Contribution ID: **92**   Type: **Poster**

Applying Transaction Processing Techniques to Large Scale Analysis

Data analysts working with large datasets require absolute certainty that each file is processed exactly once. ServiceX addresses this challenge by using well established transaction processing architectures. This system implements a fully transactional workflow powered by PostgreSQL and RabbitMQ, ensuring data integrity throughout the processing pipeline. This presentation details both the infrastructure design that enables these transactional guarantees and the techniques we used to identify and remediate intermittent transaction leaks, resulting in a reliable system that operates consistently at scale.

Significance

This talk shows how the application of tools and techniques from the broader computer science field can improve the productivity of physics analysis.

References

Experiment context, if any

Author: GALEWSKY, Benjamin (Univ. Illinois at Urbana Champaign (US))

Co-author: ONYISI, Peter (University of Texas at Austin (US))

Presenter: GALEWSKY, Benjamin (Univ. Illinois at Urbana Champaign (US))

Session Classification: Poster session with coffee break

Track Classification: Track 1: Computing Technology for Physics Research