

Application Interoperability Using Standards and the PHINMS

PHIN Conference Session 4C

May 14, 2003

W. Ted Klein

Klein Consulting, Inc.



SAFER • HEALTHIER • PEOPLE™

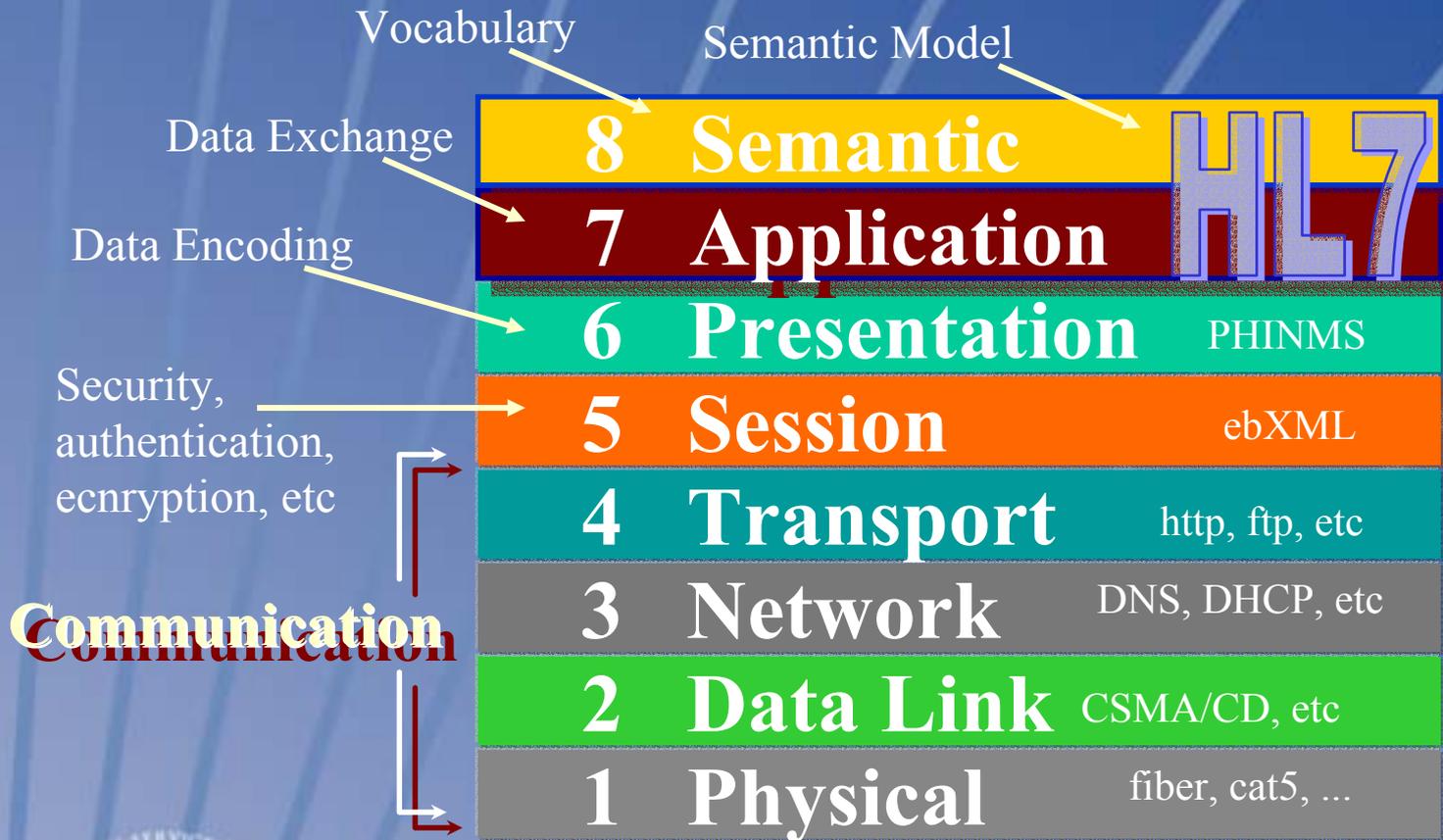


Public Health Messaging Requirements

- Routine surveillance and reporting
- BT or other unexpected event
- Must enable *interoperability* among State and Local Departments of Health and the Federal agencies and others
- Usability requires that as many things be automated as possible
 - ◆ New partners coming on line
 - ◆ Resend of queued messages



Interoperability Layers



PHINMS: ebXML Based Messaging

- Handles Physical, Network, Session, and Encryption requirements
- Built on ebXML – an International Standard
 - ◆ Draft specification for HL7 messages in ebXML underway
- Based on message-oriented transactions between a sender and a recipient
- Fully secure – used widely for financial transactions
- Handles the multiple types of authentication paradigms found in the States
- Supports XML packaging of any type of data exchange format (both HL7 versions 2 and 3)



SAFER • HEALTHIER • PEOPLE™



PHINMS Requirements

- Each pair of partners has a **route** between them
- Each route must have an associated **Collaboration Protocol Agreement** (CPA)
- Each CPA is made up of:
 - ◆ Unique IDs for sender and receiver
 - ◆ Transport protocols for sender and receiver
 - ◆ URL endpoint
 - ◆ Authentication type
- Still must setup the authentication mechanism and supporting data
- These must be setup before you can communicate!
- Yesterday's talk had more detail on PHINMS technology



Common Semantics to achieve Application Level Interoperability

Built on top of the infrastructure, it requires:

- The 'format' being used: Data Exchange formats
 - ◆ HL7, DICOM, ...
- The 'items' being communicated and they are recognized: Semantic model and Identifiers
 - ◆ Cases, contacts, exposure cohorts
 - ◆ Laboratory orders and results
 - ◆ Interventions
 - ◆ Environmental and Spatial data
 - ◆ Health alerts and Recommendations
 - ◆ Identifiers and Namespace management
- The 'words' used to communicate: Terminology
 - ◆ LOINC, SNOMED, NDC, ...



Data Exchange Formats

- HL7 is the #1 emerging standard for communicating Clinical information – implemented very widely
 - ◆ ELR and Vaccination Records are current carried in V2
 - ◆ Version 3 is the new HL7 standard
 - ★ First balloted release of version 3 will occur within the next few months – the RIM has already passed ballot
 - ★ New Notification messages are version 3
 - ◆ Uses XML and vocabulary standards
 - ◆ Note that the flexibility of HL7 requires that interfaces follow specific Implementation Guides
 - ★ Without conformance to these guides, messages may be incomprehensible
- DICOM widely used for images
 - ◆ Supported by nearly all PACs systems



Semantic Model

- Must have shared understanding of common concepts
 - ◆ What is a Case? Sample? Outbreak? Investigation?
- Concepts have explicit relationships to each other
 - ◆ Cases have contacts that yield exposure cohorts
 - ◆ Samples for Laboratory orders generate results
 - ◆ Exposure cohorts may receive interventions
 - ◆ Health alerts and Recommendations may be based upon laboratory results
- Concepts and relationships are documented in a shared semantic model
 - ◆ HL7 RIM, PHLDM

THESE ARE NOT PHYSICAL DATABASE MODELS!

(but may be implemented in one)



Standards for Semantics

- HL7 covers many of these items
 - ◆ Balloted ANSI standard
 - ◆ Standard Reference Model
- Other standards are covering other areas key to Public Health
 - ◆ OpenGIS – Environmental and Spatial Data
 - ◆ RxNorm – Standard model for drugs
 - ◆ Many others



Terminology

- Concepts represented by codes with display text
- Each code is unique in a versioned code system
- Every concept is uniquely identified by a tuple:
 - ◆ (coding system, code, version)
- Used so computers can manipulate concepts
 - ◆ Values typically stored in tables in the software
- Sets of codes defined by Value Sets
 - ◆ Used to help distribution and maintenance of codes
- Standard vocabulary reduces the huge issue of mapping
 - ◆ John will speak in more detail on this
- PHIN makes use of both standard vocabulary and CDC defined and maintained vocabulary
 - ◆ consistent with emerging CHI recommendations



Code System Usage

- Standard code systems will be employed wherever possible
 - ◆ LOINC, SNOMED, NDC, HL7, ...
- Certain concepts peculiar to Public Health will have CDC maintained code systems
 - ◆ Case Definition Rules, Case Reporting Source, ...
- Code system tables for PHIN messaging will be made available for electronic download



Code Mapping

- Almost everyone uses at least some local codes
- Interoperability requires mapping if the same codes are not used
- Codes to be used for notification messaging will be downloadable
- The electronic form can be used for mapping
- Translation/mapping products and services are commercially available
- Systems must be designed to enable mapping and new updates to codes



Infrastructure Requirements

- Standards Based
 - ◆ Web, TCP/IP
 - ◆ XML
 - ◆ HL7
- Support Secure Messaging
 - ◆ Authentication
 - ◆ Encryption
 - ◆ Non-repudiation
- Public Health Focused
 - ◆ PHLDM
 - ◆ Public Health vocabulary
 - ◆ Public Health defined messages
- Deployable and Maintainable



Standards Based

- Following standards gives the greatest opportunity for interoperability
- A large number of standards already exist
 - ◆ Technology
 - ★ TCP/IP, XML, ebXML, http, https, PKI
 - ◆ HealthCare
 - ★ HL7
 - ◆ Vocabulary
 - ★ LOINC, SNOMED, NDC, HL7 Vocabulary Services
- PHINMS is standards based



Secure Messaging

- Multiple requirements/desires
 - ◆ Access Control, Authentication
 - ◆ Encryption, non-repudiation
 - ◆ Transportability, Verifiability
- Recommend Digital Certificates
 - ◆ Access Control, non-repudiation
 - ◆ Other authentication types may be used (eg username/password)
- Recommend Public Key Infrastructure
 - ◆ Encryption and Verifiability
- PHINMS ensures transportability using all of these



Public Health Focused

- Support for Public Health Concepts
 - ◆ Based on PHLDM
 - ◆ Support Case, Investigation, etc.
- Make use of existing public health messaging standards
 - ◆ ELR, Vaccine Registry, etc.
- Must support Public Health workflow
 - ◆ Both BT and routine surveillance
- Must extend beyond Federal level
 - ◆ State and Local jurisdictions



SAFER • HEALTHIER • PEOPLE™



Deployable and Maintainable

- Deployment Ease
 - ◆ Software should be distributed on CDs with straightforward install scripts
 - ◆ States must be able to load software onto their own partners without CDC involvement
 - ◆ Partners must be able to configure easily from supplied documentation
- Maintenance Ease
 - ◆ Remote where possible
 - ◆ Simple Software and Vocabulary updates
- The number of configured and maintained CPAs should be held to a minimum



Establishing Communication

- Discover and Identify of Recipient(s)
- Establish Communications Environment
- Identify Route
- Setup Authentication Information
- Acquire Encryption Key for Recipient
- Package Message
- Send Message to Recipient



Discovery and Identification

- Recipients of message must be known
 - ◆ A priori
 - ◆ Dynamically discovered (new or temporary)
- Recipients must be identified
 - ◆ Messaging systems communicate with software instances
 - ◆ On specific computers
 - ◆ At specific network locations
 - ◆ Every software instance must be uniquely identified
- The identifiers are technical addresses used by the communications machinery



Standards and Discovery

- Emerging Standards
 - ◆ ISO 23950 – Information Discovery Standard
 - ◆ FIPS 192-1 – Establishment of Government Information Locator Service
 - ◆ OASIS/ebXML Registry Services v2.0
 - ◆ OASIS/UDDI v2 and v3 API
 - ◆ many others



SAFER • HEALTHIER • PEOPLE™



A Word About Identifiers

- HL7 version 3 has embraced ISO OIDs – Object Identifiers
 - ◆ Globally unique and widely used
 - ◆ Easily machine processable and readable
- Structured as a well formed tree
 - ◆ HL7 root: 2.16.840.1.113883
 - ◆ CDC PHIN root: 2.16.840.114222.4
- CDC will use OIDs in the PHIN for:
 - ◆ Public Health Identifier Namespaces
 - ◆ Vocabulary Constructs
 - ◆ Well known objects
 - ★ Messaging partners
 - ★ Software Packages and Services
 - ◆ Enable linkage of key Public Health concepts
 - ★ Cases, specimens, results, etc.



OIDs – consistency with HL7

■ HL7 use of OIDs

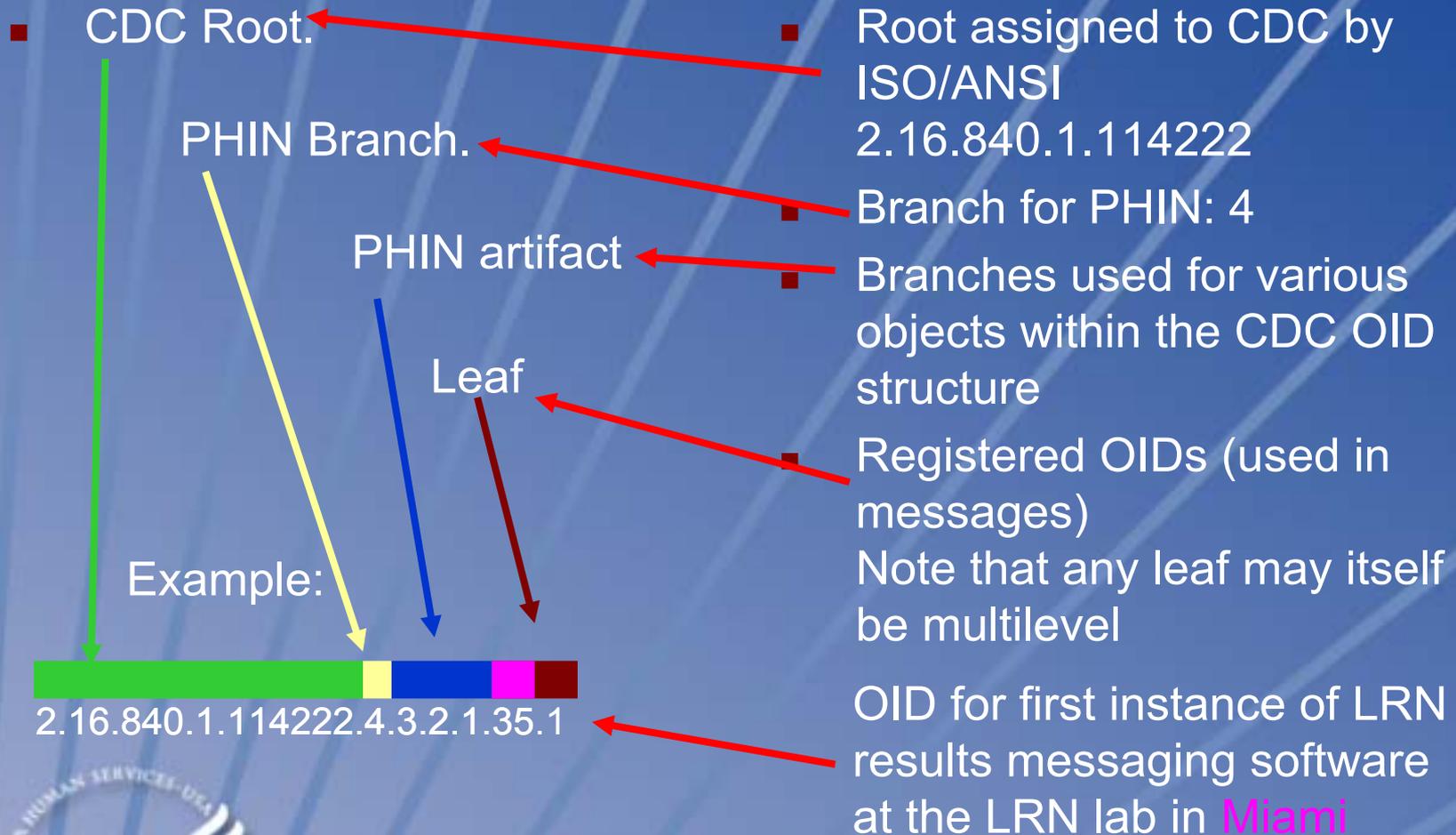
- ◆ Identifier Namespaces
- ◆ Vocabulary Constructs
- ◆ Well known objects
 - ★ HL7 Artifacts
 - ★ HL7 Organizational Bodies
 - ★ Member Organizations

■ CDC use of OIDs

- ◆ Public Health Identifier Namespaces
- ◆ Vocabulary Constructs
- ◆ Well known objects
 - ★ Messaging partners
 - ★ Software Packages
 - ★ Public Health Organizations



The CDC OID Structure



Identifying Software Instances

- PHIN Root OID
 - 2.16.840.1.114222.4
- NBS OID
 - 2.16.840.1.114222.4.3.2.4
- Nebraska DOH OID
 - 2.16.840.1.114222.4.1.168
- Nebraska PHIN suffix
 - 168
- NBS at Nebraska DOH
 - 2.16.840.1.114222.4.3.2.4.168
- Second instance
(assume #1 is pilot, #2
is production)
 - 2.16.840.1.114222.4.3.2.4.168.2
- Case ID Namespace
 - 2.16.840.1.114222.4.3.11
- Namespace for Case IDs generated by the second
instance of NBS at the Nebraska DOH:
2.16.840.1.114222.4.3.2.4.168.2.3.11



What does an OID identify?

- The PHINMS delivers messages to queues and to software packages implementing the *Services*
- These destinations must be uniquely identified
- OIDs permit precise identification and association with metadata to deliver to the correct service
 - ◆ Metadata design is underway to enable searches so that a sender may route to the right receiver service
 - ◆ Alleviate the need for States to program their own redirection software at the receiver



How to find the right OID?

- Must have the technical address of the destination
- Static addresses are published
 - ◆ State/Local DOH, federal agencies
 - ◆ Static values published in the Impl. Guides
 - ◆ But they may change on redeployment or updates
- Dynamic addresses must be 'looked up'
 - ◆ Field team systems
 - ◆ Event-in-progress systems
 - ◆ Both HL7 and CDC will have an OID registry available for lookups



Communications Environment

- The environment for this transaction needs:
 - ◆ Unique IDs for sender and receiver
 - ◆ Transport protocols for sender and receiver
 - ◆ URL endpoint
 - ◆ Authentication type and the data
 - ★ Basic authentication: username/password or custom forms-based authentication
 - ★ Digital certificates
 - ◆ Delivery addresses at receiver (Service/Action)
 - ◆ Public Key of Recipient
 - ◆ Verisign used currently as a certificate authority
- Must determine the Route or Routes for this message to discover these items!



Must Handle Multiple Routes

- Routine surveillance and reporting
 - ◆ Nebraska samples sent to South Dakota Laboratory
 - ◆ Lab results sent to both/either South Dakota/Nebraska
 - ◆ Follow-up confirmations done in other locations (eg CDC RRAT lab)
- BT or other unexpected event
 - ◆ Multiple locations involved
 - ◆ High probability targets cover multiple jurisdictions (New York City → NY, NJ, CT, NYC)
 - ◆ Capacity overflows sent to remote laboratories



A Recent Example

- The 2001 Anthrax event
- Samples taken from four locations
 - ◆ Hart Building, Washington, DC
 - ◆ AMI in Florida
 - ◆ Postal facilities in New Jersey
 - ◆ Various locations in Connecticut
- Samples were tested in over 100 laboratories
- Results were sent to many States
- For all of this information to be immediately available, a high level of routing is needed



More Routing Requirements

- Copies of messages often need to be delivered to additional recipients
- Message copies may be directed to recipient lists based upon content
 - ◆ Disease notifications get sent to different recipients depending upon the disease
- Data transforms may be required for the different copies
 - ◆ Identifiers may be made opaque
 - ◆ Some recipients might only want roll-ups



Identify Route(s)

- ebXML communication involves a single pair of participants: each of these is a single route
- Each transaction is a secure point-to-point transaction
- All potential routes for all possible messages are on the order of $n*(n-1)$ where n is the total number of public health participants (in the hundreds)
- Every route must have a profile of characteristics (the CPAs discussed earlier)
- Intelligent and dynamic routing is required to alleviate the need to manage *tens of thousands* of route profiles
- The same problem exists for any secure communications mechanism (other non-ebXML)
- Note that using Route-not-Read reduces the combinations to $2n$ (everyone can talk to the CDC)



Routing Infrastructure

- Extensions to the PHINMS are in design to extend initial routing capabilities
 - ◆ Send a message to a Group, broadcast
 - ◆ Fully encrypted multicast and forwarding
- Routing is required even if you do not use a messaging technology other than the PHINMS
 - ◆ The Uses Cases remain regardless of technology
- There will be application responsibilities defined and published
 - ◆ If you have your own application using the PHINMS you will know how to enable routing

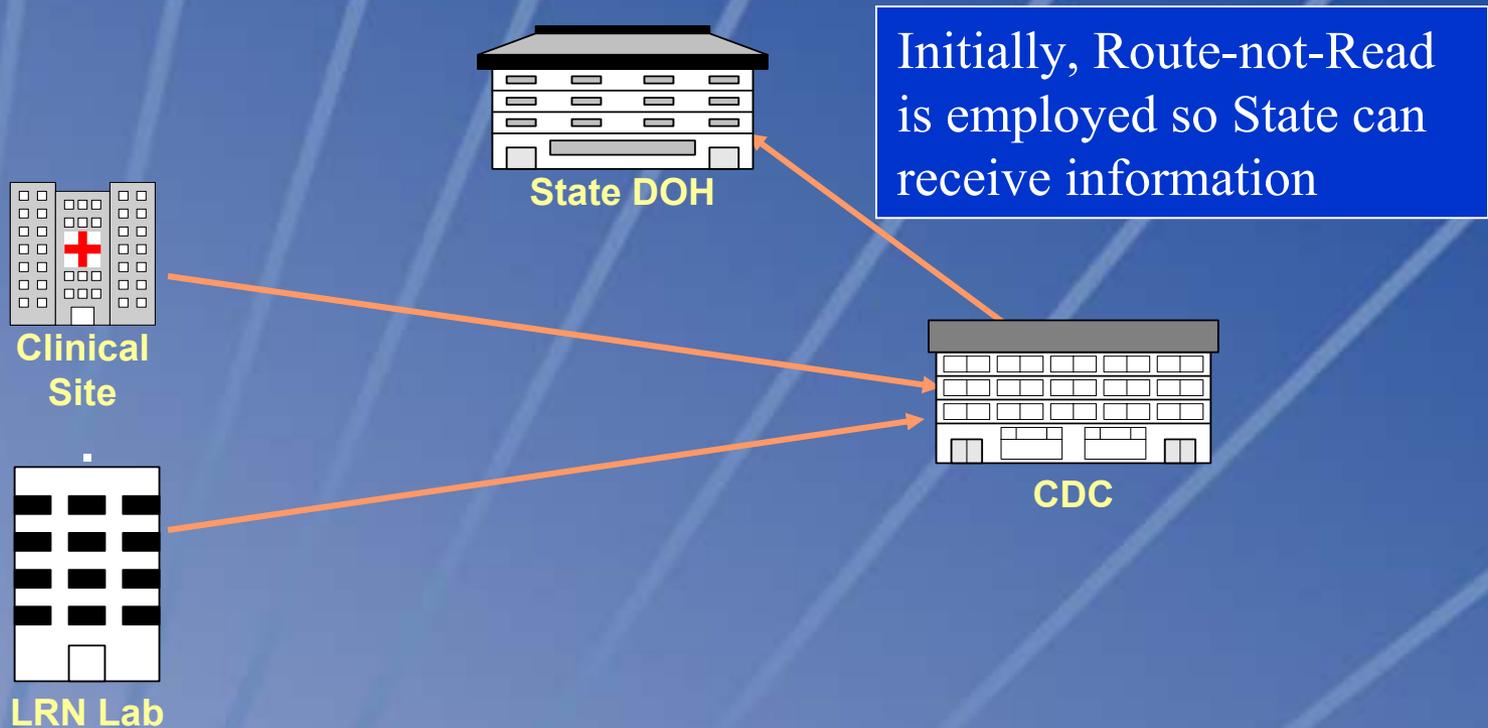


Default Routing Operations

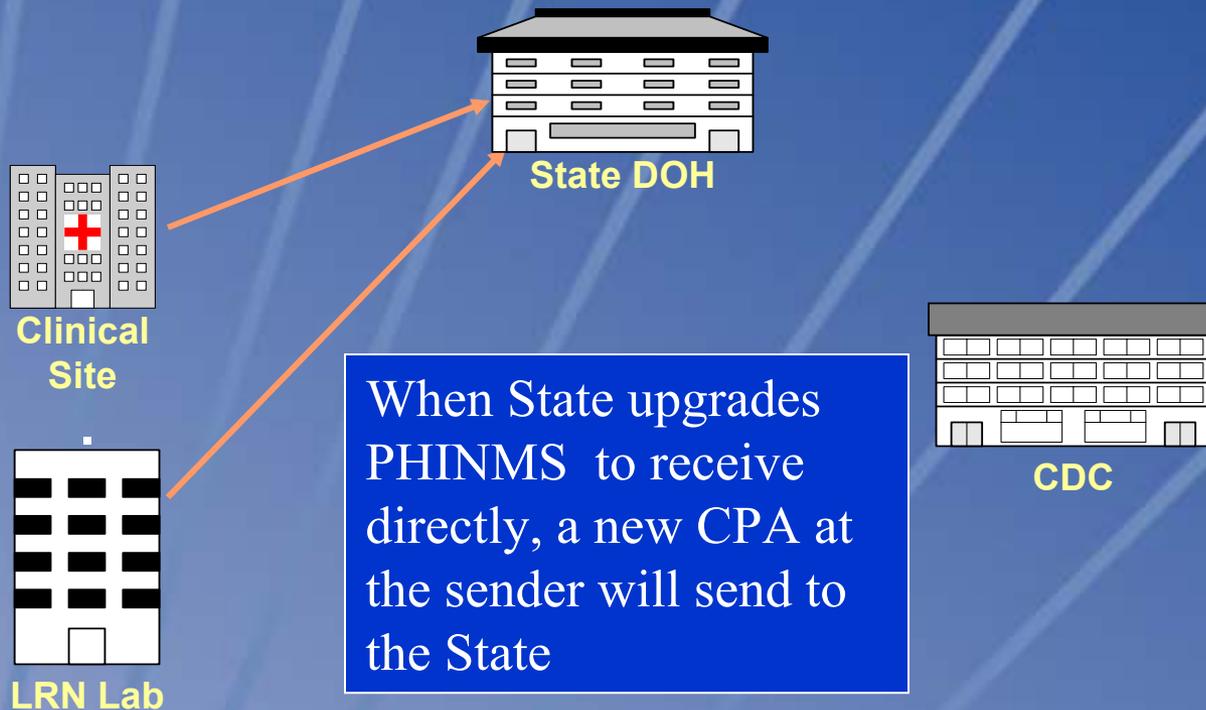
- If you have a CPA for all recipients, you can route to everyone
- If not, you can send to anyone else in the network using the Route-not-Read facility at the CDC
- At any time you may establish a CPA to send to any partner able to receive
- Design is underway to permit the infrastructure to make these determinations so that it is not a user or operator issue



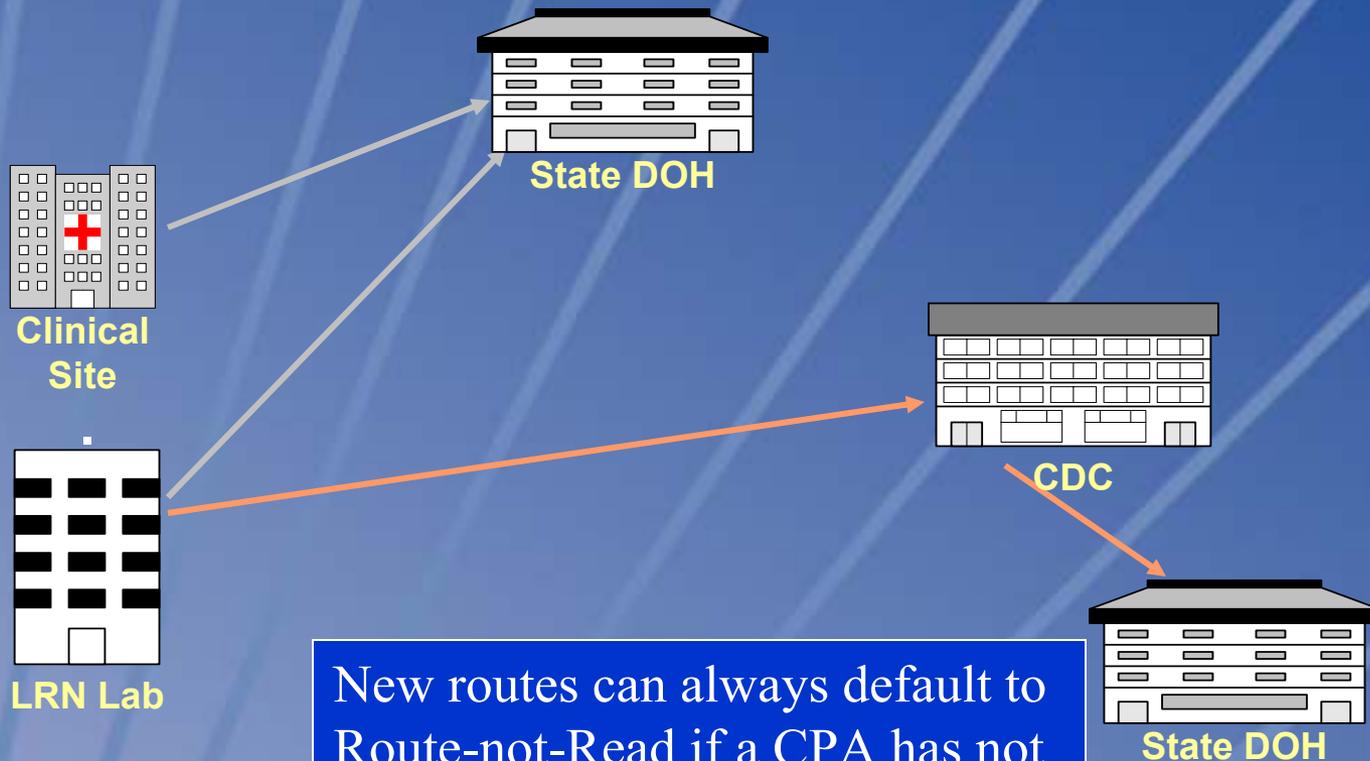
Routing Illustration



Routing Illustration



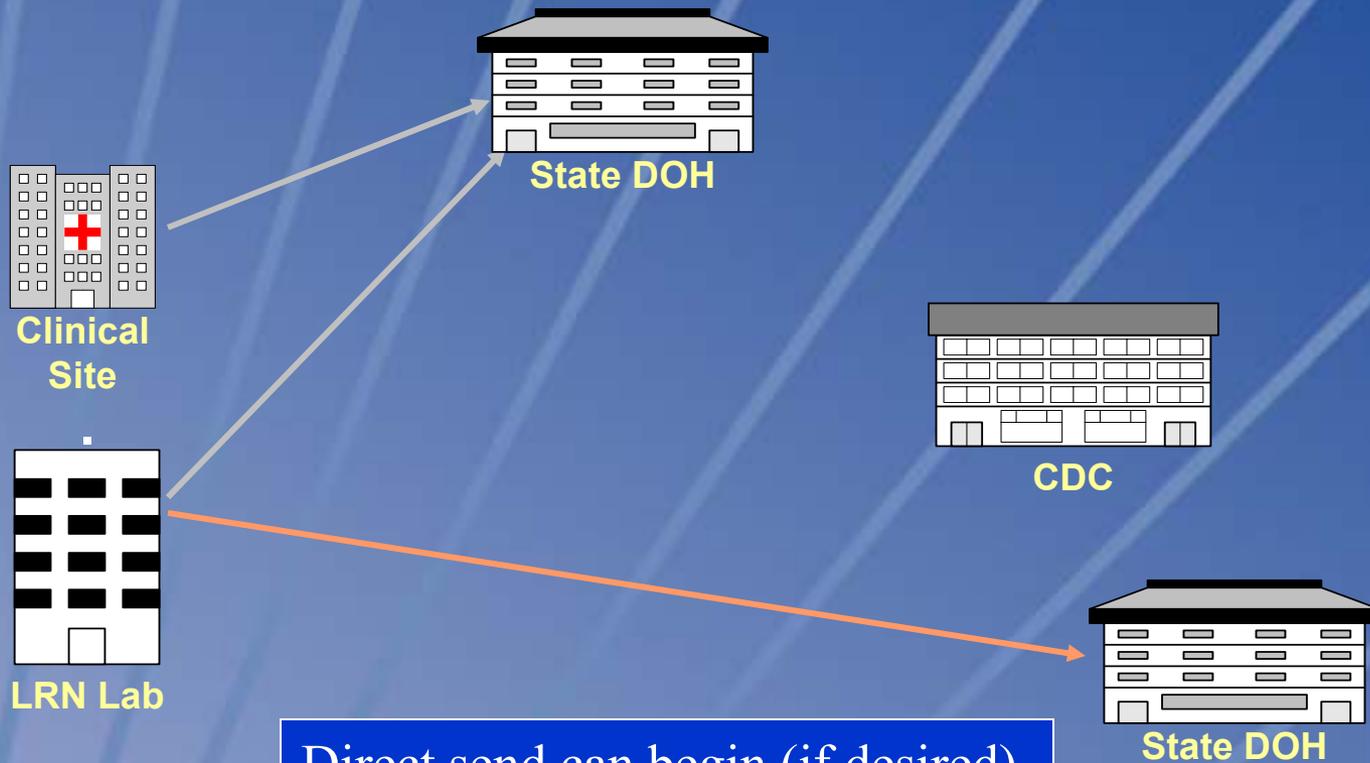
Routing Illustration



New routes can always default to Route-not-Read if a CPA has not been setup yet



Routing Illustration



Direct send can begin (if desired)
when the CPA is established



Setup Authentication Information

- States use a variety of Authentication mechanisms requiring different data
 - ◆ Username/Password
 - ◆ Client Certificate
 - ◆ Forms-based authentication
 - ◆ None (does anyone have no authentication?)
- Existing mechanisms through the SDN to get certificates
 - ◆ Uses two-factor authentication for increased security
- Without a central authentication authority, each Sender must manage credentials for each route



Acquire Encryption Key for Recipient

- All messages are sent securely and may be encrypted with the recipient's Public Key
- Currently support LDAP calls to acquire Public Key
- Note that the Private Key for each recipient must be part of the recipient setup/configuration



Package Message

- Various applications will generate data for messages in various ways
- There are tools to support data mapping between RDBMS and different message construction tools
- Various interface engine vendors are preparing to support version 3 schemas
- Many components are included in standard installations
- The completed message is passed to the PHINMS



Send Message

- A message is sent by writing the message into a queue
- The queue entry contains all the necessary identifiers for the PHINMS to use the correct CPA
- The package specifies whether the send is direct or through the CDC Route-not-Read facilities



So What Does The Recipient Do With It?

- PHINMS delivers the message to a particular Service at the recipient
- There are many parsers available for received HL7 messages, both COTS and Open Source
- NBS will parse and accept/store messages
- States may write their own recipient software to parse and store, then register the ID so others may send to it, along with the metadata to describe it properly
- Implementation guidelines for doing this will be published

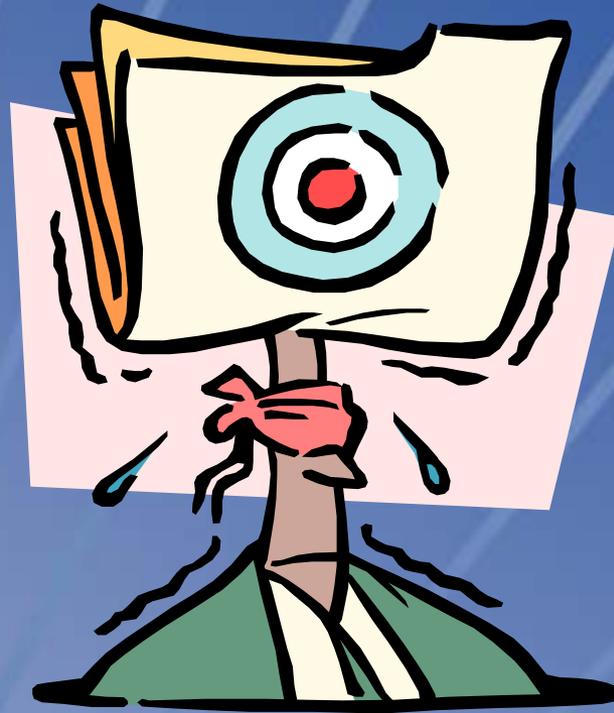


Interoperability Achieved!

- Once the received message has been parsed, the data contained is available
- As long as the structural semantics and vocabulary are shared, the message data can be directly used by the recipient
- States may use the same mechanisms to communicate with their local jurisdictions
 - ◆ This is additional routing at this level



Thank You



Questions?



SAFER • HEALTHIER • PEOPLE™

