CyberConnector: Bridging the Earth Observations and Earth Science Models

Liping Di[1], Ben Domenico[2], Xiaoqing Wu[3], Haosheng Huang[4], Quansong Tong[1], Eugene Genong Yu[1], Jia Song[1]

1. Center for Spatial Information Science and Systems, George Mason University, 4400 University Drive MSN 6E1, Fairfax, VA 22030
2. Unidata Program Center, UCAR Community Programs, 1850 Table Mesa Drive, Boulder, CO 80305
3. Department of Geological & Atmospheric Sciences, Iowa State University, 3010 Agronomy Hall, Ames, IA 50011-1010
4. DEPARTMENT OF OCEANOGRAPHY & COASTAL SCIENCES, School of the Coast & Environment, 1002-Y Energy, Coast & Environment Bldg., Louisiana State University, Baton Rouge, LA 70803

ABSTRACT

CyberConnector is to be developed and implemented as a system of services for facilitating the automatic preparation and feeding of both historic and near-real time Earth Observation customized data and on-demand derived products into Earth science models. The service system is designed to extensively adopt open geospatial standards/specifications, including the ISO geospatial data and metadata standards and standard-based geospatial web service, workflows, and sensor web technologies are the foundation. It bridges the sensors and earth science models through standard interfaces, such as Web Processing Service, Sensor Planning Service, and Catalogue Service for the Web. The standard interfaces allow the automatic handshaking between components with workflow designers and underlying workflow execution language. The system is to be demonstrated with actual earth science models: The Cloud-Resolving Model (CRM), the Community Multi-scale Air Quality Model (CMAQ), and the Finite Volume Coastal Ocean Model (FVCOM). The final platform are to be deployed and maintained for operational ingestion, discovery, access, and present geospatial models, data, and information.

OBJECTIVES

. Earth observation (EO) through sensors is the most important way to collect Earth science data. It is very common that scientists who conduct data-intensive Earth science research need to integrate and analyze multi-source EO datasets, which are typically very diverse in the collecting sensors, data formats, projections, spatial/temporal resolutions and coverage, access methods, quality, documentation, and user support. Studies showed that scientists who used satellite remote sensing data for data-intensive geospatial science applications had to spend over 60% of their time just for data acquisition, preprocessing, and post-processing. The major barriers are 1) difficulty to find and obtain the needed data from geographically distributed data sources, 2) data not in ready-to-analyze form, e.g., incompatible format, projection, and resolution among data from different sources and between the data from external sources and the in-house analysis system used by scientists, 3) the unavailability of the needed data products, e.g., further process the low-level data into higher level customized products are often needed before an application or a decision support system can use them, and 4) lack of or inadequate computing resources (both software and hardware) to handle the large volumes of data.

The overall goal is to remove or significantly lower these major barriers and to demonstrate the dramatic reduction (by at least one order of magnitude) of the time and efforts spent by scientists and engineers on preparing data for data-model inter-comparison, model V&V, and model utilization through developing and operating a sensor Web enabled EarthCube building block, the CyberConnector, to bridge EO and ESMs.

CyberConnector aims to achieve objectives: (1) Efficiently and robustly preparing and feeding EO data to ESMs, (2) Efficiency, scalability, and generality, (3) Enabling the knowledge accumulation and sharing, (4) Automatically presenting ESM outputs/results through the Web, (5) Facilitating Business Event Processing (BEP), and (6) Engaging in broad community participations in developing and using CyberConnector and sustaining its operation.

ARCHITECTURE

Adoption of consensus-based, open specifications

ISO/TC 211 standards - geographic information, methods, tools, and services

OGC specifications - geospatial Web services : Discovery of geospatial data and services - Catalogue Service for Web (CSW) provides; Data serving and accessing - Web Coverage Service (WCS) and Web Feature Service (WFS); Geospatial processing algorithms - Geospatial Web Processing Service (WPS); Discovering, acquiring, planning, processing, and accessing Sensor observations - OGC Sensor Web Enablement (SWE)

Service-oriented architecture

Service-oriented architecture (SOA)

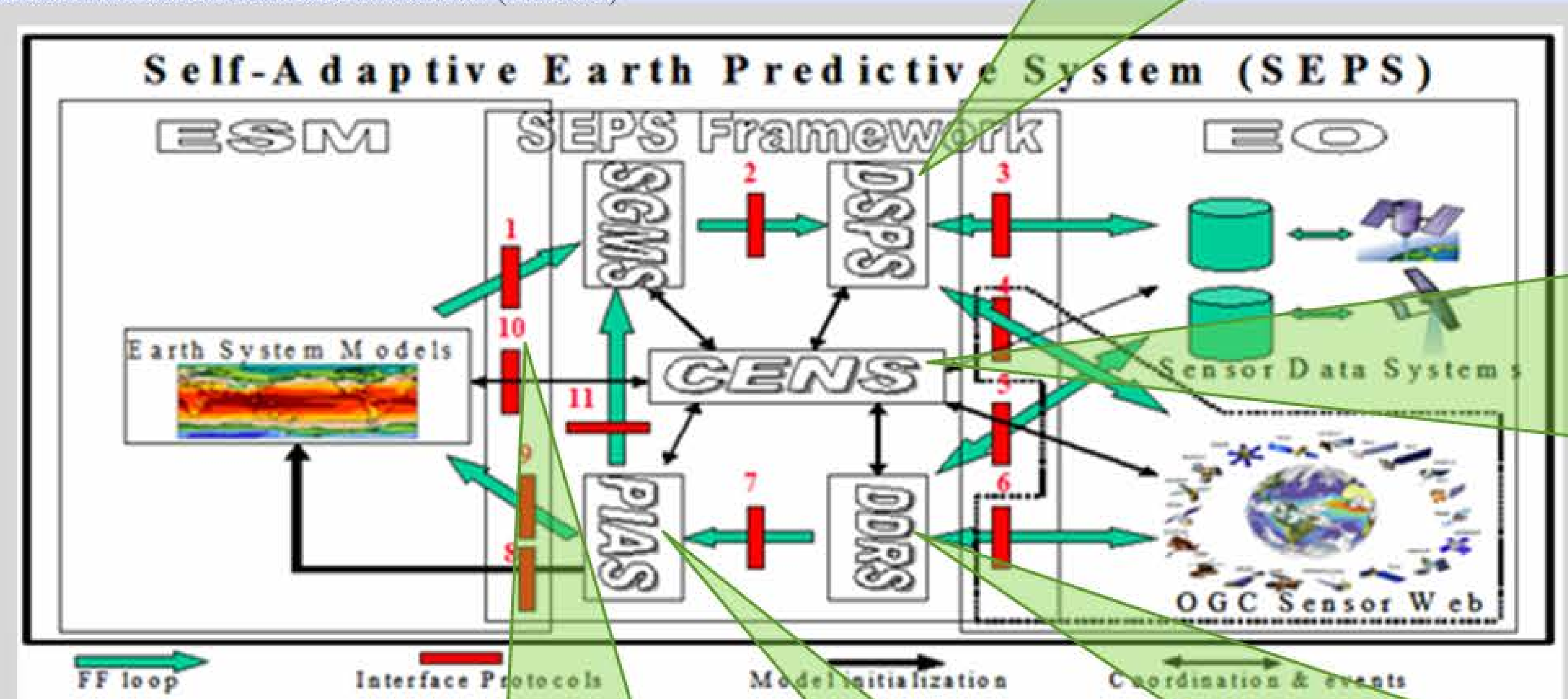
Representational State Transfer (REST) and Simple Object Access Protocol (SOAP) Web services

Geographic Resources Analysis Support System (GRASS) Web services

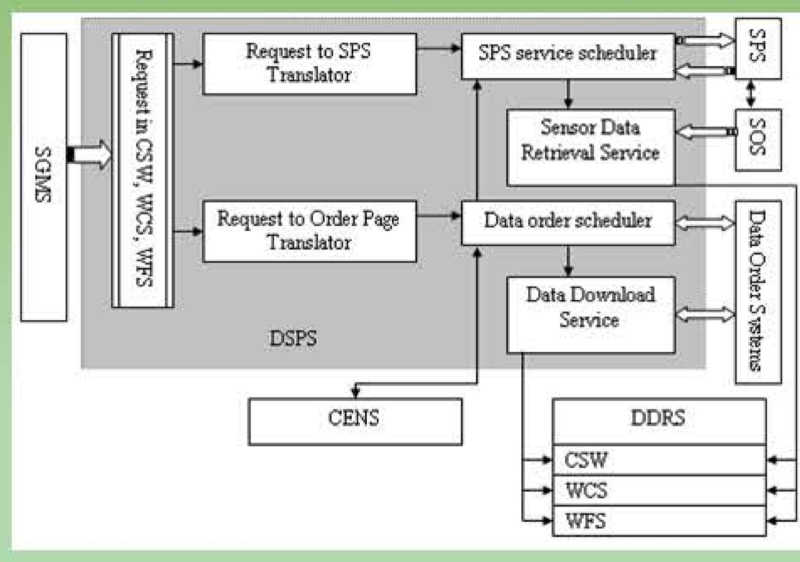
Service-composition and coordination

Business Process Execution Language (BPEL)

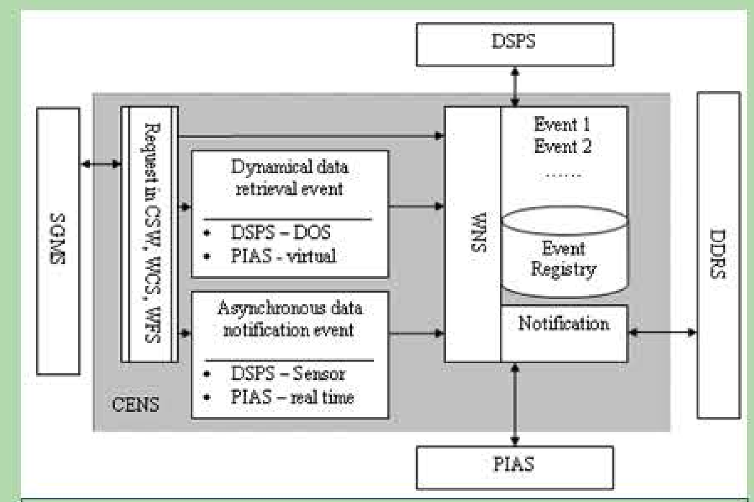
Business Process Model and Notation (BPMN)



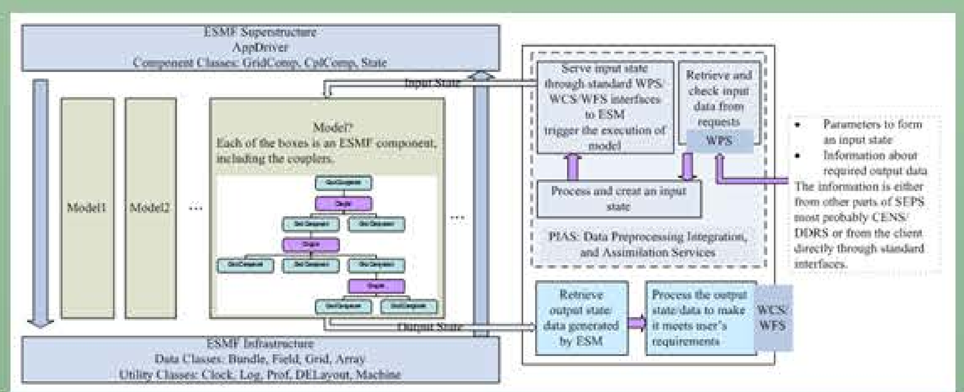
1. FTP, HTTP, WCS-T, WFS-T	5. OpenDAP, WCS, WFS	9. WCS, WFS
2. SPS	6. SOS	10. WNS
3. THREDS, ECHO, CSW	7. CSW, WFS, WCS	11. WCS, WFS
4. SPS	8. ESM state	



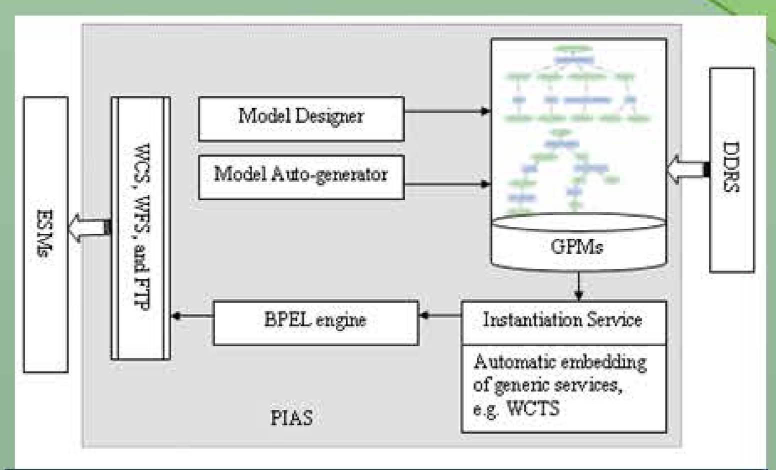
- Parse the request and translate to subscriptions for SPS
- Parse the request and translate to orders for DOS (Data ordering system)
- Schedule and submit the request/order/subscription and notify CENS the creation of events
- Parse the response and retrieve the data from SOS (Sensor Observation Service), feed these data into DDRS, and notify the result to CENS
- Parse the response and retrieve data from DOS, feed these data into DDRS, and notify the result to CENS



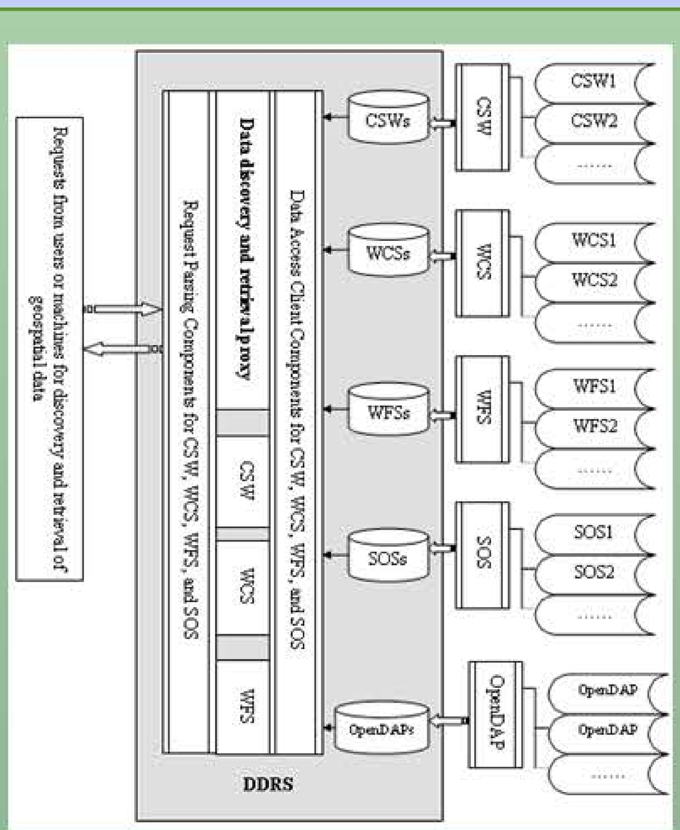
- Manages the messages past between PIAS, DDRS, and DSPS
- Coordinate the discovery, preparing, and downloading of data
- Notify and alert the the status of events to correpojnding services



- Interoperate with ESMF
- Exchange states with ESMF
- Import states: input parameters, configuration, and data
- Export states: output data and their running configuration



- Interactive model design
- Automatic service chain through ontology and artificial intelligence
- GPM cataloguing
- Instantiation abstract GPMs and generation of concrete BPEL workflow
- BPEL execution service



Discovery
- Federated CSW
- Interoperate with THREDDS

Access
- WCS
- WFS
- Interoperate with OpenDAP

CONCLUSIONS

The standard-based architecture design demonstrates advantages: (1) Interoperability: Standard geospatial services as the foundation of components support interoperability. Plug-in-and-play are made possible as long as the new components follows the supported open specifications or standards. (2) Reusability: Common interfaces assure reusability. Both existing Web services and newly developed Web services can be re-used as long as they are under the same geospatial Web service common specification. (3) Scalability: Web-based service-oriented architecture enables scalability. A Web service can be re-deployed onto many servers. Many servers can be used simultaneously if partition of tasks is properly done. For example, large processing tasks can be partitioned into small tasks in terms of spatial regions and time periods. They can be processed by multiple services and then combined by an aggregator service. This would enable the processing of large datasets in a timely fashion if we can find enough computing resources.

The further development will complete the implementation and integration of the framework. Three demonstrations, Cloud-Resolving Models (CRM), Community Multi-scale Air Quality Model (CMAQ), and Finite-Volume Coastal Ocean Model (FVCOM), will be developed as examplar cases to show how models can leverage the functionalities of the framework and facilitate the management of geospatial Web servie execution and processes.

REFERENCES

Chen, N., L. Di, G. Yu, and J. Gong, 2010. Geo-processing workflow driven wildfire hot pixel detection under sensor web environment, *Computers & Geosciences*, Volume 36, Issue 3, March 2010, Pages 362-372. doi:10.1016/j.cageo.2009.06.013

Chen, N., L. Di, G. Yu, and M. Min, 2009. A flexible geospatial sensor observation service for diverse sensor data based on Web service. *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 64, Issue 2, March 2009, Pages 234-242. doi:10.1016/j.isprsjprs.2008.12.001

Chen, N., L. Di, G. Yu, J. Gong, and Y. Wei, 2009. Use of ebRIM based CSW with Sensor Observation Services for registry and discovery of remote-sensing observations. *Computer & Geosciences*. Volume 35, Issue 2, Pages 360-372. doi:10.1016/j.cageo.2008.08.003.

Di, L., K. Moe and G. Yu., 2009. Metadata Requirements Analysis for the Emerging Sensor Web, *International Journal of Digital Earth*, Volume 2, Issue S1 2009, pages 3 - 17. doi:10.1080/17538940902866195.

Yu, G. E., Zhao, P., Di, L., Chen, A., Deng, M., & Bai, Y. (2012). BPELPower—A BPEL execution engine for geospatial web services. *Computers & Geosciences*, *47*, 87–101.

Acknowledgements

This project is funded by the National Science Foundation under the EarthCube initiative through grant ICER-1440294. The PI is Professor Liping Di (George Mason University). Co-Is are Dr. Ben Domenico (UCAR), Dr. Haosheng Huang (Louisiana State University), Dr. Daniel Tong (George Mason University), and Dr. Xiaqing Wu (Iowa State University).