



**Poseidon House
Castle Park
Cambridge CB3 0RD
United Kingdom**

TELEPHONE:
INTERNATIONAL:
FAX:
E-MAIL:

**Cambridge (01223) 515010
+44 1223 515010
+44 1223 359779
apm@ansa.co.uk**

ANSA Phase III

The Name of the Resource

Owen Rees

Abstract

In order to make resources available, or to use resources that have been made available, it is necessary to be able to identify and distinguish the resource in question. This means that both the provider and the user of the resource must be able to name the resource.

This leaves a few questions which are simple to ask but very hard to answer.

What is a resource?

What is a name?

What makes that name a name for that resource, and from whose point of view?

If I have a name for a resource, how do I arrange for you to have a name for that resource that is meaningful to you?

What can I do with the name once I have got it?

This document considers the problems being discussed by the URI working group of the IETF, and relates these to the ANSA Naming Model, and to the naming problems for object oriented distributed systems. The construction of a gateway between the World Wide Web and a CORBA based distributed system requires answers to these questions.

APM.1282.00.01

Draft

12nd August 1994

Request for Comments (confidential to ANSA consortium for 2 years)

Distribution:

Supersedes:

Superseded by:

The Name of the Resource



The Name of the Resource

Owen Rees

APM.1282.00.01

12nd August 1994

The material in this Report has been developed as part of the ANSA Architecture for Open Distributed Systems. ANSA is a collaborative initiative, managed by Architecture Projects Management Limited on behalf of the companies sponsoring the ANSA Workprogramme.

The ANSA initiative is open to all companies and organisations. Further information on the ANSA Workprogramme, the material in this report, and on other reports can be obtained from the address below.

The authors acknowledge the help and assistance of their colleagues, in sponsoring companies and the ANSA team in Cambridge in the preparation of this report.

Architecture Projects Management Limited

Poseidon House
Castle Park
CAMBRIDGE
CB3 0RD
United Kingdom

TELEPHONE UK
INTERNATIONAL
FAX
E-MAIL

(01223) 515010
+44 1223 515010
+44 1223 359779
apm@ansa.co.uk

Copyright © 1994 Architecture Projects Management Limited
The copyright is held on behalf of the sponsors for the time being of the ANSA Workprogramme.

Architecture Projects Management Limited takes no responsibility for the consequences of errors or omissions in this Report, nor for any damages resulting from the application of the ideas expressed herein.

Contents

1	1	Overview of resource naming
1	1.1	What is a resource
1	1.1.1	Definitions
1	1.1.2	Discussion focus and danger of omission
2	1.1.3	Issues for CORBA-like objects or ANSA interfaces as resources
2	1.2	Multiple locations
2	1.2.1	URN ==> URL
3	1.3	Uniformity of names
3	1.3.1	U is for Uniform
3	1.4	The type of the resource

1 Overview of resource naming

This is a collection of notes on the naming of resources in the Internet, based on my understanding of the ANSA naming model, and my reading of the archives of the mailing list <uri@bunyip.com>.¹

1.1 What is a resource

1.1.1 Definitions

The minutes from the CNIDR/URMARC meeting held on January 22, 1993 at Burnsley Hotel, Denver, CO <news:9303222215.AA22834@mocha.bunyip.com> include the following definition:

What is resource?

(entity and object will not be used by this group)
this is **not** necessarily a file
that which is to be located, identified, or described

long list of examples here

files
services
document
etc.

The discussions tend to focus on the names of various kinds, but if we adopt the view that a resource is that which is named by a URN, then a resource is whatever a naming authority decides to name.

In the ANSA naming model, the term “entity” is used for a thing named, and an entity is “any thing that may be of interest”.

These definitions are essentially equivalent, so an “entity” in ANSA terminology is a “resource” in URI/CNIDR/URMARC terminology.

1.1.2 Discussion focus and danger of omission

Much of the discussion on <uri@bunyip.com> tends to assume that the resources are “file-like”², in the sense that they are a body of data that you retrieve. This is the most common case in the WWW, although there are other kinds of resources, such as forms with side-effects.

1. This document contains some references to messages sent to that mailing list; these are of the form <news:9303222215.AA22834@mocha.bunyip.com>, for the convenience of those with access to an NNTP server that has the messages available. Names of this form are the subject of one of the debates on the mailing list. Note that

2. This was pointed out in a message on the list, but I can’t remember which one. I hope to track it down and insert the reference in a future version of the document.

This raises two related issues; considering resources as services provided at interfaces:

1. Does the resource support more than one operation?
2. Does the operation take parameters?

In principle, HTTP can support many methods (operations), of which only two (GET, POST) are used in practice¹. These two methods are typically used to invoke the same single operation of the resource.

Some resources take arguments; the most common being image maps that are sensitive to the position of a click on the image. Forms and searchable indexes also take arguments.

Most of these resources support only retrieval of data. Even if the data is subject to change, this is typically accomplished by means unrelated to the naming of the resource in the WWW, such as the updating of a file in a filesystem.

Some resources have side-effects, the most common being resources that process the arguments captured in a form. The form contains the URL for the resource, but there is nothing to prevent the resource being invoked by other means.

1.1.3 Issues for CORBA-like objects or ANSA interfaces as resources

These resources typically have several operations, the operations typically take parameters, and invoking operations affects the state of the object.

It is the object or interface that is considered to be worthy of naming in a relatively wide context, but the individual operations that are invoked.

This difference of granularity of name is an issue to be resolved.

1.2 Multiple locations

1.2.1 URN ==> URL

One of the established features of the URI work is that there may be multiple copies of a resource at different locations. Each location has its own URL, but the resource as a whole is named by a single URN.

There has been much discussion of variants and encodings in the copies of the resource, but the common assumption is that one copy of the resource at one location will be accessed when access is required.

One current proposal is that there is a service that will convert the URN that is effectively opaque to the client into a URC - a Uniform Resource Characteristic - which typically includes some number of URLs. For each URL there would be additional data describing variant, encoding and various other properties.

The common assumption is that the client retrieves the URC and then picks one of the URLs based upon the properties, should it wish to access the resource.

1. I am not aware of any HTTP clients or servers that actually support the use of other methods, whether defined in the HTTP spec, or otherwise.

This view fits well with a current problem in the WWW. There are some resources that exist in multiple locations, each with its own URL. There is no way to refer to the underlying resource itself.

The assumption that any one copy will do is consistent with the view that resources are data to be observed. The assumption will break down when the updating of resources is brought within the same scheme.

1.3 Uniformity of names

1.3.1 U is for Uniform

The need for uniformity of names arises because names are passed by means that cannot translate the name to a name appropriate to the new context. One of the requirements of URNs is that it be possible to pass them in ink on bar napkins, without benefit of automation. This has led to a great deal of discussion of character sets, encoding of whitespace, and avoidance of transcription errors.

The key issue here is that you have to construct the naming contexts in such a way that URNs have the same meaning in all the contexts that may be used to resolve them. The option of generating a new name appropriate to the context is not available.

At present URLs are used in a way that imposes this constraint - they occur in email, newsgroup messages, printed newspapers and other similar media. You get the URL that works for someone else - it had better work for you too.

Locator or addressing information retrieved by looking up a URN need not have this property, provided that it is not revealed in a such a way that people would pass it on to inappropriate contexts.

1.4 The type of the resource

Resources, even in the limited case of “file-like” resources, have a type. One of the debates (e.g. <news:93Oct22.222549pdt.2795@golden.parc.xerox.com>) is whether or not type information should appear in URLs.

Type information is necessary to make use of the outcome returned from accessing a resource; current WWW practice has various ways of determining a type to use.

1. HTTP 1.0 returns headers that include a MIME compatible Content-Type field, thus the type is returned as part of the response. The outgoing request specifies which types are acceptable.
2. The HTTP standard [BERNERS-LEE 93] says that when delivered a pre-HTTP/1.0 response, clients should assume that the response is “a document body in type text/html”, but the practice was to guess the type from the “file extension” at the end of the URL.
3. For FTP URLs, clients typically guess the type from the file extension.
4. Gopher URLs include type information. In the words of <news:9310260057.AA28871@boombox.micro.umn.edu> “the type information is (in this case) explicitly coded into the URL path as a separate item before the gopher selector.”

In these cases¹, the type is the type of the data returned as the single result of the retrieval operation defined implicitly by the access protocol. The typical types are plain text, HTML markup, PostScript, various graphic formats etc.

Type is one of the pieces of meta-information associated with a resource. Meta-information, and how it is presented and used has been the subject of discussion (e.g. <news:199310192100.AA10993@oit.oit.gatech.edu>).

1. I am not sure about Gopher, not having used it to any significant extent.

References

[BERNERS-LEE 93]

Berners-Lee, T.; "Hypertext Transfer Protocol", draft-ietf-iiir-http-00.txt, <ftp://cnri.reston.va.us/internet-drafts/draft-ietf-iiir-http-00.txt>.

