

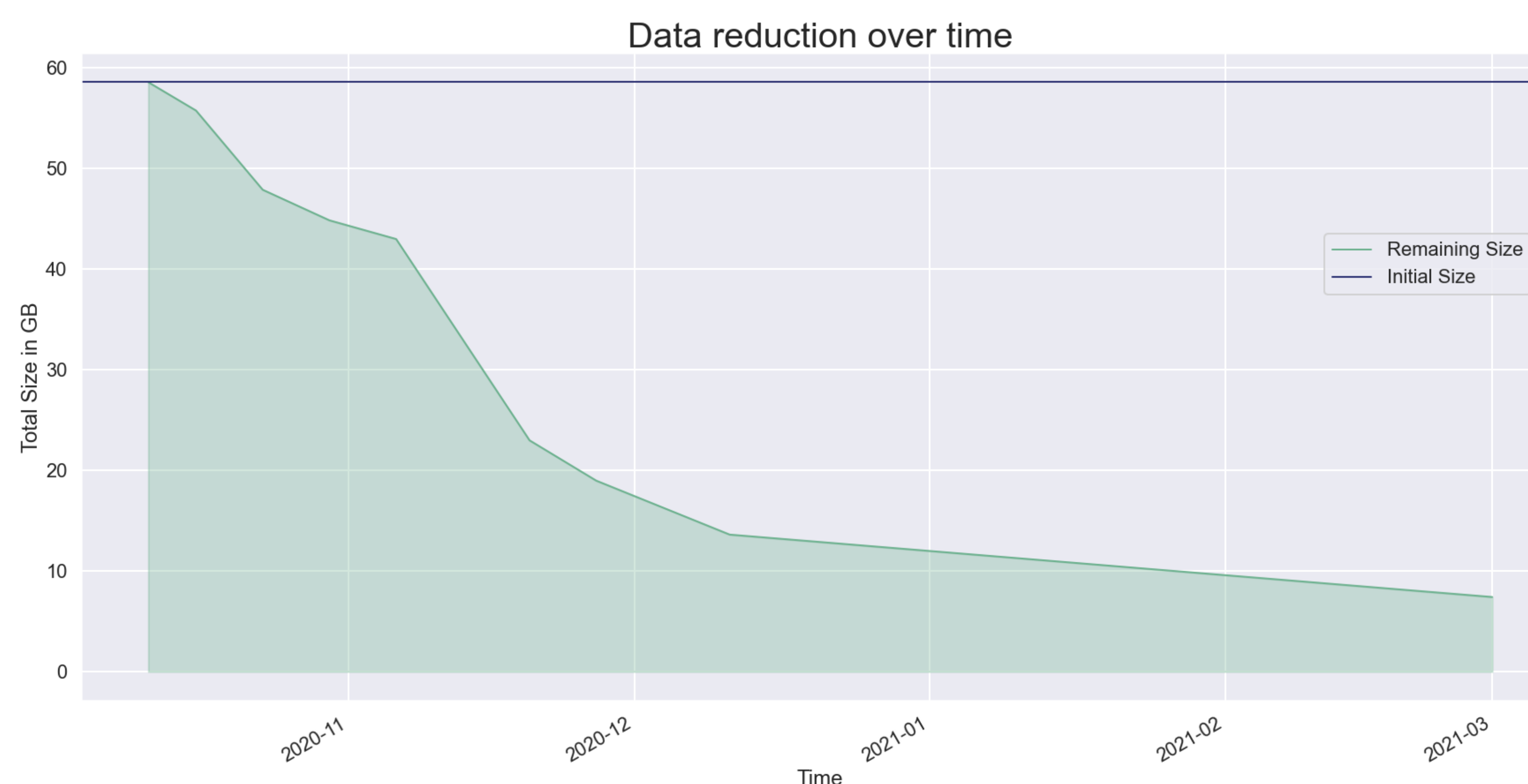
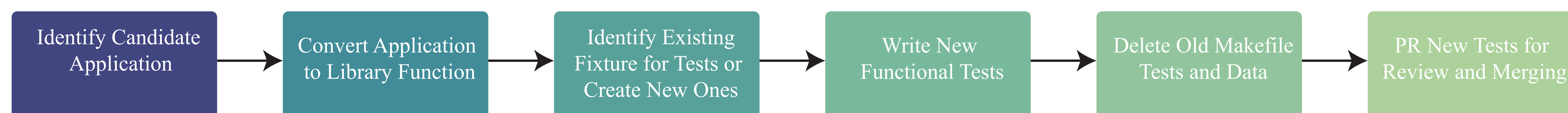
ISIS Test Data Reduction

K. Rodriguez, K. D. Lee, A. C. Paquette, A. R. Sanders and J. R. Laura, Astrogeology Science Center, 2255 N. Gemini Drive, Flagstaff, AZ, 86001 ; krodriguez@usgs.gov.

Overview

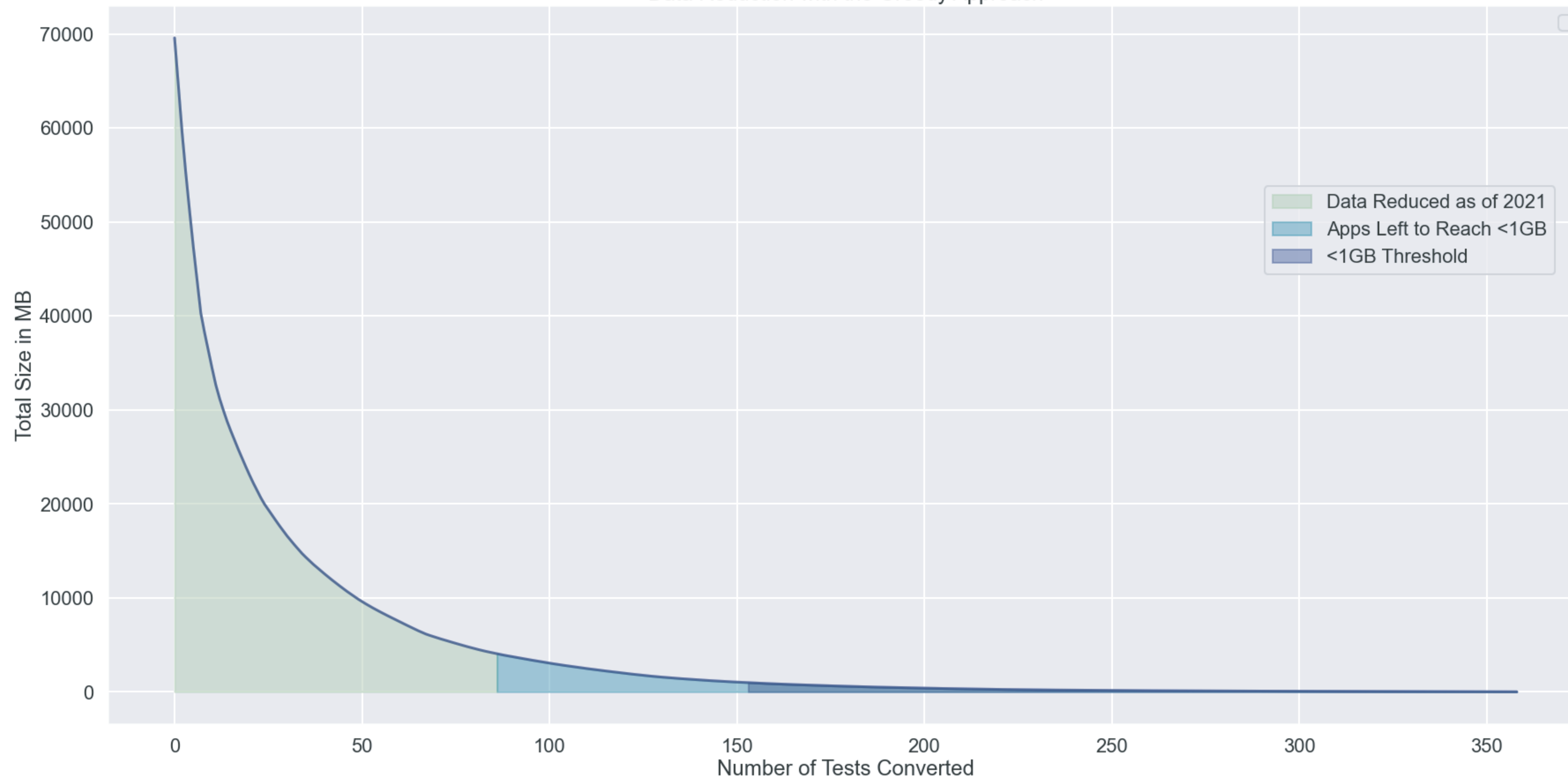
- The Integrated System for Imagers and Spectrometers (ISIS) test data area is a necessary aspect of running ISIS tests locally.
- The required data volume of ~72GB is a burden to users who wish to contribute to the ISIS code base and internal developers who maintain the tests. This is because the Makefile-based tests contain large quantities of redundant data.
- We made an effort to begin reducing the data burden by moving to Google's GTest suite for application tests, which make up a majority of the ISIS test data. By using representative, reusable data, for these tests, we can dramatically reduce the ISIS test data size.
- This is a continuing effort at USGS Astro to completely remove all tests still using the old paradigm of Makefile based tests and move remaining data into the main ISIS repository.

Application Conversion Process



Converting ISIS applications is a multi-step process that largely has been streamlined by the software team. The process requires converting ISIS applications to a callable on top of deleting the old data. This means exposing ISIS applications as part of the libisis API and opening the door to potentially expose this functionality to other languages. As the data reduction rate is slowing to a linear pace, getting the test data size to zero is a long-term goal for USGS Astro.

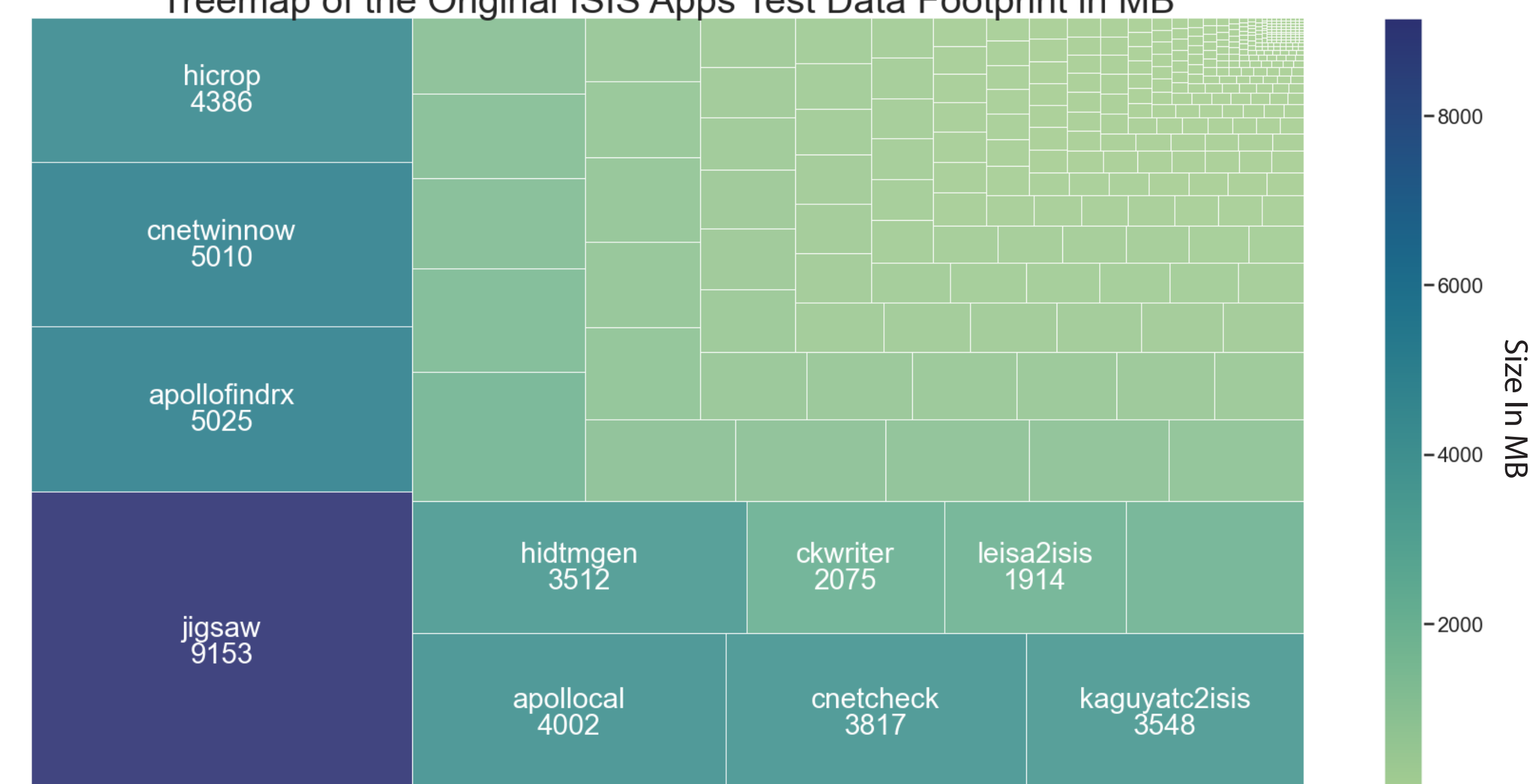
Data Reduction with the Greedy Approach



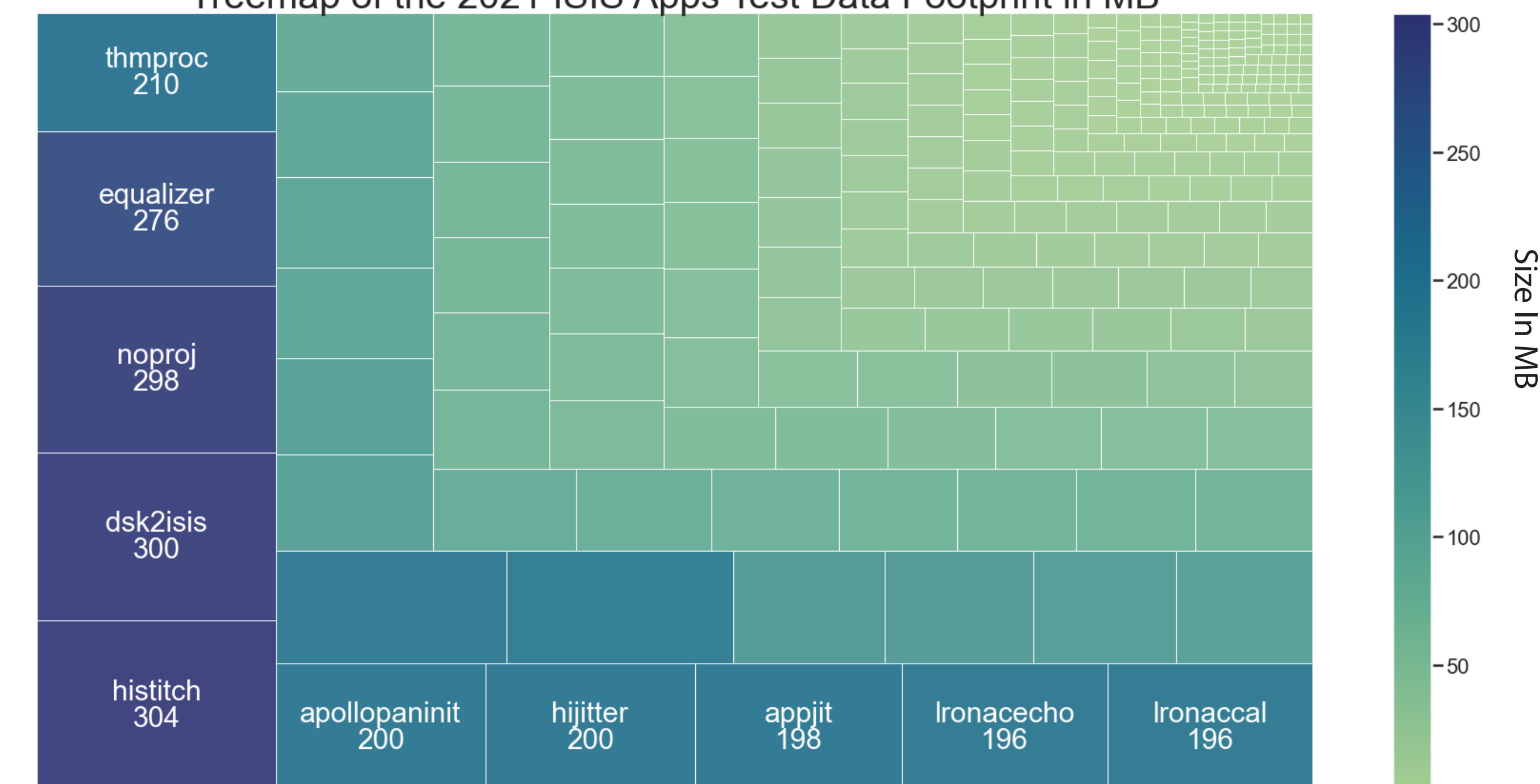
PROGRESS AS OF 2021

Here is a line plot depicting the exponential decrease in ISIS's test data volume as each of its application's test data is removed after the tests are converted to GTests with fixtures using a greedy approach (higher impact applications converted first). The green area indicates data reduction from the 86 applications that have been converted as of 05/01/2021. The dark blue area indicates the remaining 153 apps with small data volume contributions that can be moved into source control as their total contribution is <1GB. The light blue area in-between indicates the apps left to meet that threshold. We are approaching a point of diminishing returns, such that in order to reduce the data area by an additional ~6GBs requires the translation of 64 more ISIS applications. Compared to the current 86 applications already converted that contributed ~70GBs to the ISIS test data volume.

Treemap of the Original ISIS Apps Test Data Footprint in MB



Treemap of the 2021 ISIS Apps Test Data Footprint in MB



ISIS Test Data Reduction Visualized

Treemaps of the ISIS data area and how much each applications contributes to the total data volume for the ISIS test data area in MB. Very few applications contribute a disproportionate amount, with jigsaw comprising ~9GBs of the total ~72GBs of test data.