

NEPTUNE Stage I Network Architecture

P. Phibbs

NEPTUNE Canada, University of Victoria, Victoria BC V8W 2Y2 Canada

S. Lentz

Lentz Telecommunications Strategies LLC, McLean VA 22101 U.S.A

Abstract-Developed over the course of more than six years, the NEPTUNE Stage I network will be the first ocean observatory to link a wide variety of deepwater science instruments with the global Internet. The development of NEPTUNE Stage I poses many challenges. The network must deliver high bandwidth and high power, while being manageable and reliable.

NEPTUNE Stage I addresses these challenges using a combination of commercial submarine telecommunications equipment, commercial off-the-shelf equipment intended for enterprise applications, and newly developed components. A single, optically amplified fiber pair and distributed Optical Add-Drop Multiplexing (OADM) branching units provide a protected optical path to each node. Each optical path has an effective capacity of 2 Gb/s. Layer 2 switching and IEEE 802.3 ad Link Aggregation Control Protocol provide a protection mechanism and allow traffic to be shared on the two diverse paths thus making 4 Gb/s bandwidth available to each node during normal operation. The powering solution employs a constant voltage scheme capable of providing up to 10 kW at each node and over 120 kW total. An intermediate voltage of 400V is used for distribution with the node and from the node to instrument platforms.

I. BACKGROUND

The network architecture for the NEPTUNE Stage I Regional Cabled Observatory (RCO) is designed to meet the requirements of the science community both now and many years into the future. The science requirements establish that the observatory must deliver three essential functions: power, communications and timing. In addition, the observatory must employ standard interfaces, be expandable, be fault tolerant, and have a twenty-five year design life. In setting these requirements, a philosophy of encouraging new applications, rather than designing for existing applications, is followed. Accordingly, the power, communications, and timing requirements exceed what might reasonably be needed today. By making high power, high bandwidth and precision timing capabilities available, NEPTUNE is intended to

encourage new forms of data collection and ocean observing.

Transforming the science requirements to a functioning ocean observatory entails many steps, which are only briefly summarized here. Early on, a decision was taken that the observatory backbone and primary nodes should be constructed by an industrial partner, rather than by academic organizations. This decision was based on industry's ability to manage the development process, to control design risk, and thus to maintain a fixed price and scope in the event of unforeseen difficulties. This decision drove much of what followed. The original science requirements were redrafted in a format suitable for use as the technical portion of a Request for Proposal (RFP) and became known as the "Functional Requirements." A draft RFP was circulated to likely respondents and the comments received were incorporated into the final RFP document. Existing design work, reports, and evaluations were made available to respondents, who were free to incorporate this into their responses or to select alternatives. From the outset, the commercial telecommunications cable providers were deemed best positioned to respond, however most proposals were led by a systems integrator with a commercial telecom company as a subcontractor. The selection of a prime contractor to construct NEPTUNE Stage I was based on design, ability to implement, control of project costs, and consistency with NEPTUNE's long-term objectives. The RFP process concluded with the award of a contract to Alcatel Submarine Networks (ASN) in September 2005. The Functional Requirements used in the RFP then became the basis of the Technical Requirements which define the work to be performed under the contract.

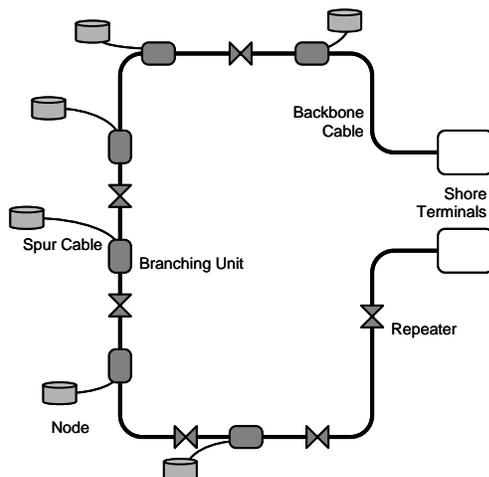
The Functional Requirements define what functions the NEPTUNE network must perform, but do not define how to deliver or achieve that performance. Accordingly, many of the design selections were made by ASN or by ASN in consultation with NEPTUNE. By accepting ASN's offer, NEPTUNE also accepted many of Alcatel's design decisions. Thus, the RFP process played an essential role in establishing the network architecture for NEPTUNE Stage I.

Before reviewing each functional area in detail, it is worth noting that each aspect of the network architecture influences the others, so the overall network architecture represents an integration of each functional area. Additionally, the ability to deploy and maintain the network are taken into consideration when making design decisions. To control cost and development risk, the selected architecture makes extensive use of Commercial-Off-the-Shelf (COTS) components, although some of these may be used in ways not originally intended. Where this is the case, a series of qualification and/or demonstration tests are conducted to verify the component will serve the intended purpose. Because it is not always apparent how a change in one functional area or component will affect the rest of the network, the final architecture was arrived at through an iterative process.

II. GENERAL ARCHITECTURE

The architecture adopted for NEPTUNE Stage I is a trunk and branch topology as illustrated in **Error! Reference source not found.** As will be shown, this architecture achieves the desired functionality for both power distribution and communications. Most importantly, it allows the trunk or backbone to be constructed exclusively from components designed and qualified for use in commercial sub-sea telecommunications systems and leverages the many years of design experience and high reliability of these components. The architecture supports up to ten primary nodes attached to the backbone; the initial implementation of NEPTUNE Stage I will have six branching units and five equipped nodes along an 825 km cable route.

The trunk or backbone cable forms a continuous loop from one landing point to another. For practical



reasons, both NEPTUNE Stage I landings are at the

Fig 1: Cabling Architecture

same location, naturally forming a ring. The network design could easily be adapted so that the backbone stretches between two landing points, with terrestrial communications channels closing the ring to provide diversity.

The network nodes are placed at the ends of branches off the main backbone cable. The current design allows a distance of up to 24 km between the backbone cable and node. The nodes incorporate newly developed components as well as COTS components originally designed for use in terrestrial telecom or data networks. Most of the design and development risk is confined to the nodes; the network architecture allows the network to continue operation even if one or more nodes are missing or faulty. The nodes are designed to be recovered and repaired if necessary. Because the entire network does not depend on any single node, nodes can be placed in areas where the risk of damage from underwater activity is relatively high, while the backbone can be placed in safer areas. Modifications, changes, and upgrades to the network can be made by replacing the nodes, while the backbone remains unaffected.

Each node provides six interface ports for connection of science platforms or extensions. Each port provides dual optical Ethernet links. A total of 9 kW of electrical power is shared among the ports; a single port can deliver the full 9 kW if desired. Electrical power is delivered at 400 VDC. In addition, two connection points to the backbone cable voltage are provided to support long extensions. Underwater wet mate connectors are used to allow Remotely Operated Vehicles (ROVs) to make the necessary connections between the node and extension cables.

III. POWER

The objective for power deliver for NEPTUNE is 10 kW per node. Of this, 1 kW is allocated to observatory operations, i.e. the communications and control equipment. This magnitude of power capability is an example of provision for future applications, but it is not difficult to envision the power requirements of several motors running winches or pumps, together with high intensity discharge lamps for cameras, when added to the cabling losses downstream from the node, totaling 10 kW. This power delivery would not be continuous, but would be highly variable, with a low steady-state load along with periods of high usage.

Both to deliver a high power capability and to deal with highly varying loads, a system of parallel powering is used. Power is applied at the cable ends, each of the nodes is a load in parallel with each of the other nodes, and sea water provides the return path. This method of power distribution is analogous to a

terrestrial power grid and is unlike a telecommunications cable system which employs serial power feeding. The benefit of serial power, with a constant current running between a positive cable end and a negative cable end, is that a cable fault to ground can be accommodated by moving the zero voltage point to coincide with the fault location. The drawback to serial power is that a power converter for a serially powered system which can provide up to 10 kW and deal with variations in the load does not appear feasible today. The scientists' demands for significant amounts of power drove NEPTUNE Stage I to a parallel configuration. Parallel feeding has the additional advantage of accommodating the trunk and branch cable topology, including additional branches to extensions beyond the nodes. The powering architecture is illustrated in Fig. 1.

NEPTUNE Stage I will employ a power feeding voltage of -10 kV applied to each cable end at a maximum current of 8 amperes. Analysis shows that this can provide 10 kW at 10 nodes, even after accounting for losses over a maximum of 1800 km of backbone cable. Medium Voltage power Converters (MVCs) at each node receive 1 to 2 A at 5 to 10 kV and provide a 400V output at up to 25 amps. The MVC incorporates a high level of internal redundancy and is designed such that most failure modes result in a reduction of power handling capability rather than a complete loss of service.

The 400V power is distributed to each of the node interfaces. 400V was chosen because it is the highest input voltage for which COTS DC-DC converters are readily available; the highest possible voltage is used to minimize power losses in the extension and distribution cables which connect science instruments to the node. The 400V is stepped down to other voltages as needed to power each science load; typical input voltages to a science instrument are 12 or 48V. All low voltage power converters are configured with up to four way redundancy.

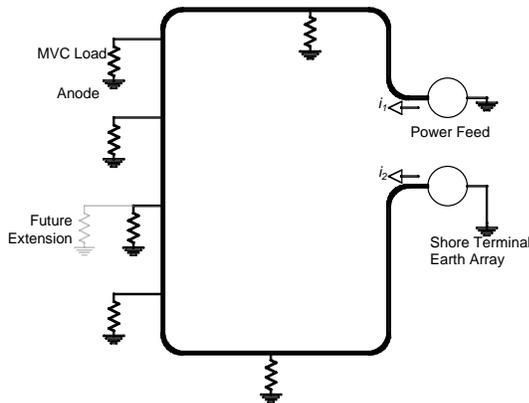


Fig 2: Parallel Power Feeding

Fig. 1: Parallel Power Feeding

The parallel power distribution deals with faults by means of Power Switching Branching Units (PSBUs) which connect each node branch to the backbone trunk cable. The PSBUs used on NEPTUNE have been modified from standard designs both to accept the higher currents and to support a switching state in which all three legs are connected. This “all connected” state is the normal state for the PSBU and allows current to flow both along the backbone and into the branch. In the event of a cable fault, the PSBUs can be reconfigured to isolate the faulty section. A faulty trunk segment can be isolated by switching the two PSBUs on either end from trunk to branch. A faulty branch can be isolated by switching the power path through the PSBU from trunk to trunk; this configuration can also isolate a node which has catastrophically failed. The same approach is used for shunt faults (short circuits to seawater) as for cable breaks. Reconfiguration does entail shutdown and restart of the network and total power delivery may be reduced until the fault is repaired, but the parallel powering scheme allows the network to remain in operation in the presence of shunt faults, cable breaks, and node faults.

IV. COMMUNICATIONS

The design of the communications architecture for NEPTUNE Stage I is driven by the need for open, standardized interfaces. Because the worldwide Internet employs TCP/IP and Ethernet has become the de facto standard for interconnecting data communications devices, these standards were selected for use in NEPTUNE. Ethernet can be readily adapted to serial data interfaces such as EIA-232 (ITU-T V.24), EIA-422, or EIA-485 as needed. Adaptation to less common standards, such as digital TV camera interfaces, is also possible.

An objective for total bandwidth of between 10 and 20 Gb/s for the entire network was seen as suitable for both present and future requirements. As with the power requirements, this is not driven by existing requirements, but rather to support future, as yet unspecified, requirements. The vast majority of science instruments produce only a small data stream; applications generate more than 100 Kb/s are rare. Even a 1 Gb/s network can accommodate hundreds of such instruments. To date, the only significant use of bandwidth proposed for NEPTUNE Stage I is High Definition Television (HDTV) cameras. Operating in native mode an HDTV camera produces a data stream of 1.5 Gb/s. Various forms of compression can be used to reduce this to 19 Mb/s, 38 Mb/s or 110 Mb/s, but these all result in some signal degradation. Lossless

compression, together with a reduction in frame rate, can reduce the bandwidth requirements to just under 1 Gb/s, allowing 1000Base Ethernet to accommodate HDTV.

COTS workgroup Ethernet switches fulfill the requirement for Ethernet interfaces and can easily handle the expected bandwidth. These devices provide 10/100/1000Base-T and some optical interfaces, a full Layer 2 feature set and may also handle Layer 3 routing. For redundancy, a pair of Ethernet switches is placed in each node. The science interface ports each contain a pair of optical Ethernet links; 100Base-T is used within the nodes

The optical links from node to shore present another challenge. Since the Ethernet switches support various optical interfaces, it would be natural to use optical Ethernet for node to node connections. However, even with Erbium Doped Fiber Amplifiers (EDFAs), the maximum reach of these links is limited to about 200-225 km, which is greater than the distance between some nodes. Adding in-line amplifiers does not increase this length, since the Ethernet receivers are not designed to deal with accumulated optical noise. Additionally, this configuration makes each node act as a regenerator for data from more distant nodes.

The solution employed by NEPTUNE is to employ COTS SONET/SDH equipment which is compatible with optically amplified links and commercial submarine telecom repeaters as in-line amplifiers. A distributed Wavelength Division Multiplexing (WDM) scheme is used to allow each node to transmit and receive on its own dedicated frequency directly to and from the shore station. Optical devices in the PSBUs drop and add the optical signals for the nodes. Optical amplifiers in the PSBUs and repeaters provide the necessary signal gain so that nodes can be any distance apart.

The resulting architecture, illustrated in **Error! Reference source not found.**, has a number of beneficial features. Because there is a signal path from each node directly to the shore station, each node is independent of every other node. This allows nodes to be bypassed or left out of the system. The philosophy of using proven, commercial submarine telecom components in the backbone is maintained. By transmitting in both directions on the backbone, a protected ring is formed. The design of the optical line, i.e. the shore terminals, repeaters, and branching units, is such that up to ten nodes can share a single pair of fibers, thus reducing the cost of cable and repeaters.

SONET/SDH transponders operating at 2.5 Gb/s were selected because they are available as COTS equipment and are designed to work over optical amplified links. The optical transmitters are narrow band lasers and the receivers have been tuned to deal with accumulated optical noise. Additionally, the transponders incorporate Forward Error Correction (FEC) to improve the system performance margin and bit error rate. Although normally used in metropolitan or regional networks, ASN has demonstrated that the transponders are suitable for use over submarine cable spans of up to 1800 km without modification. Low power utilization and compact size makes them suitable for installation in a submarine housing. Finally, the availability of 1000Base-T tributary interfaces allows them to connect directly to the Ethernet switches. SONET/SDH equipment is used because it is the only COTS equipment that could transport 1000Base Ethernet over an optical line including multiple optical amplifiers, not because it was necessary or even desirable to have a SONET/SDH layer in the network. The SONET/SDH capabilities of this equipment are not used; its primary function is to provide a transparent pipe for two 1 Gb/s Ethernet signals.

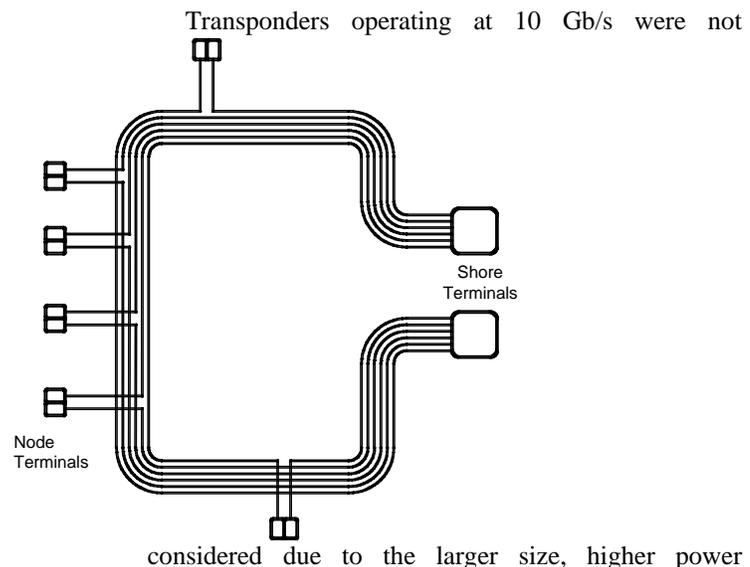


Fig 3: 2.5Gb/s Optical Channels

consumption and lower reliability as compared to 2.5 Gb/s transponders. Additionally, operation at 10 Gb/s requires dispersion compensating fiber be added at each terminal. However the optical line design is capable of 10 Gb/s operation. So, when suitable 10 Gb/s transponders become available the nodes can be upgraded to increase the available bandwidth. Because each node operates independently, this can be done on a selective basis, with only those nodes which require extra bandwidth being upgraded.

The usable bandwidth provided by the NEPTUNE Stage I architecture totals 4 Gb/s per node. One optical signal is sent each direction along the backbone to and from each node; each optical signal carries two 1000Base Ethernet channels. Since each node has its own dedicated channels, the total network bandwidth scales with the number of nodes. Thus, a ten node network would be capable of 40 Gb/s total capacity. The six node network being built will have 24 Gb/s capacity, which exceeds the original objective.

The use of IEEE 802.3ad Link Aggregation Control Protocol (LACP) allows all the available bandwidth to be shared while transparently rerouting traffic in the event of a failure. Thus, in normal operation all four 1 Gb/s Ethernet channels are available to carry traffic. In the case of a fiber break or other fault, traffic is routed on the remaining channels. As with the power delivery capability, communications capacity is reduced by a fault but the network can continue operating.

V. TIMING

The science requirements for NEPTUNE also included delivery of precision timing. The stated objective is ± 1.0 μ second error (also referred to as "dispersion") with respect to Universal Time Coordinates (UTC). To achieve this level of accuracy, some method of correcting for transmission latency, which may be as much as 10 ms, is needed. Much thought was given to providing a dedicated timing signal, but this was not an ideal solution because it added another set of long range optical transponders and would have entailed additional custom development. Instead, a solution using IEEE 1588 / IEC 61588 Precision Time Protocol (PTP) to deliver precision timing via Ethernet was selected.

PTP is capable of accuracies within several hundred nanoseconds on Local Area Networks (LANs) and when used in conjunction with switches that implement PTP in hardware (boundary clocks). Standard network switches degrade the accuracy of PTP, but early testing shows the accuracy available on NEPTUNE to be better than ± 5 μ s. This is a preliminary result and some additional improvement can be expected. Given this result was achieved with no additional complexity and

minimal additional cost, it is viewed as satisfactory for initial operation. Various methods of improving the timing performance, such as confining all PTP packets to a dedicated VLAN or dedicated Ethernet links, long term averaging, and improved local oscillators, are being considered and will be implemented if warranted. As PTP becomes more widely used, network switches with the boundary clock function can be added to the nodes, either as part of a repair or as part of an upgrade; when this is done, the expected accuracy is on the order of ± 300 ns.

Because PTP is delivered via Ethernet, precision time is available at any point on the network without the use of additional equipment. Precision time can be acquired by instruments in a variety of ways. Instruments with a hardware PTP capability will achieve the best results; PTP implemented with off-the-shelf network interface cards (NICS) will be about an order of magnitude less accurate, but is still about two orders of magnitude better than Network Time Protocol (NTP), which is typically accurate to ± 2 ms. An adapter card which implements a PTP client and outputs a pulse-per-second (PPS) signal can be provided to support older instrument designs.

VI. CONTROL

The primary function of the control system for NEPTUNE Stage I is to monitor and control power to the science interfaces. Custom designed circuit boards are used to measure current, detect current imbalances, and to interrupt power to each science interface. A pair of control processors in each node communicates with the shore terminal over the Ethernet data network. A processor based in the shore terminal communicates with the control processors in each node and provides a Service Oriented Architecture (SOA) interface to NEPTUNE's Data Management and Archiving System (DMAS).

Network management is accomplished through the use of the capabilities provide with each network element. The Ethernet switches support a command line interface and HTML Graphical User Interface (GUI); the SONET/SDH equipment, repeaters, PSBUs, and submarine line terminals employ a conventional telecom network manager.

VII. EXPANSION

Achieving the science goals for NEPTUNE requires that the network be extended and expanded. The most common extension will be from the nodes to the science instruments. Each node will support science instruments over an area ten to fifteen km across, so the most common method of expansion is to employ an

extension cable and junction box to place instruments some distance from the node. Power is delivered over these cables at 400 V and 1000Base-LX optical Ethernet is used to provide a nominal 8 km span length. Extensions of up to 100 km are possible using long-range Ethernet media converters and connecting to the 10kV power distribution plugs. Each node can support two such long extensions. Finally, the optical line design allows the backbone cable route to be extended to a length of up to 1800 km.

trunk or backbone forms a ring, providing redundant paths for both communications and power. The backbone contains components qualified for use on commercial telecom cables and provides extremely high reliability. PSBUs control power flow into the branches while passively routing optical signals to/from the branches. Precision timing is delivered using IEEE 1588 over Ethernet. The six interface ports on each node provide optical Ethernet and power at 400V. COTS equipment is used in the nodes, providing somewhat less reliability while reducing design and development costs. Extensions from the nodes will allow instruments to be placed at any location within 8-10 km from a node.

VIII. SUMMARY

The NEPTUNE Canada Stage I architecture is designed to achieve the objectives established by the science community. Power and communications are delivered over a trunk and branch architecture. The

Future capabilities include upgrade of the node-shore data links to 10 Gb/s and expansion of the backbone ring.

ACKNOWLEDGMENTS

The authors of this paper would like to thank the Canada Foundation for Innovation, the British Columbia Knowledge Development Fund, Woods Hole Oceanographic Institution, Monterey Bay Aquarium Research Institute, University of Washington, the Jet Propulsion Laboratory, Alcatel-Lucent, and its subcontractors L-3 Maripro and Texcel.