# [Создание атомов для независимых объектов](http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=211218665)

* [Организационные нюансы](http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=211218665#id-%D0%A1%D0%BE%D0%B7%D0%B4%D0%B0%D0%BD%D0%B8%D0%B5%D0%B0%D1%82%D0%BE%D0%BC%D0%BE%D0%B2%D0%B4%D0%BB%D1%8F%D0%BD%D0%B5%D0%B7%D0%B0%D0%B2%D0%B8%D1%81%D0%B8%D0%BC%D1%8B%D1%85%D0%BE%D0%B1%D1%8A%D0%B5%D0%BA%D1%82%D0%BE%D0%B2-%D0%9E%D1%80%D0%B3%D0%B0%D0%BD%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D0%BE%D0%BD%D0%BD%D1%8B%D0%B5%D0%BD%D1%8E%D0%B0%D0%BD%D1%81%D1%8B)
* [Состав атомов](http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=211218665#id-%D0%A1%D0%BE%D0%B7%D0%B4%D0%B0%D0%BD%D0%B8%D0%B5%D0%B0%D1%82%D0%BE%D0%BC%D0%BE%D0%B2%D0%B4%D0%BB%D1%8F%D0%BD%D0%B5%D0%B7%D0%B0%D0%B2%D0%B8%D1%81%D0%B8%D0%BC%D1%8B%D1%85%D0%BE%D0%B1%D1%8A%D0%B5%D0%BA%D1%82%D0%BE%D0%B2-%D0%A1%D0%BE%D1%81%D1%82%D0%B0%D0%B2%D0%B0%D1%82%D0%BE%D0%BC%D0%BE%D0%B2)
  + [AtomHeader](http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=211218665#id-%D0%A1%D0%BE%D0%B7%D0%B4%D0%B0%D0%BD%D0%B8%D0%B5%D0%B0%D1%82%D0%BE%D0%BC%D0%BE%D0%B2%D0%B4%D0%BB%D1%8F%D0%BD%D0%B5%D0%B7%D0%B0%D0%B2%D0%B8%D1%81%D0%B8%D0%BC%D1%8B%D1%85%D0%BE%D0%B1%D1%8A%D0%B5%D0%BA%D1%82%D0%BE%D0%B2-AtomHeader)
  + [SqoopImport](http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=211218665#id-%D0%A1%D0%BE%D0%B7%D0%B4%D0%B0%D0%BD%D0%B8%D0%B5%D0%B0%D1%82%D0%BE%D0%BC%D0%BE%D0%B2%D0%B4%D0%BB%D1%8F%D0%BD%D0%B5%D0%B7%D0%B0%D0%B2%D0%B8%D1%81%D0%B8%D0%BC%D1%8B%D1%85%D0%BE%D0%B1%D1%8A%D0%B5%D0%BA%D1%82%D0%BE%D0%B2-SqoopImport)
  + [HiveUpdate](http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=211218665#id-%D0%A1%D0%BE%D0%B7%D0%B4%D0%B0%D0%BD%D0%B8%D0%B5%D0%B0%D1%82%D0%BE%D0%BC%D0%BE%D0%B2%D0%B4%D0%BB%D1%8F%D0%BD%D0%B5%D0%B7%D0%B0%D0%B2%D0%B8%D1%81%D0%B8%D0%BC%D1%8B%D1%85%D0%BE%D0%B1%D1%8A%D0%B5%D0%BA%D1%82%D0%BE%D0%B2-HiveUpdate)
  + [AtomTail](http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=211218665#id-%D0%A1%D0%BE%D0%B7%D0%B4%D0%B0%D0%BD%D0%B8%D0%B5%D0%B0%D1%82%D0%BE%D0%BC%D0%BE%D0%B2%D0%B4%D0%BB%D1%8F%D0%BD%D0%B5%D0%B7%D0%B0%D0%B2%D0%B8%D1%81%D0%B8%D0%BC%D1%8B%D1%85%D0%BE%D0%B1%D1%8A%D0%B5%D0%BA%D1%82%D0%BE%D0%B2-AtomTail)
  + [Пример](http://confluence.moscow.alfaintra.net/pages/viewpage.action?pageId=211218665#id-%D0%A1%D0%BE%D0%B7%D0%B4%D0%B0%D0%BD%D0%B8%D0%B5%D0%B0%D1%82%D0%BE%D0%BC%D0%BE%D0%B2%D0%B4%D0%BB%D1%8F%D0%BD%D0%B5%D0%B7%D0%B0%D0%B2%D0%B8%D1%81%D0%B8%D0%BC%D1%8B%D1%85%D0%BE%D0%B1%D1%8A%D0%B5%D0%BA%D1%82%D0%BE%D0%B2-%D0%9F%D1%80%D0%B8%D0%BC%D0%B5%D1%80)

Организационные нюансы

На вас задача упадет либо пиьсмом от аналитика, либо в jira.

После этого необходимо создать ветку в репозитории:

<http://git.moscow.alfaintra.net/projects/DATALAKE/repos/etl/browse>

Сделать разработку и закомитить все в свою ветку

По оканчании разработки сделать пулл реквест в develop.

Состав атомов

Коллеги, атомный поток состоит из 4-ех компонентов:

1. AtomHeader - java action, который записывает метаинформацию
2. SqoopImport - sqoop action, который импортит данные
3. HiveUpdate - hive action, который обновляет данные в целевой таблице
4. AtomTail - java action, который записывает метаифнормацию и в перспективе будет обновлять метаданные Impala

AtomHeader

<java>  
       <job-tracker>${jobTracker}</job-tracker>   
       <name-node>${nameNode}</name-node>   
       <main-class>[ru.alfa.bi](http://ru.alfa.bi/).hdp.orchestration.AtomHeader</main-class>   
       <java-opts>-DATOM\_NAME=wf\_dmacrm\_productgroup\_sdim -DCTL=wf\_ctl\_dmacrm\_daily\_dictionaries -DREG=wf\_reg\_dmacrm\_daily -DDB\_CONNECTION=${oracle\_jdbc\_url} -DDB\_USER=${oracle\_jdbc\_user} -DDB\_PASSWORD=${oracle\_jdbc\_password}</java-opts> здесь заполняете первые три параметра соответственно  
       <file>/user/${user\_name}/libs/generator-1.0-SNAPSHOT-jar-with-dependencies.jar#generator-1.0-SNAPSHOT-jar-with-dependencies.jar</file>  
       <file>/user/${user\_name}/libs/ojdbc6.jar#ojdbc6.jar</file>  
</java>

SqoopImport

<sqoop xmlns="uri:oozie:sqoop-action:0.2">  
<job-tracker>${jobTracker}</job-tracker>   
<name-node>${nameNode}</name-node>   
<arg>import</arg>   
<arg>--connect</arg>    
<arg>jdbc:oracle:thin:@(DESCRIPTION = (ADDRESS = (PROTOCOL = TCP)(HOST = exa2-scan)(PORT = 1521))(ADDRESS = (PROTOCOL = TCP)(HOST = exa1-scan)(PORT = 1521))(CONNECT\_DATA = (SERVER = DEDICATED)(SERVICE\_NAME = RBPROD\_TAF)))</arg> - прописываете коннект  
<arg>--username</arg>  
<arg>${ORA\_USER\_RBPROD\_TAF}</arg>  
<arg>--password</arg>  
<arg>${ORA\_PASSWORD\_RBPROD\_TAF}</arg>  
<arg>--compression-codec</arg>  
<arg>[org.apache.hadoop.io](http://org.apache.hadoop.io/).compress.SnappyCodec</arg>  
<arg>--as-parquetfile</arg>  
<arg>--delete-target-dir</arg>  
<arg>--target-dir</arg>  
<arg>${PATH\_TO\_TMP}/rbprod\_taf/productgroup\_sdim\_tmp</arg>  
<arg>--query</arg>  
<arg>SELECT JOB\_INSERT , JOB\_UPDATE , to\_char(AS\_OF\_DAY, 'yyyy-MM-dd') as AS\_OF\_DAY, EMIX , DELETED\_FLAG , DEFAULT\_FLAG , UK , PARENT\_UK , PRODUCTGROUPTYPE\_UK , CCODE , NAME FROM dmacrm.productgroup\_sdim WHERE $CONDITIONS</arg> прописываете селект, все атрибуты с типом данных DATE должны быть обернуты в  to\_char(ИМЯ АТРИБУТА, 'yyyy-MM-dd')  
<arg>-m</arg>  
<arg>1</arg>   
<arg>--map-column-java</arg>  
<arg>JOB\_INSERT=Integer,JOB\_UPDATE=Integer,EMIX=Integer,UK=Integer,PARENT\_UK=Integer,PRODUCTGROUPTYPE\_UK=Integer</arg> мапаете все нестринговые поля.   
</sqoop>

HiveUpdate

<action name="hive2-productgroup\_sdim" cred="hive2">  
<hive2 xmlns="uri:oozie:hive2-action:0.1">  
<job-tracker>${jobTracker}</job-tracker>  
<name-node>${nameNode}</name-node>  
<jdbc-url>${hive2\_jdbc\_url}</jdbc-url>  
<script>hive/hive\_load\_no\_part.hql</script> меняете в зависимости от типа update  
<param>DATABASE=${DATABASE}</param>  
<param>TABLE\_NAME=productgroup\_sdim</param>  меняете в зависимости от имени объекта  
<param>TMP\_DATABASE=${TMP\_DATABASE}</param>  
</hive2>  
<ok to="atom-tail-productgroup\_sdim"/>  
<error to="Kill"/>  
</action>

AtomTail

<action name="atom-tail-productgroup\_sdim">  
<java>  
<job-tracker>${jobTracker}</job-tracker>  
<name-node>${nameNode}</name-node>  
<main-class>[ru.alfa.bi](http://ru.alfa.bi/).hdp.orchestration.AtomTail</main-class>  
<java-opts>-DATOM\_NAME=wf\_dmacrm\_productgroup\_sdim -DDB\_CONNECTION=${oracle\_jdbc\_url} -DDB\_USER=${oracle\_jdbc\_user} -DDB\_PASSWORD=${oracle\_jdbc\_password}</java-opts> здесь заполняете первые три параметра соответственно  
<file>/user/${user\_name}/libs/generator-1.0-SNAPSHOT-jar-with-dependencies.jar#generator-1.0-SNAPSHOT-jar-with-dependencies.jar</file>  
<file>/user/${user\_name}/libs/ojdbc6.jar#ojdbc6.jar</file>  
</java>  
<ok to="End"/>  
<error to="Kill"/>  
</action>

Пример

[workflow.xml](http://confluence.moscow.alfaintra.net/download/attachments/211218665/workflow.xml?version=1&modificationDate=1533208322000&api=v2)