



Sync Services for ADO.NET

Вересов Андрей



Зачем это нам надо?

- **Online**
 - оптимизация количества запросов к серверу
 - ускорение расчетов
- **Offline**
 - возможность работать без подключения к серверу



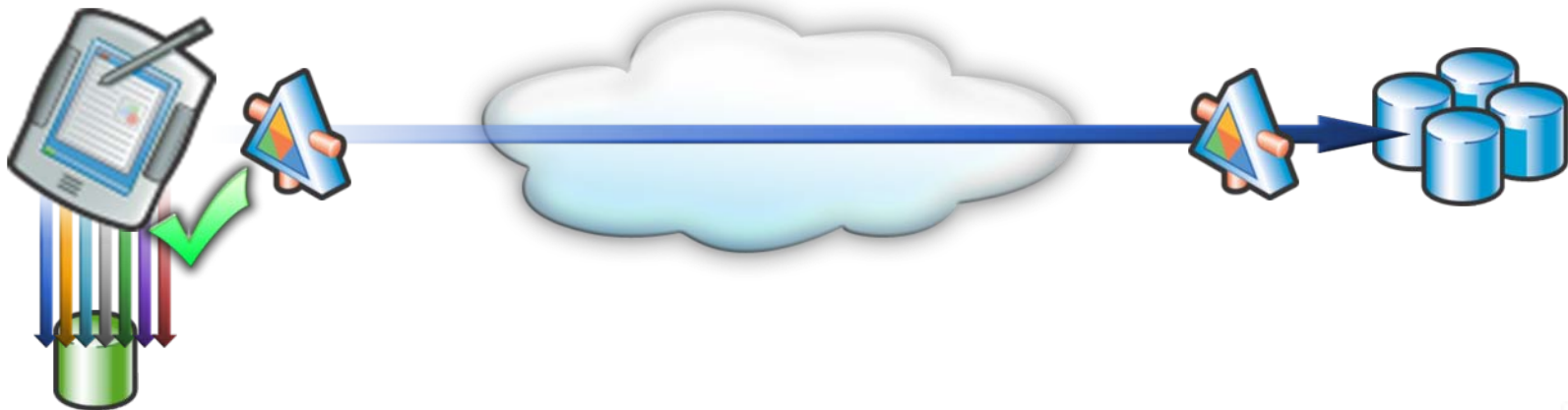


Online (обычное приложение)



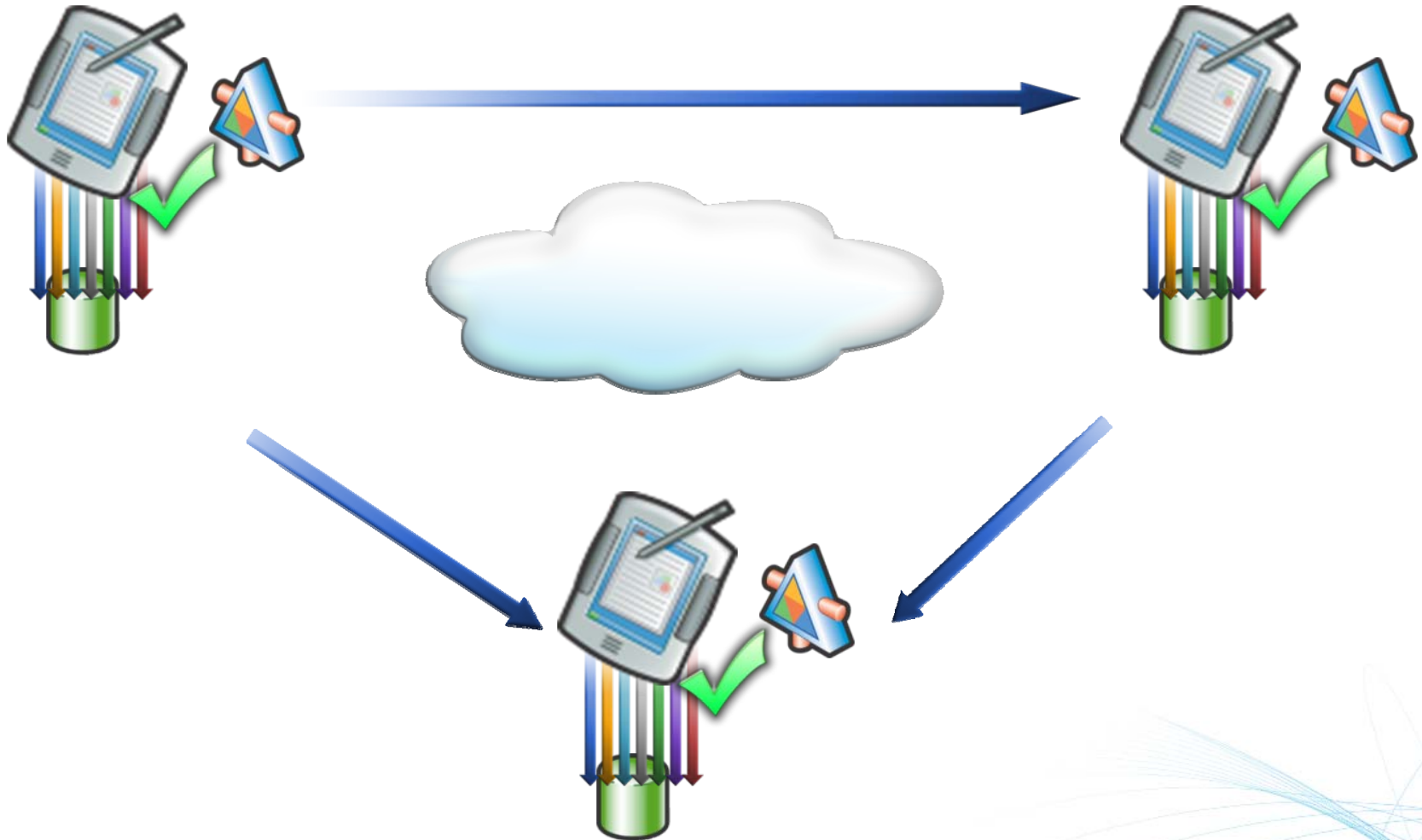


Online (кэширование данных)





P2P





Sync Framework

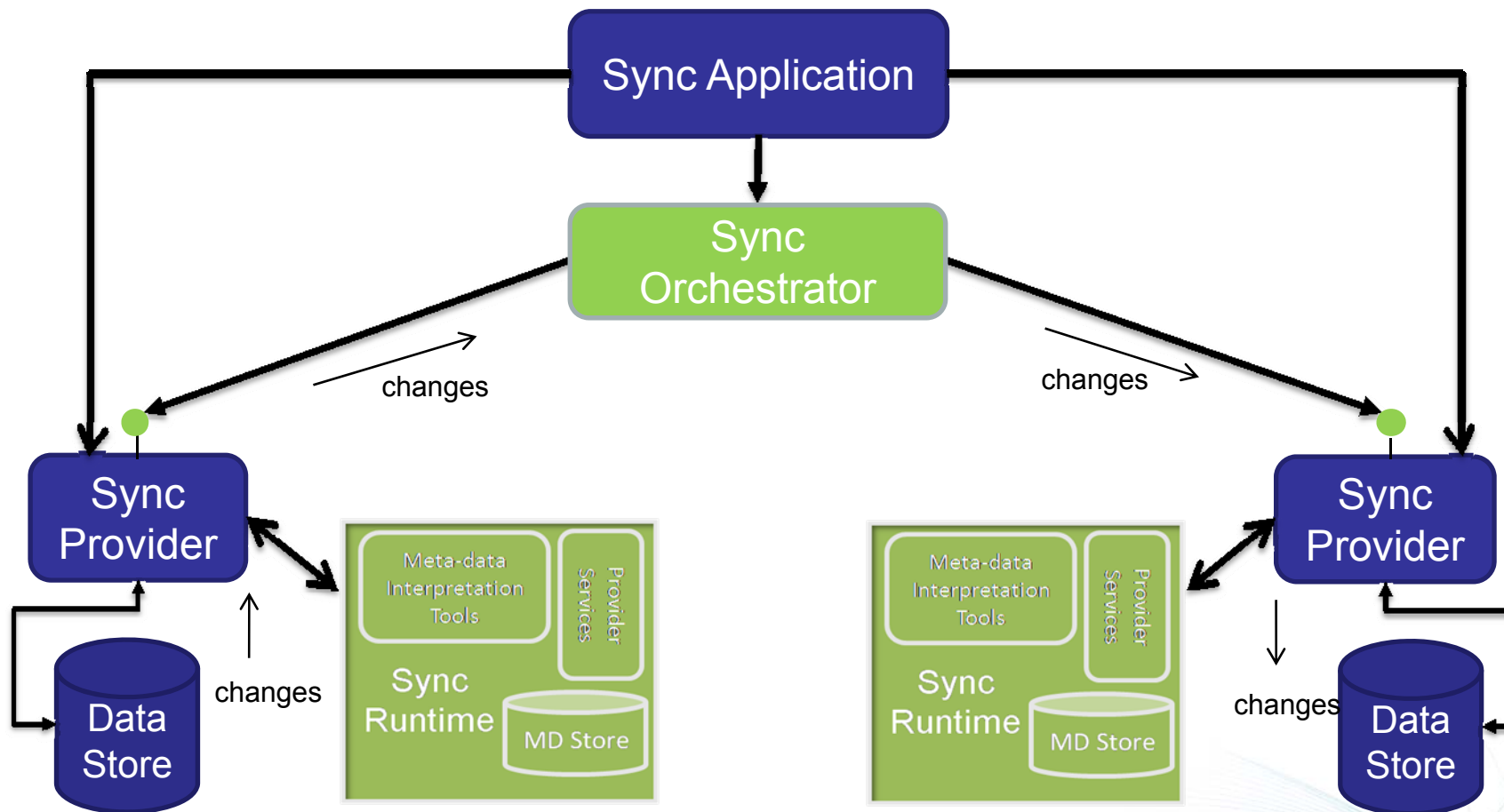
Предоставляет:

- инфраструктуру, в том числе подсистему автоматического разрешения конфликтов
- готовые модули (провайдеры) для работы с файловой системой, реляционными БД
- простоту создания собственных провайдеров
- средства интеграции

Sync Services for ADO.NET базируются на Sync Framework



Процесс синхронизации





Описание задачи

Есть веб приложение. Хочется сделать аналог, но десктопный.

- Помимо сервера БД для веб сайта есть внутренний сервер БД, где данные подготавливаются перед отправкой в БД веб сайта
- Пользователи могут добавлять свои комментарии, т.е. Данные меняются на обоих серверах
- SQL Server 2005 на сервере, SQL Server Compact 3.5 на клиенте
- Прямой доступ дать не можем, очень хотим через веб сервисы

Sync Services for
 **ADO.NET v2.0**



Sync Services for ADO.NET

- Работает как с большим сиквелом так и с SQL Server Compact
- Умеет работать с сервером БД напрямую или через WCF
- Использует DataSet
- Предоставляются шаблоны для Visual Studio
- Легко настраиваемый





Структура таблиц

```
CREATE TABLE [dbo].[Table1](  
    [Id] [int] IDENTITY(1,1) NOT NULL,  
    [CreationTimestamp] [bigint] NULL,  
    [update_originator_id] [int] NULL,  
    [LastEditTimestamp] [bigint] NULL,  
    [LastEditTimestampsource] [timestamp] NOT NULL,  
    CONSTRAINT [PK_Table1] PRIMARY KEY CLUSTERED (Id ASC))
```

```
CREATE TABLE [dbo].[Table1_Tombstone](  
    [Id] [int] NOT NULL,  
    [DeletionTimestampSource] [timestamp] NOT NULL,  
    [update_originator_id] [int] NULL,  
    [DeletionTimestamp] [bigint] NULL,  
    CONSTRAINT [PKDEL_Table1] PRIMARY KEY CLUSTERED (Id ASC))
```



Еще пара шагов: триггеры и процедура выбора записей

```
CREATE TRIGGER [dbo].[Table1_update_trigger] on [dbo].[Table1] for update  
as
```

```
    UPDATE o SET o.update_originator_id = 0  
        FROM Table1 o JOIN [inserted] i on o.Id = i.Id WHERE NOT  
UPDATE(update_originator_id);
```

```
    UPDATE o SET o.LastEditTimestamp = o.LastEditTimestampSource  
        FROM MCMatchHistory o JOIN [inserted] i on o.Id = i.Id;
```

```
go
```

```
CREATE TRIGGER [dbo].[Table1_delete_trigger] on [dbo].[Table1] for delete  
as
```

```
    insert into Table1_Tombstone (Id, DeletionDate, update_originator_id) select Id,  
getdate(), 0 from deleted
```



Еще пара шагов: триггеры и процедура выбора записей

```
CREATE procedure [dbo].[sp_new_batch_anchor] (  
    @sync_last_received_anchor binary(8) ,  
    @sync_batch_size          int,  
  
    @sync_max_received_anchor binary(8) output,  
    @sync_new_received_anchor binary(8) output,  
    @sync_batch_count        int output,  
  
    @sync_group_name         nvarchar(250))  
as  
...  
go
```



SOAP сериализация данных

Sync Services используют DataSet для предоставления набора изменений

Размер	Реальных данных	Результирующего SOAP сообщения	Соотношение
10 x 10	200	9 038	~45 раз
100 x 10	2000	60 573	~30 раз
1000 x 10	20000	577 066	~28 раз
10000 x 10	200000	5 758 852	~28 раз

* в каждой колонке двух байтное целое число

Итого для 10 000 записей ~5,49 MB данных

Если бинарная сериализация, то ~3,66 MB – не очень спасает 😊



Что если объем синхронизируемых данных большой?



- Первая синхронизация зачастую очень большая
- Произошло обновление во многих записях
- Долгое время синхронизация не проводилась



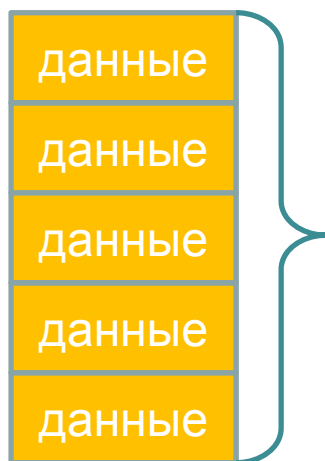
Передавать данные в виде готового *.sdf файла



Поддержка батчей



- Плохой коннект, разрывы связи
- Долгое время ожидания, можем вылететь по таймауту
- Большой размер сообщений, банально может не хватить памяти для сериализации/десериализации



Таблица

Разбить весь набор данных на несколько кусков и синхронизируем не за один раз, а за несколько запросов





Неравномерная вставка или большое кол-во удалений

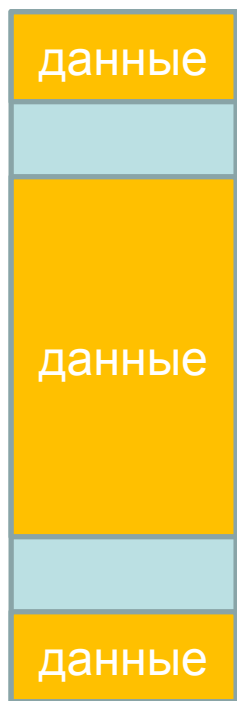


Таблица 1

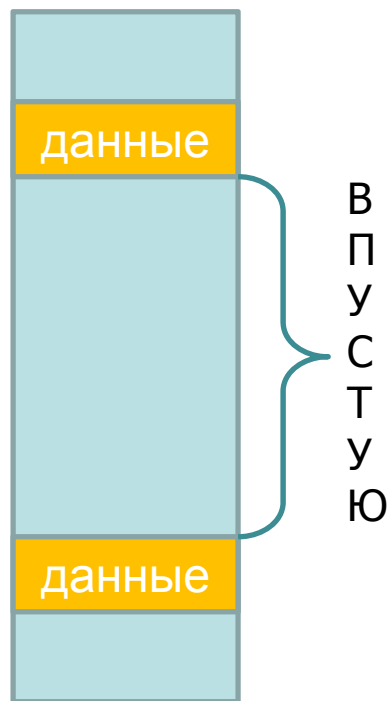
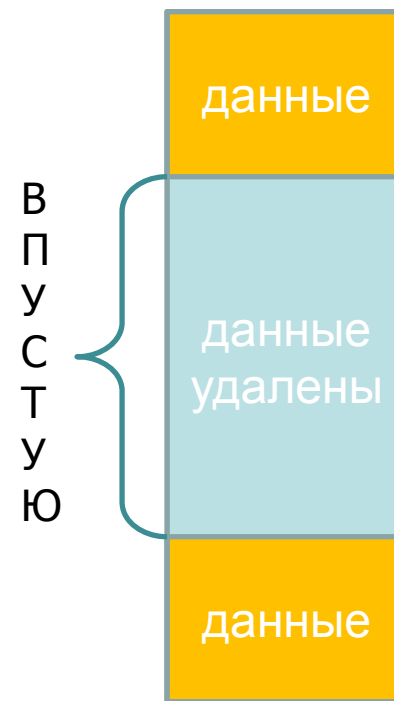


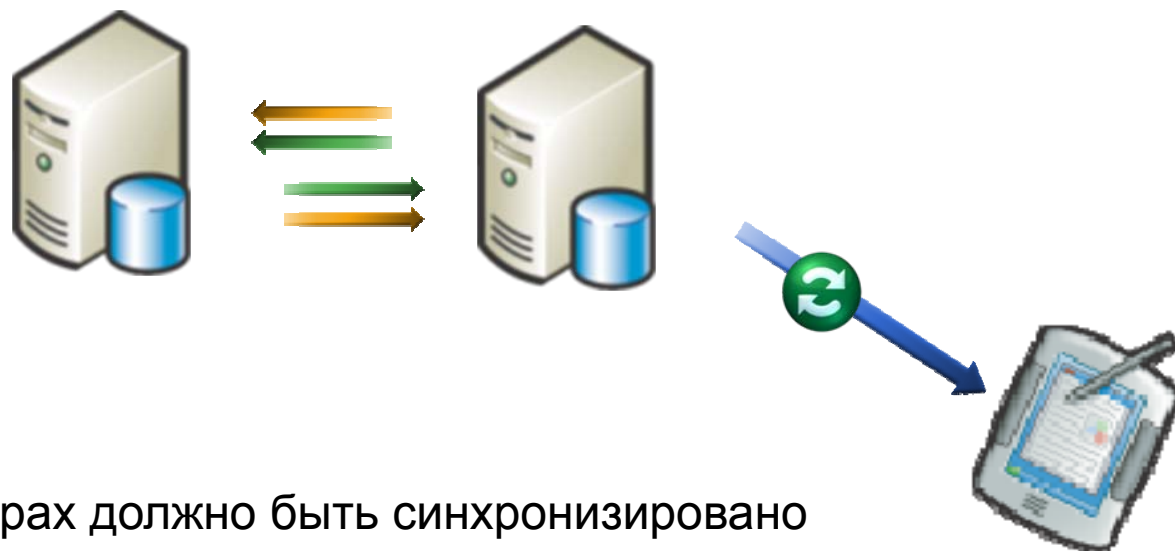
Таблица 2



Таблица



Что если данные создаются на двух серверах а синхронизируются только с одним?



DateTime

время на серверах должно быть синхронизировано

Timestamp

для каждой таблицы данные меняются только на одном сервере



Очистка таблицы удаленных записей



Со временем, размер таблицы удаленных записей растет и в какой-то момент может начать влиять на производительность



Удалять, но удалять нужно осторожно ибо кто-то может не узнать об удаленной строке



Q&A