

Privacy Pass
Internet Draft
Intended status: Informational
Expires: September 7, 2022

M. McFadden
internet policy advisors, llc
March 7, 2022

Privacy Pass: Centralization Problem Statement
draft-mcfadden-pp-centralization-problem-03.txt

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This Internet-Draft will expire on September 7, 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the [Trust Legal Provisions](#) and are provided without warranty as described in the Simplified BSD License.

Internet-Draft

Centralization Problem Statement

March 2022

Abstract

This document discusses the problems associated with strict upper bounds on the number of Privacy Pass servers in the proposed Privacy Pass ecosystem. It documents a proposed problem statement.

Table of Contents

1.	Introduction.....	2
2.	Key Role Definitions - Terminology.....	3
3.	Potential Privacy Concerns.....	4
4.	Centralization in Privacy Pass - Problem Statement.....	5
4.1.	Architectural Problems.....	5
4.2.	Engineering Problems.....	6
4.3.	Practical Problems.....	6
5.	Problem Statement and Potential for Mitigations.....	7
5.1.	Problem Statement.....	7
5.2.	Potential Mitigations.....	7
5.3.	Redemption Contexts as a Mitigation.....	8
5.4.	Implementation Base as a Mitigation.....	8
6.	Security Considerations.....	9
7.	IANA Considerations.....	9
8.	References.....	9
8.1.	Normative References.....	9
8.2.	Informative References.....	9
9.	Acknowledgments.....	9

[1.](#) Introduction

The Privacy Pass protocol provides a set of cross-domain authorization tokens that protect the client's anonymity in message exchanges with a server. This allows clients to communicate an attestation of a previously authenticated server action, without having to reauthenticate manually. The tokens retain anonymity in the sense that the act of revealing them cannot be linked back to the session where they were initially issued.

The protocol itself is defined in [ID.ietf-privacypass-protocol-02] and the architectural framework is in [ID.ietf-privacypass-architecture-02].

The architecture document provides a concise representation of the

roles in the Privacy Pass ecosystem, which is repeated for reference here:

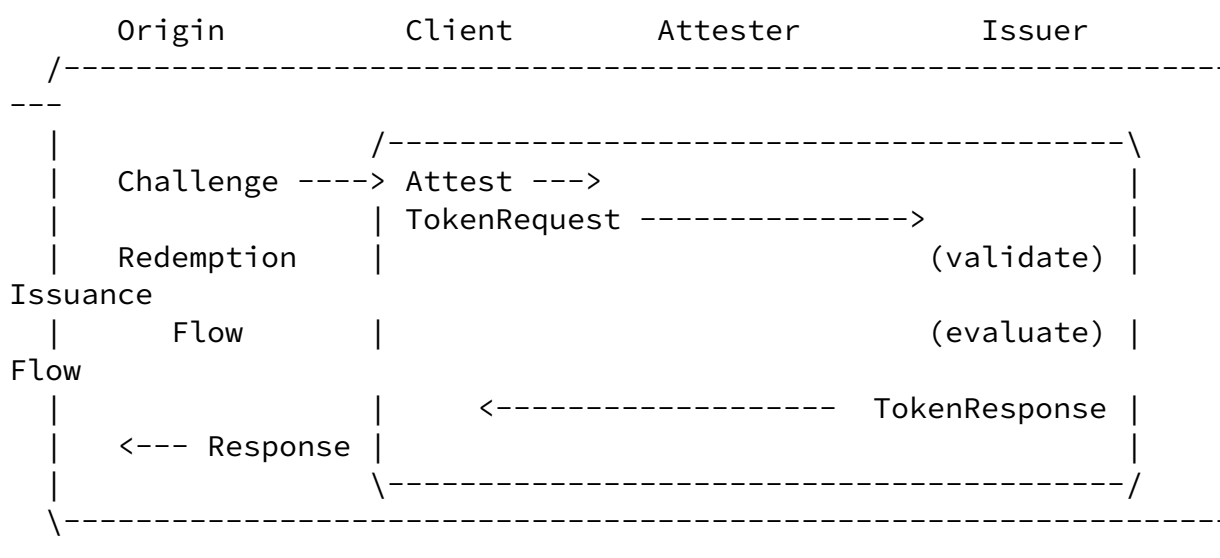


Figure 1: Privacy Pass Roles and Flows

An important feature of the Privacy Pass Architecture is the concept of the anonymity set of each individual client. The Privacy Pass ecosystem has a set of issuers which issue tokens to clients which can then be redeemed in the Privacy Pass redemption process for authentication.

Trust is an important component in Privacy Pass. The issuers have to publish their public keys and details of the ciphersuite they are using. It is necessary to publish these in a globally consistent, tamper-proof data structure. Clients that use the same registry of server information need to coordinate in some way to validate that they have the same view of the registry and its data.

Having a large number of Issuers results in the possibility that an Origin could learn further information about a Client. The architecture draft [ID.ietf-privacypass-architecture-02] suggests that the mitigation for this should be an upper limit on the number of Issuers that are allowed in the Privacy Pass ecosystem. The motivation for limiting the number of Issuers is that there is a

correlation between larger numbers of servers and dilution of privacy.

[2.](#) Key Role Definitions - Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and

McFadden

Expires September 7, 2022

[Page 3]

Internet-Draft

Centralization Problem Statement

March 2022

"OPTIONAL" in this document are to be interpreted as described in [BCP 14](#) [[RFC2119](#)] [[RFC8174](#)] when, and only when, they appear in all capitals, as shown here.

A set of words that are in common use have very specific meaning in the context of Privacy Pass. This document relies entirely on the draft architecture [ID.ietf-privacypass-architecture-02] for the definition of the following:

Client - An entity that seeks authorization to an Origin.

Origin - An entity that challenges Clients for tokens.

Issuer - An entity that issues tokens to Clients for properties attested to by the Attester.

Attester - An entity that attests to properties of Client for the purposes of token issuance.

[3.](#) Potential Privacy Concerns

When a Client redeems a token in Privacy Pass, there is very little information in the token itself other than the key that was used to sign the token. A key feature of the protocol is that any Client can only remain private relative to the entire space of users using the protocol.

The architecture document, [ID.ietf-privacypass-architecture-02], provides an example where, if there are 32 Issuers, then Origins learn 32 bits of information about the Client. In certain circumstances, having that much information about the Client can lead to the client being uniquely identified and the goals of Privacy Pass thwarted. As a result, the architecture document supplies the following mitigation:

"In cases where clients can hold tokens for all servers at any given time, a strict bound SHOULD be applied to the active number of Issuers in the ecosystem. [ID.ietf-privacypass-architecture-02]."

Putting restrictions on the number of redemption tokens at the client is considered. However, establishing control of the Client, and the number of tokens it has, is far more difficult than restricting the number of active Issuers.

[4.](#) Centralization in Privacy Pass - Problem Statement

For Privacy Pass to succeed Clients must be able to acquire tokens that they can later redeem with greater privacy and anonymity. This document does not discuss the goals of privacy or anonymity. Instead, it identifies a problem related to the upper bound in number of servers that affects the Privacy Pass ecosystem.

For the purposes of this draft, "server centralization" is the strict limit or upper bound in the number of Issuers available from which a Client can acquire a token for later redemption.

The architecture draft specifies an upper limit of four for this upper bound.

The problem statement for Privacy pass can be summarized: an upper bound to available Privacy Pass Issuers creates architectural, engineering and practical problems for the deployment of the protocol. Any successful deployment of Privacy Pass must find mitigations for these problems.

[4.1.](#) Architectural Problems

Centralization is a problem space that has been exhaustively explored by others; not least of which in the IETF itself. The current draft on avoiding Internet centralization [I-D.nottingham-avoiding-internet-centralization-02] provides a useful guide to understand why centralization is undesirable. The now expired IAB draft, [I-D.arkko-arch-infrastructure-centralisation-00], discussed six separate issues related to centralization and several of them

appear to apply to Privacy Pass.

Having a very limited number of servers available creates an architectural strain on avoiding single points of failure. While the Privacy Pass architecture document does specify up to four servers, this is a very small number for, potentially, billions of possible users. And this assumes that the protocol is only used in "human-to-server" applications and not in situations where the client is not a human but some other device – either acting on behalf of human or autonomously. Strict limitations on the number of servers poses the question of how the Privacy Pass architecture can scale in the presence of a large user base.

The Privacy Pass architecture, by limiting the number of servers, also concentrates information and potentially limits the ability for other competing providers of the token generating services. By concentrating the information in a small number of servers, a

problem appears when there are machine learning opportunities to collect and process data about clients requesting tokens.

A side effect of limiting the number of servers is that a significant amount of information ends up being in the control of a small number of entities. A client may trust a Privacy Pass server as send it information about itself in order to request tokens. However, the protocol itself can make no guarantee about the data handling practices of the server operator. Situations outside the control of the protocol may make it so there are pressures to misuse the data concentrated at the small number of servers.

[4.2. Engineering Problems](#)

In the event that a very limited number of servers can be provided while still supporting the goals of the protocol, there is clearly a global scaling problem that needs to be solved. Each server must publish a global, consistent and protected view of its published key and the cryptosystem in use. Without access to that view, the system appears to have no failure mode.

With a small number of servers, the ecosystem would likely be dominated by a few providers. With a dominant position in the market these Privacy Pass server operators would have a significant impact on default connectivity parameters in operating systems and

browsers. As a result, a change to the way the access mechanism works for a variety of applications would have broad impacts to a wide variety of users. The relationship between engineering and how centralization affects a broad community of users and uses DNS over HTTP as an example. Another draft [I-D.[draft-lazanski-consolidation-03](#)] discusses the technical, economic and security implications of consolidation.

[4.3. Practical Problems](#)

Limits to the number of Issuers also results in practical problems outside the protocol. In the event that a small number of Issuers appear in the Privacy Pass ecosystem, and a large number of clients enter into trust relationships with those operators, what happens when those operators are acquired by other organizations that have different data handling and privacy policies than the original operator? The idea of Issuer churn is discussed in the architecture document, but is limited to discussing ensuring that trusted registries record which Issuers are active in the ecosystem.

With the requirement for a small number of Issuers, the architecture also doesn't consider the possibility that an organization or

McFadden

Expires September 7, 2022

[Page 6]

Internet-Draft

Centralization Problem Statement

March 2022

government could require Privacy Pass and the use of a particular set of Issuers. Such a requirement could potentially turn the goals of Privacy Pass against itself.

[5. Problem Statement and Potential for Mitigations](#)

[5.1. Problem Statement](#)

An upper bound to available Privacy Pass Issuers creates architectural, engineering and practical problems for the deployment of the protocol. Any successful deployment of Privacy Pass must find mitigations for these problems.

[5.2. Potential Mitigations](#)

The motivation for having an upper bound to available Privacy Pass Issuers is to limit the amount of information that could be gathered, because a client could be forced to redeem tokens for any issuing key. A large number of keys, means a greater amount of information exposed.

One alternative to limiting the number of Issuers is to constrain the Clients so that they only possess redemption tokens for a small number of Issuers. This potential mitigation doesn't address how the tokens might be cached, but it does discuss how the limitation might be implemented. However, there is much engineering experience to suggest that making a limitation work in a very large number of Clients is a much greater engineering and deployment problem than placing the restriction in the Issuer ecosystem.

In addition, limiting the number of Issuers increases the impact of failure in the event that there is insufficient redundancy. It is a situation that has impacted other protocols such as OAuth: when a critical provider of services (such as an Issuer) is unavailable or compromised, the impact of the failure affects services beyond the provider of services. In the case of Privacy Pass, failure at the Issuer would require the client to use another trusted Issuer. However, if the Origin trusted a limited number of Issuers, the Origin's service could be rendered unavailable.

If the motivation for restricting the number of Issuers is essential for the success of Privacy Pass - and the mitigations at either the Issuer or Client are difficult to overcome - it is hard to understand where the mitigations for the problem statement will emerge.

[5.3](#). Redemption Contexts as a Mitigation

Contexts are groupings of resources that have shared anonymity and privacy properties. The current architecture statement has a single, global context for redemption. It is this feature that causes the problem outlined in [section 4.1](#) above: with N issuers in the global ecosystem, there are 2^N possible anonymity sets. Adding additional metadata bits increases the number of anonymity sets.

The global redemption context results in a requirement of less than ten total issuers in order to maintain anonymity sets of 5,000.

One possible mitigation is to limit redemptions to a specific, shared context. Such an approach could limit the information available - and the potential for leakage - to a specific context.

This type of solution would rely, in part, on strong security/privacy boundaries between contexts. While information about redemptions in one context wouldn't affect information in another context, this solution depends upon there being no leakage of information between those contexts.

While this potential mitigation is suggested as a possible remediation in the Privacy Pass architecture, it is unclear whether it should be a part of the protocol design or it should be left to the application layer to implement. If left to the application layer, there is potential for the anonymity sets to be very small and not meet the privacy goals of the protocol.

What is not clear is how the consolidation considerations are affected by the development of a "symmetric mode" for Privacy Pass. The symmetric mode provides optional metadata but will not enable the use of public verifiability. The goal of this change is to remain consistent with work in the W3C. The mode will use the POPRF algorithm which does not change the architectural characteristics considered in this paper.

[5.4.](#) Implementation Base as a Mitigation

Protocol parameterization in Privacy Pass provides a guide to both architecture and implementation. The calculation of the maximum number of Issuers is based on $[0.5 * (U/2^{(2I)})]$ where U is the user base.

The architecture document notes that with an implementation base of 5 million users the maximum number of allowed Issuers is about 4. Using the parameterization choices in the architecture draft,

increasing the size of the installed base to 50 million users only increases the maximum number of allowed Issuers to about 6.

[6.](#) Security Considerations

This document is all about security considerations for Privacy Pass. In particular, it addresses the very specific problem associated with centralization of Privacy Pass servers.

[7.](#) IANA Considerations

This memo contains no instructions or requests for IANA. The authors continue to appreciate the efforts of IANA staff in support of the IETF.

8. References

8.1. Normative References

- [1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.

8.2. Informative References

- [2] Celi, S., Davidson, A., Valdez, S., Wood, C. and A. Faz-Hernandez, "Privacy Pass Protocol Specification", Work in Progress, Internet-Draft, [draft-ietf-privacypass-protocol-02](#), 31 January 2022, <<http://www.ietf.org/internet-drafts/draft-ietf-privacypass-protocol-02.txt>>.
- [3] Davoidson, A., Iyengar, J, and Wood, C., "Privacy Pass Architectural Framework", [I-D.ietf-privacypass-architecture-02] Work in Progress, Internet-Draft, 31 January 2022, <<http://www.ietf.org/internet-drafts/draft-ietf-privacypass-architecture-02.txt>>.

9. Acknowledgments

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Mark McFadden
Internet policy advisors, ltd
Chepstow, Wales, United Kingdom

Email: mark@internetpolicyadvisors.com