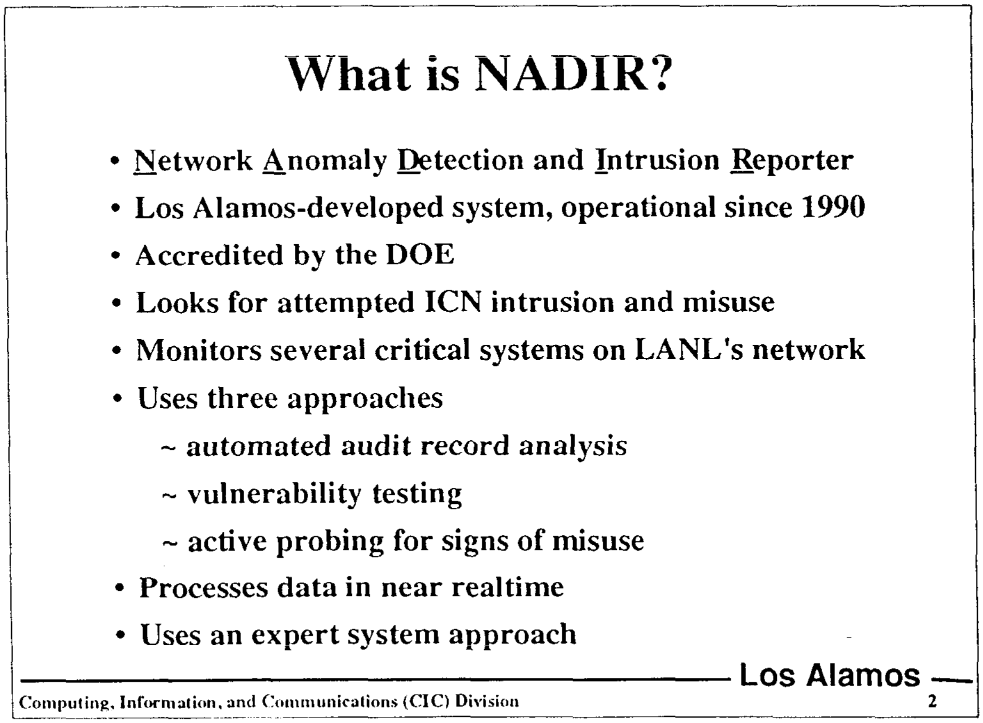


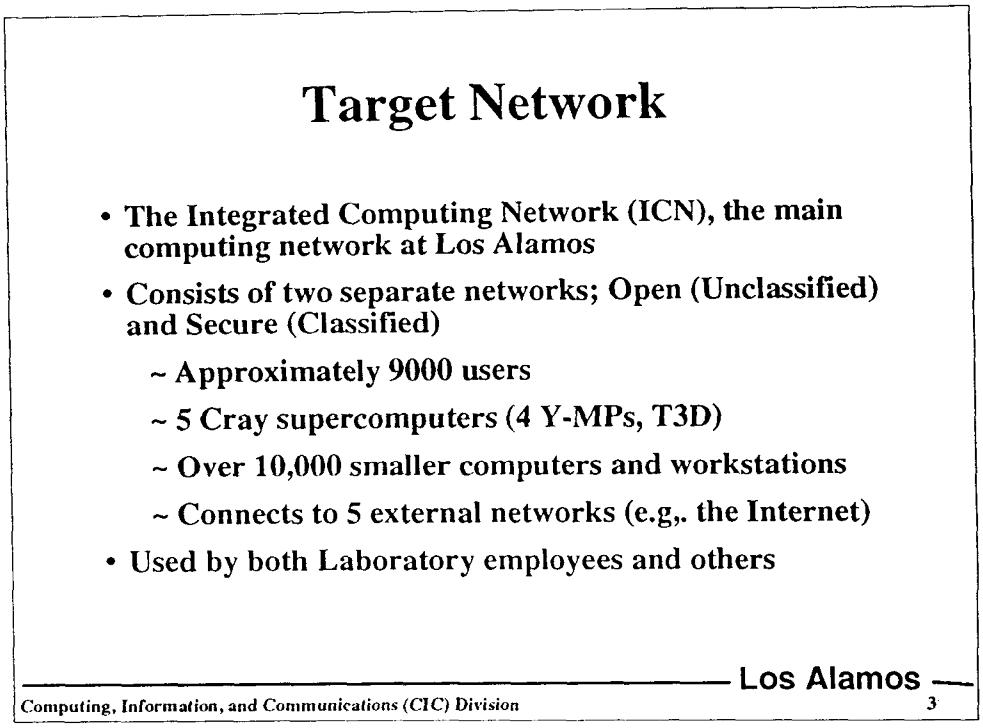
A NADIR Progress Report

Kathleen A. Jackson
Team Leader, Division Security Office
Computing, Information and Communications Division



What is NADIR?

- Network Anomaly Detection and Intrusion Reporter
- Los Alamos-developed system, operational since 1990
- Accredited by the DOE
- Looks for attempted ICN intrusion and misuse
- Monitors several critical systems on LANL's network
- Uses three approaches
  - ~ automated audit record analysis
  - ~ vulnerability testing
  - ~ active probing for signs of misuse
- Processes data in near realtime
- Uses an expert system approach



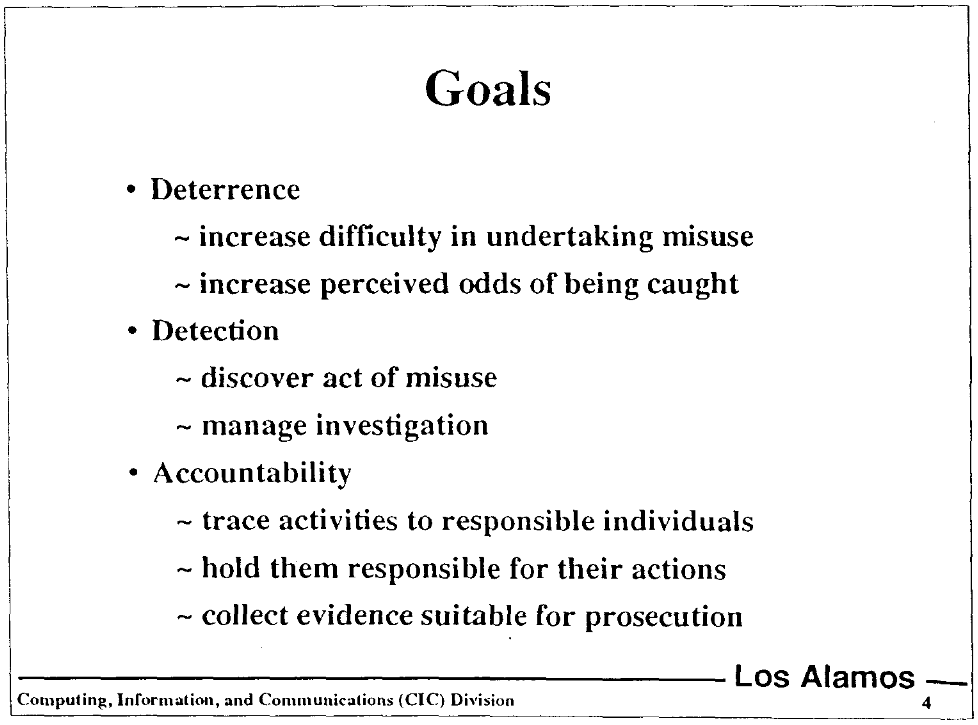
Target Network

- The Integrated Computing Network (ICN), the main computing network at Los Alamos
- Consists of two separate networks; Open (Unclassified) and Secure (Classified)
    ~ Approximately 9000 users

    ~ 5 Cray supercomputers (4 Y-MPs, T3D)

    ~ Over 10,000 smaller computers and workstations

    ~ Connects to 5 external networks (e.g,. the Internet)
- Used by both Laboratory employees and others



Goals

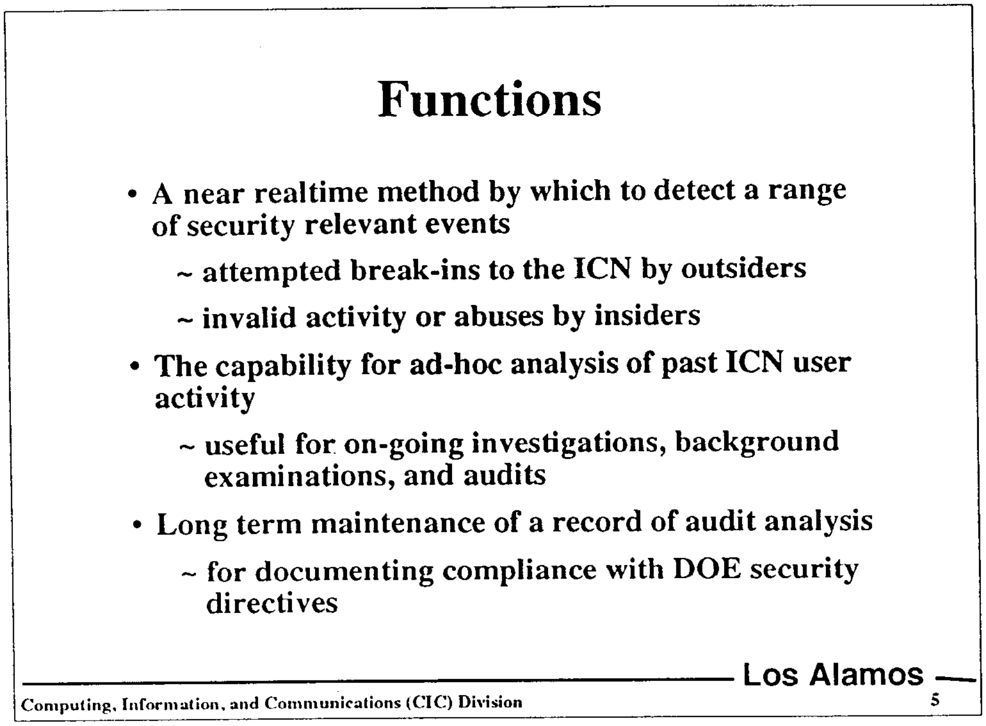
- Deterrence
    ~ increase difficulty in undertaking misuse

    ~ increase perceived odds of being caught
- Detection
    ~ discover act of misuse

    ~ manage investigation
- Accountability
    ~ trace activities to responsible individuals

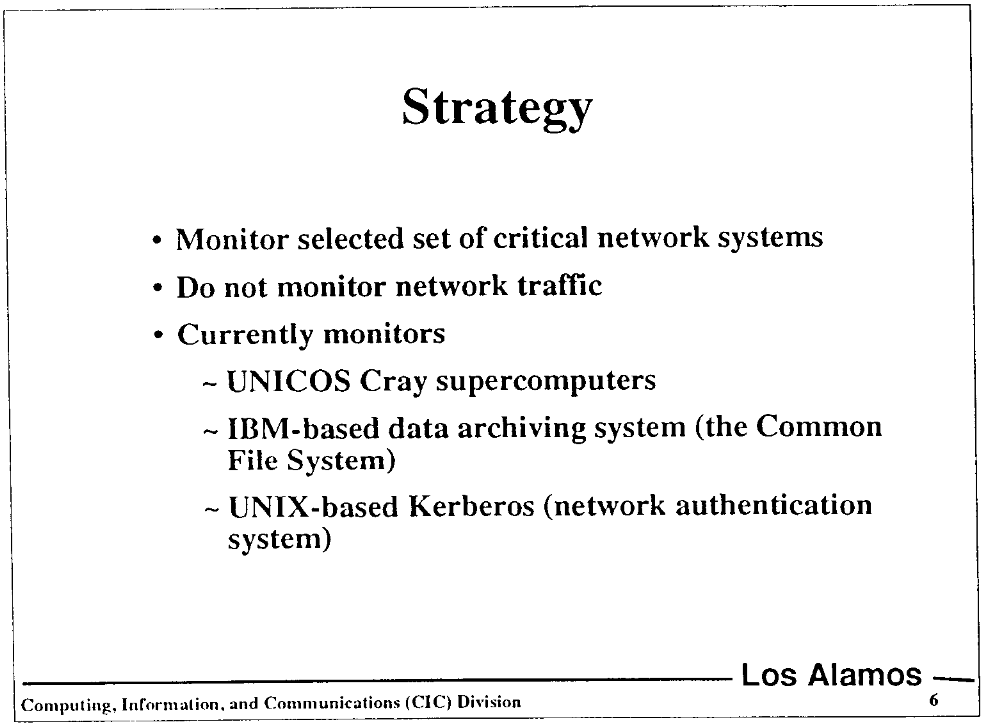
    ~ hold them responsible for their actions

    ~ collect evidence suitable for prosecution



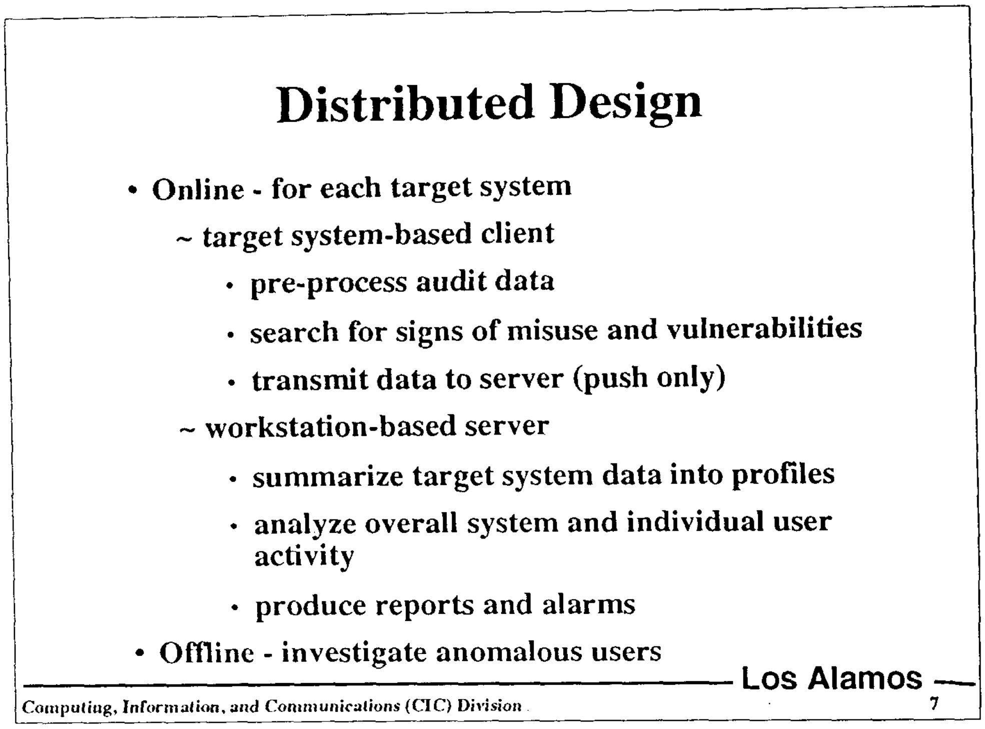
Functions

- A near realtime method by which to detect a range of security relevant events
  - ~ attempted break-ins to the ICN by outsiders
  - ~ invalid activity or abuses by insiders
- The capability for ad-hoc analysis of past ICN user activity
  - ~ useful for on-going investigations, background examinations, and audits
- Long term maintenance of a record of audit analysis
  - ~ for documenting compliance with DOE security directives



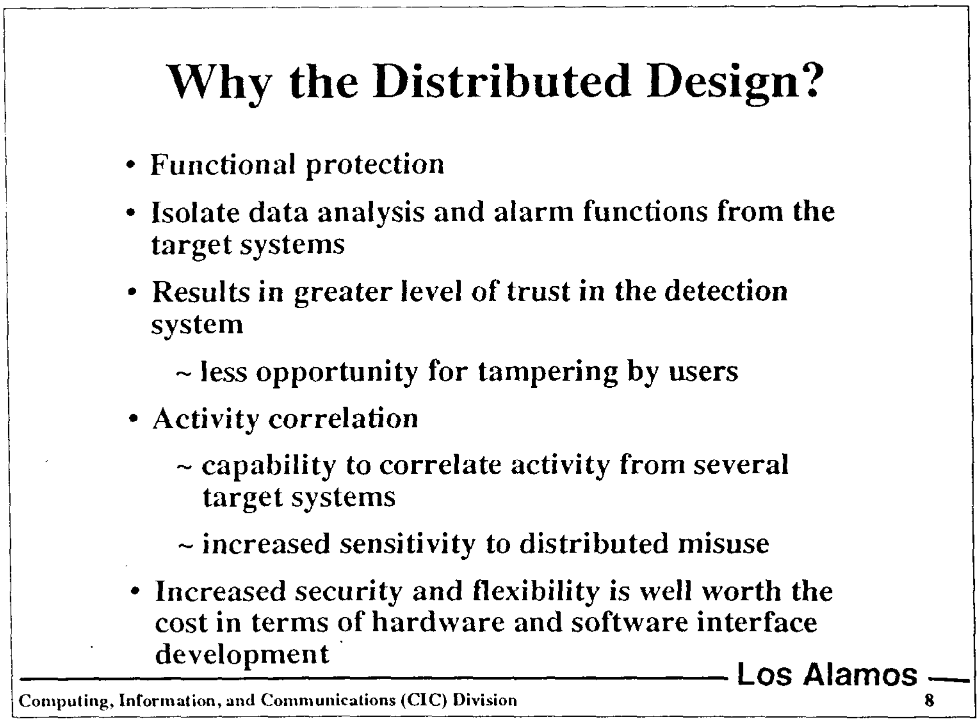
Strategy

- Monitor selected set of critical network systems
- Do not monitor network traffic
- Currently monitors
  - ~ UNICOS Cray supercomputers
  - ~ IBM-based data archiving system (the Common File System)
  - ~ UNIX-based Kerberos (network authentication system)



Distributed Design

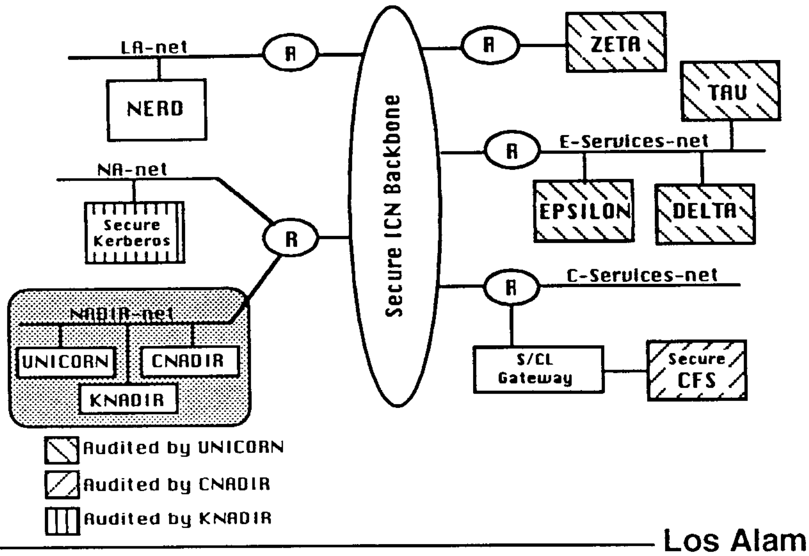
- Online - for each target system
  - ~ target system-based client
    - pre-process audit data
    - search for signs of misuse and vulnerabilities
    - transmit data to server (push only)
  - ~ workstation-based server
    - summarize target system data into profiles
    - analyze overall system and individual user activity
    - produce reports and alarms
- Offline - investigate anomalous users



Why the Distributed Design?

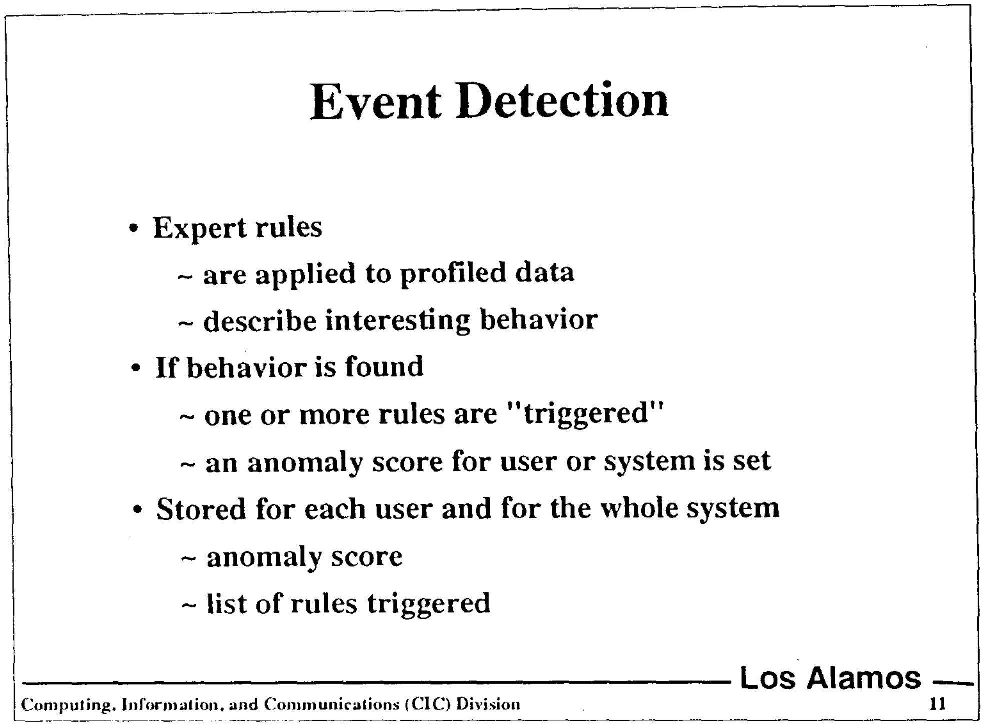
- Functional protection
- Isolate data analysis and alarm functions from the target systems
- Results in greater level of trust in the detection system
  - ~ less opportunity for tampering by users
- Activity correlation
  - ~ capability to correlate activity from several target systems
  - ~ increased sensitivity to distributed misuse
- Increased security and flexibility is well worth the cost in terms of hardware and software interface development

NADIR in the Secure ICN



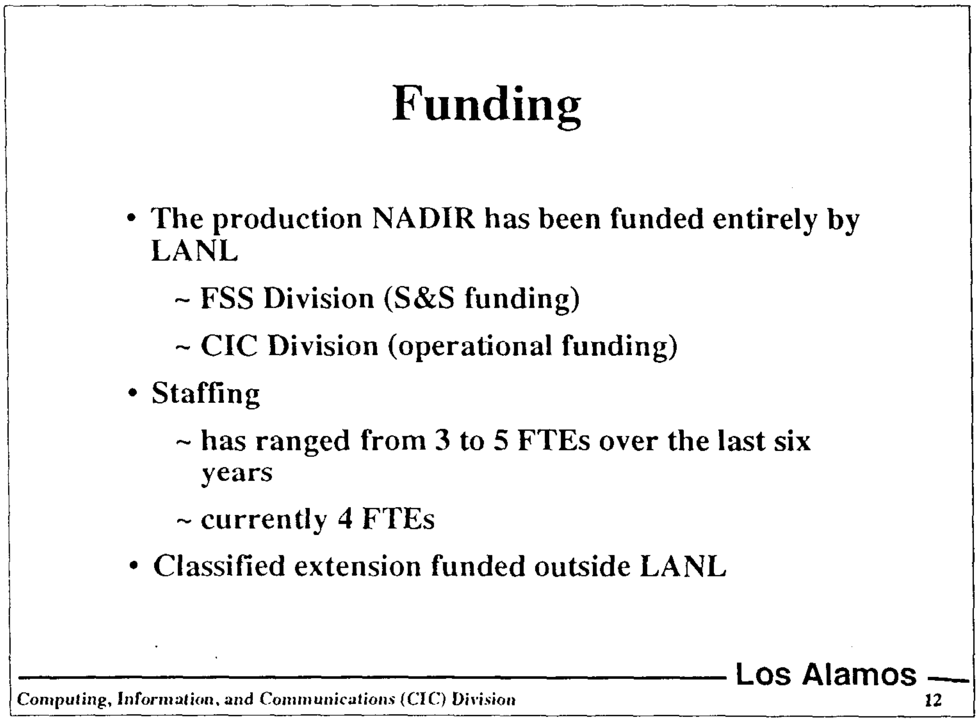
Profiles

- Profiles provide a statistical summary of activity on each target system
- Individual user profiles
  - ~ one for each system user
  - ~ activity that can be attributed to that user
- Composite (system) profile
  - ~ one for each system
  - ~ combination of all user activity on the system
  - ~ misuse not attributable to a single user
  - ~ vulnerable configuration information



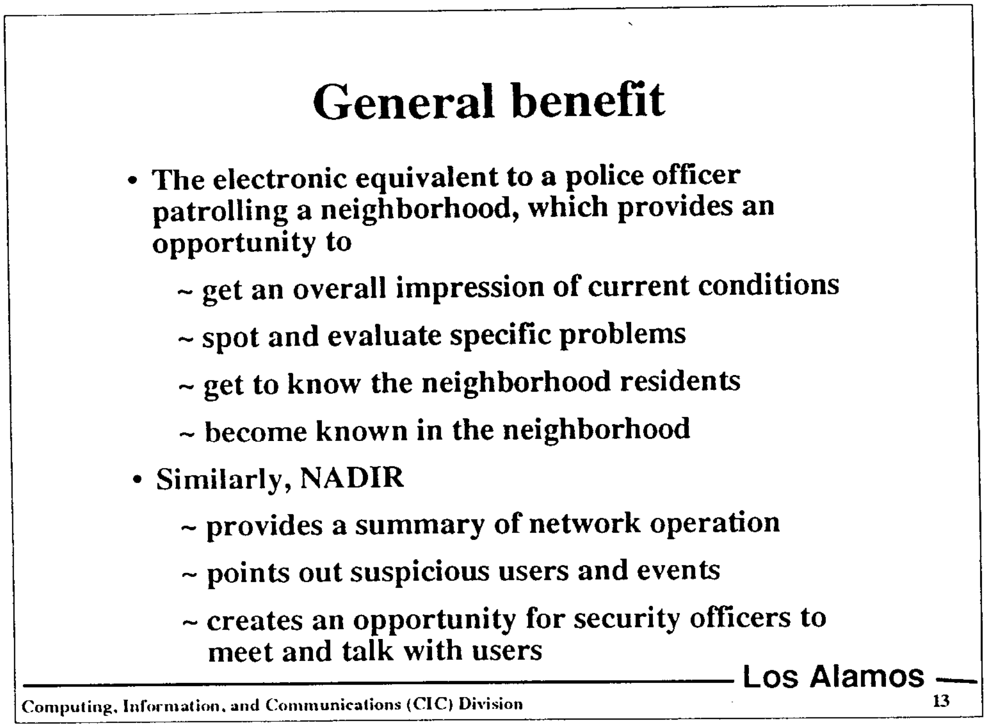
Event Detection

- Expert rules
  - ~ are applied to profiled data
  - ~ describe interesting behavior
- If behavior is found
  - ~ one or more rules are "triggered"
  - ~ an anomaly score for user or system is set
- Stored for each user and for the whole system
  - ~ anomaly score
  - ~ list of rules triggered



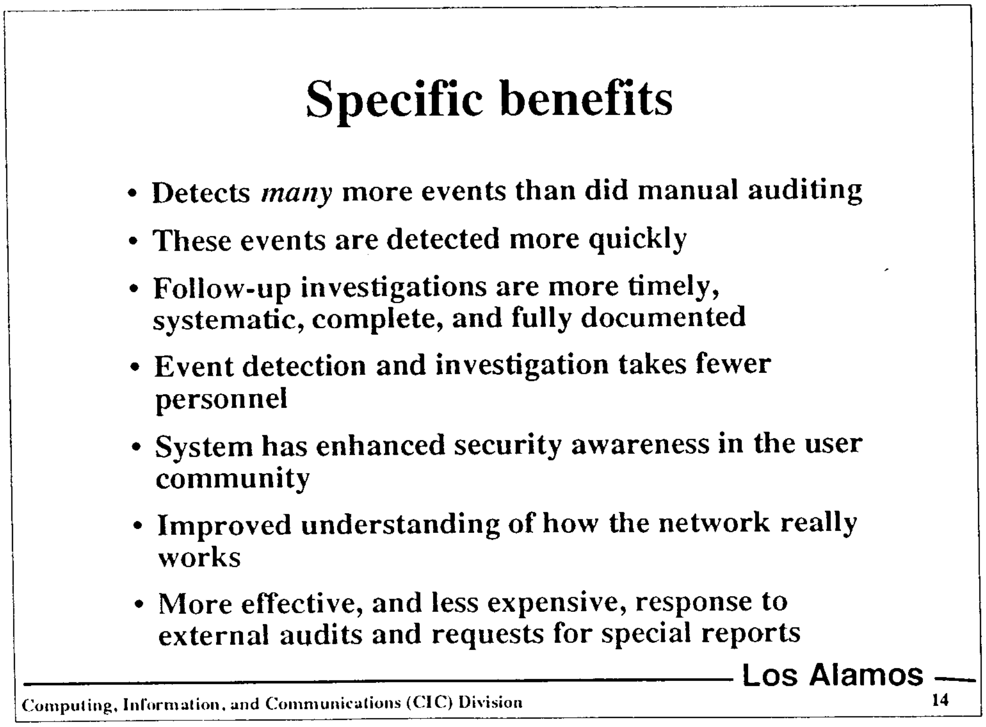
Funding

- The production NADIR has been funded entirely by LANL
  - ~ FSS Division (S&S funding)
  - ~ CIC Division (operational funding)
- Staffing
  - ~ has ranged from 3 to 5 FTEs over the last six years
  - ~ currently 4 FTEs
- Classified extension funded outside LANL



General benefit

- The electronic equivalent to a police officer patrolling a neighborhood, which provides an opportunity to
  - ~ get an overall impression of current conditions
  - ~ spot and evaluate specific problems
  - ~ get to know the neighborhood residents
  - ~ become known in the neighborhood
- Similarly, NADIR
  - ~ provides a summary of network operation
  - ~ points out suspicious users and events
  - ~ creates an opportunity for security officers to meet and talk with users



Specific benefits

- Detects *many* more events than did manual auditing
- These events are detected more quickly
- Follow-up investigations are more timely, systematic, complete, and fully documented
- Event detection and investigation takes fewer personnel
- System has enhanced security awareness in the user community
- Improved understanding of how the network really works
- More effective, and less expensive, response to external audits and requests for special reports

Attack handling

- NADIR compares individual and composite activity to typical or valid activity
- Attacks that require frequent repetition are detected easily
  - ~ by comparing current usage to normal past usage
- It also recognizes violations of computer policies
  - ~ like improper accesses
  - ~ illegal combinations of events
- Second order anomalies, like being repeatedly being "almost interesting", are missed

False positives and negatives

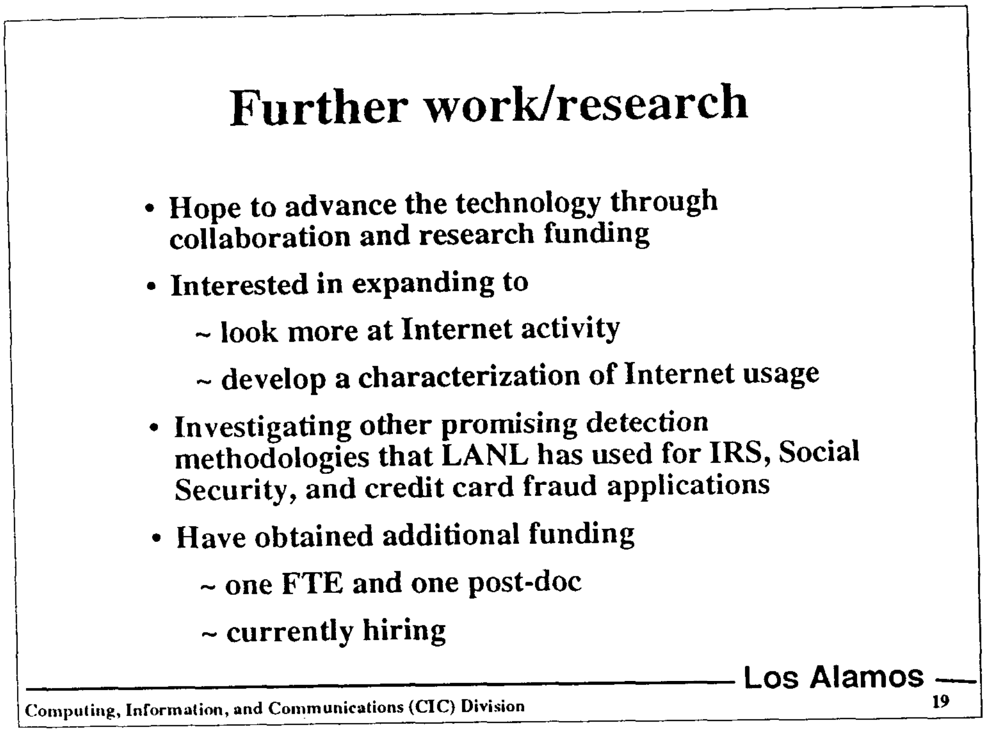
- How many false positives?
  - ~ few enough that they can easily be investigated and eliminated by a half-time investigator
  - ~ getting fewer
  - ~ invested a considerable effort to improve detection accuracy, using automated statistical tuning over a significant period of past usage
- False negatives are hard to prove
  - ~ we do not know of any significant event missed by NADIR (but found by other means) since the current system was implemented

Tuning

- NADIR was designed and tuned for the LANL user population
- We chose NOT to implement self-learning to avoid the potential weaknesses of that method
- We pre-characterize the user population, followed by periodic re-characterizations

Fielding the system

- Normal business constraints limit our ability to do everything we'd like to do
  - ~ i.e., we've never had the funding to all we'd like
- Development/maintenance costs are on-going and seemingly never ending
  - ~ monitored systems constantly change
  - ~ five workstations must be maintained/upgraded etc.
  - ~ resource intensive (3 to 5 developers/administrators)
- Running costs are low
  - ~ 9000 users on eight network systems are monitored
  - ~ with one half-time investigator
- We have a proven, well-functioning system



Further work/research

- Hope to advance the technology through collaboration and research funding
- Interested in expanding to
  - ~ look more at Internet activity
  - ~ develop a characterization of Internet usage
- Investigating other promising detection methodologies that LANL has used for IRS, Social Security, and credit card fraud applications
- Have obtained additional funding
  - ~ one FTE and one post-doc
  - ~ currently hiring