Wiring the Juan de Fuca Plate for Science: The NEPTUNE System

BRUCE M. HOWE
Applied Physics Laboratory, University of Washington, Seattle, WA 98105, USA

HAROLD KIRKHAM
Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

ALAN D. CHAVE AND ANDREW R. MAFFEI
Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA

SÉVERIN GAUDET
Herzberg Institute of Astrophysics, National Research Council, Victoria, BC V9E 2E7, Canada

ABSTRACT

NEPTUNE will “wire” the Juan de Fuca Plate in the northeast Pacific Ocean. It will enable in-depth study of life in extreme environments, plate tectonics, oceanography, and fisheries, to name but a few of the many topics. This will be accomplished by means of a multi-node submarine fiber-optic/power cable network deployed around the plate. It will provide power for the instrumentation as well as make seafloor sensors and interactive experiments easily accessible via the Internet to researchers, educators, and the public around the globe.

The NEPTUNE Project is beginning its development phase. The many and varied issues associated with the infrastructure technical design are discussed, including power, communications, and data management and archiving.

1. Introduction

The ocean sciences are on the threshold of a revolution that will fundamentally change the means by which many processes in and beneath the oceans are observed. Historically, ship-based expeditions have characterized the oceans using data collected onboard and with short-term deployments of autonomous instruments. It has become increasingly apparent that many problems cannot be addressed by time-limited visits alone. Ocean research is moving towards a mode of operation in which ship-based studies will be complemented by long-term observations from a network of coastal, regional, and global seafloor observatories.

A recent Ocean Studies Board report concluded that “Seafloor observatories present a promising, and in some cases essential, new approach for advancing basic research in the oceans.” It recommended that “NSF should move forward with the planning and implementation of a seafloor observatory program” (NRC, 2000).

Although many oceanographers are accustomed to designing autonomous instruments with power consumption limited to less than a watt, there are good reasons for developing observatories that can provide considerably more power and real-time two-way communications. Power consumption is inherently high for applications that require propulsion, pumping, lighting, heating or cooling, and acoustics. Video and acoustics require high data rates. Two-way communications are necessary for interactive experiments. Further, accurate timing is necessary for navigation and tomographic imaging. For some new technologies, the availability of substantial power and communications capability on the seafloor will reduce the development costs and time required to transfer long-term sensors from the laboratory to the oceans. The Ocean Studies Board report notes that the “projected power requirements for cabled observatories are anticipated to be 2-20 kW per node” and “The communications requirement … depends primarily on the use of video transmission and is anticipated to be, at most, less than 1 Gb/s per node” (NRC, 2000).

The NEPTUNE[1] project will deploy a fiber-optic/power cable network around and across the Juan de Fuca Plate off the west coast of the North America to study a variety of geological, oceanographic, and ecological processes.

---

[1] http://www.neptune.washington.edu

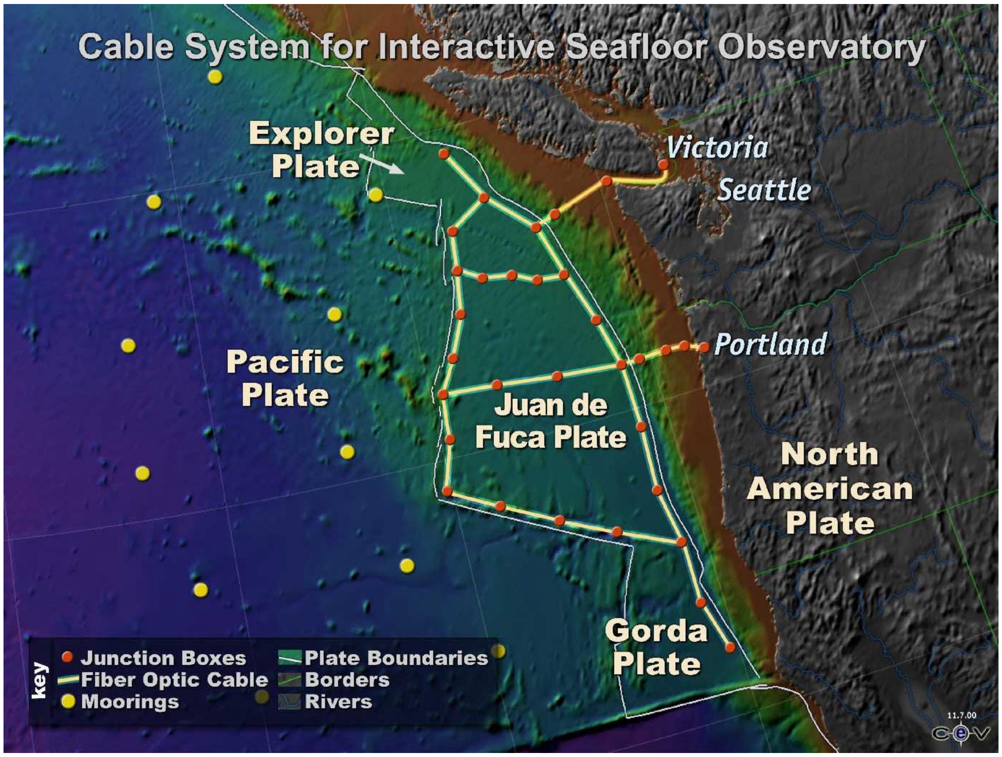

Figure 1. NEPTUNE will provide power and communications via fiber-optic/power cables to junction boxes (nodes) on the Juan de Fuca Plate and surrounding areas. By connecting sensor networks to the nodes, with extensions onshore and offshore into the interior, scientists will, for the first time, have the capability to study a host of interrelated processes with high spatial and temporal resolution over long periods of time. Large amounts of reliable power and high-bandwidth communications will be crucial elements of any such system.

NEPTUNE anticipates participating in the regional, tectonic-plate-scale component of the Ocean Observatories Initiative, a new NSF program to support long-term research observations and interactive experiments from the sea surface to the seafloor and beneath. These arrays will allow the coherent study of the behavior of the earth and ocean on time scales from seconds to decades and longer. (Delaney et al., 2000)

It is envisioned that the NEPTUNE backbone will comprise 3000 km of cable connecting about 30 evenly distributed nodes (Figure 1). Secondary branch cables will extend to any location on the plate. At each node, a junction box will provide standard power and Internet communication interfaces for scientific sensors, sensor networks, and various mobile platforms (Figure 2). The complete network will carry ~10 gigabits per second of data and deliver up to 100 kW of power with an operational life span of at least 30 years. NEPTUNE will support large community experiments, smaller experiments developed by individual investigators, and a range of innovative educational and outreach activities. The latter are especially important for educating and involving the public in the scientific research, whether in a major aquarium or a schoolhouse.

A conceptual design and feasibility study was completed in mid-2000 as part of NEPTUNE's Phase 1 effort (NEPTUNE Phase 1 Partners, 2000)[2]. The feasibility study demonstrates that there are no insurmountable technical obstacles and identifies a path for realizing NEPTUNE.

The feasibility study preparation included meetings of small *ad hoc* science working groups convened to explore the scientific opportunities provided by NEPTUNE and to provide input to the design of the system. The science working group topics were:

- cross-margin particulate fluxes,
- seismology and geodynamics,
- seafloor hydrogeology and biogeochemistry,
- ridge-crest processes,
- subduction zone processes (fluid venting and gas hydrates),
- deep-sea ecology,
- water column processes, and
- fisheries and marine mammals[3].

These topics indicate the breadth of science that will be addressed by NEPTUNE. The NEPTUNE infrastructure

---

[2] Available on the NEPTUNE web site, http://www.neptune.washington.edu

[3] White papers from each science working group are available on the NEPTUNE web site.

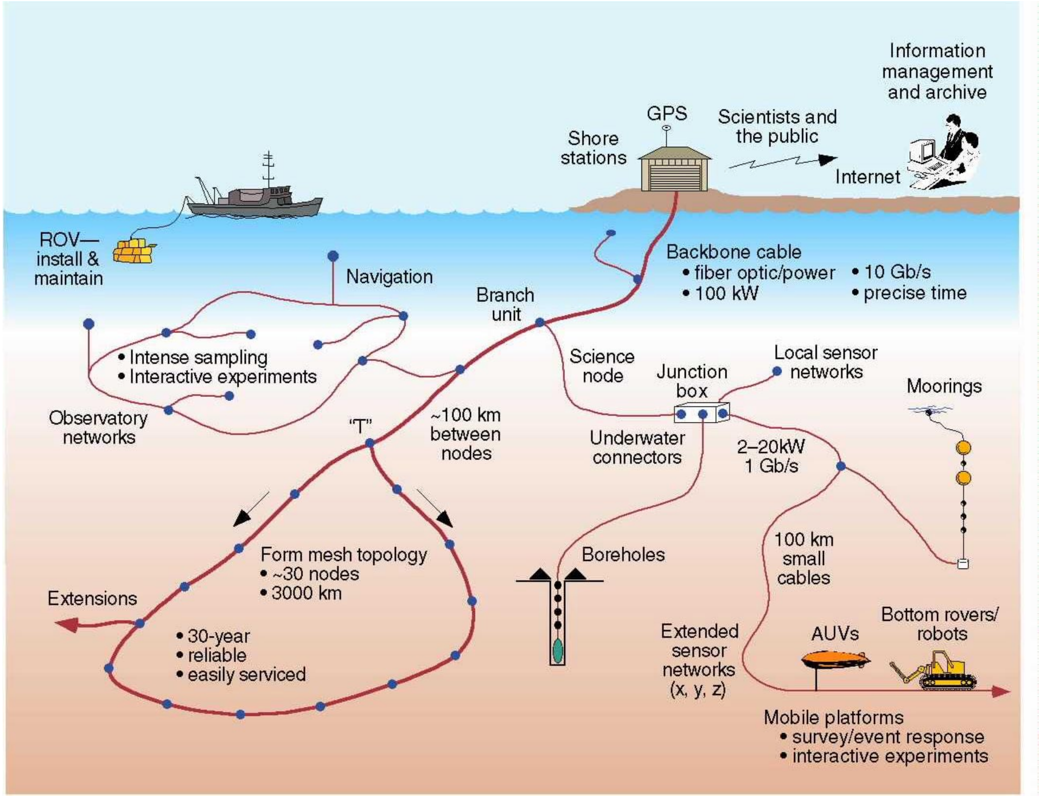

Figure 2. The essential elements of NEPTUNE. Land-based scientists and the public are linked in real time and interactively with sensors, sensor networks, and various mobile platforms, in, on, and above the seafloor. NEPTUNE's fiber-optic/power cable and associated technology provide the enabling network infrastructure.

will positively impact the way the international and national scientific communities conduct research.

NEPTUNE has now entered a period of detailed development (Phase 2) that will lead to installation (Phase 3) and operations (Phase 4) as early as 2005. The Phase 2 partners are the Institute for Pacific Ocean Science and Technology (IPOST/Canada[4]), the California Institute of Technology Jet Propulsion Laboratory (JPL), the Monterey Bay Research Aquarium Institute (MBARI), the University of Washington (UW), and the Woods Hole Oceanographic Institution (WHOI).

This paper describes our present view of the NEPTUNE infrastructure, a system that will provide substantial amounts of reliable electrical power and communications capability to support existing and future science needs. This technology development will play a crucial role in enabling the fundamental paradigm shift necessary for observatory science.

The balance of the paper is organized as follows. Section 2 gives an overview of the system followed by a description of the major components: shore stations, cable, and nodes (Sections 3-5). Then each of the subsystems we are addressing in this paper, namely power, communications, and data management and archiving, are discussed in more detail (Sections 6-8). Section 9 addresses systems issues, such as interactions between subsystems and trade studies that are necessary. Section 10 contains concluding remarks.

2. System Overview

The components and subsystems of NEPTUNE are shown in schematic form in Figures 3 and 4. The four subsystems power, communications, timing, and control, thread their way through the physical system, connecting the scientists and other users on shore to science instruments from the sea surface to the seafloor and beneath. (Here, we will group timing and control with communications.). We have chosen a mesh topology for both power and communications to meet the capacity requirements and to maximize reliability and flexibility.

A parallel power system as on land will be used. It uses a conventional submarine cable. Unlike typical undersea telecommunications systems, however, the current is not held constant. The power system is based on the use of flexible, regulated power supplies, automatically adjusting to changing load conditions. A multi-layered protection system will make the power system as fault-tolerant as possible.

---

[4] http://www.neptunecanada.com

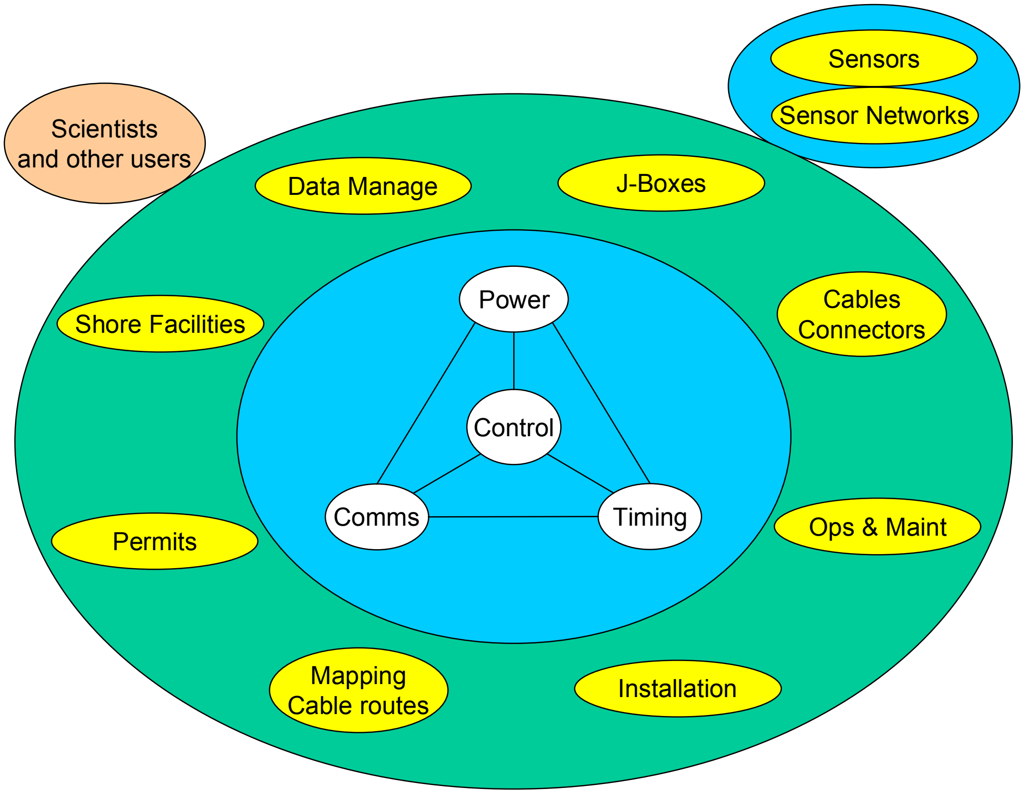

Figure 3. The major operational subsystems and the supporting components of NEPTUNE link the scientist/user to the sensors and data.

For communications, NEPTUNE will use data networking technology that is widely used to connect the Internet to laboratories, classrooms, and homes around the world. From the dual perspectives of flexibility that preserves simplicity and cost minimization, the Ethernet family is the most promising technology. NEPTUNE is similar to campus networks that use this approach and, with multiple shore connections, can benefit substantially from its automatic routing and restoration capabilities. Network security will also be an important aspect of NEPTUNE's final desgin.

A well-integrated data management and archiving system is important to the scientific and educational goals of the project, given the nature of the data produced by NEPTUNE:

- The wide variety and number of instruments,
- Heterogeneous data,
- The multi-disciplinary science goals,
- The projected terabytes-per-day data volume,
- Long life span of the project.

Data entering the data management and archive system must be properly time-stamped and packaged. At the other end, the system must allow the research and public education communities to easily access relevant archive data. In-between lies management of the data flow, catalog generation, data storage, data processing pipelines and numerous other components.

The major components of the physical system are the shore station, the cable, and the nodes. After discussing the functions of each of these components, we will address the subsystems.

3. Shore Stations

The functions of the shore stations are as follows:

- Provide stable, controlled power to the system cable,
- Provide two-way communication between the sensors and the scientists,
- Provide accurate timing signals for the system,
- Provide the data management and archiving service, and
- Control and manage the overall system.

As shown in Figure 1, it is assumed that there will be (at least) two shore stations, both to provide improved system survivability as well as increased power and communications capability. The shore station in Victoria is planned to be in the downtown area as part of a waterfront science center. The shore station in Nedonna Beach, Oregon, is somewhat remote. It is likely that there will be "network operations centers" in convenient locations, and

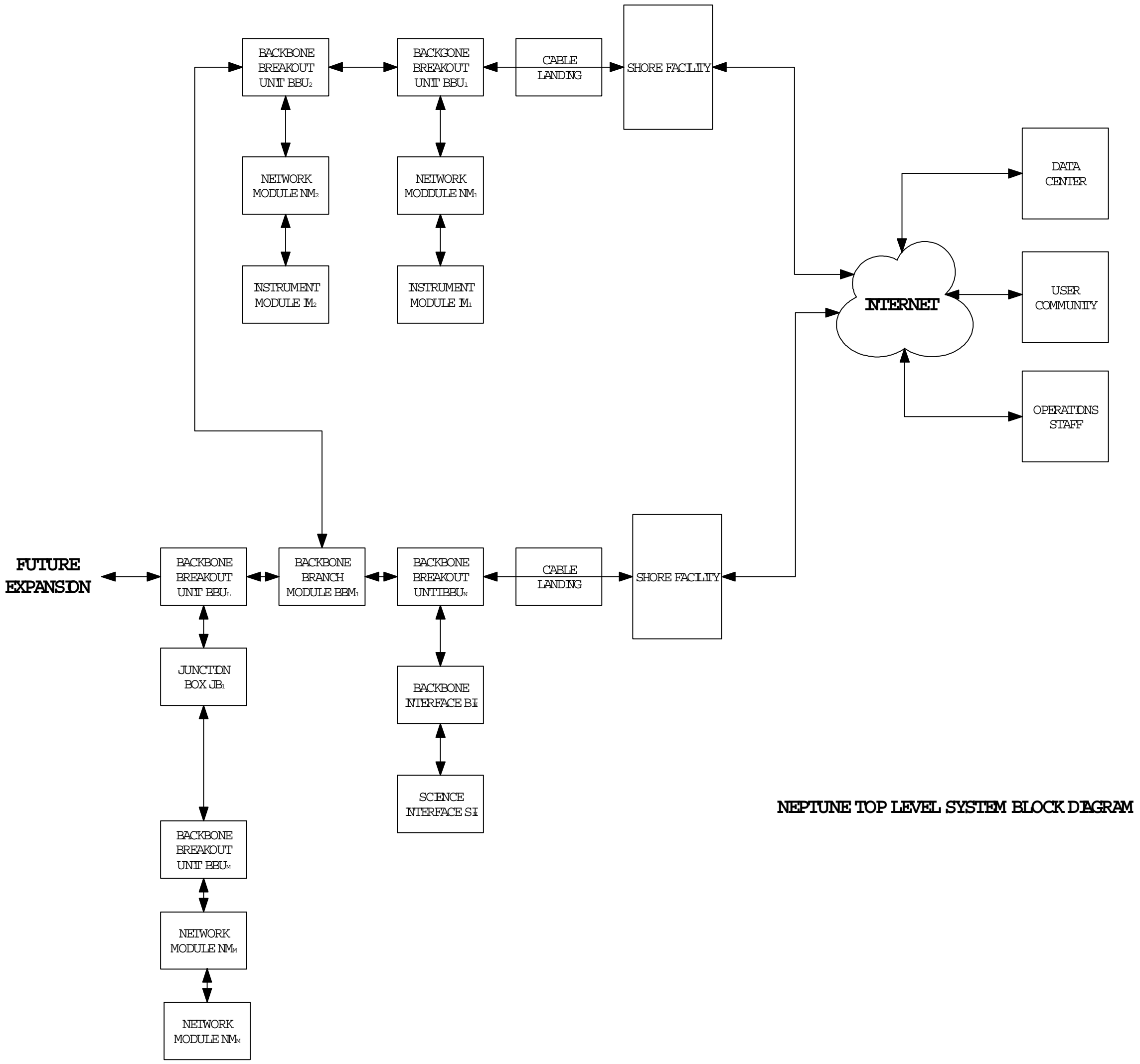

Figure 4. System-level block diagram.

some of the above functions might be physically separated from the shore station itself.

4. Cable

The function of the “backbone” cable is as follows:

- Transport power to the nodes, and
- Transport data to and from the shore stations and the nodes.

Standard “commercial off-the-shelf” (COTS) cable from the submarine telecommunications industry can be used, with a single copper conductor (1 Ω/km), a 10-15 kV rating, up to 30 optical fibers, and outside diameter under 25 mm. Because the cable has only a single conductor, sea water grounds at each node will be used.

In shallow or rough areas, the cable will have additional protection. It will be buried out to a water depth of about 2000 m to protect it from fishing trawls. (The placement of nodes as well as sensor networks on the continental shelf and slope will be challenging.)

The length of the cable network shown in Figure 1 is just over 3000 km. The cable will be the single most expensive item, at about $10/m.

5. Nodes

The functions of the nodes are as follows:

- Connect and branch the backbone cable,
- Distribute stable low voltage power and perform load management, relaying, and protection,
- Provide communications for the network backbone and the sensors and sensor networks, and
- Distribute timing signals.

The junction boxes that form the node are configured with underwater connectors that can be connected using a remotely operated vehicle (ROV). The entire system

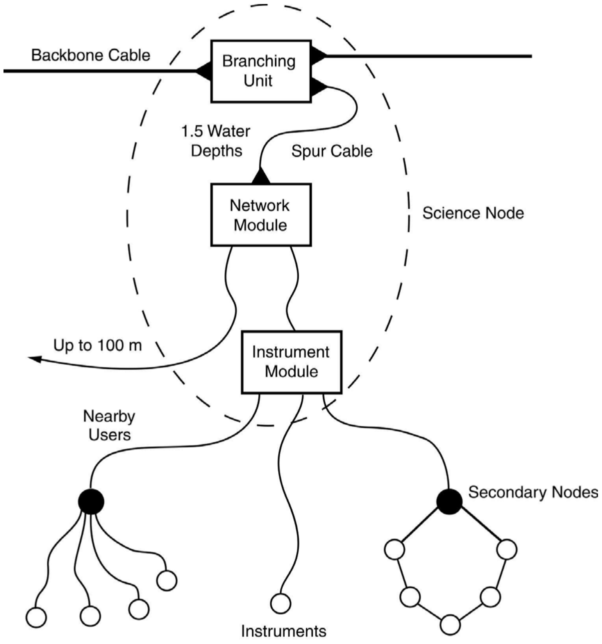

Figure 5. The conceptual layout of a NEPTUNE science node.

depends on these ROV-mateable connectors. Electrical connectors for this purpose are proven. However, underwater-mateable optical connectors are an emerging technology, and are less attractive from both a cost and a reliability perspective.

The configuration of the nodes and backbone cable will be such as to permit simple node recovery for maintenance and upgrade, using a research ship rather than a cable ship. Either a passive branch unit with a spur cable to the junction box will be used, or the junction box will be in a loop of cable, long enough to reach the surface. Such a configuration is crucial to keep maintenance costs down.

Figure 5 shows the layout of a science node using a single spur cable. The backbone fiber-optic cable contains an in-line, passive branching unit. The spur cable is 1.5 water depths long and contains two conductors and twice as many optical fibers as the backbone cable. The network module contains the high (~10 kV) voltage power supply and backbone router equipment, along with the low level control system. It can be recoverable for maintenance or upgrading using conventional research ships.

The instrument module contains the low voltage power supplies, low speed data switches, and instrument control systems. It serves as the connection point for scientific instruments. The instrument module can either be located quite close to the network module or up to 100 km away, and more than one instrument module may be attached to a network module.

An example of an in-service "node" is the Hawaii-2 Observatory (H2O) as shown in Figure 6. The junction box has recently been attached to the out-of-service Hawaii-2 analog telephone cable, half-way between Hawaii and California (Butler et al., 2000; Chave et al., 2000). The junction box is about 2 m long by 1 m wide by 1 m high and is constructed entirely of titanium and plastic for corrosion protection. It contains two pressure cases to house the system electronics. An oil-filled manifold is about 1 m off the seafloor and houses a set of wet-mateable electrical connectors to which instruments may be attached. The H2O junction box is designed to be recoverable for servicing, and plugs into the telephone cable at a termination frame. For NEPTUNE, the instrument modules are expected to be similar to the H2O design. The network mode will be permanently attached to the spur cable, via a gimbaled cable termination.

6. Power

The basic trade-offs in the power system design can be made easily. Alternating current at 60 Hz can be ruled out for the supply because of the effect of the cable capacitance. Compensating this capacitance for the length scales of NEPTUNE would cost more than the cable itself (NEPTUNE Phase 1 Partners, 2000). Low frequency alternating current (e.g., 0.1 Hz) is ruled out on the grounds that it is more complex and costly than DC. Therefore, the power system should use DC for the delivery function.

The choice between a series system (as in conventional submarine telecommunications) and a parallel system (as in terrestrial power systems) can be made on the basis of fault tolerance. Branching, and the provision of alternate circuits for power, is crucial to reliability. It has been estimated that the availability of the distribution part of the terrestrial utility power delivery system would be improved by a factor of between 5 and 10 if it were interconnected, instead of radial (Billinton and Jonnavitihula, 1996: Brown and Taylor, 1999). Since the technology to branch a series system does not exist, its development and deployment would add greatly to the system cost and complexity. Further, it can be shown that a branched series network is not capable of delivering as much power as a parallel network of the same topology; a parallel system is always more efficient. Therefore, a parallel system has been chosen.

The design approach is to maximize the amount of power that can be delivered to the science user, without a special cable. To provide less power than this amount would not significantly reduce the cost; to provide more would increase costs greatly. For NEPTUNE, the system will be able to deliver in aggregate about 100 kW to the distrib-

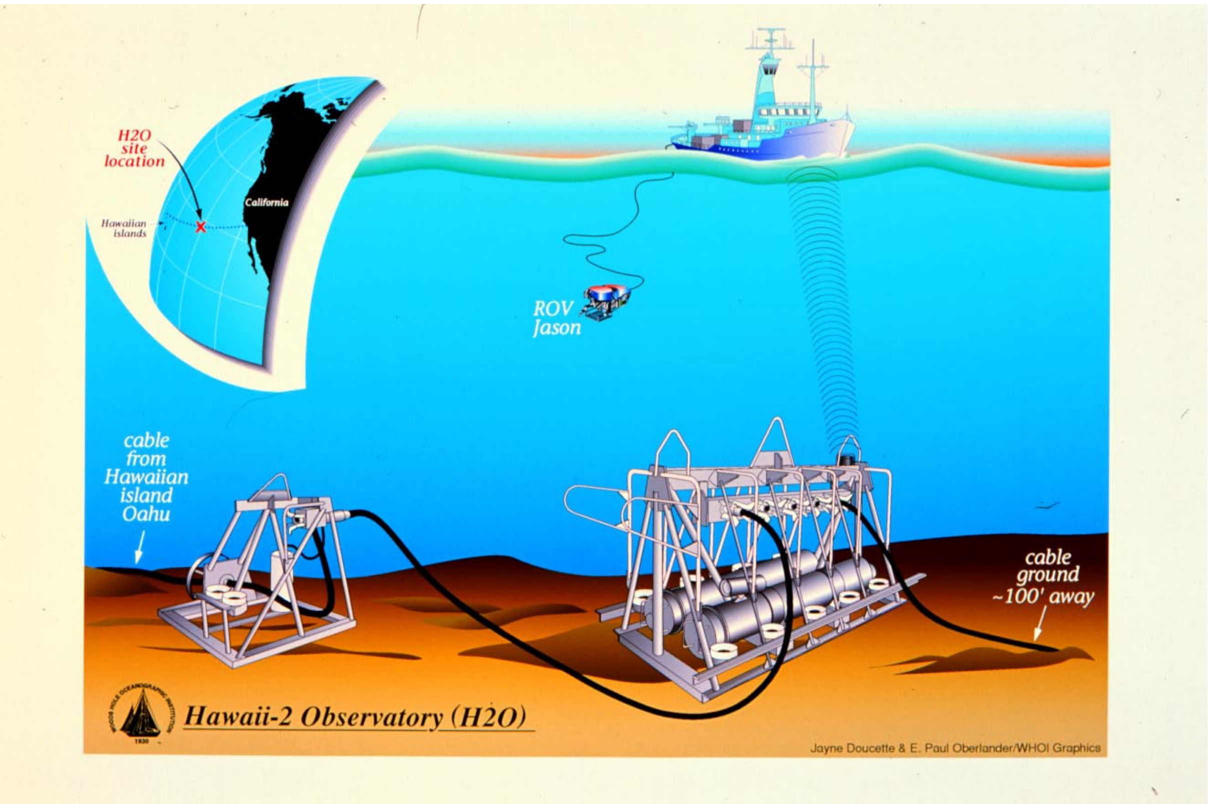

Figure 6. The Hawaii-2 Observatory (H2O) seafloor observatory system midway between Hawaii and California.

uted nodes on the seafloor with high reliability.

While some parts of the proposed power system might be available commercially-off-the-shelf (for example, the cable), a good deal of the way the system works will have to be tailored to the specific requirements of the NEPTUNE system. Indeed, power systems with many of the functional elements of NEPTUNE already exist except that they operate at much higher voltages.

Rather than go into a detailed description of all the components, we highlight those that will require the most attention.

DC/DC Converter

Within each node, the DC/DC converter must provide a stable low voltage output, given an input voltage that is varying considerably, up to 10-15 kV and 2-20 kW. The converter will be a voltage-sourced switching design. The details of the design, such as the switching frequency and the implementation of the necessary low pass input and output filters will be determined as part of the design task. Because these units form the foundation of the overall system, they will be highly reliable with redundancy built in.

DC Switchgear

Because there are no naturally occurring current zeroes, direct current is non-trivial to interrupt. This will be particularly true of the problem of interrupting faults in the NEPTUNE system, where the current available from the discharge of the cable capacitance can be large and sustained for several seconds. The approach routinely taken to implement a DC breaker is to force a current zero by means of stored energy.

Such a DC switch or breaker will be integrated with a “smart” programmable controller that would accept local measurements as well as remote commands. These “digital relays” will constitute a primary building block of the protection system and will be used, for instance, in branching units.

Protective Relaying

The purpose of the protection system is to isolate faults in the power system, so as to ensure the safety of personnel and equipment. For reliable operation, an essential part of this operation is to disconnect a faulted part of the system, while not disconnecting any parts of the system that are not faulted. In view of the difficulty (and expense) of performing repairs, such fault-tolerant operation is essential. Breaking the system up into sections is therefore an essential part of maximizing reliability.

There are several ways that faults can be detected. The simplest is based on the concept of overcurrent. A fuse is an example of an overcurrent protection device. A problem with this approach is that it does not distinguish faults that are close from faults that are remote. More than a minimum amount of the system may thus be disconnected.

By using information from several parts of the power system, the protection scheme can do a better job of discrimination between faults. For example, a distance relay gives better discrimination. In this device, a computation based on the measurement of both voltage and current is used to estimate the fault's location relative to the point of measurement. An even better protection scheme is the differential scheme. In this approach, the current at one end of a protected section is compared with the current at the other end. If there is a difference, there must be a fault between the two points of measurement. This kind of protection relies on the existence of a fairly fast communication connection between the two ends. In NEPTUNE, this connection certainly exists so that a differential protection scheme can be implemented by interfacing with the communication system. In the protection plan proposed here, an Internet-based differential relaying scheme will be the first line of defense.

For finding a fault location in a cable, a method that measures the relative arrival times of pulses of voltage or current from the fault can be used. This method relies on the availability of accurate time at both ends of the cable, which will be true for NEPTUNE (a fraction of a microsecond). Time-of-flight relaying will therefore be used as a second layer of protection.

In practice, because a fault might prevent direct communications, it would be prudent to adopt the utility practice of having several layers of protection, some of which can operate without communications. NEPTUNE's protection system will include differential, timing, distance, and overcurrent relaying. This way, the first (and best) lines of defense would be the differential scheme and the timing scheme; if the fault causes loss of communication, the distance relaying would operate; if the distance relaying failed, the overcurrent system could save the day. Since the loads are expected to be far more deterministic than those of a typical utility, overcurrent protection levels can be set quite close to the normal load values. These various levels of protection can be implemented using the digital relays mentioned before.

The task of designing and implementing the protection scheme is the largest within the power system design. Functionally, it is known from utility practice how to implement such a multi-layered protection system. The fact that the parameters being measured are direct current (not alternating) will make practically no difference. The biggest difference comes about because the DC sources (the shore station power supplies) and the system impedances provide a natural limit to any fault current magnitude, whereas on land, there is effectively no limit on current and ultra-fast fault clearing is needed.

However, the design will be new in that it is based on digital technology, and the first layer will use the Internet for communications. While digital relaying is widely used abroad, it is fairly new to the US, and relatively little US-made COTS hardware is available.

Maximum Power

While the voltage and current limits can be set simply by consideration of insulation lifetime and perhaps voltage drop, establishing a value for the maximum power that can be delivered by a network is more complex. Although undersea cables tend to be DC, the theory of maximum power has been addressed extensively only for AC networks.

As the point of maximum power transfer is approache, voltage stability manifests itself.. The symptoms of the approach of this kind of instability are a greater decrease in voltage than would be expected for a given increase in load. At the same time, the voltages around the system are hard to maintain. The problem is to determine the actual value of the maximum power transfer, bearing in mind that in the general case the load will not be the same at all the nodes, nor will it be time-invariant.

Being careful to avoid this point of instability, steady-state power flow calculations have been done to determine the maximum theoretical power that can be delivered to the planned 30-node system shown in Figure 1. For a cable resistance of 1 Ω/km and shore voltage of 10 kV, a maximum of 6.7 kW can be delivered equally to all the nodes. If the cable resistance is 0.7 Ω/km and the shore voltage is 15 kV (corresponding to a recently introduced long-haul COTS cable), the maximum power to each node is 21 kW. In practice, all nodes will not have equal loads, and one would operate at less than the theoretical maximum.

Dynamic Stability

A DC/DC converter that includes a regulator presents essentially a constant load to the source, since all load is on the regulated side of the converter. If the input voltage drops, the converter compensates and the input current increases. This means that the source will see a negative incremental resistance, and the system may go unstable.

Since the negative incremental resistance effect is a low-frequency one only (due to the regulation action), Middlebrook (1977) showed that any oscillations that occurred were likely to be at the crossover frequency of the input filter, where the filter impedance increased. He demonstrated that by choosing the filter components properly, the potential instability could be eliminated. Further, the filter design was not dissipative in the steady-state.

One study has already been done to demonstrate the stability of the system when subject to load fluctuations. A Pspice simulation of a simple three-node system was performed. The three nodes were separated by 150 km, with the first one 500 km from the shore station (cable resistance 1.5 Ω/km). The load at each node was adjusted to 13.2 kW (96% of the 13.9 kW theoretical maximum for this system). Figure 7 shows the results of one simulation with loads turning on and off at random. The response is clearly a stable, overdamped response.

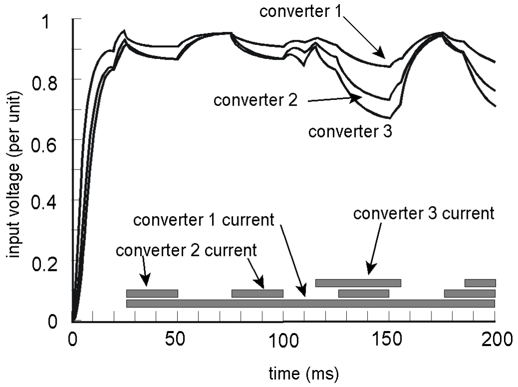

Figure 7. Results of a Pspice simulation with 3 nodes. See text for explanation.

One of the crucial elements is the input filter to the DC/DC converters; if the filter is removed, the system quickly enters an unstable oscillatory mode and subsequently crashes. This too has been simulated.

Note that this study demonstrates two aspects of the system behavior. First, it is unstable if the damping components in the converter filter are omitted. Second, it is stable with a power transfer level as high as 96% of the theoretical maximum. No unusual measures are required for such a transfer. These issues are further discussed in a recent paper (Howe, Kirkham and Vorpérian, 2000).

7. Communications

Backbone and Science Communications

For NEPTUNE, there is no need to distinguish between voice, video, and data as is common for a conventional telecommunication system. For this and other reasons, all NEPTUNE communication systems will use Internet Protocols (IP) that are now ubiquitous.

There are several standard communications technologies which might serve NEPTUNE, and selecting from them has been guided by three principles:

- NEPTUNE should use COTS components wherever possible to minimize development and acquisition costs.
- NEPTUNE should choose a technology which meets the requirements but not more.
- The communications technology must be available in a form factor and at a power consumption which is compatible with seafloor packaging.

Based on these and other considerations, Gigabit Ethernet (GbE) will be used to interconnect the NEPTUNE nodes. Ethernet is the most widely used data networking technology. GbE routing and switching hardware is readily available from many vendors. Standard converters (lasers and receivers) are also available with the ability to drive 100 km or more of single mode optical fiber. GbE hardware can serve as data concentrator/router/switch and a repeater, eliminating the need for separate optical amplifiers. Long haul 1 Gb/s full duplex communications uses a pair of optical fibers. Higher data rates are feasible using multiple fiber pairs; the NEPTUNE goal of a 10 Gb/s backbone can be met by ten pairs of fibers. Each NEPTUNE node will contain a pair of gigabit switch routers for redundancy.

A higher rate version of GbE (10 Gb/s) will be commercially produced in 2001, and will be evaluated when available. 10 GbE will have the same functionality as GbE, including range capability of 100 km on single mode fiber. The use of 10 GbE, and possibily some form of wavelength division multiplexing (WDM) could substantially reduce the fiber count in the backbone cable, with an attendant cost reduction.

The science instrument interface on NEPTUNE will be implemented using 10 and 100 Mb/s Ethernet as the communications technology. A standard user interface is being developed which will operate in a “plug and play“ mode, greatly simplifying the addition of new instrumentation to the network. This interface definition will include the incorporation of data about the data, or metadata, into the data stream.

Monitoring and Control

Monitoring and control of the NEPTUNE data network can be implemented using the simple network monitoring protocol (SNMP) served from network supervision work stations located on shore. SNMP clients can also be incorporated into the instrument module to simplify user level supervision. These monitoring and control systems will function when the data communications backbone systems are operational.

In addition, NEPTUNE will include a low level or out of band control system. This will provide full control of all router functions through a standard serial interface, and can be used to configure the network and download software to the backbone routers. The low level control system is fully redundant, and operates on a pair of dedicated optical fibers. The low level control system is being implemented using the serial ASCII interchange loop (SAIL) protocol (ANSI/IEEE standard 997-1985) operating over a fiber optic physical link.

Time Distribution

Accurate and precise time will be almost universally necessary for science experiments. Standard IP protocols are available for clock synchronization (e.g., NTP), with a few milliseconds accuracy across the NEPTUNE network.

For higher accuracy time requirements, GPS time information with an accuracy goal of 1 µs will be distributed using the low level control fibers.

Network Modeling

Efforts are currently underway at WHOI to use an advanced network modeling package (OPNET) to build a software model of the NEPTUNE network. When complete, this model will simulate network switches, routing protocols, cable failures, instrument data transport of various types (seismic, video, chemical/biological sensors, etc.), time synchronization characteristics, bandwidth growth requirements, and a number of other communications-related parameters.

Work is currently being done to complete the model of a seven-node laboratory test-bed version of the NEPTUNE network. This model draws components from a library of models for existing COTS network components for both the NEPTUNE backbone and the science junction box components. Cisco Systems is providing the hardware necessary to build the seven-node network represented by the model. By using measurements made on the seven-node network we will be able to validate the software model and thus have confidence in the full NEPTUNE network design.

8. Data Management and Archiving

The NEPTUNE Data Management and Archive Facility will provide the scientific community with on-line access to all NEPTUNE scientific data to ensure full exploitation of those data. The Facility will be an integrated part of the instrument planning, data acquisition, data processing, and data distribution processes that occur with NEPTUNE. Further, the Facility will guarantee that the valuable data sets obtained with NEPTUNE are preserved for use by future generations for research and education. A well-designed and properly implemented Data Management and Archive Facility, as part of the capabilities of NEPTUNE, will be a major contribution toward the full exploitation of the unique characteristics of the NEPTUNE data. An effective archive will boost scientific productivity and ensure that maximum value is extracted from the expensive-to-obtain data.

Data Acquisition

Protocols for packaging and transmitting data from sensors to shore stations via lower-tier junction boxes and nodes will be specified as part of the basic infrastructure. Standard protocols will be used, but will have to be documented. We will fully specify these protocols and all interface requirements for experiment designers in Phase 2.

The issue of metadata is the foundation upon which a successful data management and archiving facility will be built. Without it, data cannot be catalogued, which in turn means that it cannot be located and used effectively. Metadata generation must be automated in a way compatible with the standard sensor interfaces. The packaging of sensor data must be designed such that the relevant metadata can be stored with the data. Again, the research community must be consulted and choices made.

In addition to allowing cataloguing of data sets, metadata will be at the heart of the quality control and quality assurance plans. As the data flows through NEPTUNE systems, the quality of the metadata and the data itself must meet pre-defined standards before archiving. This will also allow problems in the data acquisition process to be flagged immediately.

Data Distribution

Phase 2 plans will draw on emerging technologies in information management and on the experience of organizations such as NRC, NASA, NOAA and IRIS, all of which manage large and heavily used databases. Scientific researchers are coming to expect user-friendly, web-based access to large data archives and NEPTUNE will follow this course.

There will be two major categories of data distribution within NEPTUNE. The first is data streams that will make raw data from the community experiments available to the full community of researchers and educators in near real time (subject only to network latency and the minimal delays associated with automated quality control). Similar mechanisms will also exist for making proprietary raw data from principal investigator (PI) experiments available in near real time to PIs and their teams. Both these types of data stream may flow directly from the shore stations.

The second is the query directed search in the archive catalogues. These web-based solutions will be crucial to successful implementation of a NEPTUNE archive, as they provide convenient access to data by all users, both new and established. In addition, the prevalence of open standards in the web community allows significant leverage to be obtained from standards developed for a large developer community, and significantly decreases the risk of technological obsolescence of NEPTUNE implementations. Increased emphasis on the web as an analysis as well as a data interchange environment will require that we consider metadata issues associated with the analysis of data as an integral part of any data management and archiving approach.

Data Handling

Given that the current estimates of NEPTUNE data flow are in the range of terabytes per day, the data handling components of the system will be designed for volume and automation. There are several key areas that will be examined in order to define these components. The system will require critical and robust criteria to determine which raw data are archived and whether processed data should be archived instead of raw data. Data compression in the form of lossy compression or subsampling algorithms will also be an issue and decisions on their use will come after consulting with the scientific community.

An important component of the data handling is the processing infrastructure. It is with this infrastructure that catalogue will be automatically built using sensor metadata; that quality assurance will be automated to raise alarms in case of problems; that raw data will be reduced for archiving (see above); that will allow archive users to specify the type and degree of processing to be done for their specific data requests; that will automate feature recognition or event detection in specific data streams. Such an infrastructure is required since users are much less interested in the actual raw data than in the *data products*. If these can be generated with consideration for data quality and the application, then the scientific community will be better served.

Data Exploration

A design goal for NEPTUNE is to provide data sets that can be easily linked and visualized in both the time and space domains to address cutting-edge problems and to allow completely new questions to be posed and answered. Database design and "data mining" techniques are subjects of active research. NEPTUNE will take full advantage of existing databases while remaining flexible enough to take advantage of new development. To ensure that the NEPTUNE catalogues meet the scientific goals, mechanisms will allow users to incorporate their NEPTUNE results and products (with sufficient quality control) into the archive. Automated data processing pipelines applying community standard algorithms in a systematic and homogeneous fashion over the whole collection of NEPTUNE data will create derived-data products and catalogues for user querying. The NEPTUNE archive will also contain links to publications based on data from community and PI experiments or from archival data.

Virtual observatories, a recent development in other disciplines[5], would enrich the data exploration environment by providing links for querying across different data collections or for exploring related data stored at other sites in a uniform manner. The NEPTUNE archive will explore establishing links with other centres (e.g. IRIS, PODAAC)

9. System Issues

The scenario published in the NEPTUNE Feasibility Report describes one of several possible designs and cost models for the backbone, science instrumentation, and other components of NEPTUNE (NEPTUNE Phase 1 Partners, 2000). The subsequent work, some of which is described here, has focused on both clarifying some aspects of the subsystems (e.g., the stability of the power system) and system aspects. We are beginning the process of analyzing the tradeoffs in the design. Throughout, we are striving for the system architecture that best addresses such factors as reliability, flexibility, operations, and maintenance. In the following we present a brief view of the work in progress, including some of the questions we are facing.

Network topology

The exact network topology needs to be defined to maximize reliability, to remove single points of failure, and maximize quality of service at all times including times when partial failures occur. For example, how can we maximize the reliability of a spur, where a single failure can compromise the rest of the system "downstream."

Cable Engineering

Cable vendors are being surveyed to identify the different possible backbone trunk cables. Tradeoff studies on the number and types of fiber, amount and structure of copper, insulation and standoff requirements, armor/protection packages, and jointing, etc., are needed. A strategy for shelf/slope cables and nodes is also required.

Nodes

A tradeoff study will be done to determine the optimal number of connecting ports on the network module and the instrument module, as well as the optimal voltages and data rates. Putting COTS components in underwater

---

[5] White Paper: A National Virtual Observatory for Data Exploration and Discovery, http://www.srl.caltech.edu/nvo/nvo7.0.pdf

pressure cases requires careful consideration in regards to thermal issues and reliability. How do mean-time-between-failure times quoted by industry for equipment normally used on land apply to this situation?

Connectors

Tradeoff studies of optical versus electrical connectors indicate that the former are prohibitively expensive, though reliability issues are not as severe as expected (but connector life cycle is limited to 100 matings). A validation/test program for connectors carrying 100/1000BaseT has been initiated.

Power

A tradeoff study will be necessary to determine whether the protection system should be based on digitally controlled breakers or on switches. The former could work under full system voltage, while the latter, which might entail fewer subsea components, would necessitate the system first to shut down and then restart.

Further, cost versus power capability tradeoff studies are being conducted.

Communications

A conceptual/functional baseline for the low-level control/timing system is being developed. Tradeoff studies of 1 and 10 Gb/s Ethernet are ongoing (the latter being an emerging technology), and this has been expanded to incorporate expected coarse WDM products.

The modeling of the communications system is on going. Once the seven-node model is validated the software components and associated assumptions will be used to populate the larger NEPTUNE model. As the project progresses we will continue to use both the seven node model and the larger model to simulate the actual operations of the NEPTUNE network in more and more detail. These models will help greatly in the design of the backbone and science nodes for NEPTUNE.

We expect the data management and the communications aspects of the power subsystem can also be incorporated into the software model as the NEPTUNE project progresses.

Data management and archiving

The studies related to data management are broadly grouped into two areas – data and topology (what gets done where). In the data area studies will be conducted to (a) determine file formats and metadata issues for different classes of sensors; (b) determine the classes of data which will be archived; and (c) estimate the initial data rates and projected rates over the lifetime of the project in order to ensure that the communications and storage infrastructure is capable of handling the data.

In the area of topology, studies will be conducted to determine (a) the cost of processing, compressing and buffering capabilities in the instrument/node versus the bandwidth and reliability of the link to the shore station; (b) the cost of data storage, raw data processing, catalogue creation, quality assurance, buffering capability and staffing at the shore station versus high bandwidth and reliability of the link to the archive centers.

Integration and Testing

The integration of the various subsystems and components into one complete system is essential. Even though individual and integrated components and subsystems will be thoroughly tested on land, a wet testbed that exercises the total system is a crucial element in the development.

A testbed is being planned for NEPTUNE in Monterey Bay. This will be crucial to NEPTUNE. In the short term, it will be used to test the overall integration of the infrastructure. In the longer term, it will be used for testing new sensors and instrumentation before they are deployed on NEPTUNE. Further, a Victoria Experimental Network UnderSea (VENUS) is being planned which will provide additional opportunities of this kind. The development of these facilities will be coordinated with the overall NEPTUNE effort.

10. Concluding Remarks

As with any similar development effort at this stage, there are unresolved issues, some which we have described here. Some, if not most, revolve around optimizing a particular system aspect given several good options to chose from. All the work done since the Feasibility Study reinforces the conclusions: NEPTUNE is technically feasible at a reasonable cost.

NEPTUNE's effectiveness is increased by a power system that can deliver high power levels, with the flexibility to grow, and the required protection elements to make it highly reliable. It will be able to deliver in aggregate about 100 kW to the distributed nodes on the seafloor with high reliability. Similarly, the communications system will be able to provide 10 Gb/s or more aggregate data rate. The data management system will provide easy access to the data for all the users.

Acknowledgments. The authors gratefully acknowledge funding from their respective home institutions as well as other sources: the National Ocean Partnership Program (NOPP), the National Science Foundation (NSF), and the National Research Council, Canada.

REFERENCES

Billinton, R., and S. Jonnavitihula, A Test System for Teaching Overall Power System Reliability Assessment, *IEEE Transactions on Power Systems*, **11**, 4, 1670-1676, 1996.

Brown, R. E., and T. M. Taylor, Modeling the Impact of Substations on Distribution Reliability, *IEEE Transactions on Power Systems*, **14**, 1, 349-354, 1999.

Butler, R., A. D. Chave, F. K. Duennebier, D. R. Yoerger, R. Petitt, D. Harris, F. B. Wooding, A. D. Bowen, J. Bailey, J. Jolly, E. Hobart, J. A. Hildebrand, A. H. Dodeman, The Hawaii-2 Observatory (H2O), *Trans. Am. Geophys. Union*, **81**, pp 157, 162-163, 2000.

Chave, A. D., F. K. Duennebier, R. Butler, R. A. Petitt, Jr., F. B. Wooding, D. Harris, J. W. Bailey, E. Hobart, J. Jolly, A. D. Bowen, and D. R. Yoerger, H2O: The Hawaii-2 Observatory, Elsevier series on Developments in Marine Technology, submitted, 2000.

Delaney, J. R., G. R. Heath, A. D. Chave, B. M. Howe, and H. Kirkham, NEPTUNE: Real-time ocean and earth sciences at the scale of a tectonic plate, *Oceanography,* **13**, 71-83, 2000.

Howe, B. M., H. Kirkham, and V. Vorpérian, Power system considerations for undersea observatories: The basic trade-offs, *IEEE Oceanic Engineering*, submitted, 2000.

Middlebrook, R.D., Input Filter Considerations and Application of Switching Regulators, *Proceedings of the Power Electronic Specialist Conference, PESC 77 Record*, 366–382, 1977.

NEPTUNE Phase 1 Partners (University of Washington, Woods Hole Oceanographic Institution, Jet Propulsion Laboratory, Pacific Marine Environmental Laboratory), *Real-time, Long-term Ocean and Earth Studies at the Scale of a Tectonic Plate: NEPTUNE Feasibility Study* (prepared for the National Oceanographic Partnership Program), University of Washington, Seattle, 2000.

National Research Council (NRC), *Illuminating the Hidden Planet: The Future of Seafloor Observatory Science*, National Academy Press, Washington D.C., 2000.