

Life Science Resource Profile BoF

Fumikazu KONISHI, Tokyo Tech.

OGF IPR Policies Apply



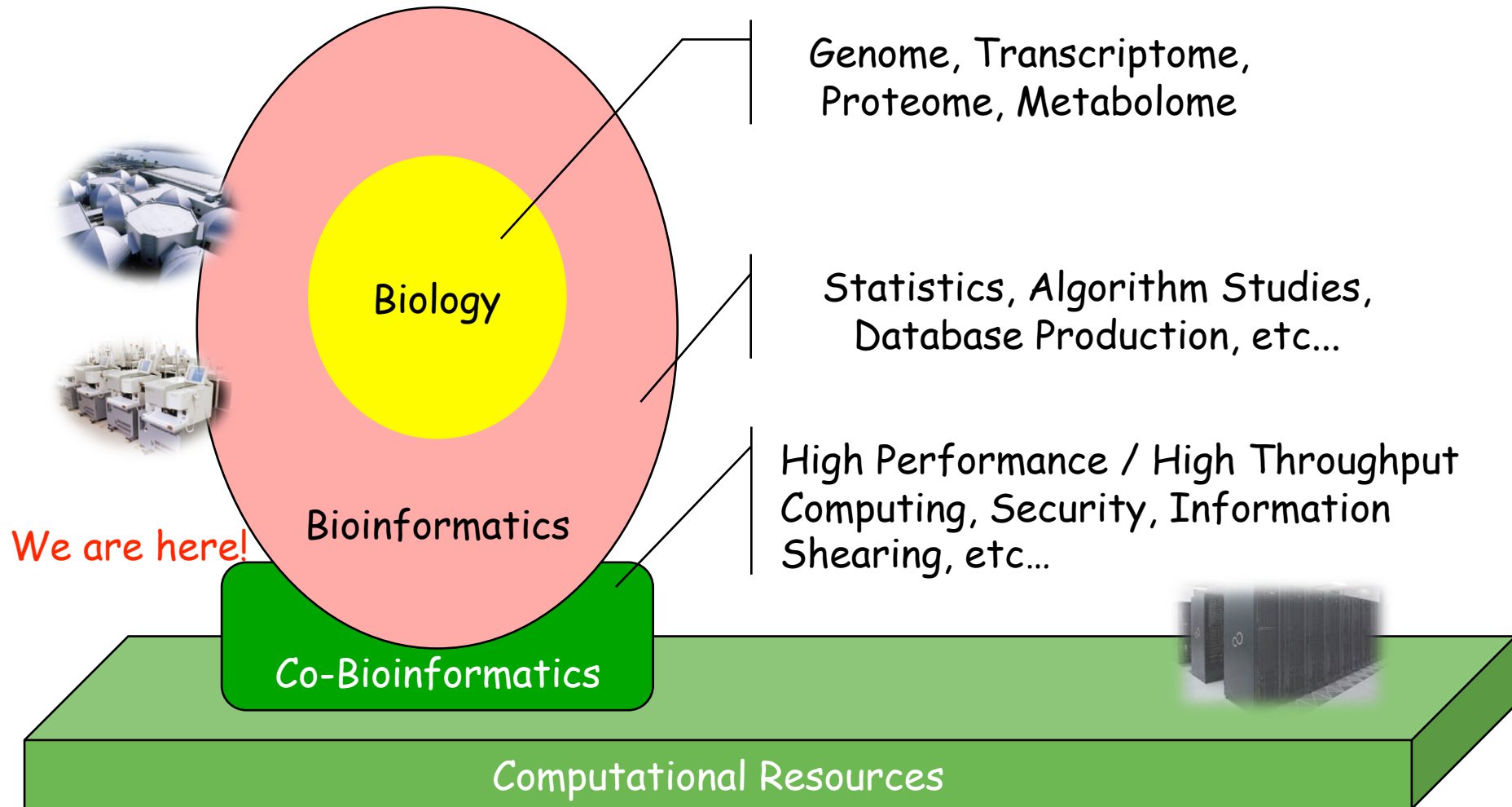
- “I acknowledge that participation in this meeting is subject to the OGF Intellectual Property Policy.”
- Intellectual Property Notices Note Well: All statements related to the activities of the OGF and addressed to the OGF are subject to all provisions of Appendix B of GFD-C.1, which grants to the OGF and its participants certain licenses and rights in such statements. Such statements include verbal statements in OGF meetings, as well as written and electronic communications made at any time or place, which are addressed to:
 - the OGF plenary session,
 - any OGF working group or portion thereof,
 - the OGF Board of Directors, the GFSG, or any member thereof on behalf of the OGF,
 - the ADCOM, or any member thereof on behalf of the ADCOM,
 - any OGF mailing list, including any group list, or any other list functioning under OGF auspices,
 - the OGF Editor or the document authoring and review process
- Statements made outside of a OGF meeting, mailing list or other function, that are clearly not intended to be input to an OGF activity, group or function, are not subject to these provisions.
- Excerpt from Appendix B of GFD-C.1: “Where the OGF knows of rights, or claimed rights, the OGF secretariat shall attempt to obtain from the claimant of such rights, a written assurance that upon approval by the GFSG of the relevant OGF document(s), any party will be able to obtain the right to implement, use and distribute the technology or works when implementing, using or distributing technology based upon the specific specification(s) under openly specified, reasonable, non-discriminatory terms. The working group or research group proposing the use of the technology with respect to which the proprietary rights are claimed may assist the OGF secretariat in this effort. The results of this procedure shall not affect advancement of document, except that the GFSG may defer approval where a delay may facilitate the obtaining of such assurances. The results will, however, be recorded by the OGF Secretariat, and made available. The GFSG may also direct that a summary of the results be included in any GFD published containing the specification.”
- OGF Intellectual Property Policies are adapted from the IETF Intellectual Property Policies that support the Internet Standards Process.

Agenda



- Introduction of Life Science Resource Profile
- Current status of Resource Namespace Service (RNS)
- Discussion

Background



Understand the present situation in Life Science Activities.



- Various type of device can produce an experimental data.
- Each device can store the data in various types of file format.

Ex. picture, sequence, structure, etc..

Bioinformatics activities will start to read and to parse those files.



© 2006 Open Grid Forum

OGF23



Thermus thermophilus HB27, complete genome.



- /pub/kegg/genes/organisms/tth/ / ftp.genome.jp
 - T.thermophilus.ent
 - t.thermophilus.genome
 - t.thermophilus.nuc
 - t.thermophilus.pep
 - t.thermophilus.pos
 - tth_enzyme.list
 - tth_ko.list
 - tth_ncbi-geneid.list
 - tth_ncbi-gi.list
 - tth_pathway.list
 - tth_pfam.list
 - tth_prosite.list
 - tth_uniprot.list
 - tth_xrefall.list

KEGG



- /genomes/Bacteria/Thermus_thermophilus_HB27/ / ftp.ncbi.nih.gov
 - NC_005835.asn
 - NC_005835.cog
 - NC_005835.faa
 - NC_005835.ffn
 - NC_005835.fna
 - NC_005835.frn
 - NC_005835.gbk
 - NC_005835.gff
 - NC_005835.ppt
 - NC_005835.rnt
 - NC_005835.rpt
 - NC_005835.val
 - NC_005838.GeneMark-2.5f
 - NC_005838.GeneMarkHMM-2.6m
 - NC_005838.Glimmer3
 - NC_005838.asn
 - NC_005838.cog
 - NC_005838.faa
 - NC_005838.ffn
 - NC_005838.fna
 - NC_005838.gbk
 - NC_005838.gff
 - NC_005838.ppt
 - NC_005838.rpt
 - NC_005838.val

NCBI

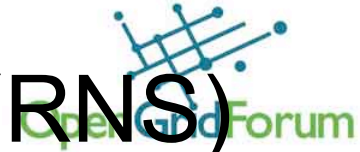


Handling Life Science Resource



- In order to handle these files in the grid environment, we need standardize the metadata of Life Science Resource.
- The activity to define a data(file) type profile for these life science resources should start in the LSG-RG.

Resource Namespace Service (RNS)

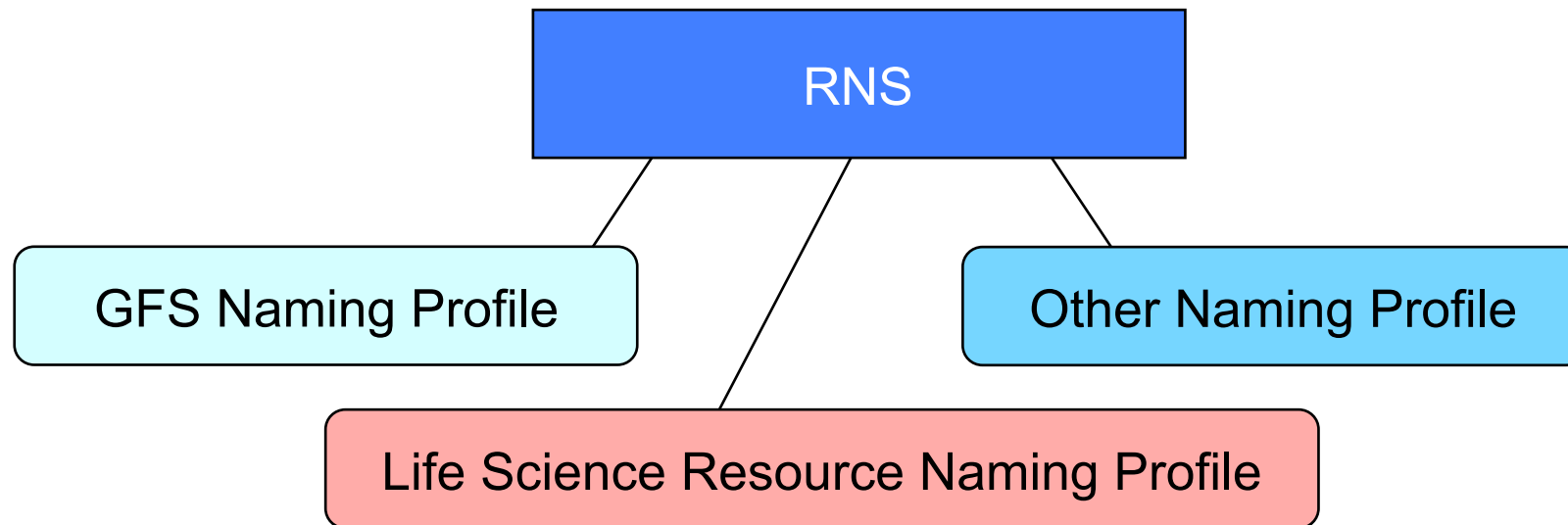


Now Publish! GFD.101

- A multi-faceted approach for addressing the needs of *access to resources* within a distributed network or grid
- Facilitates **a uniform and universal namespace** where:
 - Names resolve to ultimate target addresses
 - Names are hierarchically managed
 - Names may be used in human interface applications and browsable
- Utilizes document style messaging
- Intended to facilitate namespace services for a wide variety of Grid services
 - OGSA Basic Profile compliant Web service capable of providing namespace services for any addressable entity by registering an Endpoint Reference

Thanks, Grid File System Working Group (GFS-WG)

RNS Service Description



RNS Entry Property



All of the following properties **MUST** be implemented to represent properties of a namespace *entry* by an RNS service implementation.

QName	Description
ChildCount	Integer: Number of subentries corresponding to <i>Entry</i> , if and only if <i>Entry</i> is a <i>Virtual Directory</i> ; zero or NULL otherwise.
Description	String: Description of <i>Entry</i>
ModificationTime	DateTime (<code>xsd:dateTime</code>) representation of the last modified timestamp of <i>Entry</i>
Name	String: Representation of the human interface name of <i>Entry</i>

Optional entry property for (Endpoint Reference) Junction

EndpointReferenceList	<code>wsa:EndpointReference[]</code> : List of all Endpoint References associated with <i>Entry</i>
-----------------------	---

Optional entry property for Logical Reference Junction

LogicalName	String: Used to set , add, or retrieve the <i>logical name</i> of a <i>LogicalReference</i>
LogicalResolver	String: Used to set or add a single resolver of a <i>LogicalReference</i>
LogicalResolvers	<code>LogicalResolver[]</code> : List of all resolvers of a <i>LogicalReference</i>

We can standardize the LSR-Profile based on a Resource Namespace Service (RNS)



- RNS is currently at the publish status of OGF now.
- RNS offers a simple standard way of mapping names to endpoints within a grid or distributed network.
- RNS provides a general service for managing a hierarchical name space and an endpoint reference (EPR) .
- GFS-WG plan to publish a RNS Profile for a Grid File System for managing the namespace of federated and virtualized data from file system resources.
- We propose to work with this group and produce a LSG profile for life science data sets.

What do we do concretely?



- Listing up Life Science related files.
- Classify those files.
- Define QName (Qualified Name) for Life Science Resources.
- Make a OGF recommendation document for a Life Science Resource Profile based on RNS.
- Work towards common interoperable life science research across international Grid resources!

Life Science Resource Profile (Draft)



QName	ex.	
Source	NCBI, EBI, NIC, GSC, ...	
Application	blastall hmmer clustalW ...	
Content	GenBank, EMBL, DDBJ, PDB, Pubmed, Swiss-prot, Pfam, Prosite,	
Type	flatfile, indexfile, modelfile, mediafile, outputfile, ...	
Mode	tabler, XML, fasta, ...	
Version	1.0.1	
Identifier	NM_011952 , LSID	
Reference		

Reference



- Tuesday, May 8
 - 4:45-5:30 GFS naming profile (Exchange 4/5)

GFD-R-P.101
GFS-WG

Manuel Pereira, IBM Almaden Research Center
Osamu Tatebe, University of Tsukuba
Leo Luan, IBM Almaden Research Center
Ted Anderson, IBM Almaden Research Center
September 22, 2006
[Revised: March 1, 2007]

Resource Namespace Service Specification

Status of This Document

This document provides information to the Grid community about resource namespace services.
Distribution is unlimited.

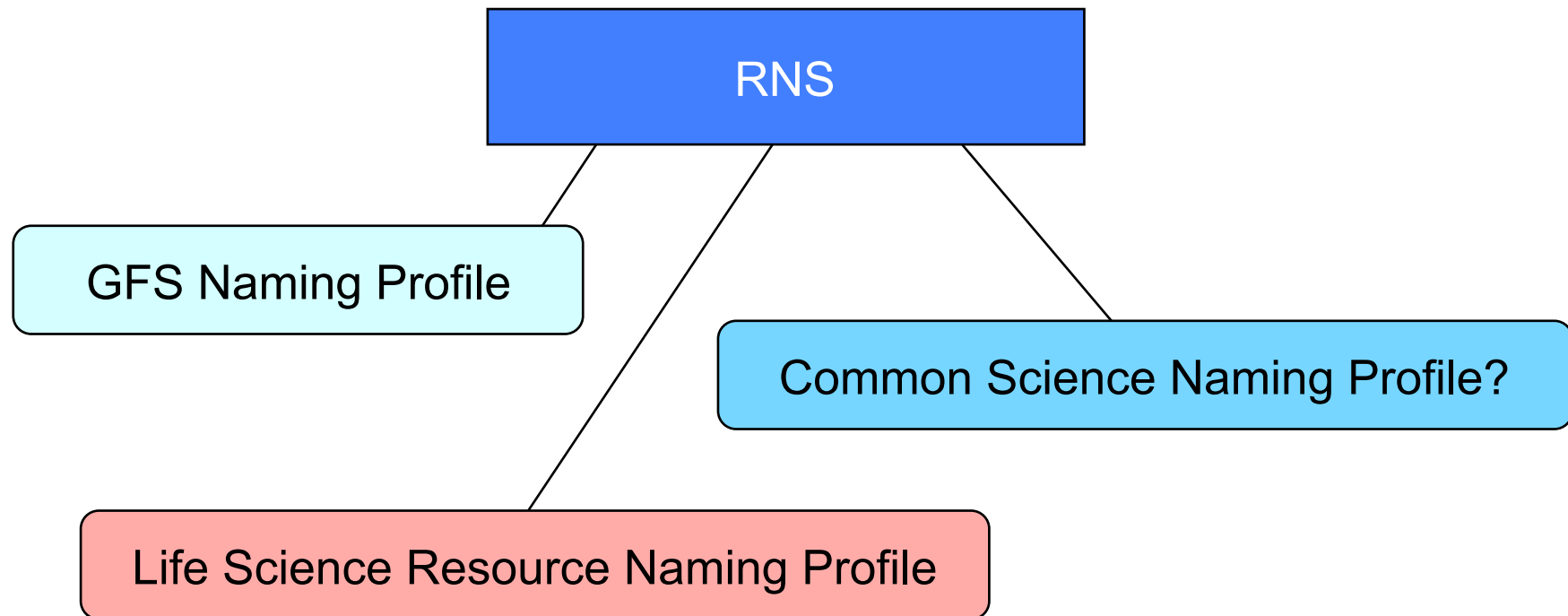
Copyright Notice

Copyright © Open Grid Forum (2006-2007). All Rights Reserved.

Abstract

This document describes the specification of the Resource Namespace Service (RNS), which offers a simple standard way of mapping names to endpoints within a grid or distributed network. It functions as a Web Services registry that provides name-to-resource mapping. Names are mapped to endpoint references and are stored in a persistent namespace repository. RNS offers hierarchical rendering of human-oriented names, which makes resource mappings easily accessible and usable by human interface applications. It can be employed to manage the namespace of federated and virtualized data, services, or effectively any resource capable of being referenced in a grid/web environment. This document proposes a set of operations

RNS Service Description



Contact



- Fumikazu KONISHI
 - Tokyo Tech.
 - Email: konishi@is.titech.ac.jp

Reference



GWD-R
Category: Recommendations

Manuel Pereira – IBM Almaden Research Center
Osamu Tatebe - University of Tsukuba
Leo Luan, IBM Almaden Research Center
Ted Anderson, IBM Almaden Research Center

GFS-WG

May 2006

Resource Namespace Service Specification

Status of This Memo

This memo provides information to the Grid community about resource namespace services. It does not define any standards or technical recommendations. Distribution is unlimited.

Copyright Notice

Copyright © Global Grid Forum (2006). All Rights Reserved.

Full Copyright Notice



Copyright (C) Open Grid Forum (2008). 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.

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