InCLUDE Data Exchange

Julia Harrell, GISP
GIS Coordinator, NC DENR
Julia.harrell@ncdenr.gov

InCLUDE Project Partners

- **NC Department of Environment & Natural Resources:** The State of NC's lead environmental agency

- **Eastern Band of Cherokee Indians:** Has Tribal lands largely in NC

- **NC Property Mapper's Association:** a non governmental professional organization of County Government property mappers

- **NC Division of Land Records:** part of the NC Secretary of State's office

- **NC Statewide Mapping Advisory Committee:**
  WGSP: Working Group for Seamless Parcels
  WGRT: Working Group for Roads & Transportation

Why use the Exchange Network to share parcel & land use data?

- Many environmental problems cross multiple jurisdictions. The business of managing and solving these problems has become complex and very information intensive.

- Environmental policy makers and other stakeholders need access to timely, accurate, and consistent data that present a complete picture of the environment. This includes property ownership and land use.

- Exchanging high quality information allows governments and regulated communities to make better planning decisions regarding the environment. Property data is a vital missing piece of the puzzle, not only for planning, but also incident response.

Benefits to NC DENR

- Most regulatory and natural resource programs need current property boundaries, ownership and land use information, which is not consistently available in a common format.

- Some of the 100 NC County systems are mainframe-based and/or extremely proprietary.

- A consistent source of seamless parcel and land use data will aid DENR in: regulatory permitting decisions, mitigation and remediation efforts, land acquisition for parks and protected areas, wildfire response activities, and regional level analysis.

Benefits to Eastern Band of Cherokee

- More easily get copes of county property data and share Trust land boundaries digitally with NC County property mapping offices.

- Access NC County land records digitally and allow the Tribal Historic Preservation Office to work more closely with the State Historic Preservation Office.

- Locate culturally important sources of natural resources, such as River Cane and Butternut, and have easy access to property owner information to facilitate permission for harvesting these resources.

Benefits to NC Counties

- Reduce the burden of filling frequent parcel data requests from state and federal agencies.

- Establish a common data exchange format for parcel data without requiring counties to change any of their back-end systems.

- Allows counties to maintain control of transforming their data and determining who should have access to it.

- Will aid in creating regional level economic development and growth plans, as well as mutual aid and incident response compacts.

Major Project Components:

InCLUDE = **In**tegrated **C**adastral and **L**and **U**se **D**ata **E**xchange. InCLUDE is the name of both the data flow and the project. Will use a GML "Simple Features" Level 1 schema, following the FGDC Parcel Working Group guidelines as closely as possible.

SPART-NC = **Spa**tial **Tr**anslator for **NC.** A secure web application allowing local governments to upload and transform their parcel and land use data to a common schema (format) for use in the InCLUDE data exchange.

NC LandGate = another secure web application for Government & Tribal users to identify and download portions of the transformed parcel & land use data from a map-centric interface.

What has changed in the project?

- Cost analysis determined that it would be more economically viable to build the data transformation application as a web-based product rather than as a desktop software application.

- Functionality analysis determined that much of the needed data transformation capabilities already existed in an open source software packaged designed to provide Spatial ETL (extract, transform & load) services: ***GeoKettle***

- Additional functionality analysis determined that much of the needed "on-demand" spatial data requests that the InCLUDE data flow requires are already supported in an open source Web Feature Service software application: ***GeoServer***
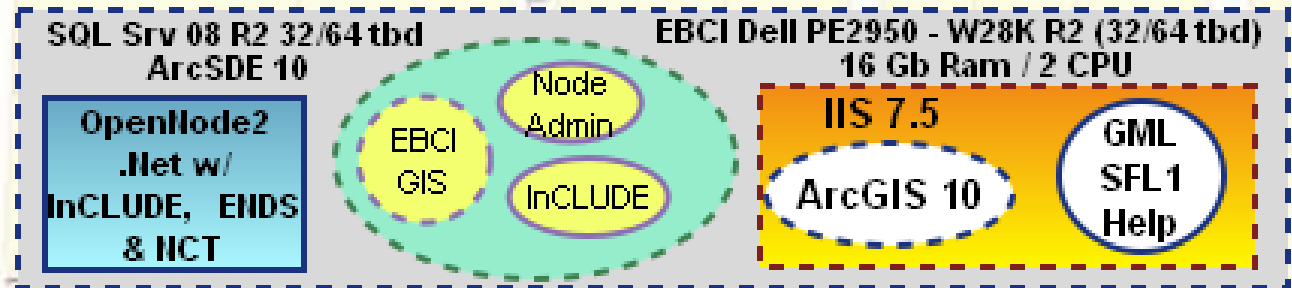
What is GeoKettle?

- A spatially-enabled version of Pentaho's Data Integration Suite

- A metadata-driven spatial ETL tool with direct execution of transformations (think Safe Software's FME, only open source)

- Support for most RDBMS formats and over 30 vector GIS formats (via OGR & GeoTools)

- Handles Geometry objects (via GeoTools & Java Topology Suite)

- Has a Javascript API for accessing Geometry Objects

- Supports topological predicates, and aggregation operators (intersect, envelope, union, etc.)

- Supports Coordinate Transformations (via GeoTools & Proj)

- Has cartographic preview and can be deployed via web services

What has not changed in the project?

- InCLUDE Plugins will still be built for both the Java & .Net flavors of OpenNode2. DENR will use Java & EBCI will use .Net

- The OpenNode2 administration database will be ported to run on the PostgreSQL + PostGIS platform for use at NC DENR.

- DENR will use the Open Source GeoServer 2.1 product as its companion WFS server and EBCI will use ArcGIS Server 10. The Open Source 52North GeoWeb Security software will be used to secure both types of WFS services.

- EBCI will use an existing instance of SQL Server +ArcSDE as their InCLUDE spatial database. ArcGIS Server will require a custom "assist layer" to provide WFS support for InCLUDE's GML Simple Features Level 1 schema.
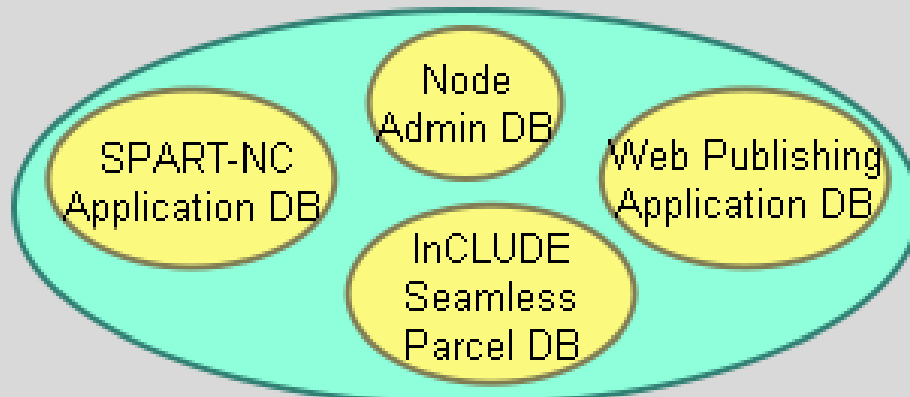
InCLUDE System Components Diagram

NC InCLUDE System
Proposed Components
(dashed = existing)

NAAS

NCID

SQL Srv 08 R2 32/64 tbd
ArcSDE 10

OpenNode2
.Net w/
InCLUDE, ENDS
& NCT

EBCI GIS

Node Admin

InCLUDE

EBCI Dell PE2950 - W28K R2 (32/64 tbd)
16 Gb Ram / 2 CPU

IIS 7.5

ArcGIS 10

GML SFL1 Help

DENR VMWare 4.5 VM, OS = W28k Server R2 (32/64 tbd) 16 GB Ram / 2 CPU

GeoKettle 3.2 (or latest)

SAN Storage - File system for uploaded data, metadata, transform recipes, etc.

Apache 2.2.17 & mod_cas

SPART-NC Web App

InCLUDE Map Web App

SPART-NC Application DB

Node Admin DB

Web Publishing Application DB

InCLUDE Seamless Parcel DB

PostgreSQLS 9.0.3 + PostGIS 1.5.2
(or latest stable release when dev. starts)

ApacheTomcat 6.0.18

52North WAS/WSS /WSC.Web

Geoserver 2.1 & AppSchema Ext.

OpenNode2 Java w/ InCLUDE, ENDS & NCT

2 / 25 / 2011

SPART-NC: User Operations

1. County GIS user logs into SPART-NC using NCID or NAAS account, uploads a Parcel shapefile & export file from the County Tax Assessor's CAMA DB.

2. Uses SPART-NC to "map" the elements in their parcel source data to the common InCLUDE schema, and the County Land Use codes to the common high-level Land Use classifications in the InCLUDE schema.

3. Saves their data transformation "recipe" for export, later re-use or further modification.

4. Executes the transformation recipe on their uploaded source data.

5. Runs a validation tool against the transformed version of the data.

6. If validation finds errors, modify recipe, re-run steps 4 & 5 until no errors.

7. Defines "sharing rules" and submits correctly transformed data to NC InCLUDE spatial DB's "seamless parcels" layer.

SPART-NC: behind the scenes

- Uploaded source shapefile & DB file ingested into NC InCLUDE's PostgreSQL + PostGIS spatial database.

- App reads source element structure out of NC InCLUDE spatial DB and transformed data structure out of InCLUDE GML schema.

- User defines which source elements correspond with transformed destination elements, maps source domain values to destination domain values in InCLUDE schema. App has Javascript string/number functions and simple conditional logic.

- App saves user defined data mappings in GeoKettle's existing xml transformation 'recipe' file format.

- App sends 'recipe' file to GeoKettle's Carte web service interface for transformation. GeoKettle's validation & reporting tools enhanced for providing user-friendly error reports. Underlying JTS library used for geometry validation.

- Transformed data passing validation loaded into NC InCLUDE spatial database as part of "seamless parcels" layer.
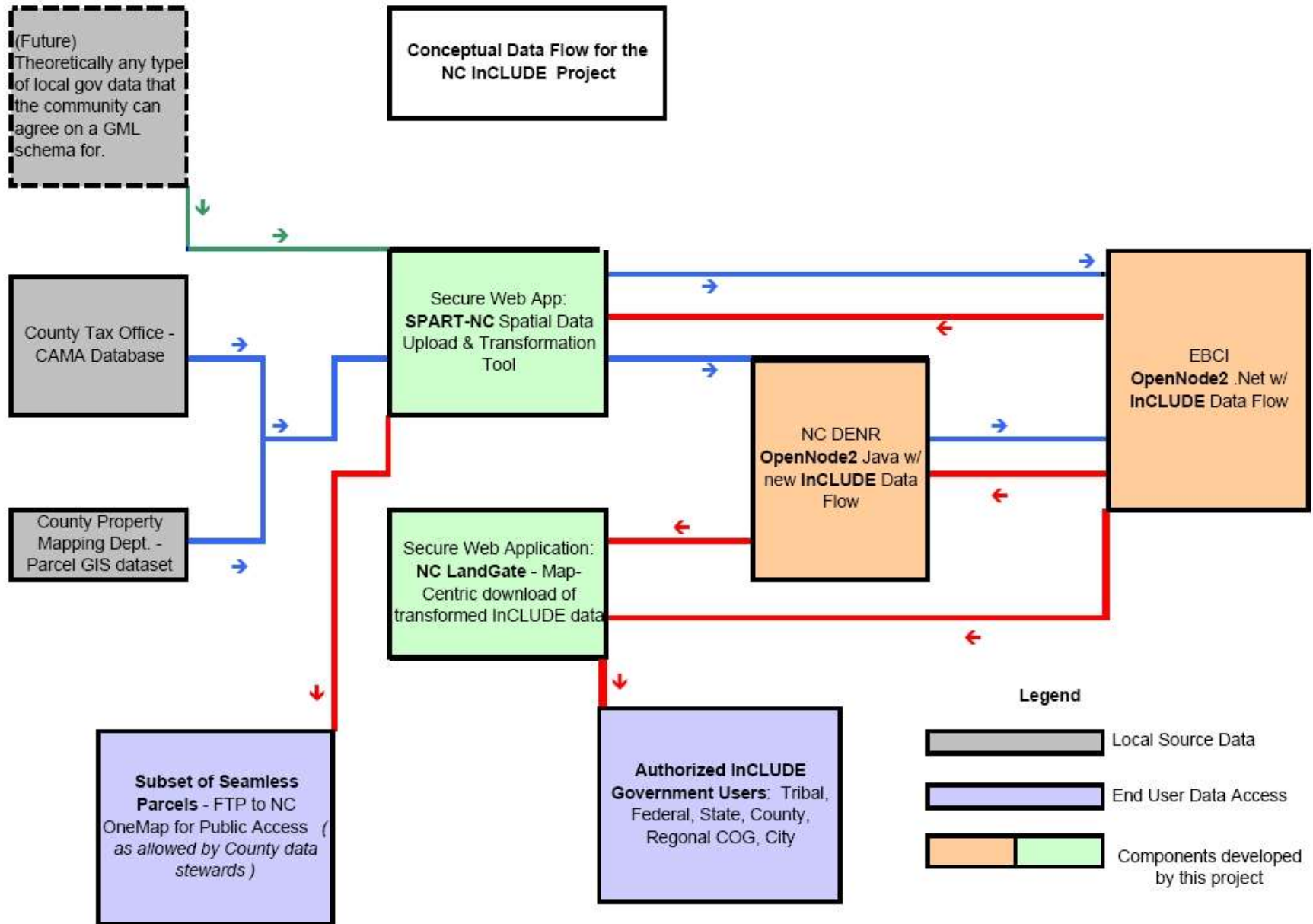
NC LandGate: User Operations

- Authorized Government or Tribal user logs into secure NCLandGate app using NCID or NAAS account.

- Selects a County or Tribal Area Subdivision Name from a pick-list, or zooms on the map interface to the location where they want to get InCLUDE parcel data by one of these methods:
  - Lat/Lon coordinate
  - NC State Plane coordinate
  - US National Grid coordinate
  - Street Address (*Mapstraction* API w/ Google / Yahoo / OSM)
  - Manual Navigation

- Methods for selecting InCLUDE parcel data to download:
  - By County or Tribal Area Subdivision Name
  - By User-Defined Bounding Box (within reason)
  - By User-Defined Center Point & Buffer Radius (within reason)
  - By 10-digit Watershed Boundary ID

NC LandGate: Behind the Scenes

- User specifies GML/Shapefile output & desired Coord. System.

- Web App takes user-defined area and formulates appropriate request, sends it to InCLUDE Node along w/ user credentials.

- InCLUDE logic checks if user has permission to access parcel data for the areas their current data request covers.

- If OK, InCLUDE plugin sends request to secured WFS service (Geoserver or ArcGIS Server), which serves the transformed "seamless parcels" data in DENR or EBCI InCLUDE spatial DBs.

- If user selected GML, the WFS returns GML in the InCLUDE schema. (Geoserver uses App Schema extension, AGS will need custom "assist" code to support SF Level 1).

- If user selected Shapefile format, parcel features with multi-valued properties (ex: a parcel with > 1 owner names) are saved in a "related" DBF file that can be joined to the shapefile.

Proposed Data Flow Diagram

(Future)
Theoretically any type of local gov data that the community can agree on a GML schema for.

Conceptual Data Flow for the NC InCLUDE Project

County Tax Office - CAMA Database

County Property Mapping Dept. - Parcel GIS dataset

Secure Web App: **SPART-NC** Spatial Data Upload & Transformation Tool

Secure Web Application: **NC LandGate** - Map-Centric download of transformed InCLUDE data

NC DENR **OpenNode2** Java w/ new **InCLUDE** Data Flow

EBCI **OpenNode2** .Net w/ **InCLUDE** Data Flow

Subset of Seamless Parcels - FTP to NC OneMap for Public Access (as allowed by County data stewards)

Authorized InCLUDE Government Users: Tribal, Federal, State, County, Regional COG, City

Legend

Local Source Data

End User Data Access

Components developed by this project

Open Source Software planned for use

Functional Area	Software Product
Presentation / Web Client	OpenLayers, MapFish, Mapstraction API (all Javascript-based) http://openlayers.org/ http://mapfish.org/ http://www.mapstraction.com/
Exchange Network	OpenNode2 + NAAS + ENDS http://code.google.com/p/opennode2/ http://www.exchangenetwork.net/
GIS Application Server	GeoServer + 52North GeoDRM services (WSS/WAS) http://www.geoserver.org/ http://52north.org/communities/security/index.html
Spatial Data ETL	GeoKettle http://www.spatialytics.org/geokettle http://www.slideshare.net/tbadard/geokettle-a-powerful-open-source-spatial-etl-tool-5193932
Database	PostgreSQL RDBMS + PostGIS spatial extender http://www.postgresql.org http://www.postgis.org
Webserver	Apache Webserver + Apache Tomcat Servlet Container http:httpd.apache.org and http:tomcat.apache.org
Documentation	Likely either Sphinx or Sandcastle http://sphinx.pocoo.org or http://sandcastle.codeplex.org

Commercial Software planned for use

Functional Area	Software Product
Presentation / Web Client	none
Exchange Network	none
GIS Application Server	ESRI's ArcGIS Server http://www.esri.com/software/arcgis/arcgisserver
Spatial Data ETL	none
Database	Microsoft SQL Server RDBMS + ArcSDE spatial extender http://www.microsoft.com/sqlserver http://www.esri.com/software/arcgis/arcsde
Webserver	Microsoft IIS http://www.microsoft.com/windowsserver2008/en/us/internet-information-services.aspx
Documentation	none