JSDL 18th January 2005

Akioka
Andreas
Donal
Ali
An
Fred
Saga
Steve

Agenda

Next Meeting: 25 Jan 2005 14:00 GMT, 07:00 MST

Look at Fred's and Igor's Proposals

[Text Slide A]

Notes from last week

---

1. We don't see where we could specify a type of application distribution, e.g. MPI, OpenMPI. Do you think this is too specific to add it to the application element?

Donal> Either it is a Resource (something which has to be there for the job to run) or it is part of the Application (an inherent aspect of the app running itself). I don't fully understand which. :^) Either way, it should be possible to fit it as an extension element into one of the two places.

Group> This can be handled with the Application type using the other value in the enumeration. It could just be an executable and then the process topology.

2. How would you suggest to specify a logical file name: using predefined syntax or using a special attribute in the FileName element? Can a whole directory be specified as well?

Group> You can specify either one.

(Action Item) We need to show an example of that in the spec. Only '/' can be used as a delimiter for directories and files in the filepath.

3. Why Source & Target in DataStaging are in one element? How to use different name for an input and output file? We would suggest to add a choice element above Source and Target or to define some separate elements: e.g. SourceFiles and TargetFiles (probably of the same type).

Donal> They are in a single element so you can have a (logical) file that is staged in, modified, and then staged out again. No file has to have both Source and Target, and I think there are use-cases for having neither (e.g. where you just want to closely control some deletion behaviour).

Group> Give more examples in the spec on how this mechanism works.

4. Why the CreationFlag element is mandatory? I think this feature is not available in many systems. Furthermore, a reasonable default value can be chosen.

Andreas> When the CreationFlag was not mandatory, behaviour was stated as implementation dependent if the flag was left undefined. The consensus was that it wasn't right to leave it implementation dependent. And since as a general approach we have chosen not to specify 'default default' values (we cannot say what is default for the grid) the flag became mandatory.
I guess the main question (for me) would be whether the values defined for this flag at the moment (including 'dontOverwrite' which is not in the spec but is in the latest schema) cover the behaviour most people would want to specify. Obviously, whether this flag is mandatory or not, it is still possible to specify a value that is not supported by many systems. That only means that the job can only be executed in the, very restricted, environment that supports that option.

(Action Item) We should probably provide informative text in the spec. on what we think is a reasonably common choice for this flag.

5. Definition 5.6.6.1:
"A Source element contains the location and may contain a user on the remote system. This file MUST be staged in from the location specified by the URL as the user on the remote host before the job has terminated."

(Action Item)Do you really mean "terminated" or it should be "started"?

6.What is expected as operatingSystemDesc? Is it human readable description?

Donal> I think that's one of the items we hope to borrow from CIM.

(Action Item) All element that are of type Description is human readable and should be handled as such.

7. I saw the uses-cases containing descriptions of a job consisting of multiple processes and threads

requiring multiple resources and/or processors. I guess that specification of alternative configurations is not possible, e.g. 4 nodes 4 processors each OR 1 node, 16 processors?

(Action Item) We need to show this example using the schema in the spec.

8. I didn't find a specification of the queue name in the new schema. I think it might be sometimes useful similarly as an implicit specification of a specific host.

Group> Again we have removed this because it is part of the scheduling document that is outside of the jsdl language.

(Action Item) Show how a queue can be specified in an example outside of the jsdl language and as an extension to jsdl.

9. Regarding your question in the spec about the webService value of the ApplicationTypeEnumeration, we think this is definitely needed (invoking an existing WebService).

(Action Item) Put Fred's example into the spec.

(Action Item) Discuss Fred's example next week. We need to discuss whether what we currently haved sufficient.

10. Do limits mean that if they are exceeded during application execution the application must be terminated?

Donal> They're meant to be interpreted as what the user want's his POSIX system limits to be.

(Action Item) Do we need to specify "hard" and "soft"? What does it mean to be a hard limit what does it mean to be a soft limit. Or it the limit cannot be enforced then the consuming system must reject the document.

11. Within the FileSystem element there is a sub-element MountPoint. Who should specify this? A user? I think a more common use case is that users use predefined variables (e.g. home, tmp) to specify paths that are relative to these variables (but mount points depend on a local system).

Donal> IIRC, the user can specify it and the system can either ensure that the FS is mounted, or it can just check to see if it is mounted and throw the job out if it isn't.

(Action Item) We need more text and some examples. We also need show how to specify disk space separately from filesystem, maybe? We need to discuss this in the email list and in the next meeting. (Andreas to send email)

12. I asked in the previous email about defining software dependencies, e.g. necessary libraries and you requested some use-cases. Simple examples are as follows:
- interactive application may need additional software (e.g. VNC) to enable users to access remotely application's user interface
- graphical application may need the OpenGL library to run
- Java applications need Java Virtual Machine installed Therefore, in my opinion such an element would be useful and general. Of course, we can add it using the extension possibility.

(Action Item) Show examples in the spec on how to specify these use cases. An additional use case would be a script that requires java.

[Text Slide B]

Actions:
=======

1) Place an example in the Spec for directory delimiting
CF: From point 2 in last weeks minutes
(Andreas)

2) More exaples for File staging - we have some of these on GridForge, they need to be merged in.
(Andreas, Steve)

3) Add to spec that overwrite is a commonly used value for CreationFlag
CF: From point 4 in last weeks minutes
(Does anyone have a good reason for making this compulsory? or can we make it optional and implementation dependent?)
(Steve to post to list)

4) Check Darren has done item 5 from last week

5) Steve to complete action 6 from last week

6) Andreas to do item 7 from last week

7) Steve to add f2f slides into introductory matter in spec (cf point 8 from last week)

8) Action Fred's ideas for more thought (cf 9)

9) Donal to look into (posix) limits (cf10)

10)

====
Donal K. Fellows wrote:

> Ariel Oleksiak wrote:
>
> > 11. Within the FileSystem element there is a sub-element MountPoint. Who
> > should specify this? A user? I think a more common use case is that users

> use predefined variables (e.g. home, tmp) to specify paths that are relative
> to these variables (but mount points depend on a local system).

> IIRC, the user can specify it and the system can either ensure that the
> FS is mounted, or it can just check to see if it is mounted and throw
> the job out if it isn't.

> But MountPoint is a mandatory element. How to specify required disc space
> without knowledge about mount points on local systems?

> IIRC, the mount point does not need to be associated with an explicit
> location. (I hope someone else will try to answer this question!)

=====

There are 3 scenarios here:

1. User specifies that certain filesystems must be available at a certain locations
2. User specifies certain filesystems must be available but doesn't care where, and the information is passed back somehow, e.g., environment variables.
3. User only cares that sufficient DiskSpace is available because there is access to some other mechanism that will deploy/configure the resource before the job is executed.

At the moment we clearly support (1) and somewhat obsurely (2), and perhaps even (3) as Donal mentions. If we agree that these 3 scenarios are equally valid I wonder if we shouldn't make the rules clearer. For example:
- make the MountPoint optional
- if the MountPoint is specified then the filesystem MUST be available at that location
- if the MountPoint is not specified then the filesystem MUST be available at some location and there MUST be an environment variable (subject to the rules of ApplicationType) defined by the execution system with the same name as the "FileSystem id". The value of this environment variable MUST be the location of the filesystem.

And for case (3) we could allow the DiskSpace element to appear directly under Resource.

Andreas

Action: Have two tags MountPoint and MountSource both of which are optional.

11) To provide examples of software resource reqirements and where they should go in the document. Probably place these in as extension elements for now (cf 12)

[Text Slide C]

Fred's & Igor's webService example

Attached is an example. The logic is this:

1. Given the message form the "job server" will look for deployed Web services which can accept this message. <wsa:Action> and the body of the Soap message make that unique for the function requested. If necessaary, "job server" can use WS-MetadataExchange to retrieve the WSDL of a WS and match the message form and action.

2. Once known resource can be selected according to restrictions.

3. SOAP message is sent to the matching WS.

4. WS replies to the original sender (thanks to WS-Addressing headers and logic defined in that spec).

Looks good to me.

```
<?xml version="1.0" encoding="UTF-8"?>
<jsdl:JobDefinitions
 xmlns:jsdl="http://www.gridforum.org/JSDL"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:schemaLocation="http://www.gridforum.org/JSDL/jsdl.xsd">

  <jsdl:JobIdentification>
    <jsdl:JobName>WebServicesJob</jsdl:JobName>
    <jsdl:JobAnnotation>
 This is the Web services job submitted through JSDL.
    </jsdl:JobAnnotation>
    <jsdl:JobAnnotation>
 This sends a SOAP message to any recipient which can accept it.
    </jsdl:JobAnnotation>
  </jsdl:JobIdentification>

  <!-- this indicates requirements which a SOAP message recipient must comply with -->
```

```
  <jsdl:Resource name="SomeComputer">
    <jsdl:CPUCount>2</jsdl:CPUCount>
    <jsdl:CPUSpeed>2Ghz</jsdl:CPUSpeed>
    <jsdl:PhysicalMemory>1024Mb</jsdl:PhysicalMemory>
  </jsdl:Resource>

  <jsdl:Application>
    <jsdl:Message>
      <!-- note that this is an unecrypted/unsigned message, but it may be so! -->
      <S:Envelope xmlns:S="...SOAP1.2..." xmlns:wsa="...WS-Addressing..." xmlns:wse="...WSS..."
       xmlns:math="http://schemas.xml.org/math">
 <S:Header>
     <wsa:To>urn:com:myorg:math:anyhost</wsa:To>
     <wsa:ReplyTo>http://reports.myorg.com/resultsConsolidator</wsa:ReplyTo>
     <wsa:MessageId>...UUID...</wsa:MessageId>
     <wsa:Action>http://wsdl.xml.org/math/MonteCarloSimulation/PassRequest</wsa:Action>
     <wsse:SecurityContext>
  ... WSS context with, say, X.509 token in it ...
     </wsse:SecurityContext>
     ... other custom headers ...
 </S:Header>
 <S:Body>
     <math:MonteCarloSimulationPass>
  <math:MatrixEncoded>... base64Binary data ...</math:MatrixEncoded>
  ... other parameters ...
     </math:MonteCarloSimulationPass>
 </S:Body>
      </S:Envelope>
    </jsdl:Message>
  </jsdl:Application>
</jsdl:job>
```

[Text Slide D]