

Groupe de travail Réseau
Request for Comments : 1736
Catégorie : Information

J. Kunze, IS&T, UC Berkeley
février 1995
Traduction Claude Brière de L'Isle

Recommandations fonctionnelles pour les localisateurs de ressource de l'Internet

Statut de ce mémoire

Le présent mémoire apporte des informations pour la communauté de l'Internet. Il ne spécifie aucune sorte de norme de l'Internet. La distribution du présent mémoire n'est soumise à aucune restriction.

1. Introduction

Le présent document spécifie un ensemble minimum d'exigences pour les localisateurs des ressources de l'Internet, qui portent les informations de localisation et d'accès des ressources. Des exemples typiques de ressources incluent des documents accessibles par le réseau, des bases de données WAIS, des serveurs FTP, et des destinations Telnet.

Les localisateurs peuvent s'appliquer aux ressources qui ne sont pas toujours ou ne sont jamais accessibles par le réseau. Des exemples de ces derniers incluent des êtres humains et des objets physiques qui n'ont pas d'instanciation électronique (c'est à dire, des objets sans une existence complètement définie par des objets numériques tels que des fichiers de disques).

Un localisateur de ressource est une sorte d'identifiant de ressource. D'autres sortes d'identifiants de ressource permettent que des noms et des descriptions soient associés aux ressources. Un nom de ressource est destiné à fournir un moyen stable de se référer à une ressource longtemps après que la ressource elle-même a été déplacée ou peut-être n'a plus d'existence. Une description de ressource comporte un corps de méta informations pour aider à la recherche et à la sélection de la ressource.

Dans le présent document, un localisateur de ressource Internet est un localisateur défini par une norme de localisation de ressource de l'Internet. Une norme de localisation de ressource conjuguée à une description de ressource et une norme de désignation de ressource spécifie une infrastructure complète pour la dissémination d'informations fondées sur le réseau. Les mécanismes pour transposer les localisateurs en noms, et en identifiants descriptifs sortent du domaine d'application du présent document.

2. Vue générale du problème

Les fournisseurs de ressources d'informations fondées sur le réseau exigent une méthode de description de la localisation et de l'accès à leurs ressources. Les utilisateurs des systèmes d'informations ont besoin d'une méthode par laquelle le logiciel client puisse interpréter les descriptions d'accès et de localisation des ressources en leur nom d'une façon relativement transparente. Sans une telle méthode, transparente et largement répartie, l'accès ouvert aux informations sur l'Internet serait difficile sinon impossible.

2.1 Définition du localisateur général de ressource

Les exigences énumérées dans le présent document imposent des restrictions aux localisateurs généraux de ressources. Pour mieux comprendre ce qu'est le localisateur de ressource de l'Internet, la définition suivante de localisateur général donne quelques indications.

Définition : Un localisateur de ressource général est un objet qui décrit la localisation d'une ressource.

Cette définition admet délibérément de nombreux degrés de liberté afin d'englober les recherches plus approfondies du débat à long terme sur les normes de localisation de ressource. Vaste comme il est, cet espace de problème est une toile de fond utile pour la discussion sur les exigences (qui est faite plus loin) qui génère un espace de problème plus restreint et plus gérable. Une norme de localisation de ressource restreint encore l'espace en appliquant des exigences supplémentaires.

Considérons la définition selon quatre parties : (1) un localisateur de ressource général est un objet (2) qui décrit (3) la localisation (4) d'une ressource.

2.1.1 Un localisateur général de ressource est un objet...

L'objet pourrait être une structure de données complexe. Il pourrait être une séquence d'octets contigus. Il pourrait être une paire de coordonnées de latitude-longitude, ou une carte routière en trois couleurs imprimée sur papier. Il pourrait être une séquence de caractères qui sont imprimables.

2.1.2 ...qui décrit

Dans le cas le plus général, il y a de nombreuses façons dont un localisateur de ressource pourrait décrire la localisation. Il pourrait employer une description en langage graphique ou naturel. Il pourrait être fortement codé ou compressé. Il pourrait être légèrement codé et directement compréhensible par les humains. La description pourrait être une hiérarchie multi niveaux avec une sémantique commune à chaque niveau. Elle pourrait être une hiérarchie multi niveaux avec une sémantique commune aux seuls deux premiers niveaux, où la sémantique au dessous du second niveau dépend de la valeur donnée au premier niveau. Ce ne sont que quelques unes des possibilités.

2.1.3 ...la localisation

Un localisateur de ressource décrit une localisation mais ne garantit jamais que l'accès peut être établi. Bien que l'accès soit souvent désiré lorsque les clients suivent les instructions de localisation données dans un localisateur de ressource conforme, la ressource n'a pas besoin d'exister encore, ou déjà. Bien sûr, elle peut ne jamais exister, quand bien même le localisateur continuerait de décrire une localisation où une ressource pourrait exister (par exemple, elle pourrait être utilisée comme un fourre-tout où la disponibilité de la ressource va dépendre d'un événement tel qu'un paiement).

De plus, la nature de certaines ressources potentielles, en particulier des êtres animés ou des objets physiques sans instanciation électronique, rend dans certains cas l'accès réseau non significatif ; de telles ressources ont des localisateurs qui impliqueraient un accès qui ne serait pas par le réseau, mais là encore, l'accès n'est pas garanti.

2.1.4 ...d'une ressource.

Une ressource peut être de nombreuses choses. Au delà des ressources qui ne sont pas sur le réseau ou des ressources non électroniques qu'on vient de mentionner, des exemples familiers sont un document électronique, une image, un serveur (par exemple, FTP, Gopher, Telnet, HTTP) ou une collection d'éléments (par exemple, un menu Gopher, un répertoire FTP, une page HTML). D'autres exemples accompagnent des protocoles multi fonctions tels que Z39.50, qui peut effectuer un accès réseau en un seul aller-retour, des approfondissements de recherche en mode session, et un feuilletage d'indice.

2.2 Producteurs et interpréteurs de localisateurs de ressource

La question de l'analyse est centrale dans la discussion des exigences des localisateurs. C'est la capacité pour un agent de reconnaître ou comprendre un localisateur en tout ou en partie. La discussion peut être aidée par une claire distinction des deux principales actions associées aux localisateurs.

Les localisateurs de ressource sont à la fois produits et interprétés. Les producteurs sont liés par les normes de localisation de ressource qui sont à leur tour liées par les exigences formulées dans le présent document. Les interpréteurs de localisateurs ne sont pas liés par les exigences, ils en sont les bénéficiaires.

2.2.1 Interpréteurs de localisateurs de ressource

Un localisateur de ressource est interprété par des agents d'interprétation, qui dans le présent document sont simplement appelés des interpréteurs. Les interpréteurs peuvent être des êtres humains ou des logiciels. Le long du chemin qui mène à l'établissement de l'accès sur la base des informations contenues dans un localisateur, un ou plusieurs interpréteurs peuvent être employés. Certains exemples de plusieurs interpréteurs qui traitent un seul localisateur illustrent le concept qu'un localisateur de ressource peut n'être compréhensible qu'en partie par chacun des différents interpréteurs, mais est compréhensible dans sa totalité par une combinaison d'interpréteurs.

Dans le premier exemple, un logiciel interpréteur reconnaît assez d'un localisateur pour comprendre à quel agent externe il doit le transmettre. Ici, l'agent externe peut être un usager et le localisateur le numéro de téléphone d'une bibliothèque ; le logiciel transmet le localisateur simplement en l'affichant. L'agent peut être une couche logicielle du réseau qui est spécialisée dans un certain protocole de communications ; une fois que le service est reconnu, le localisateur lui est retransmis avec une demande d'accès.

Dans un autre exemple, un interpréteur humain peut aussi reconnaître assez d'un localisateur pour comprendre où le transmettre. Ici, la personne peut être un usager qui reconnaît comme tel le numéro de téléphone d'une bibliothèque mais qui ne comprend pas les informations de localisation qui sont codées dedans ; la personne le transmet à un employé de la bibliothèque (un agent externe) qui sait comment établir l'accès à la ressource de la bibliothèque.

Un pré requis pour interpréter un localisateur est de comprendre quand l'objet en question est en fait un localisateur, ou contient un ou plusieurs localisateurs. Certaines contraintes d'environnement rendent facile la réponse à cette question, par

exemple, au sein des ancres HTML ou des éléments de menu Gopher. Des environnements moins contraints, comme ceux où il n'y a pas de texte, rendent la réponse plus difficile sans des hypothèses bien définies. Une norme de localisateur de ressource doit faire ces hypothèses de façon très explicite.

2.2.2 Producteurs de localisateurs de ressource

Les localisateurs de ressource sont produits de nombreuses façons différentes, souvent par un agent qui les interprète aussi. Le fournisseur d'une ressource peut produire un localisateur pour elle, laissant ce localisateur dans des endroits où il est destiné à être découvert, comme une page HTML, un menu Gopher, ou une annonce dans une liste d'un message électronique.

Les non producteurs de ressources peuvent être les producteurs majeurs de localisateurs ; par exemple, un logiciel client de la Toile mondiale produit des localisateurs en traduisant des localisateur de ressource étrangers (par exemple, des éléments de menu Gopher) dans son propre format. Certaines bases de données de localisateurs (par exemple, Archie) ont été entretenues par des processus automatiques qui produisent des localisateurs pour des centaines de milliers de ressources FTP qu'il "découvrent" sur l'Internet.

Les utilisateurs sont les producteurs majeurs de localisateurs de ressources. Un utilisateur qui en construit un pour le partager avec d'autres est responsable de sa conformité aux normes de localisateur. Parfois, un usager compose un localisateur de ressource dans un souci de politesse et le soumet au logiciel client avec l'intention d'établir l'accès. Cet usager est un producteur, dans un certain sens, mais si le localisateur est pour sa consommation personnelle, l'usager n'est pas tenu par les exigences. En fait, certains logiciels clients peuvent offrir le service de traduire des localisateurs abrégés non conformes entrés par les usagers dans des instructions d'accès réussies ou dans les localisateurs conformes (par exemple, en ajoutant un nom de domaine à un nom d'hôte non qualifié).

2.3 Unicité des localisateurs de ressource

Le sujet de l'exigence "d'unicité" pour les localisateurs de ressource a été beaucoup discuté. Le présent document considère les aspects suivants de l'unicité, mais les rejette délibérément en tant qu'exigences. Il incombe à une norme de localisateur de ressource qui traite de ce sujet d'être claire sur les aspects dont elle traite.

2.3.1 Unicité et copies multiples d'une ressource

L'exigence d'unicité peut imposer que n'existe aucune copie identique d'une ressource. Le présent document ne formule pas une telle exigence.

2.3.2 Unicité et accès déterministe

Une exigence d'unicité pourrait imposer que la ressource à laquelle une tentative a donné accès soit aussi le résultat identique de toute autre tentative réussie. Le présent document ne formule pas une telle exigence, pas plus qu'il ne définit de "caractère identique". Il n'est pas approprié qu'une norme de localisateur de ressource définit de "caractère d'identité" parmi les ressources.

2.3.3 Unicité et localisateurs multiples

Une exigence d'unicité pourrait imposer qu'une ressource n'ait pas plus d'un localisateur sauf si ces localisateurs sont le même. Le présent document ne formule pas une telle exigence, pas plus qu'il ne définit ce que peut être "le même" parmi les localisateurs (ce qu'une norme pourrait faire en utilisant, par exemple, des règles de canonisation).

2.3.4 Unicité, ambiguïté, et multiplicité d'objets par accès

Une exigence d'unicité pourrait imposer qu'un localisateur de ressource identifie exactement un objet par opposition à plusieurs objets. Le présent document ne donne pas de définition générale de ce que constitue un objet, plusieurs objets, ou un objet consistant en plusieurs objets.

3. Accès et disponibilité de la ressource

Un localisateur ne garantit jamais l'accès, mais l'établissement de l'accès est de loin l'application la plus importante qui est visée par un localisateur de ressource. Bien qu'il soit considéré comme inélégant d'annoncer un localisateur pour une ressource qui ne sera jamais accessible (que ce soit une ressource "en réseau" ou non) il est normal que l'accès à une ressource échoue à un taux qui augmente avec l'âge du localisateur utilisé.

L'accès à une ressource peut échouer pour de nombreuses raisons. Les fournisseurs affectent fondamentalement l'accessibilité en déplaçant, remplaçant, ou supprimant les ressources avec le temps. La fréquence de tels changements dépend entre autres choses de la nature de la ressource et des pratiques des fournisseurs de service. Un localisateur qui se conforme à une norme de localisation mais échoue pour une de ces raisons est appelé "invalide" pour les besoins du présent

document ; le terme de localisateur invalide ne s'applique pas aux localisateurs mal formés ou non conformes. Les normes sur la dénomination des ressources traitent du problème des localisateurs invalides.

Les politiques ordinaires de prise en charge par les fournisseurs peuvent être cause que des ressources sont inaccessibles durant des périodes prévisibles (par exemple, à certaines heures de la journée, ou certains jours de l'année) ou durant des périodes de lourde charge du système. Des restrictions de droits d'accès impossibles à exprimer dans un localisateur affectent aussi l'accessibilité pour certaines populations d'utilisateurs. Une lourde surcharge du réseau peut aussi empêcher l'accès. Dans de telles situations, le présent document dit qu'une ressource est "indisponible". Un localisateur peut être à la fois valide et identifier une ressource qui est indisponible. Les normes de description des ressources traitent, entre autres choses, de certains aspects de la disponibilité d'une ressource.

En général, la probabilité qu'un localisateur de ressource donné conduise à un accès réussi décroît avec le temps, et dépend de conditions telles que la nature de la ressource, des politiques de prise en charge du fournisseur, et de la charge du réseau.

4. Liste d'exigences pour les localisateurs de ressource Internet

Cette liste d'exigences s'applique à l'ensemble des localisateurs généraux défini au paragraphe 2.1. Le sous-ensemble résultant, appelé localisateurs Internet dans le présent document, devra être précisé par une norme de localisation des ressources de l'Internet. Certaines exigences concernent le codage de localisateur alors que d'autres concernent la fonction de localisateur.

Une exigence qui provient de la liste du projet d'origine a été abandonnée après qu'une longue discussion a révélé qu'elle était impossible à satisfaire en pratique. Elle déclarait que le logiciel devait, avec un haut degré de fiabilité, reconnaître les localisateurs Internet dans certains environnements relativement non structurés, tels que au sein d'un texte ASCII.

4.1 Les localisateurs sont transitoires

La probabilité avec laquelle un localisateur de ressource Internet donné conduit à un accès réussi décroît avec le temps. Des schémas plus stables d'identifiants de ressource sont traités dans les normes de désignation de ressources et sortent du domaine d'application d'une norme sur la localisation de ressource.

4.2 Les localisateurs ont une portée mondiale

L'espace de nom de localisateur de ressource inclut le monde entier. La probabilité de réussite d'un accès en utilisant un localisateur Internet ne dépend en aucune façon, modulo la disponibilité de la ressource, de la localisation géographique ou Internet du client.

4.3 Les localisateurs sont analysables

Les localisateurs Internet peuvent être scindés en parties constitutives complètes suffisantes pour que les interpréteurs (logiciels ou humains) tentent d'y accéder si ils le désirent. Bien que ces exigences ne lient pas les interpréteurs, trois points méritent d'être soulignés :

4.3.1 Un certain type de localisateur peut être analysable même si un interpréteur donné ne peut pas l'analyser.

4.3.2 Analysable par les utilisateurs n'implique pas une analyse directe par des usagers non entraînés.

4.3.3 Un localisateur donné n'a pas besoin d'être complètement analysable par un interpréteur quelconque si une combinaison d'interpréteurs peut l'analyser complètement.

4.4 Les localisateurs sont directement distingués des identifiants de désignation et de description qui peuvent occuper le même espace de nom.

Durant une période de transition (d'une durée éventuellement indéfinie) d'autres sortes d'identifiants de ressource sont supposés coexister dans les structures de données avec les localisateurs Internet.

4.5 Les localisateurs sont "faciles à transporter"

Les localisateurs Internet peuvent être transmis d'un usager à l'autre (par exemple, via la messagerie électronique) à travers les protocoles de communication standard de l'Internet sans perte ni corruption d'informations.

4.6 Les localisateurs sont transcrivables en langage humain

Les usagers peuvent copier les localisateurs Internet d'un support à un autre (comme de la voix au papier, ou du papier au clavier) sans perte ni corruption des informations. Ce processus n'est pas obligatoirement confortable.

4.7 Un localisateur Internet consiste en un paquetage d'un service et d'un paramètre opaque

Le paquetage de paramètres n'a de signification que pour le service avec lequel il est apparié, où un service est une méthode d'accès abstraite. Une méthode d'accès abstraite peut être un outil logiciel, une institution, ou un protocole réseau. Le paquetage de paramètres peut être des instructions d'accès spécifiques du service. Afin de protéger le développement créatif de nouveaux services, il y a une classe extensible de services pour lesquels aucune sémantique de paquetage de paramètre commune à travers les services ne peut être supposée.

4.8 L'ensemble de services est extensible.

De nouveaux services peuvent être ajoutés à l'avenir.

4.9 Les localisateurs ne contiennent pas d'information sur la ressource autres que celles requises par le mécanisme d'accès

L'objet d'un localisateur Internet est seulement de décrire la localisation d'une ressource, et non pas d'autres propriétés comme son type, sa taille, sa date de modification, etc. Celles-ci et ses autres propriétés relèvent d'une norme de description de ressource.

5. Considérations pour la sécurité

Bien que les exigences n'aient pas d'implications directes pour la sécurité, les applications qui se fondent sur la conformité aux normes peuvent avoir besoin de prendre en compte deux faiblesses potentielles. D'abord, comme les localisateurs sont éphémères, un client qui utilise un localisateur invalide pourrait involontairement obtenir l'accès à une ressource qui n'est pas la cible espérée. Par exemple, lorsque un nom d'hôte n'est plus enregistré pendant un certain temps, puis est enregistré à nouveau, un localisateur qui n'était plus valide durant cette période peut à nouveau conduire à une ressource, mais peut-être une qui prétend seulement être la ressource originale.

Ensuite, comme un localisateur consiste en un service et un paquetage de paramètres, une liberté de traitement potentiellement énorme est autorisée, selon le service individuel. Un serveur est vulnérable si il ne restreint pas de façon convenable ses paramètres d'entrée. Par exemple, un serveur qui annonce des localisateurs pour certains objets du système de fichiers local peut ouvrir par inadvertance une porte à travers laquelle on peut accéder à d'autres objets du système de fichiers.

Un client est aussi vulnérable si il ne comprend pas les limitations du service qu'il utilise. Par exemple, un client qui fait confiance à un localisateur pour obtenir une ressource d'une source incertaine pourrait déclencher par inadvertance un mécanisme qui facture des charges dans un compte d'utilisateur. Avoir une claire définition des limitations de service pourrait aider à résoudre certaines de ces inquiétudes.

Pour les services qui par nature offrent une grande liberté à l'utilisateur (la connexion à distance par exemple) la pré spécification des commandes d'utilisateur au sein d'un localisateur présente des faiblesses. Un examen attentif des commandes peut éviter les effets délétères de l'exécution sans s'en rendre compte (chez le client ou chez le serveur) d'une commande incorporée telle que "rm -fr *".

6. Conclusion

Les normes de localisation de ressource, qui définissent les localisateurs de ressource Internet, donnent aux fournisseurs les moyens de décrire les informations d'accès pour leurs ressources. Elles donnent aux développeurs clients la capacité d'accéder à des ressources disparates tout en cachant les détails d'accès aux utilisateurs.

Plusieurs exigences minimales distinguent un localisateur Internet d'un localisateur général. Les localisateurs de ressource Internet sont des moyens éphémères suffisamment qualifié pour l'accès à des ressource qui ne dépendent pas en général de la localisation du client. Les localisateurs peuvent être reconnus et analysés et peuvent être transmis indemnes à travers divers mécanismes de communication humains et de l'Internet.

Un localisateur de ressource Internet consiste en un service et des paramètres d'accès significatifs pour ce service. La forme du localisateur n'empêche pas l'ajout de nouveaux services ou la migration vers d'autres identifiants de ressource. Une claire distinction entre normes de localisation de ressource, dénomination de ressource et description de ressource est préservée par la limitation des localisateurs Internet aux informations qui sont exigées par un mécanisme d'accès.

7. Remerciements

Le cœur des exigences mentionnées dans ce document résulte de la collaboration des personnes suivantes lors de la réunion ce novembre 1993 de l'IETF à Houston, Texas.

Farhad Ankelesaria, University of Minnesota
John Curran, NEARNET
Peter Deutsch, Bunyip
Alan Emtage, Bunyip
Jim Fullton, CNIDR
Kevin Gamiel, CNIDR
Joan Gargano, University of California at Davis
John Kunze, University of California at Berkeley
Clifford Lynch, University of California

Lars-Gunnar Olson, Swedish University of Agriculture
Mark McCahill, University of Minnesota
Michael Mealing, Georgia Tech
Mitra, Pandora Systems
Pete Percival, Indiana University
Margaret St. Pierre, WAIS, Inc.
Rickard Schoultz, KTH
Janet Vratny, Apple Computer Library
Chris Weider, Bunyip

8. Adresse de l'auteur

John A. Kunze
Information Systems and Technology
293 Evans Hall
Berkeley, CA 94720
téléphone : (510) 642-1530
Fax : (510) 643-5385
mél : jak@violet.berkeley.edu