ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-THPDP016

FULL STACK PERFORMANCE OPTIMIZATIONS FOR FAIR OPERATION

A. Schaller, H. Hüther, R. Müller, A. Walter
GSI Helmholtzzentrum für Schwerionenforschung, Darmstadt, Germany

Abstract

In the last beam times, operations reported a lack of performance and long waiting times when performing simple changes of the machines' settings. To ensure performant operation of the future Facility for Antiproton and Ion Research (FAIR), the "Task Force Performance" (TFP) was formed in mid-2020, which aimed at optimizing all involved Control System components. Baseline measurements were recorded for different scenarios to compare and evaluate the steps taken by the TFP. These measurements contained data from all underlying systems, from hardware device data supply over network traffic up to user interface applications. Individual groups searched, detected and fixed performance bottlenecks in their components of the Control System stack, and the interfaces between these individual components were inspected as well. The findings are presented here.

INTRODUCTION

At GSI, the Control System is currently being modernized, also to support the future FAIR accelerators. The new Control System layout is described in Fig. 1.

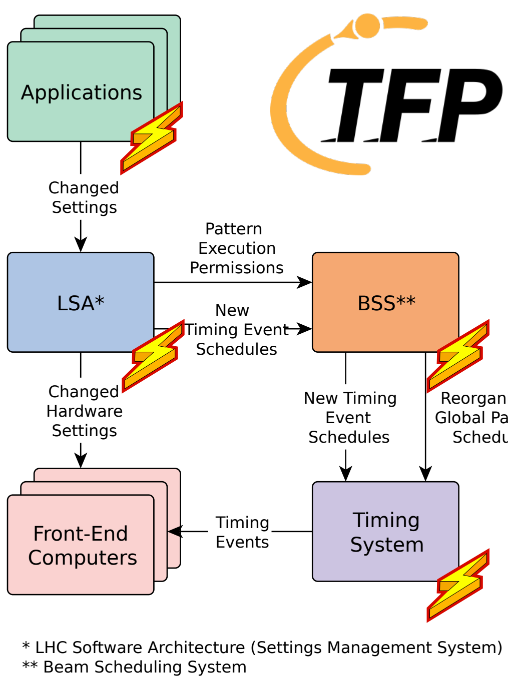

Figure 1: Control System Stack at GSI and FAIR. The parts marked with a lightning bolt were optimized by the Task Force Performance.

The LHC Software Architecture (LSA) [1] internally uses a hierarchical description of physics parameters to hardware parameters [2]. Processing the parameter hierarchy has already been sped up when the first larger contexts were set up at GSI in 2017, for example by parallelizing the calculations done in LSA [3]. After the parallelized calculation was used in production, focus changed to implement new features like the Storage Ring Mode [4]. Since most necessary technical features are available for productive usage now, the Controls department got the mandate by management to prioritize performance improvement in 2020.

ESTABLISHMENT OF THE TASK FORCE PERFORMANCE

The Control System stack involved in settings and timing changes of the machine should be sped up after operations reported poor performance and severely reduced usability due to long waiting times for trims (setting changes). Therefore, the Task Force Performance (TFP) was established with members from different departments and groups [5]. The project lead was taken by the chief architect of the Controls department, while the individual Controls groups delegated different developers to fully work for the TFP. Operations, as well as the different machine departments, delegated specialists to support the TFP in implementing the various scenarios like machine setup or beam manipulation.

TIME FRAME AND MILESTONES

The first action was to plan a time frame for when the TFP could take its baseline measurements, and for when the performance optimizations should be done to perform the next beam time without interruptions. The kick-off meeting took place at end of April, while the milestone "Ready for Integration" was planned for the end of October, and "Ready for Production" was scheduled for the end of November 2020. During the first weeks, a measurement concept was established in the Control System software stack to measure the duration of its different functions.

KICK-OFF

Operators and machine experts provided different scenarios, from simple setting changes to complex machine setup procedures, that should be optimized by the TFP. For these scenarios, baseline measurements under real production conditions were taken in June 2020, directly after the beam time 2019/2020 had ended.

Goals

The main goal for the TFP was to substantially improve performance for the most relevant use cases, so the facility can be operated efficiently. Identifying potential performance issues that may become relevant for FAIR (i.e. the new synchrotron SIS100 and beyond) was the secondary objective.

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-THPDP016

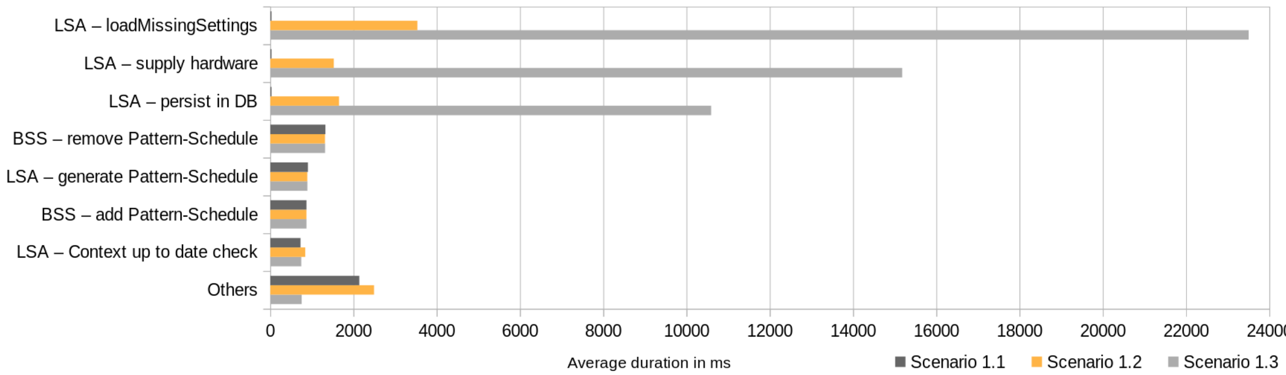

Figure 2: SIS18 Baseline Measurements (before optimizations).

Scenario-Based Approach

A naming convention *Scenario x.y*, where *x* represents the type of modification and *y* the grade of complexity, was set up. Three different scenarios were defined:

- Scenario 1.y: Regular Trim by Operations
- Scenario 2.y: Trim by Beam-Based Feedback
- Scenario 3.y: Storage Ring Manipulation by Operations

While implementing the measurements for each scenario, it quickly became apparent that *Scenario 1* and *Scenario 2* triggered the same operations in the Control System stack. The differentiation between manual and automatic trims was therefore dropped, and *Scenario 2* was not explicitly investigated further. For each scenario, multiple levels of complexity were considered.

Scenario x.1 represents a "minimal hierarchy" trim which results in changes to a single kicker device. This mainly affects data supply to hardware and data supply to the timing system via the Beam Scheduling System (BSS) [6]. This minimal hierarchy trim leads to hardly any calculations in LSA. Operations emphasized that this scenario is a basic task done multiple times during daily operations and should not take a noticeable amount of time.

Scenario x.2 represents a "medium hierarchy" trim. This is realized by a change of a physics parameter that triggers calculations for multiple hardware devices via the hierarchy [1]. In addition to *Scenario x.1*, this also contains more complex calculations done in LSA, and more data to be supplied.

Scenario x.3 is the most complex scenario, the "large hierarchy" trim, which was realized by changing the target energy value. This caused settings for the entire context to be recalculated, leading to large-scale calculations done in LSA that affect nearly the entire hierarchy, and results in a substantial re-supply of timing and hardware devices.

Measurement Setup

All measurements were performed by accessing LSA directly to trim settings, keeping track of the execution time of specific sub-tasks as well as the total execution time for the calculation and subsequent data supply to underlying systems and hardware. In order to reduce outliers and make measurements comparable, each measurement was repeated five times for warming up the systems and machines, and then five more times for recording the actual measurements. From these recordings, the minimal, maximal, average and median execution time were used as performance indicators. This setup could then be repeatedly executed to measure individual changes done by the TFP and compare the results with the baseline to be able to immediately judge the effect of a change made. To be able to start as soon as possible with the optimization works, *Scenario 1* was first implemented only at the synchrotron SIS18, see Fig. 2, and not yet at the storage ring ESR. The measurements of the baseline and of the scenario with its complexity levels were then analyzed by the TFP members. Different actions were derived for the individual groups.

ITERATION 1: READY FOR INTEGRATION

As seen in Fig. 2, there are time-consuming parts that grow with the hierarchy complexity, e.g. *LSA - load Missing Settings*, but also constant parts like *BSS - add Pattern-Schedule*. The TFP members decided to look at the constant parts first, using *Scenario 1.1* to quickly see the effects of changes. Afterwards, *Scenario 1.2* was used to optimize the hierarchy-dependent parts. The total execution times after Iteration 1 are compared to baseline in Fig. 3. Performance for these scenarios could be sped up by a factor of about 7.1 (*Scenario 1.1*), 5.2 (*Scenario 1.2*) and 2.7 (*Scenario 1.3*).

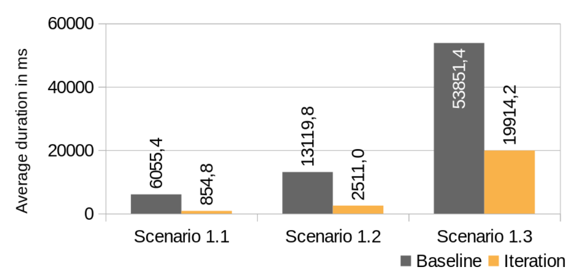

Figure 3: SIS18 - Baseline compared to Iteration 1 (total execution time).

During this phase, the *Scenarios 1.y* were also implemented at ESR. Even though there is no baseline measurement for *Scenarios 1.y* at ESR and therefore no direct com-

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-THPDP016

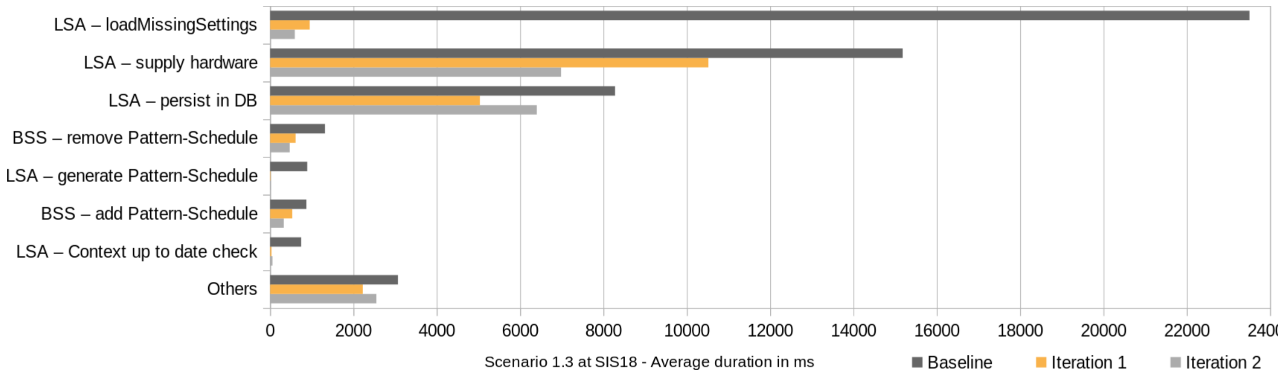

Figure 4: Average execution times for *Scenario 1.3* at SIS18.

parison of execution times, there are many indicators that ESR also benefited from this iteration's optimizations.

Additionally, the storage ring-specific *Scenarios 3.1* and *3.2* were set up at ESR as well. These new *Scenarios 3.y* covered the setting changes via online beam manipulation [4]. The manipulation feature is used to influence the beam while it is stored, whereas regular trims are performed between beam executions. As manipulations are limited to non-timing settings, *Scenario 3.3* is not possible.

ITERATION 2: READY FOR PRODUCTION

During the second iteration, the most time-consuming tasks done in LSA and underlying systems were further inspected and optimized.

The optimization results for those tasks in the context of *Scenario 1.3* over the iterations are shown in Fig. 4. The loss of performance on the "LSA - persist in DB" task between iterations 1 and 2 can be explained by the increased amount of data stored in the database, because the search overhead grows as well. Unfortunately, the database performance itself could not be optimized due to personnel availability. Therefore, the persistence speedup was mainly achieved by minimizing the amount of data that gets stored.

The total execution times for all complexity levels of *Scenario 1* over the iterations are visualized in Fig. 5.

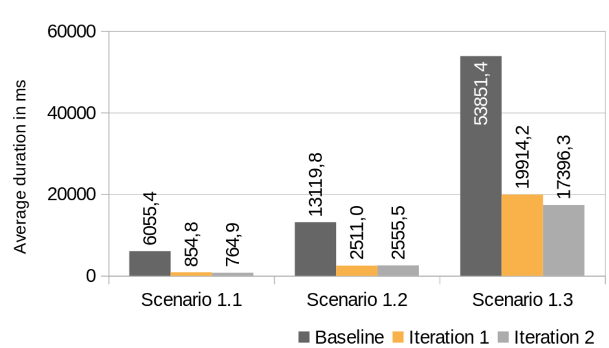

Figure 5: SIS18 - Baseline compared to Iteration 1 and 2 (total execution time).

Additionally, technical limitations have been reduced. Before this iteration, it was only possible to change the settings for one beam at a time. Setting changes for other beams were queued and processed sequentially, potentially causing long waiting times before even very simple setting changes could be processed. The TFP members managed to allow setting changes for different beams simultaneously. As the operation of multiple different beams in parallel is the most common operation mode at GSI [7], this new feature resulted in a huge performance gain.

Another technical limitation was to always stop the machine's beam execution while changed settings were supplied to the hardware. This was done as a safety measure to make sure settings were consistent within a beam's lifespan. However, there were some cases in which stopping beam execution was actually unnecessary, because accepting the newly calculated settings instantaneously would not be harmful to the beam. This limitation can now be circumvented with the so called "Bypass Trim" that does not stop the machine's execution but directly sets new settings to the devices. The "Bypass Trim" is limited to scalar settings that do not affect the timing execution behavior of any component in the Control System stack or devices. This feature can be explicitly enabled by operators when they deem it safe to use. The speedup for a "Bypass Trim" can be seen in Fig. 6.

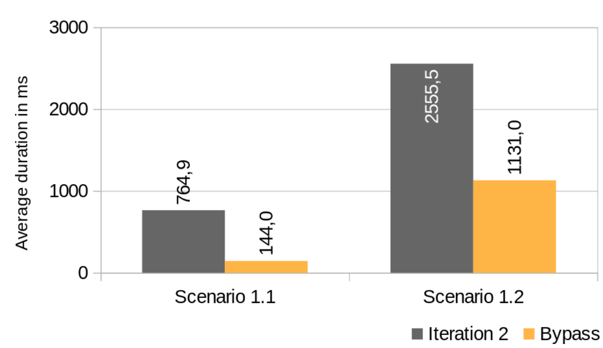

Figure 6: SIS18 - Iteration 2 compared to Bypass Trim (total execution time).

In addition to the described optimizations, TFP members found that further efforts on the level of LSA and the underlying systems did not make sense from a cost-benefit

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-THPDP016

point of view, as the expected higher efforts in refactoring and architectural changes would not be worth the expected speedup. So there were further optimization possibilities identified, but the estimated effort was too high to the low expected performance gain.

Instead, the applications used to change the machines' settings came into focus. They had been considered to be of lower priority in comparison to the underlying systems. Since all setting changes by an application have to pass LSA and the underlying systems this has been considered as the bottleneck and was analyzed first. Now, the main application used to change the machines' settings was inspected in more detail.

As shown in Fig. 7, it turned out that the application indeed played a big part of the total execution time as perceived by the operators, even after the other parts of the Control System stack had been optimized.

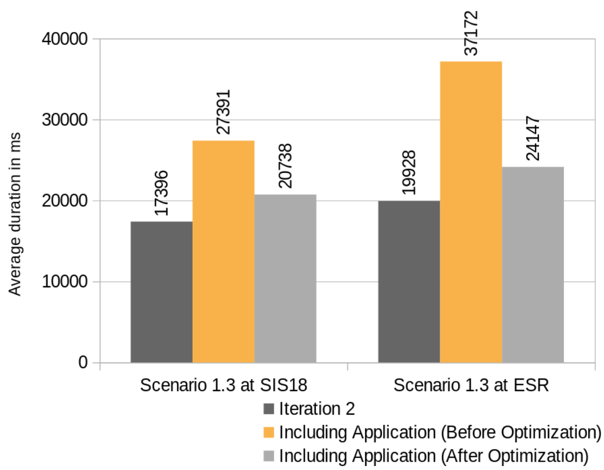

Figure 7: *Scenario 1.3* - Iteration 2 combined with application before and after optimization (total execution time).

The time of about 10 s was mainly spent in reading the updated settings back from LSA, and displaying them in a preprocessed form. Some tasks were done multiple times, or much more data was loaded than necessary, because the app uses generic widgets that update themselves and notify others. By adding some more logic, some duplicated task executions and unnecessary data retrievals could be removed. Also, the creation of different tabs for displaying groups of related settings could be done in parallel, instead of sequentially as before. Another performance boost was to hide some tabs that are rarely used, so that they are only processed when they are actually displayed.

COLLABORATION CONSTRAINTS

Since LSA was initially implemented at CERN and is now developed in collaboration with GSI, any changes made on the common code base had to be verified and accepted by CERN. CERN found several cases where an improvement made at GSI led to a performance degradation at CERN since the system is set up and used in a slightly different way on both institutes. To find a solution satisfying the needs of both institutes, in several cases multiple implementations were tested and compared. In other cases the code base was made extendable to support institute specific implementations.

IMPLEMENTED OPTIMIZATIONS

During both iterations, TFP members implemented different optimizations in their Control System components. The most effective optimizations were to reduce the memory consumption, so that there are less garbage collections and less data to search. Also reducing I/O operations, e.g. database accesses, had a noticeable performance boost. Adding parallelization where possible as well as the usage of more performant data structures are some more advanced optimizations that were implemented especially in the application and LSA. The most invasive optimization was to refactor the API between Control System components. This reduced the amount of data needed to be serialized, sent over network and deserialized again. As an optimization with some drawbacks, caches were implemented on several places. Mostly used for immutable data, caches can further reduce I/O and computationally intensive operations. For mutable data, caches were rarely used because of their drawbacks such as obsolescence.

FINAL RESULT

The primary goal to improve performance for currently most relevant use cases has been achieved.

The average speedup for *Scenario 1.1* for example is about a factor of 7.9, or for a "Bypass-Trim" 42, which makes the Control System much more responsive for the operators. For the most complex *Scenario 1.3*, the average speedup has a factor of 3.1. "Bypass-Trims" are not possible in this scenario, since it always contains changes for timing-relevant settings. The time spent in the application directly before and after each trim has also been reduced significantly by a factor of about 4.

Positive feedback from the operators has been reported after the next beam time in early 2021.

By implementing simultaneous trims for parallel beams and "Bypass Trims", the secondary goal can also be seen as fulfilled. No further technical issues or architectural problems encountered, so operating for FAIR is considered possible without the necessity for further such a big performance optimization.

OUTLOOK - ONGOING INVESTIGATION

Maintaining performance will be an ongoing activity to keep the Control System as performant as possible. The test setup for the TFP measurements is still available and can be used to verify that new developments do not negatively influence the Control System's performance. For LSA, a dedicated performance test suite has been created that focuses on specific LSA-internal methods.

New features as well as additional machines, e.g. UNILAC, which is currently still operated with the old legacy Control System or the new SIS100, will introduce new chal-

ISBN: 978-3-95450-238-7 ISSN: 2226-0358 doi:10.18429/JACoW-ICALEPCS2023-THPDP016

lenges. Ideally, they will be implemented with respect to performance from the beginning.

In summary, it can be said that cross-departmental collaboration was necessary and really beneficial at this point. Working exclusively on this topic over such a long period of time has created awareness among everyone, so that performance will always be in mind when new features are developed.

REFERENCES

[1] D. Jacquet, R. Gorbonosov, and G. Kruk, "LSA - the High Level Application Software of the LHC - and Its Performance During the First Three Years of Operation", in *Proc. ICALEPCS'13*, San Francisco, CA, USA, Oct. 2013, paper THPPC058, pp. 1201–1204.

[2] D. Ondreka, J. Fitzek, H. Liebermann, and R. Mueller, "Generic Settings Generation for FAIR: First Experience at SIS18", in *Proc. IPAC'15*, Richmond, VA, USA, May 2015, pp. 156–158. doi:10.18429/JACoW-IPAC2015-MOPWA027

[3] A. Schaller, J. Fitzek, D. Lorenz, and F. Wolf, "Optimized Calculation of Timing for Parallel Beam Operation at the FAIR Accelerator Complex", in *Proc. ICALEPCS'17*, Barcelona, Spain, Oct. 2017, pp. 411–415. doi:10.18429/JACoW-ICALEPCS2017-TUPHA019

[4] R. Mueller, J. Fitzek, H. Hüther, H. Liebermann, D. Ondreka, A. Schaller, A. Walter, "Supporting Flexible Runtime Control and Storage Ring Operation with the FAIR Settings Management System", in *Proc. ICALEPCS'21*, Shanghai, China, Oct. 2021, pp. 768–773. doi:10.18429/JACoW-ICALEPCS2021-WEPV047

[5] H. Hüther, J. Fitzek, Dr. O. Geithner, Dr. F. Herfurth, Dr. C. Hessler, Dr. S. Litvinov, Dr. B. Lorentz, S. Krepp, R. Mueller, Dr. D. Ondreka, A. Schaller, Dr. J. Stadlmann, Dr. R. Steinhagen, Dr. M. Steck, A. Walter, "Task Force Performance Project Report", *GSI annual report 2021*, GSI, Darmstadt, Germany, to be published. doi:10.15120/GSI-2021-01005

[6] S. Krepp, J. Fitzek, H. C. Hüther, R. Mueller, A. Schaller, and A. Walter, "A Dynamic Beam Scheduling System for the FAIR Accelerator Facility", in *Proc. ICALEPCS'21*, Shanghai, China, Oct. 2021, pp. 138–142. doi:10.18429/JACoW-ICALEPCS2021-MOPV013

[7] H. C. Hüther, J. Fitzek, R. Mueller, and A. Schaller, "Realization of a Concept for Scheduling Parallel Beams in the Settings Management System for FAIR", in *Proc. ICALEPCS'15*, Melbourne, Australia, Oct. 2015, pp. 434–436. doi:10.18429/JACoW-ICALEPCS2015-MOPGF147