

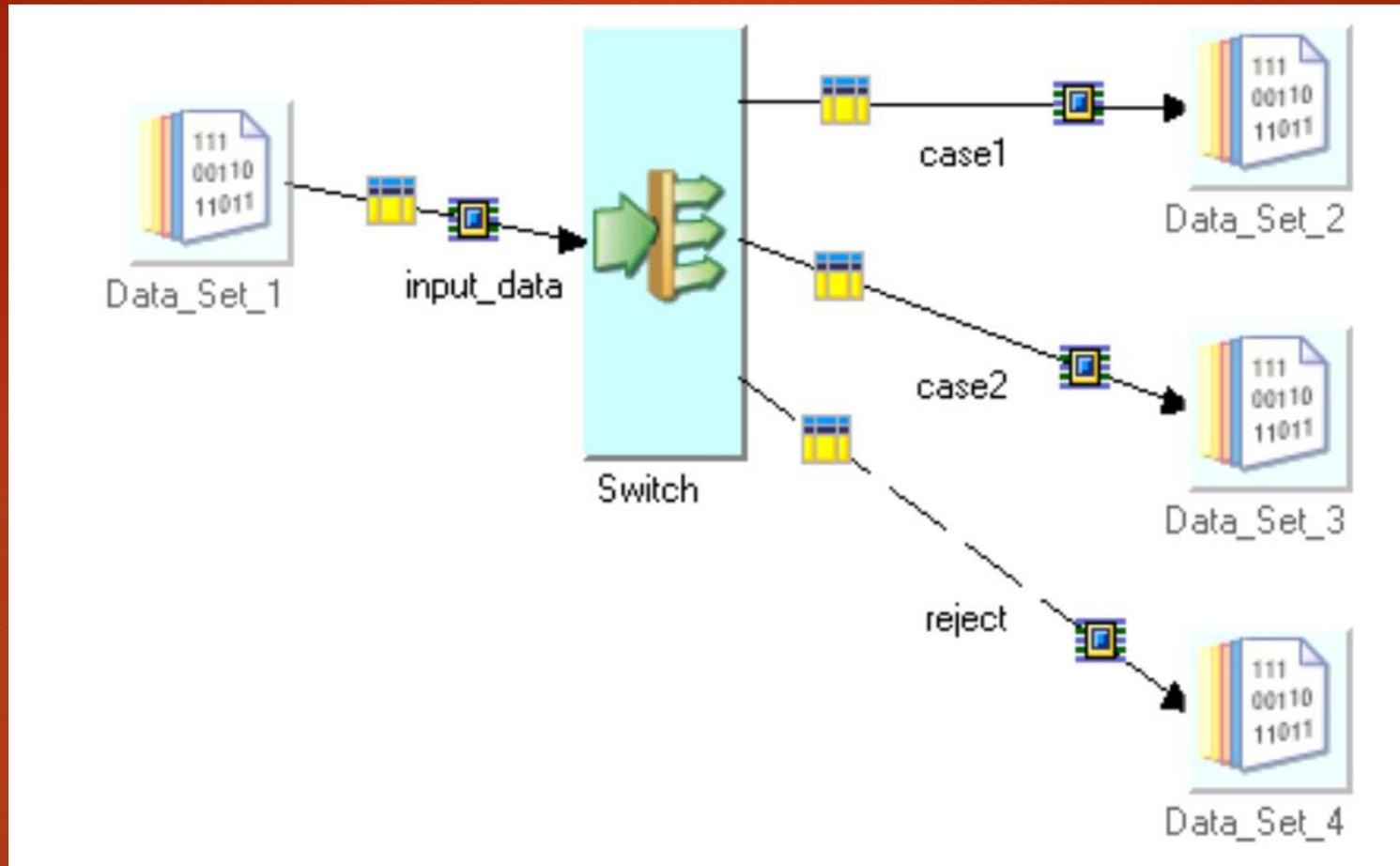
Switch Stage





Switch Stage это stage обработки. Он может иметь одну входную ссылку, а на выходе может иметь до 128 выходных ссылок. Switch stage принимает один набор данных в качестве входных данных и присваивает входной набор в выходной набор данных на основе значения поля селектора. Switch stage выполняет операцию, аналогичную в языке C операцию, которая осуществляет переход к той ветви программы, для которой значение переменной-селектора или выражения совпадает с указанным. Строки, которые не удовлетворяют ни одному выражению, выводятся в reject link (отверженная ссылка). Все остальные попадают в Case link.

Пример Switch Stage



Switch Stage состоит из трех частей:

- ▶ **Stage Page:** указываются общую информацию и характеристики данного компонента.
- ▶ **Input Page :** указываются подробная информация о входном наборе , из которого выбираются строки.
- ▶ **Output Stage:** указываются сведения об обработанных данных, которые поступают на ВЫХОД.



- ▶ Обязательно нужно использовать свойство селектора указать входной столбец, по которому и будет произведен выбор `switch`. Вы также можете указать является ли столбец чувствительным к регистру или нет. Другие свойства зависят от того, какой режим вы выбрали:



- ▶ Если вы выбрали свойство **User-defined mapping mode**, находящимся под **User-defined Mapping category** укажите выражение **Case** в свойствах.
- ▶ В категории **Options** выберите если не найдена характеристика, какое действие нужно произвести. Если значение столбца не соответствует ни одному из случаев, также нужно указать действие(будь-то удаление строк или переход по **reject** ссылке).

Stage Page

На вкладке **General** можно указать дополнительное описание **stage**. На вкладке **Properties** можно указать, что **switch stage** делает. Вкладка **Advanced** позволяет указать этапы выполнения. В **Link Ordering** вкладке указывается порядок вывода обработанных ссылок. **NLS Locale** позволяет выбрать **locale**(языковой стандарт) другой, чем по умолчанию проекта, чтобы определить правила сортировки.

Properties Tab

- ▶ Вкладка **Properties** позволяет Вам определять свойства, которые определяют что на самом деле делает **Stage**. Некоторые свойства обязательны, хотя у многих есть настройки по умолчанию. Свойства без дефолта параметры настройки кажутся в цвете предупреждения (красными по умолчанию) и становятся черными , когда Вы поставяете значение для них.

Input Category

- ▶ Selector

Указывает столбец ввода, который применяется к Switch.

- ▶ Case Sensitive

Указывает, является ли столбец чувствительным к регистру или нет.

- ▶ Selector Mode

Указывает, как вы собираетесь определять case заявления в switch.
Выбирать между:

User-defined Mapping. По умолчанию и означает, что вы должны предоставлять явные сопоставления case значения выходов. Если вы используете этот режим, вы укажите выражение User-defined Mapping category.

Auto. Может быть использована там, где существует несколько различных селекторов.

User-defined Mapping Category

▶ Case

Это свойство появляется, если вы выбрали режим выбора пользователя. Укажите выражение Case в собственном случае. Выражение имеет следующий формат:

Selector_Value[= Output_Link_Label_Number]

Например, в случае заявления:

1990=0

1991

1992

1993=1

1994=1

приведет к строкам, содержащим даты 1990 1991, или 1992 году в колонки селектор маршрутизации выходного **case 0**, а в строках, содержащих в период с 1993 по 1994 года направляется на выходное звено **case 1**.

Options Category

► If not found

Указывает действие, выполняемое, если строка не соответствует ни одному из **case** заявления. Не появится, если вы выберете режим селектор. В противном случае, выбирать между:

1. **Fail.** Причины сбоя задания.
2. **Drop.** Удалить строки.
3. **Output.** Направить строки в **Reject links**.

Advanced Tab

Вкладка “Дополнительно” позволяет указать некоторые необязательные параметры.

► **Режим выполнения** .

Этап может выполняться в параллельном режиме или последовательном режиме. В параллельном режиме ввода данных обрабатывается с помощью доступных узлов , как указано в файле конфигурации, а также любым узлом ограничений , указанных на Advanced вкладке. В последовательном режиме весь набор данных обрабатывается узлом проводника.

► **Режим сочетаемости** .

Параметр задан по умолчанию автоматически, что позволяет InfoSphere® DataStage® комбинировать операторы, которые лежат в параллельных стадиях таким образом , что они работают в том же самом процессе , если это целесообразно для данного типа сцены.

Advanced Tab

- ▶ **Сохранение секционирования.**

Этот **Propagate** является свойством по умолчанию. Он принимает **Set** или **Clear** от предыдущей стадии. Вы можете явно выбрать **Set** или **Clear** . При **Set** следующий этап работы будет сохранен в разделе, при **Clear** наоборот удален.

Link Ordering Tab

- ▶ Эта вкладка позволяет указать ссылки на этапы, которые имеют более чем одно звено и где указание ссылки является обязательным.

NLS Locale Tab

- ▶ Появляется, если вы **NLS** включена в системе. **NLS** позволяет посмотреть текущая конвенцию и выбрать другую при необходимости. Вы также можете использовать параметр задание локали, или перейдите к файлу, который определяет настраиваемые правила сортировки.

Inputs Page

Общая (**General**) вкладка позволяет указать дополнительное описание линии ввода. Разметка (**Partitioning**) вкладка позволяет определить, как входящие данные разделены перед тем, как начнет работу **Stage**. Столбцы (**Columns**) вкладка определяет определения столбцов входящих данных. Дополнительно (**Advanced**) вкладка позволяет изменить настройки по умолчанию для буферизации входного звена.

Partitioning Tab

Partition вкладка позволяет указать подробную информацию о том, как входящие данные разделены или собраны , прежде чем начнет работу **Stage**. Она также позволяет указать, что данные должны быть отсортированы перед тем, как работать дальше.

Методы секционирования:

- ▶ **(Авто)** . Пытается выработать наилучший метод разделения в зависимости от режимов выполнения текущего и предыдущих этапах, и сколько узлов указаны в файле конфигурации. Этот метод разметки по умолчанию для **Switch stage**.
- ▶ **Весь(Entire)** . Каждый файл , записанный получает весь набор данных.
- ▶ **Hash** . Записи хэшируются на разделы на основе значения ключевого столбца или столбцов , выбранного из **доступных** списка.
- ▶ **Modulus** . Записи разделены с помощью функции модуля на ключевого столбца , выбранного из **доступных** списка. Это обычно используется для разделения на полях тегов.
- ▶ **Random** . Записи разделены случайным образом , на основании выходного сигнала от генератора случайных чисел.
- ▶ **Same** . Консервированные разбиение уже на месте.
- ▶ **DB2®** . Репликация метод разбиения DB2 конкретной таблицы DB2. Требуется дополнительные свойства , которые будут установлены. Доступ к этим свойствам, нажав на кнопку Свойства.
- ▶ **Диапазон(Range)** . Деление набора данных на приблизительно равные перегородками размера на основе одного или нескольких ключей разделения. Диапазон разделения часто является шагом предварительной обработки для выполнения общего рода на множестве данных. Требуется дополнительные свойства , которые будут установлены. Доступ к этим свойствам, нажав на кнопку Свойства.

Методы сбора :

- ▶ **(Авто)** . Это метод сбора по умолчанию . Обычно, когда вы используете автоматический режим, Info Sphere DataStage с нетерпением прочитает любую строку из любого входного раздела , как он станет доступным.
- ▶ **Заказанный (Ordered)** . Читает все записи из первого раздела, а затем все записи из второго раздела, и так далее.
- ▶ **Round Robin** . Читает запись из первого раздела ввода, то из второго раздела, и так далее. После достижения последнего раздела, оператор начинается заново.
- ▶ **Сортировка слиянием** . Читает записи в порядке , основанном на одном или нескольких столбцах записи. Для этого необходимо выбрать для сбора ключевой столбец из списка доступных.

Outputs Page

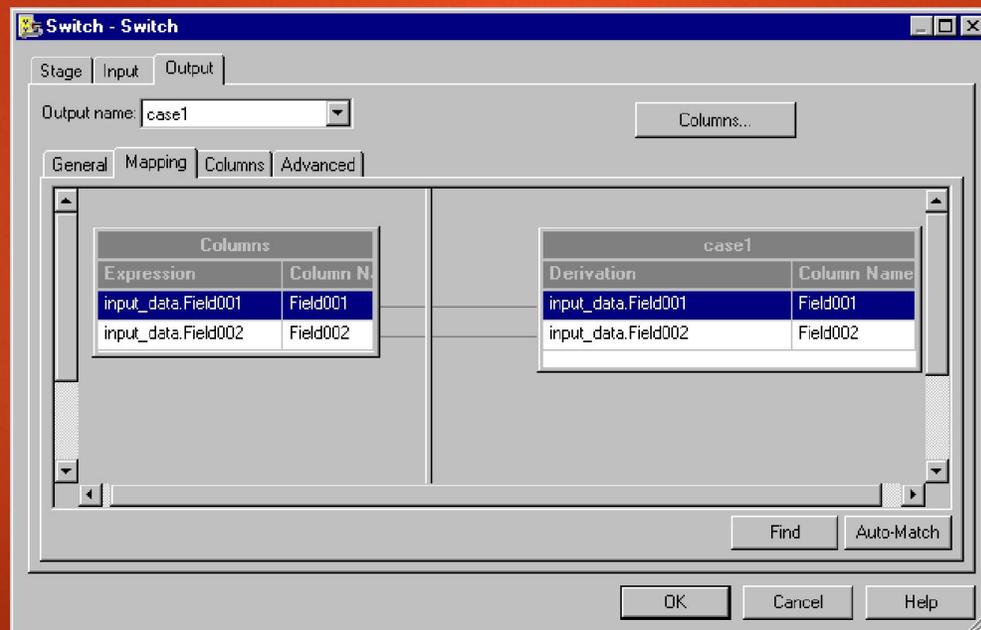
- ▶ Страница **Output** позволяет указать подробную информацию о данных , выводимых из **Switch**.
- ▶ **Switch Stage** может иметь до 128 исходящих ссылок.
- ▶ **General** вкладка позволяет указать дополнительное описание выходного звена. **Columns** вкладка определяет определения столбцов данных. **Mapping** вкладка позволяет определить взаимосвязь между колонками, поступающих на этапе **Switch stage** и столбцы вывода. **Advanced** вкладка позволяет изменить настройки по умолчанию для буферизации выходного звена. Также может иметь **reject** ссылку строк, которым было отказано. Нужно выбрать ссылку с которой вы хотите работать в окне **Output name**.

Mapping Tab



- ▶ В **Switch stage** вкладка Mapping позволяет указать, каким образом столбцы выводятся.

- ▶ На левой панели отображаются входные столбцы или генерируемые столбцы. Они только для чтения и не могут быть изменены на этой вкладке.
- ▶ На правой панели отображаются выходные столбцы для каждой ссылки. Оно имеет поле происхождения, где можно указать, как выводится столбец. Вы можете заполнить его путем перетаскивания входных столбцов с помощью системы автоматического соответствия.



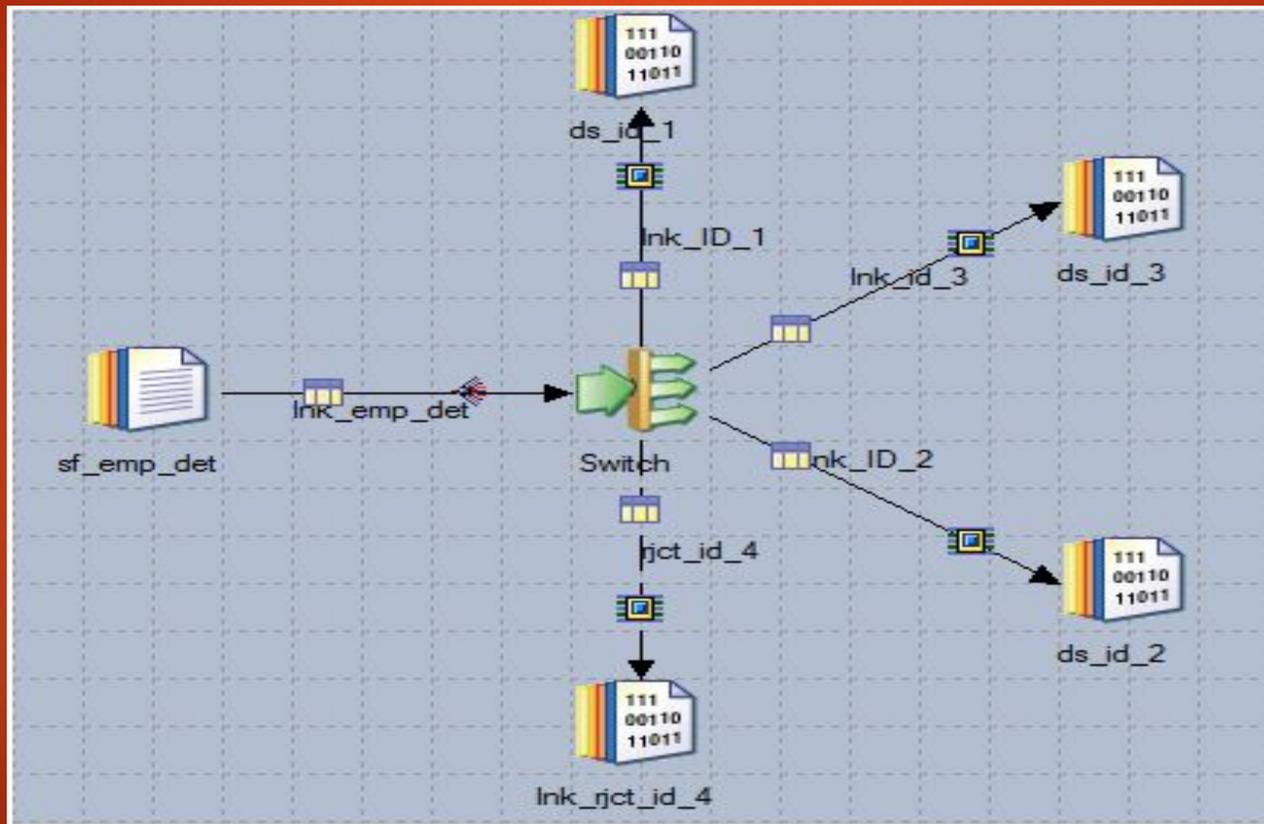
Reject Link

- ▶ Вы не можете изменить ссылки **Reject link**. Связь использует определения столбцов по ссылке отвергая строки данных.

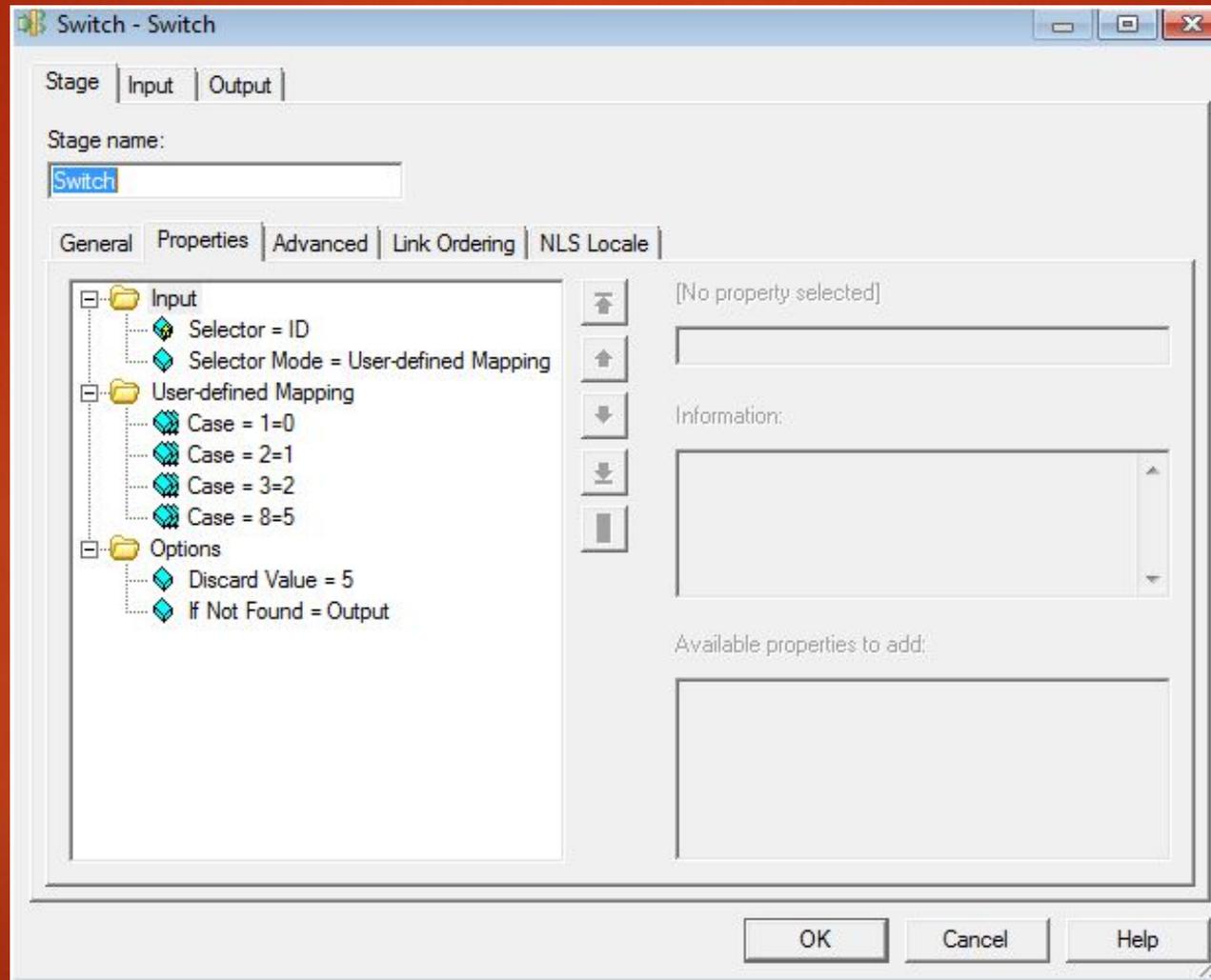
Рассмотрим данную таблицу для примера

ID	First Name	Last Name	Location	Network ID	EmailID
1	Jach	Simmons	LA	JS524	Letsc@gmail.com
2	Shumas	Jane	Las	Sj145	Jaene@ymail.com
3	Jonty	Waughn	Syd	JW927	JontyW@sdbh.com
3	Suhana	Safar	Mexico	SS99	Sas@gmail.com
8	Abc	Cde	Mex	ab254	lmn@gmail.com
4	Shumas	Jane	Las	Sj145	Jaene@ymail.com

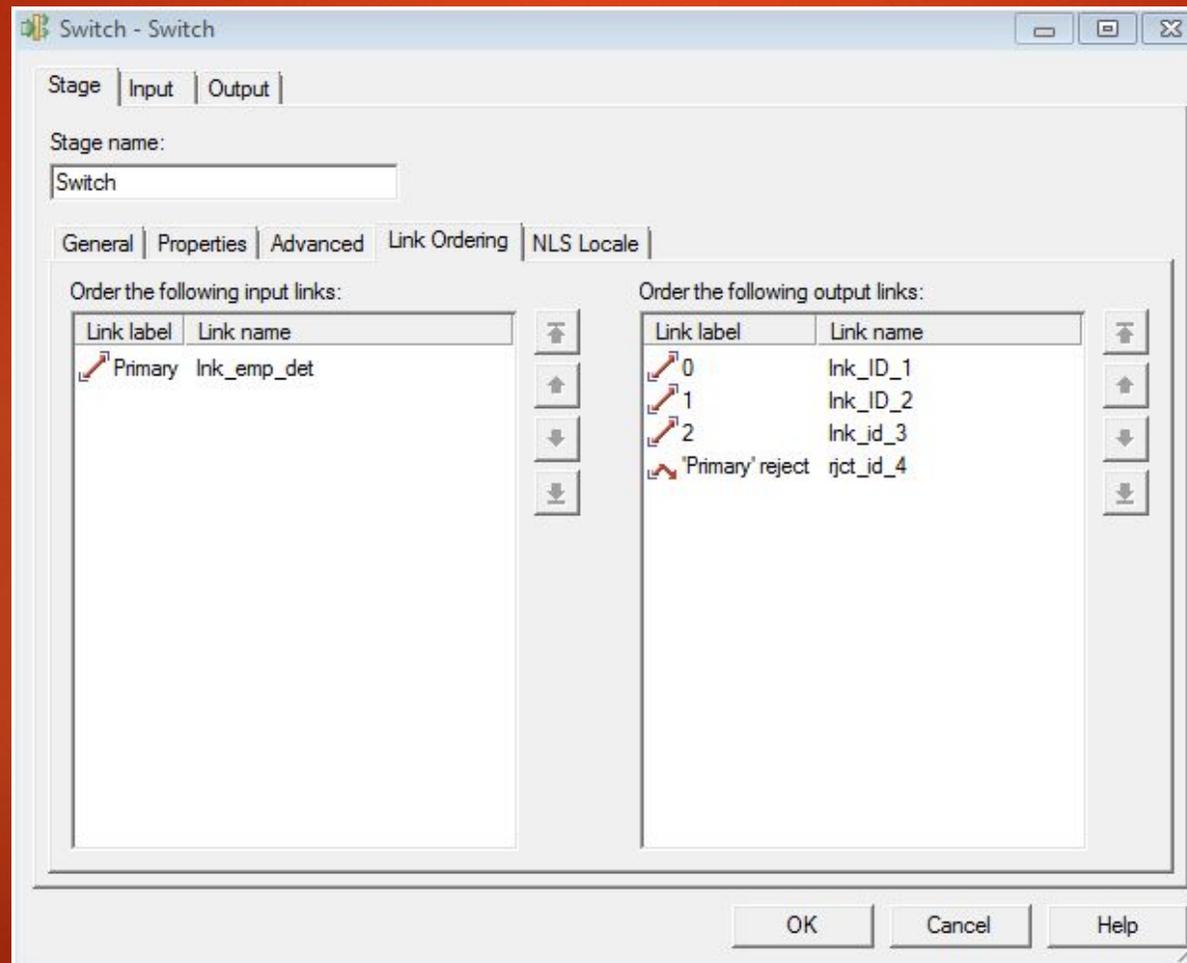
Создадим данную
структуру



Далее настроим конфигурации **Switch Stage**



Вкладка ссылка



Результат выполнения Switch Stage

1. Link **Ink_ID_1** for which 'Case=1=0'.

ID	FWName	Lname	Loc	LanID	Email
1	Jach	Simmons	LA	J5824	Jebec@gmail.com

2. Link **Ink_ID_2** for which 'Case=2=1'.

ID	FWName	Lname	Loc	LanID	Email
2	Shumas	Jane	Las	Sj148	Jeane@gmail.com

3. Link **Ink_id_3** for which 'Case=3=2'.

ID	FWName	Lname	Loc	LanID	Email
3	Suhana	Safar	Mexico	SS39	Sas@gmail.com
3	Jonny	Naught	Syd	J5927	Jonny@edbh.com

4. For Reject link **Ink_rjct_id_4** :As we didn't mention any case statement for ID=4, it will get route to the reject link.

ID	FWName	Lname	Loc	LanID	Email
4	Shumas	Jane	Las	Sj148	Jeane@gmail.com

5. For 'Case=8=5': Record for ID=8 get discarded.