



UNIVERSITY OF
SOUTHERN CALIFORNIA



LOCKHEED MARTIN



Successful Deployment of Mesh Routers for Trans-Continental Urban Resolve Experiments

Thomas D. Gottschalk, Robert F. Lucas, Gene Wagenbreth, John J. Tran and William E. Helfinstine
CACR - Caltech, ISI - Univ. of Southern California and Info. Systems - Lockheed Martin
tdg@cacr.caltech.edu, rflucas@isi.edu, genew@isi.edu, jtran@isi.edu and bill.helfinstine@lmco.com



Outline of Presentation

- **Problem and Background**
- **The Mesh Router Development**
- **Integrating the Mesh Router into Urban Resolve**
- **Generalizations for Fault Tolerant Operations**
- **Dynamic Router- Client Reconfiguration**
- **Shadow Routers**
- **Implementation issues**
- **Conclusion**

Problem and Background

- JFCOM J9 is conducting Urban Resolve
- Very large-scale, entity level battlefield simulation
- Joint Urban Operations involved human participants
- Geographically distributed across the United States
- Requires a communication network that is:
 - real time
 - interest-filtered
- Integration of the MeshRouter handles the WAN
- MeshRouter developed for use in the two Linux clusters
- DoD High Performance Computing Modernization Program
- Extending it for use in WAN brought opportunities to include:
 - Lower latency
 - Effective use of network bandwidth
- New set of challenges

The Mesh Router Development

- **MeshRouter provides interested-filtered exchanges**
- **Used in very large-scale battlefield simulations**
- **Important aspects of the framework are as follows:**
 - **Isolation of application-specific components**
 - **Carefully factored objects in overall software framework**
 - **Routers manage message traffic according to interest**
 - **Support for a wide variety of communication topologies**
 - **Links individual routers into full communications networks**
- **Following diagram shows tree router configuration**

Communications network for a Tree router network

TreeRouter Test Configuration

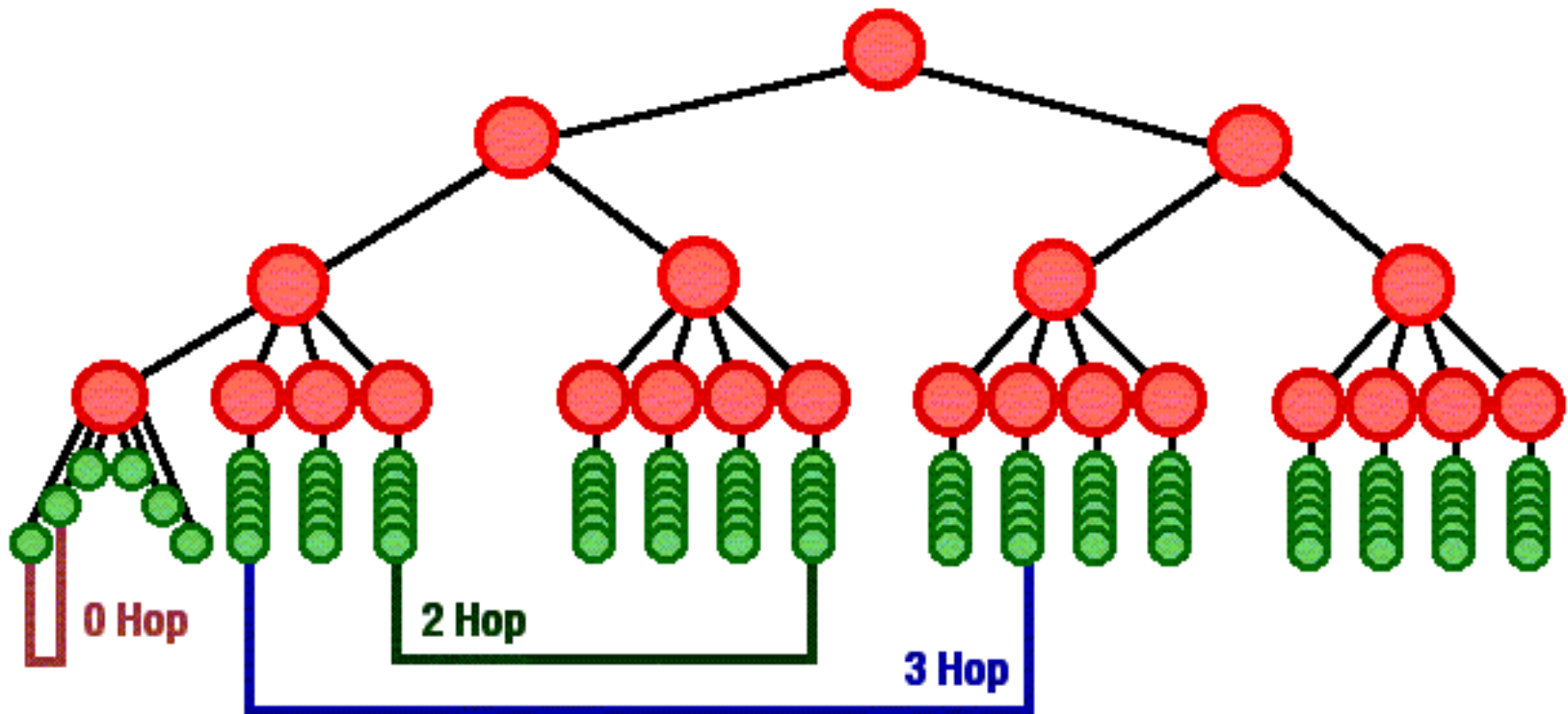


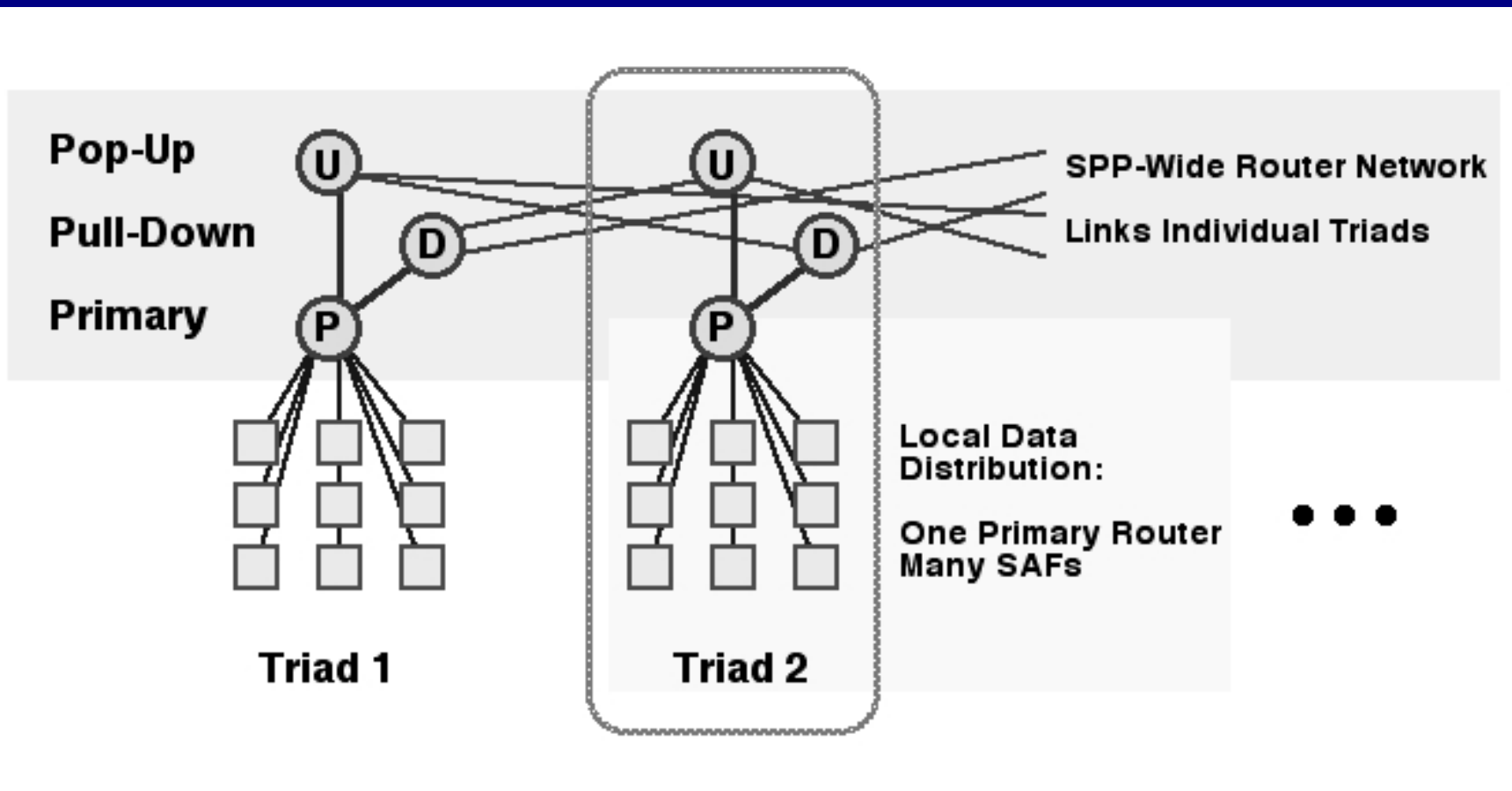
Diagram Explanation

- The “N-Hop” labels are a function of the number of routers used
- Application processes are organized into fixed groups
- They are associated with a given router
- All application communications go through low-level routers
- A hierarchical tree of additional router processes provides communications among the low-level {Router,Client} groups.
- Tree configurations have a number of standard problems
- Most notably, as the overall simulation size increases, the depth of the tree increases, meaning that:
 - The communication volume and processing load at the highest-level (root) router increases.
 - The longest communications path within the network, and hence latency, also increases.

SF Express as a Progenitor

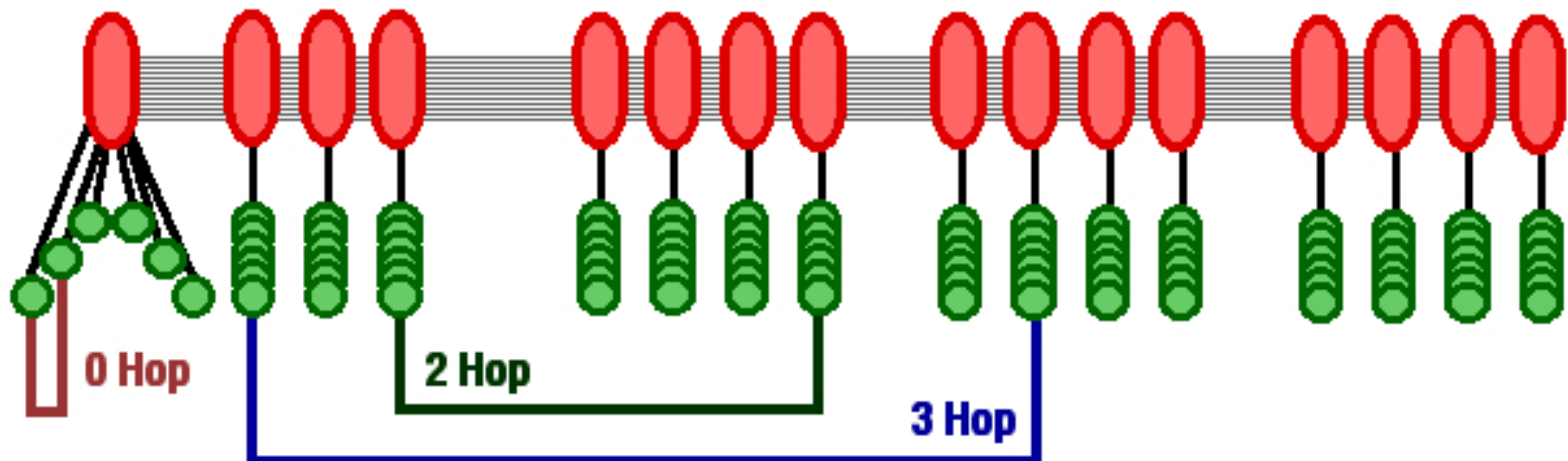
- **Alternative communications architecture is based on the “Synthetic Forces Express” (SFExpress)**
- **A scalable implementation of ModSAF**
- **Done for DARPA in 1995-1997**
- **Collections of SAF simulators exchange messages**
- **Specified Primary Router processors**
- **Additional “Pop-Up” and “Pull-Down” router layers**
- **Provided scalable, interest-limited messages**
- **SFExpress network is illustrated on the next slide**

Schematic illustration of the SFExpress communications architecture



One dimensional mesh topology within the MeshRouter architecture

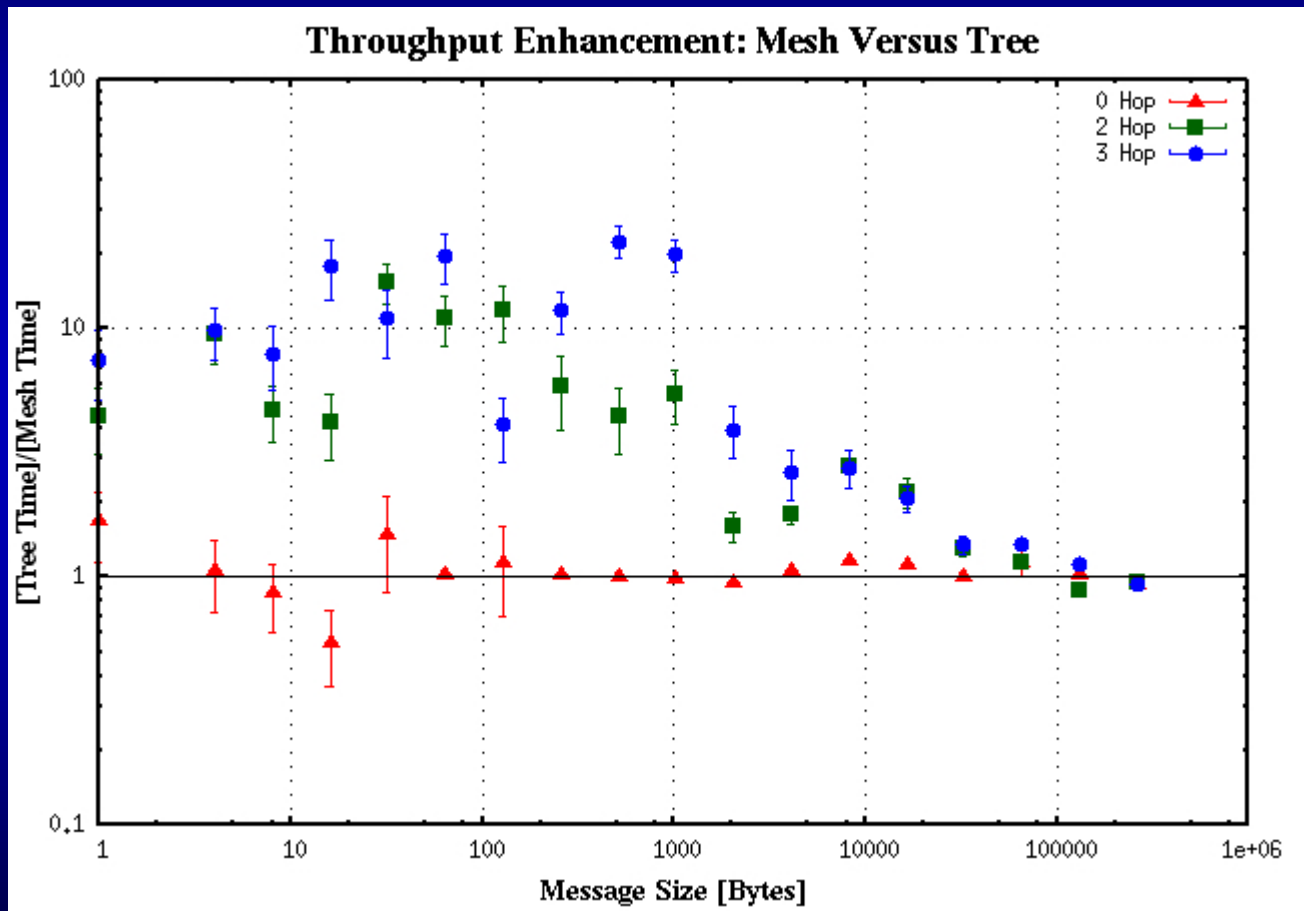
MeshRouter Test Configuration



Mesh Router Performance

- **Architecture is demonstrably scalable**
- **When connected in a one-dimensional mesh**
 - maximum number of communications links is three
 - independent of the overall simulation problem size
- **MeshRouter framework supports a straightforward generalization of the SFExpress framework**
- **See next slide**
- **Three distinct routers for the triads instantiated as three distinct router processes**
- **Performance comparisons for the Tree Router and Mesh Router**
- **Studies involved comparisons of message exchange times**
- **Observed ratios of message times are shown in next slide**
- **Improved performance due to the shorter message paths**

Ratios (Tree/Mesh) of mean communications times versus message sizes



Integrating the Mesh Router into Urban Resolve

- **WAN used for Urban Resolve experiments shown in next slide**
- **Three primary sites are:**
 - JFCOM in Suffolk, VA,
 - TEC at Ft. Belvoir, VA
 - SPAWAR in San Diego, CA
- **Also two Linux clusters at:**
 - ASC in Ohio
 - MHPCC in Hawaii
- **Five sites are connected by a high-level tree, rooted at JFCOM site-specific tree networks.**
- **Not feasible to join all the processors at the five sites**
 - Only a limited number of (secure) physical communications links joining the distinct sites.
 - Only a small number of the individual processors within the SPP assets can be accessed directly from the outside world.

Software routing topology for the urban Resolve experiments



Generalizations for Fault Tolerant Operations

- Router networks are not fault-tolerant
- Failures of router processes disrupt communications
- For the tree-router failures at root router divide simulation
- Failures on a mesh-router triad are more localized
- Two distinct strategies for tolerating router failures:
 - Strategy 1: Dynamic Reconfiguration.
 - Strategy 2: Shadow Routers
- Reconsider and generalize manner links are
 - specified
 - initialized

Dynamic Router- Client Reconfiguration

- Triad/router nodes are aware of already connected clients
- Simplest procedure for the director:
 - Assign the newly received communications request
 - Use the least-loaded triad within the system.
- There are several obvious advantages to this procedure:
 - No backup router triads (“hot spares”)
 - Provides dynamic balancing of the communications load
- Director processed in this scheme
 - Must be moderately “intelligent”
 - Receiving and maintaining health and load information

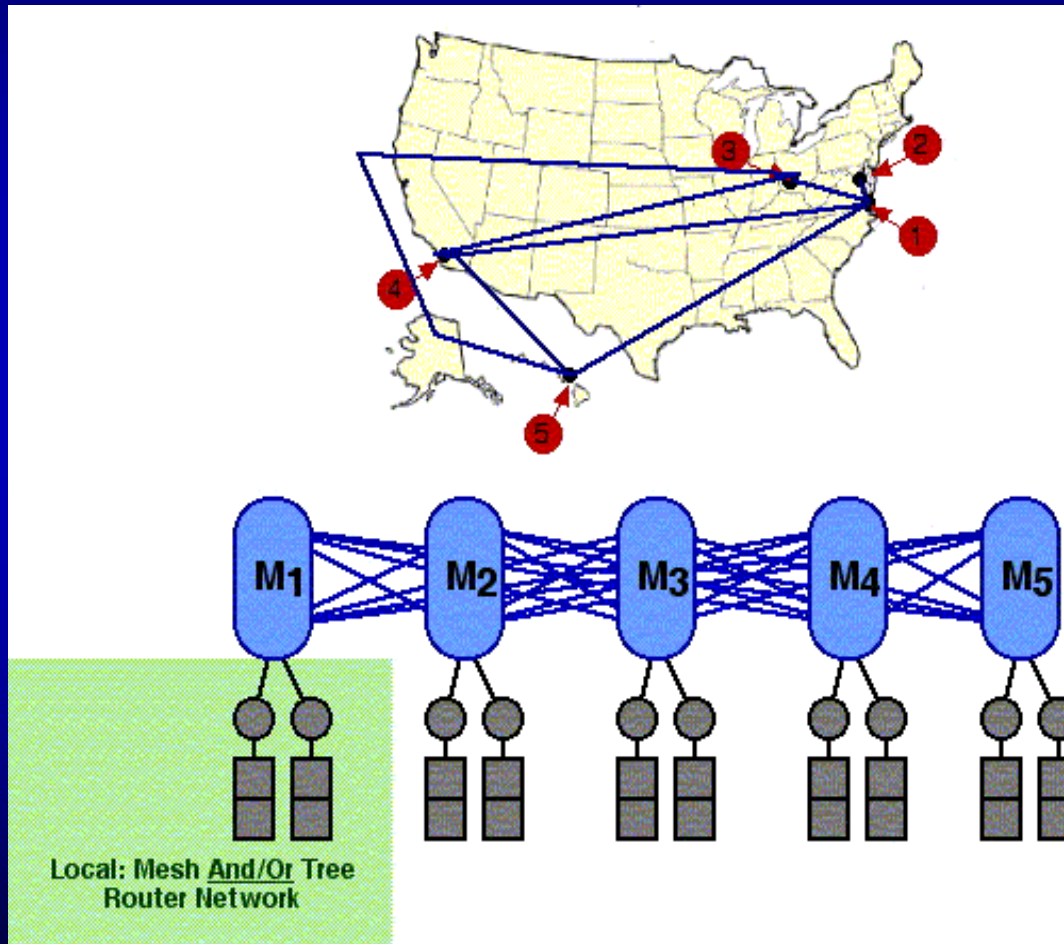
Shadow Routers

- Previous model not applicable for WAN communications
- The reasons are obvious:
 - Limited number of physical communication links
 - An SPP has only a limited processors that can be remotely accessed
- In these situations, the nominal Site \leftrightarrow Triad associations of the nominal network must be maintained
- Easiest implementation of fault-tolerant communications comes through
 - Replicated/redundant triads at each site
 - Known as Shadow Routers

Implementation issues

- **Strategy for fault tolerant networks involves both**
 - **Dynamic reconfiguration**
 - **Shadow routers**
- **Initial challenge involves extending RTI-s to recognize**
 - **Directors**
 - **Shadow routers**
- **Client has to reconnect following the loss of its router**
- **Has to declare its entire interest space.**
- **Once changes are made, will have fault tolerance**
- **RTI “Fire and forget” message passing paradigm means**
 - **There will be packet losses when routers fail**
 - **The number should be small**

Schematic illustration of a hybrid "TransContinental Mesh" communications framework



Conclusions

- MeshRouter successfully integrated in Urban Resolve
- Operation of routers indistinguishable from previous network
- Mesh enables J9 to maximize use of network bandwidth
- Reducing communication latency among participants
- MeshRouter and “tree” network now available
- Next improvement should be to make it fault tolerant
- Minimize loss of experiment time due to the failure of routers
- Dynamically reconfiguring router-clients suitable for LAN/SPPs
- Shadow routers appear to be a better solution for WAN

Government Support

This material is based on research sponsored by the Air Force Research Laboratory under agreement number FA8750-05-2-0204. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.