iCORE: A GEOINT Processing Framework and Incubator

GSAW, March 3, 2011

Thomas Kibalo, Principal Engineer/Scientist
Dr. B. Scott Michel, Department Director
Dr. Craig A. Lee, Senior Scientist, lee@aero.org
Nehal Desai, Engineering Specialist, Nehal.Desai@aero.org
Jacob Everist, Member of the Tech. Staff, Jacob.S.Everist@aero.org
The Aerospace Corporation

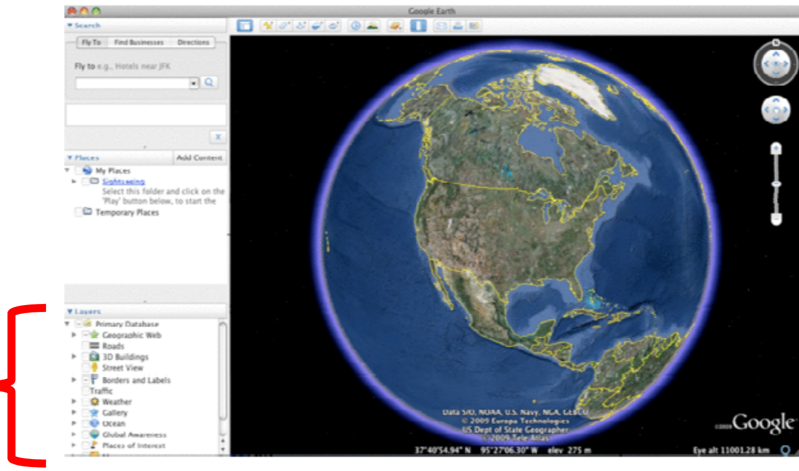
Analytical Environment –The Current Paradigm

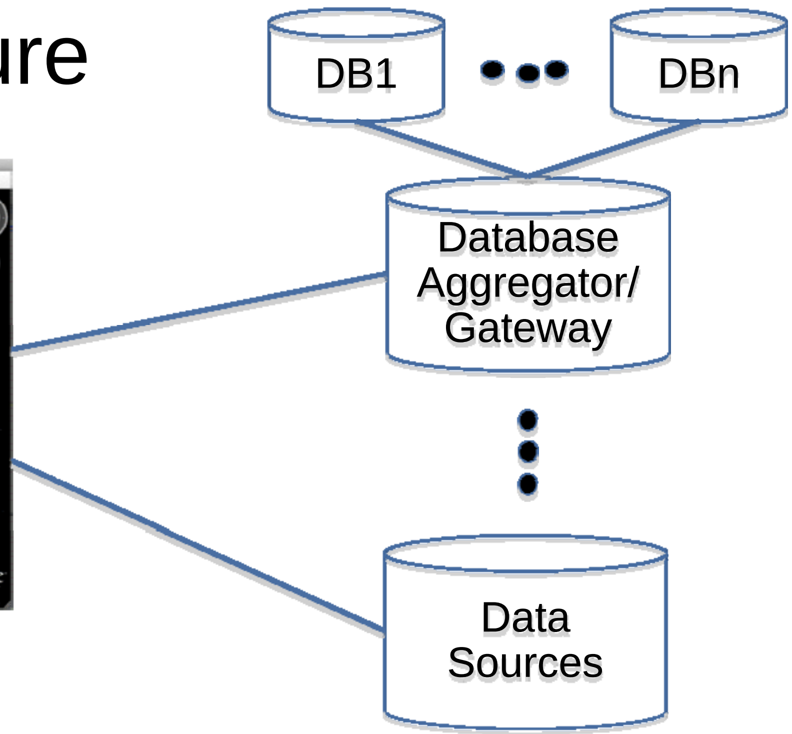
- *“In today’s analytic environment the majority of time is spent on “process” rather then critically thinking about the intelligence question and its analytic solution”*
- The Environment
  - – Compute Power limited to user’s local workstation
  - – Applications limited to specific data types and functions
  - – Limited access to analytic experts
  - – Limited sharing /collaboration
- The Process
  - – Unnecessary retrievals, voluminous data pulls
  - – Too much time spent on mechanics of data access rather than desired analysis

Vision for the Future

- Shift emphasis from retrieval to analysis
  - *Hide the complexities of data retrieval while facilitating the analytical process*
- Move to fully distributed, virtualized architecture, i.e., *clouds*
  - *Provide as much data access as possible, and as much processing power as necessary to answer the analytical query at hand*
- Document and share the analytic tradecraft
  - *Publish, share and reuse data products and tools among experts*
- *Leverage NRO "CORE" -- M. Rothman, SATL Lab*
  - *Consolidated Operations Research Explorer*
  - *A thin client visualization tool for all geo-referenced metadata*
- *Integrate iCORE -- "Intelligent CORE"*
  - *Add analytical capabilities to CORE using cloud-based processing engines for model-based analytics, i.e., workflows*

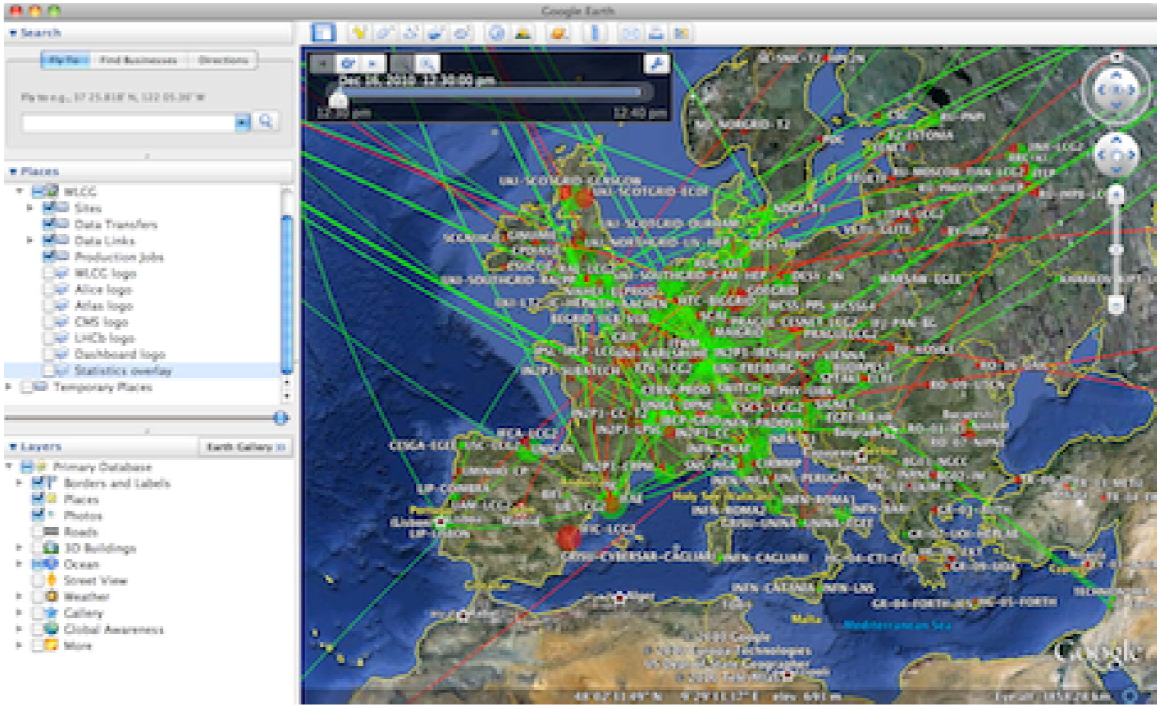
Basic CORE Architecture





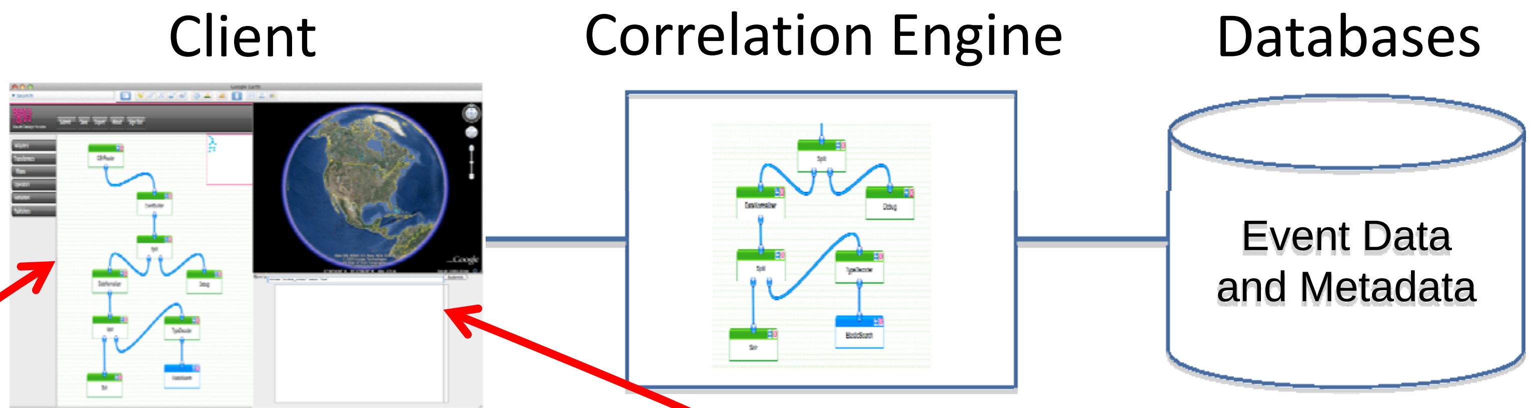
- Google Earth-based tool for accessing and displaying data from dozens of databases
- Available data sources given on the indented “tree” menu
  - *When user selects a check-box, GE retrieves that data from the remote data source using KML “network links” and displays it*
- CORE is ~41k SLOC of KML, in addition to remote scripts for actual data access

A “CORE” Issue: *Information Overload!*



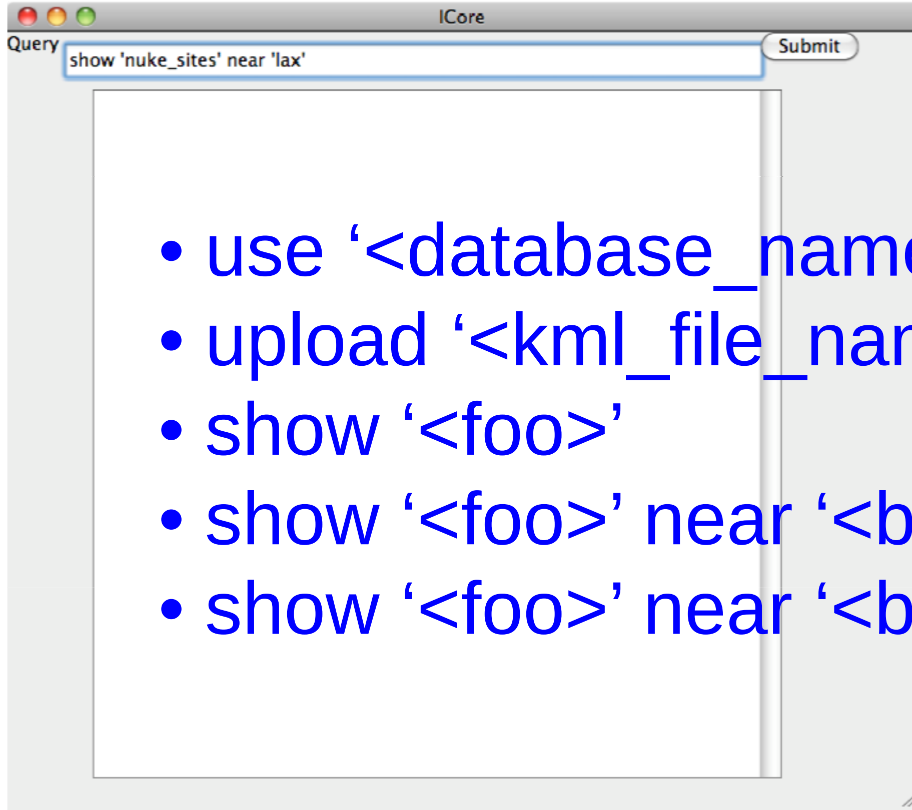
- Too easy to get overwhelmed with too much data
- Difficult to know what data is available in the “tree” menu
- CORE has no analytical capabilities
- Analysts need tools to help them “connect the dots”
- ***Intelligent CORE: iCORE***

Basic iCORE Architecture

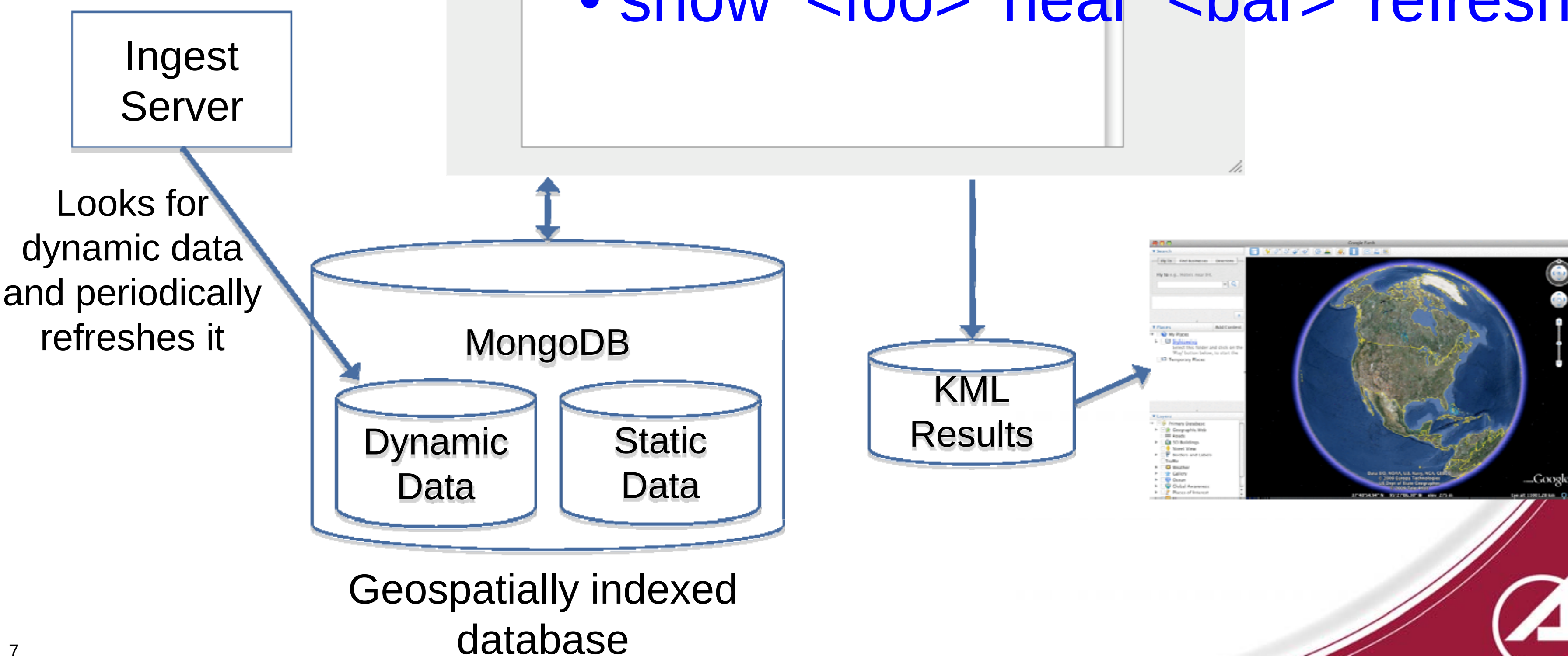


- Google Earth augmented with
  - *Domain Specific Language pane*
    - Simple geospatial query language, e.g., "X near Y"
    - *Word completion* based on data available over KML links
  - *Visual Programming pane*
    - Analyst can "wire together" processing chains based on palette of available functions to access, process, and display data
    - Workflow -- model-based analytics -- run in dynamically hosted *Correlation Engines*

Domain Specific Language and Text Completion Area

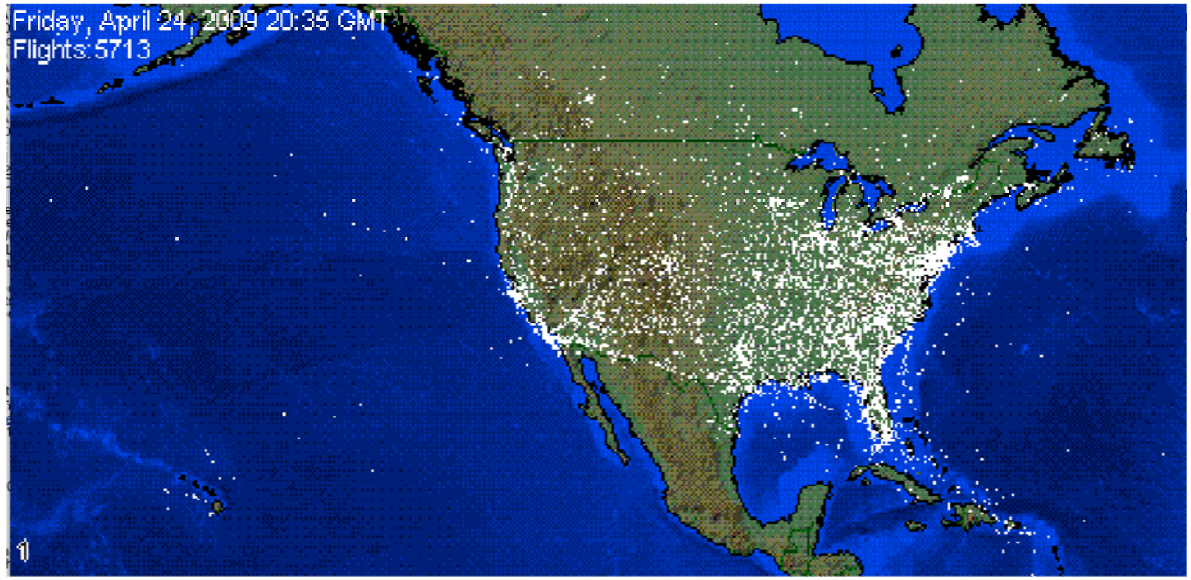


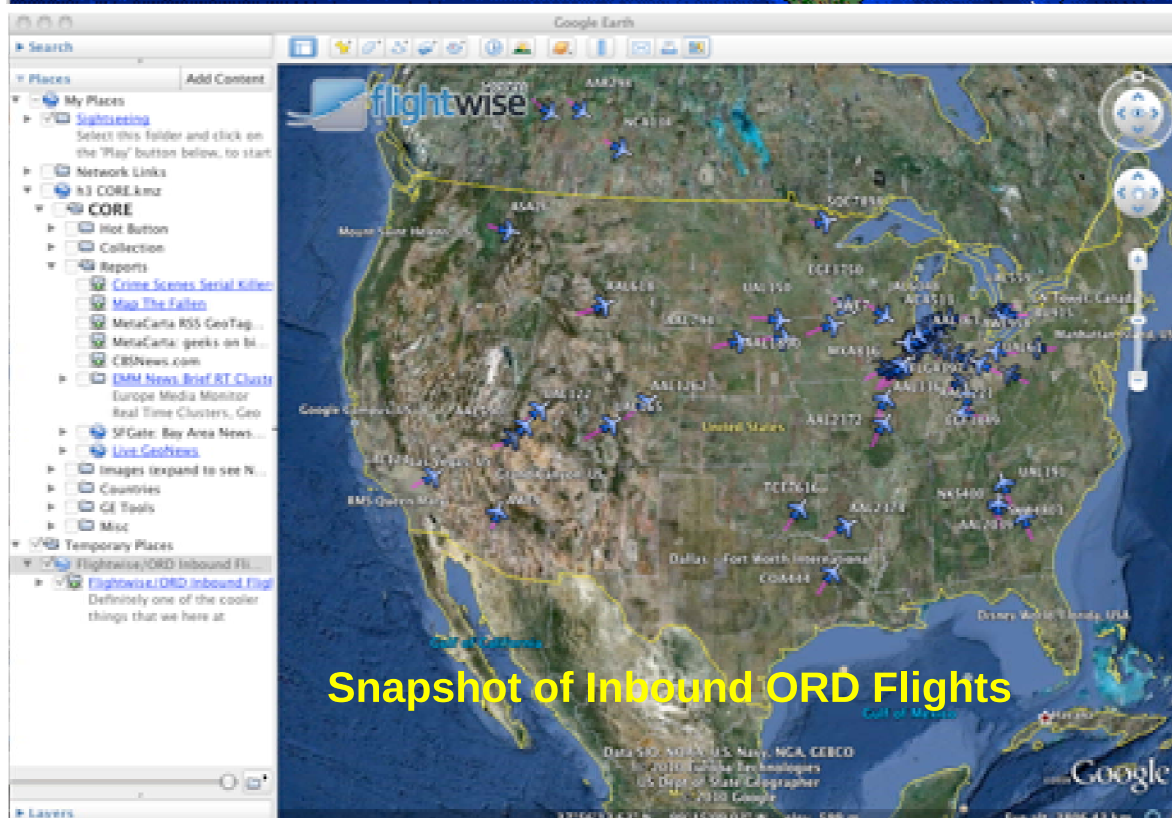
- use ‘<database_name>’
- upload ‘<kml_file_name>’
- show ‘<foo>’
- show ‘<foo>’ near ‘<bar>’
- show ‘<foo>’ near ‘<bar>’ refresh



An Example: Proximity Detection, i.e., “No-Fly Zone”

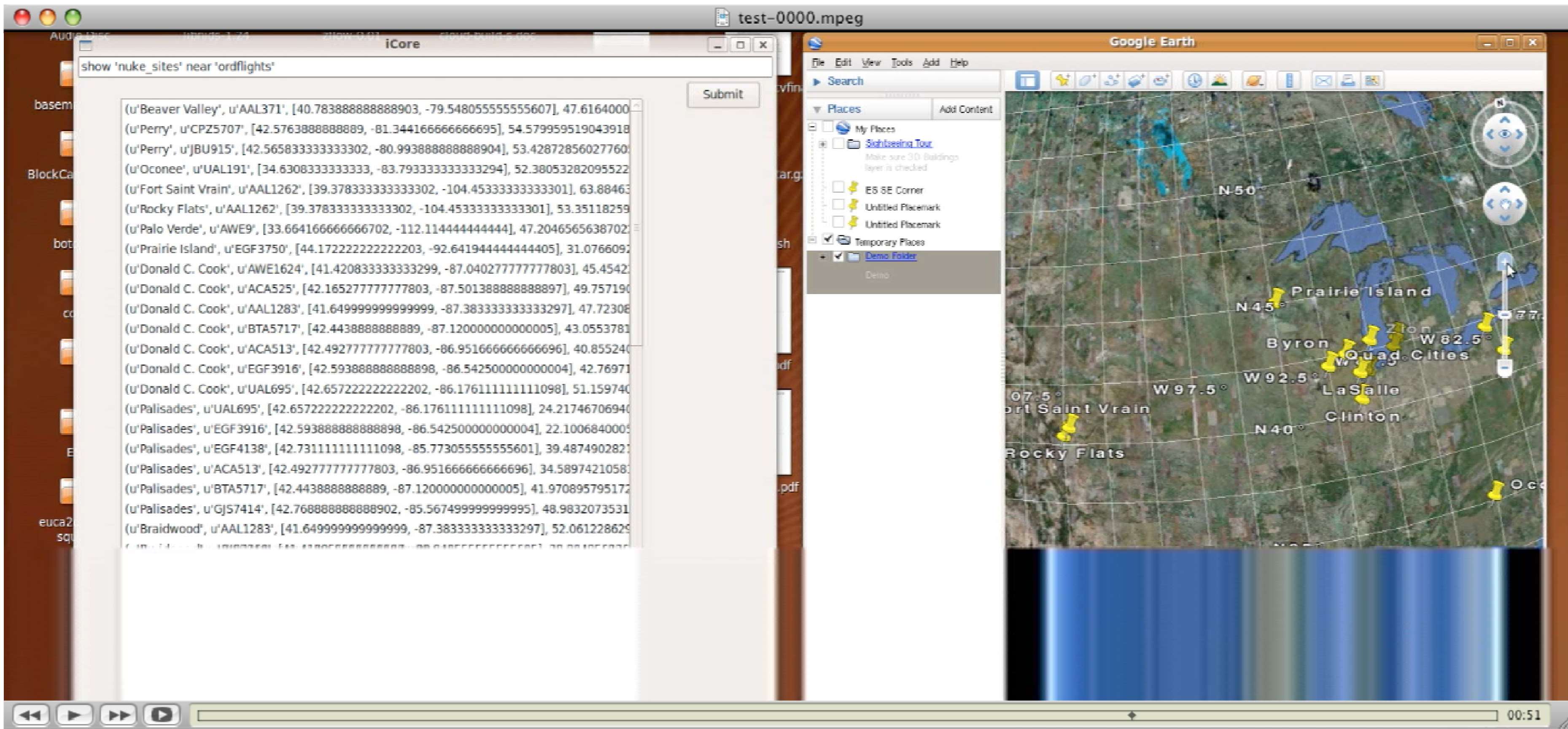
- *Correlate live CONUS flight paths with No-Fly Zones*
  - *Nuclear power plants, military installations, etc.*
- *Thousands of flights at any one time*
  - *Current flight status available on-line*
  - *Update periodically, e.g., every 60 seconds*
  - *Compare all flights against all No-Fly Zones to identify incursions*





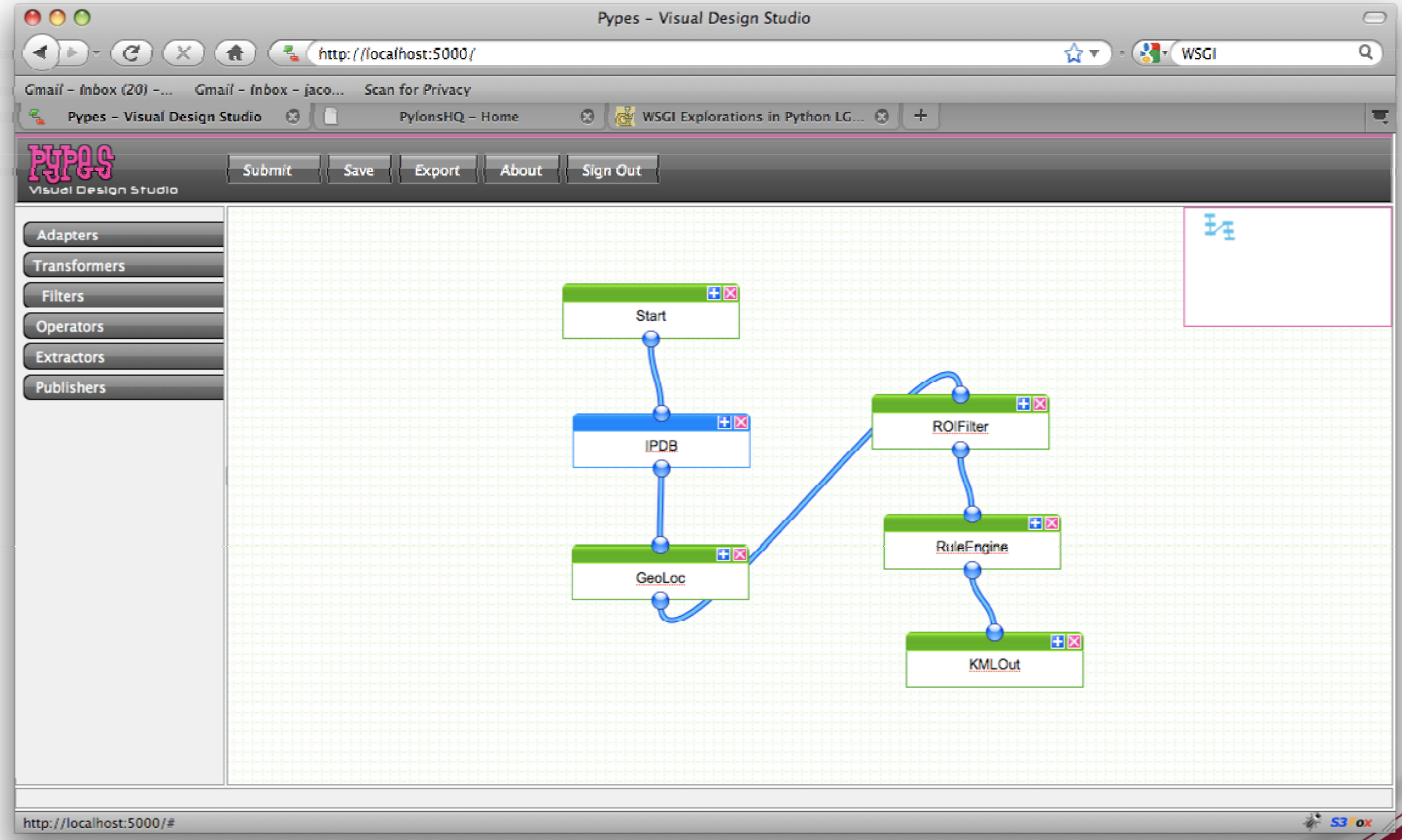
The iCORE Proximity Detection Demo

“ show ‘nuke_sites’ near ‘ordflights’ ”



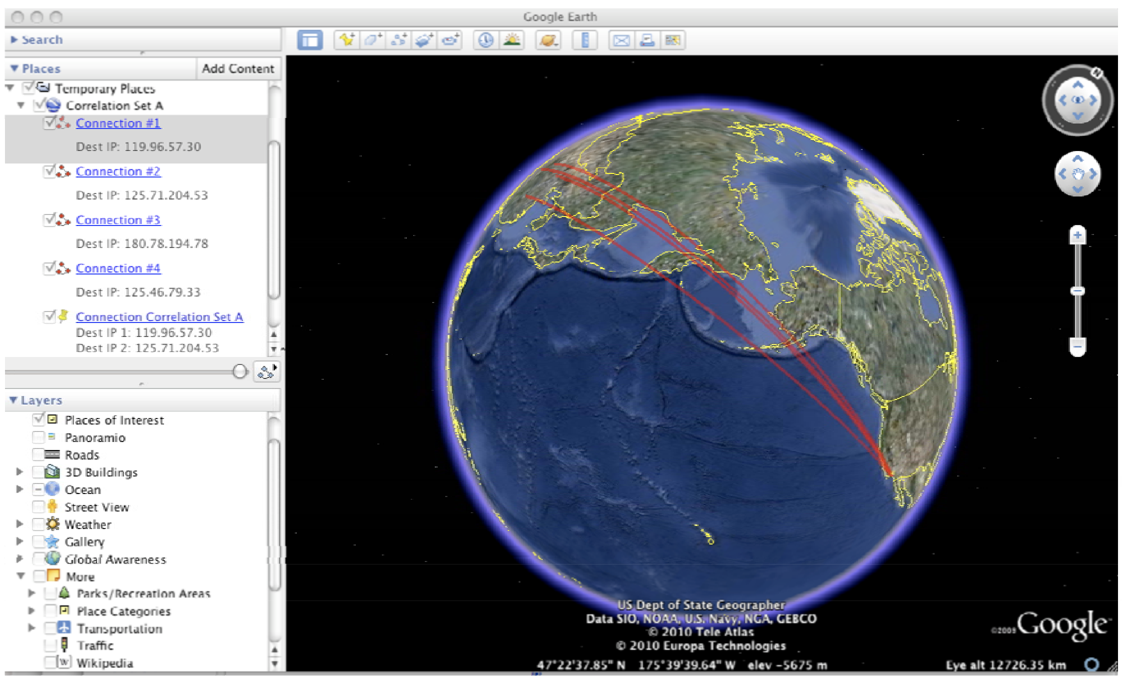
The iCORE Visual Programming Interface

- *Pypes*: python-based visual programming tool
- Palette of functions available to compose CE semantics
- Can be used as a user interface and workflow manager

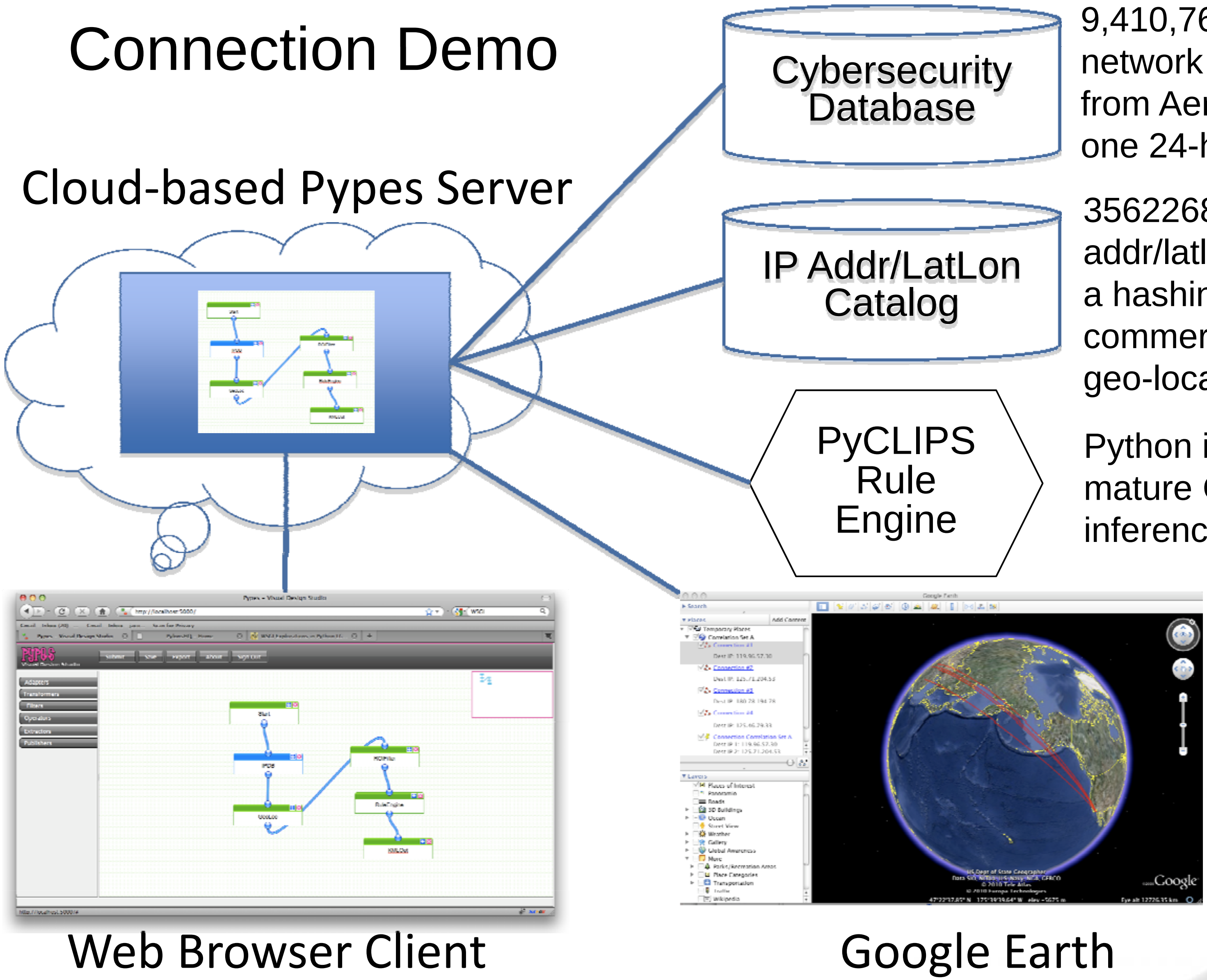


An Example: Network Connections

- Identify connections in a network connection log DB to a given Region of Interest during a given Time Period
- Use Rule Engine to identify “Connection A” events followed by “Connection B” events within a given *Δt*



The iCORE Network Connection Demo



mySQL DB with 9,410,769 outgoing network connections from Aerospace in one 24-hour period

3562268 specific IP addr/latlon mappings in a hashing scheme for commercial, city-level geo-location

Python interface to mature CLIPS inference engine

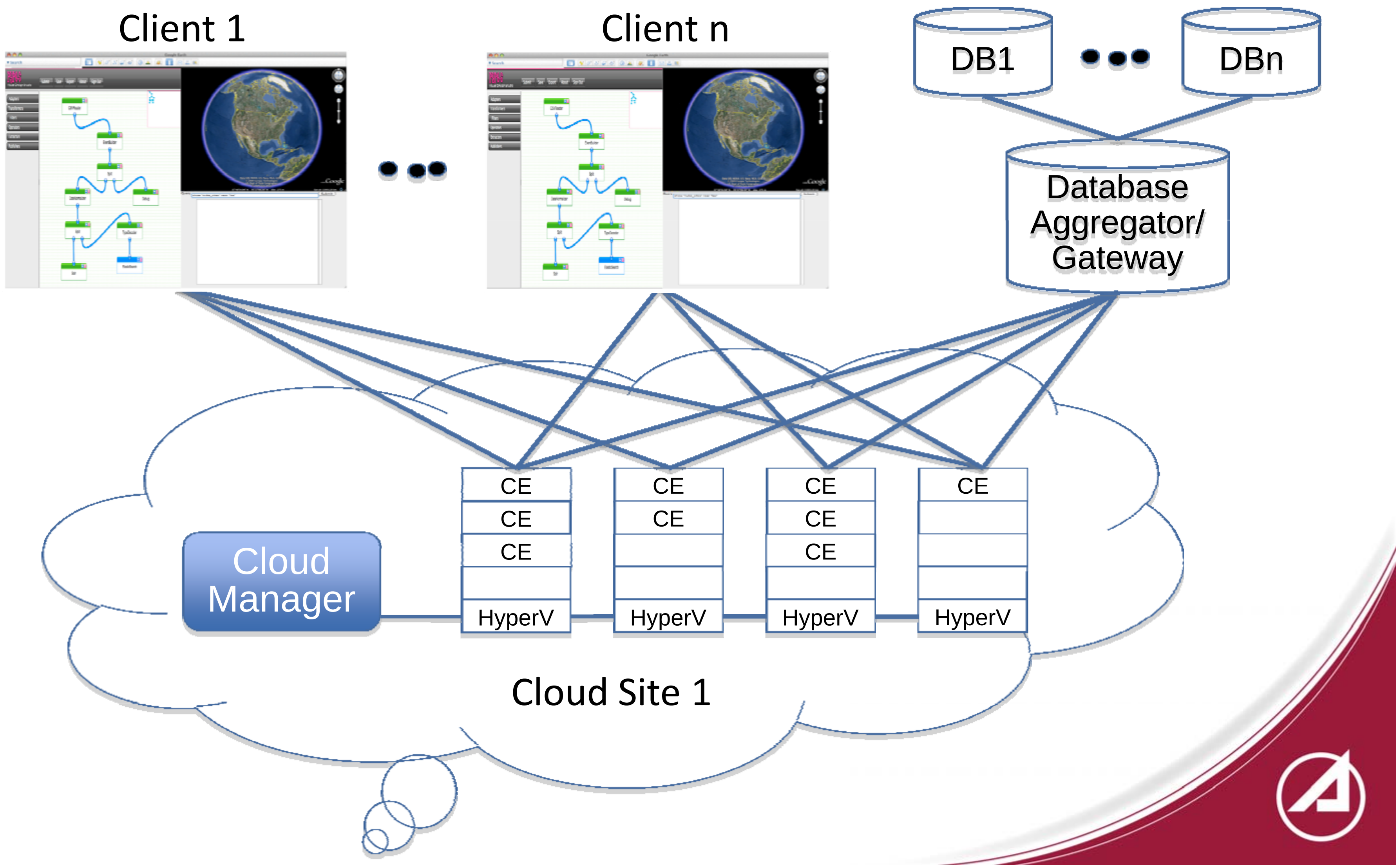
The iCORE Framework as an *Incubator* to Explore Issues in End-User Capabilities *and* IT Infrastructure

- Query Semantics
  - *Spatial, temporal, event types, fusion queries*
- Cloud Computing Resources
  - *Correlation Engines are spun-up on-demand*
- Autonomic Cloud Workflows
  - *Automatic management of application and cloud performance*
- Disadvantaged Users
  - *User in the field on a mobile device with low bandwidth has on-demand access to massive data and processing power*
- Distributed Data Management
  - *Enforce data policy across sites*
  - *Analyst recommender systems*
- Security and Virtual Organizations
  - *Users can only see and operate on data they are permitted to see: role-based authorization*

Aerospace Eucalyptus Private Cloud

Eucalyptus: open source Amazon EC2/S3 API cloud originally from UCSB
Intel Nehalem boxes, 168 cores total
Head Node: 8 cores, 16GB
Worker Nodes: (1x) 16 cores, 33GB, (6x) 24 cores, 33GB

Ultimate Goal: iCORE Released to a Beta User Community Running "At Scale" on a Cloud



Conclusions

- We have developed iCORE as a Framework and an Incubator for exploring analytical tools and computing infrastructure for geospatial intelligence
  - *Geospatial query language, disadvantaged users, ...*
  - *Cloud computing, autonomic control, security & governance, …*
- In close collaboration with the original CORE team
- Demonstrated at GEOINT 2010, New Orleans
- Potential exists for integration with other systems
  - *DIB, DCGS-IC, DSS*
- Much more in-depth information is available
- Contact us: lee@aero.org