

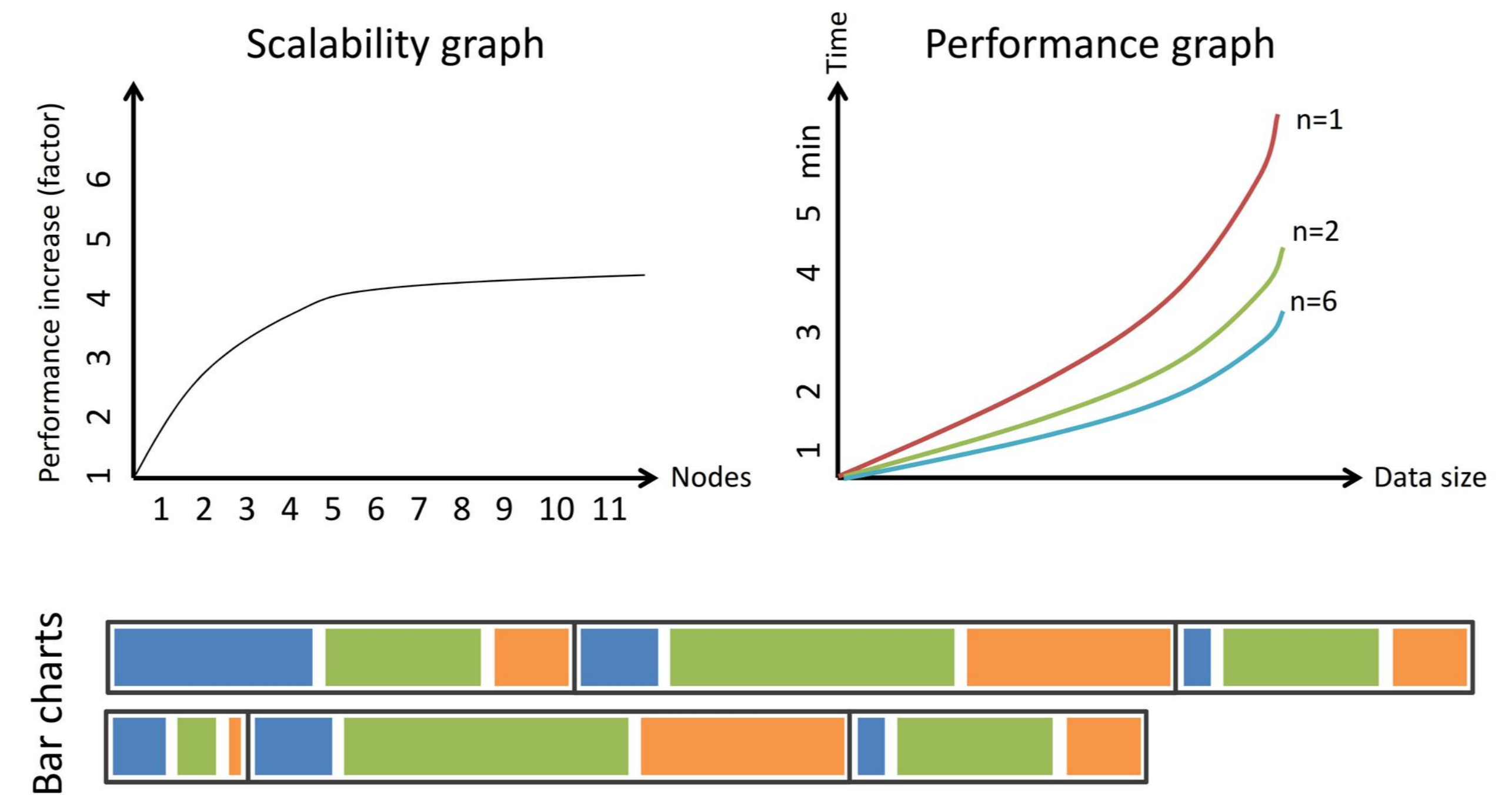
Introduction

IQmulus is a knowledge-driven processing system for large geographic data that produces several software and infrastructure components that require quantitative performance testing, thereby measuring and monitoring the progress toward meeting the project goals. The scalability tests are designed to assess the **innovative processing services**, their composition in **logical workflows**, the **cloud based infrastructure** and the **visualization clients**. A high degree of automation was aspired in the implementation of the test infrastructure so to facilitate the repetitive testing of services and entire workflows with minimal human interaction. The focus of this presentation is on the presentation of the testing infrastructure and the assessment of processing services.

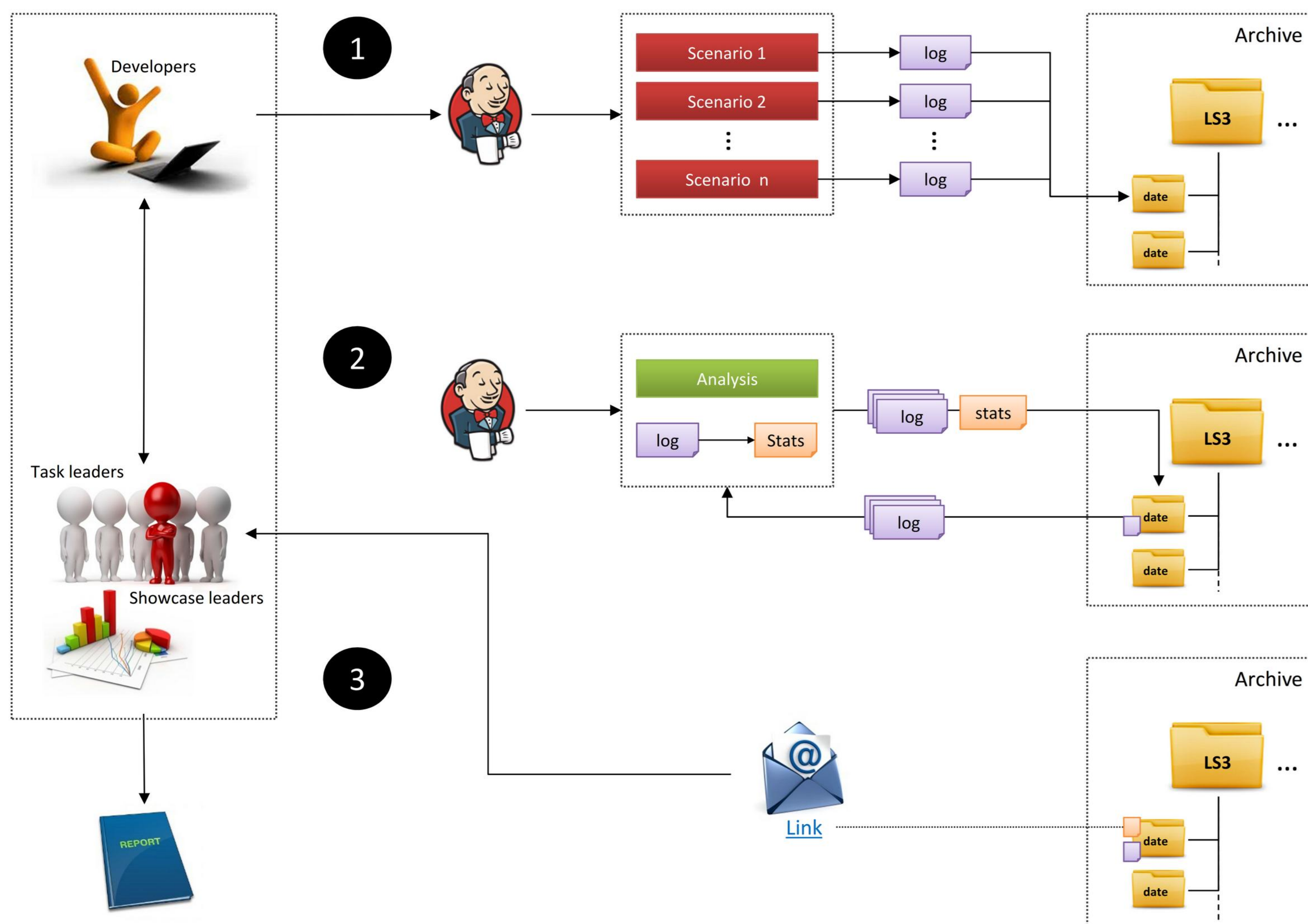
Information products

For the sake of the analysis, each service is divided up into reading, processing and writing stages. This subdivisions allow not only the comparison of total execution times but present a more detailed and informative picture of a given service's behavior in a set of scenario of data size and computational power. Bar charts are produced to quantify the difference between two or more scenarios in terms of absolute and relative execution times with the reading, processing and writing stages (marked blue, green and orange, respectively, here). The results highlight the performance delta between the reading, processing and writing stages with varying data size and computational power and thus identify bottlenecks.

Similarly, entire workflows (i.e. a chain of logically connected services) are tested and evaluated. The comparison of stage execution times produces information on the scalability barrier in terms of parallelized service execution. This effect is also known as Amdahl's law and states that the maximum performance improvement is achieved when scalable components are reduced to a negligible minimum and the execution times are determined solely by the sequential components of the test. No increase in parallelization will result in better performance results from that barrier on. The approximation to the Amdahl boundary is visualized using the scalability graph to the right below that charts a performance increase factor against the parallelization (i.e. the node count). A levelling of the chart line implies that the boundary has been met. Finally, charting the overall test duration against data size for all computational scenarios illustrates the linear or non-linear performance behavior of the processing services tested.



Infrastructure



All services and workflows are tested in the IQmulus project cloud environment and, where feasible, in other cloud and cluster environments. Every possible combination of a predefined set of input data sizes and computational configuration is tested. Data sizes are selected individually for each service or workflow. Computational scenarios themselves are a set of nodes, CPU cores and available RAM.

The testing infrastructure was implemented using Jenkins. Every processing service and workflow comes with a defined test plan that specifies which datasets are to be tested against which computing configurations. They are triggered according to schedule, upon service updates or manually. Standardized project specific logging of all services and system control components ensure all necessary information such as timestamps, IDs, system performance, data sizes, progress, etc. is recorded continuously and stored in a file-based archive for analysis upon the completion of the scalability test.

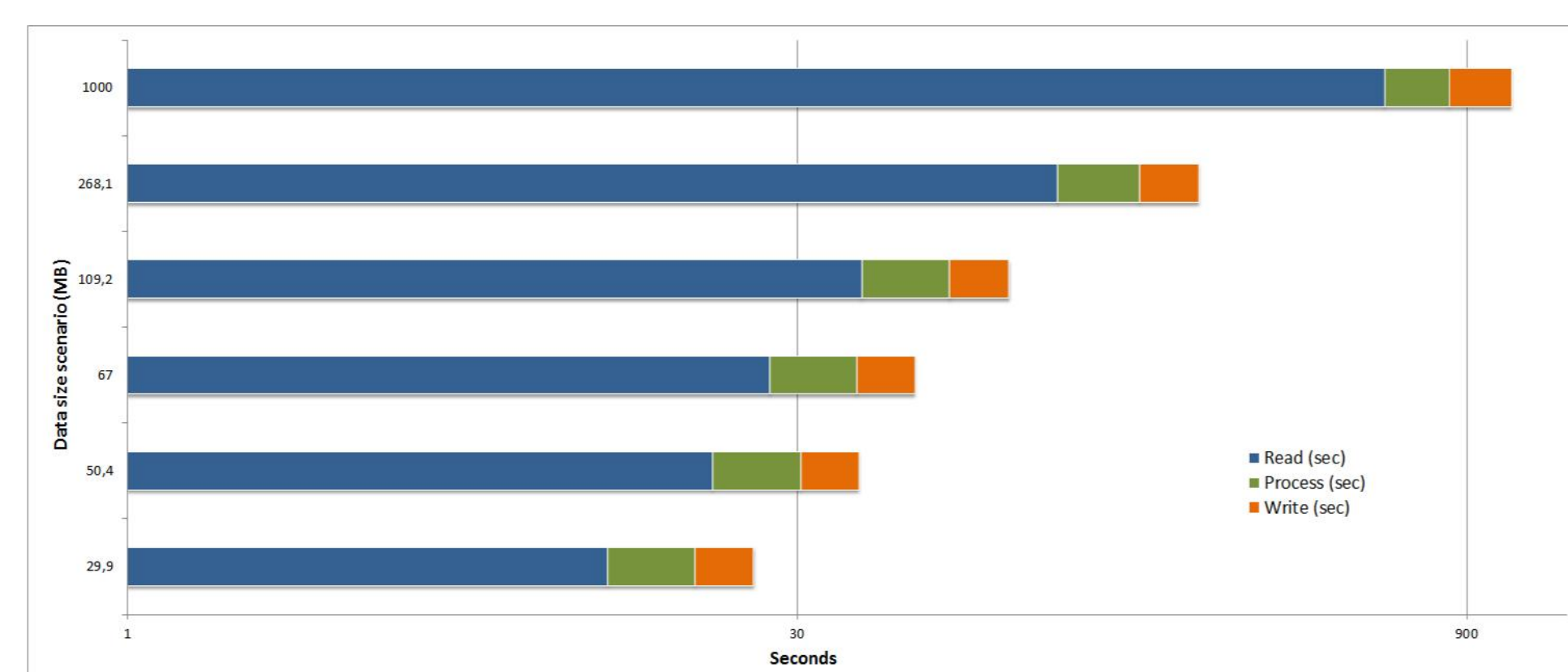
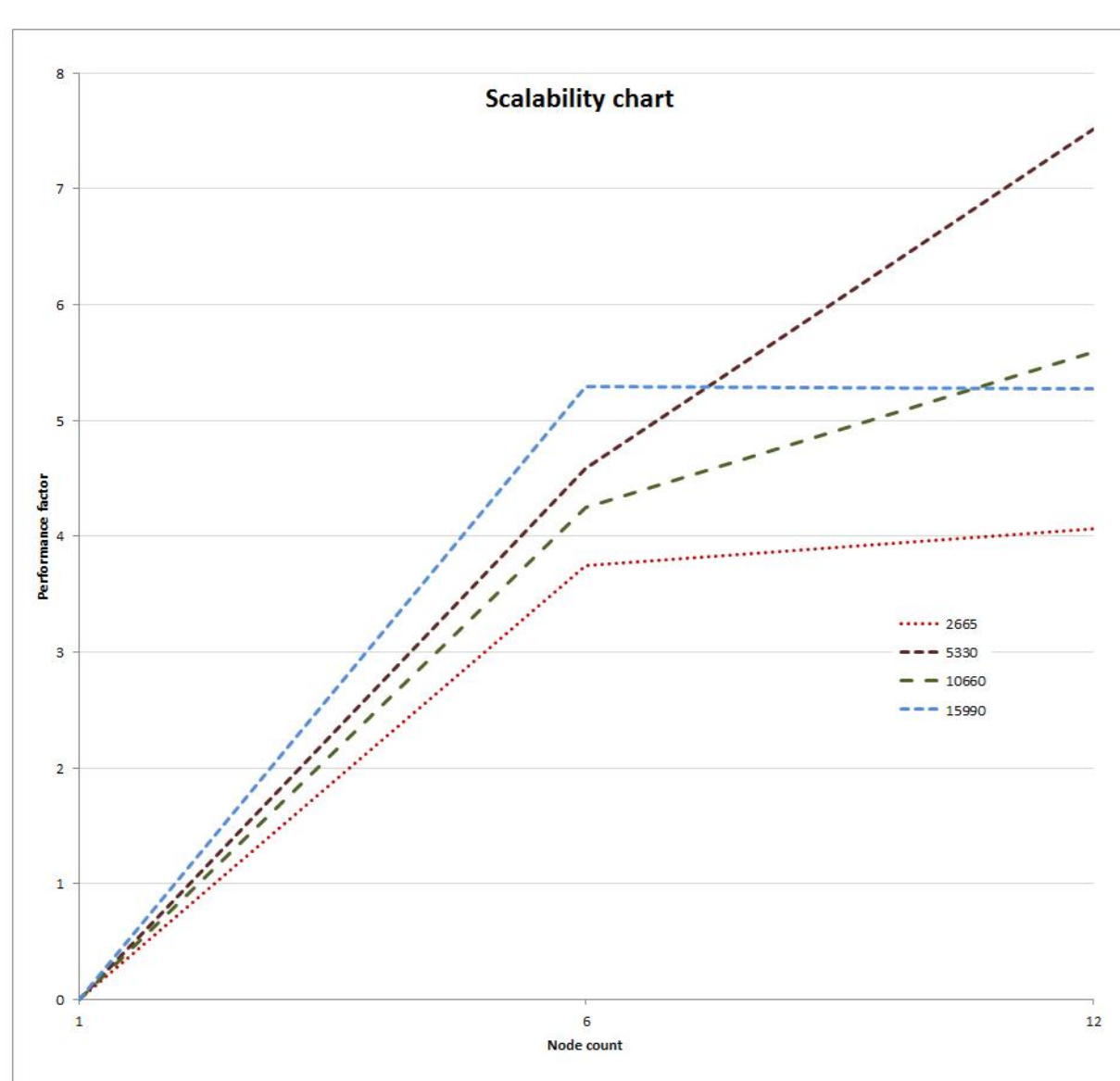
Once the test has concluded, Jenkins executes the IQmulus Analytics script that parses all log files stored in the test archive and produces a standardized and exhaustive set of statistics and charts. The results are added to the archive and made available for project internal use. All logs and results of previous runs are preserved to allow for service development reviewing.

Finally, the results of service or workflow tests are made available to the project staff concerned with development oversight as well as the respective developers themselves. The analysis results are reviewed and interpreted, resulting in corrective actions and reruns of the scalability test. The infrastructure thus facilitates the smooth and effective circulation of service updates and near time performance results.

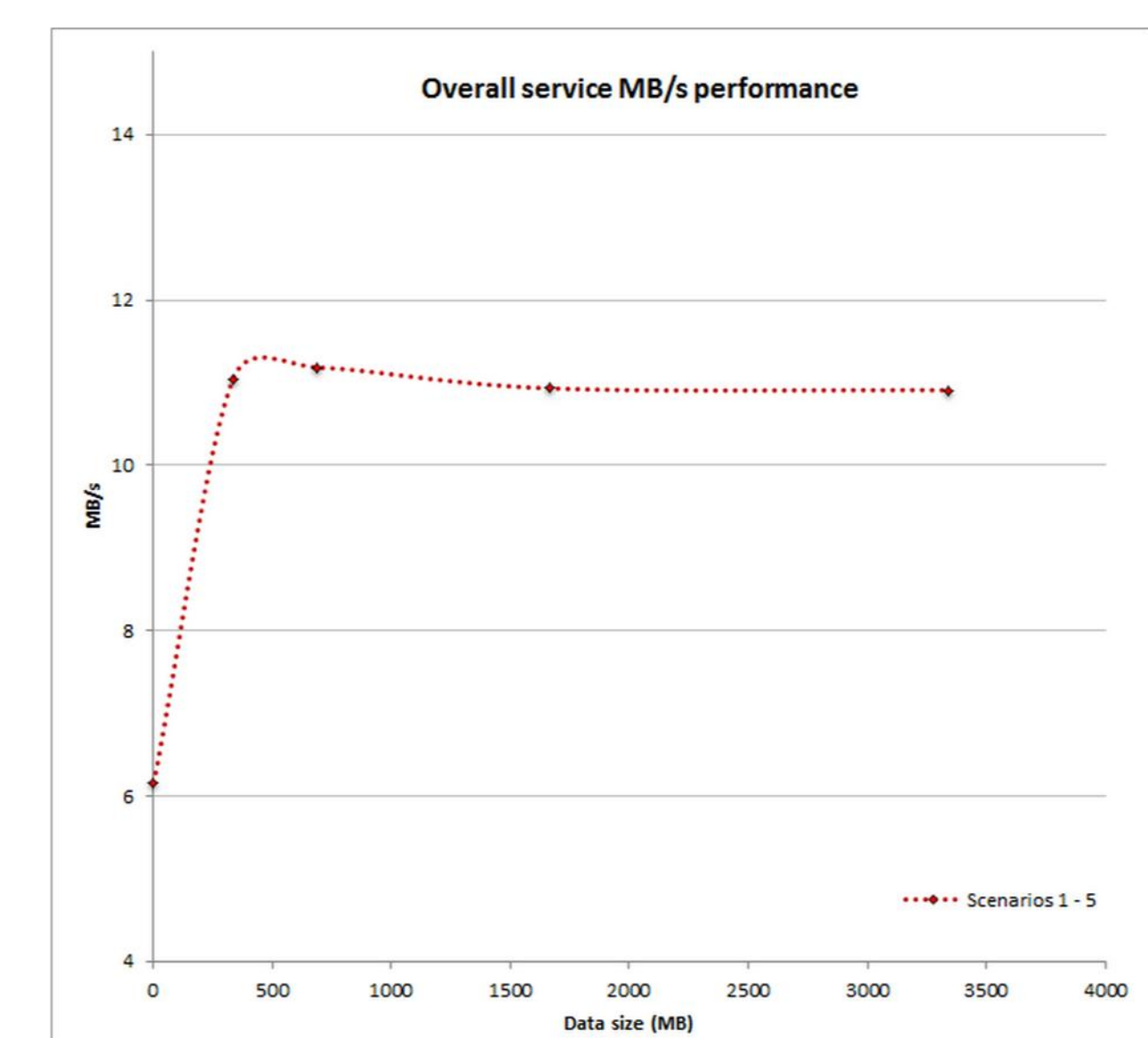
Initial results

Initial results are currently limited to a number of available services that primarily focus on the processing of laser point clouds. The following results represent the initial test outcomes for a small subset of the IQmulus service portfolio and are intended to provide an insight into the progression of the IQmulus project.

Service 68 (Tree crown recognition): The service recognizes tree crowns from mobile mapping point clouds. The scalability graph implies that this service experiences significant performance improvements through parallelization and the Amdahl boundary appears not to have been reached at 12 nodes.



Service 80 (Water logging detection): Testing service 80 with varying data sizes on a single node results in the commonly observed exponential increase of processing time. Note that the time scale on the x-axis is logarithmic. Note also that reading times increase at a significantly greater rate than those of the processing and writing stages do.



Service 35 (Resampling of point clouds): The service recognizes tree crowns from mobile mapping point clouds. The performance graph demonstrates the constant processing performance across increasingly larger data sizes. The common decrease in processing speed is not evident across the tested data sizes.