Leidos Data Exchange Framework (LDEF)
Leidos created the Leidos Data Exchange Framework (LDEF) to provide healthcare organizations with a robust software solution to develop, deploy, operate, and manage messaging and data exchange between disparate systems. Leidos took the latest and best open-source components available today and combined them into a single, integrated framework. The result is a powerful, flexible, and cost-effective alternative to the existing commercial interface engine/integration tools. Our unique integration and distributed deployment methodology of these open-source software components will provide healthcare organizations with the key capabilities they need for an interface engine solution at a reasonable price.

The design objectives for LDEF were to build a highly scalable solution that not only could easily grow to meet a healthcare organization’s increasing data exchange needs, but could also operate and be managed as a single, comprehensive system. With the increasing mission-critical nature of healthcare data exchange, the solution had to offer high-availability/fault-tolerant operation with automatic failover, and it had to be configurable to have no single point of failure. In addition, the system needed to be “cloud-ready” to give users the widest possible options for implementation. LDEF’s unique and highly distributed “nodal” deployment model allows it to address all of these important requirements.

As the name suggests, our solution is a framework—i.e., a set of concepts, tools, and components that together allow rapid, secure, and tested solutions for the creation of applications. LDEF was designed to manage all types of “message-based data”—a term we use in the broadest sense. In this context, a “message” is a payload of data. We support data types ranging from true message protocols (e.g., Health-Level 7 – HL7 and Clinical Document Architecture forms, or CDA) to data-transfer protocols (e.g., Secure File Transfer Protocol, or SFTP) to database protocols (e.g., SQL).

POSSIBLE USE CASES
LDEF was designed to provide solutions that are able to address a wide variety of data exchange use cases, including:

- **A data conversion engine** to reformat data extracts from one system and convert them into load files for another;
- **Batch or real-time system-to-system integration**;
- **An intelligent front end** to identify, transform, and send data to specialty systems like registries, quality improvement, and business analysis systems;
- **A gateway** to securely connect a sending or receiving system to external networks, including standard networks like eHealth Exchange via NwHIN CONNECT or Direct;
- **A data message broker** to intelligently receive, validate, and route messages from a central gateway to the appropriate receiving system; and
- **An extract, transform, and load (ETL) solution** that can parse messages and file the data extracted into the target database.
LDEF is deployed and managed from a version-controlled repository as a collection of virtual machine images. This repository may be physically located at Leidos, the client’s infrastructure, or both. This repository ensures version control down to the source code and library level for all released code, meaning that any release may be regenerated at any time—days, months, years—after the release and be identical in every way to the original instance. A link to the repository is as simple as a secure URL.

The deployed components of LDEF can be provided on Windows or Linux systems as run-time packages. These components may also be distributed as virtual images. The virtual image solution provides larger systems with greater flexibility in a post-deployment scenario that requires rapid and automated expansion for managing dynamic message loads.

Leidos understands that a client may already have an existing hardware and software footprint. Our ability to function across most hardware and operating systems makes LDEF more able to work within existing
footprints, allowing the client to minimize operational risks associated with footprint expansions. Proper sizing of an LDEF installation, relative to physical hardware and needed framework nodes, relies on the client’s use cases. This approach allows us to provide the correct amount of redundancy, failover, fault tolerance, and throughput for the client’s individual needs.

To illustrate the system requirements in our demonstration environment, we used LDEF to build an application to detect certain infections based on electronic lab results. For that use case, a single instance of each component deployed on a Dell Latitude Laptop with an i5 Intel chip and running SELinux CentOS 6.4 was able to receive 200 messages per second, evaluate which messages had C-Diff or MRSA, create CDA output for the identified cases, and record 5 instances of the message at the 5 different states of transformation for purposes of traceability.

In practice, we prefer to use virtual images with smaller granularity. This granularity means a higher number of virtual images than will ultimately be needed; however, the excess allows the Leidos team to fully understand the loading of the individual components, which in turn helps us correctly tailor the system configuration to the environment and reduce system resources down the road. The virtual images may be provided in one of four possible formats:

- SELinux Kernal Virtual Machine (KVM)
- Oracle VM VirtualBox
- VMware
- Raw Virtual Image – a standard image format.

Regardless of the number of images, LDEF includes a centralized management center that allows administrators to view and manage all images (i.e., the whole solution) as one system. The centralized approach, with our proactive monitoring, allows the administrator the freedom to effectively ignore the actual images and think of the complete system as a single component.

LDEF may be deployed in different arrangements. The three most common are:

- A “back-end” system completely behind the firewall;
- A system located entirely outside the firewall (i.e., in the “DMZ”) with National Security Agency (NSA)-level security; and
- A system with nodes both in the DMZ and behind the firewall.

Whatever version of LDEF the client selects, the basic system architecture remains the same. It also allows the combination of components, which may be replicated and/or reordered as needed to meet a client’s use case:

These components may be replicated and/or reordered as needed to meet a client’s use case:

- Gateway
- Queuing
- Worker
- Data Store
- Monitoring

The Leidos team will design a solution specifically tailored to the customer’s needs.

DISASTER RECOVERY / BUSINESS CONTINUITY

When it comes to messaging, updating the data store regularly is all that is needed to retain all states of the system. This plan allows LDEF to provide disaster recovery (DR) sites for continuity of business in near-real-time by providing “Cassandra” nodes in the production environment and at the synchronized DR site. These nodes
will be kept in sync by the run-time host until it becomes unavailable, in which case the DR site will resume operations at the current state of the production system. Designed with this type of use in mind, Cassandra provides rapid system-level communications using the Gossip protocol between active nodes.

**DATA FLOW ORCHESTRATION**

LDEF begins orchestrating a data flow by using one of over 200 different communications endpoints in its library to connect to the source system and acquire the data. LDEF coordinates the services used to validate, filter, and transform the data, then uses our rules engine to apply business rules (see Figure 2).

LDEF provides a queuing mechanism to assist in performance, monitoring, and data flow between nodes. In short, we can pull external data, receive external data in real-time on the wire, or read data from numerous source types internally. From this initial stream of data LDEF then orchestrates various flows of data applying business logic, transformations, analysis, and more. The data are then pushed to the target output system(s) using any endpoint protocol from our library as defined by the client in formats required by the target application. We are able to provide guaranteed message delivery (GMD) on the receiving and the exiting end-points for those clients able to realize the benefits of GMD.

LDEF is composed of the following major integrated functional components (see Figure 2):

- A **web portal** for system-wide monitoring and management, with separate operational and executive dashboards that provide real-time status of the system and possibly client applications that use the framework (as defined by the client’s use cases).
An enterprise service bus (ESB) that allows the integration of applications in complex landscapes while providing a strong client-server for distributed computing.

An orchestration engine with an integrated graphical development environment for constructing routes to manage message processing and data transformation through the system. These routes are used by “worker” nodes to produce defined functional processes within the framework. The worker nodes perform their work by calling Service Oriented Architect (SOA) services while defining workflows and providing parallel processing capability.

An extensive library of over 200 data communications endpoints to provide secure connectivity to a wide variety of new and legacy systems.

An integrated rules engine to provide a unified and system-wide platform for managing rules, workflow, event management, and complex event processing (CEP).

A comprehensive queuing and load-balancing component to manage message flows seamlessly across the distributed system.

A persistence component that uses a No-SQL database to provide audit logging, traceability, and synchronized access to data across the system.

Healthcare-specific messaging libraries and services to support the validation and transformation of standard healthcare message formats.

CENTRALIZED MANAGEMENT

In designing LDEF, our architects recognized one of the biggest limitations of many interface engine solutions: When multiple instances are needed to meet processing requirements, each instance of the engine must be monitored and managed individually. Regardless of the size of the system or number of interfaces deployed, LDEF makes the real-time status of both system performance and message-processing activity available throughout the system. The console provides two web-portal-based dashboards—one for operations staff and one for the business/executive staff.

The operations dashboard (Figure 3, next page) provides key system status and performance data on all system resources. Standard metrics include the following:

- Operational status of each node or virtual machine;
- CPU and memory use of each node;
- Number of messages processed at key points in the route;
- State and/or status of processes internal to nodes.

This list includes the metrics provided out-of-the-box; if the client desires, any metric available in the system can appear on the business or operational dashboards, including use-case-specific values and operating system information. The timescale for presenting this data can also be changed to focus on current status or to view trends over time. For more details on the customizable LDEF dashboard, please refer to Appendix A.
The business dashboard is similarly customizable and configurable to provide information to either to executives or other business users interested in a specific interface or use case. The system performance data are generally limited to an overall “red/yellow/green” status of the interface, but they can include as much detail as needed to address the organization’s monitoring requirements.

AUDITING

Our solution provides several audit features, all of which may be viewed by approved users on request through a graphical user interface:

- Logging
  - User login information
  - User activity – e.g., data modification, submitting of original information, attempts to use features outside of the user’s rights
  - Errors at the component, application, system levels

Figure 3. The Operational Dashboard Provides a Comprehensive View of the Entire System
LDEF retains a “versioned” copy of all transformations—i.e., when a message or inbound data stream is modified, we save the original as well as a copy of every new state of the information as it evolves during processing. This approach allows the client and Leidos staff to better understand an issue that may arise from the improper implementation of a rule or of business logic. It also provides the ability to reissue any previous version of the message. This ability can help reduce the costs of resubmitting data in the event of an issue.

**ALERTING**

LDEF’s integrated rules engine lies at the heart of our alerting functionality. This component spans the entire system, constantly evaluating message content and system performance data to make decisions based on user-defined business rules. When the system identifies an event or complex event that meets the criteria, it can trigger an alert and deliver it according to the user-determined method—for instance, by email to the appropriate staff, via SMS messages, or as alerts displayed on one of the system’s management dashboards. Events and complex events may also initiate additional process flows in the system to provide automated responses to identified facts.

**SCALABILITY**

LDEF’s distributed design makes the system highly scalable. The system is deployed as a collection of specialized nodes. All nodes update and coordinate with one another so that the overall solution acts as a single, integrated virtual system. As volumes or workloads increase, message-processing rates are maintained by adding additional nodes, and all system-wide resources—the rules engine, persistence database, the system-management functionality—are immediately available to the new resources.

LDEF uses a publish-and-subscribe mechanism to keep all system segments aware of the current state of the architectural footprint. For example, if LDEF starts another orchestrated process engine—per client-defined rules—the new engine “publishes” that it is now “live.” The relevant components, which will have “subscribed” beforehand to be notified of such events, will then receive notice to be aware of the new addition to the architecture. The model works the same way to reduce the footprint once the spike subsides. With this system, LDEF knows at all times what the active footprint is and how to distribute the workload effectively and efficiently.

Similarly, the No-SQL data store expands horizontally rather than vertically—so more data requires more distributed space, not larger computers. One of the main benefits of this approach is that scalability is limited only by the number of virtual machines that can be allocated to the system and the number of physical resources available to be assigned to these virtual machines. This horizontal scalability is achieved by pushing data across numerous nodes and applying an internal protocol, Gossip, to keep all nodes informed of the locations of data. We say “locations” because our No-SQL data store distributes the data to more than one node; that way, the data are available from more than one location. This approach provides better throughput: if one location with the data is busy, the new request is passed to an available node. This approach also avoids having a single point of failure—if one node with the data goes down, other nodes are still running and available for access.
DATA SECURITY

One of LDEF’s strongest features is its security model (Figure 4). Our library of data communications endpoints includes a full complement of secure, encrypted transport protocols, including HTTPS, SFTP, and SSH Server.

To address security for all data in transit through LDEF, nodes are created as Linux appliances under SELinux. This system delivers the highest level NSA- and Department of Defense–approved security. At activation, the SELinux firewall assumes that all ports are closed. By opening only those ports needed for LDEF’s internal communications and restricting all open ports to predefined IP addresses known to LDEF, the system tightly controls access to data. Ports can also be set to receive data only from certain systems, further narrowing the “window of opportunity” for access violations. In addition, the SELinux image of each node is encrypted when offline, a precaution that provides additional security for both the operating system and the data. The run-time hard disk can also be encrypted, providing an additional layer of operating-system security for all data.

HIGH AVAILABILITY

In addition to providing exceptional scalability, LDEF’s highly distributed architecture ensures high-availability operation of all interfaces by providing automated failover system nodes to compensate for either hardware or software failures. Nodes can be deployed on the same server, on multiple local servers, or on a combination of local and remote servers for disaster recovery and continuity of operations. Because all nodes are working in coordination, if one or more nodes fail, the system keeps working: remaining resources automatically take over, and processing continues (albeit with somewhat reduced performance). This failover
process also preserves data stability. LDEF manages data across multiple nodes, with data redundancy. This system allows remaining nodes to always have access to a full copy of the system’s data.

**SHORT CYCLE TIMES FOR DEVELOPMENT**

LDEF’s full-featured graphical development environment, shown in Figure 5 gives the client’s integration team the ability to quickly create new orchestration routes to link inbound and outbound endpoints and to define workflow paths for message routing, filtering, validation, and transformation services. This graphical output is also a valuable graphic for documentation and for communicating to both the integration team and management how the system is designed to function.

![Figure 5. Graphical Route Development Drag-n-Drop Environment](image)

**HEALTH INFORMATION EXCHANGE (HIE) DATA STANDARDS**

Providing effective interoperability across diverse organizations and systems is a primary use case for any integration engine. LDEF can provide the functionality needed to support and manage a wide variety of secure data-transfer protocols, validation and parsing of different standard message formats, and data-content mapping and code translation between coding systems (e.g., from local to target-system standards.) The system can also integrate with health information exchange (HIE) resources, including enterprise community patient indexes, document repositories, and external gateways (e.g., NwHIN CONNECT, Direct) to enable effective interoperability across entities both inside and outside an HIE.

LDEF’s comprehensive integration environment can also simplify and reduce the demand for redundant data exchange. LDEF’s integrated nature as a highly distributed virtual solution means that all messages flowing through the system are available to any messaging process within it. Therefore, data from a single interface to an EMR system can be transformed and reformatted to support multiple outbound interfaces to a variety of HIEs and other reporting agencies.
ARCHIVE AND RESTORE

Both online (i.e., active data store) and offline (i.e., inactive data store) backups are possible for the LDEF solution. Leidos will work with the client to design an effective archive and recovery solution. Because online backups can slow production data storage during the backup process, Leidos architects typically recommend offline backups for scheduled backups. However, we will configure online backups for any scenarios that require this approach and, when done in conjunction with a DR site, we are able to provide this feature with a very small impact on the production system.

Cassandra is a No-SQL database that represents data in a flat-file type of structure; this technology results in a simple restoration process. Given the sensitive nature of health data, we recommend that data extracted for restoration be encrypted. Leidos can assist the customer in configuring the solution for restoring data and encryption where necessary. Encryption is fully supported as part of the LDEF tool set using the SELinux encryption processes.

DATA VERIFICATION

Data verification is a built-in function of the LDEF rules engine. The client provides the needed business logic for data verification as part of the gathering of requirements, and the Leidos team will create the business rules to implement it. Leidos staff will work with the subject matter experts to create and validate all verification processes.

DATA VALIDATION

Data validation is managed by our rules engine. Clients may view the associated rules for verification and validation at any time. The customer can be trained to create these rules with in-house or contractor staff. Leidos and third-party resources are available for training the customer’s staff.

MESSAGE CONFORMANCE VALIDATION

For message-conformance validation, we also use the rules engine and libraries to apply standards, local codes, and specifications. As messages are validated, they are allowed to proceed through the process flow. If a message does not meet the required validation, it is saved, and an error log is populated with the message-ID and the reason for failure. This system allows the message to be reissued once the validation problems are resolved.

CONCLUSION

The Leidos Data Exchange Framework is a new approach to providing healthcare organizations with the integration functionality they need to quickly establish and effectively manage their growing HIE environment. We look forward to the opportunity to install this solution in a reference implementation.
Appendix A

Note: All features listed in this Appendix are accessible only to staff members via their explicit login rights as defined by a customer-defined administrator.

A DETAILED LISTING OF FEATURES PROVIDED BY LDEF MONITORING AND OPERATIONAL CONTROL

1. No-SQL Data Store
   - Create Cluster: create a cluster.
   - Feedback: an online form that sends your evaluation of OpsCenter or any comments to us.
   - Report: information about clusters that OpsCenter manages in PDF format.

   This sub-system is divided into these main functional areas:
   - Overview—Survey each cluster’s Dashboard in this condensed view. Displayed when multiple clusters are present.
   - Dashboard—View graphs of the most commonly watched Cassandra performance metrics.
   - Cluster—See your cluster from different perspectives and perform certain maintenance operations on cluster nodes.
   - Performance—Monitor a number of Cassandra cluster performance metrics. Real-time and historical performance metrics are available at different granularities: cluster-wide, per node, or per column family.
   - Schema—Create and manage keyspaces and the column families within them.
   - Data Explorer—Browse through column family data.
   - Event Log—View the most recent OpsCenter log events, such as OpsCenter startup and shutdown.

2. VM Node Management
   - Create VM images.
   - Install new VM images.
   - Manage VM images across multiple machines.
   - Establish VM state.
   - Version control VM images.
   - Connect and manage VM instances.
   - Extend VM Disk Allocations.
   - Manage VM Networking.
   - Document and manage all VM configurations.
   - Deploy Images.
3. General Systems Monitoring

- Graphical Interface of all LDEF nodes, message counts, status of node, relationships to other nodes, and more.
- Graphical Interface of all LDEF physical and virtual servers with status icons in sub-real-time (i.e., seconds to sub-seconds of actual state).
- Events, Notifications, and Alerts for system issues to client designated staff via physical devices as defined by client (e.g., e-mail, cell phone, etc.).
- Plots and reports based on time windows and metric type.
- The ability to monitor, event, notify, alert, and automate where appropriate any electronically known fact.

ABOUT LEIDOS

With expertise in National Security, Health, and Engineering, Leidos will continue SAIC’s proud legacy of securing the future of our families, our communities, and our world. Our diverse knowledge base and involvement in multiple disciplines and industries make us unique. We use our insights in new ways and expand our capabilities to evolve the industries we serve.

For more information, visit leidoshealth.com.

CONTACT US

770.986.3526
interoperability@leidos.com

11951 Freedom Drive   /   Reston, VA 20190

© LEIDOS. ALL RIGHTS RESERVED.