

Автоматизация процессов работы с хранилищами данных на платформе Micro Focus Vertica

Продукты компании EasyData

Алексей Константинов

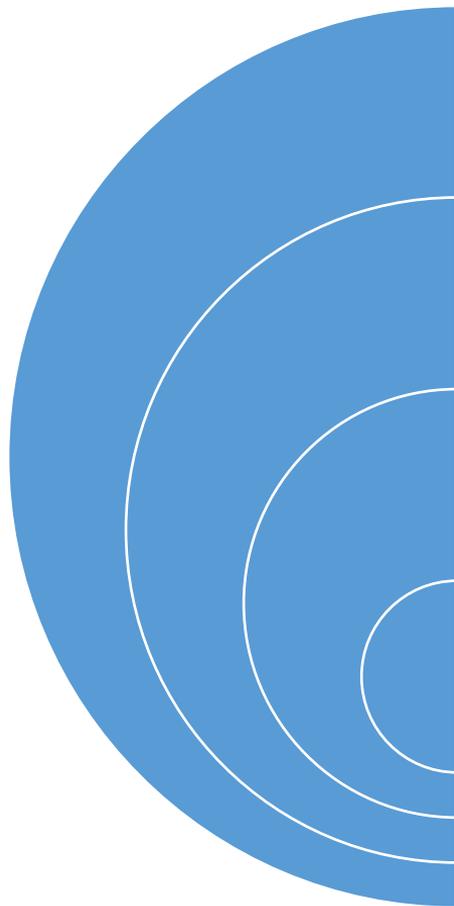
О компании EasyData

- ✓ С 2013 г. реализует проекты на BigData
- ✓ Многолетний подтвержденный золотой статус Micro Focus Gold Big Data Specialist
- ✓ Разработчик крупных проектов в области хранилищ данных на платформе Micro Focus Vertica и Cloudera Hadoop
- ✓ Сертифицированный тренер и аудитор решений ХД на платформе Micro Focus Vertica
- ✓ Статус глобального технологического партнёра Vertica в категории «Developer/Modelling/IDE»
- ✓ Разработчик open source, freeware и коммерческого ПО автоматизации ETL/ELT/Cron процессов

Техническая поддержка Vertica

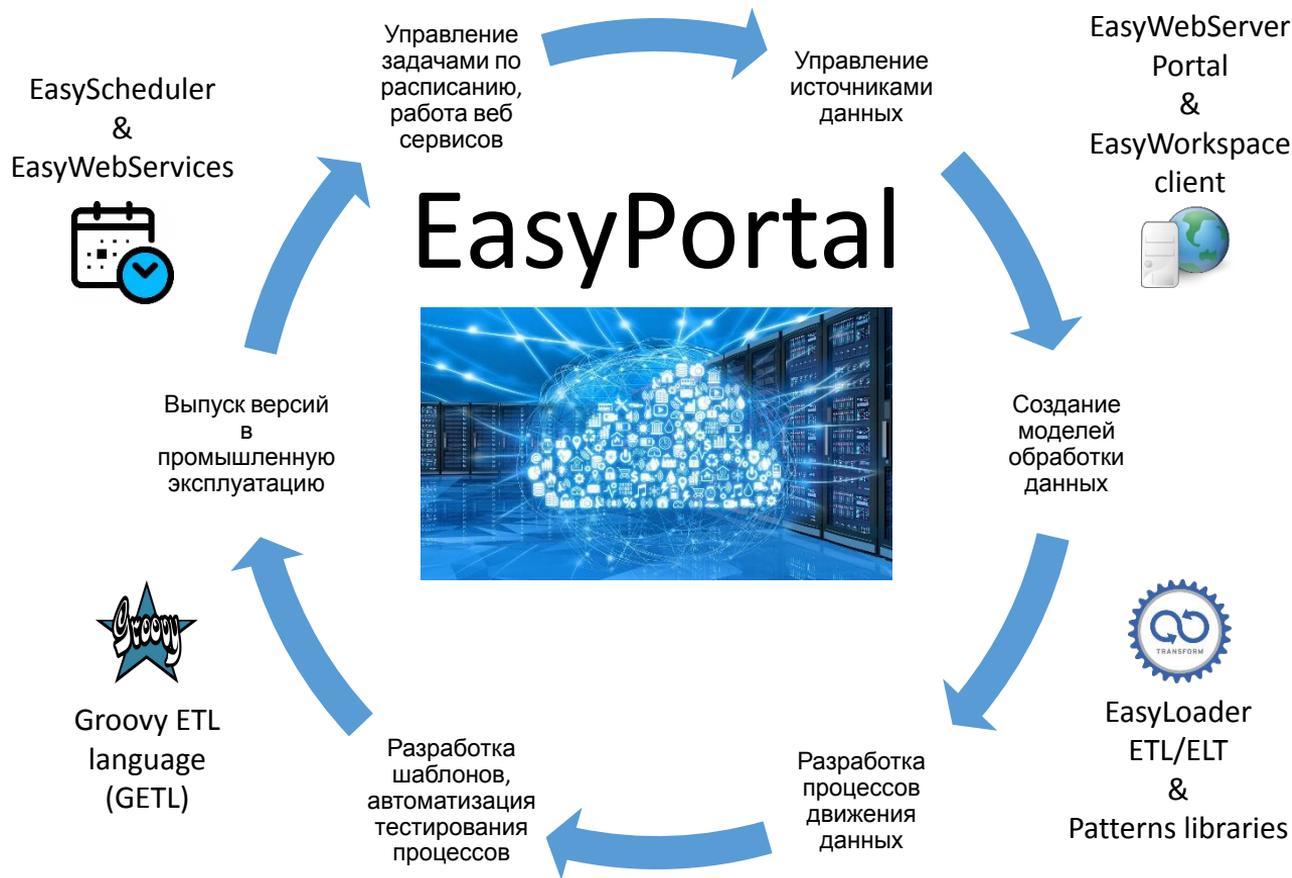


Автоматизация ХД Vertica

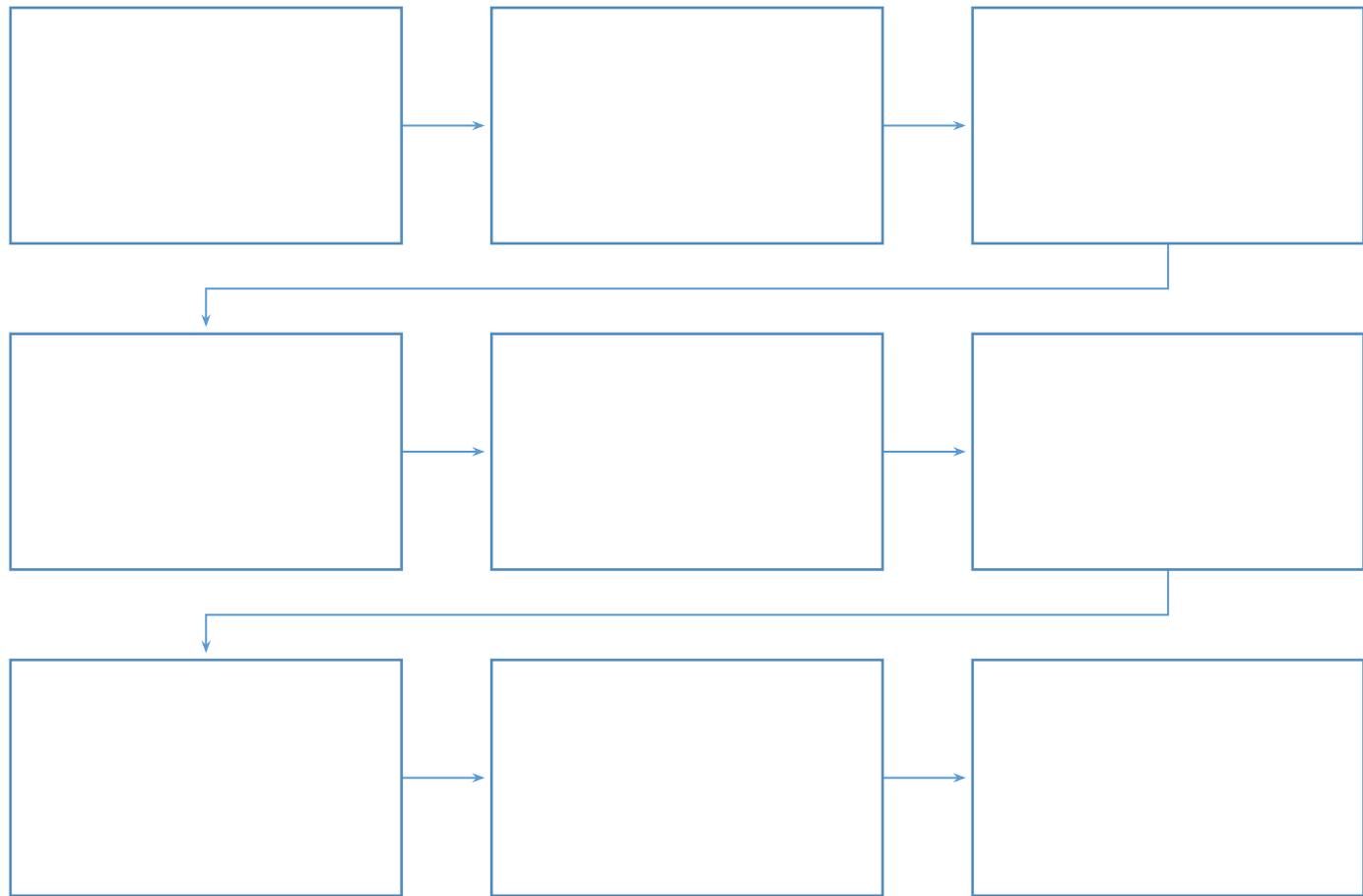


Yota [250 тб]	<ul style="list-style-type: none">• ETL процессы захвата и доставки данных• ELT процессы расчета витрин• Репликация OGG• Управление и мониторинг задачами на серверах эксплуатации ETL/ELT/Hadoop
Humans [10 тб]	<ul style="list-style-type: none">• ETL процессы захвата и доставки данных• ELT процессы расчета витрин• Репликация OGG• Управление и мониторинг задачами на серверах эксплуатации ETL/ELT/Hadoop
Трансфин М [1 тб]	<ul style="list-style-type: none">• ETL процессы захвата и доставки данных• ELT процессы расчета витрин• Репликация OGG• Управление и мониторинг задачами на серверах эксплуатации ETL/ELT/Hadoop
Ситилинк [15 тб]	<ul style="list-style-type: none">• Оптимизация работы Vertica• Синхронизация стендов разработки и промышленной эксплуатации

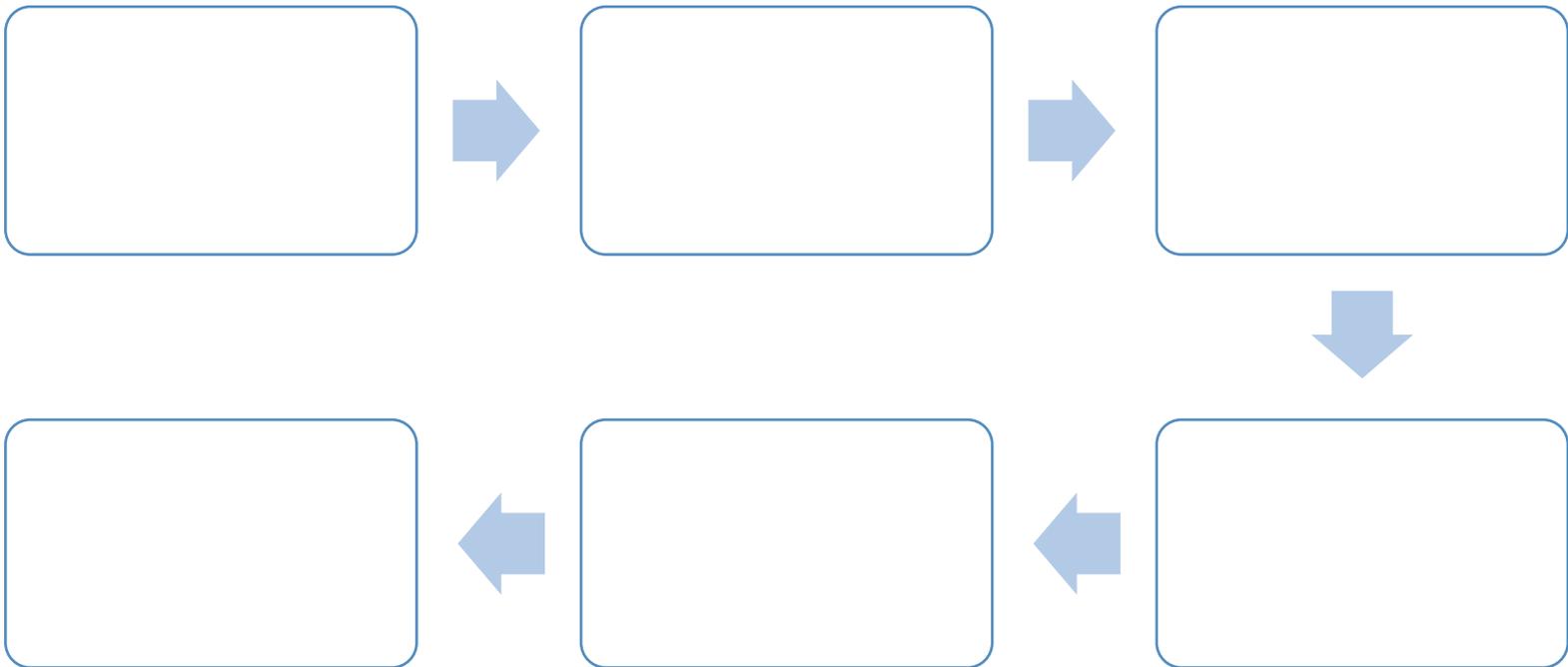
Экосистема продуктов EasyData



Работа с EasyPortal



Инициализация данных из источников в таблицы ХД



Инициализация данных

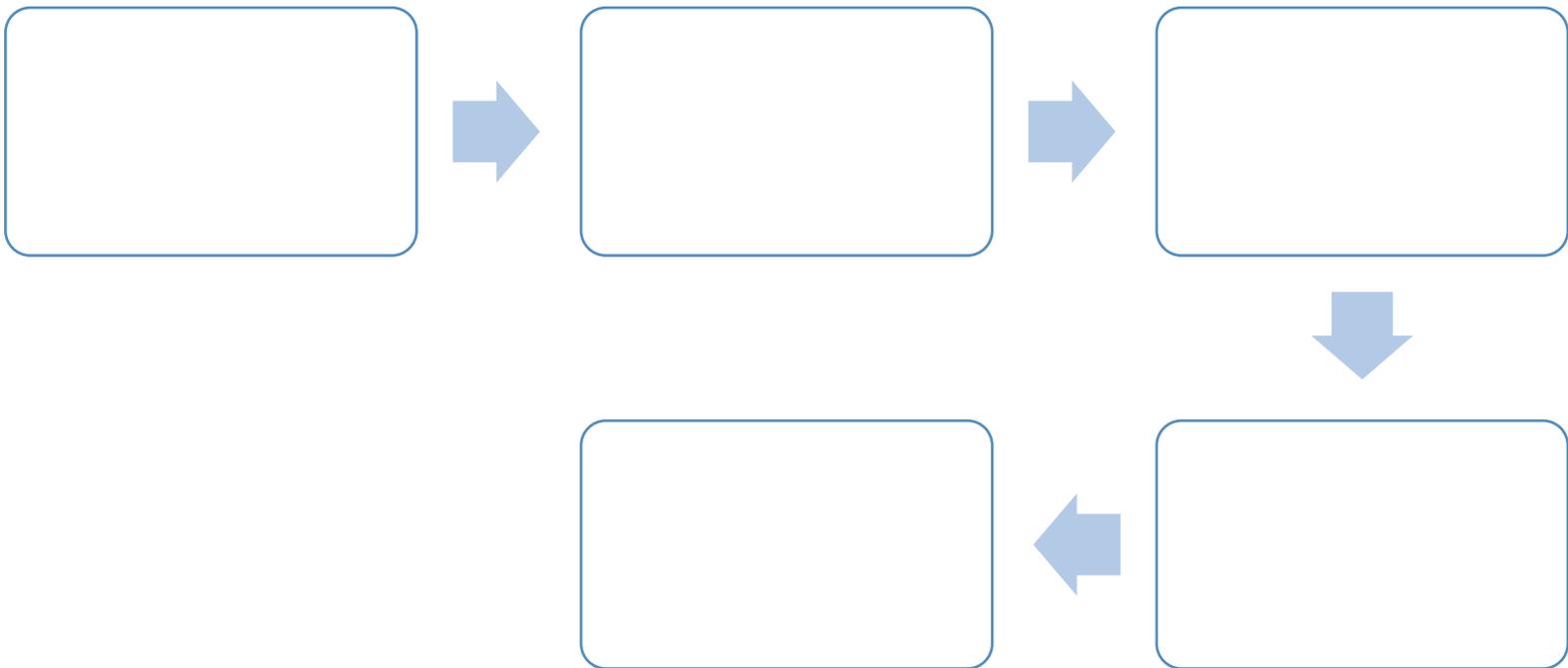
Поддержка
выгрузки хинтом
partition

Не требуется
заново выгружать
с источника в
случае ошибки
загрузки в ХД

Раздельные
транзакции
выгрузки и
загрузки на
источнике и
приёмнике

Дополнительная
логика
инициализации на
языке ХП

Репликация данных источников в таблицы ХД



Репликация из источников

Не более 9000
полей в таблице

Для таблиц без РК
поддерживаются
только операции
вставок записей

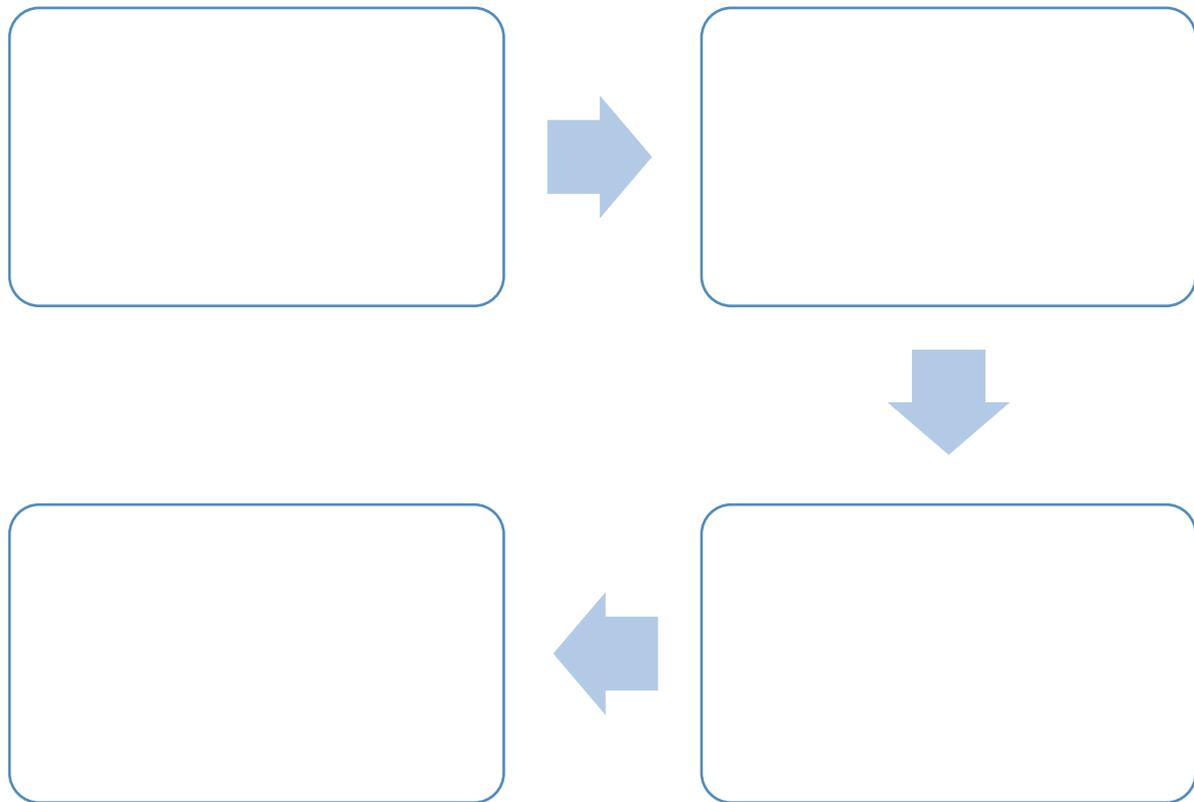
Можно
перегрузить заново
с нужной scp файлы

Горизонтальное
масштабирование:
выполнение на
всех нодах кластера

Оптимизация
выполнения INSERT,
UPDATE, DELETE
операторов

Возможность
анализа ошибок
логов репликации

Инкрементальный захват данных с источников



Инкрементальная загрузка

Точка захвата может
быть числовым или
дата-время полем



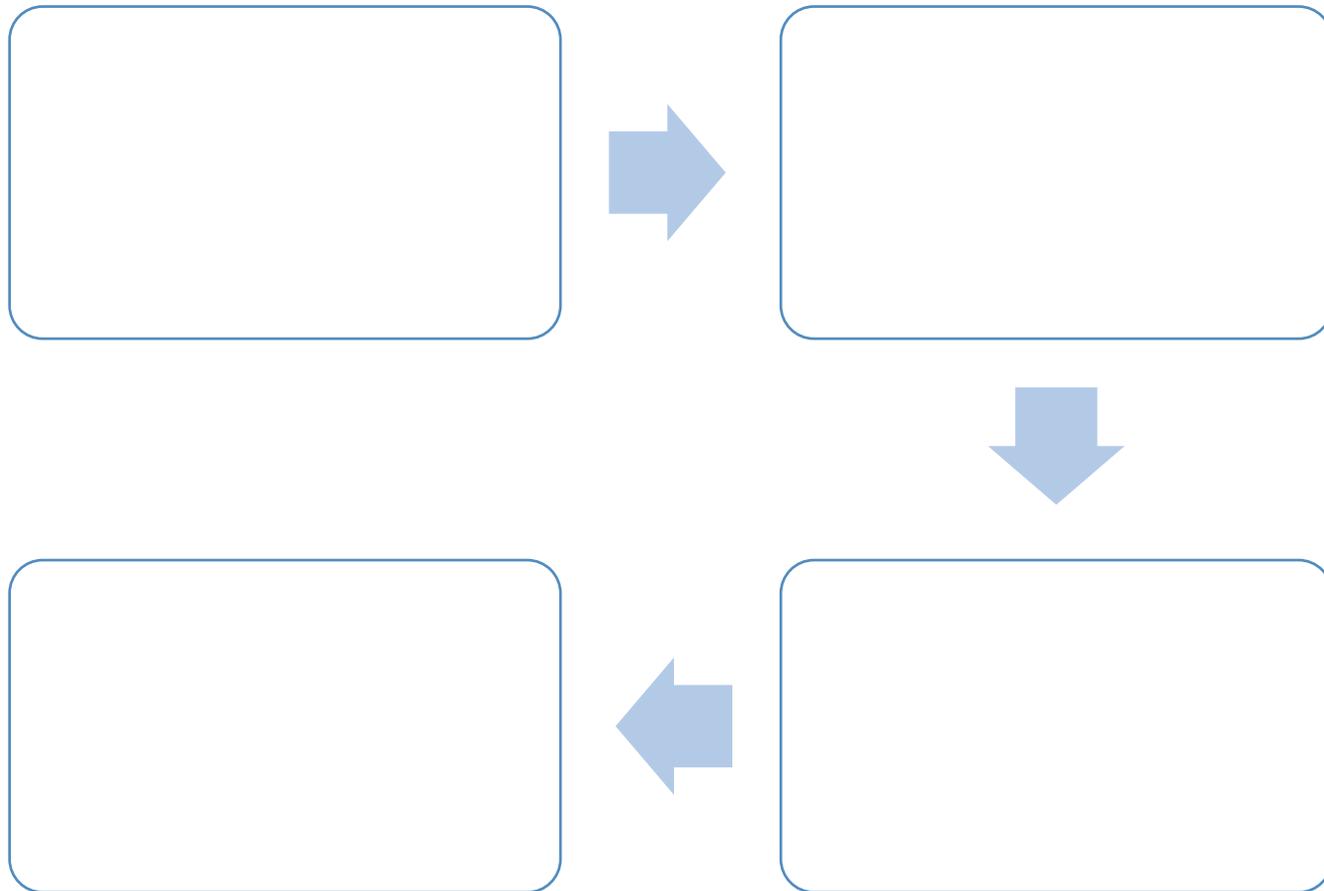
Можно
перегрузить заново
с нужной точки
файлов



Раздельные
транзакции
выгрузки и загрузки
на источнике и
приёмнике

Дополнительная
логика после загрузки
на языке ХП

Расчеты витрин



Преимущества EasyPortal



Преимущества EasyScheduler

WebCenter

Спасибо за внимание!

Алексей Константинов



+7 499 390 88 92



easydata.ru



info@easydata.ru