Integration of Kepler with ROADNet: Visual Dataflow Design with Real-time Geophysical Data

Fricke, T.T.†,Δ and B. Ludäscher ‡, I. Altintas‡, K.G. Lindquist, T.S. Hansen†, A. Rajasekar‡, F.L. Vernon†, J. Orcutt†*

Abstract

The ROADNet project concentrates real-time data from a wide variety of signal domains, providing a reliable platform to store and transport these data. Ptolemy is a general purpose visual programming environment in which work flows on data streams can be constructed by connecting general purpose components. The Kepler scientific workflow system extends Ptolemy to approach design and automation of scientific data analysis tasks. In this work we discuss our integration of ROADNet (and the Antelope platform on which ROADNet is based in part) with the Ptolemy environment.

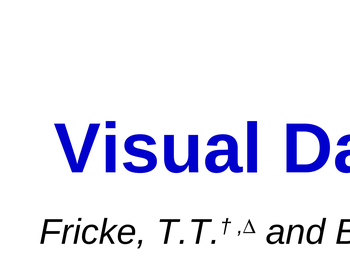
We have produced interface components that allow someone using the Kepler scientific workflow system to readily use ROADNet data resources. Presently we have working components to gather real-time waveform and image data from ROADNet object ring buffers, and we are working to provide the ability to perform Datascope database queries from Kepler. The Kepler project, including the Antelope interface, is entirely free and open-source, and will run on any platform where Java is available.

We discuss existing applications in addition to possible future directions, such as coherent array processing, event detection, and online stream processing. A major advantage of the Ptolemy environment is the ease with which it may be used for rapid prototyping of analyses by even inexperienced users. For instance, a user can drag-and-drop an Orb Waveform Source component and several general purpose analysis and display components, connect them visually, and immediately perform an analysis on real-time data.

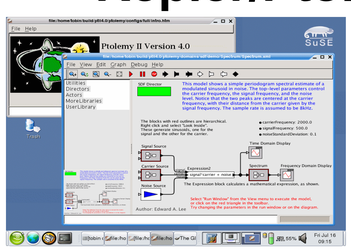
ROADNet/Antelope

The **ROADNet** system concentrates data from many different sources into a general purpose system. A crucial element of ROADNet is the use of orb servers, "object ring buffers" which concentrate and distribute realtime data. Orbservers and other elements of ROADNet are part of the **Antelope real-time system** by Boulder Real-Time Technologies, Inc.

http://roadnet.ucsd.edu/



Kepler/Ptolemy



Kepler is a system for graphically designing and automating data flow models that encapsulate "scientific workflows." Kepler is based on UC Berkeley's **Ptolemy II** system for heterogeneous, concurrent modeling and design. Both Kepler and Ptolemy II are distributed as free software.

http://kepler-project.org/

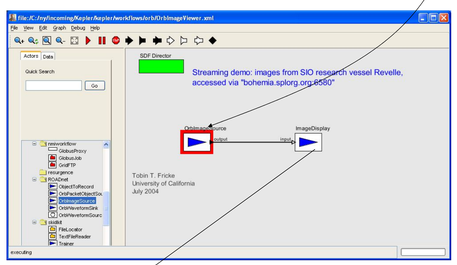
Our project is to integrate these two systems, providing access to ROADNet resources from Kepler.

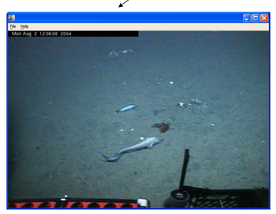
Examples of Current Applications

Remote Video. The JASON submersible, teathered to the R/V Revelle, explores the sea floor off the coast of the Aleutian Islands. Images and other data are transmitted in real time to the ROADNet system, based in San Diego.

With the ROADNet actors we have developed, a Kepler model in New York can connect to the public Orb in San Diego; the model receives and processes real-time data (in this case, images) from the JASON submersible in Alaska. The Kepler system consists entirely of free software.

ORB
Antelope Object Ring Buffer
- event driven
- reliable, fault tolerant
- message-oriented
- multicast
- content-addressed
- simple interface





This simple model requests and immediately displays frames from one of JASON's cameras.

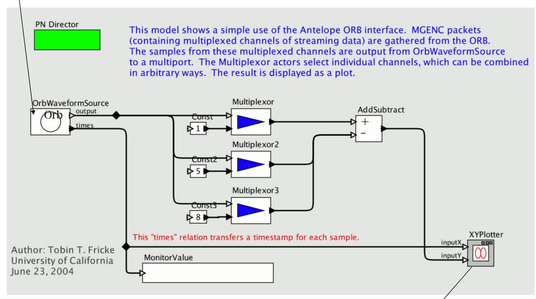
Prototype signal processing applications. Laser strain gauges in the California desert transmit their signals to the ROADNet system.

"*What we'd like to display is series #2 minus #6 and #9, #6 and #9 being the two main `correction' series. Once these series are subtracted (as you would in a spreadsheet, data sample by data sample) the result ought to be a convincing earthtide signal…*

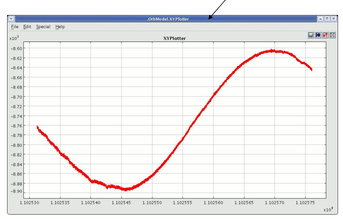
For a feeling on this, we in San Diego are moving up and down about a foot, twice a day, from the center of the earth; the strainmeter measures the stretching of the surface as this happens."

– **Frank Wyatt** suggesting an example application

We can quickly construct a Kepler model to encapsulate this task. The "OrbWaveformSource" connects to the Orb and acquires the necessary data stream.



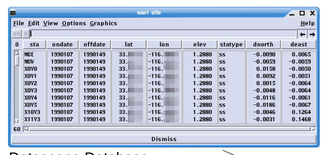
Kepler/Ptolemy contains many general-purpose components for data display and manipulation. Here we use a plotter; we could also use **OrbWaveformSink** to send the computed signal back to the Orb.



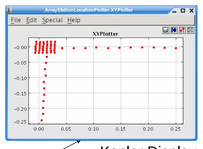
Build models that process real-time data using drag-and-drop components.

Database connection

A full Kepler-ROADNet integration requires not just connectivity to Orbs (real-time data) but also to relational databases. Kepler already has support for many databases (via JDBC); we are working on an interface to Datascope.



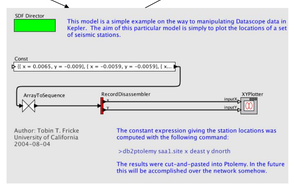
Datascope Database



Kepler Display

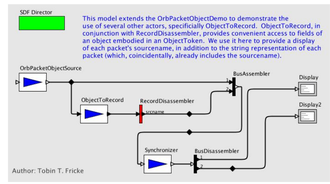
Using existing tools, we can transform the results of a Datascope database query into XML and then transform these results into a form suitable for Kepler. The data may then be used in a Kepler model.

A next step is to automate this.



Kepler Model

Packet Handling



We're not restricted to processing Orb data as timeseries. We may also operate on a higher level, manipulating orb *packets* in arbitrary ways. This has potential applications in prototyping Orb packet routing systems.

Conclusions

- The Kepler-Antelope interface makes it easy to manipulate realtime data in an intuitive manner, without writing code.
- Kepler models that send their output back to the Orb may be a useful means of expressing a stream processing task.
- Future work could implement more sophisticated signal processing (leading to applications such as coherent array processing), distributed processing, and modelling of higher level tasks such as event detection and packet routing.

† Institute of Geophysics and Planetary Physics, UCSD (IGPP)
‡ San Diego Supercomputer Center (SDSC)
**Lindquist Consulting*
*Δ now University of Rochester; contact **tobin@pas.rochester.edu***