

Конфигурационно-управляемая система потоковой обработки данных

Данилов П.Ю.¹

Научный руководитель – канд. техн. наук, доцент Логинов И. П.¹

¹Университет ИТМО

p.danilov@itmo.ru

Введение

Потоковая обработка данных является ключевой технологией построения современных распределённых систем реального времени. Фундаментальные модели и архитектурные принципы stream-processing подробно рассмотрены в работах T. Akidau [1] и обзоре M. Hirzel [2], где показана роль событийной семантики, управления состоянием и низколатентной обработки в промышленной аналитике. Одновременно в исследованиях по эволюции стриминговых систем [3] подчёркивается рост сложности сопровождения таких решений. Это формирует потребность в архитектурных подходах, снижающих трудоёмкость модификации потоковых сервисов, в том числе за счёт декларативного и конфигурационного управления логикой обработки.

Основная часть

В работе предложена архитектура конфигурационно-управляемой системы потоковой обработки данных, ориентированной на унификацию и преобразование входящих событий в распределённом учётном контуре. В отличие от классических моделей, где бизнес-логика жёстко реализуется в коде потокового приложения [1], разработанный подход отделяет инфраструктурный конвейер обработки от правил трансформации данных.

Логика преобразования описывается декларативно через конфигурации и интерпретируется во время выполнения, что позволяет изменять сценарии обработки без перекомпиляции системы. Архитектура включает:

- унифицированный конвейер обработки событий;
- механизм декларативного описания правил трансформации;
- подсистему валидации конфигураций;
- модуль расширения для интеграции в распределённые платформы.

Предложенное решение сочетает принципы потоковых вычислений с моделью конфигурационного управления, что соответствует тенденциям повышения абстракции и модульности в распределённых системах. Практическая реализация подтверждает применимость подхода для построения сервисов конвертации событий в промышленной среде.

Выводы

Разработанная архитектура демонстрирует, что интеграция декларативного управления правилами обработки в потоковую систему позволяет снизить сложность модификации бизнес-логики и повысить адаптивность решения при сохранении преимуществ stream-processing модели. Полученные результаты подтверждают перспективность конфигурационно-управляемых архитектур для распределённых событийных систем. Дальнейшие исследования могут быть направлены на формализацию модели описания правил и сравнительный анализ производительности конфигурационного и традиционного программного подходов.

Литература

1. Akidau T. et al. The dataflow model: a practical approach to balancing correctness, latency, and cost in massive-scale, unbounded, out-of-order data processing //Proceedings of the VLDB Endowment. – 2015. – Т. 8. – №. 12. – С. 1792-1803.
2. Hirzel M. et al. A catalog of stream processing optimizations //ACM Computing Surveys (CSUR). – 2014. – Т. 46. – №. 4. – С. 1-34.
3. Fragkoulis M. et al. A survey on the evolution of stream processing systems //The VLDB Journal. – 2024. – Т. 33. – №. 2. – С. 507-541.