1 Vision for the Platform (0.5 page)

We propose to coallesce and leverage existing data sets using international standards for data interchange to be shared by several domains of knowledge, including astrophysics, oceanography, remote sensing, fisheries, emergency management and robotics. This platform would allow for the exciting possibility of combining data and algorithms across these different domains. Example use cases include:

1. A forest fire-fighter consulting her smartphone before approaching the fire to see an immersive view of when the fire is at this moment from near real-time remote sensing combined with GIS data of the terrain. Perhaps more importantly, she can also see where the fire will be in five or ten minutes by harnessing simulations of the propagation of the fire using the most recent data on the fire, wind and the forest.

2. The development of a robotic emergency rescue first responder equipment through a simulation of the design in the actual terrain of interest. With parallel computation, thousands of potential designs could be tested and modified using a genetic algorithm in a matter of minutes to find the optimal design.

3. Management of natural resources combined with new techniques such as fracking combined with climate change, potentially taking into account the subtle impact on agriculture at a global scale.

4. A public health researcher correlates medical outcomes in a private database against geographical data to discover new trends underlying disease, without gaining access to any private data.

2 Expected Impacts and Added Value (1 page)

Many Canadian built research tools either involve domain specific geographical data (e.g. Ocean Networks Canada) or specialized provisioning of distributed computational resources (e.g. CANFAR). Currently, to interact with this data or perform computations, one not only has to be an expert in the particular sub-discipline but also in each interface of the underlying tool or the data representation. We believe these efforts are now mature enough to compose their collective APIs into one standard interface available for users from around the world to collaborate through a web browser.

The key tenet of the proposed platform will be to exploit existing tools, data sets, and standards to provide an interactive interface for industry stakeholders and the public. This will enable multi-disciplinary and multi-sectoral networks of stakeholders to create strategic international partnerships. Not only will exchanges be accelerated, but existing Canadian efforts will be better internationalized. Promotion of standards would dramatically increase access to existing infrastructure and data, foster collaboration among international stakeholders and ultimately yield new understanding of our environment, natural resources, more rapid development of robotic technology and better situational awareness for defense and emergency management.

CHORUS will enable partners to deploy new sophisticated information products take advantage of the new data that will be available in the near future (such as constellations of high resolution SAR and Optical satellites). The enhanced cloud platform will have the capability to perform big data analytics and generate powerful new information products based on the analytics results. We envision the technology to have multiple international market opportunities.

The first is to exploit the powerful new information products that can be generated from big data analytics combined with the huge data archive. These opportunities include creating services around these information products and/or selling the information products to service companies. Big data analytics is a rapidly growing market segment that is growing at a 26.4% compound annual rate and is projected to be \$41.5 B annually by 2018 (around the time this project will be underway and the technology is ready for exploitation). This dwarfs the global earth observation market which is currently around \$1.9 B annually and projected to get to $3B annually by 2020. Enhancements made to CHORUS will allow competing directly in the big data analytics market with a powerful cloud based system. Therefore, the market potential for this technology is very high, and it only takes a fraction of a percent of market share to see significant revenues.

Another application of this technology is to take CHORUS and deploy this entire ecosystem on to a private cloud infrastructure that can be sold to customers around the world as a stand alone product. This targets mainly large government organizations and institutions and also large companies (e.g., insurance companies) that are interested in the cloud based computing capability but want to have the system behind their own firewall so they have complete control of the security. The value of this technology is very high to each of these customers (i.e., easily in the 10âĂŹs of million dollars) which is what it would cost for them to develop this capability plus it would cause them significant delays in acquiring the capability if they developed it themselves. Therefore, a price of several million dollars per instance is readily defendable. Since there are hundreds of organizations around the world that would be highly interested in this type of capability, the revenue potential is greater than $100M.

3 Model for Collaboration (1 page)

By offering an open and collaborative platform, CHORUS will strengthen, expand and intensify meaningful international collaborations across a variety of knowledge domains, including astrophysics, oceanography, remote sensing, fisheries, emergency management and robotics. The need for a common infrastructure, in particular for applications that crosscut both the knowledge domains and datasets involved, will build links and ensure shared benefits between national and internation institutions alike. We have successfully assembled leading practioners from both academic and industry stakeholders involved in extracting information from existing Earth Observation platforms.

4 Strategic Plan (2 pages)

gathering of existing tools: SEDRIS, HLA, VO, MAST interface

gathering of data sets and algorithms (new stakeholders)

development of interoperability tools and standard

identification of infrastructure partners (google, amazon, microsoft): where will the data go, how to get computing, how to put it all together

The objective of CHORUS would be to create a collaborative, international development environment for algorithms and tools designed to harvest information from shared earth observation data repositories. Sharing applications within this environment has the following advantages:

1. It allows the system to move the code to the data, which speeds up the processing and reduces the cost;
2. It can make the applications available to other clients and users;
3. It encourages collaboration within the CHORUS community, possibly including crowd-sourcing.

Tasks will include:

1. Developing standard workflow practices: for example, users first find data through the use of the catalog API, then assign the selected dataset to a working set, and then launch individual tasks/algorithms or a chained set of tasks/algorithms.
2. Designing automated and efficient mechanisms to deploy algorithms and applications as tasks within the system. These would for example take advantage of Hadoop and Spark for mass batch processing and analysis.
3. Automatically selecting CPU or GPU (CUDA) optimized code for analysis tasks.

4. Exploring strategies such as containerized runtime virtualization (for example using Docker) to run algorithms at scale through the CHORUS environment.

5. Automating how algorithms can be re-factored for speed, for example through feature engineering and with reference to the machine learning literature.

6. Assessing the trade-off space between standardization of workflows and optimization of workflows for data throughput, with reference to various classes of problems.

7. Designing and optimizing the user experience for interacting with and sharing workflow containers within the ecosystem developers and users.

8. Experimenting with âĂIJcloud sourcingâĂİ of applications. This might for example use a revenue-sharing model similar to AppleâĂŹs âĂIJApp StoreâĂİ.

The vision for this project is to provide critical new capabilities based on what is now redibly available Earth Observation datasets. The goal is to take CHORUS to a stage where the technology is ready to be commercialized by industrial partners. Modern constellations of commercial surveillance satellites that are now being developed are of great interest to both civilian and defence agencies around the world. Such constellations will consist of many satellites with frequent revisit capability and will generate huge amounts of data. To meet the requirements of many users, those high volumes of data must first be translated into geometrically and radiometrically corrected imagery, analysed to extract the embedded information, and then delivered around the world in minutes. This capability can only be achieved through Cloud Computing on systems such as CHORUS.

This work builds on existing Geospatial Cloud Platforms and tools that currently enable the ability to generate information products suitable for such tasks as search and rescue and related defense applications that include big data analytics applications. In addition, the IP that is developed will have high commercialization potential and the technology will be taken to a stage that it is ready to be commercialized by industrial partners.

The most critical metric for success will be the engagment factor from the mulitple stakeholders involved. If CHORUS is successful, then it will attract additional datasets, algorithms, and information products from all of the knowledge domains involved. This method of crowd sourcing will also identify internationally-important knowledge gaps and the shared platform will be the foundation upon which we can build new collaborations.

Development of CHORUS will adhere to an Agile Methodology. Agile iterative approach that builds software incrementally from the start of the project, instead of trying to deliver a completed platform near the end of 4 years. In each iteration, stakeholders will be asked to come to a consensus on key elements of functionality to support. This approach will allow the project to most easily adapt to new opportunities and unexpected challenges.

5 Proposed Team (2 pages)

The partnership consists of representatives from UVic, UBC, SFU, Compute Canada, Ocean Networks Canada, the Space Telescope Science Institute, International Virtual Observatory Association, and industry partners from Urthecast and Magnetar Games. The key individuals highlighted in this brief LOI are all Canadians with established extensive international collaborations. Consolidating efforts through CHORUS will intensify these relationships with partners internationally and further expand them to include more knowledge domains.

UrtheCast will provide multispectral remote sensing data and have strong expertise in GIS. Magnetar games provide expertise on immersive user interfaces and augmented reality. We envision developing web-based prototypes allowing a non-expert citizen scientist to not only access the data but to interact with the data as an expert would, creating new simulations and perhaps more importantly to combine data from different domains to generate new understanding.

Yvonne Coady

Yvonne leads the systems research group at the University of Victoria, exploring cloud-based application infrastructure and scientific visualization. Her group has had impact in optimizations and efficiencies afforded by new hardware and infrastructure in a wide range of scientific applications, including the Thirty Meter Telescope and Near Field Tsunami Detection and Warning Systems. As a co-recipiant of the University of Victoria's Knowledge Mobilization Award, and a co-winner of Johnson & Johnson's Cognition Challenge, she has been both locally and internationally recognized for her participation in Knowledge Translation activites.

Jeremy Heyl

Tania Lado Insua

Tania has investigated environmental impacts of oil-spills on mussel communities using population genetics techniques, and holds a Diploma of Advanced Studies in Marine Biology and Aquaculture from the University of Vigo, in collaboration with the University of Puerto Rico, MayagÃijez. She obtained MS and PhD degrees in Ocean Engineering from the University of Rhode Island with research focusing on applying models of sediment physical properties to past and present climate change. Her most recent research includes international collaborations on diverse research topics such as environmental impact evaluation for renewable energies, paleoceanography, physical properties of the sediment, geohazards monitoring, seafloor observatories and paleoclimate.

Brian Thom

BrianâĂŹs research focus is on the political, social and cultural processes that surround Indigenous people's efforts to resolve Aboriginal title and rights claims and establish self-government. His written work explores the interplay of culture, power and colonial discourses in land claims negotiations, and examines the political and ontological challenges for Indigenous people engaged with institutions of the state. He is principle investigator for the research project Innovations in Ethnographic Mapping and Indigenous Cartographies, currently funded by SSHRC and Google. This project grapples with the practical problem of implementing socially and politically powerful mapping initiatives which can effectively visualize and communicate indigenous peoplesâĂŹ knowledge and experience of the land. Through cutting-edge immersive, networked, multimedia cartographic systems, this project offers a response to the critique that the dots and lines of conventional ethnographic mapping reduces and essentializes indigenous territoriality and senses of place.

Keith Beckett

Keith is known as an energetic, highly-productive software, systems and project engineer, successfully tackling extremely challenging problems, bringing out-of-the-box thinking to the table and getting the job done. With a vast experience in leading large and small system development teams, architecting customer-focused solutions, systems and software engineering, critical analysis and research. His specialties include: Systems Engineering, Software Engineering, Computer-based Systems, Numerical Analysis and Methods, Image and Signal Processing, Remote Sensing using both Optical and SAR Sensors, Modeling and Simulation, High-Performance Computing, and Cloud Computing.

Duncan Suttles

6 Management and Governance