GWD-R
Usage Records Working Group
<a href="http://www.psc.edu/~lfm/Grid/UR-WG">http://www.psc.edu/~lfm/Grid/UR-WG</a>
http://www.psc.edu/~lfm/Grid/UR-WG/Usage Record XML Format.pdf</a>

#### **Usage Record -- XML Format**

#### Status of This Memo

This memo provides information to the Grid community in the area of usage records and accounting. Distribution is unlimited.

#### Copyright Notice

Copyright © Global Grid Forum (2003). All Rights Reserved.

### **Abstract**

This document describes a common format with which to exchange basic accounting and usage data over the grid. This record format is intended to facilitate the sharing of usage information among grid sites, particularly in the area of job accounting. The usage record is represented in an XML format. This document does not address how these records should be used, nor does it attempt to dictate the format in which the accounting records are stored at a local site, rather it is simply meant to be a common exchange format. Furthermore, nothing is said regarding the communications mechanisms employed to exchange the records, i.e. transport layer, framing, authentication, integrity, etc.

## **Table of Contents**

A	.bstract.		1		
T	able of	Contents	1		
1.	. Intro				
2.	Conventions Used in this Document.				
	2.1	Encoding	2		
	2.2	XML Conventions			
	2.3	Schema Organization and Namespaces	4		
	2.3.	Empty and whitespace only values	4		
	2.3.2	2 Time values	4		
	2.3.		4		
	2.4	Table Column Interpretations	5		
3.	. Usa	geRecord Element	5		
4.	. Usa	ge Record Properties	7		
	4.1	Simple Usage Record Properties	9		
4.2		Consumable Resources	16		
	4.3	Extension Element	16		
	4.4	ConsumableExtension Element	16		
5.	. Exa	mples	16		

5.1 Sample Job Usage Record	16			
6. Security Considerations				
Author Information				
Intellectual Property Statement				
Full Copyright Notice	17			
References				

### 1. Introduction

In order for resources to be shared, sites must be able to exchange basic accounting and usage data in a common format. This document describes an XML-based format for usage records. The record format is intended to be specific enough to facilitate information sharing among grid sites, yet general enough that the usage data can be used for a variety of purposes - traditional usage accounting, charging, service usage monitoring, performance tuning, etc.

## 1.1 Encoding

#### 1.1.1 XML Structure

This specification uses XML Schema documents conforming to the W3C XML Schema Specification [XML\_Schema] and normative text to describe the syntax and semantics of XML encoded usage records.

## 1.1.2 Transport

This specification defines the structure of information that may be represented in a compliant XML document. No requirements are placed on the encoding of this document for a particular transport. Therefore, instance documents may be represented in ASCII or Unicode text. Further, we envision that many of the systems using this data definition will be OGSA compliant systems and therefore preferences to the http/ https protocols may occur. However, a usage record may be communicated via any lower level transport that is acceptable to the using parties.

# 1.2 Extensibility

The XML formats for representing Usage Records have been designed with consideration given to extensibility for implementation specific requirements. However, the use of extensions may reduce interoperability and therefore the introduction of extensions SHOULD be carefully considered.

# 2. Conventions Used in this Document

# 2.1 Key Words

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this specification are to be interpreted as described in IETF RFC 2119 [RFC2119]:

"they MUST only be used where it is actually required for interoperation or to limit behavior which has potential for causing harm (e.g. limiting retransmission)"

These key words are thus capitalized when used to unambiguously specify requirements over features and behavior that affect interoperability and security of implementations.

# 2.2 Schema listings

### Listings from the Usage Record Schema appear like this

In case of disagreement between the schema file and this specification, the schema file takes precedence.

#### 2.3 XML Conventions

### 2.3.1 Element names

- 1. Element names SHALL be in UCC convention (example: <UpperCamelCaseElement/>.
- 2. Capitalization of element names in external specifications SHALL remain consistent with the initial specification. (example: ds:KeyInfo).
- 3. Acronyms SHOULD be avoided.
- 4. Underscores ( ), periods (.) and dashes(-) MUST NOT be used.
- 5. Element names MUST comply with all XML Schema specific naming rules.

#### 2.3.2 Attribute names

Attribute names SHALL be in LCC convention (example: <UpperCamelCaseElement lowerCamelCaseAttribute="attributevalue"/>.

- Attribute names SHALL be in LCC convention (example: 
   UpperCamelCaseElement lowerCamelCaseAttribute="attributevalue"/>)
- 2. Capitalization of attribute names in external specifications SHALL remain consistent with the initial specification.
- 3. Acronyms SHOULD be avoided.
- 4. Underscores ( ), periods (.) and dashes(-) MUST NOT be used.
- 5. Attribute names MUST comply with all XML Schema specific naming rules.

#### 2.3.3 Enumerated attribute values

- 1. Attributes of type enumeration SHALL have values in LCC convention (example: <UpperCamelCaseElement enumAttribute="valueOne"/>).
- 2. Capitalization of enumerated attribute values in external specifications SHALL remain consistent with the initial specification.

## 2.3.4 Empty and whitespace only values

\*\*\*We should define here how empty and whitespace is expected to be treated.

#### 2.3.5 Time values

All time values have the type **dateTime**, which is built in to the W3C XML Schema Datatypes specification and MUST be expressed in UTC form.

Leap seconds??

Supporting time resolution finer than milliseconds.

## 2.3.6 Comparing Usage Record Values

\*\*\*We should define here we expect values to be compared when type string/etc

# 2.4 Schema Organization and Namespaces

The usage record structures are defined in a schema [URWG-XSD] associated with the following namespace:

http://www.gridforum.org/2003/ur-wg

The digital signature components are defined in a schema [XML-SIG] associated with the following namespace: This schema is imported into the URWG schema to directly use its definitions.

http://www.w3.org/2000/09/xmldsig#

All simple data types referenced in this document are built into the W3C XML Schema Datatypes specification. When referenced in this document, this namespace is associated with the prefix xsd

http://www.w3.org/2001/XMLSchema

# 2.4.1 Schema Header and Namespace Declarations

The following schema fragment defines the XML namespaces and other header information:

```
<xsd:schema
attributeFormDefault="qualified"
elementFormDefault="qualified"
targetNamespace="http://www.gridforum.org/2003/ur-wg"
xmlns:ds="http://www.w3.org/2000/09/xmldsig#"
xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:urwg="http://www.gridforum.org/2003/ur-wg">
...
</schema>
```

### 3. Global Element Definitions

The global definitions are those that may be used compliantly with this specification as root elements for an XML document or which may be used as extension points within or included within other XML Schema definitions.

# 3.1 UsageRecord Element

The *UsageRecord* element encapsulates a single Usage Record. The UsageRecordType complex type dictates the particular structure of this element. All specific usage record elements should extend or restrict this element. Any structure that contains usage record information should reference the UsageRecord element so that extensions or restrictions of the element are automatically handled. This element should contain all the information that is generic to a usage record and addressed by this specification. Any extensions to or restrictions of the UsageRecord should not alter this generic structure.

```
<element name="UsageRecord" type="urwg:UsageRecordType"/>
```

## 3.2 JobUsageRecord Element

This element definition establishes the structure of job usage record as it derives from the generic UsageRecordType.

\*\*\* The definition for each of these (and other property) elements is in progress.

# 4. Global Attribute Definitions

Global attributes are attributes that are common to many of the element definitions contained in this specification. By defining these attributes globally, we ensure a standard mechanism for expressing the appropriate information. Each element that references a global attribute must specify whether its use is required, optional or prohibited. Each element must also identify the default attribute value to assume, if applicable, when the attribute is omitted from the instance document.

## 4.1 Description

This attribute provides a mechanism for additional, optional information to be attached to a Usage Record element. The value of this attribute MAY provide clues to the semantic context to use while interpreting or examining the value of the owning element.

```
<xsd:attribute name="description" type="xsd:string" />
```

#### 4.2 Units

This attribute represents the unit of measure that should be applied to the value of its owning element. As an example, memory can be expressed in bytes, kilobytes, megabytes and gigabytes. The selected representation for these units of measure are: B, KB, MB and GB respectively. Each element that references this global attribute MUST define its use as required and identify the appropriate default value.

```
<xsd:attribute name="units" type="xsd:string" />
```

#### 4.3 Metric

This attribute identifies the type of measurement used for quantifying the associated resource consumption if there are multiple methods by which measure resource usage. As an example, disk usage may be reported as total, average, minimum or maximum usage. However, this attribute MUST not attempt to differentiate between requested and utilized quantities of resource usage within a single record, even if this differentiation is pertinent to the final assessed charge,. Each element that references this global attribute MUST define its use as required and identify the appropriate default value.

```
<xsd:attribute name="units" type="xsd:string" />
```

# 5. Global Type Definitions

# 5.1 UsageRecordType Complex Type

This complex type definition establishes the structure of the generic usage record. The properties defined in version XXX of the Usage Record Natural Language Document form the basis for the definition of the UsageRecordType components.

\*\*Need to talk about the ordering of the elements, why some are in a choice and the number of occurrences....

```
<xsd:element ref="urwg:CpuTime" minOccurs="0" maxOccurs="1"/>
            <xsd:element ref="urwg:NodeCount" minOccurs="0" maxOccurs="1"/>
            <xsd:element ref="urwq:Processors" minOccurs="0"</pre>
maxOccurs="1"/>
            <xsd:element ref="urwg:EndTime" minOccurs="0" maxOccurs="1"/>
            <xsd:element ref="urwg:StartTime" minOccurs="0" maxOccurs="1"/>
            <xsd:element ref="urwq:MachineName" minOccurs="0"</pre>
maxOccurs="1"/>
            <xsd:element ref="urwq:SubmitHost" minOccurs="0"</pre>
maxOccurs="1"/>
            <xsd:element ref="urwg:Queue" minOccurs="0" maxOccurs="1"/>
            <xsd:element ref="urwg:ProjectName" minOccurs="0"</pre>
maxOccurs="unbounded"/>
           <xsd:element ref="urwg:Host" minOccurs="0"</pre>
maxOccurs="unbounded"/>
            <xsd:element ref="urwg:Charge" minOccurs="0" maxOccurs="1"/>
            <xsd:element ref="urwg:Status" minOccurs="0" maxOccurs="1"/>
    </xsd:sequence>
    </xsd:complexType>
```

## 5.2 domainNameType Simple Type

This simple type definition establishes the structure of a Fully Qualified Domain Name in string format. Any valid value provided follows the format of a fully qualified domain name (FQDN) as defined by RFC 1034:

FQDNs can be up to 255 characters long and can contain alphabetic and numeric characters and the "-" and the "." characters.

\*\*\* The definition other types is TBD.

# 6. Usage Record Common Properties

The complete specifications for each element that MAY appear in the generic usage record element are presented below. A quick reference table summarizes these elements in Appendix A.

#### 6.1 JobName

This element specifies the name of the job or application that generated this usage.

description

This optional attribute provides additional information about job name.

## 6.2 Charge

This element specifies the total charge, in the generating system's allocation unit, associated with the usage represented by this structure. The charge MAY be reported without any additional information regarding the usage that generated this charge. However, the value reported MAY include premiums or discounts assessed on the actual usage represented within this record. Therefore, the reported charge may not be directly reconstructed from the specific usage reported.

#### units

This attribute specifies the unit of measurement in which the charge for usage is reported. There are no values that must be supported by implementations.

#### 6.3 Status

This element specifies a completion status associated with the usage. For example, this may represent the exit status of an interactive running process or the exit status from the batch queuing system's accounting record. There is no defined semantic meaning attached to a particular value and thus is implementation specific.

#### description

This optional attribute may provide information about the meaning associated with the reported status.

#### 6.4 Network

This element specifies the network usage.

```
description
```

This optional attribute may provide information about the type of network usage reported.

#### units

This attribute specifies the unit of measurement in which network usage is reported. The values that MUST be supported for this attribute are:

- 1. B bytes
- 2. KB Kilobytes
- 3. MG Megabytes
- 4. GB Gigabytes

The default value for the units attribute of the Network element is MB.

#### metric

This attribute specifies the metric for the reported network usage. The values that MUST be supported for this attribute are:

- 1. average
- 2. total
- 3. max
- 4. min
- 5. requested

The default value for this attribute is average.

#### 6.5 Disk

This element specifies the disk storage for this usage.

```
description
```

This optional attribute may provide information about the type of disk usage reported. For example, it may be scratch space, network storage, etc.

#### units

This attribute specifies the unit of measurement in which disk usage is reported. The values that MUST be supported for this attribute are:

- 1. B bytes
- 2. KB Kilobytes
- 3. MG Megabytes
- 4. GB Gigabytes

The default value for this attribute is MB.

#### metric

This attribute specifies the metric for the reported disk usage. The values that MUST be supported for this attribute are:

- 1. average
- 2. total
- 3. max
- 4. min
- 5. requested

The default value for this attribute is average.

# 6.6 Memory

This element specifies the memory usage.

description

This optional attribute may provide information about the type of memory usage reported. For example, it may be virtual memory, paged memory, shared memory, dedicated memory, etc

#### units

This attribute specifies the unit of measurement in which memory usage is reported. The values that MUST be supported for this attribute are:

- 1. B bytes
- 2. KB Kilobytes
- 3. MG Megabytes
- 4. GB Gigabytes

The default value for this attribute is MB.

#### metric

This attribute specifies the metric for the reported memory usage. The values that MUST be supported for this attribute are:

- 1. average
- 2. total
- 3. max
- 4. min
- 5. requested

The default value for this attribute is average.

#### 6.7 Walltime

This element specifies the wall clock time that elapsed.

```
description
```

This optional attribute provides additional information about the walltime reported

```
units
```

This attribute specifies the unit of measurement in which walltime is reported. The values that MUST be supported for this attribute are:

- 1. seconds
- 2. minutes
- 3. hours

The default value for this attribute is seconds.

# 6.8 CpuTime

This element specifies the CPU time used, summed over all processes associated with this usage.

description

This optional attribute provides additional information about the cpu time reported

units

This attribute specifies the unit of measurement in which CPU time is reported. The values that MUST be supported for this attribute are:

- 1. seconds
- 2. minutes
- 3. hours

The default value for this attribute is seconds.

#### 6.9 NodeCount

This element specifies the number of nodes used. The definition of a node may depend upon architecture.

```
description
```

This optional attribute provides additional information about the node count reported, for example it may report architecture (cluster vs. smp) associated with this usage record.

## 6.10 Processors

This element specifies the number of processors used or requested. A processor definition may depend on machine architecture

```
description
```

This optional attribute provides additional information about number of processor reported, for example it may report machine architecture associated with this usage record.

#### 6.11 EndTime

This element specifies the time at which usage completed. The end time may depend upon the queue system, file staging etc.

```
description
```

This optional attribute provides additional information about end time reported.

### 6.12 Start Time

This element specifies the time at which usage completed. The end time may depend upon the queue system, file staging etc.

description

This optional attribute provides additional information about end time reported.

### 6.13 Machine Name

This element specifies the name of the machine on which a job ran.

description

This optional attribute provides additional information about machine.

#### 6.14 Host

This element specifies the name of the host on which a job ran.

description

This optional attribute provides additional information about host.

primary

This attribute indicates whether this host acted as the primary host for the execution that incurred this usage. This attribute contains a Boolean value which MUST be true if this host was the primary host. The default value for this attribute is false.

#### 6.15 SubmitHost

This element specifies the name of the host from which a job was submitted.

description

This optional attribute provides additional information about host.

### 6.16 Queue

This element specifies the name of the queue from which the job executed or was submitted.

description

This optional attribute provides additional information about host.

## 6.17 ProjectName

This element specifies the name/identifier of the project associated with this usage. Some accounting systems define this as the ACID, which is often the same as the users Unix group from /etc/passwd. The project is also referred to as the effective gid under which the job consumed resources on some systems.

#### description

This optional attribute provides additional information about project, for example a human readable project name.

\*\*\*Additional Element definitions from NLD go here

# 7. Examples \*\*\*Need English verbage for each example

# 7.1 Sample Usage Record

# 7.2 Sample JobUsage Record

```
<StartTime>2003-02-14T13:10:30Z</StartTime>

<EndTime>2003-02-14T15:08:23Z</EndTime>

<Charge units="USD">12.75</Charge>

<Status>Completed</Status>

</UsageRecord>
```

# 8. Security Considerations

There are no explicit security considerations for these requirements.

## **Author Information**

Rebekah Lepro NASA Ames Research Center rlepro@mail.arc.nasa.gov (650)604-4359 Scott M. Jackson Pacific Northwest National Laboratory scott.jackson@pnl.gov (509) 376-2205

# **Intellectual Property Statement**

The GGF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the GGF Secretariat.

The GGF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this recommendation. Please address the information to the GGF Executive Director

# **Full Copyright Notice**

Copyright (C) Global Grid Forum (2003). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the GGF or other organizations, except as needed for the purpose of developing Grid Recommendations in which case the procedures for copyrights defined in the GGF

Document process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the GGF or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE GLOBAL GRID FORUM DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE."

## References

[KEYWORDS] S. Bradner, "Key Words for Use in RFCs to Indicate Requirement Levels", <u>RFC 2119</u>, March 1997.

[XML\_CONV] "I-X and <I-N-CA> XML Conventions".

[FED\_XML] "U.S. Federal XML Guidelines".

[XML] Bray, T., et al, "Extensible Markup Language (XML) 1.0 (Second Edition)", 6 October 2000.

[XML\_SCHEMA] D. Beech, M. Maloney, N. Mendelshohn, "XML Schema Part 1: Structures Working Draft", April 2000.

# Appendix A

# **Table Column Interpretations**

In the property tables, the columns are interpreted to have the following meanings:

Element Name: Name of the XML element (xsd:element)

Type: Data type defined by Sections 1 & 2 and the W3C XML Schema

specification

Description: Brief description of the meaning of the property

# **Common Usage Record Properties**

Simple (unstructured) usage record properties are enumerated in Table 1.

Table 1 Simple Usage Record Properties

Table 1 Shiple Osage Record Properties				
Element	Type	Description		
Name				
Charge	xsd:positiveInteger	Total charge for this usage		
CPUDuration	xsd:positiveInteger	CPU Time used in seconds, summed over all		
		processes in the job		
EndTime	xsd:timestamp	Date and time when usage ended		
MachineName	xsd:string	Name of the machine or cluster on where usage		
		was incurred		
NodeCount	xsd:positiveInteger	Number of nodes used		
ProjectId	xsd:string	Project or account name		
Queue	xsd:string	Queue name		
StartTime	sd:timestamp	Date and time when usage started		
Status	xsd:integer	Completion Status		
SubmitHost	xsd:string	The hostname from which request was submitted		
WallDuration	xsd:positiveInteger	Wallclock duration of usage in seconds		
		(sometimes referred to as WallClockTime)		