

УДК 004.896

РАЗРАБОТКА ИНСТРУМЕНТАЛЬНЫХ СРЕДСТВ ДЛЯ ИСПЫТАНИЙ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ВЫПУСКА МЕТАЛЛУРГИЧЕСКОЙ ПРОДУКЦИИ

¹Чирышев Ю.В., ¹Аксенов К.А., ¹Сысолетин Е.Г., ²Белан С.Б., ²Перескоков С.А.

¹ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина», Екатеринбург, e-mail: wiper99@mail.ru;

²ЗАО «Ай-Тек», Москва, e-mail: belan@i-teco.ru

В статье рассматриваются вопросы разработки программно-инструментальных средств, которые используются для подготовки и проведения опытной эксплуатации автоматизированной системы выпуска металлургической продукции (АС ВМП). Основной задачей таких инструментов является сбор и первичное наполнение хранилища данных тестируемой системы. Авторами работы сформулированы основные принципы построения таких инструментов и приведена структура разработанной системы, основанная на ETL архитектуре. На основе кроссплатформенных решений создана инструментальная автоматизированная подсистема восстановления последовательности событий (АС ВПС), произошедших в ходе выпуска единиц металлургической продукции. Описана методика и результаты загрузки данных в автоматизированную систему выпуска металлургической продукции. В АС ВМП были разработаны имитационные мультиагентные модели по совершенствованию и оптимизации технологических, логистических и бизнес-процессов в металлургической отрасли, которые были разработаны на основе загруженных данных.

Ключевые слова: ETL-процесс, автоматизированная информационная система, хранилище данных, опытная эксплуатация, металлургическая промышленность, надежность информационных систем, большие данные

DEVELOPMENT OF SOFTWARE TOOLS FEATURE FOR TESTING METALLURGICAL ENTERPRISE INFORMATION SYSTEM

¹Chiryshv Yu.V., ¹Aksenov K.A., ¹Sysoletin E.G., ²Belan S.B., ²Pereskokov S.A.

¹Ural Federal University named after First President of Russia B.N. Yeltsin, Yekaterinburg, e-mail: wiper99@mail.ru;

²I-Teco, Moscow, e-mail: belan@i-teco.ru

This paper presents the development of software tools that are used for testing metallurgical enterprise information system. These tools have been developed feature for data collection and initial loading into a data warehouse. Authors of a study made some principles based upon ETL processes that use for software development. The proposed approach is implemented in the event sequence assembly subsystem of the metallurgical information system. The technique and results of loading data into the target metallurgical enterprise information system are described. Examples of simulations models of metal industry process optimization and improvement are provided based on multi agent resources conversion process approach.

Keywords: ETL processes, automated information system, data warehouse, operation testing, metal industry, software reliability, big data

Одной из задач, решаемых на этапе подготовки автоматизированной информационной системы (АИС) к вводу в действие, является обеспечение качественных и количественных показателей функционирования АИС согласно техническому заданию (ТЗ). При разработке промышленных систем проверка на работоспособность и соответствие ТЗ осуществляется на этапах предварительных испытаний и опытной эксплуатации. В процессе испытаний учитывается широкий круг требований, обеспечивающих единый критерий качества системы и влияющих, в конечном счете, на работу системы в реальных условиях промышленной эксплуатации [2]. Разумеется, для каждой системы требования будут различными, так как должна учитываться специфика

конкретного предприятия и функциональность самой АИС. Несмотря на это, можно выделить несколько основных требований, общих для всех систем:

- 1) хранение и использование информации;
- 2) выполнение заданных функций системы;
- 3) интеграция и организация информационных потоков системы;
- 4) надежность и применимость системы.

Поскольку основной функцией любой АИС является хранение и работа с огромными объемами данных в течение довольно длительного периода времени, то необходимой основой для проведения таких испытаний являются инструментальные средства, позволяющие осуществить сбор и наполнение хранилища данных АИС.

В данной работе рассмотрена архитектура и принцип работы автоматизированной системы восстановления последовательности событий (АС ВПС), произошедших в ходе выпуска единиц металлургической продукции – это программный комплекс, который разработан и используется как инструмент, облегчающий процесс подготовки и проведения опытной эксплуатации АИС «Автоматизированная система выпуска металлургической продукции (АС ВМП)» [1, 5–6, 8–9]. Источниками данных АС ВПС являются наборы технологических дампов металлургического предприятия. Информационное обеспечение представляет собой реляционную базу данных Sybase.

Загрузка данных как ETL-процесс

Основной целью разработки АС ВПС является решение задачи по первичному наполнению хранилища данных АС ВМП сразу после развёртывания системы с целью дальнейшей аналитической обработки и комплексного тестирования системы. В первичное наполнение включается нормативно-справочная информация, заранее известные объекты и ресурсы, а также информация о процессах и параметрах, связанных с выпуском металлургической продукции. К первой группе относятся справочники параметров, единиц измерения и ресурсы, соответствующие структуре предприятия, такие как подразделения, агрегаты, процессы на агрегатах [3]. Эта информация требует редкой актуализации или вовсе не требует таковой. Ко второй группе относятся данные, соответствующие процессам деятельности предприятия, к ним относится информация о технологических, логистических и организационных процессах, а также значения параметров, определяющих состояния и условия протекания процессов, характеристики исходного сырья и готового продукта.

Анализ производственных процессов и структуры исходных данных металлургического предприятия [1, 3–4] показал, что извлечение данных и перенос их в хранилище данных (ХД) АС ВМП с целью дальнейшей аналитической обработки связаны с рядом проблем, основными из которых являются нижеследующие:

- исходные данные не соответствуют формату данных, допустимому для загрузки в ХД;
- данные в ряде таблиц являются «грязными» (содержат ошибки, которые мешают их корректному переносу (пропуски данных или наличие фиктивных значений));

– объем базы данных велик, что влечет накладные расходы на поиск, проверку и сортировку содержимого базы и, соответственно, на процесс импорта данных в целом.

Поэтому для переноса исходных данных из дампов предприятия в ХД АС ВМП следует использовать специальный инструмент, который обеспечивает последовательное извлечение данных и преобразование их в единый формат, поддерживаемый ХД. Очевидно, что поскольку ХД системы построено на определенной для металлургического предприятия модели данных, а интерфейс загрузки/выгрузки данных является событийным [3], то и процесс импорта должен разрабатываться с учетом всех особенностей используемой модели.

В данной работе задача импорта рассматривается с точки зрения процесса консолидации данных. Консолидация – это комплекс методов и процедур, направленных на извлечение данных из различных источников, обеспечение необходимого уровня их информативности и качества, преобразование в единый формат, в котором они могут быть загружены в хранилище данных или аналитическую систему. Комплекс программных средств, основанный на консолидации данных, имеет обобщенное название ETL (extraction, transformation, loading – извлечение, преобразование, загрузка [4]). Сам процесс переноса данных и связанные с ним действия называются ETL-процессом, а соответствующие программные средства – ETL-системами.

Независимо от особенностей построения и функционирования ETL-система обеспечивает выполнение трех основных этапов процесса переноса данных (ETL-процесса):

- 1) извлечение данных, на этом этапе данные извлекаются из одного или нескольких источников и подготавливаются к преобразованию;
- 2) преобразование данных, производятся преобразование форматов и кодировки данных, а также их обобщение и очистка;
- 3) загрузка данных, осуществляется запись преобразованных данных в систему хранения.

Существует несколько подходов к реализации процессов ETL. Общепринятый подход состоит в извлечении данных из систем источников, размещении их в промежуточной области дисковой памяти, выполнении в этой промежуточной области процедур преобразования и очистки данных, а затем загрузки данных в ХД [4].

Размещение извлеченных данных в промежуточной области означает запись данных в БД или файлы дисковой подсистемы. Альтернативным подходом к реализации процесса ETL является выполнение преобразований в оперативной памяти и непосредственную загрузку в ХД [4].

Преобразование данных в оперативной памяти выполняется быстрее, чем при размещении их предварительно на диске. Однако применение такого подхода лимитируется размером порции загружаемых данных. Если размер порции загружаемых данных достаточно большой, то необходимо использовать промежуточную область.

Архитектура подсистемы восстановления последовательности событий

На рис. 1 представлена функциональная структура системы и ее взаимодействие с внешними системами.

Подсистема восстановления последовательности построена по ETL-архитектуре. Ядром подсистемы является набор импортеров (программных пакетов), обеспечивающих формирование одностороннего потока данных из сервера баз данных металлургического предприятия в модуль обмена данными с автоматизированными системами предприятия (ОДАСП). В архитектуре выделены следующие компоненты:

Модуль взаимодействия с сервером металлургического предприятия. Выгрузка информации из БД осуществляется посредством механизма JDBC. Этот механизм представляет собой набор классов, написанных на языке Java и обеспечивающих связь АС ВПС с СУБД металлургического предприятия посредством выполнения SQL-запросов. Для стыковки реляционной модели данных с объектно-ориентированным представлением данных используется

механизм «мапинга», обеспечивающий преобразование табличных данных в объектное представление [4]. Соответственно любой импортер может читать данные из дампов металлургического предприятия и работать с ними как с обычными объектами.

Модуль преобразования данных. Основной задачей, решаемой на данном этапе, является процесс преобразования данных к виду, соответствующему структуре данных ОДАСП. Независимо от характера выгружаемых данных из базы металлургического предприятия, будь то данные о процессах, единицах продукции или параметрах, в конечном счете, каждым импортером формируются «события» [3]. Перед формированием исходных данных в пакеты событий происходит три стадии обработки:

1. *Временное согласование данных.* Данный шаг предполагает выстраивание данных в хронологическом порядке – от прошлого к настоящему, который соответствует последовательности технологических процессов получения чугуна, стали, проката и т.д. Такое упорядочивание необходимо для выстраивания генеалогии единиц продукции в том порядке, в котором они формировались в ходе металлургических процессов [3].

2. *Исправление и очистка данных.* Данный шаг связан с тем, что при переносе информации в ХД приходится сталкиваться с потоками «грязных» данных, которые могут стать причиной некорректных результатов анализа в дальнейшем.

3. *Поиск дубликатов.* На данном этапе решается задача согласования загружаемых данных с данными, которые уже загружены в ОДАСП. Такая проверка необходима, чтобы при многократном запуске процедуры импорта контролировать те данные, которые на данный момент уже загружены в хранилище.



Рис. 1. Структура АС ПВС

Интерфейсный модуль взаимодействия. Функцией данного модуля является обеспечение взаимодействия АС ВПС с сервером ОДАСП. Интерфейсный модуль осуществляет формирование требуемого запроса к серверу – «события» в терминах ОДАСП, его кодирование и отправку с использованием протоколов HTTP, SocketIO и UDP.

Загрузка данных проводится с помощью автоматических сценариев, эмулирующих работу смежных информационных систем предприятия. Сценарии написаны на языке высокого уровня Java и соответствуют конверторным процессам и процессам переработки металла сталеплавильного и прокатного передела металлургического предприятия. В основе эмуляции заложены следующие принципы:

- 1) главной составной частью процесса загрузки данных является сценарий (последовательность технологических действий по выпуску металлургической продукции);
- 2) сценарий соответствует стадии переработки металла в рамках одного цеха или агрегата металлургического предприятия;
- 3) сценарий состоит из нескольких шагов.

Каждый шаг соответствует определенной технологической операции: выплавка, доводка, вакуумирование, разливка стали, порезка слэбов т. д.

Поскольку перемещение огромного массива данных из источников в модуль ОДАСП является достаточно медленным процессом, важно учитывать его производительность. Для повышения производительности загрузки данных в АС ВПС реализованы классы, обеспечивающие многопоточный режим работы сценариев. Сценарии позволяют загружать либо все данные, хранящиеся в БД, либо выборочные, например, определенную марку стали

или данные за некоторый промежуток времени. Такой выбор глубины выгрузки исторических данных и их состав обеспечивает компромисс между объемом выгружаемых данных и их ценностью с точки зрения дальнейшего анализа. Состав импортируемых данных, а также величина временного интервала определяются в соответствии с пользовательскими настройками.

В целом процесс импорта включает две фазы: этап первичной загрузки данных и этап потоковой загрузки данных. Первый этап предназначен для начального импорта данных о структуре предприятия. Он позволяет занести начальную структуру (список цехов, агрегатов в каждом цехе) для корректного запуска сценариев импорта данных. После загрузки структуры предприятия происходит процедура регистрации справочников параметров. Обе процедуры выполняются автоматически. Второй этап – основной режим работы, связанный с непосредственным выполнением сценариев по импорту данных.

Результаты исследований и их обсуждение

Загрузка данных осуществлялась с нескольких рабочих станций, на которых были установлены компоненты АС ВПС. Поскольку каждый сценарий импорта является лишь частью полного цикла выпуска продукции и эмулирует стадии получения и переработки металла только в рамках одного передела предприятия, каждая станция выполняла загрузку данных за определенный временной интервал. В качестве временного отрезка был выбран один год работы металлургического предприятия. Такое решение позволило в отведенные для загрузки сроки

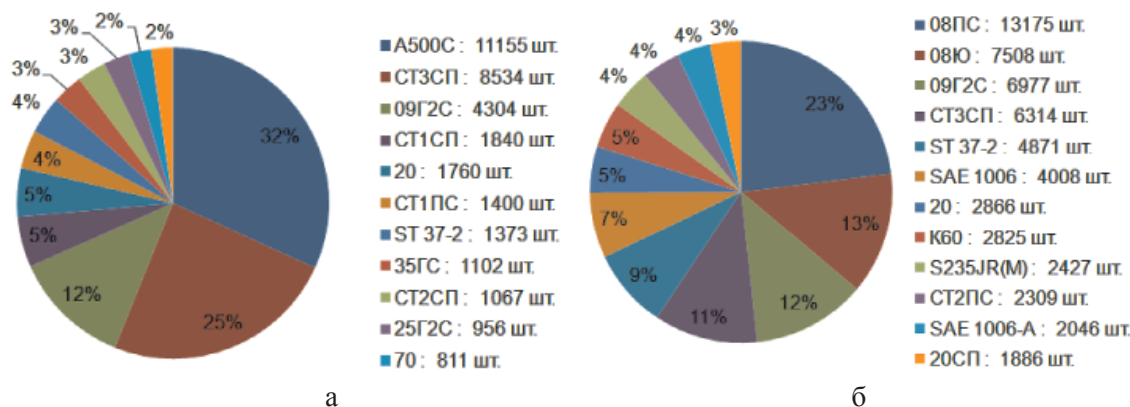


Рис. 2. Диаграмма плавок:
а – электросталеплавильного производства; б – кислородно-конвертерного производства

(720 часов) и без привлечения большого количества единиц вычислительных средств импортировать в целевую систему исторических данных в общем объеме 1,5 ТБ, что эквивалентно объему данных, сгенерированному информационными системами металлургического предприятия за три года работы. На рис. 2 приведен количественный состав загруженных данных (плавов) с процентным отношением марок стали для конвертерного и электростале-плавильного производства.

Кроме данных о выплавке при проведении экспериментов загружались данные о следующих стадиях получения и переработки металла: производство сортовой и слэбовой заготовки для дальнейшего передела в сортовой и листовой прокат; производство горячекатаного и холоднокатаного проката; производство горячеоцинкованного проката и проката с полимерным покрытием.

Заключение

Разработка и эксплуатация автоматизированной системы восстановления последовательности событий показала полную приемлемость системы для нужд проведения предварительных испытаний и опытной эксплуатации АС ВМП. На основе данных, загруженных средствами АС ВПС, был реализован и апробирован ряд решений, в первую очередь имитационных моделей [1, 5–7, 10], направленных на оптимизацию технологических и логистических процессов металлургического предприятия. Кроме этого, применение данного средства позволило получить статистические характеристики для подтверждения качественных и количественных показателей надежности АС ВМП.

Работа выполнена в рамках договора № 02.G25.31.0055 (проект 2012-218-03-167) при финансовой поддержке работ Министерством образования и науки Российской Федерации.

Список литературы

1. Аксенов К.А., Антонова А.С., Спицина И.А., Сысолетин Е.Г., Аксенова О.П. Разработка автоматизированной системы анализа, моделирования и принятия решений для металлургического предприятия на основе мультиагентного подхода // Автоматизация в промышленности. – М., 2014. – № 7. – С. 49–53.
2. Бородин А.М., Мирвода С.Г., Поршневу С.В. Особенности тестирования устойчивости к сбоям корпоративных информационных систем методом генерирования отказов // Современные проблемы науки и образования. – 2014. – № 5. – URL: www.science-education.ru/119-14997 (дата обращения: 18.11.2016).
3. Сысолетин Е.Г., Аксенов К.А., Круглов А.В. Интеграция гетерогенных информационных систем современного промышленного предприятия // Современные проблемы науки и образования. – 2015. – № 1–1. – URL: <http://www.science-education.ru/ru/article/view?id=19030> (дата обращения: 18.11.2016).
4. Туманов В.Е. Проектирование хранилищ данных для систем бизнес-аналитики (учебное пособие) // Международный журнал экспериментального образования. – 2013. – № 5. – С. 74–76.
5. Aksyonov K., Bykov E., Sysoletin E., Aksyonova O., Nevolina A. Integration of the Real-time Simulation Systems with the Automated Control System of an Enterprise // International Conference on Social Science, Management and Economics (SSME 2015) Guangzhou, China, May 09–10, 2015, WOS:000361112900160. – P. 871–875.
6. Aksyonov K., Bykov E., Sysoletin E., Aksyonova O., Goncharova N. Perspectives of Modeling in Metallurgical Production // International Conference on Social Science, Management and Economics (SSME 2015) Guangzhou, China, May 09–10, 2015, WOS:000361112900161. – P. 876–880.
7. Aksyonova O.P., Wang Kai. The method of planning and bottleneck analysis of enterprise project portfolio // Microwave & Telecommunication Technology: 24th Int. Crimean Conference (CriMiCo'2014). 7–13 September. – Sevastopol. – Vol. 1. – P. 439–440.
8. Borodin A. et al. On design of domain-specific query language for the metallurgical industry // International Conference: Beyond Databases, Architectures and Structures. – Springer International Publishing, 2015. – P. 505–515.
9. Borodin A., Mirvoda S., Porshnev S. Database Index Debug Techniques: A Case Study // International Conference: Beyond Databases, Architectures and Structures. – Springer International Publishing, 2016. – P. 648–658.
10. Nevolina A.L., Aksyonova O.P., Smoliy E.F. Development of simulation systems and decision-making method in the logistic field // Microwave & Telecommunication Technology: 24th Int. Crimean Conference (CriMiCo'2014). 7–13 September. – Sevastopol. – Vol. 1. – P. 435–436.