

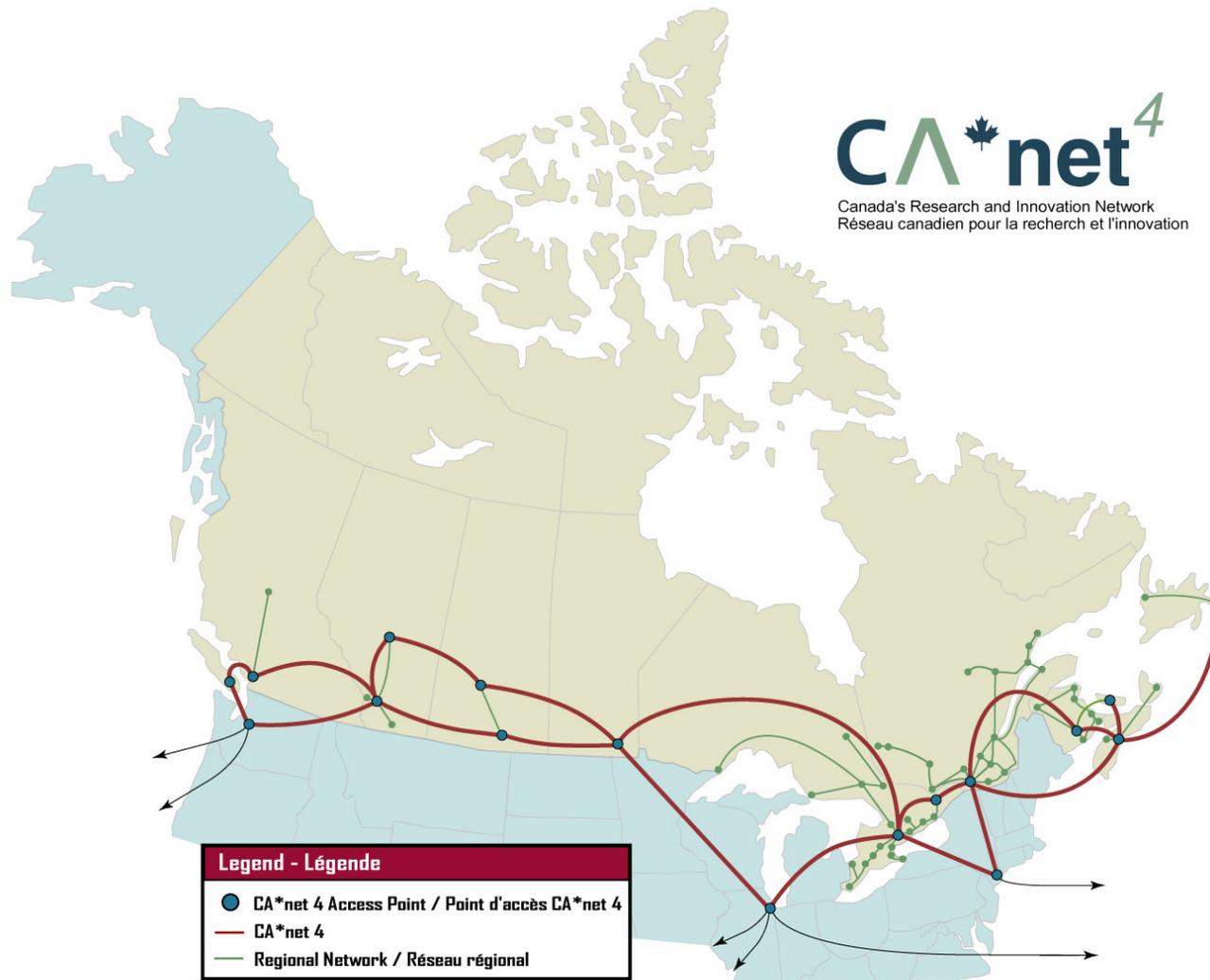


IBM Canada Ltd.

Grid Project in Canada

Dominic Lam Ph.D.
National eSystem Manager
dlam@ca.ibm.com

Grid Essential: High Speed Network in Canada



Winter / Hiver 2003

Potential Grid Projects in Canada

- **CERN Low level trigger data to U of Alberta**
 - Initially streaming data rates 1 Gbps moving to 10Gbps later in the year
- **Canadian virtual observatory**
 - .5 Tbyte per day to U of Toronto and U of Hawaii
 - 250 Mbps continuous streaming from CCD devices
- **Neptune – Canada (and US?) under sea laboratory – multiple HDTV cameras and sensors on sea floor**
- **Canada Light Source Synchrotron – remote streaming of data acquisition to U of Alberta and 20 other sites**
 - 2 to 5 Gbps continuously
- **Canadian remote Nano and micro electronics laboratories**
- **Westgrid projects**

Common Problem in Large Scale Projects

- **Many large sensor and instrumentation, major research equipment and cyber-infrastructure to be managed**
- **Most are still designed using “stove pipe” network and software database architectures**
 - Little thought is given to module reuse, common libraries and/or repositories
- **Potential cost over run due to complexity, software bugs etc**
- **Traditional locked in designs that limit future scalability, extensibility and reconfiguration**

Potential Solution

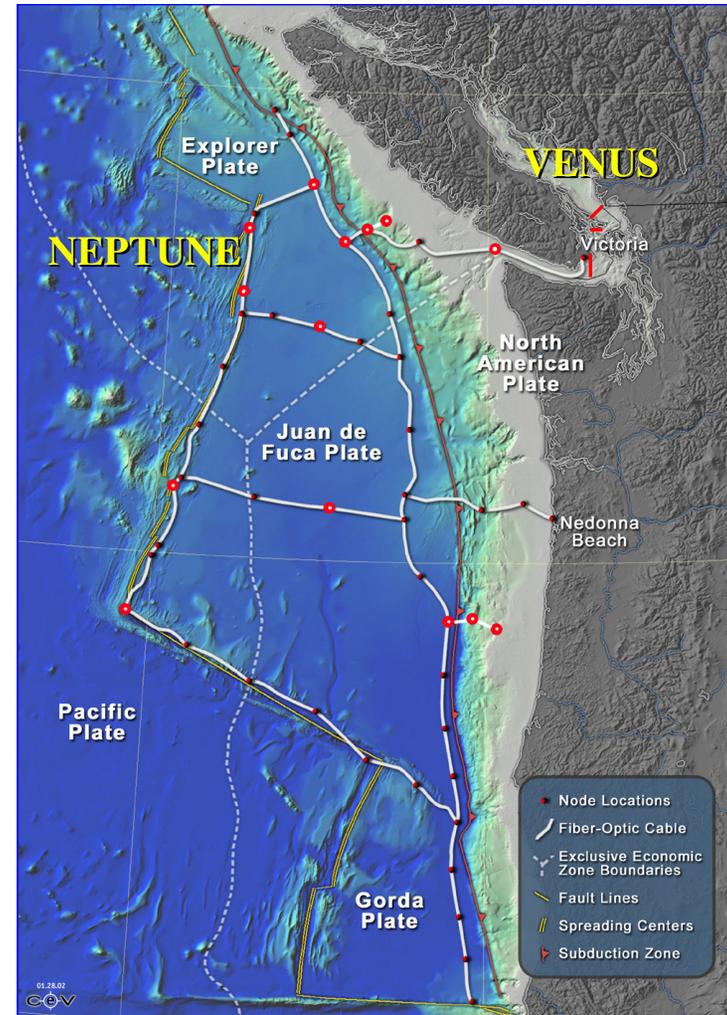
- **To develop a new architectural process for large systems software and network design methodology and implementation for major research projects such as Neptune, CLS, SNO, ATLAS and other cyber-infrastructure projects using latest web services and web service workflow architectures**
- **Create a repository of common web services for science, industry and education**
 - Data reformatting services, data syntactic services, wsfl abstraction services, data archiving services, etc
- **To add network lightpath services to repository of common web services**

NEPTUNE: The Cabled Ocean Laboratory

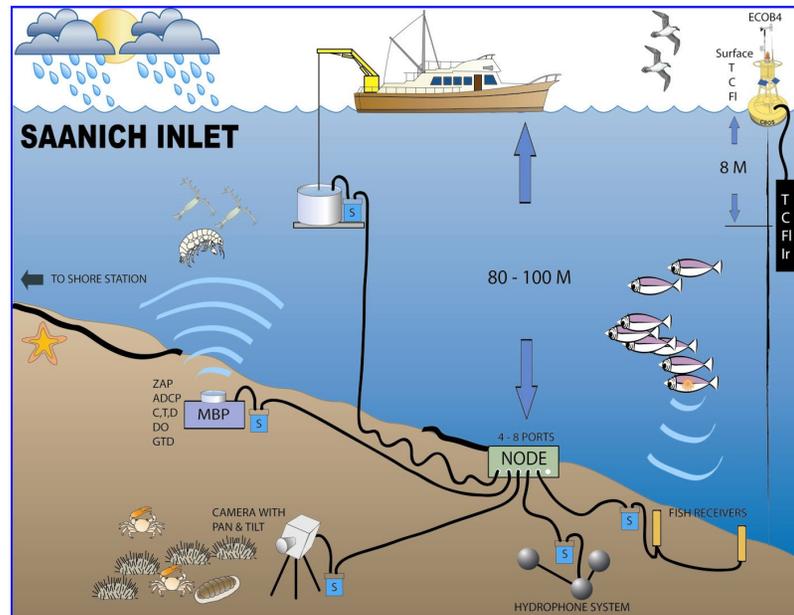
North-east Pacific Time-series Undersea Networked Experiments

- **Research areas include:**
 - Plate tectonic processes and earthquake dynamics
 - Dynamic processes of fluid fluxes and gas hydrates in the sea bed
 - Regional oceanic/climate dynamics and effects on the marine biota
 - Deep-sea ecosystem dynamics
 - Engineering and computational research

- **Delivering power and communications to instruments on and in the seafloor and through the water column**



The Enabler: VENUS & NEPTUNE System Design



- **Power & Communications Cables**
- **Seafloor nodes/arrays for Instruments**
- **Shore station**
- **Operation Centre**

- **Instruments**

- Seismometer
- Hydrophone
- Acoustic Transponders
- GPS
- HDTV & Still Cameras
- Crawler

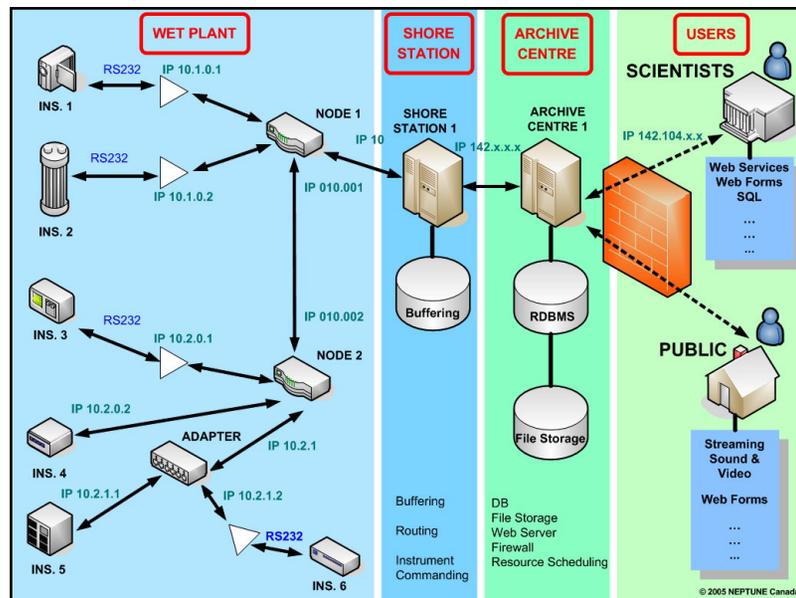
- **Data Management and Archive System (DMAS)**

- Data Capture and Retention
- Instrument Control and Programming
- Scheduling and Execution of crawlers and instruments
- Resource scheduling and assignment (crawlers, cameras)

Grid Web Service: NEPTUNE CIIP Project

To adopt (and further refine) a generic web-service enabled telemetry and control system for VENUS and NEPTUNE

- Embrace open standards to further innovation, collaboration and knowledge sharing (Open Source, Linux, OASIS)
- Migration from proprietary or legacy instrument control schemes (RS-232, RS-485, LECIS, SCADA)
- Abstract and service-enable system end-points (instruments, PC control stations, compute clusters, data management or analytical applications)
- Introduce a loosely-coupled integration fabric which places the burden on the “service bus” rather than the end-points (SOA, ESB, SOAP/XML)



- Exploit reliable asynchronous messaging or synchronous messaging as a means for instrument control or data interchange
- Research driven process and service choreography (BPEL, BPEL4WS)
- Adopt self describing data packets that can be inspected or transformed in transit (XML)
- Provide dynamic and ever-evolving experimental design or conditional handling support via content based routing for alerting, workflow, and event management

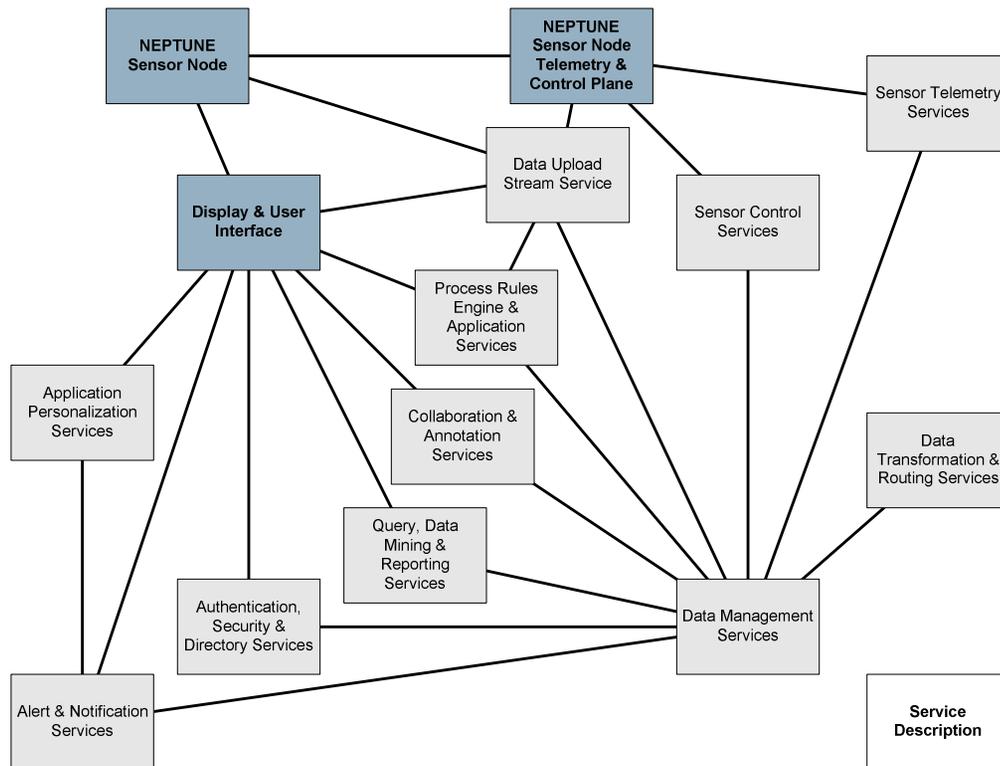
Many Industries Share NEPTUNE's Objectives

- **The “Extended Internet” will connect information systems to physical assets, products, and devices**
- **Complex business processes often need to connect to resources and systems beyond the world of IT and computer users. They link network endpoints like RFID, telematics, sensor networks, along with the necessary network protocols - bandwidth**
- **It also involves the sorting, sifting, and analysis of data gathered by networks of intelligent devices. Data analysis and business intelligence software, linked with device or sensor networks, helps organizations deal with the onslaught of data that such networks will create.**
 - Partners Healthcare equips discharged wound patients with digital cameras, allowing nurses to view after-care problems (like gangrene) via pictures downloaded over the Internet, avoiding costly home visits
 - Delta Air Lines plans to use RFID luggage tracking to handle the 0.7% of bags that are misdirected each year, hoping to cut some of the nearly \$100 million it spends to find those 800,000 lost bags.
 - Norwich Union uses vehicle telematics to adjust insurance premiums based on where customers actually drive their cars, rather than just on where they live
 - Temperature sensors in Union Pacific rail cars automatically alert repair crews when refrigeration goes on the blink, improving the quality of its produce delivery.

A Common Vision: An IBM Primer on SOA

- A *set of architectural principles* and patterns which address characteristics such as modularity, encapsulation, loose coupling, separation of concerns, composition, choreography, and single implementation
 - An *architectural style* in which a service provider offers a service description, a service requestor finds the description and binds to the implementation often through a service directory
 - As a *programming model* complete with standards, tools, methods, and technologies such as web services
 - A *linkage between business functions and aligned IT services* that collectively support changeable business functionality in an environment where changes in non-functional requirements often imply changes in service providers
- **Specifically an SOA is an enterprise-wide IT architecture that promotes:**
 - **Loose coupling** - implying the existence of a formal, well-defined interface contract (in terms of interface contract, not time)
 - **Location transparency** - a service user (consumer) should not hardcode any endpoint information.
 - **Protocol independence** - a service can be reached via several transport protocols without having to be recompile/recompile the service invocation logic.

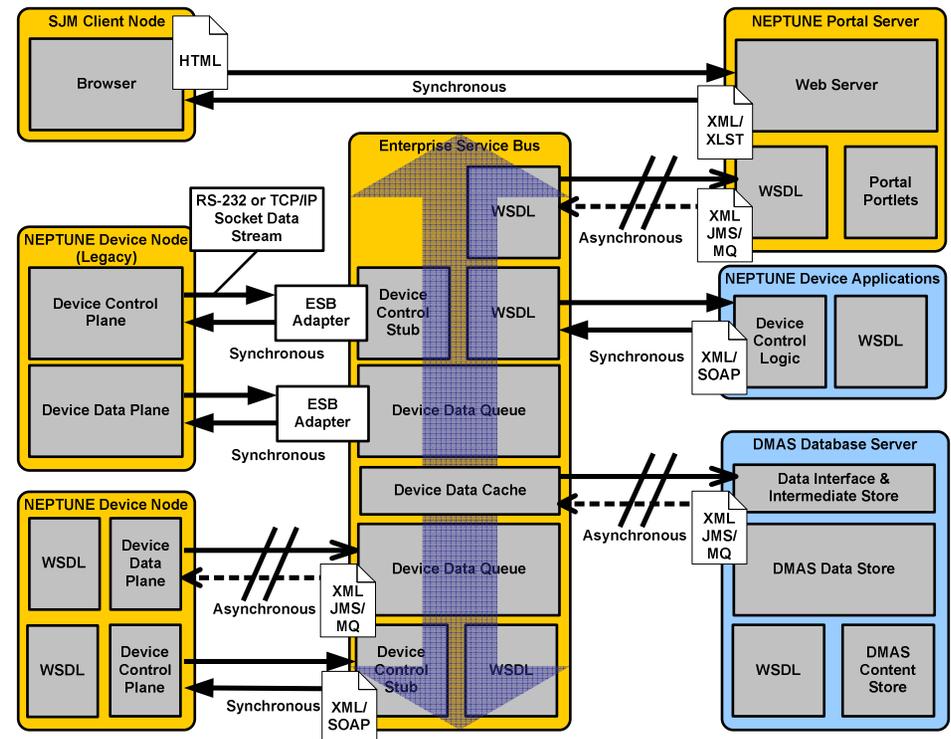
NEPTUNE: A Coarse-Grained Services View



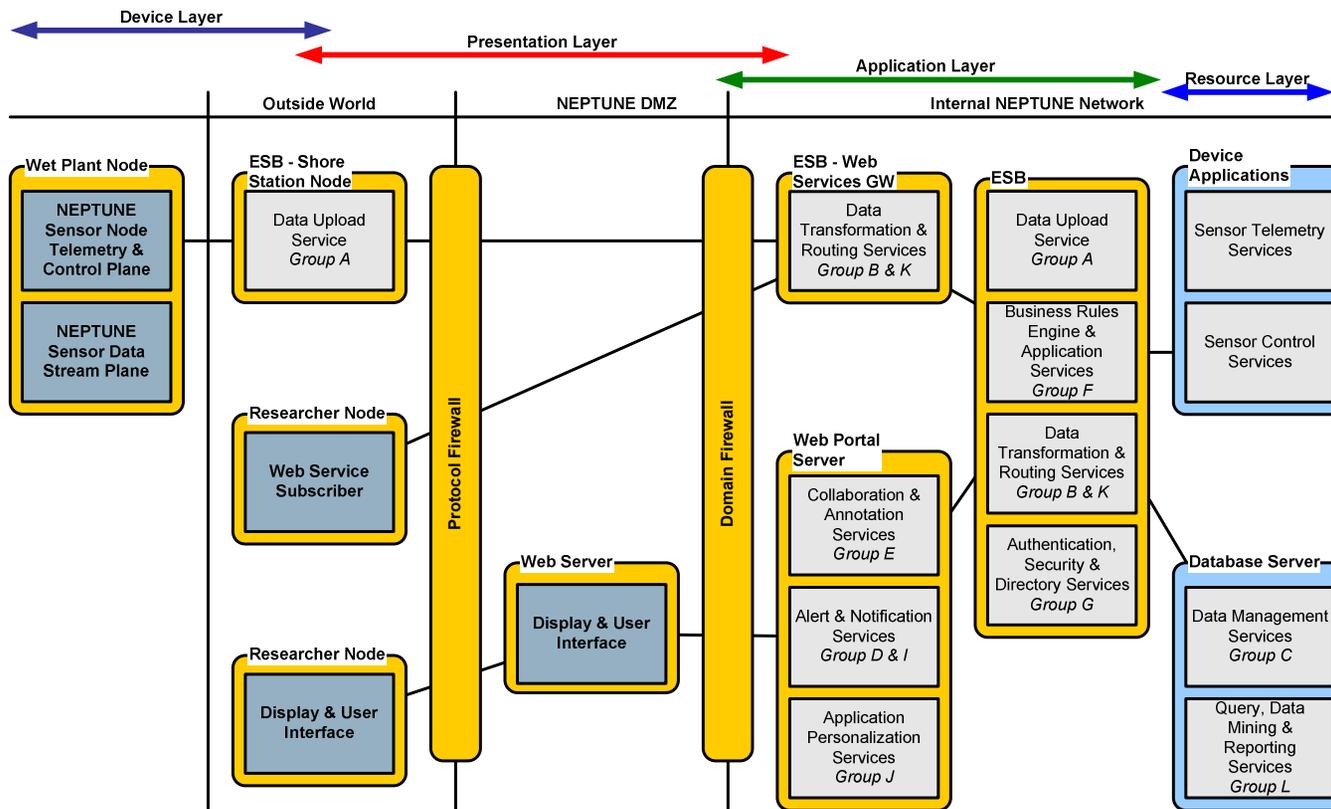
- **NEPTUNE will be based on a publish/subscribe broker:**
 - Managing the flow of information from remote devices to any Operation Centre or Researcher application
 - Command and control applications out to remote devices in the field
- **This model is built around a central broker and a number of clients. The broker acts as a go-between:**
 - An agent that matches subscribers (researchers) to information with publishers of information (sensors) that's relevant to them (hydrophone, seismic etc)
 - Clients can be publishers of, and/or subscribers to, data and can range from big academia-based servers to hand-held pervasive computing devices, such as Palm handhelds, or unattended remote telemetry devices.

Transformed into a Node Interaction View...

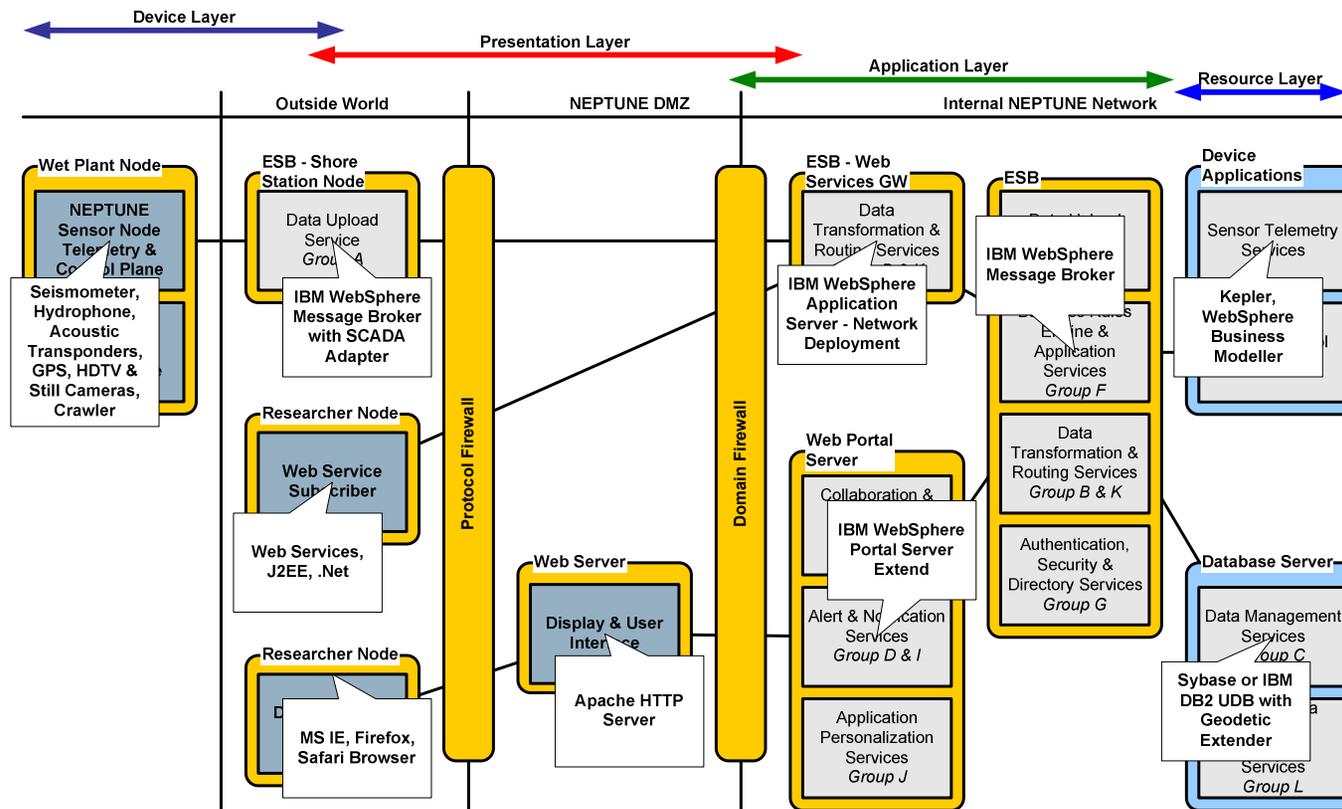
- **The system will be based on the IBM WebSphere family of products**
- **A possible message flow would be as follows:**
 - Telemetry data comes from a variety of remote monitoring sensors, transmitting via serial or TCP/IP interfaces over the fiber optic cables
 - The remote device sends data via an ESB Adapter (MQIsdp), reporting that a seismic event has occurred. The shore station, itself a distributed ESB node, transforms the message into an canonical XML format using a pre-defined namespace
 - The data is routed from the broker within the Operation Centre, where ESB rules act on the XML document to make it available via a Web Services interface
 - The data can then be “pushed” or “pulled” to the desired database or back-end system and may also trigger a Researcher’s portal web page refresh or RSS update



...and Finally a Run Time Deployment View



...and Finally a Run Time Deployment View



NEPTUNE Solution Software Components

WebSphere Message Broker

- **An advanced ESB which provides Web Services connectivity and non standard interface connectivity**
- **Any-to-any data transformation**
- **Customized communications**
- **Content-based routing**
- **Message warehousing**
- **Web Services mediation**
- **Complex event processing**
- **Higher-speed Publish/subscribe**
- **Store-and-forward**
- **Specialized transports and transport switching**
- **Message replay**
- **WebSphere MQ, JMS, SOAP**
- **Transactionality**

WebSphere Application Server

- **J2EE Application Server**
- **JMS engine embedded in the application server kernel**
- **Real-time and Multicast transports**
- **Web Services Gateway**

IBM DB2 UDB with Geodetic Extender

- **Manages geospatial information referenced by latitude-longitude coordinates**
- **Offers storage and retrieval capabilities for geographic data**
- **Supports global geographic queries without the limitations inherent in map projections**
- **Helps to ensure precision and accuracy—engineered (from inception) to treat the earth as a continuous globe rather than a flat plane bounded by edges**

NEPTUNE Solution Software Components, cont.

WebSphere Portal Server Extend

- **Connectivity and integration to allow access to event data, analytical applications or external news feeds**
- **Presentation and administration to enable computing-desktop customization to match researcher work patterns and needs**
- **Content management that provides an interface to store data and the structure to quickly access and associate that data with the experimental processes that leverage it**
- **Document management that enables portal users to share, search, organize and edit word-processing, spreadsheet and presentation files**
- **Research-process integration and orchestrated workflow capabilities that connect people and applications at a process level**

IBM WebSphere Business Integration Modeler Spatial Integration Adapter

- **Business Process Execution Language (BPEL) — optimized to export models to BPEL, Web Services Description Language (WSDL) and XML Schema Definition Language (XSD) file formats, and to import to IBM WebSphere Studio Application Developer Integration Edition to develop applications from the models**
- **MQ Workflow Flow Definition Language (FDL) — optimized to export models to FDL format, giving companies a fast start to IBM WebSphere MQ Workflow as an automated workflow solution**
- **Spatially enable your research applications**
- **Publish industry-standard geographic information system (GIS) Web services with ease**