

WebSphere Application Server

Unit 3

JSR-352 Concepts

© 2017 IBM Corporation

Objective of This Unit







This is based on the JSR-352 Specification, which can be found here:

https://www.jcp.org/en/jsr/detail?id=352

JSRs: Java Specification Requests JSR 352: Batch Applications for the Java Platform

Stage	Access
Maintenance Release	Download page
Maintenance Review Ballot	View results
Maintenance Draft Review	Download page



A Very Useful Document

http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102706

WebSphere Application Server

Understanding Java Batch (JSR-352)

This document can be found on the web at: www.ibm.com/support/techdocs Search for document number WP102706 under the category of "White Papers" This document explains the concepts and some of the details of the JSR-352 specification.



JSR-352 is a Standard, Which Means Programs Written to the Standard are Portable







The things you'll learn about in this workshop are nearly all IBM operational enhancements built around the core JSR-352 standard

- The jobOperator implementation: REST interface, batchManager, batchManagerZos
- Job logging, batch events, z/OS SMF records, multi-JVM design, etc.

The application has no direct awareness of any of that. The code has no specific requirement needed to use any of that.



JSR-352 Overview



Our Picture from Unit 1 - Overview



4. Job Operator

This interface defines how to submit and control jobs. This workshop focuses on the IBM enhancements around this job operator definition.

1. Java Batch Program

You write this based on the defined JSR-352 requirements. This is packaged as a servlet (WAR) and deployed into Liberty just like any other application would be deployed.

2. Job Step Programming Types

Two job step types defined:

Chunk: the looping model we most often associate with batch processing. This includes functions such as checkpointing, commits and rollbacks, and job restarts.

Batchlet: a simple "invoke and it runs" model. This is useful for non-looping functions such as file FTP steps.

3. Job Specification Language (JSL)

An XML specification to describe the batch job: the steps, the Java programs that implement the steps, and the flow of steps within the job.

5. Job Repository

The specification calls for a repository to track job submissions and results, but leaves it to the vendor to implement. We'll use IBM DB2 z/OS in workshop.



Some Key Terminology from the JSR-352 Specification



1. Job

A "job" encapsulates all the artifacts of a given batch process.

2. Job Step

A "step" implements a particular portion of your batch job. Your job may have 1 to many steps.

3. JSL

The Job Specification Language describes the components of the job, and defines the flow of execution (called "orchestration").

4. Job Instance

When a job is started, a "job instance" is created and assigned an instance ID.

Within a job instance a "job execution" is created and is assigned an execution ID. Under normal circumstances the execution and instance complete and the job completes. But if the job execution is stopped or failed, then you can *restart* within the same job instance. A new "job execution" is created.

If your job completes successfully and you start it again some time later, it gets a new Job Instance ID and a new Job Execution ID.



Two Step Types – Batchlet and Chunk

"Batchlet"

Start

process()

Your business logic processing is done here

This may contain whatever you require to implement the processing for this job step operation.

stop()

Spec says a well-behaved Batchlet will respond to stop() from the JobOperator



"Chunk" ItemReader readItem() Implement the data read activity here. It reads one data "record" at a time. ItemProcessor processItem() Implement the data process activity here. This acts upon the data object passed to it by the ItemReader.

ItemWriter

Start

writeItems()

Implement the data write activity here. This writes the set of data objects passed to it at checkpoint or end-of-data.

End

8



Example: Batchlet Step Outline Generated by WDT Tooling

```
package com.ibm.test;
import javax.batch.api.Batchlet;
public class MyBatchlet implements Batchlet {
    /**
    * Default constructor.
   public MyBatchlet() {
       // TODO Auto-generated constructor stub
/**
     * @see Batchlet#stop()
   public void stop() {
       // TODO Auto-generated method stub
/**
     * @see Batchlet#process()
   public String process() { 
       // TODO Auto-generated method stub
return null;
```

The IBM WebSphere Development Tool (WDT) plugin to Eclipse has code to support JSR-352 programming.

This is what the a batchlet step looks like when it's first generated.

Your batchlet step code goes here

Your batchlet stop processing code goes here

The code for ItemReader, ItemProcessor, and ItemWriter is similar ... but a bit longer and with more methods as per the spec requirements



Overview of Chunk Processing



1. The Read / Process loop

This loop processes until either a checkpoint interval is met (more on this coming up), or the reader returns a null, which means end-of-records.

The item returned from processItem() is added to a list of items which is eventually passed to the ItemWriter and written.

2. The Checkpoint / Write loop

If a checkpoint interval is reached (more on this coming up), control goes to the ItemWriter. The container passes the ItemWriter a list of items to write. Here your code may either iterate through the list, or perform a bulk insert if the output data resource permits that. At this point a transactional commit is processed if data resource supports transactional context.

3. The End of Records final write and exit

At some point you'll run out of records your ItemReader reads from, and that will trigger a final call to the ItemWriter and a final commit.

Some detail not shown:

- The open() method in both the ItemReader and the ItemWriter, which the container calls at the start
- The close() method in both the ItemReader and the ItemWriter, which the container calls at the end
- The checkpointInfo() method of both ItemReader and ItemWriter, which is used to maintain information about where within the records the last read and write was accomplished
- The transaction wrapper maintained by the container for transactional resources



Chunk Step Checkpoint Control



Your chunk code does not handle iterative looping, and does not do checkpoint processing. That is the role of the Batch Container.

Batchlet steps are another matter ... if you're looping in there, that's up to you.

1. The "batch container"

This is the JSR-352 implementation inside the Java EE 7 runtime.

2. checkpoint-policy

This is defined on the <chunk> element of the JSL, and you may either specify "item" (container handles) or "custom" (your own code handles)

3. item-count and time-limit

"item-count" specifies the number of iterations of ItemReader / ItemProcessor for a checkpoint interval (or end-of-data reached). "time-limit" specifies a time interval. This can be specified with itemcount. If time-limit reached before item-count, then checkpoint taken.

4. skip-limit and retry-limit

These control the behavior when configured skippable or retryable exceptions are encountered. This allows a job step to survive occasional or intermittent errors you've configured to skip and retry.



Status Codes: "Batch Status" and "Exit Status"

Container sets

this for each step and the

job overall



"Batch Status"



- STARTED
- STOPPING
- STOPPED
- FAILED
- COMPLETED
- ABANDONED

"Exit Status"

Under the control of the application

If the application does not explicitly set, then the "exit status" ends up being what the container sets for the "batch status"

Otherwise, the application can override and set its own exit status

Container classes: JobContext and StepContext

Two container-supplied classes that provide methods accessible to applications to get various context information, and to set context values. For example:

Job: getJobName(), getProperties(), getBatchStatus(), getExitStatus(), setExitStatus()
Step: getStepName(), getProperties(), getBatchStatus(), getExitStatus(), setExitStatus()

• This is useful in controlling flow of "orchestration"

We're going to look at "orchestration" – the flow of steps within a job – in detail. Key point here: the JSR-352 specification provides a way to control the flow based on the exit status of jobs. This is very similar to z/OS JCL condition codes. The Job Specification Language (JSL) provides an "On (exit) To (next)" mechanism.



Job Step "Flows" and "Splits"



1. Flow

From the spec: "A flow defines a sequence of execution elements that execute together as a unit."

2. Split

From the spec: "A split defines a set of flows that execute concurrently. A split may include only flow elements as children."

This picture shows two splits, but you may have more if you wish

Key Points:

- This is optional; you don't have to utilize this
- These flows and splits are defined in the Job Specificaton Language (JSL) file; they are not part of the Java code itself.
- The flows and splits for a given job operate within the same JVM; it will utilize separate threads. This is not "step partitioning," which can parallelize across JVMs.

IBM WebSphere Liberty Batch z/OS

IBM

Job Step "Partitioning"

From the spec: "A partitioned step runs as multiple instances of the same step definition across multiple threads, one partition per thread."



Key Points:

- This is optional; you don't have to utilize this
- Not all processing lends itself to effective partitioning. The data may be arranged in such a way that serial processing is better.
 However, if applicable, this can be an effective way to shorten processing time by parallelizing the processing.
- With IBM Liberty Java Batch, the partitions may run on multiple threads in the same server, or on threads across JVMs.
- This does not happen by magic. The spec: "Each partition needs the ability to receive unique parameters to instruct it which data on which to operate."

You must have knowledge of your data layout, and you must code your partition step to accept and operate on the data-range parameters.

The details of this is beyond the scope of this workshop

It is great technology, but at some point we have to draw a line to keep this workshop 2 days.



Job Specification Language (JSL)

This defines and controls the execution of the job

The Role of the Job Specification Language (JSL) File





The JSL file tells the story about what the batch job is comprised of, and how it is to run:

- Number of steps
- Step type (batchlet or chunk)
- Java class files that implement the steps
- Chunk type checkpoint policies
- Splits and flow processing
- Other control information externalized from the application

Syntax is XML; the elements are defined in the JSR-352 standard specification document

JSL is typically packaged with the batch application, but IBM Liberty Java Batch supports separate file pointed to at time of submission (called "inline JSL")



JSL Editor in the WebSphere Developer Tool (WDT) for Eclipse



There is IBM tooling support for Java Batch development. The Techdoc shown below provides details on installing into Eclipse.

Included is a "JSL Editor," which can help you design the job and the associated JSL.

The JSR specification document has further details on the JSL elements. Many other resources exist for details on the XML.

> Use this to create the JSL initially. If you wish to modify, then do so manually, or use the JSL Editor.

Techdoc http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102639



Real JSL – the JSL Packaged with the SleepyBatchlet Sample Program



1. One Step

This sample job has one step, which is the minimum.

2. Step is a batchlet

That one step is defined as a "batchlet"

3. The Java class that implements the job step

The ref= points to the Java class for the batchlet

4. One property is defined

You can pass in the number of seconds to "sleep"



A Hypothetical Two-Step Job



© 2017 IBM Corporation

Hypothetical with conditional next ...

Another Hypothetical ... Showing Conditional "Next" Processing



Yet Another Hypothetical ... Split and Two Flows







Same Hypothetical with the Generated JSL





"Listeners" – Job, Step, Chunk, ItemRead, ItemProcess, ItemWrite, Skip, and Retry

"Listeners" are callable interfaces behind which you may implement your own code to get control at various points in a job process: start of job, end of job, start of step, end of step, beginning of chunk, end of chunk, etc.



Example: The Job Listener Generate Stub Code





The batch container will call these interfaces at the beginning and the end of the job You may implement any processing here that you wish. Your code gets control, does what it does, and returns



Passing Parameters to the Batch Program



Summary





This is a big topic, and our objective here was to provide an understanding of the essential framework

We did not get into actual Java coding, but we did highlight how the WDT tool generates the stub classes for you to complete.

The key things to take-away:

- Two programming models: batchlet and chunk
- Job orchestration is accomplished with Job Specification Language (JSL)
- Your job can be simple (one step) or sophisticated (splits and flows)
- You can pass parameters into jobs from job submission

A Techdoc to be aware of: "Understanding Java Batch (JSR-352)"

http://www.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP102706



Screen Shots of JSL Editor

This illustrates how the "split and two flow" example is built



A Hypothetical Job We'll Use to Illustrate WDT and JSL Editor



1. Start with a batchlet step

We start with a batchlet step for no reason other than it's relatively easy to illustrate.

2. Start a split with two flows

A split may contain only flows, not steps outside of a flow. So we'll create two flows under the split and populate those flows with steps.

3. Flow with two steps, both chunk

For this flow we'll show two steps in the flow. We'll make them both chunk steps for illustration.

4. Flow with one step, which is a batchlet

For the other flow we'll have only one step to show that scenario. We'll make it a batchlet just to show a flow can contain either chunk or batchlets.

5. Finish with a final chunk step

We'll bring the split together with a final chunk step. After that the job ends.

We're not going to show Java coding. We're just going to show what it looks like to compose a job in the JSL editor, and use the generated JSL to explain the elements.



Starting Out ... Creating the Initial Step





Creating the Batchlet



Create the split ...



Create the Split





Add Flows Under the Defined Split





Add Steps to the Flows



Next ... add a "chunk" under each of those steps

Add Chunk to the First Step in the Flow

© 2017 IBM Corporation

IBM

Resolve the References for the ItemReader of the Chunk Step

Result: Generated Java for the ItemReader Class that Implements that Reader

```
*MyBatchJob.xml
                  🚺 Flow1Step1.java 🔀
     package com.ibm.test;
  2
  30 import java.io.Serializable;
     public class Flow1Step1 implements ItemReader {
  6
  7
         /**
  80
  9
          * Default constructor.
 10
          */
 110
         public Flow1Step1() {
12
             // TODO Auto-generated constructor stub
 13
 14
 150
         /**
          * @see ItemReader#readItem()
 16
 17
          */
△18⊖
         public Object readItem() {
F19
             // TODO Auto-generated method stub
 20
                 return null:
 21
 22
         /**
 230
 24
          * @see ItemReader#open(Serializable)
 25
          */
△26⊝
         public void open(Serializable arg0) {
27
             // TODO Auto-generated method stub
 28
 29
 300
 31
          * @see ItemReader#close()
 32
          */
△33Θ
         public void close() {
234
             // TODO Auto-generated method stub
 35
 36
 370
         /**
 38
          * @see ItemReader#checkpointInfo()
 39
          */
△40⊝
         public Serializable checkpointInfo() {
P41
             // TODO Auto-generated method stub
 42
                 return null:
 43
```

Result is an editor with Java framework for a ItemReader class. The required interfaces are pre-populated; you code from there.

Do the same for ItemProcessor and ItemWriter. Result:

Add Chunk to Second Step in the Flow and Resolve the References

We're now starting to repeat the same steps over and over, so we won't show detail

Not yet completed, but the process is as we've already illustrated: just point and click and resolving package and class references

Add "Next" References so Flow Through Job Specified in the JSL

1. Set "Next" on Step1 to Indicate "Split1"

2. Split will go down each defined flow in the split

You *do* define a "Next" on the Split, but not to the flows. You define "Next" to indicate where to go *after the split comes back together*. See #4 below.

3. Set "Next" on Flow1Step1 to Indicate "Flow1Step2"

4. Set "Next" on split to go to "FinalStep"

Detail	5	
ID*:	Split1	
Next:	FinalStep	

All the Steps, Splits and Flows ... and the Generated JSL

