Virtual Sensor Web Infrastructure for Collaborative Sensing (VSICS)

ESTC
June 26, 2008

Prasanta Bose and Ryan Timmons
Advanced Technology Center
Lockheed Martin Space Systems Company
(LMSSC, ATC)

Outline

- **Motivation and challenges**
- Technical approach
  - Objectives, conops, and architecture
  - Software implementation
- Use cases
  - Lightning prediction
  - Fire risk warning
- Conclusions and future work

Decadal Missions Motivate Effective Management of Sensors and Observations

- Sensor proliferation across land, air, and space

  "*The constant dilemma of the information age is that our ability to gather a sea of data greatly exceeds the tools and techniques available to sort, extract, and apply the information we've collected" - Jeff Davidson, MBA & author of Breathing Space*

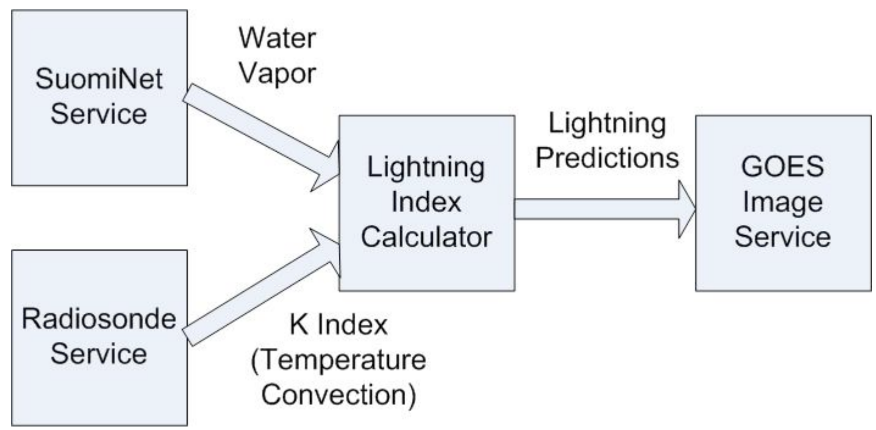
- Catastrophic events such as Katrina, Summit fire, and others cost life, property, and impact the earth we live in

- Need a sensorweb infrastructure that enables timely integration of sensors and observations via
  - Creation of services providing location-specific information
  - Flexible discovery and workflow based composition of multi-domain services
  - Distributed and robust execution of such workflows for on-demand predictions

On-Demand Location-Based Lightning Prediction is Key to Mitigation of Severe Weather Effects

- Probable lightning strikes key to predicting
  - Weather delays in launch and flight operations
  - Associated severe weather events such as tornadoes
  - Ignitions of wildfires
- Theory and models for lightning prediction
  - Involve multiple types of atmospheric data across domains – water vapor and K Index (measure of convection from temperature at different altitudes)
  - Can be modeled as workflow
- Need a sensorweb infrastructure that enables
  - Creation of services encapsulating local sensors providing location-specific information
  - Flexible discovery and workflow based composition of multi-domain services
  - Distributed and robust execution of such workflows for on-demand lightning predictions

$$Index = \frac{1}{1 + \exp(c_0 + c_1(PWV) + c_2(9h\Delta PWV) + c_3(KIndex))}$$



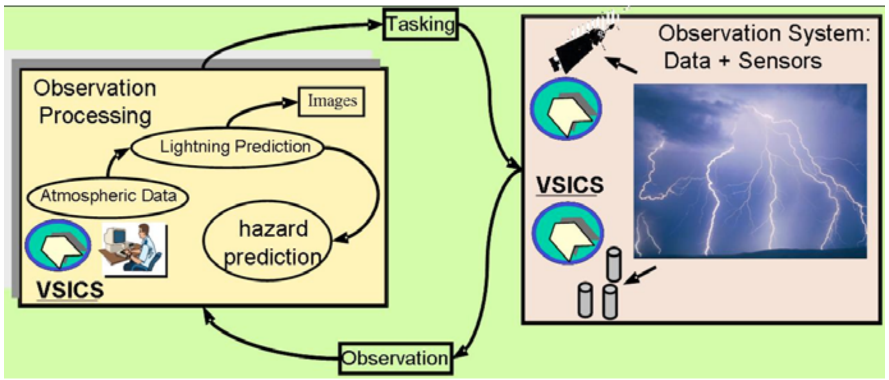
Challenge is workflow-based dynamic composition and robust coordination of disparate and distributed information providing services of physical sensors

Outline

- Motivation and challenges
- **Technical approach**
  - **Objectives, conops, and architecture**
  - **Software implementation**
- Use cases
  - Lightning prediction
  - Fire risk warning
- Conclusions and future work

VSICS Objective and Technical Approach

- *Objective:* Develop robust and scalable sensorweb middleware infrastructure that enables
  - *Flexible composition of distributed services* driven by tasks
  - *Decentralized coordination of services* for scaleable, event-driven execution and adaptive performance in response to evolving context
  - *Semantic descriptions of services* based on functional and non-functional attributes for optimal management

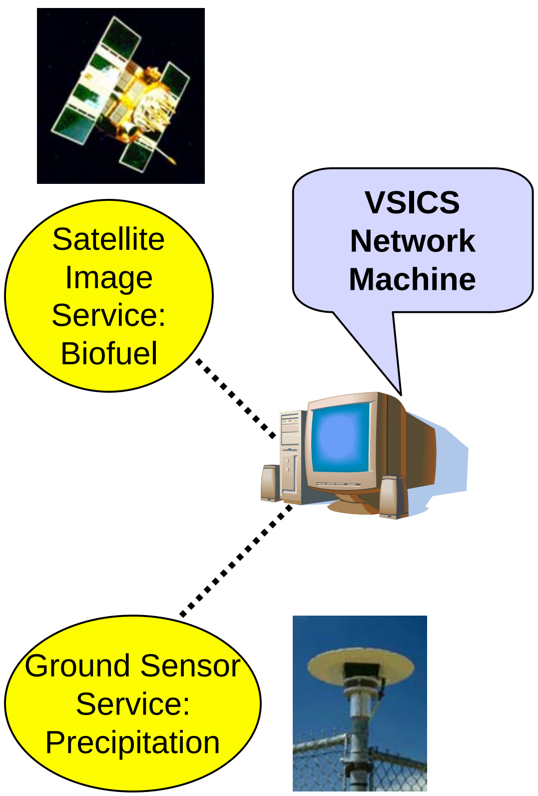


- *Technical Approach: VSICS architecture is based on*
  - *Services:* Abstraction of system that provides data of scientific value.
    - Sensor services providing raw data or modeling/prediction techniques
    - Defined primarily by data types of inputs and outputs
  - *Workflows:* Data-centric coordination of distributed services
    - Collection of services that together perform a prediction or monitoring task
    - Service links based on data inputs and outputs, using publish/subscribe model
  - VSICS middleware allows users to register and discover services, assemble them into task specific workflows executed to provide timely information

Concept of Operations (A Day in the Life of VSICS)

1. Provision of Services

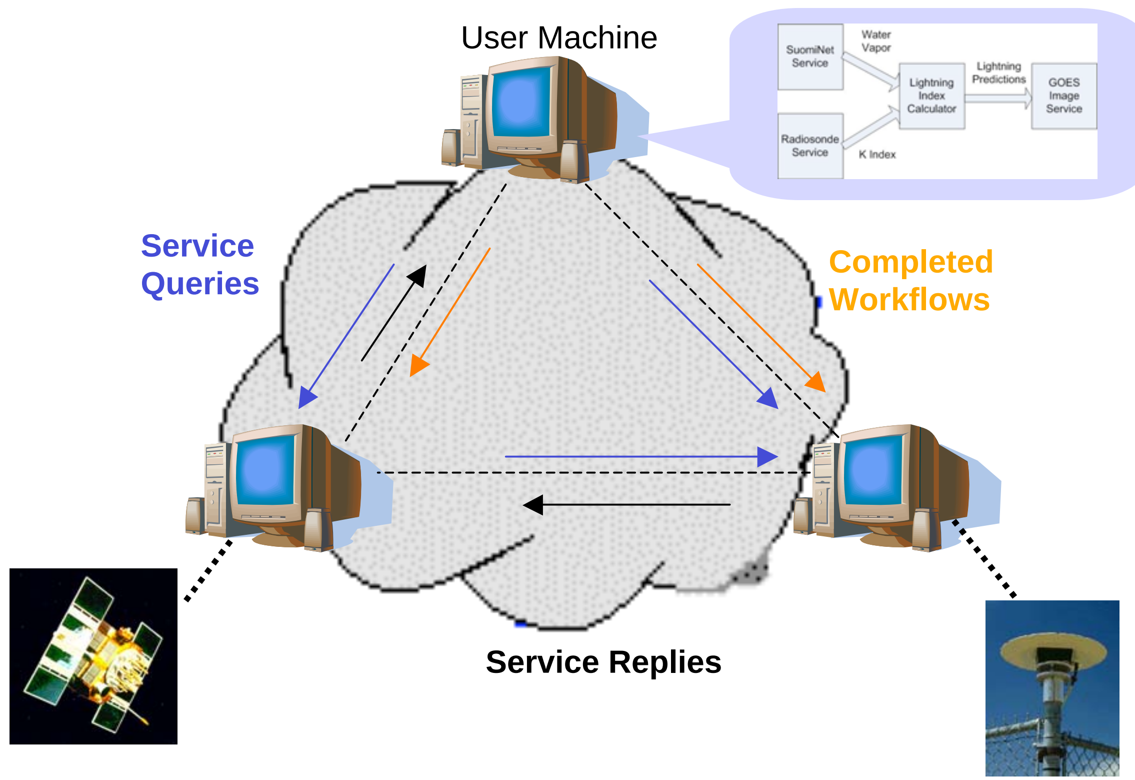
Service providers register services for use on the VSICS network



2. Discovery and Workflow instantiation

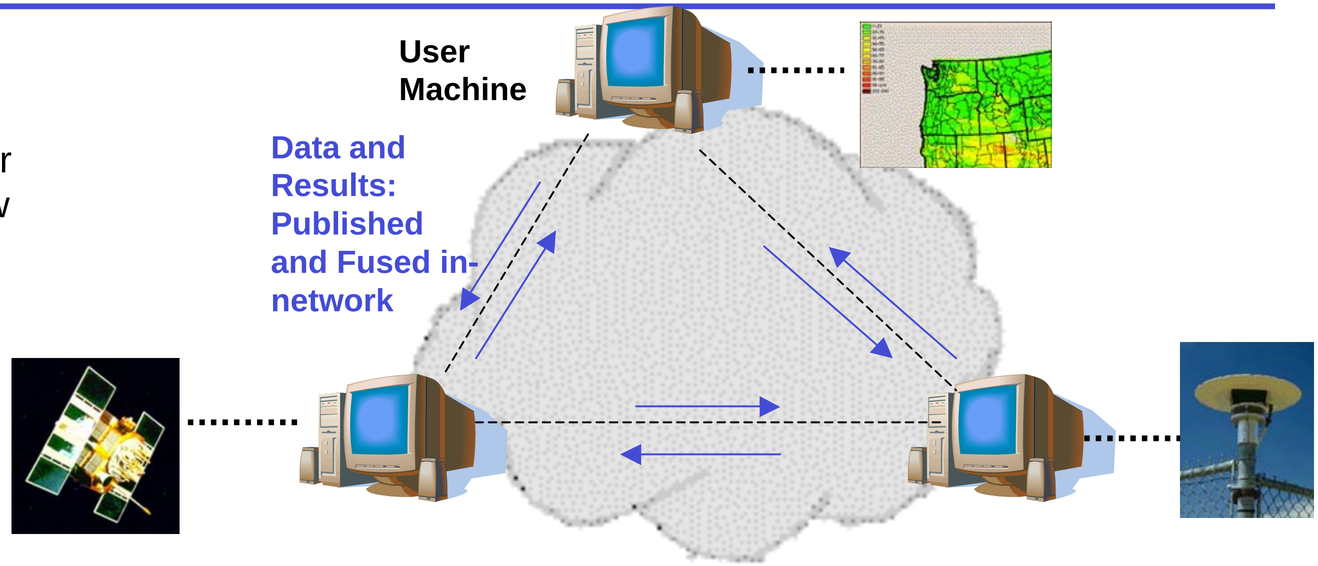
2.1 User defines partial fusion/processing workflow to realize his task needs, control, and display of results

2.2 System searches VSICS network for services that meet workflow requirements, returns a set of workflows ranked by nonfunctional attributes score

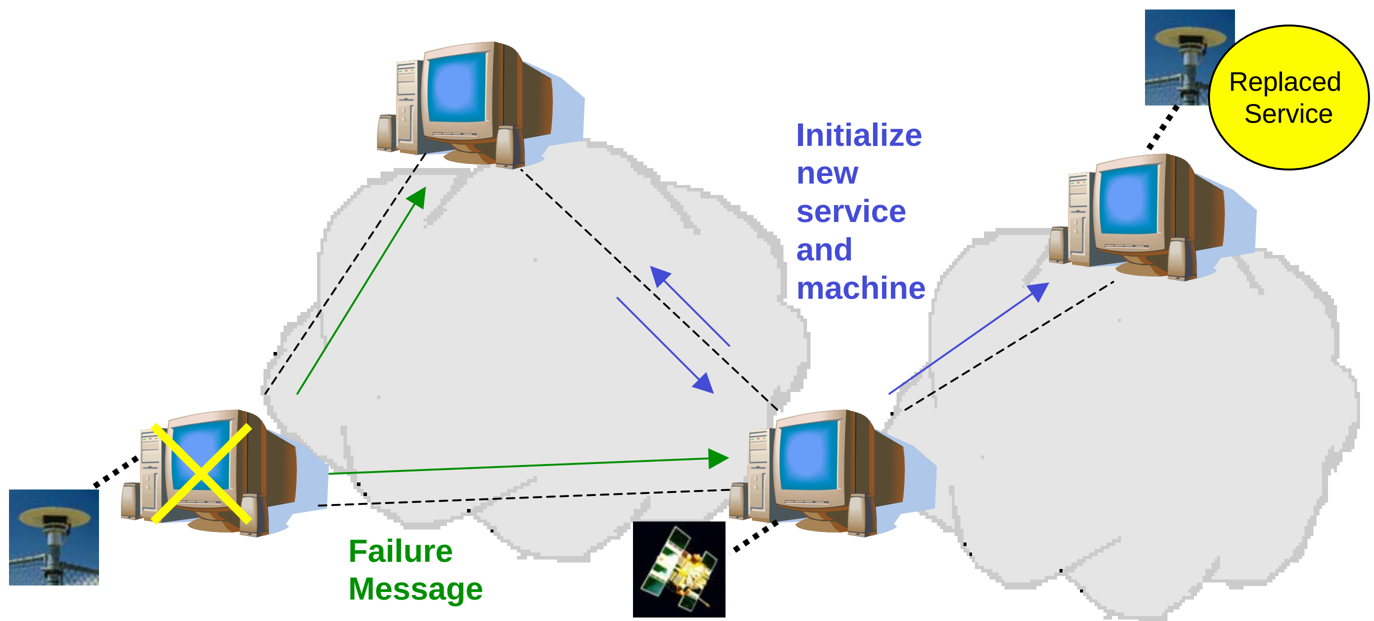


VSICS Concept of Operations (Contd.)

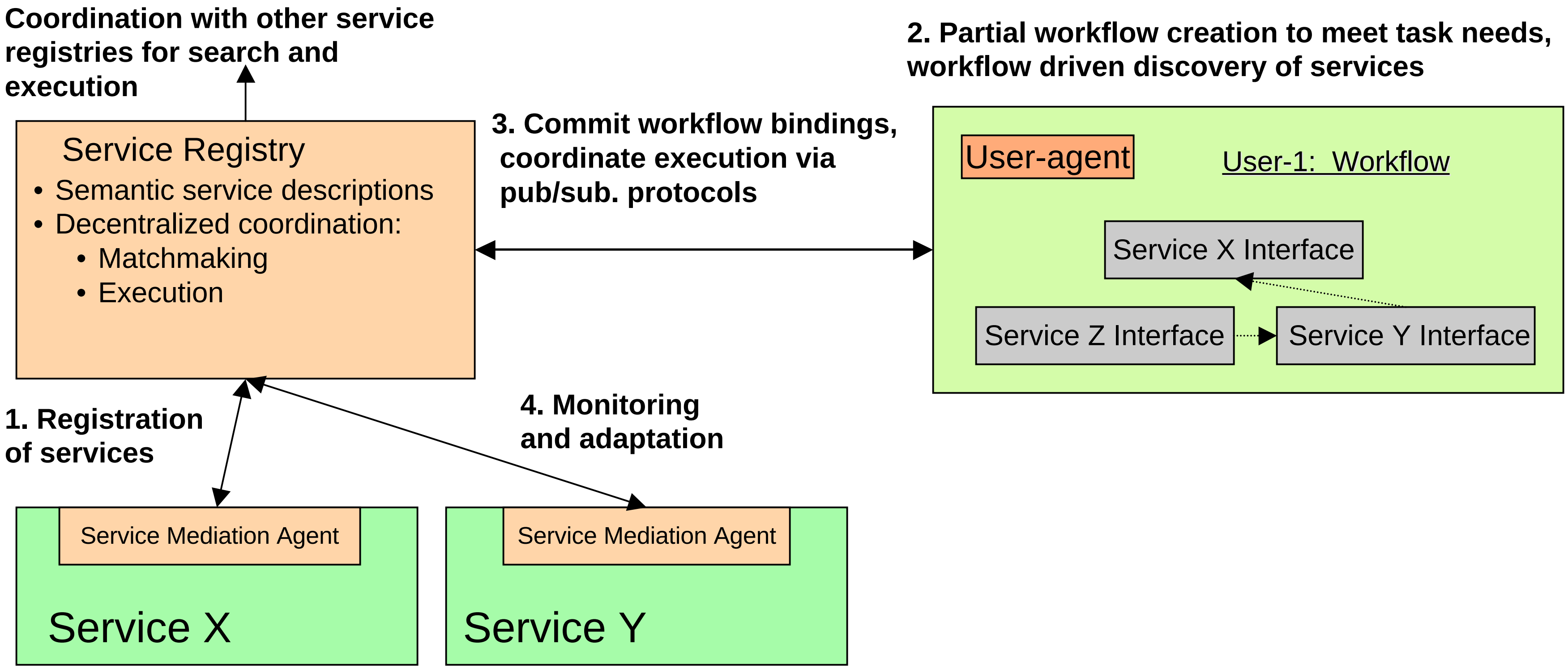
3. User commits desired workflow specific service bindings consistent with their priorities and utility; workflow executes, system returns results to user



4. During execution, VSICS middleware automatically reconfigures the workflow in case of service failure, overload, or network problems; new services added into workflow replacing unusable ones



VSICS Architecture Performs Decentralized Coordination of Distributed Services for Scalability



- **Decentralized Architecture requires:**
  - **Network infrastructure for implementing peer-to-peer distributed collaboration via discovery of services**
  - **Protocols for communication of data during execution**
  - **Event-driven coordinated execution of services as per workflow semantics**

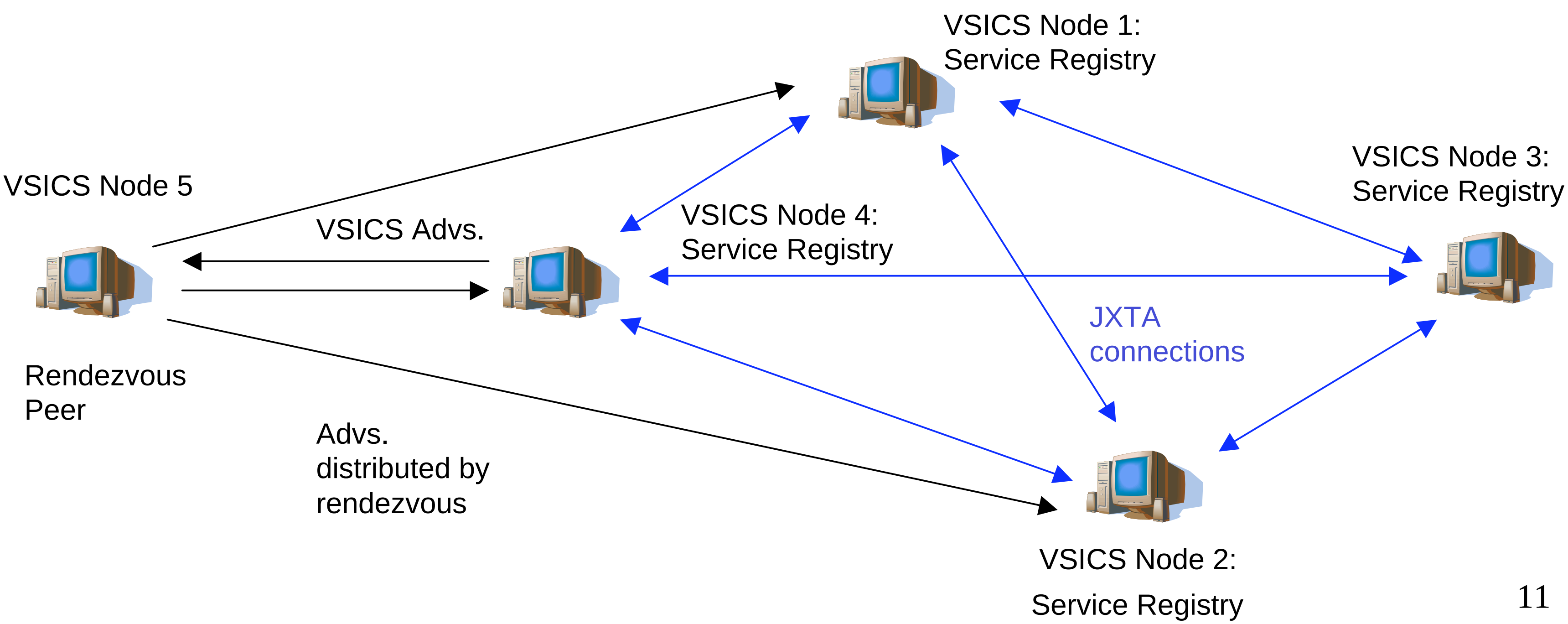
Decentralized coordination leads to scalable VSICS but introduces challenges of distributed workflow and resource management

Outline

- Motivation and challenges
- Technical approach
  - Objectives, conops, and architecture
  - **Software implementation**
    - **Distributed peer-to-peer computing and collaboration infrastructure**
    - **Message and protocol design for content-driven composition**
    - **Service registration, publish/subscribe based workflow composition and distributed execution framework**
- Use cases
  - Lightning prediction
  - Fire risk warning
- Summary and future work

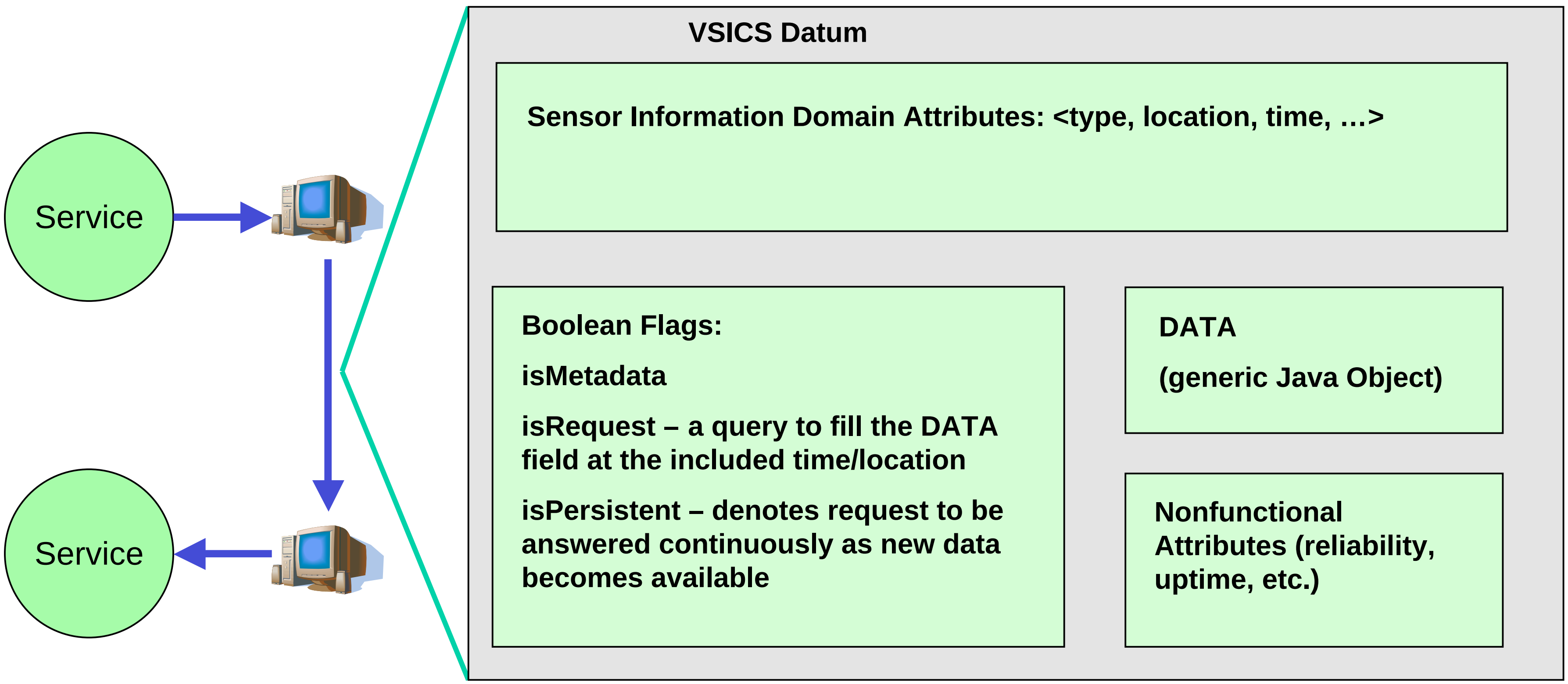
VSICS Software for Distributed Computing Exploits JXTA Peer-to-Peer Networking Protocols

- Each computer running service registry is peer; rendezvous used to guarantee peer visibility/connectivity
- JXTA's customizable advertisement system allows each peer to find all operational VSICS registries
- Registries discover each other to form a decentralized network
- Services registered in each peer-node get coordinated by corresponding coordinators for workflow binding and distributed execution
- Publish/subscribe protocols used for coordination is realized by the JXTA advertisement and response handling mechanisms



VSICS Software for Content-driven Composition and Adaptation

- VSICS Datum serves as universal container for service metadata and data
- Common tags for earth science/weather domain specific services: <time, position on the earth's surface, non-functional attributes, datatype attributes>
- Service interfaces register data types they consume and provide

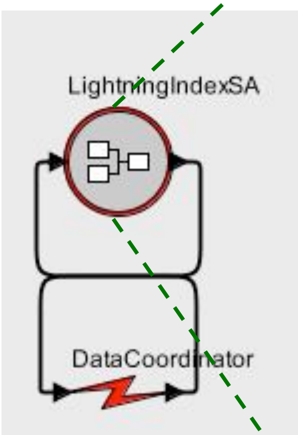


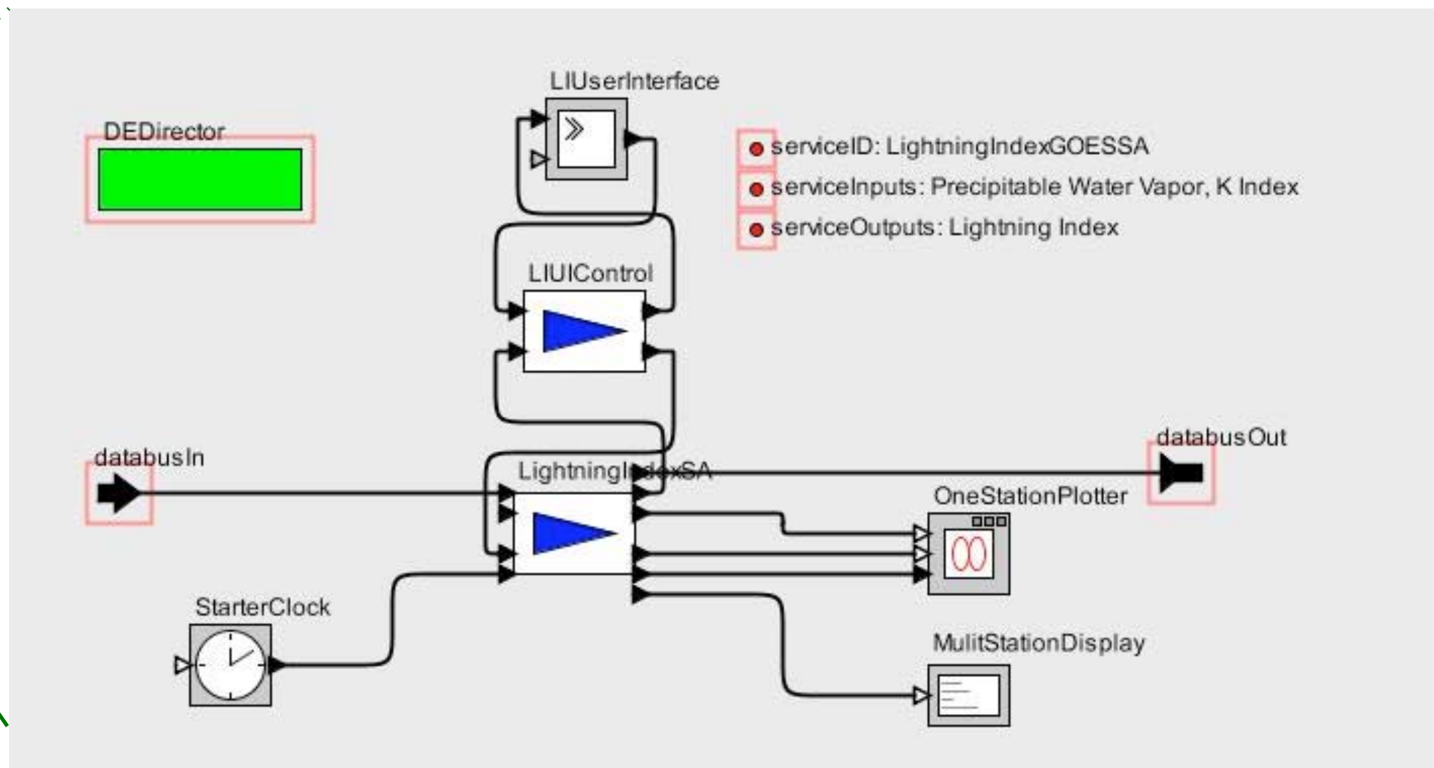
Content-Driven Flexible Composition Achieved via Semantically Tagged Messages and Service Registration

VSICS Software For Distributed Workflow Execution Coordination

- Extends the Ptolemy modeling and execution environment (http://ptolemy.berkeley.edu/ptolemyII/)
  - Ptolemy’s framework of actor and dataflow-oriented programming realizes VSICS’s workflow driven service coordination
  - VisualSense editor provides visual representation of workflows
  - Our Service Registry actor connects to the JXTA network for distributed workflow execution

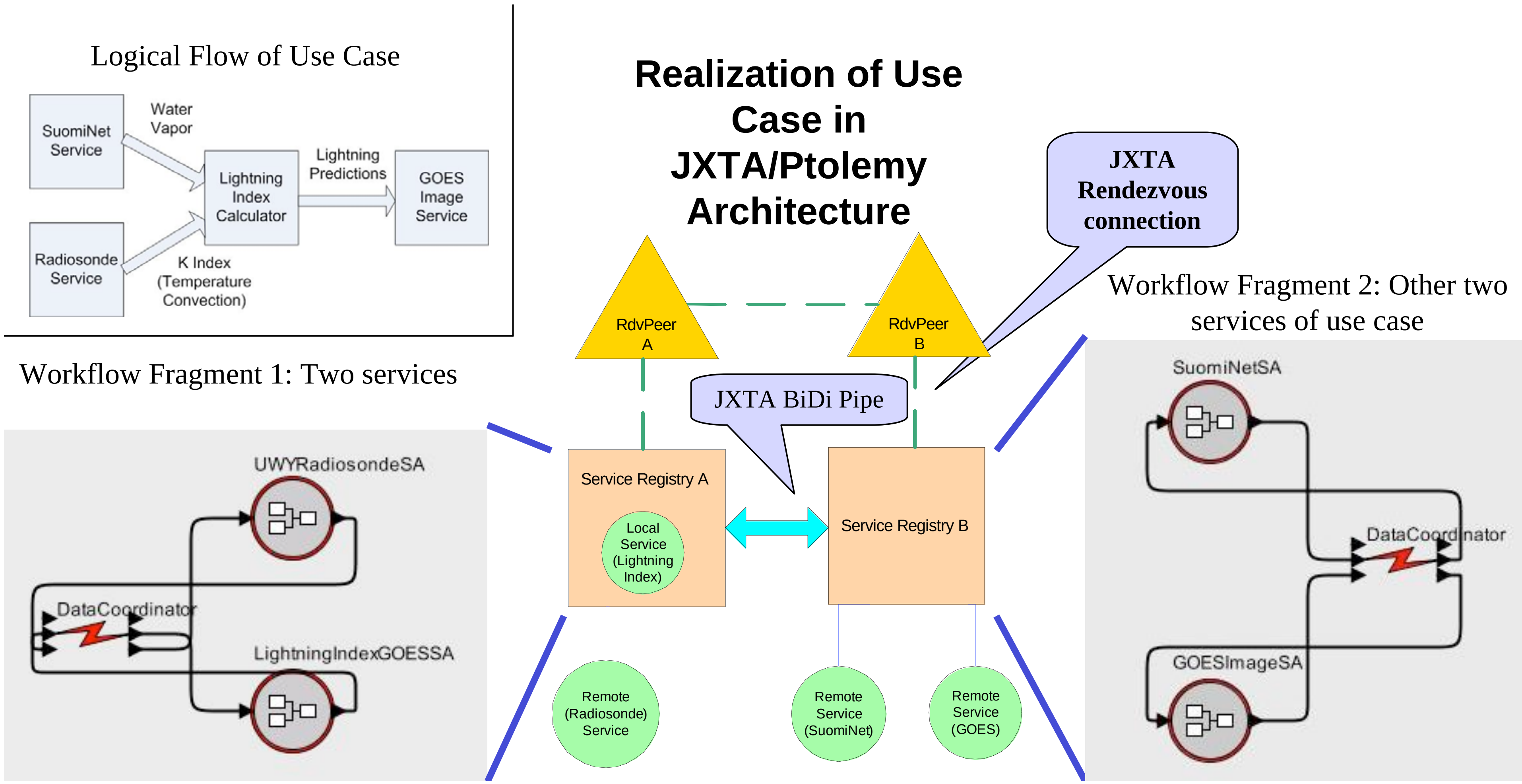
Example of a VSICS service built in Ptolemy



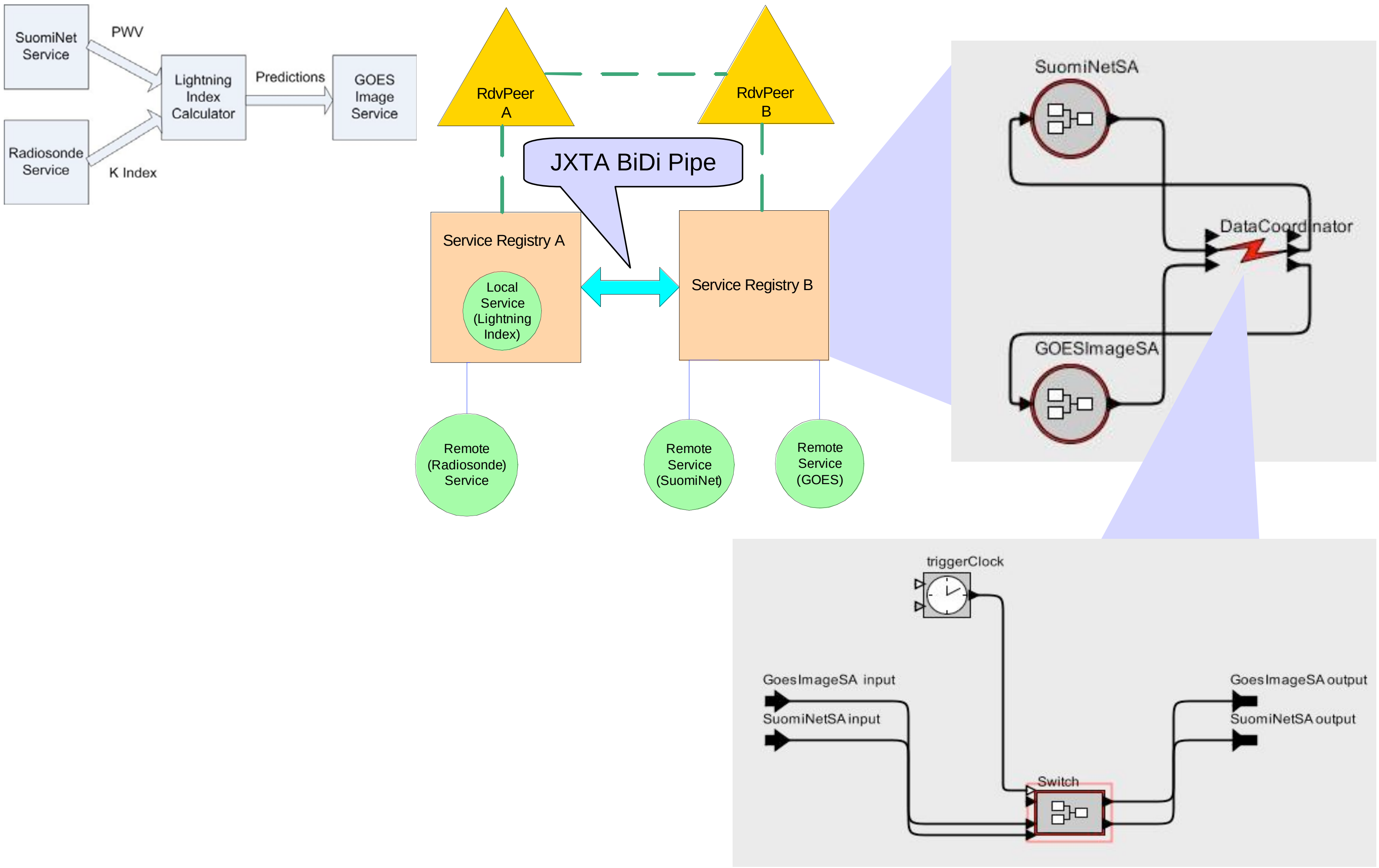


VSICS Software for Distributed Workflow Execution Coordination (Contd.)

- Workflows split into pieces depending on location of constituent services
- Each partial workflow has ***VSICS data coordinator*** managing communication according to publish/subscribe model
- Each registry node runs Ptolemy instance that executes any number of partial workflows



VSICS Software for Distributed Workflow Execution Coordination (Contd.)

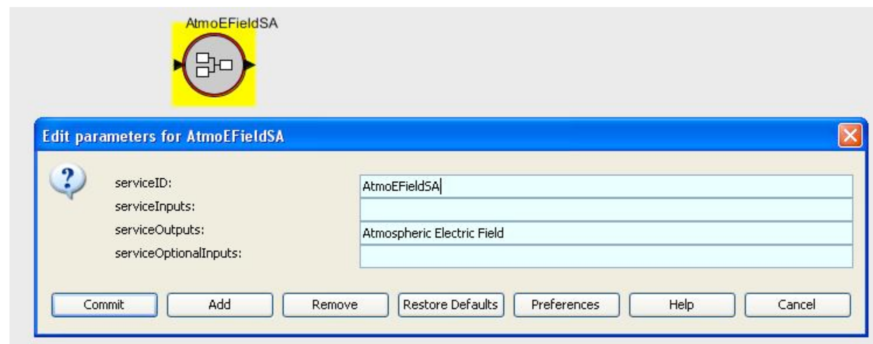


VSICS is Highly Extensible: Addition of User Provided Services

- Templates provided so those with data sources/predictors can rapidly create Ptolemy actors enabling their use in VSICS
- Architecture depends on services having common interface
- Runtime interface allows other users to quickly discover and utilize new services
- Example: Constructing service to provide atmospheric electric field data (can be used by lightning predictor)

1. Assumed starting point: Program that gives current electric field at some set of locations

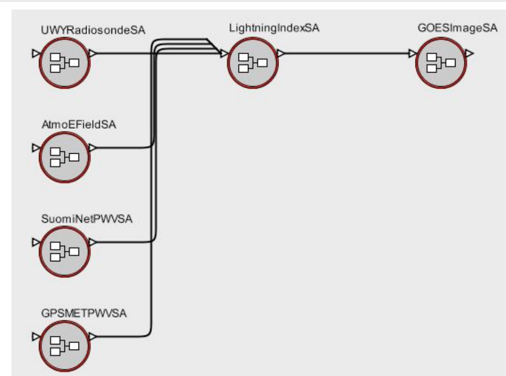
2. Define VSICS service interface from template – given a location and time range, this service will return the maximum electric field



3. Open VSICS service base code and quickly define behavior using API of existing implementation

```
public void fire() throws IllegalActionException {
    super.fire();
    //get input token
    if (input.getWidth() > 0 && input.hasToken(0)) {
        ObjectToken ot = (ObjectToken)input.get(0);
        VSICSDatum vd = (VSICSDatum)ot.getValue();
        DataDescription type = vd.getDataDescription();
        //TODO: insert cases here
        if (type.equals(DataDescription.ATMO_EFIELD) &&
                !vd.isMetadata() && vd.request) {
            vd.request = false;
            vd.data = source.getAtmosphericElectricFieldMax(
                    vd.location, vd.time , vd.endTime);
            sendDatum(vd);
        }
    }
}
```

4. Add service to registry and rerun search to add it to workflow



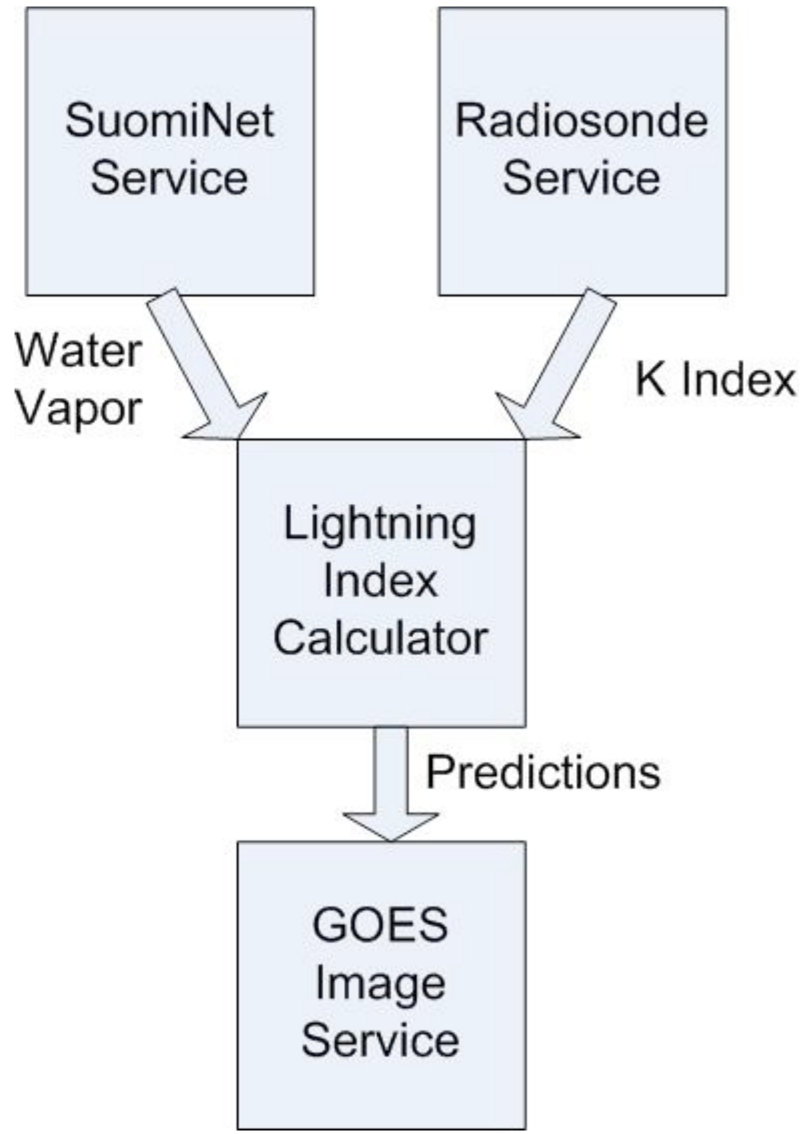
Researchers can easily add new services to VSICS sensorweb for discovery, composition and usage by other researchers

Outline

- Motivation and challenges
- Technical approach
  - Objectives, conops, and architecture
  - Software implementation
- **Use cases**
  - **Lightning prediction**
  - **Fire risk warning**
- Conclusions and future work

Lightning Prediction Use Case Overview

- Involves two sources of atmospheric data
- Predicts first strike in near future
- Predictions used to selectively retrieve satellite imagery of cloud conditions during and leading up to thunderstorms
- User begins by selecting the (end product) image service - VSICS discovers others to make it work
- Workflow runs in realtime, making predictions and downloading images as data becomes available

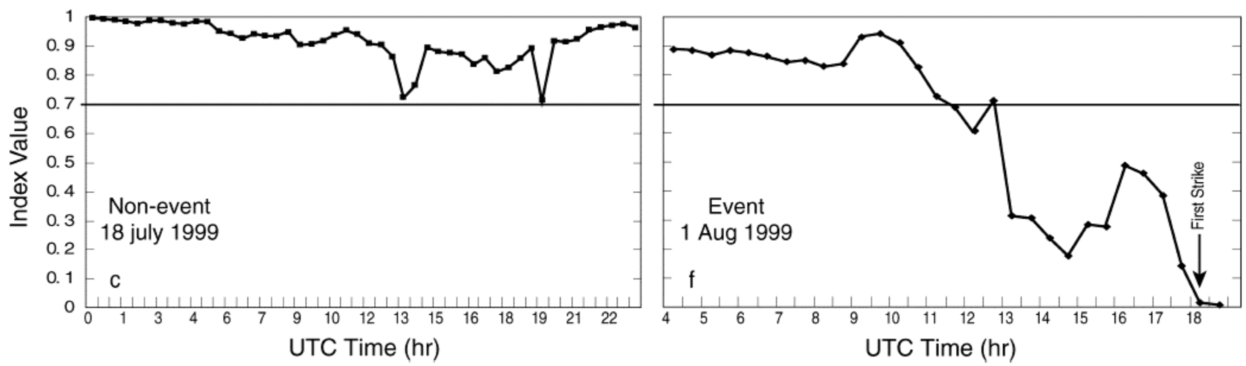


Use Case: Lightning Prediction Model

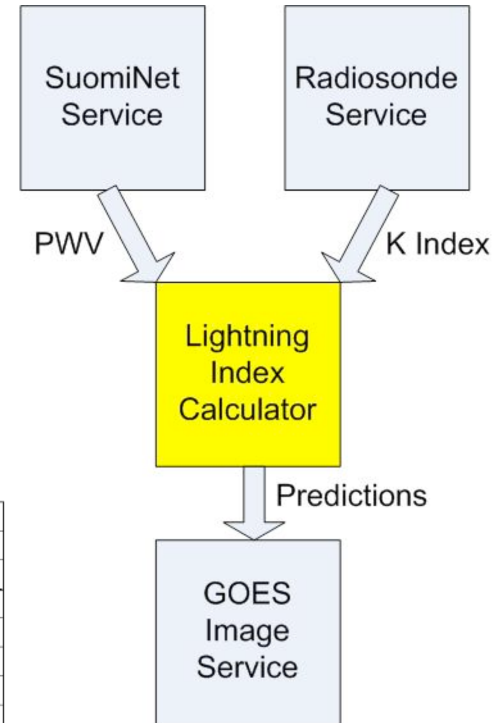
- Adapted from work by Mazany, et. al.
- Weighted logistic regression on 3 factors – precipitable water vapor (PWV), 9-hour ΔPWV, and *K* index

$$Index = \frac{1}{1 + \exp(c_0 + c_1(PWV) + c_2(9h\Delta PWV) + c_3(KIndex))}$$

- Index drops below 0.7 -> predicts first strike 1.5-12 hours from present
- Authors applied at Cape Canaveral
- We apply at many locations



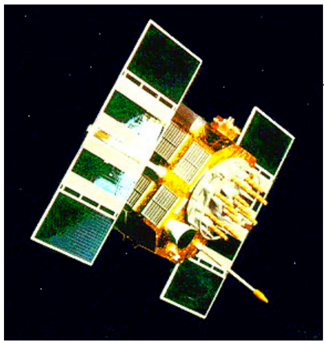
From "A Lightning Prediction Index that Utilizes GPS Integrated Precipitable Water Vapor", R. Mazany, S. Businger, S. Gutman, W. Roeder.

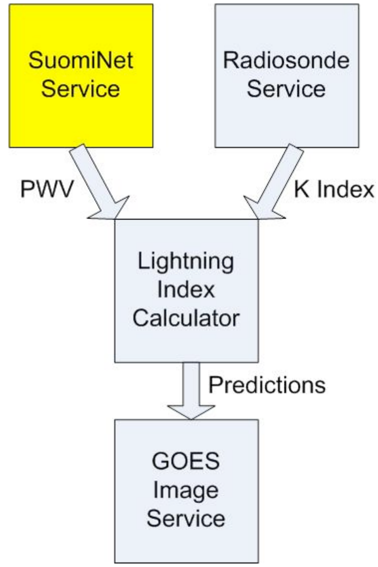


Theory of lightning events provides a foundation for systematic integration of low-level services for high-level decisions

VSICS Integrates SuomiNet Precipitable Water Vapor (PWV) Data Source

- Data service run by University Corporation for Atmospheric Research since 2000
- Measures by delays in GPS satellite signals
- Data every ½ hour at more than 300 North American sites

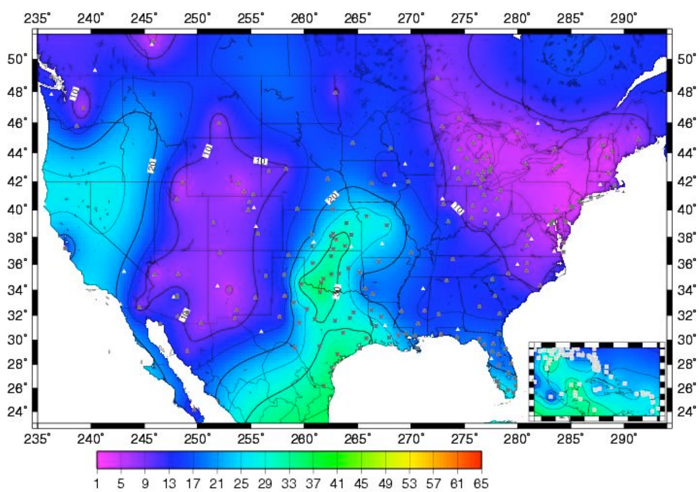




GPS receiver measures refractive delay in satellite's signal

Right: Sample map of SuomiNet data

PWV 15h-16h 03/17/08

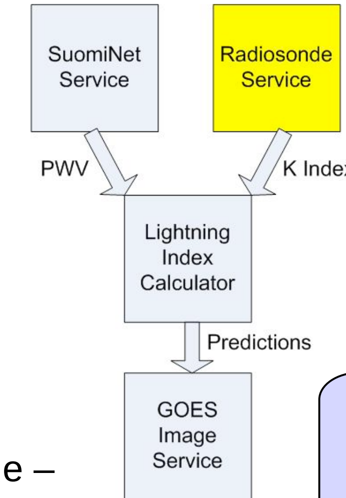


Example of data format

Site	PWVmidTim	Duration	PW	FMerr
SSSS	YYYMMDD/HHMM	MIN	[mm]	[mm]
ACU5	20080317/0415	30	8.2	0.1
ACU5	20080317/0445	30	8.2	0.1
AL20	20080317/0415	30	12.4	0.2
AL20	20080317/0445	30	13.8	0.2
AL40	20080317/0415	30	10.3	0.2
AL40	20080317/0445	30	10.2	0.2
AL70	20080317/0415	30	10.2	0.2
AL70	20080317/0445	30	10.0	0.2
AL90	20080317/0415	30	11.3	0.2
AL90	20080317/0445	30	11.1	0.1

VSICS Integrates Univeristy of Wyoming Radiosonde / K index Archive

- K index = $T_{850} - T_{500} + T_{d850} - (T_{700} - T_{d700})$ (d for dewpoint temperature; numbers = pressure in millibars)
- 850mb ≈ 4800 ft, 700mb ≈ 10000 ft, 500 mb ≈ 18000 ft
- Higher values indicative of thunderstorms
- University of Wyoming maintains data archive of 100 North American stations since 1997
- Data collected every 12hrs



Map of radiosonde stations

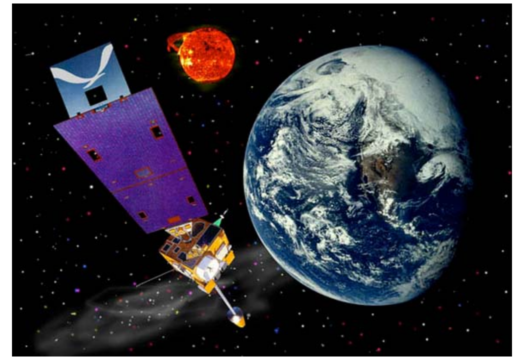
Example radiosonde – small sensor package attached to balloon

Example of data format

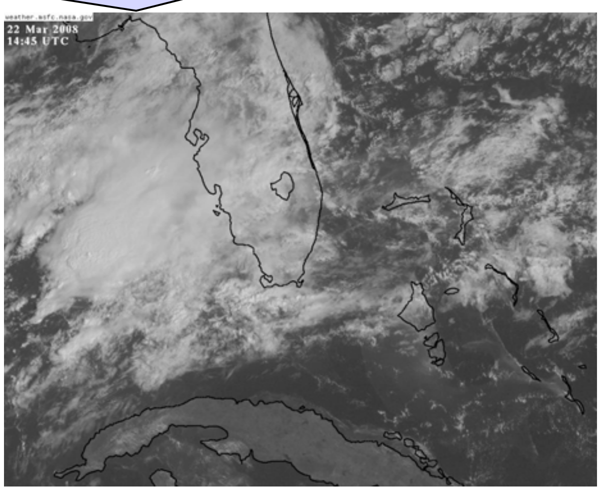
```
-----------------------------------------
   PRES   HGHT   TEMP   DWPT
    hPa      m      C      C
-----------------------------------------
 1000.0    189
  925.0    831
  920.0    872    0.4   -3.7
  914.0    925    1.0   -3.7
  881.3   1219    1.1   -6.9
  872.0   1305    1.2   -7.8
  850.0   1511   -0.3   -9.3

  Station identifier: BOI
      Station number: 72681
    Observation time: 080317/1200
    Station latitude: 43.56
   Station longitude: -116.22
   Station elevation: 872.0
     Showalter index: 14.86
        Lifted index: 15.68
virtual temperature: 15.81
         SWEAT index: 46.99
             K index: -12.10
```

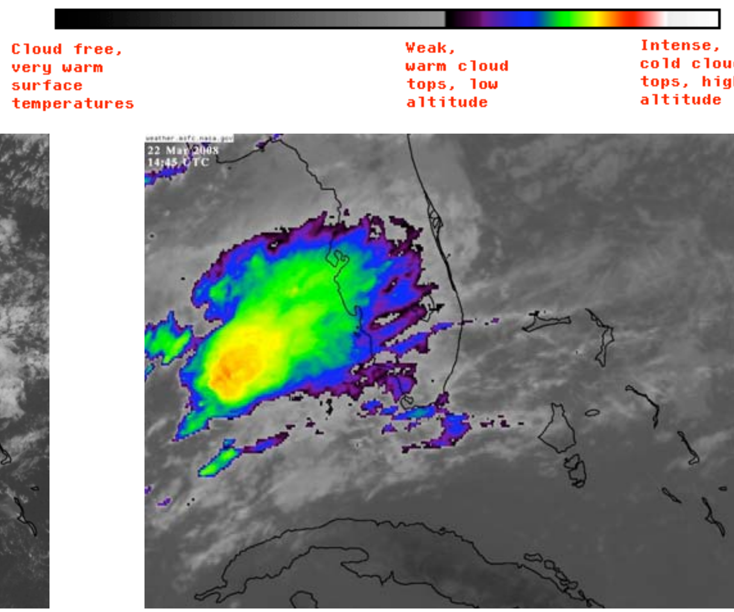
VSICS Integrates GOES Satellite Imagery

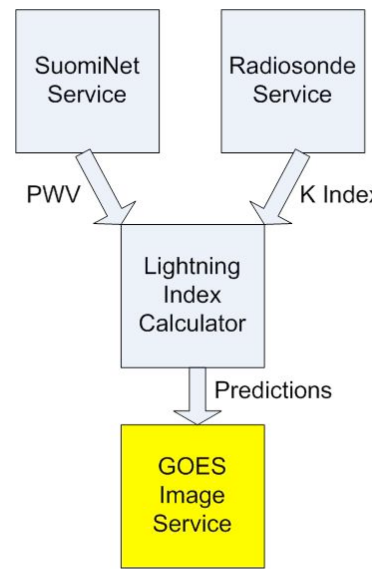
- Visible and infrared images from GOES I-M imager available for any continental US location
- Images useful to researchers investigating what cloud formations lead to thunderstorms and associated severe weather (tornadoes, wildfires)
- Volume of data is enormous and human analyst time is limited
- VSICS enables automatic filter to get images associated with lightning predictions



GOES visible image during thunderstorm in Gulf of Mexico





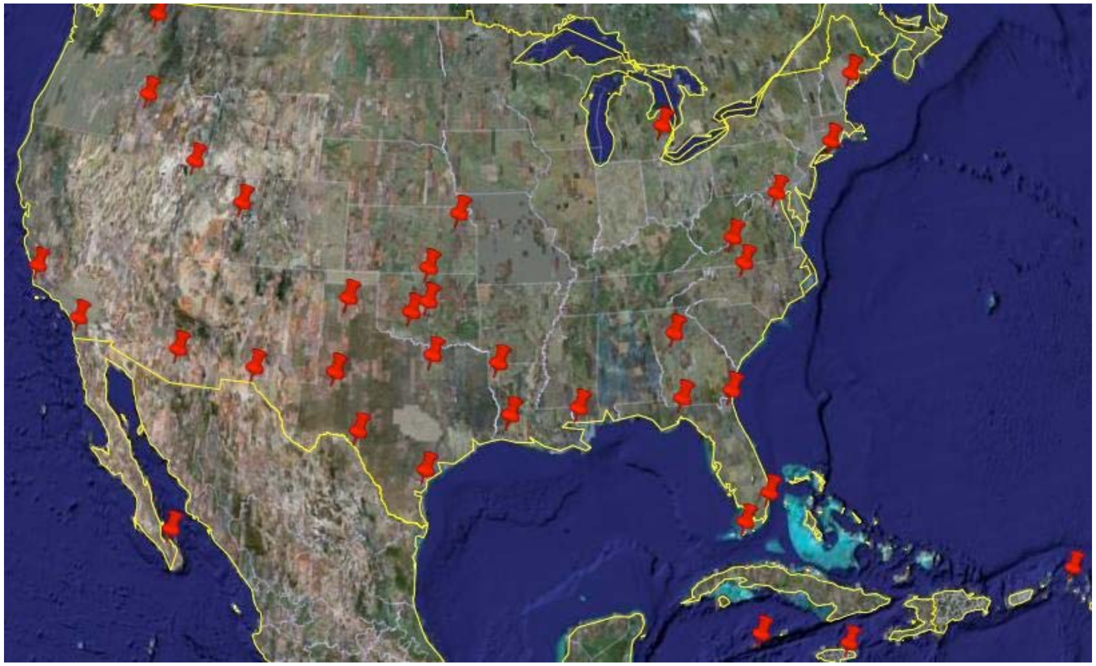


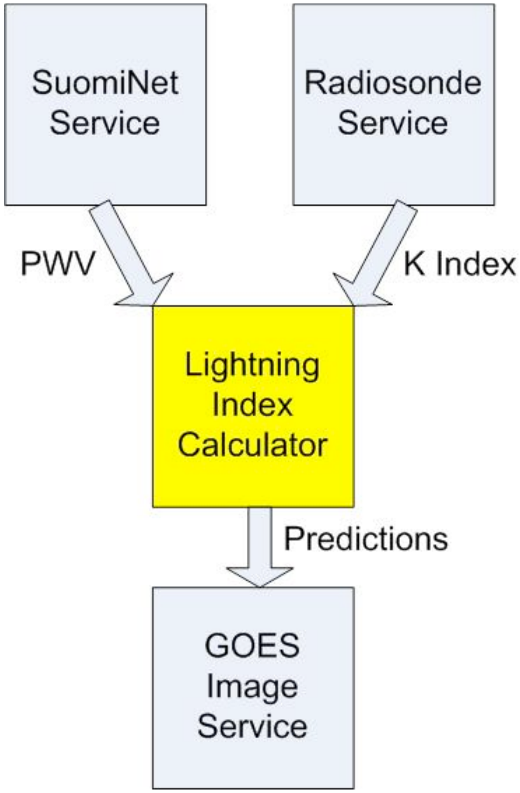
VSICS middleware enables flexible composition of disparate services that were not designed for such an integrated virtual sensorweb service

GOES IR image from same event

Data Fusion in Use Case

- Lightning prediction index allows determination of what satellite images are useful
- VSICS common interfaces and middleware make integration of data sources straightforward
- Lightning prediction service requests data from PWV and K index sources
- Results sent to GOES service which logs data and automatically downloads images where/when lightning predicted
- Done at 35 locations where the two data sources have stations within 20 miles

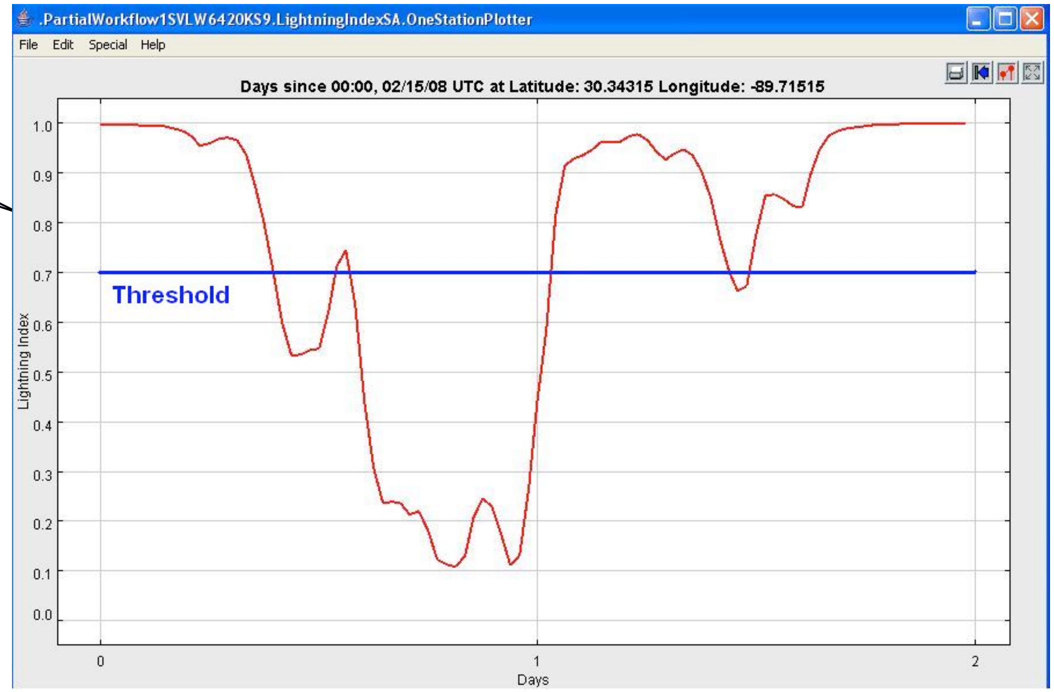




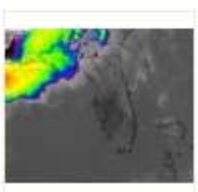
VSICS allows model execution in multiple locations with data sources not previously used together

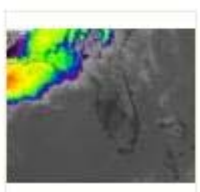
Workflow Execution in VSICS: Sample Results

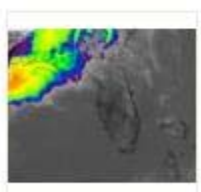
Graph of index for a single station over two-day period

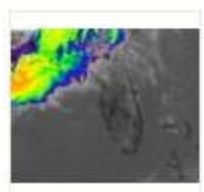


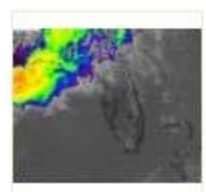
Time-stamped series of GOES images centered on lightning prediction location; workflow produces these image streams for each one


2008_0503_1710.jpg


2008_0503_1726.jpg


2008_0503_1741.jpg

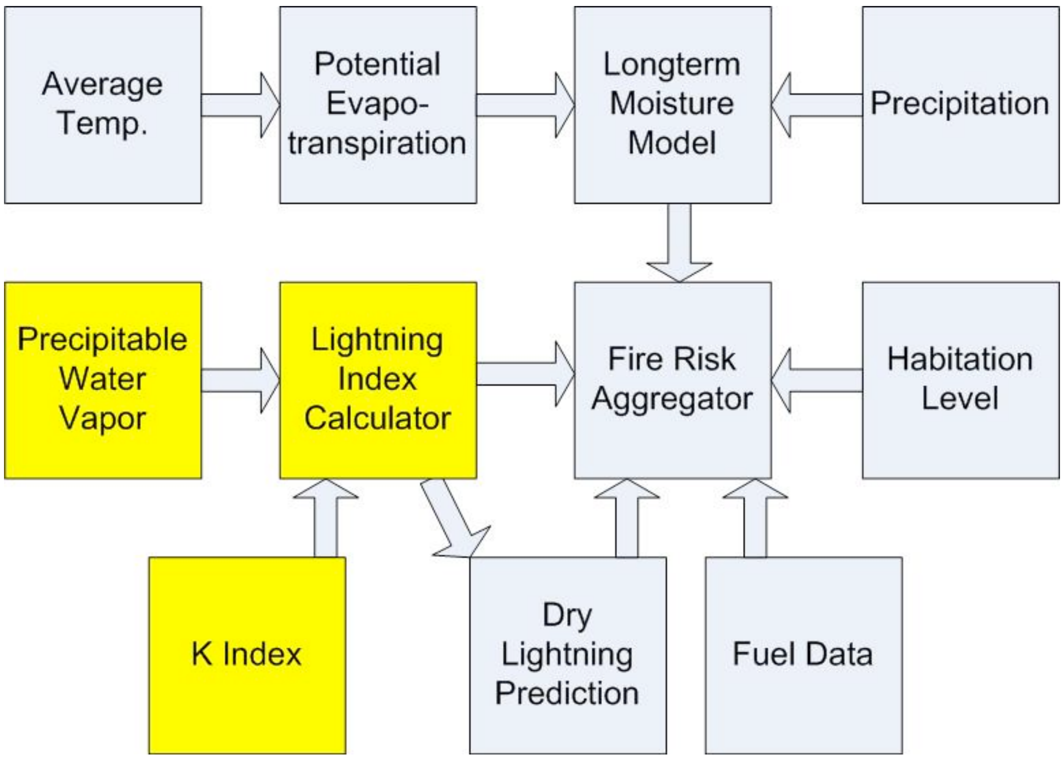

2008_0503_1802.jpg


2008_0503_1827.jpg

Results continuously updated at all locations automatically

Use Case 2: Wildfire Risk Assessment Model

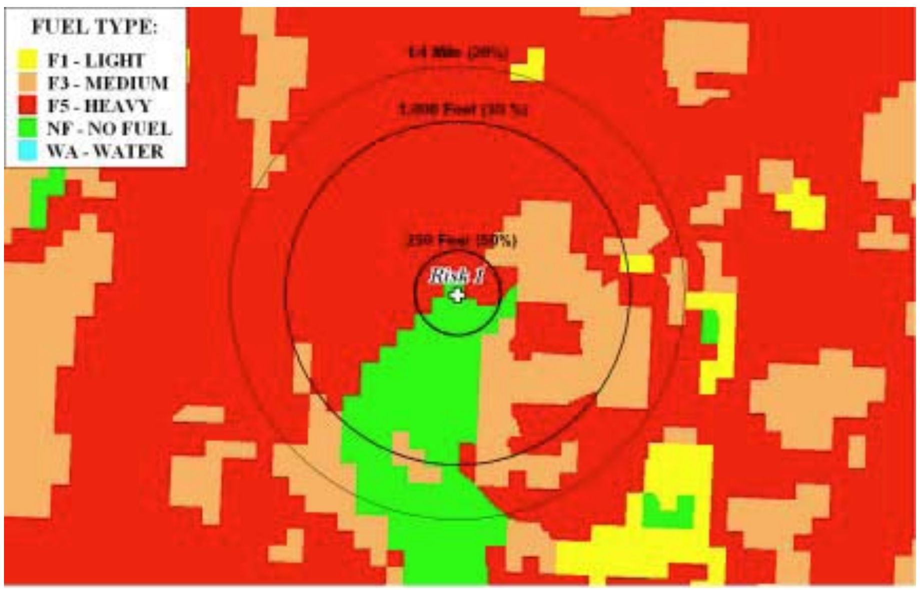
- Combines lightning prediction services (highlighted) with others to aggregate risk factors for wildfires and their impact

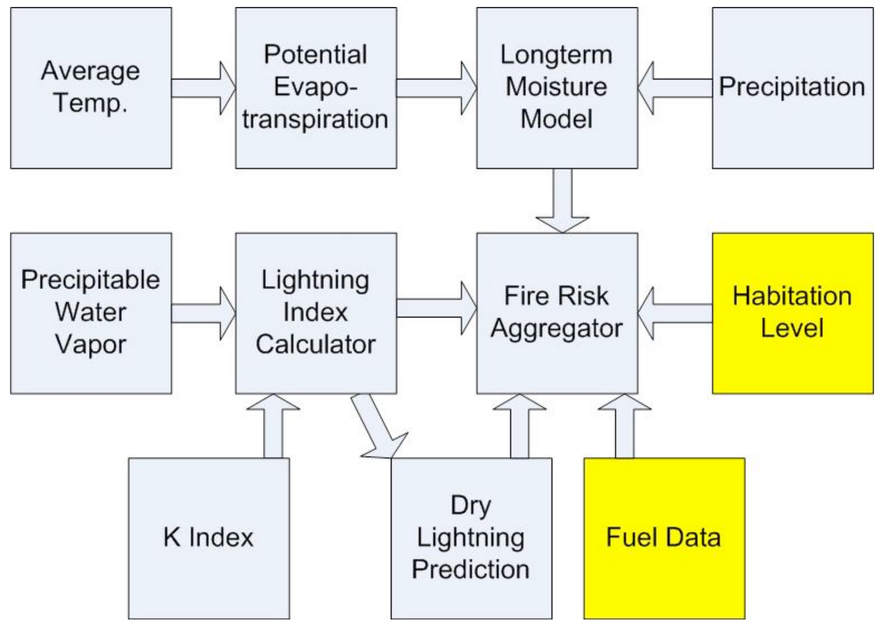


Wildfire risk prediction model involves complex multiscale fusion of information: multiple domains, temporal and spatial extents

Data Services: Habitation and Fuel

- Estimate hazard to lives/property from potential fires using 2000 US Census
- Eventually will compile average fuel levels for prediction regions in 9 western states using ISO FireLine
- Shows flexibility of VSICS by using data from commercial source as well as scientific

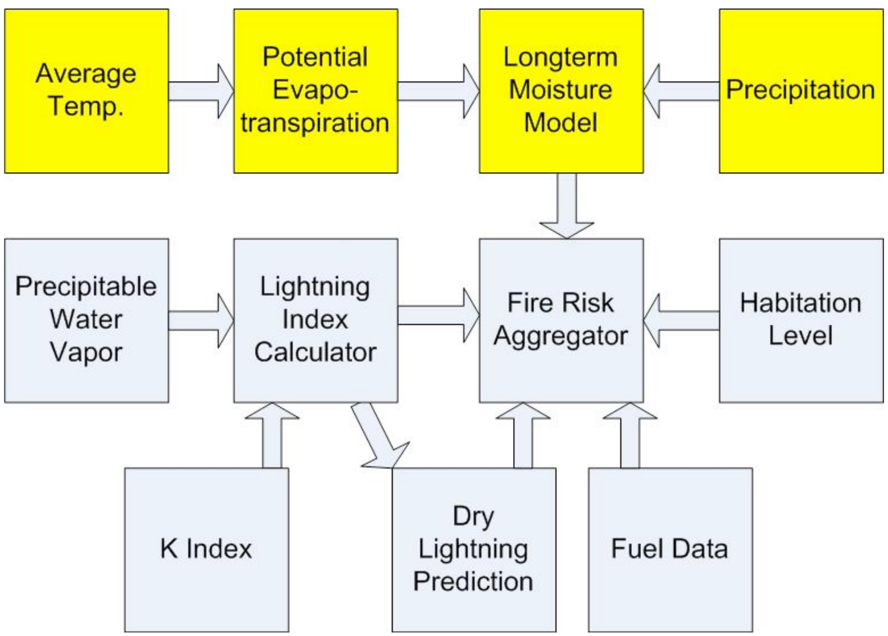




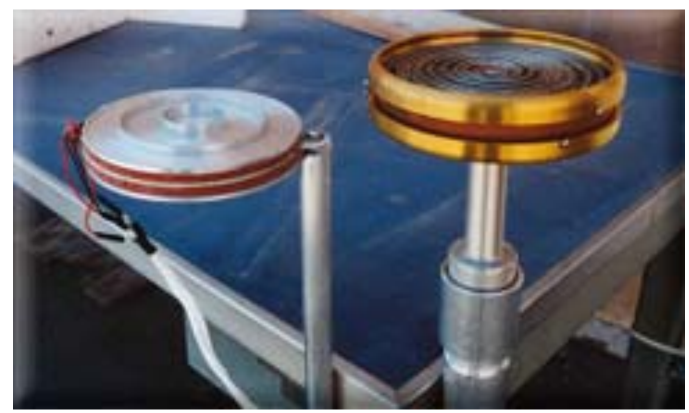
Example of FireLine fuel map, from
http://www.iso.com/dloads/CA_Fire.pdf

Integration of Precipitation and Temperature into Moisture Model

- Use C.W. Thornthwaite's model of potential evapotranspiration (PE) calculated from latitude and temperature
- Temperature obtained from SuomiNet and GPS-MET (another GPS PWV source, run by NOAA since 2001)
- Precipitation data from NOAA's Climate Prediction Center
- Add precipitation and subtract PE over multiple months to get accumulated moisture surplus or deficit



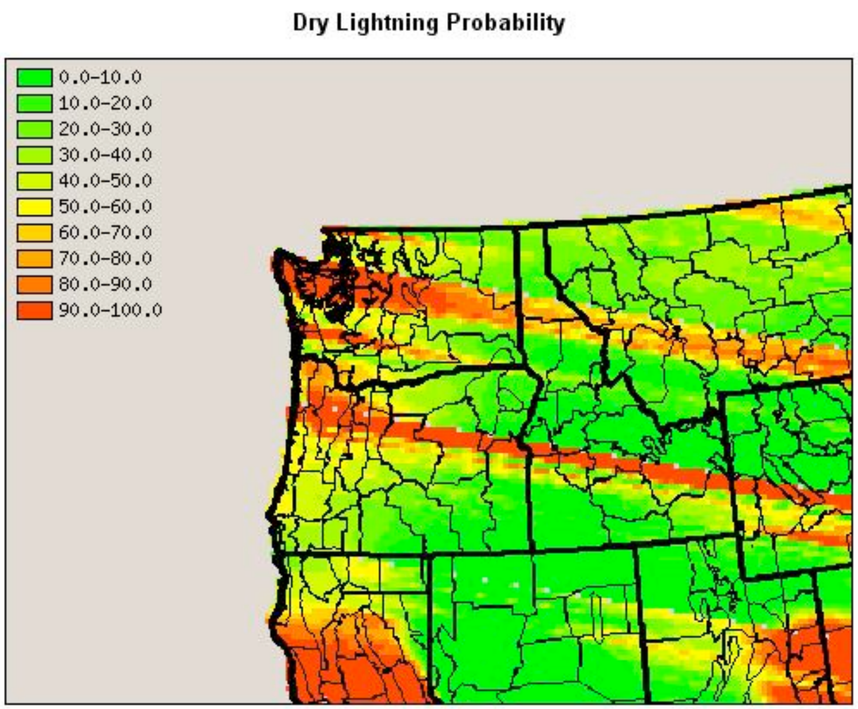
Precipitation Sensor



Model fragments involving spatio-temporal integration of information sources require generalized workflow in VSICS

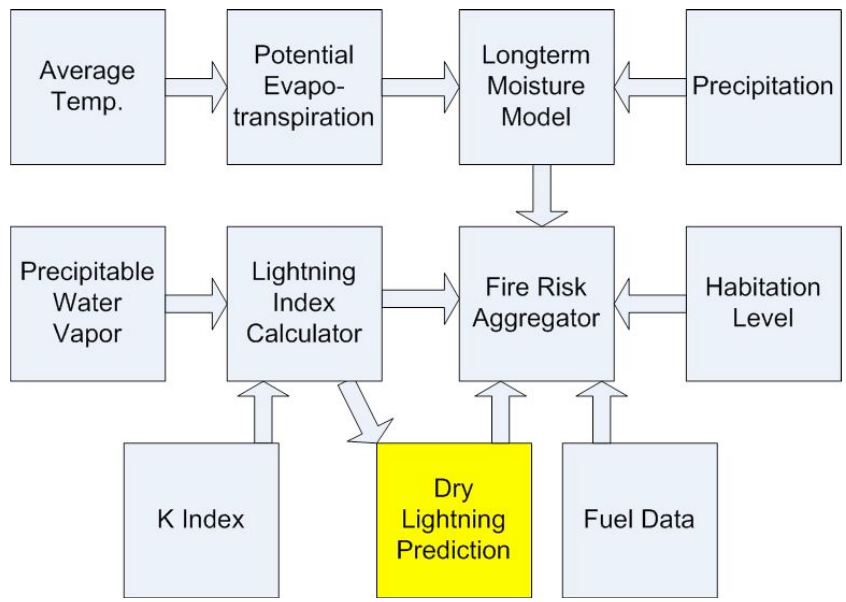
Integration of External Prediction Service: Dry Lightning

- Forest Service AirFire team provides probability of dry lightning (less than 2.5mm precipitation that day) which is more likely to start fires



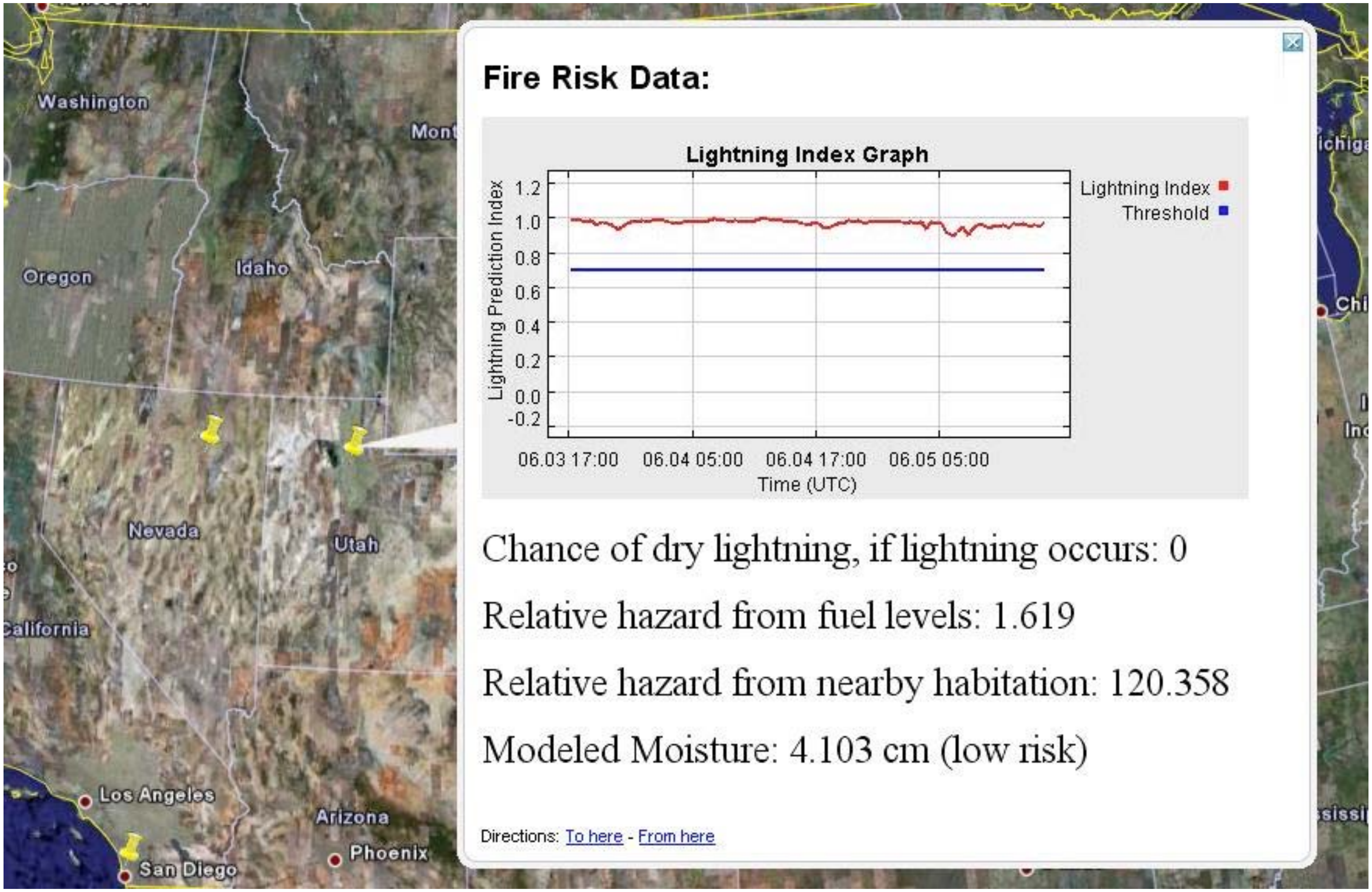
http://www.fs.fed.us/pnw/airfire/sf/

We use our approximation until their service is restored.



Data Fusion: Wildfire Risk Aggregation and Display

- Combines all available data services at 52 locations in the US.



Outline

- Motivation and challenges
- Technical approach
  - Objectives, CONOPS and architecture
  - Software implementation
- Use cases
  - Lightning prediction
  - Fire risk warning
- **Summary and future work**

Evaluation of Ease of Use

User Activity	Effort Required	Assumptions
Configuration of service for integration into VSICS	1-15 hours, depending on complexity/format conversion	Java compatible, knowledge of native data format of source
Registering service in catalog, specification of pub/sub attributes	1 hour	Semantics tags, basic experience with Ptolemy
Task-specific workflow creation and configuration	Up to 20 minutes	Existence of required services and understanding of workflow semantics

Summary

- VSICS is a scalable distributed sensorweb infrastructure that
  - Enables flexible multiscale composition of services across domains, space and time for workflow-driven coordination that realizes task objectives
  - Facilitates autonomous and robust management of sensorweb resources for timely decision making in response to natural hazards such as wildfire and storms
  - Exploits semantic descriptions of services and messages to include functional attributes for flexible and context-aware management
  - Is implemented using scalable JXTA mechanisms and Ptolemy framework for publish/subscribe driven distributed execution of workflows

- Core VSICS functionality has been displayed via demonstrations
  - Lightning prediction use case demonstration
    - Combines three services not previously integrated to implement existing prediction method at many locations
    - Runs continuously, accepting and acting on data as it becomes available
  - Fire prediction use case demonstration
    - Extends lightning prediction use case
    - Adds information on long and short term moisture, fuel, and habitation

Future Work

- Usability by analysts and science PIs
  - Improvements to service authoring
  - Development of client software, better user interface
- Flexible and dynamic service composition
  - Bayesian models for decision making with uncertainity
  - Long duration workflows for validation and refinement of models, climate monitoring
- Coordination for optimality and robustness
  - Implementation of QoS attributes and metrics
  - Reconfiguration to handle sensor and middleware failures
- Efficiency of execution
  - Bundling of information by common times and locations to reduce transport overhead
- Wildfire applications and beyond to Decadal missions
  - Integration of more data services with wildfire use case
  - Development of workflows for other priorities in Decadal Survey