The GONG++ DATA Processing Pipeline

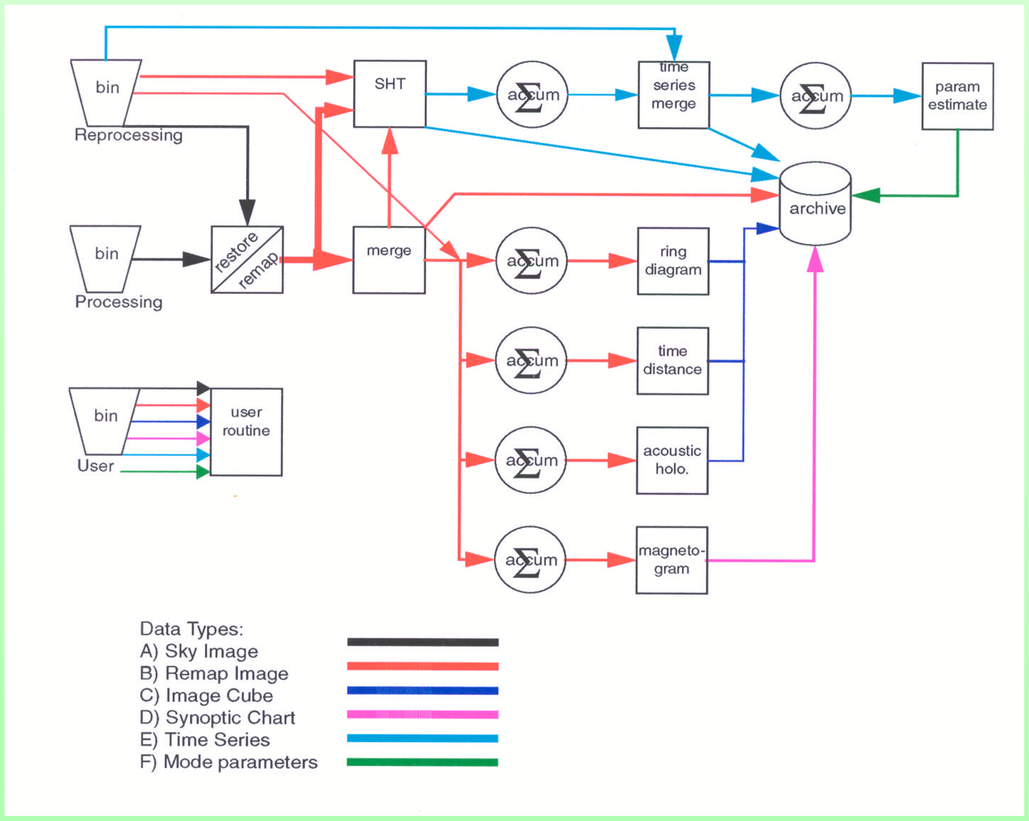
F. Hill,, J. Bolding, C. Toner, T. Corbard, S. Wampler, B. Goodrich, J. Goodrich, P. Eliason– National Solar Observatory

Introduction

The GONG++ data processing pipeline is currently under construction. This system will produce the GONG++ science data products, including subsurface synoptic flow maps, subsurface sunspot flow maps, farside images, tracked and remapped data cubes, ring diagrams, time-distance diagrams, holographic images, high-degree mode parameter estimates, merged velocity images, merged magnetic field images, and magnetic synoptic charts. In order to deal with the increased complexity of the data processing task compared to the earlier GONG Classic data, we have adopted an automated pipeline approach to the system. Here we describe the software and hardware and present our experience in the integration of a "mini-pip

Hardware

- Sunfire 4800
- Twelve 900-MHz CPUS
- 24 GB RAM
- Six T3 disk arrays, capacity 4.7 TB
- L180 Tape Library, capacity 18 TB on 180 LTO cartridges



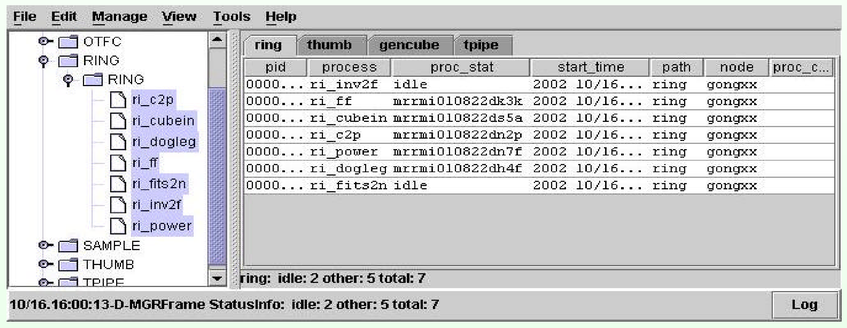
A flow chart/block diagram of the GONG++ data processing pipeline. The square blocks represent analysis modules. The arrows represent the flow of the data, which will be controlled by the OPUS pipeline package and the Veritas hierarchical storage management system. The circles represent the accumulator stages that monitor the temporal coverage of the data products and initiate processing steps as needed. The trapezoids represent the input to the pipeline, and the cylinder represents the on-line archive.

Software

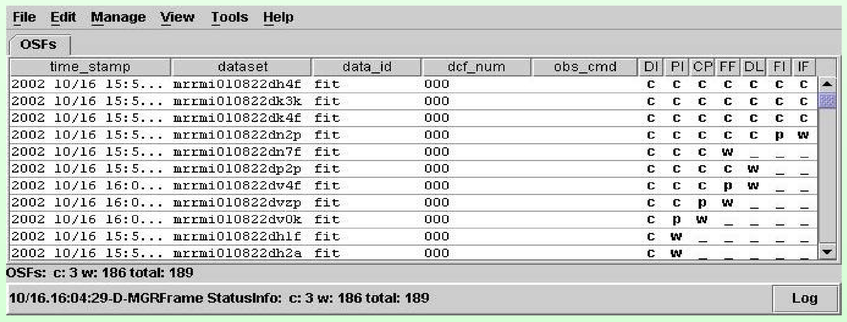
- OPUS pipeline framework from STScI to automate the data flow
- Veritas tape library and file management
- In-house and community analysis modules in any language
- "Accumulator" modules to control processing schedule

OPUS

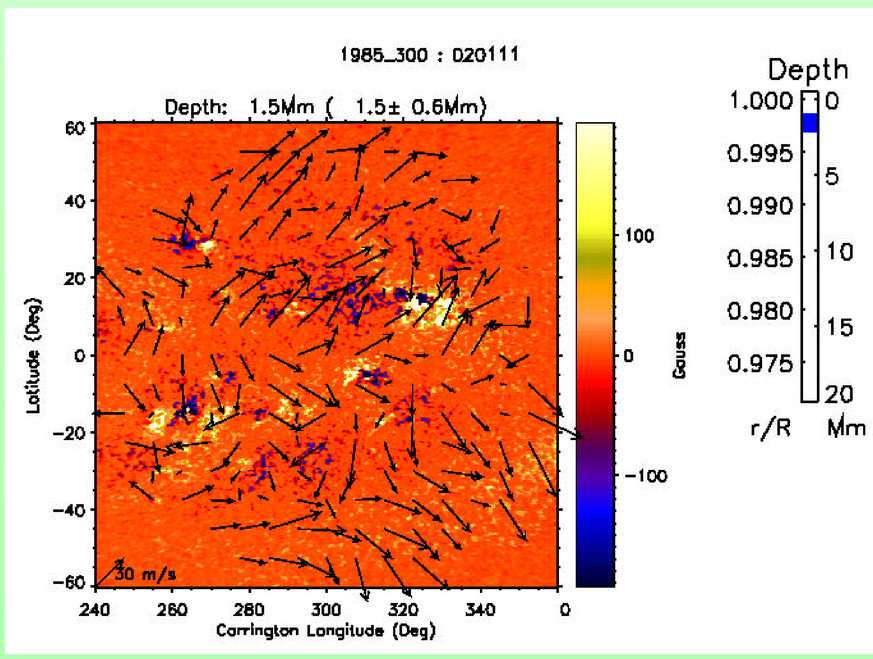
OPUS was originally developed at the Space Telescope Science to process HST telemetry input to a standard output. However, it is actually a general-purpose powerful tool that provides a fully distributed pipeline processing system for any series of applications. The OPUS environment allows the running of multiple instances of multiple processes in multiple pipelines on multiple nodes. It is ideally suited for GONG++ processing, particularly local helioseismology reduction which typically occurs on hundreds of small areas on the disk.



The OPUS Process Manager Screen. This shows an operational ring diagram fitting pipeline composed of seven steps while it is processing the 189 areas of a dense pack. In this snapshot, two of the processes are idle, while the remaining five are working on five different files in the pack.



The OPUS Observation Manager screen. This shows a portion of the list of the 189 files/regions in the dense pack. On the right are seven columns, one for each of the seven steps in the pipeline, showing the status of the file as it moves through the steps. In these columns, the letter c indicates completion, p indicates processing, and w shows a file waiting to begin the step. Here, 3 of the 189 files have completed the entire processing chain.



An example dense-pack flow map from a single day of merged GONG+ images, produced through the ring fitting pipeline.

Mini-Pipeline Experience

We found that the construction of a mini-pipeline to produce flow maps from tracked remapped data cubes was easy to implement. Since OPUS only requires that the code be executed via a Unix shell script, it was simple to link together analysis modules written in IRAF, Fortran 90, Fortran77, and C. The pipeline makes full use of all available processors that are assigned to the task without the need to parallelize the code. The next major task will be the integration of OPUS and the Veritas HSM system.

Acknowledgments

This work utilizes data obtained by the Global Oscillation Network Group (GONG) project, managed by the National Solar Observatory, which is operated by AURA, Inc. under a cooperative agreement with the National Science Foundation. The data were acquired by instruments operated by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísico de Canarias, and Cerro Tololo Interamerican Observatory.