Informatica Power Center. Нефункциональные требования к разработке потоков в IPC.

Оглавление

[1. Основные понятия. 3](#_Toc464055723)

[2. Правила разработки карты проливки в IPC Designer 3](#_Toc464055724)

[2.1 Разработка source 3](#_Toc464055725)

[3. Правила разработки карты проливки в IPC Workflow Manager 4](#_Toc464055726)

[4. Предварительные настройки при установке IPC клиента 5](#_Toc464055727)

[4.1 Предварительные настройки 5](#_Toc464055728)

[4.2 Установка IPC клиента 6](#_Toc464055729)

[5. Полезные скрипты 7](#_Toc464055730)

[5.1 Список скриптов, через которые можно мониторить проливки IPC 7](#_Toc464055731)

[5.2 Список скриптов, через которые можно мониторить проливки IPC 8](#_Toc464055732)

# Основные понятия.

IPC (Informatica Power Center) - система для разработки и поддержки ETL процессов от источника к базе-приемнику.

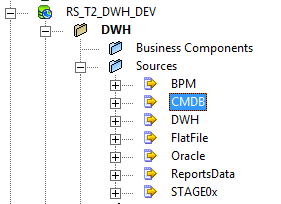
Power Center Designer - модуль в IPC, предназначенный для разработки карты проливки для потоков.

Power Center Workflow Manager - модуль в IPC, предназначенный для разработки самих потоков и настройки расписания

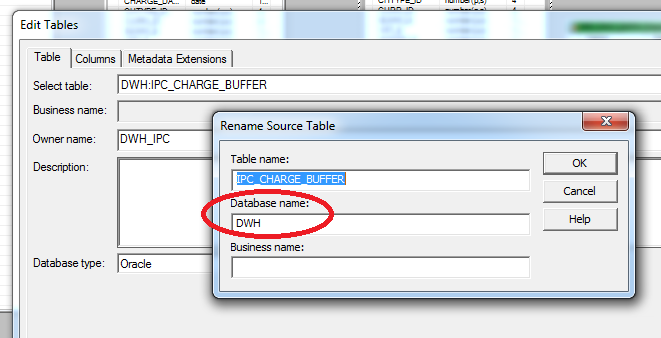
# Правила разработки карты проливки в IPC Designer

## 2.1 Разработка source

В IPC Designer все источники разбиты на базы:



Соответственно все источники сохраняем в ту базу, в которой они лежат:



# Правила разработки карты проливки в IPC Workflow Manager

При разработке потоков в WF необходимо придерживатся следующих правил:

* **Всегда ставим галочку «Fail parent if this task fail»** - когда галочка стоит, то Workflow всегда падает при падении сессии.
* **Не ставим галочку «Suspend on error»** - когда галочка стоит, то Workflow будет требовать ручной остановки и перезапуска. Т.е. при падении он становится unscheduled
* У каждого workflow-а log-файл должен соответствовать названию workflow-а
* У каждой сессии log-файл должен соответствовать названию сессии
* В разных папках не должно быть workflow-ов с одинаковыми названиями
* Не ставим галочку “Run on Integration Service initialization”
* Reject-ы при выполнении сессий недопустимы
* Сессии и workflow-ы должны запоминать не менее 5 последних логов загрузки. Настраивается в Edit Workflow- Properties
* Трансформации с TracingLevel, отличным от Normal, на боевом репозитории недопустимы
* Все Workflow при запуске расписания, должны использовать штатные объект Schedulers
* Все сессии должны начинаться с префикса “s\_M\_”
* Все Workflow должны начинаться с префикса “WF\_”
* При разработке маппингов необходимо использовать единый префикс. Все маппинги должны начинаться с “M\_”

**!!! Длительность отработки каждой сессии в потоке не должна превышать 1 часа. Любую заливку данных с большим объемом необходимо разбивать на инкременты. В противном случае есть вероятность потребить большое кол-во UNDO или TEMP, и нарушить отработку других процессов в базе.**

Для разработки заливки потоков в DWH из разных источников используется репозиторий RS\_T2\_DWH\_DEV. Соответственно для тестовых проливок используется схема DWH\_DEV. Миграция потоков в репозиторий DWH\_PROD выполняется стандартными средствами импорт/экспорт объектов в Repository Manger. И при этом схема на которую проливаются таблицы, переключается в DWH\_IPC. Для этого в потоке должен быть создан парамет $$TARGET\_TABLE, который должен получать название схемы из файла параметров или из настроечной таблицы.

Для организации последовательного запуска и исключения высокой нагрузки необходимо использовать объект таймер. Задержка используется в таймере и берется из таблицы. Cледующий lag проставляется в таблице select \* from dwh\_ipc.ipc\_tables\_timer.

Также для уменьшения нагрузки на сервер большим количеством Workflow можно использовать создание Worklet-ов.

При постановке потоков на расписание в базе создан набор преднастроенных scheduler-ов, которые необходимо использовать. Их список указан внизу. Название соответствует периодичности запуска (список пополняемый):

Run\_WF\_every\_15\_minutes

Run\_WF\_Every\_1\_Day\_03

Run\_WF\_Every\_1\_Day\_13\_00

Run\_WF\_every\_1\_hour

Run\_WF\_every\_1\_minute

Run\_WF\_every\_30\_minutes

Run\_WF\_every\_5\_minutes

Run\_WF\_every\_month\_04

Run\_WF\_every\_week\_mon

# Предварительные настройки при установке IPC клиента

## Предварительные настройки

Для получения доступа к IPC необходимо обратиться к администраторам IPC.

Дистрибутив IPC клиента вложен на ресурсе [\\corp.tele2.ru\folders \Software\Informatica\IPC client\IPC Client 9.6.1](file:///\\corp.tele2.ru\folders%20\Software\Informatica\IPC%20client\IPC%20Client%209.6.1).

При установке клиента IPC необходимо прописать ip адрес серверов информатики в  файлик hosts . Файл расположен в директории C:\Windows\System32\drivers\etc.

10.0.2.20 node01

10.0.2.22 node02

10.77.251.99 tnode01

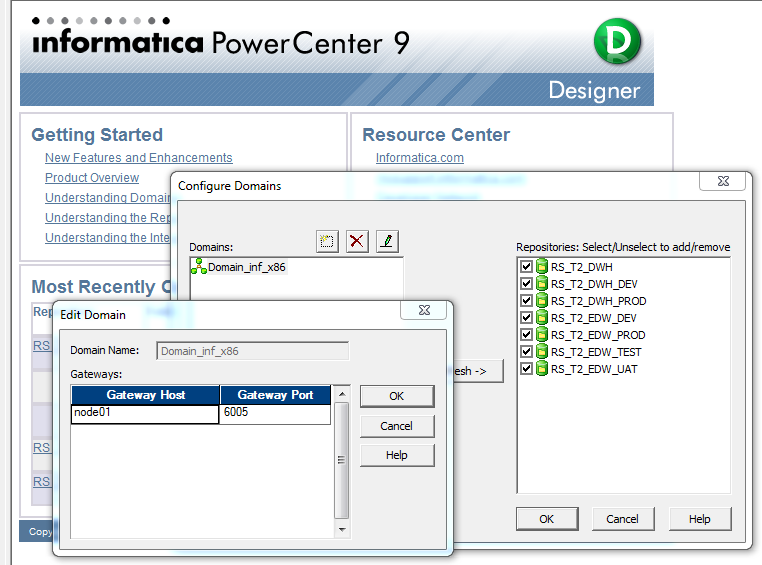
При первом входе необходимо сконфигурировать домен к подключению. Для этого нажимаем Repository – Configure Domain , прописываем параметры и выбираем репозитории к которым предоставил доступ администратор.

Host: node01

Port: 6005

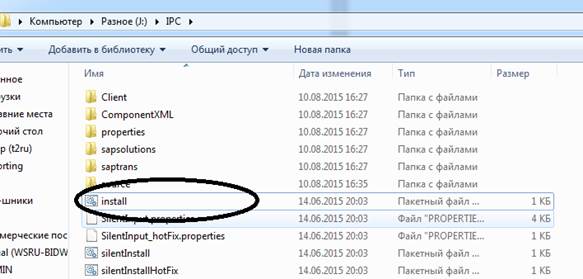
Domain: Domain\_inf\_x86

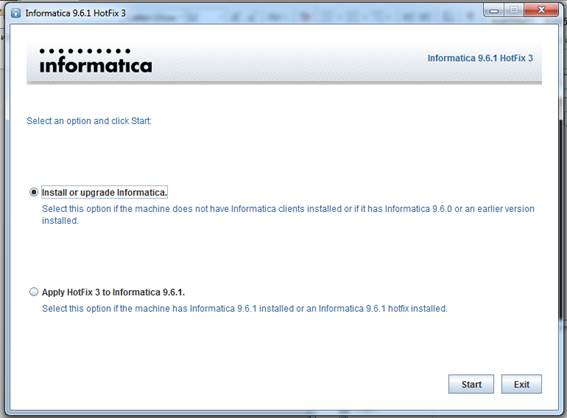
Также администратор предоставит вам доступ к папкам, в которых можно вести разработки.



## Установка IPC клиента

Запускаем пакетный установщик  install.bat в папке дистрибутива.





Выбираем установку.  Жмем несколько раз Next, и  устанавливаем IPC.

После этого обновляем до 9.6.1.

# Полезные скрипты

## Технические учетки

Все объекты метаданных IPC лежат в базе DWH в соответствующих схемах:

|  |  |
| --- | --- |
| INFA\_REP\_X86\_DWH | тестовый репозиторий. Используется для административных целей. Не для разработки потоков. |
| INFA\_REP\_X86\_DWH\_PROD | схема продуктивного репозитория DWH (RS\_T2\_DWH) на x86 |
| INFA\_REP\_X86\_DWH\_DEV | схема разработческого репозитория проекта DWH ( RS\_T2\_DWH\_DEV) на x86 - схема тестового репозитория проекта EDW ( RS\_T2\_EDW\_DEV) на x86 |
| INFA\_REP\_X86\_EDW\_DEV | схема разработческого репозитория проекта EDW ( RS\_T2\_EDW\_DEV) на x86 |
| INFA\_REP\_X86\_EDW\_TEST | схема тестового репозитория проекта EDW ( RS\_T2\_EDW\_DEV) на x86 |
| INFA\_REP\_X86\_EDW\_UAT | схема UAT репозитория проекта EDW ( RS\_T2\_EDW\_DEV) на x86 |
| DWH\_DEV | схема для разработки потоков в DWH\_DEV |
| INFORMATICA | техническая учетка для проливок из DWH в продуктивном контуре в RS\_T2\_DWH\_PROD |
| INFORMATICA\_RO | техническая учетка для проливок из DWH конкретно для EDW репозиториев в RS\_T2\_EDW\_PROD |
| INFORMATICA\_DEV | техническая учетка для проливок из DWH конкретно для DEV контура |
| DWH\_IPC | продуктивная схема для переноса потоков в DWH\_PROD |
| INFORMATICA\_DWH\_ROLE | Роль объединяющая INFORMATICA и INFORMATICA\_DEV.  **!!! При разработке потоков в RS\_T2\_DWH\_DEV гранты предоставляем этой роли.** |

При разработке потоков, если в качестве источника или приемника используется база DWH, то для заливки данных гранты на DDL операции раздаются схемам и ролям INFORMATICA\_DWH\_ROLE (для репозиториев DWH) и INFORMATICA\_RO (для EDW репозиториев).

## Список скриптов, через которые можно мониторить проливки IPC

Лог отработки потоков. Вместо INFA\_REP\_X86\_DWH\_PROD подставляется название нужной схемы репозитория.

*--1 Проверка отвалов прогрузок потоков WF IPC*

SELECT WORKFLOW\_NAME,

SERVER\_NAME,

RUN\_ERR\_CODE,

RUN\_ERR\_MSG,

SUBJECT\_AREA,

START\_TIME,

END\_TIME,

decode(RUN\_STATUS\_CODE,

1,'Succeeded',

2,'Disabled',

3,'Failed',

4,'Stopped',

5,'Aborted',

6,'Running',

15,'Terminated',

'Error. This Status is Unknown') Status

from INFA\_REP\_X86\_DWH\_PROD.REP\_WFLOW\_RUN v

where START\_TIME >= trunc(sysdate)

and v.WORKFLOW\_NAME='WF\_IPC\_PAY\_DOC'

*-- and END\_TIME <= sysdate*

*-- and v.RUN\_STATUS\_CODE != 1--NOT Succeeded*

*--2 Проеврка отработки сессий потоков WF IPC*

select t.ACTUAL\_START, t.RUN\_STATUS\_CODE, t.\*

from INFA\_REP\_X86\_DWH\_PROD.REP\_SESS\_LOG t

where*-- t.SESSION\_instanct\_NAME like '%s\_M\_IPC\_PAY\_DOC\_STAGE04%'*

*-- and*

t.SESSION\_INSTANCE\_NAME = 's\_M\_IPC\_PAY\_DOC\_STAGE04'

and

t.ACTUAL\_START > TRUNC(sysdate )

*--and t.RUN\_STATUS\_CODE != 1*

order by t.ACTUAL\_START;

*--3 Конфигурационная таблица , содержащая параметры загрузки данных.*

select \* from dwh\_ipc.ctl\_table\_parameters\_list

select \* from dwh\_dev.ctl\_table\_parameters\_list

*--4* Пакеты предназначенные для выполнения различных DDL-операций с target-таблицами. Например, если перед заливкой партиционированой таблицы необходимо выполнить добавление партиции или очистка методом truncate, то для этого используются соответствующие процедуры пакетах ниже. Также если есть необходимость в написании новой DDL-операции, то эти процедуры добавляются также сюда. Изменения в пакетах должны происходить в следующем порядке сначала в dwh\_dev потом в dwh\_ipc.

dwh\_dev.pkg\_ipc\_etl

dwh\_ipc.pkg\_ipc\_etl