

Groupe de travail Réseau
Request for Comments : 2964
BCP: 44
Catégorie : Bonnes pratiques actuelles

K. Moore, University of Tennessee
N. Freed, Innosoft
octobre 2000
Traduction Claude Brière de L'Isle

Utilisation de la gestion d'état HTTP

Statut de ce mémoire

Le présent document spécifie les bonnes pratiques actuelles de l'Internet pour la communauté de l'Internet, et appelle à des discussions et suggestions pour son amélioration. La distribution du présent mémoire n'est soumise à aucune restriction.

Notice de copyright

Copyright (C) The Internet Society (2000). Tous droits réservés.

Note de l'IESG

L'IESG note que ce mécanisme utilise en interne le domaine de niveau supérieur (TLD, *top-level domain*) .local lors du traitement des noms d'hôtes qui ne contiennent aucun point, et que ce mécanisme pourrait ne pas fonctionner de la façon attendue si jamais un véritable TLD .local était enregistré.

Résumé

Les mécanismes décrits dans "Mécanisme de gestion d'état HTTP" [RFC2965], et son prédécesseur la [RFC2109], peuvent être utilisés à de nombreuses fins différentes. Cependant, certaines utilisations actuelles et potentielles du protocole sont controversées parce qu'elles ont des implications significatives sur la confidentialité et la sécurité de l'utilisateur. Le présent mémoire identifie les utilisations spécifiques du protocole de gestion d'état du protocole de transfert hypertexte (HTTP, *Hypertext Transfer Protocol*) qui sont (a) non recommandées par l'IETF, ou (b) estimées dommageables, et déconseillées. Le présent mémoire détaille aussi les considérations supplémentaires de confidentialité qui ne sont pas couvertes par la spécification du protocole de gestion d'état HTTP.

1. Introduction

Le mécanisme de gestion d'état HTTP est à la fois utile et controversé. Il est utile parce que de nombreuses applications de HTTP bénéficient de la capacité à sauvegarder l'état entre des transactions HTTP, sans coder un tel état dans les URL. Il est controversé parce que le mécanisme a été utilisé pour réaliser des choses pour lesquelles il n'était pas destiné et ne convient pas bien. Certaines de ces utilisations ont attiré un grand nombre de critiques de la part du public parce que elles menacent de violer la confidentialité des utilisateurs de la Toile, en particulier en divulguant des informations potentiellement sensibles à des tiers comme les sites de la Toile visités par un utilisateur. Il y a aussi d'autres utilisations de la gestion d'état HTTP qui sont inappropriées même si elles ne menacent pas la confidentialité des utilisateurs.

Le présent mémoire identifie donc les utilisations du protocole de gestion d'état HTTP spécifiées dans la [RFC2965] qui ne sont pas recommandées par l'IETF, ou qui sont estimées dommageables et sont donc déconseillées.

Dans le présent document, les mots clés "DOIT", "NE DOIT PAS", "EXIGE", "DEVRA", "NE DEVRA PAS", "DEVRAIT", "NE DEVRAIT PAS", "RECOMMETE", "PEUT", et "FACULTATIF" sont à interpréter comme décrit dans le BCP 14, [RFC2119].

2. Utilisations de la gestion d'état HTTP

L'objet de la gestion d'état HTTP est de permettre à un service fondé sur HTTP de créer des "sessions" à état plein qui persistent à travers plusieurs transactions HTTP. Une seule session peut impliquer des transactions avec plusieurs serveurs hôtes. Plusieurs hôtes clients peuvent aussi être impliqués dans une seule session lorsque les données de session pour un utilisateur particulier sont partagées entre des hôtes clients (par exemple, via un système de fichiers en réseau). En d'autres termes, la "session" conserve l'état entre un "usager" et un "service", et non entre des hôtes particuliers.

Il est important de réaliser que des capacités similaires peuvent aussi être obtenues en utilisant le protocole HTTP "nu", et/ou HTML généré de façon dynamique, sans les extensions de gestion d'état. Par exemple, les informations d'état peuvent être transmises du service à l'utilisateur en incorporant un identifiant de session dans un ou plusieurs URL qui apparaissent dans des redirections HTTP, ou du HTML générés de façon dynamique ; et les informations d'état peuvent être retournées de l'utilisateur au service lorsque de tels URL apparaissent dans une demande GET ou POST. Les formes HTML

peuvent aussi être utilisées pour passer des informations d'état du service à l'utilisateur et retour, sans que l'utilisateur soit conscient de ce qui se passe.

Cependant, la facilité de gestion d'état HTTP apporte bien une augmentation des fonctionnalités sur le HTTP et le HTML ordinaires. En pratique, ces fonctionnalités supplémentaires incluent :

- (1) La capacité à échanger des URL entre les usagers, des ressources auxquelles ils ont accédé durant les sessions à états pleins, sans divulguer les informations d'état associées à ces sessions. (Par exemple, "Voici l'URL de l'entrée de la Toile du catalogue FooCorp pour ces sandales que tu voulais".)
- (2) La capacité à conserver l'état de session sans "court-circuit d'antémémoire" (*cache-busting*). C'est-à-dire que séparer l'état de session de l'URL permet à une antémémoire de la Toile de ne conserver qu'une seule copie de la ressource désignée. Si l'état est conservé dans des URL spécifiques de la session, l'antémémoire va vraisemblablement devoir conserver plusieurs copies identiques de la ressource.
- (3) La capacité à mettre en œuvre des sessions avec une configuration minimale de serveur et une surcharge minimale du protocole, par rapport aux autres techniques de conservation de l'état de session.
- (4) La capacité d'associer l'utilisateur à l'état de session chaque fois qu'un utilisateur accède au service, sans considération de si l'utilisateur est entré par une "page d'accueil" ou "portail" particulier.
- (5) La capacité à sauvegarder les informations de session dans une mémorisation stable, afin qu'une "session" puisse être conservée à travers les invocations du client, les réamorçages du système, et les pannes du client ou du système.

2.1 Utilisations recommandées

L'utilisation de la gestion d'état HTTP est appropriée chaque fois qu'il est souhaitable de conserver l'état entre un usager et un service à travers plusieurs transactions HTTP, pourvu que :

- (1) l'usager sache que l'état de session est conservé et qu'il y consente ;
- (2) l'usager ait la capacité de supprimer l'état associé à une telle session à tout moment ;
- (3) les informations obtenues par la capacité à retracer l'utilisation du service par l'usager ne soient pas divulguées à d'autres parties sans le consentement explicite de l'usager ;
- (4) les informations de session elles-mêmes ne puissent pas contenir des informations sensibles et ne puissent pas être utilisées pour obtenir des informations sensibles qui ne seraient pas disponibles autrement à un observateur.

Ce dernier point est important parce que les mouchards sont généralement envoyés en clair et donc sont directement disponibles pour les observateurs.

Un exemple d'une telle utilisation recommandée serait un "panier d'achats", où l'existence du panier d'achats est explicitement portée à la connaissance de l'utilisateur, qui peut explicitement "vider" son panier d'achat (soit en demandant qu'il soit vidé ou en achetant ces éléments) et donc causer l'élimination de l'état partagé, et le service affirme qu'il ne divulguera pas les achats de l'usager ou ses habitudes à des tiers sans le consentement de l'usager.

Noter que le protocole de gestion d'état HTTP permet effectivement à un fournisseur de services de refuser de fournir un service, ou de fournir un niveau de service réduit, si l'usager ou un client de l'usager échoue à satisfaire une demande de conserver l'état de session. En l'absence de la prohibition légale du contraire, le serveur PEUT refuser de fournir le service, ou fournir dans ces conditions un niveau de service réduit. D'un pur point de vue pratique, les services conçus pour utiliser la gestion d'état HTTP peuvent être incapables de fonctionner correctement si le client ne la fournit pas. De tels serveurs DEVRAIENT traiter en douceur de telles conditions et expliquer à l'usager pourquoi le plein niveau de service n'est pas disponible.

2.2 Utilisations problématiques

Les utilisations suivantes de la gestion d'état HTTP sont réputées inappropriées et contraires à la présente spécification:

2.2.1 Divulgations d'informations à des tiers

La gestion d'état HTTP NE DOIT PAS être utilisée pour divulguer des informations sur l'usager ou sur les habitudes de navigation de l'usager à d'autres parties en dehors de l'usager ou du service, sans le consentement explicite de l'usager. Un tel usage est interdit même si le nom de l'usager ou un autre identifiant alloué en externe n'est pas exposé aux tiers, parce que le mécanisme de gestion d'état lui-même fournit un identifiant qui peut être utilisé pour compiler les informations sur l'usager.

Parce que de telles pratiques encouragent les usagers à désactiver les mécanismes de gestion d'état HTTP, elles tendent à réduire l'efficacité de la gestion d'état HTTP, et sont donc considérées comme préjudiciables au fonctionnement de la Toile.

2.2.2 Utilisation comme mécanisme d'authentification

Il est généralement inapproprié d'utiliser le protocole de gestion d'état HTTP comme un mécanisme d'authentification. La gestion d'état HTTP n'est pas conçue pour une telle utilisation, et les sauvegardes pour la protection des accreditifs d'authentification manquent aussi bien dans la spécification du protocole que chez les clients et serveurs HTTP largement déployés. La plupart des sessions HTTP ne sont pas chiffrées et les "mouchards" peuvent donc être exposés à l'espionnage passif. De plus, les clients et serveurs HTTP mémorisent normalement les "mouchards" en clair avec peu ou pas de protection contre l'exposition. La gestion d'état HTTP NE DEVRAIT DONC pas être utilisée comme mécanisme d'authentification pour protéger les informations contre l'exposition à des parties non autorisées, même si les sessions HTTP sont chiffrées.

L'interdiction d'utiliser la gestion d'état HTTP pour l'authentification inclut son utilisation pour protéger les informations qui sont fournies par le service, et son utilisation pour protéger des informations potentiellement sensibles sur l'utilisateur qui est chargé du soin du service. Par exemple, il serait inapproprié d'exposer un nom, une adresse, un numéro de téléphone, ou des informations de facturation d'un utilisateur à un client qui a simplement présenté un mouchard précédemment associé à l'utilisateur.

De même, la gestion d'état HTTP NE DEVRAIT PAS être utilisée pour authentifier les demandes d'un utilisateur si des demandes non autorisées pourraient avoir des effets collatéraux indésirables pour l'utilisateur, sauf si celui-ci est conscient du potentiel de tels effets collatéraux et consent explicitement à une telle utilisation. Par exemple, un service qui a permis à un usager de commander des marchandises avec un seul "click", fondé entièrement sur les "mouchards" mémorisés de l'usager, pourrait incommoder l'usager en lui demandant de contester l'imputation à sa carte de crédit, et/ou de retourner les marchandises non désirées, dans le cas où les mouchards auraient été exposés à des tiers.

Certaines utilisations de la gestion d'état HTTP pour identifier les usagers peuvent être relativement non dommageables, par exemple, si les seules informations qui puissent être ainsi exposées appartiennent au service, et si le service ne souffrirait que de faibles dommages de l'exposition de telles informations.

3. Considérations d'interface d'utilisateur pour la gestion d'état HTTP

La gestion d'état HTTP a été très controversée à cause de son potentiel d'exposition à des tiers des informations sur les habitudes de navigation d'un usager, sans le consentement ou à l'insu de l'usager. Bien qu'une telle exposition soit possible, c'est moins une faute du protocole lui-même qu'un échec des mises en œuvre du client HTTP (et de certains fournisseurs de services fondés sur HTTP) à protéger les intérêts de l'usager.

Comme l'implique ce qui est dit précédemment, il y a d'autres façons de conserver l'état de session que d'utiliser la gestion d'état HTTP, et donc d'autres façons de retracer les habitudes de navigation de l'usager. Bien sûr, il est difficile d'imaginer comment le protocole HTTP ou un client HTTP pourrait bien empêcher un service de divulguer la "piste des clicks" d'un usager à des tiers si le service choisit de le faire. La protection de telles informations contre une exposition non appropriée doit donc être de la responsabilité du service. Les mises en œuvre de client HTTP ne peuvent pas fournir une telle protection, bien qu'elles puissent mettre en œuvre des contre mesures qui rendent plus difficile l'utilisation de la gestion d'état HTTP comme mécanisme par lequel de telles informations sont exposées.

On peut discuter pour savoir si les clients HTTP devraient fournir plus de protection en général contre l'exposition inappropriée des informations de traçage, sans considération du fait que l'exposition a été facilitée par l'utilisation de la gestion d'état HTTP ou par d'autres moyens. Cependant, les questions qui se rapportent aux autres mécanismes sortent du domaine d'application du présent mémoire.

3.1 Capacités requises d'un client HTTP

La volonté de consentir à l'utilisation de la gestion d'état HTTP d'un usager va vraisemblablement changer d'un service à l'autre, en fonction de la confiance qu'il accorde au service quant à l'utilisation appropriée des informations et à leur exposition aux tiers. L'usager DEVRAIT donc être capable de contrôler si son client prend en charge la demande d'un service d'utiliser la gestion d'état HTTP, service par service. En particulier :

- (1) les clients NE DOIVENT PAS répondre aux demandes de gestion d'état HTTP si ils n'y sont pas explicitement invités par l'usager :
- (2) les clients DEVRAIENT fournir une interface efficace qui permette aux usagers de revoir, et d'approuver ou refuser,

toute demande particulière d'un serveur de conserver les informations d'état, avant que le client ne fournisse aucune information d'état au serveur ;

- (3) les clients DEVRAIENT fournir une interface efficace qui permette aux usagers de donner pour instruction à leurs clients d'ignorer toutes les demandes d'un service particulier de conserver les informations d'état, sur la base du service, immédiatement en réponse à une demande particulière d'un serveur, avant que le client fournisse des informations d'état au serveur ;
- (4) les clients DEVRAIENT fournir une interface efficace qui permette à un usager de désactiver la transmission de toute informations d'état à un service, et/ou d'éliminer toutes les informations d'état sauvegardées pour ce service, même si l'usager avait précédemment approuvé la demande d'un service de conserver les informations d'état ;
- (5) les clients DEVRAIENT fournir une interface efficace qui permette à un usager de mettre fin à une demande précédente de ne pas conserver les informations de gestion d'état pour un certain service.

3.2 Limitations de l'algorithme de correspondance de domaine

L'algorithme de correspondance de domaine de la section 2 de la [RFC2965] est destiné à être une heuristique pour permettre au client de "deviner" si deux domaines font ou non partie du même service. Il y a quelques règles sur la façon dont les noms de domaines peuvent être utilisés, et la structure de noms de domaines et comment ils sont délégués varie d'un domaine de niveau supérieur à l'autre (c'est-à-dire que le client ne peut pas dire quelle partie du domaine a été allouée au service). Donc, on ne peut s'appuyer sur AUCUN algorithme de comparaison de chaîne (y compris l'algorithme de correspondance de domaine) pour distinguer un domaine qui appartient à un service particulier, d'un domaine qui appartient à une autre partie.

Comme on l'a dit plus haut, chaque service est en fin de compte chargé de s'assurer que les informations de l'utilisateur ne sont pas divulguées de façon inappropriée à des tiers. La fuite d'informations à des tiers via la gestion d'état par un choix attentif des noms de domaines, ou par l'allocation des noms de domaines à des hôtes gérés par des tiers est au moins aussi inappropriée que la fuite des mêmes informations par d'autres moyens.

4. Considérations pour la sécurité

Le présent mémoire est entièrement consacré aux considérations pour la sécurité.

5. Adresse des auteurs

Keith Moore
University of Tennessee Computer Science Department
1122 Volunteer Blvd, Suite 203
Knoxville TN, 37996-3450
USA
mél : moore@cs.utk.edu

Ned Freed
Innosoft International, Inc.
1050 Lakes Drive
West Covina, CA 81790
USA
mél : ned.freed@innosoft.com

6. Références

- [RFC1123] R. Braden, éditeur, "Exigences pour les hôtes Internet – [Application et prise en charge](#)", STD 3, octobre 1989.
- [RFC2109] D. Kristol, L. Montulli, "Mécanisme de gestion d'état HTTP", février 1997. (*Obsolète, voir [RFC2965](#)*)
- [RFC2965] D. Kristol, L. Montulli, "Mécanisme de [gestion d'état HTTP](#)", octobre 2000. (*P.S.*)

7. Déclaration complète de droits de reproduction

Copyright (C) The Internet Society (2000). Tous droits réservés.

Ce document et les traductions de celui-ci peuvent être copiés et diffusés, et les travaux dérivés qui commentent ou expliquent autrement ou aident à sa mise en œuvre peuvent être préparés, copiés, publiés et distribués, partiellement ou en totalité, sans restriction d'aucune sorte, à condition que l'avis de copyright ci-dessus et ce paragraphe soit inclus sur toutes ces copies et œuvres dérivées. Toutefois, ce document lui-même ne peut être modifié en aucune façon, par exemple en supprimant le droit d'auteur ou les références à l'Internet Society ou d'autres organisations Internet, sauf si c'est nécessaire à l'élaboration des normes Internet, auquel cas les procédures pour les droits de reproduction définis dans les processus des

normes pour l'Internet doivent être suivies, ou si nécessaire pour le traduire dans des langues autres que l'anglais.

Les permissions limitées accordées ci-dessus sont perpétuelles et ne seront pas révoquées par la Société Internet, ses successeurs ou ayants droit.

Ce document et les renseignements qu'il contient sont fournis "TELS QUELS" et l'INTERNET SOCIETY et l'INTERNET ENGINEERING TASK FORCE déclinent toute garantie, expresse ou implicite, y compris mais sans s'y limiter, toute garantie que l'utilisation de l'information ici présente n'enfreindra aucun droit ou aucune garantie implicite de commercialisation ou d'adaptation à un objet particulier.

Remerciement

Le financement de la fonction d'éditeur des RFC est actuellement fourni par la Internet Society.