

DMAS Advisory Committee Report
June 16 /2009

This report was largely based on contributions made by the DMAS committee at the May 25 meeting that took place in Toronto, ON.

DMAS Advisory Committee Meeting Members

Present (Advisory Board)

Bill St. Arnaud, Chief Research Officer, CANARIE, (Chair)
Bruce Spencer - Research Officer, Internet Logic, NRC Institute for Information Technology
Eric Tsang - Associate Director, Business Development, Communication Research Centre
Pierre Quesnel - Head, Systems and Networks Management, NRC Canada Institute for Scientific and Technical Information

Those present from UVic (UVic)

Martin Taylor - President & CEO, Ocean Networks Canada by conference call (part of meeting only)
Benoît Pirenne - Associate Director, Information Technology, NEPTUNE Canada
Mairi Best - Associate Director, Science, NEPTUNE Canada
Eric Guillemot - Manager, Software Development, NEPTUNE Canada
Chris Barnes - Director, NEPTUNE Canada (early part of meeting only)

Overview

On May 25, 2009, the NEPTUNE Canada DMAS committee was convened to review progress of the DMAS project and to discuss future options and issues with respect to security, production, future directions, funding strategies and comments on Oceans 2.0

Overall the committee was impressed with the progress made on the DMAS system and NEPTUNE Canada staff were able to address satisfactorily issues and concerns raised at the previous DMAS meeting.

Most of the discussion focused on funding issues and future directions for DMAS and Oceans 2.0

Specific comments and recommendations:

1. It is clear that the biggest challenge facing NEPTUNE/VENUS is ongoing operational funding. The committee discussed at length various options. One suggestion is for NEPTUNE/VENUS to explore how they can be a possible platform in measuring Canada's possible compliance requirements under the upcoming Copenhagen UNPCC treaty negotiation which may set new targets in ocean acidification in addition to atmospheric temperature increase. As well NEPTUNE/VENUS may want to explore how they may get funding from Pacific Carbon Trust and Environment Canada to support some of their work in this area. The committee recognizes that one of the challenges facing NEPTUNE/VENUS is that it sees itself as, and was funded as, a

research facility as opposed to a platform that helps validate and confirm government policy in areas of security, tsunami detection etc. Although one can understand the liability concerns of tsunami detection, positioning themselves as a policy testing and verification system with respect to compliance for upcoming climate treaties, especially in the area of ocean acidification may not be too far a stretch from their research objectives. Note that NEPTUNE/VENUS should talk to Environment Canada Ottawa offices who are responsible for the Copenhagen UNPCC negotiations. Ocean acidification, from a policy and treaty perspective is likely to be in the purview of Environment Canada as opposed to DFO.

Knowledge translation from the NEPTUNE/VENUS/ONCEE programs to public policy is a major objective of ONC, and its importance for strengthening the case for sustained operating funding is fully recognized. To this end, the ONC Board has approved the creation of an ONC Public Policy Advisory Committee to advise ONC management on effective ways to strengthen the links with relevant federal and provincial agencies with ocean-related mandates. The ONC Board itself has several members who are/have been in relevant senior positions in government (Minister/DM/ADM level). Within ONC, Martin Taylor and Brian Bornhold are currently in discussion with senior government staff to identify short and longer term opportunities for applications of the NEPTUNE/VENUS/ONCEE results to priority issues on the public policy agendas within and across federal and provincial departments and agencies.

2. Another area of possible new revenue opportunities is to repurpose DMAS as a tool for universities and businesses to undertake Green House Gas (GHG) baseline measurement and ongoing monitoring. Flue stacks are one of the major sources of GHG emissions at most universities and instrumenting and monitoring these devices, plus computers servers and other major GHG sources, may help universities and businesses reduce their CO₂ footprint. As well there could be substantial revenue for DMAS in terms of carbon offsets to allow universities to license the software. The challenge for many carbon offsetting programs like the Pacific Carbon Trust is to find ways to spend their money on highly qualified carbon reduction solutions. NEPTUNE may want to look to apply to upcoming CANARIE Green IT pilot program in pursuit of this objective.

DMAS thanks the committee for what is a really intriguing possibility of development going forward. The opportunities with using the Pacific Carbon Trust to finance such a development would seem to require quite a bit of consultation with the somewhat complex structure of this new provincial organization.

A quick assessment of the technology available in the area of networks of CO₂ sensors did not reveal much that would be available at the present time, hinting at a potentially untapped market. Unfortunately and because of so many other deadlines and on-going activities with VENUS and NEPTUNE, there was no time to prepare a meaningful GreenIT proposal to CANARIE by the submission deadline. The new ONC business development staff will nevertheless endeavour to explore this avenue in the months to come.

3. Another area of long range discussions is the role of NEPTUNE in terms of Oceans 2.0 and the development of tools to assist researchers. Although Oceans 2.0 and related work is a critical technology, the committee questioned whether it should be seen to be driven by the research community as opposed to NEPTUNE staff? For example US

Cyber-infrastructure and CANARIE NEP programs have always promoted the concept of Virtual Organizations (VOs), which are formalized organizations of researchers, who then direct the development of such tools as Oceans 2.0. A good example is the HUBzero program at Purdue U. This issue reflects an age old debate as to whether NEPTUNE/VENUS should be a simple platform that collects data, like a telescope, and that any value added services such as Oceans 2.0, hydrophone array processing software, etc should be directed by a formal group of PIs (ie VO) funded by PI research dollars, rather than NEPTUNE operating dollars. The committee understands this a complicated issue and there is no clear answer, but in general feels that the more the research community is seen advocating and funding the development of these tools through funding programs like CANARIE NEP, the clearer the distinction in roles between NEPTUNE DMAS as a research platform versus NEPTUNE DMAS as a set of specific research tools. One suggestion is that the development of research tools like Oceans 2.0, hydrophone array processing, etc. be undertaken by ONCEE as opposed to NEPTUNE DMAS.

Several comments come to mind here:

- 1. HUBzero is indeed a system that has many similarities to what the Oceans 2.0 environment is preparing to offer: group work capabilities, code upload, use of Grid computing etc. Where Oceans 2.0 differs is that it offers the ability to directly use archived and real-time data in user code, has features to control underwater assets in real-time, etc.*
- 2. Community-driven requirements are fine in principle, but not always easy to come by. Even when they form virtual organizations, the communities do not always know what is possible with today's technologies and simply resort to very conservative approaches.*
- 3. The Web was invented at a research organization but not necessarily defined by a group of potential users. It was a born out of technical abilities and imagination of a few, as the evolution of cruder, pre-existing tools.*

It appears that, as we move rapidly into the realm of data analysis as opposed to just data collection and re-distribution, we have to ask ourselves about the possible role of ONCEE/NC/V in driving progress. One could argue indeed that this task should be the sole responsibility of scientists, perhaps through the various Science Advisory Committees of the projects: they could propose which advanced IT-related program is most urgent to undertake. ONCEE could support a call for proposals from the community, the selection of the best data analysis project and fund the realization of its outcome. However, ONCEE does not have significant funds to support large grant applications.

- 4. The committee recommends that all racks and equipment at shore station have seismic mounts. Experience from the Richter 4.5 earth quake in Santa Cruz several years indicates that it does not take much shaking to loosen circuit boards and/or connections. A minor earth quake near Port Alberni could easily disrupt primary data collection even if the circuit to Victoria is still operational.*

DMAS Systems and Operations have given this recommendation serious thoughts. An offer was received for the insertion of a seismic platform under four racks and it appears that the cost of the retrofit of the existing installation would amount to about \$13K. The issue is that only two of the racks have any significant equipment inside and only one hosts the data collection and buffering functions. In consultation with the engineers in charge of the shore station, it was recommended against this

new expense as in the event of an earthquake having a significant impact on the shore station equipment, it is likely that other elements of the system will be affected more. The more fragile elements of the system would be disk drives and those are configured in a redundant manner, mitigating the impact of the failure of one or more of them.

5. The committee recommends that all GUIs for managing power systems, DMAS, etc should be designed to be viewable and accessible through iPhone or equivalent. This will allow staff on call, (or management) to access control systems from their iPhone anywhere in the world. It is becoming an increasingly popular practice for network systems managers. The committee understands that some GUIs may be based on X- Windows etc which will make this more challenging than those based on Java or Linux. Note that VNC can be run easily on iPod and iPhone devices. This means you may not need to write a new front-end. You would just install a VNC server on the computer you want to control remotely and a VNC client on the iPod/iPhone to connect to it.

DMAS Systems and Operations are consolidating alerts coming from the various subsystems of NEPTUNE Canada, VENUS, DMAS, Alcatel-Lucent or other manufacturers onto a Network Management System that is making extensive use of SNMP traps. Alerts are then immediately relayed to mobile phones with which, in some circumstances, initial verifications or restart of some subsystems can be made. However, extensive debugging is not usually possible as fault cause isolation and subsequent correction of the issue may require sifting through large amounts of log entries or engineering data.

6. In light of the many ground fault problems to date the committee recommends that data on this phenomena be carefully tracked in the DMAS and that new estimates on MTBF (Mean time to failure) and MTTR (Mean time to repair) be assessed after a year or two of operation.

DMAS will detect events taking place on the underwater system and record them. So there will be an ability to search those events later on and generate statistics on demand.

7. For security and reliability purposes the committee recommended that NEPTUNE setup a secondary web site with a separate DNS name and IP address. In the event of a DDoS attack or Kaminsky DNS spoof against UVic, NEPTUNE researchers across the country would still have access to NEPTUNE instruments and databases through this site. One possible solution is to install load balancer hardware in front of the web service front end (Apache). CISTI has implemented F5 Big-IP load balancers that work very well. (See <http://www.f5.com/products/big-ip/>). A Gartner study on load balancers has also been sent to Benoît Pirenne.

The DMAS Systems team has been very active in pursuing the set up of a Disaster Recovery/Business Continuity Plan (DR/BCP) for the past years. Negotiations with a site at a remote location are now completed and we are happy to report that the University of Saskatchewan in Saskatoon (UofS) will be hosting a complete copy of the NEPTUNE and VENUS data on its WestGrid storage. Moreover, a number of servers are being added to the raw storage with the view to implement a Business Continuity Plan. Following the committee's recommendation, we are implementing the provision of most DMAS user services from either site through a load balancing system. The entire set up should be in place by the end of October. The data

replication between UVic and UofS will take place through the provision of a dedicated lightpath, courtesy of CANARIE. A formal MOU to define this relationship and activity has been drafted.

8. The committee recommends that NEPTUNE or ONCEE staff should start compiling some data dictionaries or ontologies using OWL or RDF to cross reference DMAS meta tags to common XML mark languages used by other disciplines such as Geography XML, Environment XML, GeoXML, sensor XML etc. This would aid researchers in interpreting and converting data sets for their discipline. We recognize the efforts of the interoperability workshop held last November for setting up this ontology. The data modeling efforts will be important. And challenging. Compared with astronomy, ocean observatories have much more dynamic data to capture: it changes rapidly and has many more modalities. The connection with Google Earth 5.0 could lead to a vast audience. We also recognize the efforts toward allowing scientists ability to run MatLab code on your servers as potentially creating an audience, much more specialized. It may be possible to create interest through these efforts but a lot of time could be spent servicing the scientific and public audiences and it seems to be a question for the larger organization and ONCEE. Suggest that some direction be sought.

It is not really DMAS's intention to produce new ontologies. Where it can be done, we will endeavour to re-use some that already exist. As an example, in the framework of a new collaborative agreement between NANOOS (the Northwest Association of Networked Ocean Observing Systems) and VENUS, a review and translation of this organization's ontology for our system will allow us to deliver NANOOS-compliant data.

Also, through the CANARIE Network Enabled Platform NEP-10 project (Oceans 2.0), we are looking at following the lead of DFO's IMDS organization, which collates a lot of the marine ontologies and recommends standards supported at the UNESCO level. Those are likely the ones we'll cross-match initially with.

9. In terms of future membership of the DMAS committee, NEPTUNE may want to consider inviting a participant from OOI or USCS Cal-It2 who are responsible for the OOI cyber-infrastructure.

That can certainly be arranged and we will contact them in time for the next DMAS Advisory Committee Meeting. Two names come to mind: Matthew Arrott, project manager of the OOI-CI and Dr. Oscar Schofield, project scientist.

10. Storage strategies were also discussed. Neptune DMAS projects it will need 50 Terabytes of storage per year. Various alternatives were evaluated and DMAS selected Data Direct Networks as a provider. This addresses one copy of the data, but the hardware and location for the second copy are to be determined. This storage will cause an increase of Neptune's carbon footprint. It is recommended that alternative storage solutions for the second copy be investigated. Amazon is an example, but it is not deemed to be carbon friendly at this time.

Indeed, the selection of the UofS site for the second distant copy of the NEPTUNE and VENUS data is one that will indeed increase the carbon footprint of the projects, as electricity in Saskatchewan is generated by coal-fired power plants, though Saskatchewan is also looking at nuclear options.

An alternative solution would be to move the entire equipment back to the BC interior, near a BC hydro green power plant. The cost of the move would however be

significant as we would have to support the monthly hosting costs of the privately owned data centre such as that of RackForce.

11. There was a discussion regarding service level agreements (SLA). It is understood that equipment problems are difficult to predict and repair, so a SLA covering these factors is not feasible. However, hours of coverage and response time to an incident can be predicted. It is recommended that the SLA response times be defined to set appropriate expectations with researchers.

DMAS acknowledges the recommendation and will publicize the expected response times to various issues or enquiries. The SLA details will depend in a large part on the way operational support will be organized. Negotiations with UVic Human Resources and the staff are looking at a 5x10 + on-call roster arrangement.

12. Video recognition : Bill recommended looking at the Hub for ideas.
<https://hubzero.org/home>

A pilot project with IBM aims at streaming video data from DMAS to this vendor's "System S" processing infrastructure, fitted with video pattern recognition software. It would indeed be possible to make the tools available through an interface such as HUBzero or the Oceans 2.0 platform. We will look into this option together with IBM.