

Groupe de travail Réseau
Request for Comments : 2056
Catégorie : En cours de normalisation
Traduction Claude Brière de L'Isle

R. Deneberg, Library of Congress
J. Kunze, University of California, San Francisco
D. Lynch, SilverPlatter Information Ltd.
novembre 1996

Localisateurs de ressource uniformes pour Z39.50

Statut du présent mémoire

Le présent document spécifie un protocole de l'Internet en cours de normalisation pour la communauté de l'Internet, et appelle à des discussions et suggestions pour son amélioration. Prière de se référer à l'édition en cours des "Normes officielles des protocoles de l'Internet" (STD 1) pour connaître l'état de la normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

1. Introduction

Z39.50 est un protocole de restitution d'informations qui ne tient pas bien dans un modèle de restitution conçu principalement autour d'une recherche sans état des données. Son modèle est plutôt une enquête générique de l'utilisateur orientée session, de tâche en plusieurs étapes, chaque étape pouvant être suspendue temporairement pendant que le serveur demande des paramètres supplémentaires au client avant de continuer. Certaines, aucune, ou toutes ces interactions client/serveur peuvent exiger la participation de l'utilisateur client, en fonction du seul logiciel du client (le protocole lui-même n'ayant pas une telle exigence).

D'un autre côté, la restitution de données "bien connues" peut être effectuée en une seule étape, c'est-à-dire avec une session Z39.50 dégénérée consistant en exactement une demande et réponse de recherche du protocole. Au delà du sous service de recherche basique, il y a plusieurs sous services auxiliaires (par exemple, examen, établissement/suppression de résultats). Parmi les fonctions couvertes par la combinaison des sous services, deux fonctions centrales émergent qui sont traitées de façon appropriée par deux schémas d'URL distincts : l'URL Session et l'URL Restitution.

Utiliser deux schémas au lieu d'un seul fait une distinction critique entre un URL Session Z39.50, qui ouvre une session de client initialisée pour une utilisation interactive par l'utilisateur, et un URL Restitution Z39.50, qui ouvre et ferme une session de client pour restituer un élément d'information spécifique. Faire cette distinction au niveau du schéma permet à l'interface d'utilisateur de la refléter à l'utilisateur, sans exiger que l'interface d'utilisateur analyse des parties par ailleurs opaques de l'URL (ce qui est cohérent avec la pratique actuelle).

2. Concepts de base

Cette section décrit brièvement l'utilisation de la terminologie spécifique de Z39.50 au sein des définitions d'URL ci-dessous : spécifiquement, les termes "base de données", "elementset", "recordsyntax", et "docid".

Le protocole Z39.50 spécifie diverses opérations de restitution d'informations dont les deux plus basiques sont Search et Present. Dans une opération Search, un client fournit un critère de recherche et indique une base de données (ou plusieurs bases de données) à rechercher sur le serveur. Le résultat essentiel d'une opération Search est un ensemble de résultats qui est créé chez le serveur, consistant en pointeurs sur les enregistrements de la base de données choisie.

Z39.50 modèle des enregistrements de base de données, abstrait des enregistrements de base de données, et restitue des enregistrements. Un enregistrement de base de données est une unité d'informations dans une base de données, représentée dans une structure de données locale du serveur. Un enregistrement abstrait de base de données est une représentation abstraite de ces informations, où le client et le serveur partagent une compréhension commune de la représentation. Cela permet que les éléments logiques soient adressés et sélectionnés pour le transfert, via une spécification d'ensemble d'éléments, ou, comme utilisé ci-dessous via un "ensemble d'éléments" (*elementset*). Un enregistrement de restitution est l'ensemble des éléments sélectionnés mis en paquet dans une structure exportable, par l'application d'une "syntaxe d'enregistrement" (*recordsyntax*).

Une opération Search résulte donc en entrées pointant sur des enregistrements de la base de données ; via une opération Present, un client demande un enregistrement de restitution, correspondant à un enregistrement de la base de données, correspondant à une entrée dans l'ensemble résultant. Le client indique la composition et le format de l'enregistrement de restitution en spécifiant respectivement un elementset et une recordsyntax.

Un cas particulier de recherche de Z39.50 est celle d'un "élément connu", lorsque un client a l'intention qu'une recherche identifie un seul enregistrement connu de la base de données, ou "document" (pour les besoins de l'illustration, on suppose que l'enregistrement de la base de données correspond à un document) et de plus, le client connaît un identifiant pour le document qui peut être utilisé pour effectuer cette recherche d'élément connu. Dans ce cas, cet identifiant est souvent appelé un identifiant de document, ou "docid".

3. URL de session Z39.50

L'URL Session de Z39.50 peut être décrit de façon informelle comme fournissant le mécanisme pour faire passer l'utilisateur à une application client Z39.50.

- Le paramètre Hôte est exigé.
- L'accès est facultatif et est 210 par défaut.
- Tous les autres paramètres sont facultatifs.
- Le client Z39.50 va commencer une session à l'hôte/accès spécifié (autrement, il n'a pas besoin de commencer explicitement une session, mais peut à la place utiliser une session déjà ouverte sur le même hôte/accès).
- Une base de données doit être incluse si un docid est inclus.
- Si un docid est inclus, le client va effectuer la recherche spécifiée (de la même manière que pour l'URL de restitution, spécifié ci-dessous).
- Si un docid n'est pas inclus, et si d'autres paramètres (autres que hôte/accès) sont spécifiés, le client peut utiliser ces paramètres comme "indications". Divers clients peuvent choisir de les traiter comme des exigences, ou des préférences, ou de les ignorer.
- Dans tous les cas (qu'une recherche soit effectuée ou non) le client va laisser la session Z39.50 ouverte pour que l'utilisateur fasse des restitutions, de nouvelles recherches, etc. (C'est la principale distinction avec l'URL Restitution qui laisse au choix du client de garder ou non la session Z39.50 ouverte.)

4. URL Restitution de Z39.50

L'URL Restitution de Z39.50 est destiné à permettre d'utiliser une session Z39.50 comme mécanisme de transfert transparent pour restituer un objet d'information spécifique. Un client Z39.50 utilise les informations de l'URL pour formuler une demande de recherche (*Search*). La réponse de recherche du serveur indique combien d'enregistrements satisfont la demande. Si le nombre d'enregistrements correspondants n'est pas égal à un, la restitution est considérée comme un échec, et le comportement de l'application du client n'est pas défini. Si le nombre d'enregistrements satisfaisants est égal à un, le serveur peut avoir inclus l'enregistrement désiré dans la réponse de recherche. Sinon, le client demande la transmission de l'enregistrement avec une demande Present. Après que le client a reçu l'enregistrement spécifié, il peut clore immédiatement la session Z39.50, ou la garder ouverte pour des restitutions ultérieures.

- Hôte est exigé.
- Accès est facultatif est prend par défaut 210.
- Une base de données est exigée.
- La signification d'un URL de restitution sans docid est indéfinie.
- Le docid est placé dans une interrogation de type 1, comme terme unique, dans le format général (tag 45), en utilisant l'attribut Bib-1 établi, avec une valeur d'attribut Use de docid, et un attribut structure de URx. La chaîne docid est définie par le serveur et complètement opaque pour le client.
- Si le nom d'ensemble d'élément (esn, *element set name*) n'est pas spécifié, c'est le choix du client. Si l'esn est spécifié, il devrait être utilisé soit dans la demande Search pour la valeur de small- et/ou medium-set-element-set-names ou dans une demande Present qui suit une Search. Ces termes et leur signification sont définis dans la norme Z39.50 [2].
- Si la syntaxe d'enregistrement (rs, *record syntax*) n'est pas spécifiée, c'est le choix du client. Si une ou plusieurs syntaxes d'enregistrement sont spécifiées, le client devrait en choisir une (de préférence la première de la liste de celles qu'il prend en charge) et l'utiliser dans une demande Search ou Present comme valeur de PreferredRecordSyntax.

5. BNF pour les URL Z39.50

Les URL Session et Retrieval de Z39.50 suivent la syntaxe commune de schéma Internet telle que définie dans la RFC1738, "Localisateurs de ressource uniformes (URL)" [1]. Dans la définition, les littéraux sont cités avec des "guillemets", les éléments facultatifs sont inclus entre [crochets], "|" est utilisé pour désigner des solutions de remplacement, et des éléments peuvent être précédés de <n>* pour désigner n ou plus répétitions de l'élément suivant ; par défaut, n est 0.

```

z39.50url = zscheme "://" host [":" port]
           [ "/" [database *["+" database]
                ["?" docid]]
           [";esn=" elementset]
           [";rs=" recordsyntax *["+" recordsyntax]]]

zscheme   = "z39.50r" | "z39.50s"
database  = uchar
docid     = uchar
elementset = uchar
recordsyntax = uchar

```

Les futures extensions de ces URL seront de la forme [;motclé=valeur].

Les définitions sont tirées de la RFC1738. Entre la version du projet Internet et la RFC1738 publiée, deux modifications pertinentes ont été faites : '=' a été déplacé de la classe de caractères <extra> à la classe <réserve>, et <national> a été retiré de <non réservé>. Ni <national> ni <ponctuation> ne sont mentionnés dans le présent document ni dans la RFC1738.

```

lowalpha  = "a" | "b" | "c" | "d" | "e" | "f" | "g" | "h" | "i" | "j" | "k" | "l" | "m" | "n" | "o" | "p" | "q" | "r" | "s" | "t" | "u" | "v"
           | "w" | "x" | "y" | "z"
hialpha   = "A" | "B" | "C" | "D" | "E" | "F" | "G" | "H" | "I" | "J" | "K" | "L" | "M" | "N" | "O" | "P" | "Q" | "R" | "S" |
           | "T" | "U" | "V" | "W" | "X" | "Y" | "Z"
alpha     = lowalpha | hialpha
digit     = "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9"
safe      = "$" | "-" | "_" | "." | "+"
extra     = "!" | "*" | "" | "(" | ")" | ";"
national  = "{" | "}" | "|" | "\" | "^" | "~" | "[" | "]" | ""
ponctuation = "<" | ">" | "#" | "%" | "<"
réserve   = ";" | "/" | "?" | ":" | "@" | "&" | "="
hex       = digit | "A" | "B" | "C" | "D" | "E" | "F" | "a" | "b" | "c" | "d" | "e" | "f"
escape    = "%" hex hex
non réservé = alpha | digit | safe | extra
uchar     = unreserved | escape
xchar     = unreserved | reserved | escape
digits    = 1*digit

```

6. Considérations pour la sécurité

Les deux schémas d'URL Z39.50 sont soumis aux mêmes implications de sécurité que le schéma général d'URL [1], de sorte que les précautions usuelles s'appliquent. Cela signifie, par exemple, qu'un localisateur peut ne plus pointer sur l'objet sur lequel il était prévu à l'origine. Cela signifie aussi qu'il serait possible de construire un URL de telle sorte qu'une tentative d'effectuer une opération idempotente sans dommage telle que la restitution d'un objet va en fait causer la survenance d'une opération distante potentiellement dommageable.

7. Remerciements

Le groupe de mise en œuvre de Z39.50 a fourni la substance du présent document.

8. Références

- [1] T. Berners-Lee, L. Masinter, M. McCahill (éditeurs), "[Localisateurs uniformes de ressource](#) (URL)", RFC1738, décembre 1994. (*P.S., Obsolète, voir les RFC4248 et 4266*)
- [2] ANSI/NISO Z39.50-1995, "ANSI Z39.50: Information Retrieval Service and Protocol", 1995. <ftp://ftp.loc.gov/pub/z3950/>
- [3] ANSI/NISO Z39.50-1992, "ANSI Z39.50: Information Retrieval Service and Protocol", 1992. <ftp://ftp.cni.org/pub/NISO/docs/Z39.50-1992/www/Z39.50.toc.html> ; (aussi disponible sur papier auprès de Omnicom)

Information Service, 115 Park St., SE, Vienna, VA 22180 USA).

9. Adresse des éditeurs

Ray Deneberg
Library of Congress
Collections Services
Network Development/MSO
Washington DC 20540
téléphone : (202) 707-5795
fax : (202) 707-0115
mél : ray@rdn.loc.gov

John A. Kunze
Center for Knowledge Management
University of California, San Francisco
530 Parnassus Ave, Box 0840
San Francisco, CA 94143-0840
téléphone : (415) 502-6660
fax : (415) 476-4653
mél : jak@ckm.ucsf.edu

Denis Lynch
SilverPlatter Information Ltd.
10 Barely Mow Passage
Chiswick, London W4 4PH
U.K.
téléphone : +44 (0)181-995-8242
fax : +44 (0)181-995-5159
mél : DenisL@SilverPlatter.com

Appendice Exemples d'URL Z39.50

Un URL de base de session Z39.50 que pourrait utiliser un client pour ouvrir une connexion avec le catalogue de l'union MELVYL "cat" à l'Université de Californie est :

`z39.50s://melvyl.ucop.edu/cat`

Un URL qui ouvrirait la base de données des magazines MELVYL juste assez longtemps pour aller chercher un article tiré du volume 30, numéro 19 d'un périodique hypothétique pourrait ressembler à :

`z39.50r://melvyl.ucop.edu/mags?elecworld.v30.n19`

Un dernier exemple nous montrera une autre restitution d'un URL que pourrait utiliser un client pour demander un enregistrement complet (ensemble d'élément "f") dans la syntaxe MARC d'une base de données hypothétique appelée TMF au CNIDR :

`z39.50r://cnidr.org:2100/tmf?bkirch_rules__a1;esn=f;rs=marc`

Comme dans l'exemple précédent, la partie de la chaîne après le '?' est déterminée par le serveur. Dans cet exemple, le serveur fonctionne sur l'accès non standard 2100.