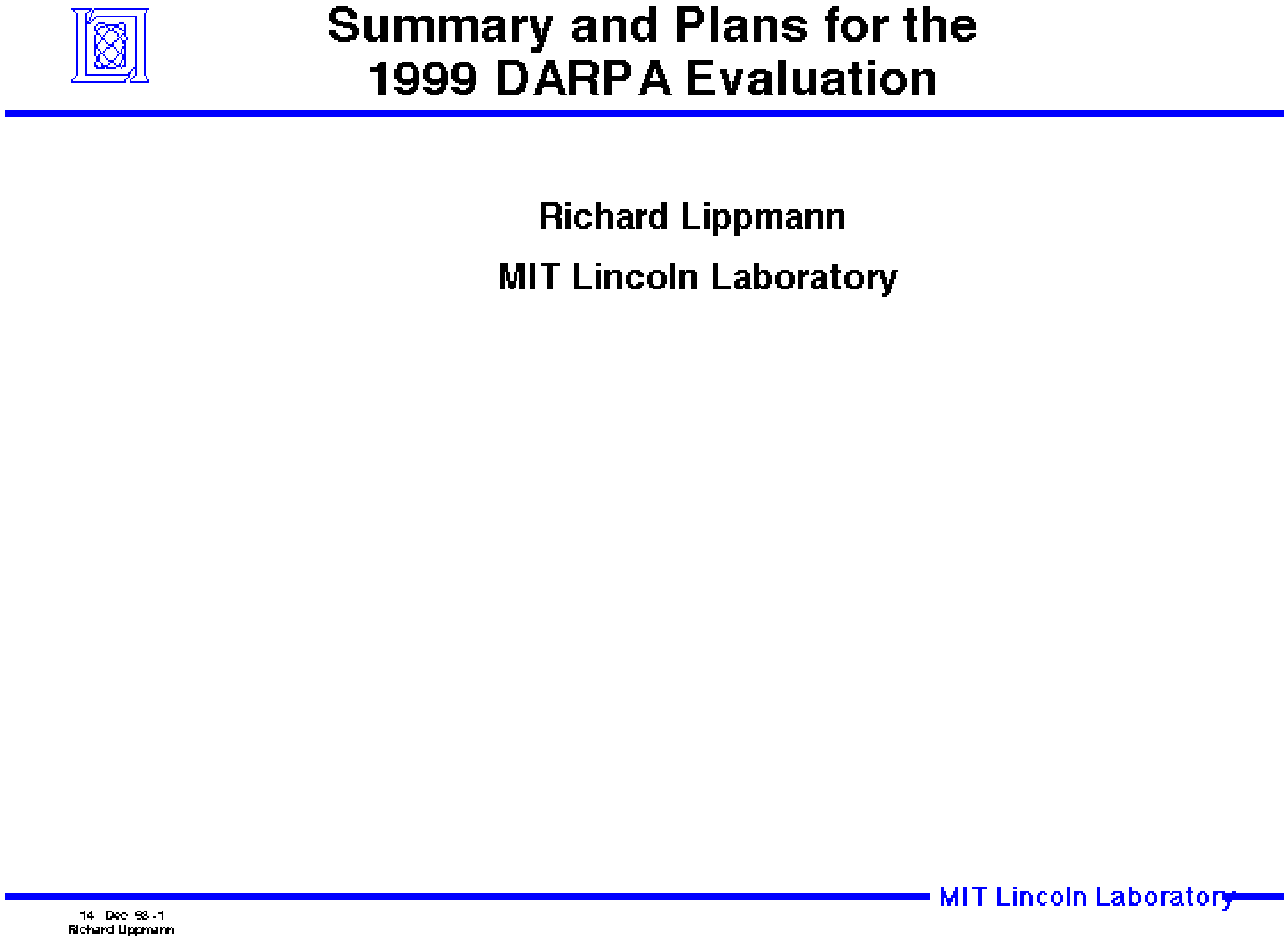


Summary and Plans for the 1999 DARPA Evaluation

Richard Lippmann

MIT Lincoln Laboratory

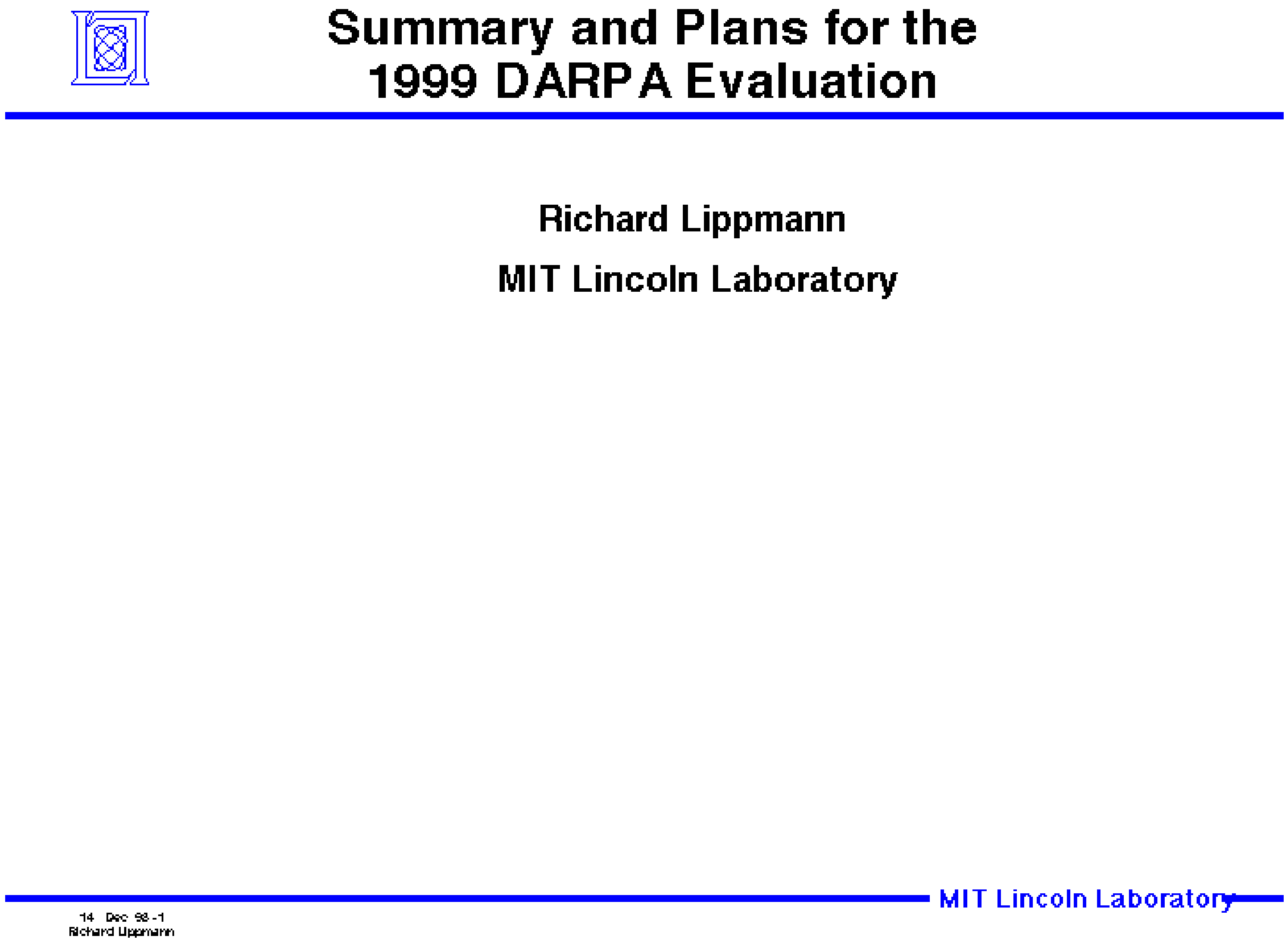
MIT Lincoln Laboratory

14 Dec 98 -1
Richard Lippmann

Slide 1 of 8

Notes:

In this talk I will first give an overview of the evaluation and then talk about traffic generation and attack development. I will then illustrate the type of data being generated, provide an example of a simple evaluation performed with this data, summarize the project, and discuss future plans.



Summary and Plans for the 1999 DARPA Evaluation

Richard Lippmann

MIT Lincoln Laboratory

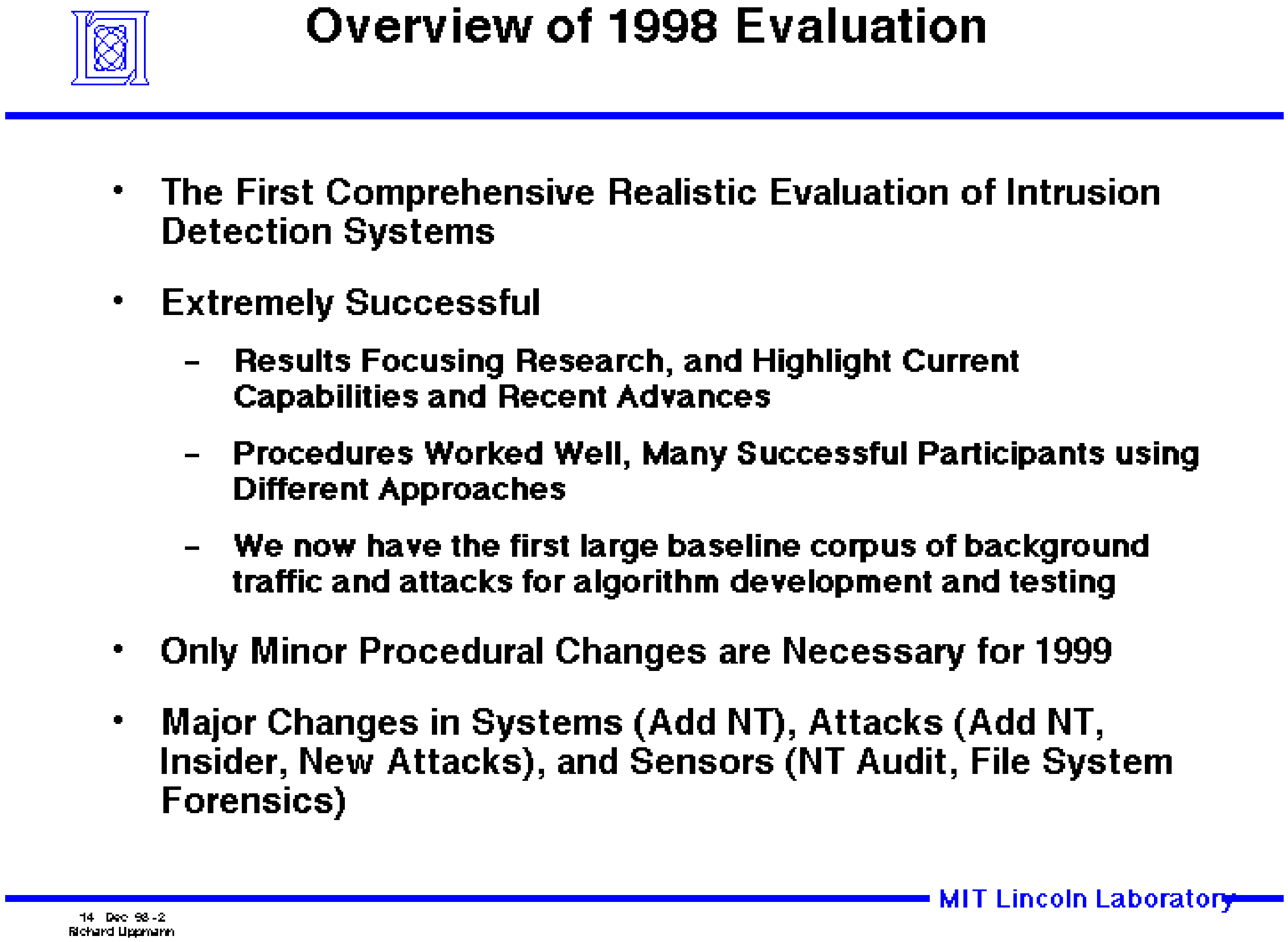
MIT Lincoln Laboratory

14 Dec 98 -1
Richard Lippmann

Slide 1 of 8

Notes:

In this talk I will first give an overview of the evaluation and then talk about traffic generation and attack development. I will then illustrate the type of data being generated, provide an example of a simple evaluation performed with this data, summarize the project, and discuss future plans.



Overview of 1998 Evaluation

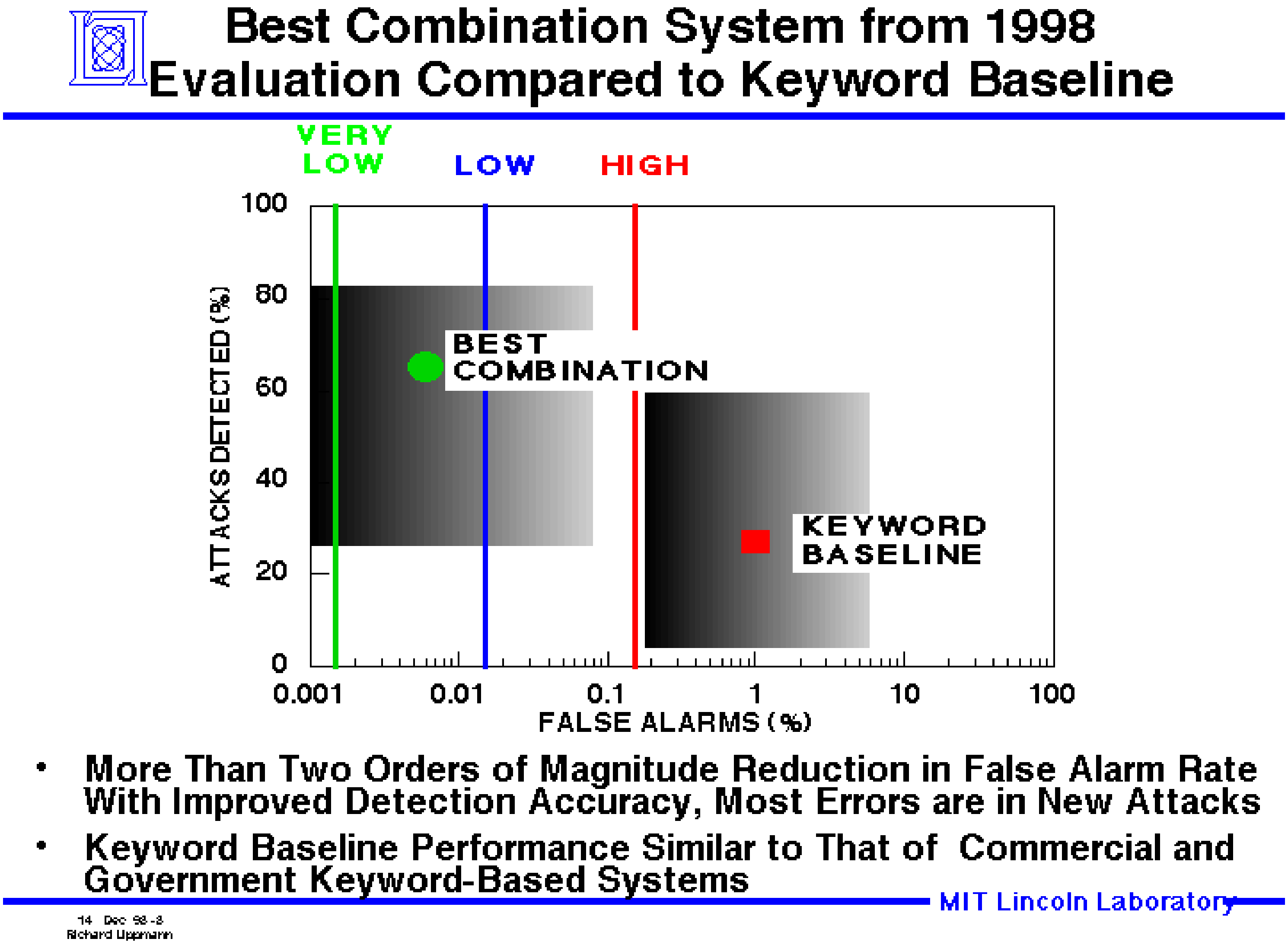
- The First Comprehensive Realistic Evaluation of Intrusion Detection Systems
- Extremely Successful
  - Results Focusing Research, and Highlight Current Capabilities and Recent Advances
  - Procedures Worked Well, Many Successful Participants using Different Approaches
  - We now have the first large baseline corpus of background traffic and attacks for algorithm development and testing
- Only Minor Procedural Changes are Necessary for 1999
- Major Changes in Systems (Add NT), Attacks (Add NT, Insider, New Attacks), and Sensors (NT Audit, File System Forensics)

MIT Lincoln Laboratory

14 Dec 98 -2
Richard Lippmann

Notes:

The 1998 evaluation was very successful in providing a realistic measuring stick for intrusion detection systems. The process has succeeded in focussing research in areas in most need of development and it has revealed the merits and drawbacks of many ID systems that use vastly different methods of detection. Minor changes that are necessary for 1999 include addition of NT and some modifications to the scoring technique.



Best Combination System from 1998 Evaluation Compared to Keyword Baseline

- More Than Two Orders of Magnitude Reduction in False Alarm Rate With Improved Detection Accuracy, Most Errors are in New Attacks
- Keyword Baseline Performance Similar to That of Commercial and Government Keyword-Based Systems

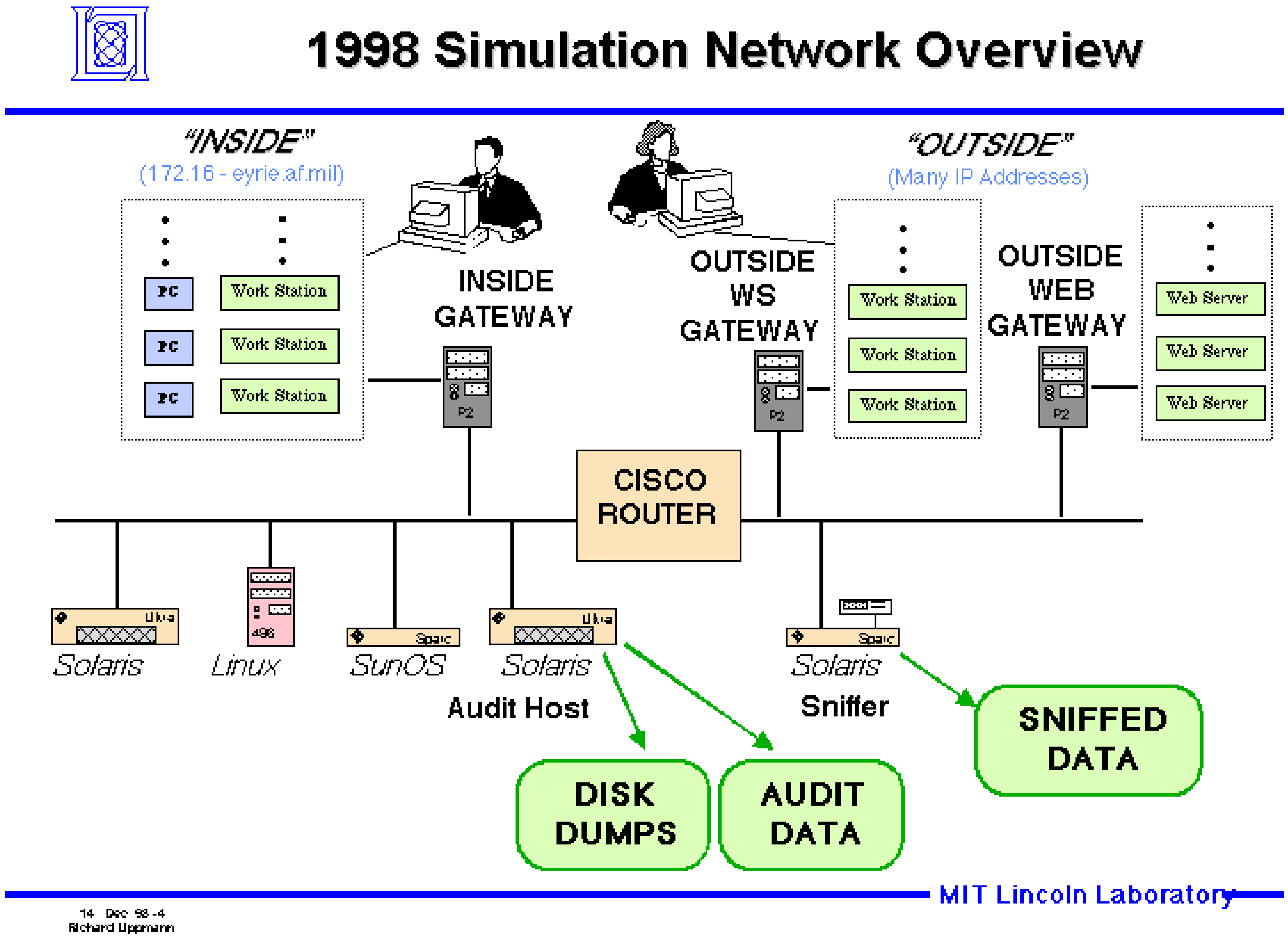
MIT Lincoln Laboratory

14 Dec 98 -3
Richard Lippmann

Slide 3 of 8

Notes:

The best combination of new generation systems has much better detection rates than the currently used keyword spotting systems, with false alarm rates more than two orders of magnitude lower.



1998 Simulation Network Overview

"INSIDE"
(172.16 - eyrie.af.mil)

PC
PC
PC
Work Station
Work Station
Work Station

INSIDE GATEWAY

P2

OUTSIDE WS GATEWAY

P2

"OUTSIDE"
(Many IP Addresses)

Work Station
Work Station
Work Station

OUTSIDE WEB GATEWAY

P2

Web Server
Web Server
Web Server

CISCO ROUTER

Ultra
Solaris

486
Linux

Sparc
SunOS

Ultra
Solaris
Audit Host

Sparc
Solaris
Sniffer

DISK DUMPS

AUDIT DATA

SNIFFED DATA

MIT Lincoln Laboratory

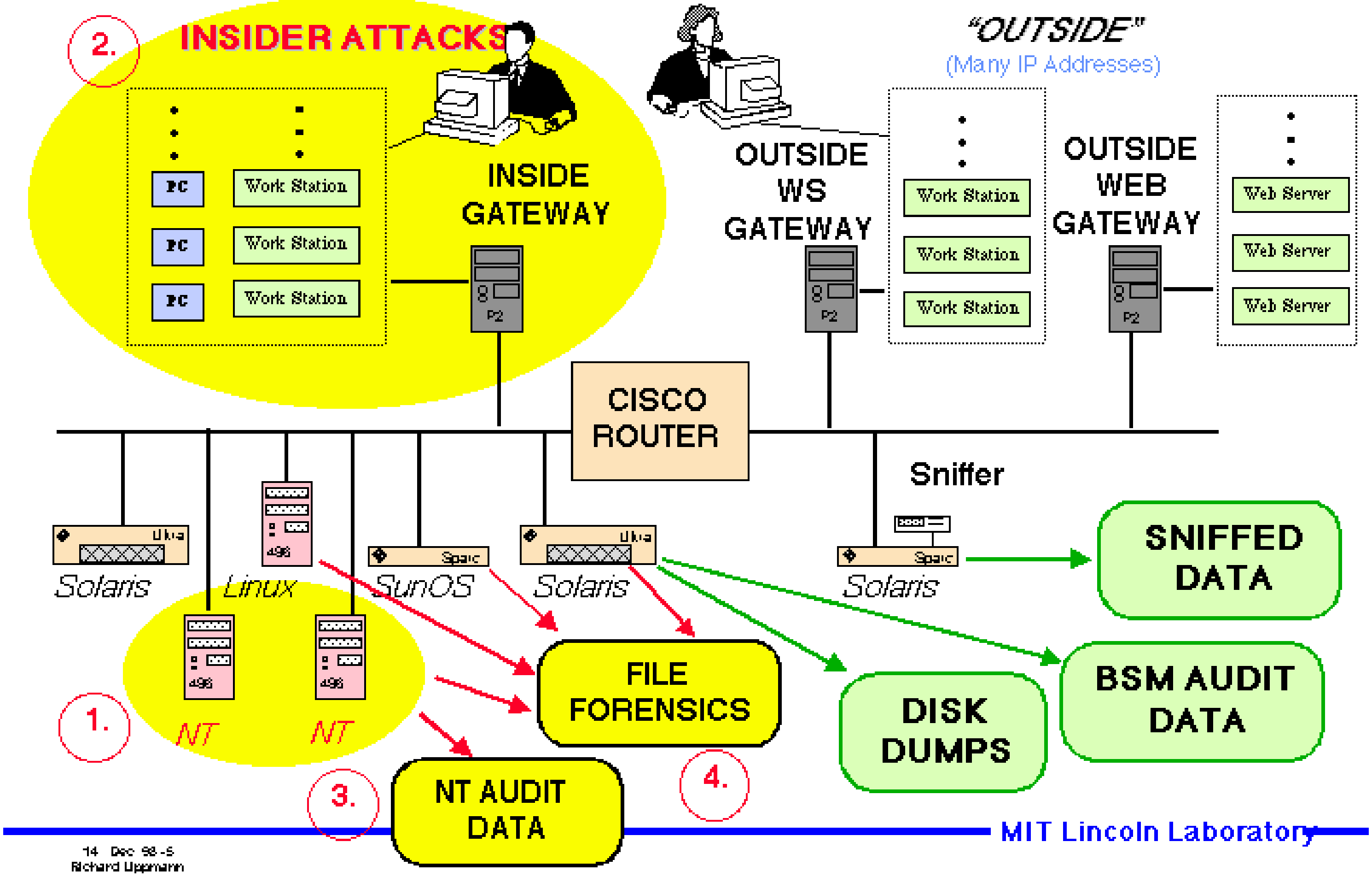
14 Dec 98 -4
Richard Lippmann

Slide 4 of 8

Notes:

The physical network used for the simulation included an inside and outside component separated by a router. The outside includes two workstations which simulate gateways to a virtual outside internet. One workstation simulates many workstations using custom software modifications of the Linux kernel provided by the Air Force ESC group. One gateway leads to roughly 100 workstations and the other leads to 1000’s of web sites with actual content that is updated daily. The inside includes victim machines of many types (e.g. Linux, Solaris, Sun OS) and a gateway to many other inside workstations. Data is collected from the inside victim running Solaris and from an outside sniffer.

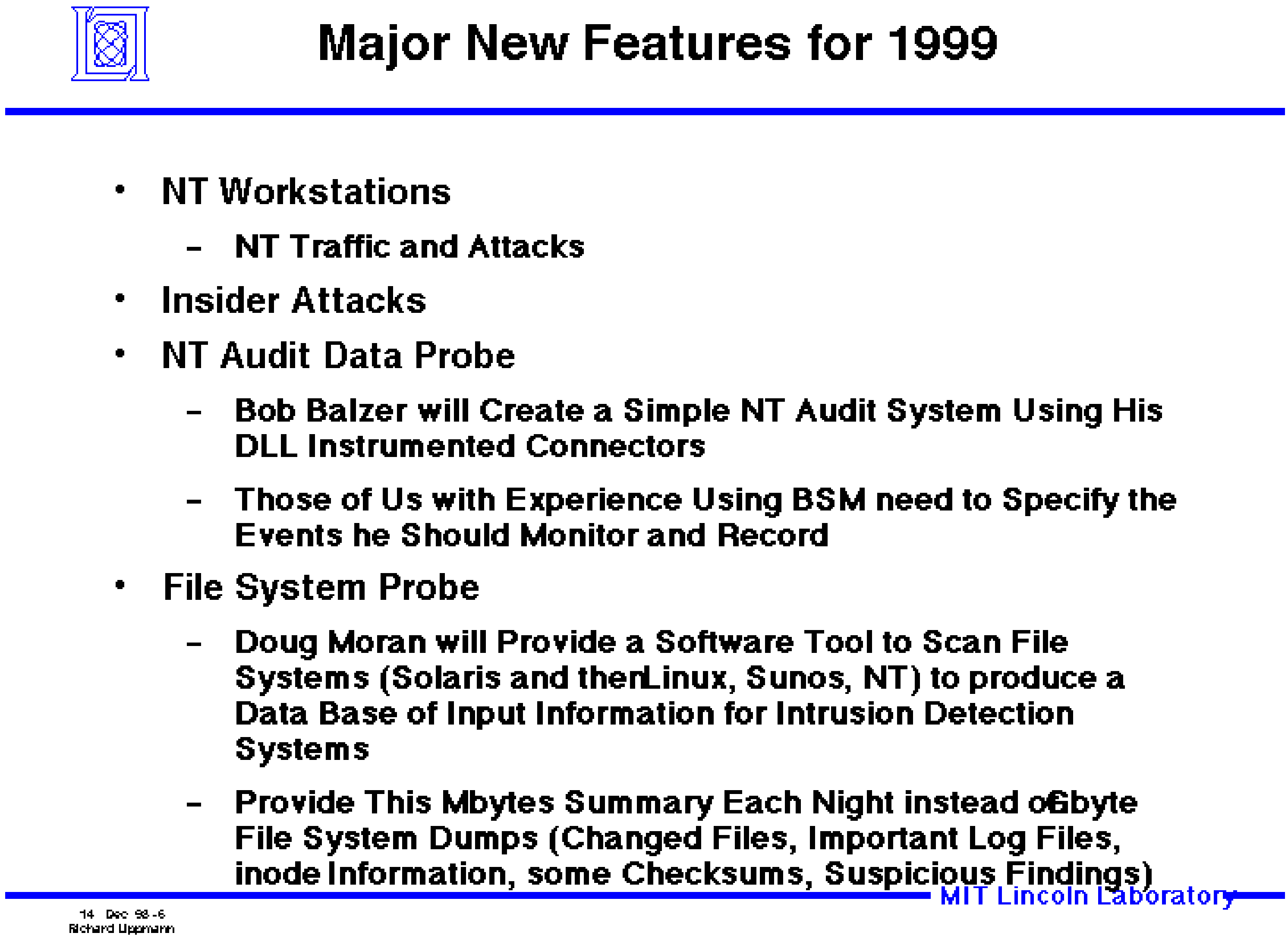
New Features in 1999 Evaluation



Slide 5 of 8

Notes:

The physical network used for the simulation included an inside and outside component separated by a router. The outside includes two workstations which simulate gateways to a virtual outside internet. One workstation simulates many workstations using custom software modifications of the Linux kernel provided by the Air Force ESC group. One gateway leads to roughly 100 workstations and the other leads to 1000's of web sites with actual content that is updated daily. The inside includes victim machines of many types (e.g. Linux, Solaris, Sun OS) and a gateway to many other inside workstations. Data is collected from the inside victim running Solaris and from an outside sniffer.



Major New Features for 1999

- NT Workstations
  - NT Traffic and Attacks
- Insider Attacks
- NT Audit Data Probe
  - Bob Balzer will Create a Simple NT Audit System Using His DLL Instrumented Connectors
  - Those of Us with Experience Using BSM need to Specify the Events he Should Monitor and Record
- File System Probe
  - Doug Moran will Provide a Software Tool to Scan File Systems (Solaris and thenLinux, Sunos, NT) to produce a Data Base of Input Information for Intrusion Detection Systems
  - Provide This Mbytes Summary Each Night instead ofGbyte File System Dumps (Changed Files, Important Log Files, inode Information, some Checksums, Suspicious Findings)

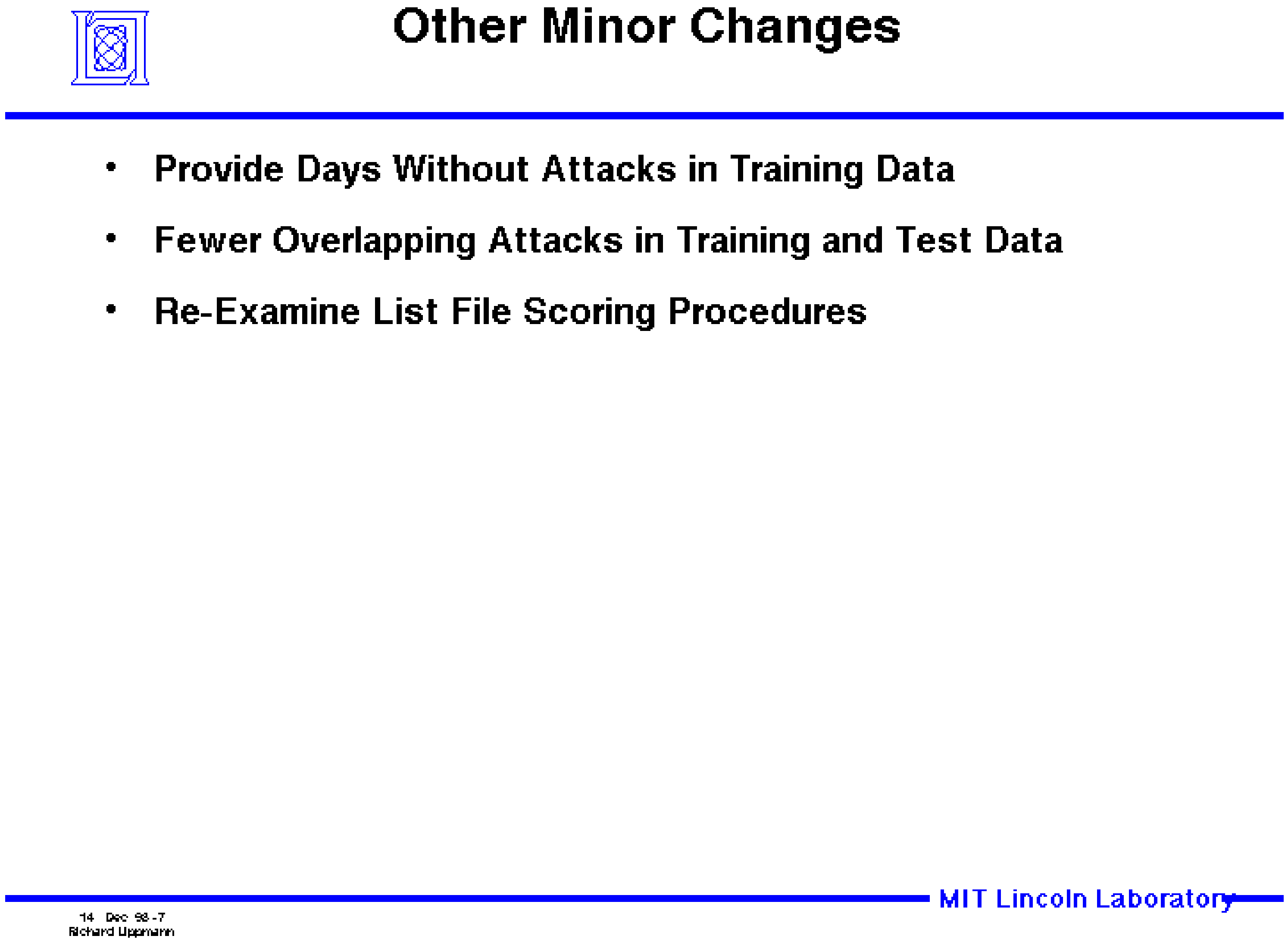
MIT Lincoln Laboratory

14 Dec 98 -6
Richard Lippmann

Slide 6 of 8

Notes:

Main new features for 1999 evaluation include NT work stations, insider attacks, and the addition of new sensors to provide different types of data.



Other Minor Changes

- Provide Days Without Attacks in Training Data
- Fewer Overlapping Attacks in Training and Test Data
- Re-Examine List File Scoring Procedures

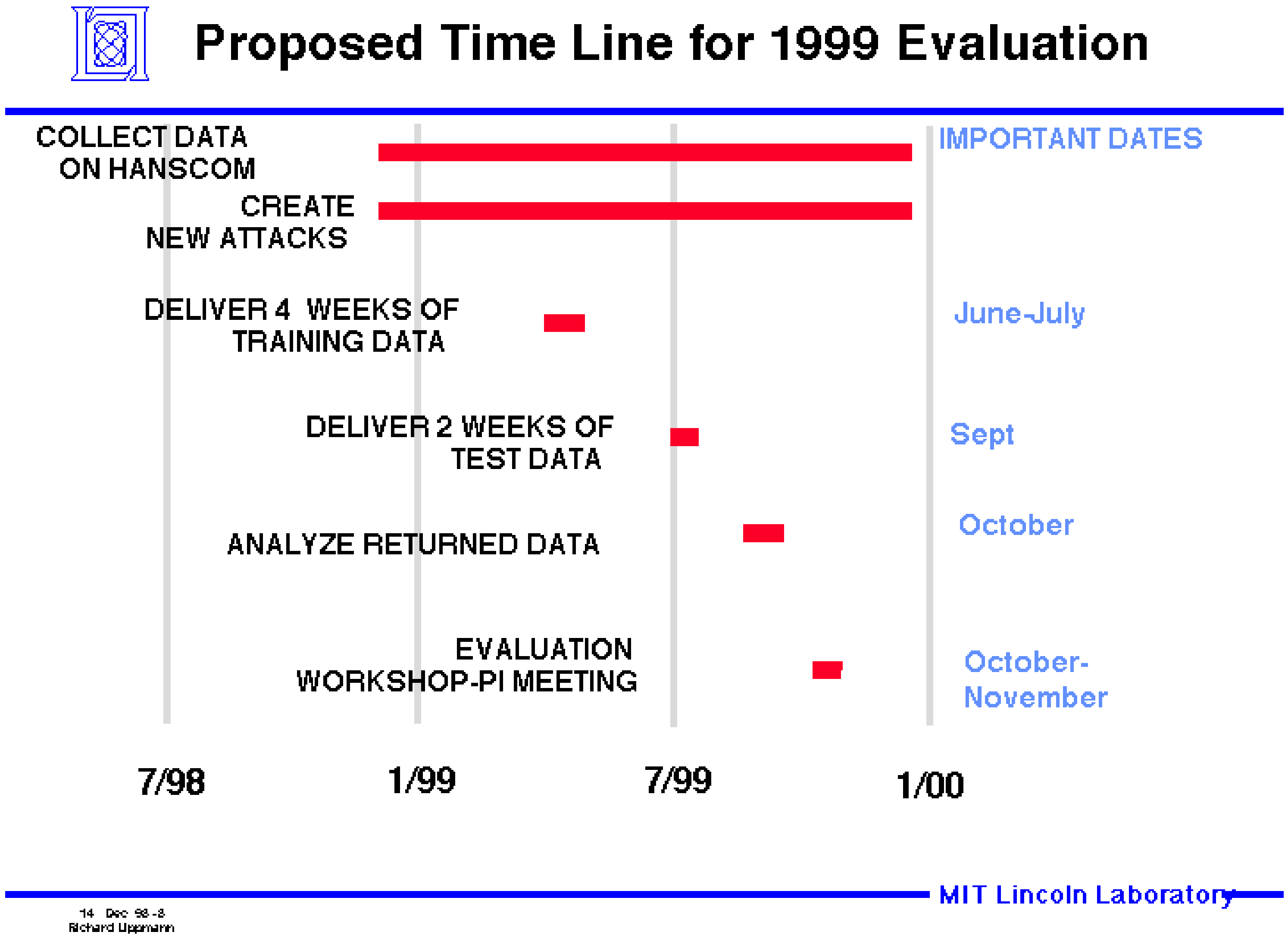
MIT Lincoln Laboratory

14 Dec 98 -7
Richard Lippmann

Slide 7 of 8

Notes:

Other minor changes include providing some days without attacks in the training data so anomaly detection systems can train more easily, and a reexamination of the scoring procedure.



Proposed Time Line for 1999 Evaluation

MIT Lincoln Laboratory

14 Dec 98 -8
Richard Lippmann

Slide 8 of 8

Notes:

This shows the time line for the evaluation. Important off-line components include the delivery of training data in July and August, the delivery of test data in September, and analysis of returned results in October.