Frequency of Execution Metrics for GBT Observations

Author: Soumya Ghosh
Advised By: Nicole Radziwill

Abstract

In this report we present our experiences with developing a scheduling block based prototype system for lending objectivity to the GBT downtime, prioritization of fixes and optimizing GBT scientific return.

1 INTRODUCTION

The Green Bank Telescope (GBT) is the world's largest fully steerable radio telescope. It is operated by the National Radio Astronomy Observatory (NRAO) at Green Bank, WV. Operating such a sophisticated ensemble of devices (which together form the telescope) is an expensive affair. Optimizing scientific return from the telescope thus becomes a critical issue. This involves recovering from failures quickly and understanding the quality of the observations being made. We could further equate optimizing scientific return with **operational availability**, provided we define operational availability as the observer being able to satisfy their intent (**what** they need) at the time the observation is executed (**when** they need it).

In order to increase operational availability, we need to record, characterize and draw inferences from three relevant parameters namely Instrumentation Health, Raw Data Quality and Derived data Quality.

a) **Instrumentation Health –** This refers to the health of the devices involved in a particular observation. It includes both the control system and the observation system (the antenna and its supporting hardware, managers, and application components).
b) **Raw Data Quality -** Raw data refers to the data that is collected from GBT observations before it is pre-processed, calibrated, or reduced. Raw data quality is essentially a preliminary measure, which indicates whether the data is fit for further processing.
c) **Derived Data Quality -** After raw data is collected, it is then evaluated by the astronomer who may detect other quality issues which would not be readily apparent in the raw data. At this time, the resultant data is interpreted to see that it matches scientific intent. Derived data quality measures the degree to which the derived data matches the scientific intent.

Over the summer we worked on recording and characterizing Instrumrentation(device) health data and assesing **downtime**(the time when the system was not operationally available). Researchers at NRAO are currently following this up by the more complex problems of characterizing raw and derived data quality parameters. In the years to follow, a complete system for quality management for the GBT would be developed.

2 RELEVANT WORK and RESULTS

Our main objective was to create a repository of instrumentation health parameters, which would provide objective decision support to the assessment of device downtimes. It is hoped that this would result in effective prioritization of fixes.

We approached the problem of recording relevant device (instrumentation) parameters from a scheduling block (SB) point of view. This makes sense since the GBT has transitioned to a SB based system and soon all observations done on the GBT would have to be SB based. We identified and extracted SB metadata to keep track of what was observed and how well it was observed.
The control system of the GBT can be modeled as a finite state machine. By tapping this state information and cross correlating it with the scheduling block metadata we were successfully able to measure some very interesting device parameters. Of note amongst these are : **per scheduling block downtime and execution time, per backend downtime and attempted use time, per receiver downtime and attempted use time.** Interestingly we were also able to measure various usage statistics. We now know the number of times a particular scan type, a particular observation type was executed in a monitored time frame. We also have statistics about individual scheduling blocks, the total number of times a SB was executed, number of successful executions, number of failures due to user aborts, number of failures due to system screw ups.

A quality database was developed to capture these parameters. The database is a MySql database and currently has ten tables. The database is currently populated by mock data from the GBT simulator. We expect the database to be populated with real data as soon as the metadata mining scripts are incorporated in the GBT system as daemons.

We have also developed scripts to glean the quality database for the most relevant parameters (user defined) and produce a report. We envision this report being used by the planning committee in making critical instrumentation related decisions.

A screenshot of the captured parameters outputted as a text file has been attached at the end of this report. (This is a raw test version, a spiffed up, wiki enabled version would soon be in place)

3 CONCLUSIONS

We took the first steps towards developing an integrated quality management system for the GBT. Our system is currently creating a repository of instrumentation parameters. It is hoped that this would prove effective in providing critical decision support and allocating scarce development resources. Further work needs to be done on capturing certain device parameters and on drawing inferences from the wealth of statistical information being captured.

```
Start Time 2005_08_04_16:01:16
End Time 2005_08_05_10:51:53

    total idle time:          32.285505
    total active time:        8.864454
    total latency time:       1.714928

SCHEDULING BLOCK STATS

Number of Scheduling Blocks executed in this time period : 10

Stats for SB 127
Number of successful executions : 0
Number of aborted executions : 1
Number of hung up executions : 0
Total Number of executions : 1
Stats for SB 17
Number of successful executions : 2
Number of aborted executions : 3
Number of hung up executions : 0
Total Number of executions : 5
Stats for SB 177
Number of successful executions : 0
Number of aborted executions : 1
Number of hung up executions : 0
Total Number of executions : 1
Stats for SB 53
Number of successful executions : 0
Number of aborted executions : 1
Number of hung up executions : 0
Total Number of executions : 1
Stats for SB 48
Number of successful executions : 0
Number of aborted executions : 1
Number of hung up executions : 0
Total Number of executions : 1
```

Fig: A screenshot of parameters mined for the automated observing report.