**00-TD性能解析入库程序开发文档**

**浪潮通信信息系统有限公司文档修订记录**

|  |  |  |  |
| --- | --- | --- | --- |
| **版本** | **日期** | **更改人** | **描述（注明修改的条款或页）** |
| V0.5 | 2013-1-22 | 武玉刚 | 文档创建 |
|  |  |  |  |
|  |  |  |  |

# 概述

**程序名称：**fpf-xml2db-td-pm

**程序用途：**该程序用于实现TD-SCDMA各厂家性能数据的采集、解析、入库工作。

**开发语言：**J2SE-1.5

**开发工具：**Eclipse

**源码位置：**https://218.57.135.34/develop/iNOS/branch/00-部门工作/优化分析产品部-平台产品部/02-独立工具/16-工具-fpf-xml2db-td-pm/源码

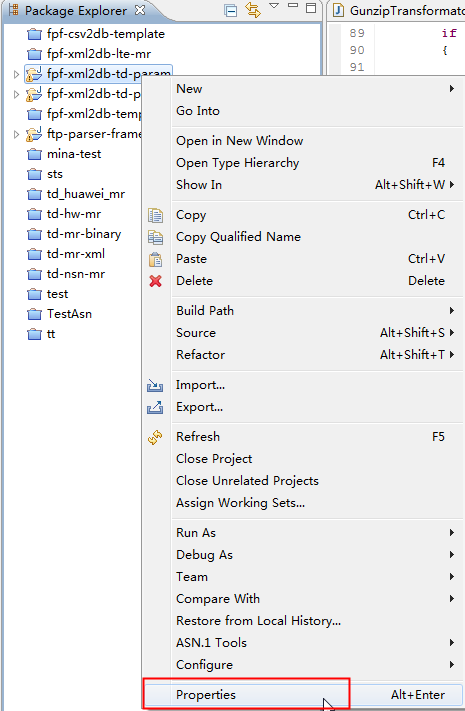
# 依赖条件

该程序依赖于FPF框架(ftp-parser-framework)开发实现，开发之前首先要下载FPF框架并在eclipse中将其添加为该程序的依赖项目，步骤如下：

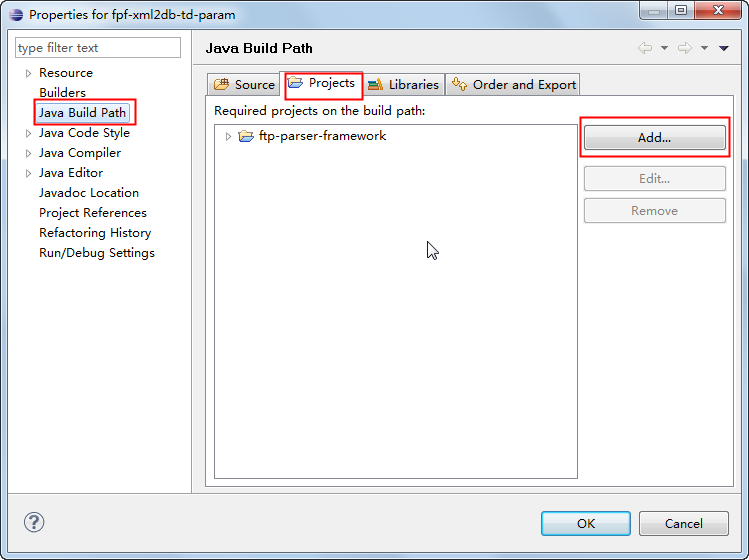
1. **从svn上下载FPF框架代码并建立工程。**

FPF源码位置：https://218.57.135.34/develop/iNOS/branch/00-部门工作/优化分析产品部-平台产品部/02-独立工具/10-工具-FtpParserFramework/源码

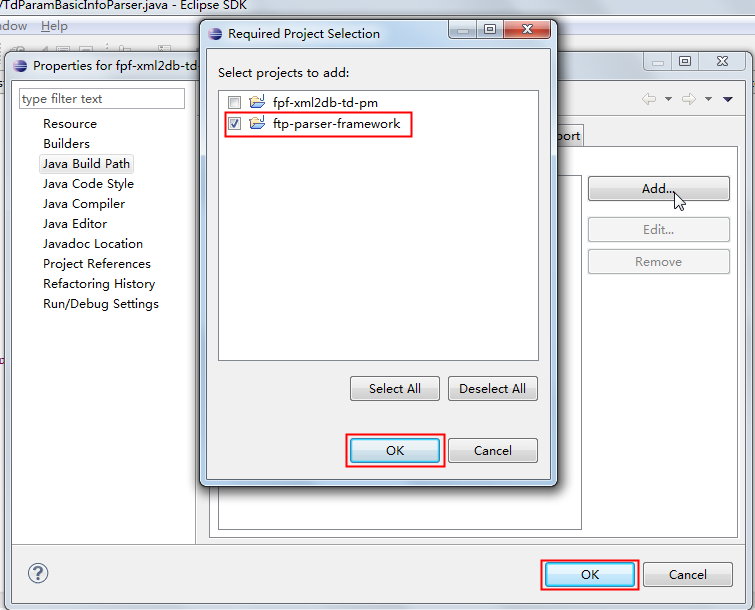
1. **确保FPF工程无异常的情况下，将其添加为该项目的依赖工程**。操作步骤如下：
   1. 选中该项目->右键->Properties



* 1. 在弹出框中选中Java Build Path->Projects->Add



* 1. 在弹出框中选中ftp-parser-framework->OK



# 程序流程

1. **文件下载流程**：遵循FPF中的FTP文件下载流程。
2. **文件解析流程**：总体流程遵循FPF框架中的文件解析流程。具体到TD性能的解析共经过四个步骤，文件解压、TD-PM-文件预处理、Xml2SimpleXml文件转换、SimpleXml2Db文件转换及入库，说明如下：
   1. **文件解压：**

对应类：com.inspur.ftpparserframework.transformator.GunzipTransformator

来源：FPF通用转换器

作用：将通过gzip压缩后的厂家原始性能文件解压缩，文件后缀名增加.unzip，如果原始文件不是压缩文件，则直接在文件名后面加后缀.unzip。

* 1. **TD-PM-文件预处理：**

对应类：com.inspur.nwom.td.param.transformator.TdPmPreTransformator

来源：TD性能解析专用转换器

作用：

(1)获取文件所属厂家的中文名、厂家英文名、厂家UK

(2)获取各厂家原始测量项及其顺序并按此顺序构造其对应的数据库字段及其顺序

(3)替换华为、诺西两个厂家不规范的命名空间

(4)获取文件所属OMC名称

(5)构造文件解析所需的xslt文件名

(6)处理完之后文件后缀名增加.pre

* 1. **Xml2SimpleXml文件转换**

对应类：com.inspur.ftpparserframework.transformator.Xml2SimpleXmlTransformator

来源：FPF通用转换器

作用：将复杂格式的xml文件根据xslt样式表转换为简单格式的xml文件，该简单格式的xml文件可做数据入库使用，处理完之后文件后缀名增加.xml2xml。

SimpleXml文件格式示例：

* 1. **SimpleXml2Db文件转换及入库**

对应类：com.inspur.ftpparserframework.transformator. SimpleXml2DbTransformator

来源：FPF通用转换器

作用：将简单格式的xml文件通过sqlldr入库。

# 开发扩展点

该程序在FPF框架的基础上共开发了2个扩展点，说明如下：

1. **TD性能文件基本信息扩展点**

对应类：com.inspur.nwom.td.param.extend.TdPmBasicInfoParser

作用：根据TD性能文件名构造基本信息对象(BasicInfo)。

说明：基本信息对象各字段的取值情况如下：

Domain = 无线

Technology = Constant.TECH\_TD\_SCDMA

DataType = Constant.DATATYPE\_PM

NeType = Constant.NETYPE\_OMC

TimeType = Constant.TIMETYPE\_15MIN

NeName = neName

Vendor = vendor

DataStartTime = startTime

DataEndTime = endTime

ReportTime = reportTime

其中蓝色字体的字段是固定值，紫色字体是需要根据各厂家的文件不同来分别提取的，**如果新增其他厂家的性能解析则需要扩展getVendor方法。**

1. **TD-PARAM-文件预处理扩展点**

请见上文中的描述。

**如果新增其他厂家的性能解析则需要扩展getVendorCnName、getVendorUk、getStandVendorName方法。**

# 关键参数设置参考

1. **java虚拟机最大内存数（ftp-parser-framework/bin/startup.sh）**

设置Xmx2G，TD性能文件结构比较简单，对内存占用不大，这里只开了2G，如果有条件当然是开的越大越好。

1. **文件并发解析数（SystemConfig.xml/parser/threadpool）**

由于TD性能文件格式较简单但是文件数量比较多，因此在进行解析时并发控制在10个。

1. **原始文件保存天数（SystemConfig.xml/parser/fileSavedDays）**

设置为7，表示保存7天的原始文件。

1. **ftp采集轮询间隔时间（SystemConfig.xml/ ftp/scanInterval）**

设置为30，表示保存30秒即半分钟，因为TD性能文件是15分钟粒度的数据，其文件生成很频繁，因此缩短轮询时间保证文件及时下载。

1. **ftp采集最近多长时间内生成的文件（SystemConfig.xml/ ftp/fileLastModifyTime）**

设置为3600，表示下载近3600秒即1小时内生成的文件，不重复下载。