

报表数据源

集算器 · 润乾软件出品





报表数据源面临的问题



复杂逻辑SQL难写

- SQL无序，集合化不彻底，过程化计算困难
- 很多项目不允许报表存储过程
- 语法不通用，难以移植



多样数据源不好处理

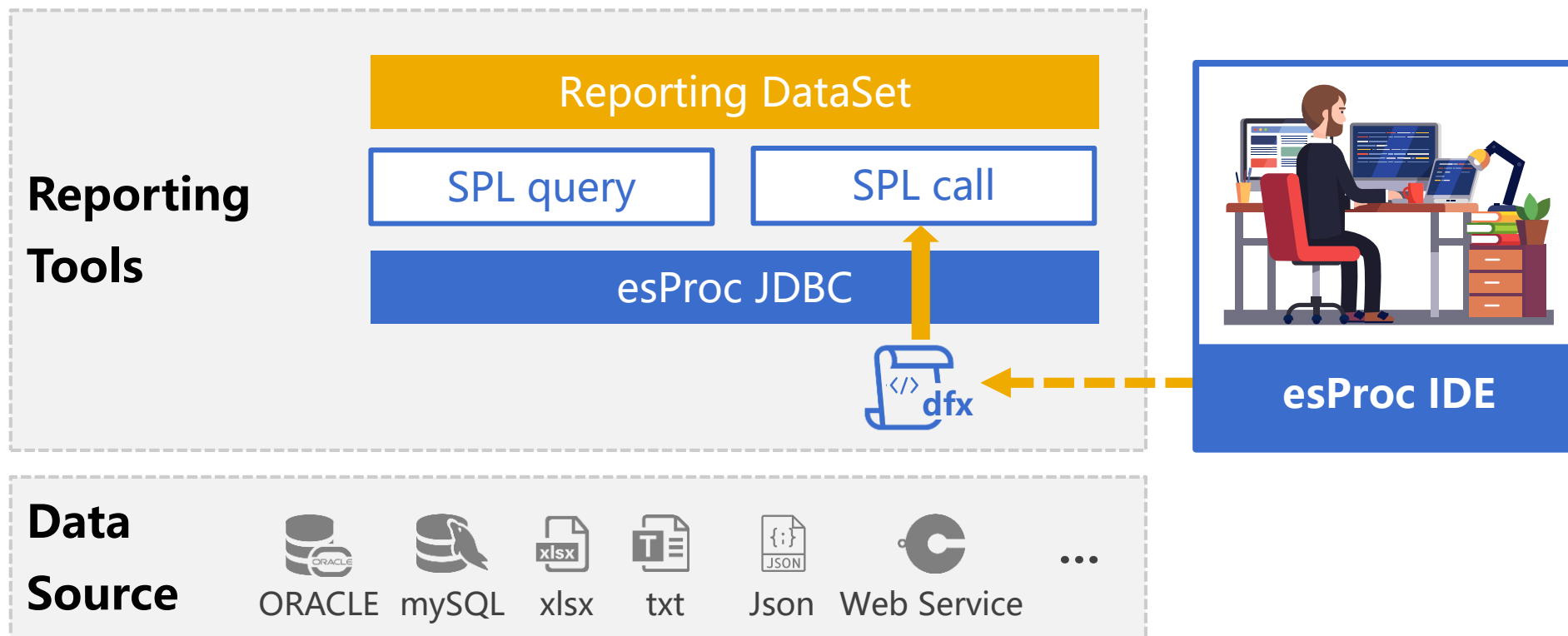
- 特殊数据源如Excel、NoSQL
- 跨源跨库
- 有些报表工具只支持单数据源



自定义数据源(Java)更难写

- 缺乏结构化类库，代码繁琐
- 耦合性高，维护困难

引入计算层 — 集算器



说明：SPL是esProc采用的脚本语言，dfx为脚本文件，相当于外置存储过程，报表通过JDBC接口传入SPL语句实现结构化计算或执行脚本。



简化复杂SQL/存储过程

? 某支股票最长连续涨了多少交易日

```
1 select max(continuousDays)-1
2 from (select count(*) continuousDays
3       from (select sum(changeSign) over(order by tradeDate) unRiseDays
4             from (select tradeDate,
5                   case when closePrice>lag(closePrice) over(order by tradeDate)
6                       then 0 else 1 end changeSign
7                   from stock) )
8       group by unRiseDays)
```

	A
1	=stock.sort(tradeDate)
2	=0
3	=A1.max(A2=if(closePrice>closePrice[-1],A2+1,0))

SPL

语法体系更容易描述人的自然思维!

SQL



思考：按照自然思维怎么做？



➤ 算法外置减少存储过程

采用存储过程实现数据准备算法，会造成报表与数据库的耦合问题

- 存储过程和报表的存放位置不同，导致对应难度很大
- 存储过程修改需要分配相应的数据库权限，存在安全隐患
- 存储过程容易被其他应用使用，造成多个应用间的耦合
- 很多项目出于安全和权限的目的，不允许为报表建立存储过程

使用集算器替代存储过程完成报表数据准备，**会极大减少存储过程**，算法外置后与报表模板一起存放管理，完全归属于应用本身，**降低报表与应用其他部分或其他应用的耦合**

避免复杂JAVA编程



集算器采用**集合化**语法，代码要比没有直接提供结构化计算的JAVA更加**短小**

写的更快更短

- 集算器基于Java提供了更高层的类库和方法

容易理解和排错

- 伪实代码的比例大约是只有1:1.5，大多数报表数据准备算法可以在**一个屏幕内**显示出来
- 一个页面内能看到更多代码，能更完整地理解代码的含义与排错



解释执行降低应用耦合度

使用JAVA和集算器准备报表数据源会有以下不同：

JAVA

模块化困难

Java程序必须和主应用一起编译打包，耦合度高

难以热切换

Java编写的报表数据准备算法有修改后会导致整个应用重新编译部署，很难做到热切换

缺乏类库

JAVA缺乏结构化和半结构化数据计算的类库，所有算法都需要硬编码

集算器

模块化简单

集算器脚本文件可以和报表模板一起管理维护，从而使报表功能模块化

容易热切换

集算器是解释执行的语言，很容易做到热切换

类库丰富

更丰富的语法和类库，让结构化数据的计算更有效率



多样性数据源支持

报表工具的计算能力难以胜任多样性数据源

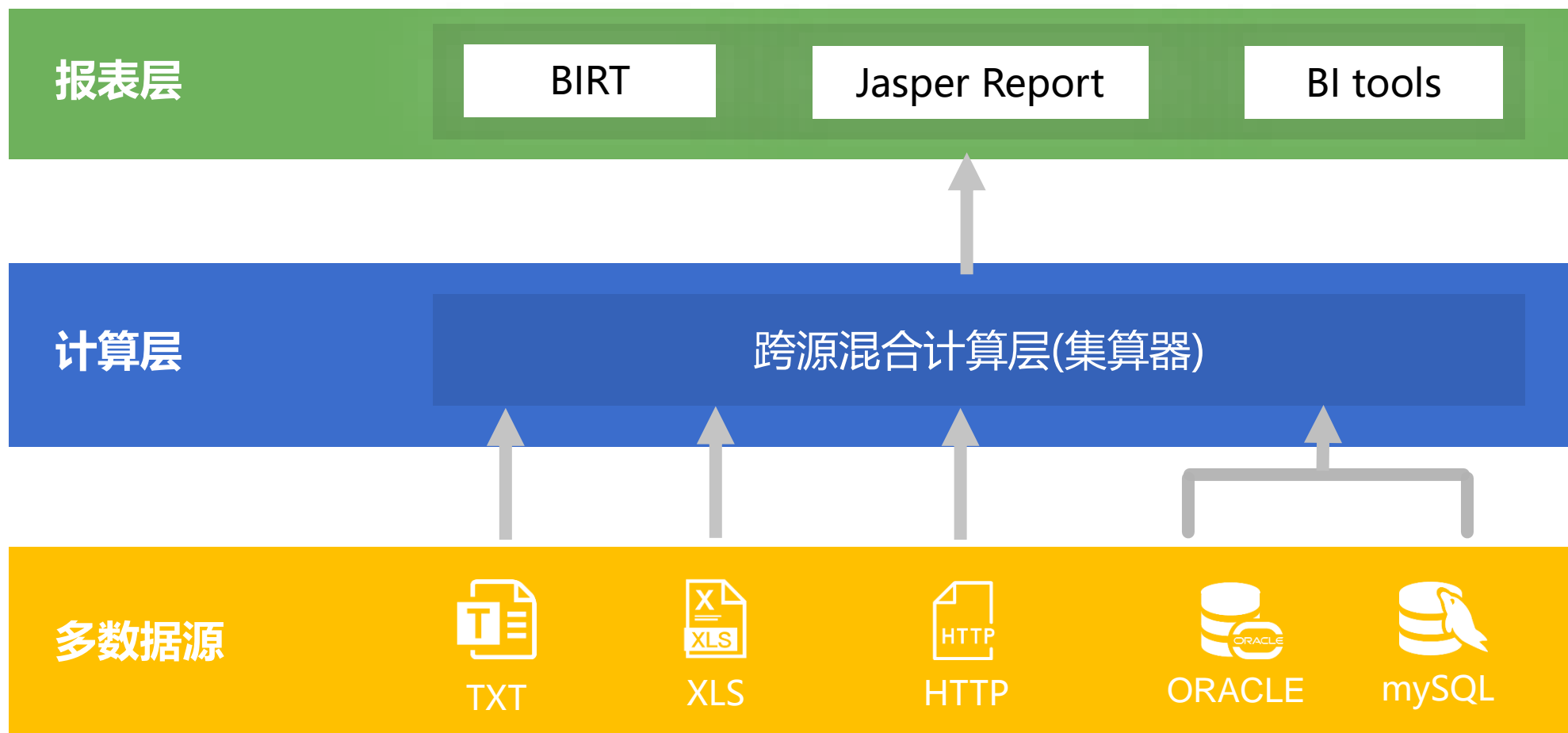
计算层处理多样性数据源

- RDBMS: Oracle、MS SQL Server、DB2、Informix、MySQL、PostgreSQL
- NOSQL: MongoDB、Redis、Cassandra、ElasticSearch
- Hadoop: HDFS、HIVE、HBase
- 应用软件: SAP ECC、BW
- 文件: Excel、Json、XML、TXT
- 其他: Http Restful、Web Services、OLAP4j、...





直接使用多数据源及跨库计算





动态数据源/集

动态数据源

根据参数动态切换数据库 `${pds}.query("select * from T where F=?",pF)`

动态数据集

动态生成算法，动态拼接SQL

通用代码实现算法，移植时无需修改

	A	
1	<code>=sums.array().("sum("+~+"") as "+~).string()</code>	/把a,b变成sum(a) as a,sum(b) as b
2	<code>=db.query("select G,"+A1+" from T group by G")</code>	

结果集容量控制

	A	B	
1	<code>=db.cursor("select * from T")</code>	<code>=A1.fetch(1000)</code>	
2	<code>if B1.fetch@0(1)</code>	<code>>B1.insert(0,"继续")</code>	/未完成则插入标记
3	<code>>A1.close()</code>	<code>return B1</code>	



并行取数提升性能

数据库JDBC性能较差，报表性能又严重依赖于取数环节；集算器可以采用多线程并行的方式同时建立多个数据库连接从数据库分段取数，可以获得数倍性能提升

	A	B	C
1	fork 4	=connect(db)	/分4线程，要分别建立连接
2		=B1.query@x("select * from T where part=?" ,A1)	/分别取每一段
3	=A1.conj()		/合并结果

集算器作为BIRT数据源-动态解析CSV



BIRT representation

Showing page 1 of 1

dataset_date	measurement	managedelement	content	last_minute_cpu	last_5_minute_cpu	last_30_minute_cpu	last_hour_cpu
2015-05-26 21:45:00.0	cpqHoCpuUtilTable	nltcom04	20:7:5:5	20	7	5	5
2015-05-26 21:45:00.0	cpqHoCpuUtilTable	nltcom04	17:8:5:5	17	8	5	5
2015-05-26 21:45:00.0	cpqHoCpuUtilTable	nltcom01	20:4:7:7	20	4	7	7
2015-05-26 21:45:00.0	cpqHoCpuUtilTable	nltcom03	8:12:5:5	8	12	5	5
2015-05-26 21:45:00.0	cpqHoCpuUtilTable	nltcom02	11:5:5:6	11	5	5	6

esProc calculation

	A
1	=file("D:\DummyData.csv").import@t(dataset_date,measurement,managedelement,content:string;";")
2	=file("D:\Mapping.csv").import@t(,";")
3	=A2(1).content.split(";")
4	=A3.to(2).("cols("+string(#+1)+"):"+~).string()
5	=A1.derive((cols=string(content).split(";"))(1):\${A3(1)},\${A4})

Data sources

DummyData.csv

```
1 dataset_date,measurement,managedelement,content
2 2015-05-26 21:45:00.0,cpqHoCpuUtilTable,nltcom04,20:7:5:5
3 2015-05-26 21:45:00.0,cpqHoCpuUtilTable,nltcom04,17:8:5:5
4 2015-05-26 21:45:00.0,cpqHoCpuUtilTable,nltcom01,20:4:7:7
5 2015-05-26 21:45:00.0,cpqHoCpuUtilTable,nltcom03,8:12:5:5
6 2015-05-26 21:45:00.0,cpqHoCpuUtilTable,nltcom02,11:5:5:6
```

Mapping.csv

```
1 managedelement,content
2 nltcom,last_minute_cpu_utilization_kpi:last_5_minute_cpu_utilization_kpi:last_30_minute_cpu_utilization_kpi:last_hour_cpu_utilization_kpi
```



集算器作为JasperReport数据源-跨库关联

Jasper representation

OID	EID	NAME	DEPT	AMOUNT
10262	8	Megan	Marketing	583.199999392
10265	2	Ashley	Finance	1170.0
10268	8	Megan	Marketing	1098.0
10276	8	Megan	Marketing	400.0

esProc calculation
Cross-database associative filtering

	A
1	=mysql.query("select OID, EID, AMOUNT from orders")
2	=pg.query("select EID,NAME,DEPT from employee")
3	=A1.switch(EID,A2:EID)
4	=A3.select(["Marketing","Finance"].pos(EID.DEPT))
5	=A4.new(OID,EID.EID:EID,EID.NAME:NAME,EID.DEPT:DEPT,AMOUNT)

Data sources

MySQL table:orders(partial field)

OID	EID	AMOUNT
10248	5	428.0
10249	6	1842.0
10250	4	1523.4999898076...
10251	3	624.94999976083...

PG table:employee(partial field)

EID	NAME	DEPT
1	Rebecca	R&D
2	Ashley	Finance
3	Rachel	Sales
4	Emily	HR

外部JAVA程序通过JDBC调用集算器脚本



JDBC class stored procedure calls SPL script file

```
...
Connection con = null;
Class.forName("com.esproc.jdbc.InternalDriver");
con= DriverManager.getConnection("jdbc:esproc:local://");
// Calling stored procedures , CountName is the file name of dfx
st =(com. esproc.jdbc.InternalCStatement)con.prepareCall("call
CountName()");
// Execute stored procedures
st.execute();
//Get result set
ResultSet rs = st.getResultSet();
...
```

JDBC query files directly using SQL

```
...
Connection con = null;
Class.forName( "com.esproc.jdbc.InternalDriver" );
con= DriverManager.getConnection( "jdbc:esproc:local://" );
// Calling stored procedures , CountName is the file name of dfx
st =(com. esproc.jdbc.InternalCStatement)con.createStatement();
//Query files using SQL,get result set
ResultSet rs = st.executeQuery( "$select name,count(*) from
/home/user/duty.txt group by name");
...
```



资源



- 常用计算的SPL代码 <http://doc.raqsoft.com.cn/esproc/coderefer/>
- 在 Birt 中实现交叉表的动态分组 <http://c.raqsoft.com.cn/article/1560936881610>
- Birt 如何实现不规则月份统计 <http://c.raqsoft.com.cn/article/1552378480299>
- Jasper 等工具制作报表的几个难题 <http://c.raqsoft.com.cn/article/1554711055465>
- 集算器助力 ireport <http://c.raqsoft.com.cn/article/1560233466960>
- 解决 BIRT 报表特殊布局的若干示例 <http://c.raqsoft.com.cn/article/1573958206309>
- 解决 BIRT 动态数据源的若干示例 <http://c.raqsoft.com.cn/article/1576544515636>
- BIRT 调用 SPL 脚本 <http://c.raqsoft.com.cn/article/1544076096840>
- JasperReport 调用 SPL 脚本 <http://c.raqsoft.com.cn/article/1544077307265>
- 安装使用、免费授权 <http://c.raqsoft.com.cn/article/1571895350771>