

Оптимизация обновления информационной базы

Информация для технических специалистов



- Долго проходит монопольный этап обновления
 - на больших базах может не завершаться за двое суток, т.е. нельзя перевести работу на новую версию за выходные
- Отложенные обработчики обновления выполняются последовательно
 - пока полностью не обновлены одни данные, обработка других даже не начинается
- Зависимость обработчиков друг от друга не позволяет даже теоретически выполнять их в несколько потоков без риска некорректного обновления
- Нет информации об объеме необновленных данных
- Многократная повторная запись одних и тех же объектов разными обработчиками обновления
- Отсутствует контроль работы пользователя с необновленными данными
 - пользователь может изменить данные, сделав их обновление невозможным

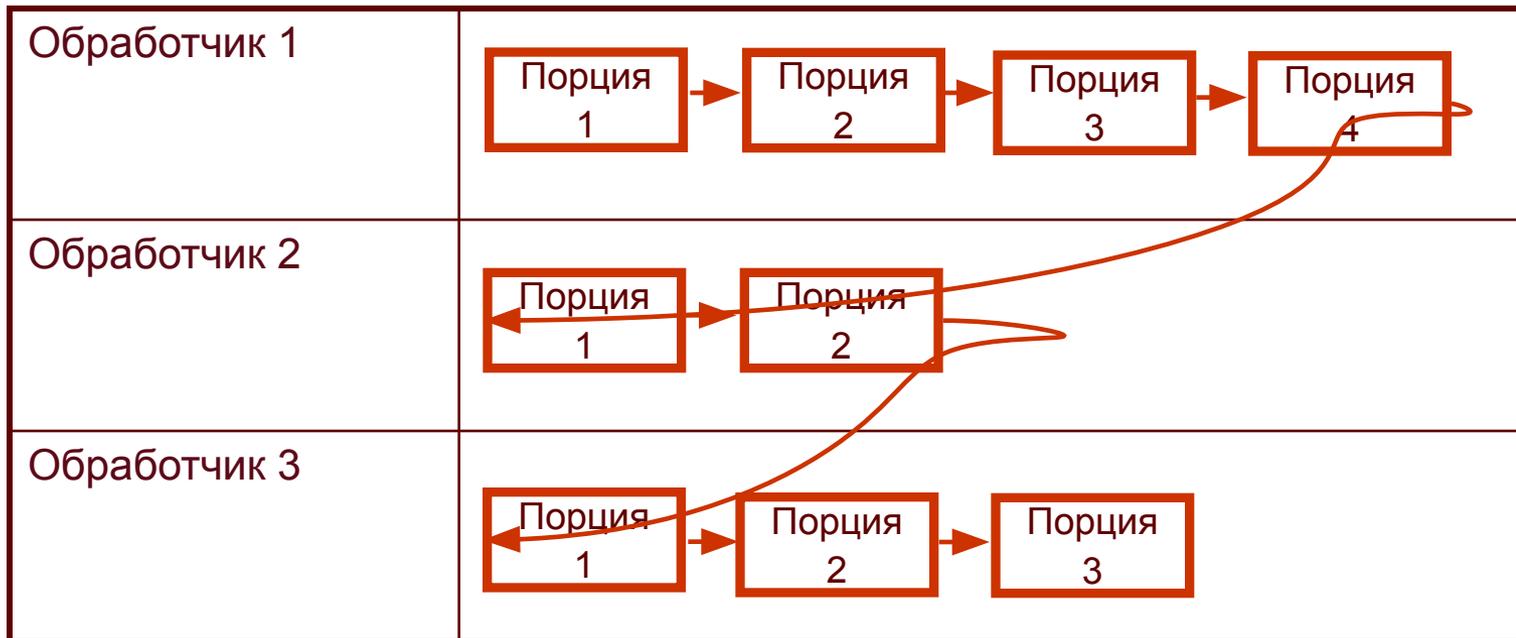


- Радикальное сокращение количества монопольных обработчиков обновления
 - если раньше монопольно обновлялась НСИ и данные текущих периодов, то теперь монопольно обновляется только НСИ
- Параллельный режим выполнения отложенных обработчиков обновления
 - обработчики обновляют данные порциями, после обработки одной порции данных первым обработчиком, происходит обработка первой порции второго обработчика и т.д. – т.к. выборка данных в каждой порции упорядочена по убыванию дат, данные всех разделов примерно с одинаковой скоростью обновляются от текущих к архивным
- Все данные, которые предстоит обновить, перед обновлением регистрируются на специальном плане обмена **Обновление информационной базы**
 - обработчики перед выполнением явным образом проверяют, обновлены или необходимые для их работы данные – это делает обработчики независимыми друг от друга, их можно запускать в несколько потоков, они не испортят друг другу данные (сам запуск в несколько потоков сейчас не реализован)
 - всегда можно простым и универсальным запросом выяснить сколько и каких данных еще не обновлено (для этого предусмотрен специальный API)
 - на основе этой информации реализована блокировка работы пользователя с необновленными объектами
- Переписаны обработчики обновления, многократная запись объектов практически исключена



Параллельный режим выполнения отложенных обработчиков обновления

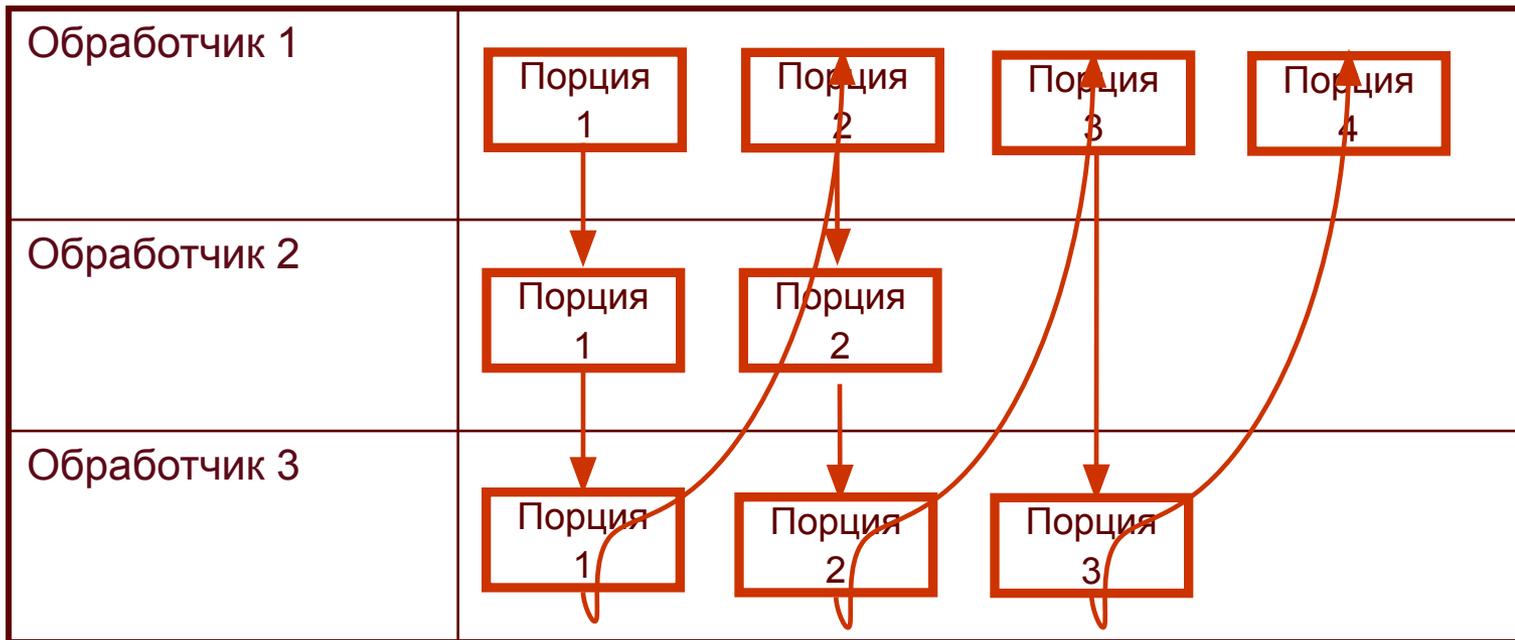
- Раньше все обработчики выполнялись последовательно





Параллельный режим выполнения отложенных обработчиков обновления

- Теперь обработчики выполняются параллельно





Параллельный режим выполнения отложенных обработчиков обновления

- Параллельный режим выполнения вкпе с правилом «обновление идет от текущих данных к архивным» позволяет
 - например, кладовщику не ждать пока обновятся данные, необходимые, например, кассиру
 - максимально снизить дискомфорт пользователей при работе с базой, находящейся в процессе обновления
 - т.к. пользователи чаще работают с текущими данными, чем с архивными, вероятность, что пользователи столкнутся с необновленными данными становится тем меньше, чем дольше идет обновление



Очереди отложенной обработки данных

По разным причинам написать обработчики полностью независимыми не представляется возможным, поэтому реализован механизм очередей. Рассмотрим его на примере

Дано

- Есть два обработчика – «Обновление реквизитов Документа» (Обработчик 1) и «Обновление движений Документа по Регистру, в т.ч. по данным обновленных реквизитов» (Обработчик 2)
- По сути обработчиков конкретный документ должен сначала быть обработан Обработчиком 1, а затем Обработчиком 2

Реализация

- После анализа Обработчику 1 присвоена очередь №1, а Обработчику 2 – очередь №2
- После выполнения всех монопольных обработчиков запускаются процедуры регистрации, написанные отдельно Обработчика №1 и Обработчика №2. Процедуры регистрации анализируют состояние базы с регистрируют к обработке:
 - Обработчик 1 на узле по очереди №1 служебного плана обмена Обновление информационной базы регистрирует все Документы, в которых нужно обновить реквизиты
 - Обработчик 2 на узле по очереди №2 регистрирует все регистраторы, по которым нужно обновить движения
- Монопольная часть обновления завершается, пользователи могут входить в программу и работать. В фоне стартует отложенное обновление
 - Из данных на узле по очереди №1 выбирается первая порция документов для обработки обработчиком 1. Данные обрабатываются, при этом удаляется регистрация Документов на узле по очереди №1
 - Запускается обработчик №2, который выбирает первую порцию регистраторов для переформирования движений. При этом выбираются регистраторы, которые зарегистрированы на узле по очереди №2 и которые не зарегистрированы (или по которым уже удалена регистрация) по очереди №1. При выполнении обработчика 2 удаляется регистрация на узле по очереди №2



Очереди отложенной обработки данных

Особенности использования очередей отложенной обработки данных

- очереди присваиваются независимо от версии, в которой выполняется обработчик. Версия влияет на то, будет выполняться обработчик при переходе между конкретными сборками. Все отобранные обработчики запускаются в порядке очередей
- независимо от очереди обработчики выполняются параллельно – сначала Порция 1 обработчика очереди 1, затем – Порция 1 обработчика очереди 2 и т.д.
 - т.к. во всех выборках данные упорядочены по убыванию дат, велика вероятность, что уже при первом запуске обработчика очереди №2 данные, необходимые ему, будут уже обновлены обработчиками очереди №1
- в одной очереди регистрируются несколько обработчиков (сейчас при переходе на новую подредакцию выполняются около 230 обработчиков, при этом используются около 20 очередей)
 - обработчики одной очереди точно не зависят друг от друга ни по читаемым, ни по изменяемым данным
 - если обработчики меняют один тип данных, они точно располагаются в разных очередях (при этом фактически они могут и не записывать одни и те же данные)
- для построения очередей используется новый функционал СППР (будет добавлен в публикуемую версию СППР несколько позже), который на основании заполненной разработчиками информации о читаемых и изменяемых обработчиками данных присваивает ему очередь
 - это означает, например, что конкретный номер очереди одного и того же обработчика может меняться при добавлении других обработчиков



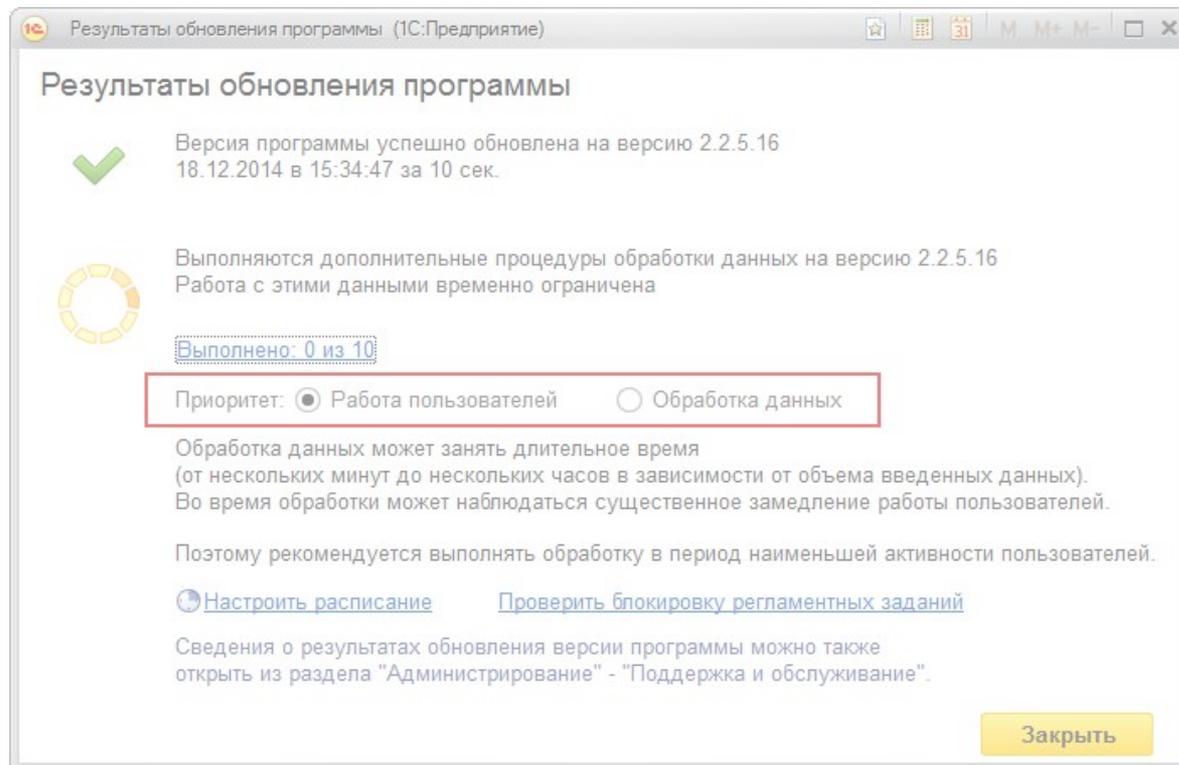
- На время отложенного обновления блокируются
 - формы объектов и формы записей независимых регистров сведений
 - блокируются только те объекты, которые еще не обновлены (т.е. по которым еще есть регистрация на узлах плана обмена «Обновление информационной базы»)
 - формы таких объектов открываются в режиме «Только просмотр», пользователю сообщается список обработчиков, которые еще не обработали этот объект
 - запись необновленных объектов
 - блокируется запись всех типов объектов, если их запись может помешать работе процедур обновления
 - работа офф-лайн-механизмов с необновленными данными, например:
 - механизм автоматического формирования расходных ордеров на товары будет формировать ордера только по тем распоряжениям на отгрузку, по которым полностью завершено обновление регистра «Товары к отгрузке»
 - проводки регламентированного учета будут формироваться только по полностью обновленным документам



Управление процессом обработки данных

Повышение приоритета обновления

- по умолчанию перед обработкой очередной порции данных делается пауза для уменьшения нагрузки на сервер, чтобы не замедлять работу пользователей
- теперь можно динамически этим управлять. Например, можно на ночь повышать приоритет обновления данных, а утром возвращать приоритет работе пользователей





- Появилась возможность повышать приоритет отдельных обработчиков данных
 - при этом нужно понимать, что повышение приоритета только одного обработчика, без повышения приоритета обработчиков, от которых этот обработчик зависит, может не дать ожидаемого эффекта



Особенности обновления РИБ (только в УТ 11)

- Полный РИБ
 - обновляется конфигурация главного узла, выполняются монопольные обработчики обновления
 - данные выгружаются в подчиненные узлы
 - в подчиненном узле загружается новая конфигурация и результаты выполнения монопольных обработчиков в центральном узле
 - в подчиненном узле запускаются монопольные обработчики обновления, которые в основном обрабатывают в холостую, изменяя только те данные, которые успели ввести между сеансами обмена
 - в подчиненном узле запускаются отложенные обработчики обновления, кроме тех, для которых определено, что они выполняются только в центральном узле (таких обработчиков очень мало, в основном это обработчики, которые создают объекты с новыми значениями ссылок)
 - в подчиненном узле, так же как и в центральном, работают процедуры блокировки данных на время обновления
 - данные, которые обновляются только в центральном узле так же блокируется – информация о том, что они уже обработаны передаются в подчиненный узел вместе с данными обмена.



- РИБ с фильтрами – отличия от полного РИБ
 - всё обновление (и монопольное, и отложенное) проходит только в главном узле
 - в первом, после обновления главного узла, сообщении обмена в подчиненный узел передаются информация обо всех объектах, которые будут обновляться. На основе этой информации в подчиненном узле регистрируются данные в плане обмена «Обновление информационной базы»
 - механизм блокировки данных в подчиненном узле работает так же, как в главном
 - информация о том, что они уже обработаны передаются в подчиненный узел вместе с данными обмена



- По данным наших замеров (в т.ч. на базах предоставленных пользователями), монопольная стадия обработки данных при обновлении теперь даже на больших базах проходит не дольше, чем за 3 часа
- Выполнение отложенного обновления ускорилось в разы
- С одной стороны дискомфорт пользователей, связанный с работой в базе, в которой не завершилось отложенной обновление, сведен к минимуму, с другой стороны система контролирует, что работа пользователей не помешает выполнению процедур обновления
- Новый механизм обновления имеет большой потенциал по расширению средств диагностики хода процесса обновления, т.к. всегда известен пул необработанных объектов
- Обновление корректно обрабатывает в РИБ



Планы по развитию механизма

- Большая часть доработок механизмов обновления сейчас реализована по месту в ERP (и соответственно в КА 2, и УТ 11) – доработки будут перенесены в БСП
- Будет реализовано информирование пользователей отчетов о чтении необновленных данных
- Следующие доработки мы планируем включить в конфигурацию, но уже сейчас их достаточно просто реализовать на конкретных внедрениях
 - Расширение средств диагностики хода обновления
 - Реализация возможности запуска отложенного обновления в несколько потоков

