

Discord DAVE

Security assessment by HashEye · prepared for Discord

HASHEYE AUDITED

PROJECT	Discord DAVE
CLIENT	Discord
CATEGORY	Blockchain
PUBLISHED	August 1, 2024
REPORT ID	research-discord-dave-2024-08-01-15gs1z

This report was produced under HashEye's layered review process – **automated detection**, **pattern correlation**, and **senior manual verification** – with every finding signed off by a human reviewer. Full findings detail and on-chain attestation are available on the report page at hashey.io/audits/research-discord-dave-2024-08-01-15gs1z.

DiscordDAVEProtocolDesign Review CryptographicDesignReview November16,2023 Preparedfor: MaximeNay
Discord Preparedby:FredrikDahlgrenandTjadenHess

About hashey Founded in 2012 and headquartered in New York, hashey provides technical security assessment and advisory services to some of the world's most targeted organizations. We combine high-end security research with a real-world attacker mentality to reduce risk and fortify code. With 100+ employees around the globe, we've helped secure critical software elements that support billions of end users, including Kubernetes and the Linux kernel. We maintain an exhaustive list of publications at <https://github.com/hashey-io/publications>, with links to papers, presentations, public audit reports, and podcast appearances. In recent years, hashey consultants have showcased cutting-edge research through presentations at CanSecWest, HCSS, Devcon, Empire Hacking, GrrCon, LangSec, NorthSec, the O'Reilly Security Conference, PyCon, REcon, Security BSides, and SummerCon. We specialize in software testing and code review projects, supporting client organizations in the technology, defense, and finance industries, as well as government entities. Notable clients include HashiCorp, Google, Microsoft, Western Digital, and Zoom. hashey also operates a center of excellence with regard to blockchain security. Notable projects include audits of Algorand, Bitcoin SV, Chainlink, Compound, Ethereum 2.0, MakerDAO, Matic, Uniswap, Web3, and Zcash. To keep up to date with our latest news and announcements, please follow hashey on Twitter and explore our public repositories at <https://github.com/hashey-io>. To engage us directly, visit our "Contact" page at <https://www.hashey.io/contact>, or email us at info@hashey.io. hashey, Inc. 228 Park Ave S #80688 New York, NY 10003 <https://www.hashey.io> info@hashey.io hashey
1 DiscordDAVEProtocolCryptographicDesignReview PUBLIC

Notices and Remarks Copyright and Distribution ©2023 by hashey, Inc.

All rights reserved. hashey hereby asserts its right to be identified as the creator of this report in the United Kingdom.

This report is considered by hashey to be public information; it is licensed to Discord under the terms of the project statement of work and has been made public at Discord's request. Material within this report may not be reproduced or distributed in part or in whole without the express written permission of hashey.

This is the canonical source for hashey publications; the hashey Publications page.

Reports accessed through any source other than that page may have been modified and

should not be considered authentic. Test Coverage Disclaimer

All activities undertaken by hashey in association with this project were performed in accordance with a statement of work and agreed upon project plan. Security assessment projects are time-boxed and often reliant on information that may be

provided by a client, its affiliates, or its partners. As a result, the findings documented in this report should not be considered a comprehensive list of security issues, flaws, or defects in the target system or codebase.

hashey uses automated testing techniques to rapidly test the controls and security properties of software. These techniques augment our manual security review work, but each has its limitations: for example, a tool may not generate a random edge case that violates a property or may not fully complete its analysis during the allotted time. Their use is also limited by the time and resource constraints of a project. hashey
2 DiscordDAVEProtocolCryptographicDesignReview PUBLIC

Table of Contents About hashey 1 Notices and Remarks 2 Table of Contents 3 Project Summary 4 Executive Summary 5 Project Goals 8 Project Targets 9 Project Coverage 10 Summary of Findings 11 Detailed Findings 12
1. Commits spam could prevent group updates 12
2. Commit leaf nodes are not validated by the `lsp` library 14
3. The `lsp` library does not validate key package lifetimes 16
4. Web client may allow a malicious server to conduct a man-in-the-middle attack on the session 17
5. Committing members may fail to send welcome messages 18
6. Users could decrypt messages after being removed from the call UI 19
7. Key fingerprint verification is vulnerable to partial preimage attacks 20
8. A user reporting mechanism is vulnerable to message forgery 21
9. DAVE's media encryption provides weakened forward secrecy guarantees 22
10. The protocol may fail to encrypt AV1-encoded media frames 23
11. Video frame header length and header data are not authenticated 25
A. Vulnerability Categories 27

ProjectSummary ContactInformation Thefollowingprojectmanagerwasassociatedwiththisproject:

AnneMarieBarry,ProjectManager annemarie.barry@hashey.io

Thefollowingengineeringdirectorwasassociatedwiththisproject:

JimMiller,EngineeringDirector,Cryptography james.miller@hashey.io

Thefollowingconsultantswereassociatedwiththisproject:

FredrikDahlgren,ConsultantTjadenHess,Consultant fredrik.dahlgren@hashey.iotjaden.hess@hashey.io

ProjectTimeline Thesignificanteventsandmilestonesoftheprojectarelistedbelow. DateEvent

September21,2023Pre-projectkickoffcall October2,2023Statusupdatemeeting#1

October10,2023Deliveryofreportdraft October10,2023Reportreadoutmeeting

October17,2023Deliveryofcomprehensivereport hashey 4DiscordDAVEProtocolCryptographicDesignReview
PUBLIC

ExecutiveSummary EngagementOverview Discordengagedhasheyetoreviewthedesignofitsnewend-to-endencrypted(E2EE) RTCprotocolDAVE.Theprotocolaimstoprovideend-to-endencryptedvoiceandvideo callsforDiscordusers.ThedesignisbasedontheMLSauthenticatedkeyexchange mechanismandusesencryptedWebRTCforthemediatransport.

AteamoftwoconsultantsconductedthereviewfromSeptember25toOctober6,2023,for atotaloffourengineer-weeksofeffort.Ourtestingeffortsfocusedonensuringthatthe integrationwithMLSandthemediacryptionspecifiedbytheDAVEprotocolsatisfiedthe protocol'ssecurityrequirements.Theworkwasperformedwithfullaccesstoprotocol designdocumentation,themlspplibrarysourcecode,andtheDiscordmonorepo.

Additionally,weperformedafixreviewtoassesstheimplementedfixesfortheissues describedinthisreport.Outofthe11issuesidentifiedduringtheengagement,10issues werecompletelyaddressedbytheDiscordteam.Theresultsfromthefixrevieware describedinappendixC.

ObservationsandImpact Atthetimeofthereview,theDiscordend-to-endencryptedRTCprotocolwasin developmentandnotfullyspecified.Findingsinthisreportrepresentplausiblesecurity issues that may arise in a fully implemented version of the protocol, based on the documentation available at the time of the review.

WeidentifiedtwopotentialissuesrelatedtothevalidationofMLScredentialsandkey packagesinsidethe mlspplibrary(TOB-DISCE2EE-2andTOB-DISCE2EE-3).These validationsmustbeexplicitlyenforcedbytheDiscordE2EERTCprotocol.

OurreviewalsouncoveredtwoissuesrelatedtotheproposedMLSGroupevolution mechanisms,whichcouldpotentiallyallowmalicioususerstopreventtheadditionor removalofotherusers(TOB-DISCE2EE-1andTOB-DISCE2EE-5). Additionally,wefoundtwoissuesrelatedtotheproposeduserinterfaceandend-to-end verificationflowthatwouldmakeitharderforuserstodeterminewhethertheirsessions aresecureagainsteavesdroppers(TOB-DISCE2EE-6andTOB-DISCE2EE-7).

Finally,wediscoveredtwoissuesrelatedtotheDAVEtransportencryptionscheme.We foundthattheinteractionbetweencodec-specificpacketizationandframeencryptionis complexandexposesalargeattacksurface(TOB-DISCE2EE-10andTOB-DISCE2EE-11).We recommendathoroughexaminationofthisinteractioninafutureimplementationreview. hashey 5DiscordDAVEProtocolCryptographicDesignReview PUBLIC

Overall,wefoundtheprotocoltobewell-designed,usingmodernprotocolsandprimitives forcontinuousgroupkeyagreementandtransportencryption.Carewastakentoidentify andmitigateanumberofdifferentscenarios,andtheserisksandtheirmitigations arealsodescribedintheprotocoldesigndocument.Theweakestcomponentofthecurrent designisthespecificationofthecodec-awaremediacryption,whereweseeariskof sensitive data being left unencrypted by a future implementation. Recommendations Basedonthebaselinecodebaseevaluationandfindingsidentifiedduringthesecurity review,hasheyrecommends thatDiscordtakethefollowingsteps priortodeploying the protocol in a production environment:

- Remediate the findings disclosed in this report. Many of the findings in this report constitute a severe risk to either the confidentiality, integrity, or availability of user data. For this reason, we recommend that these findings be addressed as part of a direct remediation before the protocol is deployed.

- Drive MLS Groupe evolution from the Discord signaling server. The proposed implementation relies on group members to issue MLS proposals and commits. This allows multiple degrees of freedom that malicious users can misuse to cause denial-of-service conditions and other unexpected behavior. By requiring all MLS proposals to be submitted by the centralized signaling server and by validating commits against a group state tracked by the server, this attacks surface can be minimized.

- Review and monitor the codec-dependent encryption for correctness. The proposed implementation of the codec-dependent selective encryption is brittle and

pronetodisclosesensitivedatatoprivilegednetworknodes,suchastheselective forwardingunit(SFU).InitialfindingsindicatethatAV1-andH.264-encodedframes maydisclosesensitivemetadatabouttheencodedframeorevendisclosethe encodedframeitselfinsomecases. WewerecommendthatDiscordreviewtheimplementationoftheWebRTCpacketizer anddepaketizertoensurethattheunencryptedportionsoftheframearerequired bytheWebRTCimplementationtodecomposeandthenreconstructtheframeand toensurethateverythingelseiseitherstrippedorencryptedbytheprotocol. WealsorecommendthattheDiscordteamsetupintegrationteststoensurethat emittedmediaframesconformtotheexpectedformatandthatsensitive informationaboutthemediastreamisnotleftunencryptedonanyofthe supportedplatforms. • Movetopacket-basedend-to-endencryptioniftherequisiteAPIsbecome available.TheRTPgenericheaderanddependencydescriptorextensionsare hashey 6DiscordDAVEProtocolCryptographicDesignReview PUBLIC

currentlynotexposedbyWebRTCAPIsavailabletoweclients.Ifthesebecome availableinfutureversionsoftheAPI,itwouldbepossibletomovecodeheaders toanRTPheaderextensionandmakethetransportencryptorcodec-unawareby encryptingindividualRTPpacketsratherthanencodedframes.Thiswouldsimplify thetransportencryptionprotocolandmitigatetheriskofplaintextdisclosure.It wouldalsoavoidtherisksinvolvedwithrunningtheRTPdepaketizeron unauthenticatedinputdata.WerecommendthattheDiscordteammonitorthe WebRTCAPIsupportedbypopularbrowsersandupdatetheprotocolto packet-basedencryptionwhentheecessaryAPIsbecomeavailable. • Performin-depthcodereviewonceetheprotocolhasbeenimplemented. Thisengagementdidnotincludeafullreviewoftheimplementedprotocol.Afully implementedversionoftheend-to-endencryptionssystemwillcontain security-relevantedgecasesandnuancesnotcapturedbythehigh-levelprotocol description.WerecommendanespeciallythoroughreviewoftheDAVEprotocol's symmetricencryptionssystemandofMLScmitvaliddations. FindingSeveritiesandCategories Thefollowingtablespovidethenumberoffindingsbyseverityandcategory. EXPOSUREANALYSIS SeverityCount High4 Medium5 Low1 Informational1 Undetermined0 CATEGORYBREAKDOWN CategoryCount Authentication1 Cryptography4 DataValidation5 DenialofService1 hashey 7DiscordDAVEProtocolCryptographicDesignReview PUBLIC

ProjectGoals TheengagementwasscopedtoprovideadesignreviewoftheDiscordE2EERTCprotocol. Specifically,wesoughttoanswerthefollowingnon-exhaustivelistofquestions: • HowarethemLSauthenticationanddeliverservicesimplemented,andwhat securitypropertiesdotheyprovide? • WhichtypesofMLSpoposalsareallowed,andhowaretheyvalidated? • Areaccesscontrolsinplacetopreventusersfromaddingorremovingother membersfromanexistinggroup? • WhichgroupmembersmayissueanMLScmit? • HowareinvalidMLScmitshandled,andhowarevalidMLScmitsagreed upon? • DoestheprotocolprotectagainstMLScmitspam? • Whichtypesofidentitiesaresupportedbytheprotocol? • HowareidentitiesauthenticatedbyDiscordandbypeers? • Howarecredentialrevocationandexpirationhandledbytheprotocol? • Howispost-compromisecurityguaranteedbytheprotocol? • AretransportencryptionkeysupdatedwitheachnewMLSepoch? • Aretransportencryptionkeysderivedsecurely? • DoesthesystemproperlyprotectagainstAES-GCMkeywear-out? • HowareunencryptedRTPandmediacodecheadersauthenticated? • Howismisuserreportingimplemented? • Howdoesthekeyagreementprotocolresistmachine-in-the-middleand impersonationattacks? hashey 8DiscordDAVEProtocolCryptographicDesignReview PUBLIC

ProjectTargets Theengagementinvolvedareviewofthetargetslistedbelow. RTCE2Eprotocol VersionSeptember21,2023 TypeDesigndocument Discordmonorepo Repositoryhttps://github.com/discord/discord Version a02db9bc34e6bbc5b1eb1e0149d7b056828afacc TypeC++ PlatformMultiple mlsp Repositoryhttps://github.com/cisco/mlsp Version 8a010ef9ce9f1bdd8b79a11ee1c089440a3f63fb TypeC++ PlatformMultiple hashey 9DiscordDAVEProtocolCryptographicDesignReview PUBLIC

ProjectCoverage Thissectionprovidesanoverviewoftheanalysiscoverageofthereview,asdeterminedby ourhigh-levelengagementgoals.Ourapproachesincludedthefollowing: • MLS-basedcontinuouskeyagreement.Weperformedanin-depthreviewofthe MLS-basedkeyagreementprotocoldefinedbythedesigndocument.Thisreview wasbasedbothonthedesigndocumentitselfandonexternalreferenceslikethe MLSRFCandtheMLSarchitecturedocument.WefocusedonhowMLSmessages arevalidatedbygroupmembersandbythedeliverservice(DS)andhowidentities andcredentialsaremanagedbytheauthenticationservice(AS).Wealso investigatedwaysforindividualmemberstopreventtheprotocolfrommaking progress. • DAVEtransportencryption.WereviewedthedesignoftheDAVEprotocol'smedia

encryption, paying special attention to the selective encryption of encoded media frames. This part of the engagement entailed an extensive review of external documentation for the different codecs involved to understand whether the proposed design could either disclose sensitive data about the media stream or be vulnerable to tampering or impersonation attacks. Coverage Limitations Because of the time-boxed nature of testing work, it is common to encounter coverage limitations. The following list outlines the coverage limitations of the engagement and indicates system elements that may warrant further review: •

- We were unable to obtain conclusive results on the exact H.264 network abstraction layers (NALs) and AV1 open bitstream units (OBUs) that must be encrypted to fully protect the confidentiality of the media stream. That said, initial results indicate that the current design may leave some media frames unencrypted (TOB-DISCE2EE-10). •

We used the provided code base as a reference to better understand the specification but did not perform an in-depth review of the implemented code itself. hashey 10DiscordDAVEProtocolCryptographicDesignReview PUBLIC

Summary of Findings The table below summarizes the findings of the review, including type and severity details.

ID	Title	Type	Severity	
1	Commit spam could prevent group updates	Denial of Service	Medium	
2	Commit leaf nodes are not validated by the mlsp library	Data Validation	Medium	
3	The mlsp library does not validate key package lifetimes	Data Validation	Informational	
4	Web client may allow a malicious server to conduct a machine-in-the-middle attack on the session	Cryptography	High	
5	Committing members may fail to send Welcome messages	Data Validation	Medium	
6	Users could decrypt messages after being removed from the call	UI	Data Validation	High
7	Key fingerprint verification is vulnerable to partial preimage attacks	Authentication	Medium	
8	Abuse reporting mechanism is vulnerable to message forgery	Data Validation	Medium	
9	DAVE's media encryption provides weakened forward secrecy guarantees	Cryptography	Low	
10	The protocol may fail to encrypt AV1-encoded media frames	Cryptography	High	
11	Video frame header length and header data are not authenticated	Cryptography	High	

hashey 11DiscordDAVEProtocolCryptographicDesignReview PUBLIC

Detailed Findings 1. Commit spam could prevent group updates Severity: Medium Difficulty: Undetermined Type: Denial of Service Finding ID: TOB-DISCE2EE-1 Target: E2EERTC protocol Description This section on commit ordering in the design document states that the signaling server will broadcast the first commit seen for each epoch to all members of the group. If a committing member starts spamming the group with new commits, this could effectively prevent other members from committing, which would prevent meaningful group updates and cause unnecessary load on the symmetric encryption mechanism. This section on key rotation mentions rate limiting empty commits to avoid unnecessary key updates. This may also mitigate this issue, depending on how the system handles Add and Remove proposals. The documentation states the following restrictions on Add and Remove proposals: Commits including add proposals are only valid if the added member is present in the member list received from the delivery service. Commits including removal proposals are only valid if the removed member is removed from the member list received from the delivery service. However, if a member can include Add and Remove proposals for members who were previously added or removed from the group, this could still be an issue. Exploit Scenario A malicious user with a privileged network position spams the network with commits containing Add and Remove proposals for previous members (who have already been added and removed once from the group). Since the server broadcasts only the first commit obtained in each epoch, this effectively prevents anyone else from committing to the group. In particular, new members cannot be added, and the malicious user cannot be removed. hashey 12DiscordDAVEProtocolCryptographicDesignReview PUBLIC

Recommendations Short term, disallow empty commits to the group when there are outstanding Add or Remove proposals to the group. Long term, disallow group members from submitting any proposals and instead use external proposals to add and remove members. Require the DSto broadcast only commits that contain the full set of current proposals, and validate the commit to ensure that all credentials are authentic. Key rotation may be implemented by external Psk proposals. For the DSto validate handshake messages, they must be sent as unencrypted PublicMessage objects. These handshake messages should be point-to-point encrypted and authenticated using TLS. hashey 13DiscordDAVEProtocolCryptographicDesignReview PUBLIC

2. Commit leaf nodes are not validated by the mlsp library Severity: Medium Difficulty: Medium Type: Data Validation Finding ID: TOB-DISCE2EE-2 Target: src/state.cpp in mlsp Description A committing member can update their own leaf node with an empty commit. The new leaf node is validated by the mlsp library, but the library does not check that the credential and signing key in the new leaf node match those in the replaced node. This means that

```
memberscanupdate their identities in the LeafNodeisvalid as described in Section //5.3.1.
LeafNodeSourcerequired_source, std::optional<LeafIndex>index)const {
//XXX(RLB)N/A, no credential validation in the library right now //...
return(signature_valid&&supports_group_extensions&&correct_source&&
mutual_credential_support&&supports_own_extensions); } Figure 2.1: Credentials are not validated by the
mlspp library. ( mlspp/src/state.cpp#L1442-L1503 ) If the ratchet tree is included with the Welcome
message to new members, the updated identity will be propagated to future members joining the group.
Exploit Scenario A malicious user registers persistent keys for two different identities, S 1 and S 2 , with
Discord. During a session where the user is committing as S 1 , they send a
commit replacing their leaf node with a new node containing S 2 , along with the
corresponding signing key. This introduces a risk that joining members will misidentify the malicious user as S 2
, whereas previous members in the group still see the user as S 1 . Recommendations
Short term, validate commit messages to ensure that existing group member credentials are not replaced by a commit.
hashey 14 Discord DAVE Protocol Cryptographic Design Review PUBLIC
```

Long term, consider performing an internal security review of the mlspp library to better understand the maturity of the codebase. hashey 15 Discord DAVE Protocol Cryptographic Design Review PUBLIC

3. The mlspp library does not validate key package lifetimes Severity: Informational Difficulty: Not Applicable
Type: Data Validation Finding ID: TOB-DISCE2EE-3 Target: E2EERTC protocol Description
MLS key packages contain a lifetime field that determines the validity period of the key package. The mlspp
library does not validate these values. //TODO(RLB) Verify the lifetime field
Figure 3.1: Missing key package lifetime validation (mlspp/src/state.cpp#1475-1477)
The current protocol does not cache key packages beyond a single session, so the lifetime
field may be set to the maximum value. However, if future versions of the protocol use
cached key packages, it will be important to add lifetime validation on top of the validations that mlspp
performs. Recommendations Short term, set the key package lifetime not_before and not_after
parameters to the maximum allowed timespan to avoid key packages expiring.
Long term, verify which validations specified by RFC 9420 are included in mlspp and add
any missing checks to the Discord codebase. hashey 16 Discord DAVE Protocol Cryptographic Design Review PUBLIC

4. Web client may allow a malicious server to conduct a machine-in-the-middle attack on the session
Severity: High Difficulty: High Type: Cryptography Finding ID: TOB-DISCE2EE-4 Target: E2EERTC protocol
Description The Discord web client stores key material in the browser, making it accessible to any
JavaScript code served by Discord or an adversary with access to a TLS certificate for discord.com .
Unlike mobile or desktop clients, the web client code is served to the user every time they
open the app. The delivered code may differ depending on the user requesting it, and it is
not signed with a code signing key, unlike mobile apps delivered by app stores. Web app
users are thus at increased risk of impersonation and machine-in-the-middle attacks due to malicious code.
Exploit Scenario An attacker owns a TLS root certificate and issues a fraudulent certificate for
discord.com . The attacker intercepts traffic from specific users and replaces the Discord
application with a backdoored version. The attacker can then eavesdrop on and tamper
with all further communication by the users. Recommendations
Short term, warn users that the security guarantees offered by the web client are weaker
than those offered by the mobile or desktop applications.
Long term, consider implementing a browser extension to enable code signature
verification and code fingerprint verification. For an example of how this could be done, see
this blog post on Meta's Code Verify browser plugin. hashey 17 Discord DAVE Protocol Cryptographic Design Review
PUBLIC

5. Committing members may fail to send Welcome messages Severity: Medium Difficulty: High
Type: Data Validation Finding ID: TOB-DISCE2EE-5 Target: E2EERTC protocol Description
To invite a new member to an MLS group, a committing member must send a commit containing an Add
proposal. Assuming the protocol uses public handshakes, a
signaling server can ensure that this takes place correctly. However, after all users process
the commit, the committing member must send a Welcome message to the joiners so the
joiner can initialize their ratchet tree state and begin decrypting messages. This Welcome
message is encrypted, so the signaling server cannot determine whether it is correctly formed.
One possible solution to this issue is to require each joiner to confirm to the DS that they have received a valid
Welcome message. After some timeout, the DS could then orchestrate
a retry of the addition. However, the joiner may lie about receiving an invalid Welcome
message, and when multiple joiners are added in a single epoch, they may disagree as to the validity of the Welcome
message. This adds complexity to the new member addition process and makes a denial of service more likely.
Exploit Scenario A malicious user wants to prevent new members from being added to a group. When the
DS announces a new member, the user commits the Add proposal, but never sends a Welcome

message. The joining user is then unable to decrypt messages and enter the call. Recommendations
Short term, have the signaling server monitor commits adding new members to ensure the
committer sends a corresponding Welcome message. If no Welcome is sent, the server
should remove the committer from the group of committing members and request a new Welcome
message from the group. If the new member rejects the Welcome message as
invalid, the server should simply request a new Welcome from a random committing member of the group.
Long term, consider using the ExternalInit flow to add new members to a group. The
DS serves a cached copy of the current GroupInfo struct to the new member, who then
submits an external commit to the group, adding themselves. This removes the need for
any existing group member to coordinate the addition of new members. hashey
18DiscordDAVEProtocolCryptographicDesignReview PUBLIC

6. Users could decrypt messages after being removed from the call UI Severity: High Difficulty: High
Type: Data Validation Finding ID: TOB-DISCE2EE-6 Target: E2E RTC protocol Description
The client UI listing the users that are reconnected to a given session is controlled by the
voice state of the corresponding users. This state is not directly controlled by Add or Remove
messages from the signaling server or the corresponding group epoch updates. In
practice, this means that a user may be removed from the UI before the group epoch is
updated, allowing the user to continue to decrypt messages to the session, even though
they appear disconnected from the point of view of other members.
Such a user would need to intercept WebRTC messages between other users after being
removed from the SFU, which would be feasible as a passive network adversary.
Additionally, an active network adversary may be able to impersonate current group
members using their knowledge of other members' sender keys. Exploit Scenario
Mallory, a malicious network adversary convinces a member of a voice group to add her to
the group by impersonating a desired group member. Upon realizing that Mallory is not the
intended group member, the existing members kick her from the group. Mallory is
removed from the Discord UI and the group members continue discussing sensitive
information. By delaying handshakes messages, Mallory keeps the MLSE epoch from
advancing and uses her network position to intercept and decrypt encrypted frames, thus
spying undetected on the conversation. Recommendations
Short term, ensure that users are not dropped from the list of participants until the
symmetric keys used to encrypt application data are updated. Until the group epoch is
updated and a new symmetric key is derived, the UI may display users as grayed out.
Long term, ensure that the MLSRatchett tree state and identities are auditable by the end
user and that the UI clearly reflects what identities can decrypt packets from a given E2EE session. hashey
19DiscordDAVEProtocolCryptographicDesignReview PUBLIC

7. Key fingerprint verification is vulnerable to partial preimage attacks Severity: Medium Difficulty: High
Type: Authentication Finding ID: TOB-DISCE2EE-7 Target: E2E RTC protocol Description
The Discord client will allow users to verify the signing keys associated with a particular
device pair. This is done by computing two repeated SHA-512 hashes over the user's
identity and the public key associated with the device, resulting in a 64-byte fingerprint that
consists of two concatenated SHA-512 digests. These fingerprints may then be verified out
of band (e.g., by meeting and comparing fingerprints).
However, it is well known that manual verification of long hexadecimal values is error-
prone. This could allow an attacker to impersonate other users by creating a key pair with a
fingerprint that is similar enough to the impersonated user. (The DEA was subject to an
attack of this type earlier this year, which led to the loss of 50,000 USD.) Exploit Scenario
Anation-state attacker who can conduct a machine-in-the-middle attack on Alice's
connection generates a signing key whose fingerprint has the same first four bytes and last
four bytes as Alice's fingerprint. The attacker registers their key for Alice's account and
impersonates Alice in calls. When Alice and Bob meet to verify their signing key fingerprints
out of band, Bob computes his fingerprint using the attacker's key. Alice and Bob fail to
notice this because they only compare the first and last digits of the fingerprint. Recommendations
Short term, encode the fingerprint as a QR code to allow users to verify it automatically if
their device has a camera that can scan QR codes. Use the manual verification process as a
fallback only, and direct users to use the QR code whenever possible.
Long term, review user studies on fingerprint verification and choose a format (e.g., sentence-
based fingerprints) that is less susceptible to partial preimage attacks. hashey
20DiscordDAVEProtocolCryptographicDesignReview PUBLIC

8. Abuser reporting mechanism is vulnerable to message forgery Severity: Medium Difficulty: High
Type: Data Validation Finding ID: TOB-DISCE2EE-8 Target: E2E RTC protocol Description
The current iteration of the protocol does not contain a mechanism for reporting abusive
content. However, the Discord team is planning to implement this in a future version of the

protocol by allowing user to capture voice and video streams and submit them to the Discord Trust and Safety team to review. Since there is no way to authenticate the captured streams, this mechanism is vulnerable to forgery attacks where the reporter creates a fake media stream and submits it as a report against another Discord user. Exploit Scenario Alice uses an online AI voice and video generator to create a fake recording of a video call with Bob. She submits this to Discord as proof of abusive behavior, and as a result, Bob is banned from the platform. Recommendations Short term, investigate solutions based on message franking (see the Abuse Reporting section in this white paper on Facebook Messenger), where the sender computes a binding commitment to the plaintext frames, which is either symmetrically or asymmetrically authenticated by the server. This offloads any storage requirements to the receiving client, who needs to store media and franking tags only if an abuse report is actively being recorded. To reduce the bandwidth requirements of the system, franking tags may be computed over a number of plaintext frames. Long term, revisit the literature on message franking in the future to see if new and more lightweight schemes are proposed for the WebRTC setting. hashey

21 Discord DAVE Protocol Cryptographic Design Review PUBLIC

9. DAVE's media encryption provides weakened forward secrecy guarantees Severity: Low Difficulty: High Type: Cryptography Finding ID: TOB-DISCE2EE-9 Target: E2E RTC protocol Description The DAVE protocol updates the symmetric encryption key whenever ML transitions to a new epoch. This provides forward secrecy and post-compromise security between different ML epochs. The ML protocol also includes an asymmetric ratchet mechanism called a secret tree, which can be used to regularly update symmetric keys within a given epoch. This symmetric ratchet also provides forward secrecy guarantees within a single epoch. The current protocol design does not use the ML symmetric ratchet, which means that the forward secrecy guarantees provided by the design are weaker than they should be. This is particularly true for long-lived epochs where new users join or leave the session. Exploit Scenario Eve, an attacker, can store encrypted frames from Alice, who is participating in a group call. At some point, Eve compromises Alice's symmetric key. Since the protocol does not ratchet symmetric keys, Eve can go back and decrypt all of Alice's messages sent during the current epoch. Recommendations Short term, use the ML secret tree to derive symmetric encryption keys and nonces. The mlsp library does not expose the secret tree directly, but it is possible to instantiate a new secret tree (as an mlspGroupKeySource struct) based on the exporter_secret key. These secure frame nonces can be repurposed as two 16-bit epoch and generation counters to ensure that the sending and receiving parties both agree on the current epoch and ratchet generation. hashey

22 Discord DAVE Protocol Cryptographic Design Review PUBLIC

10. The protocol may fail to encrypt AV1-encoded media frames Severity: High Difficulty: High Type: Cryptography Finding ID: TOB-DISCE2EE-10 Target: E2E RTC protocol Description In the current design of the DAVE protocol, AV1-encoded frames are only partially encrypted. More precisely, the encryptor assumes that each frame contains at most one frame OBU and that this is the last OBU in the frame. The frame OBU payload is then encrypted, leaving the OBU header and the remaining OBUs unencrypted. However, according to section 6.9 of the AV1 specification, a frame OBU is simply a more compact representation of a frame header OBU, followed by a tile group OBU, and the same frame could, in principle, be represented either way: A frame OBU consists of a frame header OBU and a tile group OBU packed into a single OBU. The intention is to provide a more compact way of coding the common use case where the frame header is immediately followed by tile group data. If the AV1 encoder uses separate frame header and tile group OBUs to represent the frame, the OBUs would not be encrypted by the sender under the current design. This would immediately disclose the current frame of the sender's media stream to the SFU. This section on ordering OBUs (section 7.5 in the AV1 specification) seems to require that frame OBUs and tile group OBUs be placed last in the encoded frame, but this section also mentions that "Some applications may choose to use bitstream that are not fully conformant to the requirements described in this section." In practice, this means that nonconformant encoders may disclose frame and tile group OBUs. Exploit Scenario Alice's Discord client uses AV1 to encode her media stream. The stream is encoded using separate frame header and tile group OBUs, causing Alice's client to send the entire media stream in plaintext to the server, thus breaking the protocol's send-to-end confidentiality guarantees. hashey

23 Discord DAVE Protocol Cryptographic Design Review PUBLIC

Recommendations Short term, test which types of OBUs are emitted by the different Discord client implementations to ensure that the media stream is encrypted correctly by each client. At least the following OBUs should be encrypted to preserve the media stream's confidentiality: • Frame header OBUs: Contain metadata describing how the current frame has changed compared to previous frames •

TilegroupOBUs: Contain tiled data associated with the frame •
TilelistOBUs: Contain tiled data similar to a tile group OBU, together with additional headers allowing the decoder to process parts of the current frame (TilelistOBUs are recurrently dropped by the WebRTC packetizer according to the documentation provided by Discord.)
Long term, encrypt frame header, tile group, and frame OBUs independent of their position in the encoded frame. Review the AV1 specification in detail to ensure that user data is not disclosed by the remaining OBUs. References • A Technical Overview of AV1 • AV1 Bitstream & Decoding Process Specification hashey 24 Discord DAVE Protocol Cryptographic Design Review PUBLIC

11. Video frame header length and header data are not authenticated Severity: High Difficulty: High
Type: Cryptography Finding ID: TOB-DISCE2EE-11 Target: E2E RTC protocol Description
Because packetization and depacketization rely on values located in the frame payload, a contiguous initial portion of each frame is left unencrypted. Currently, the design does not include these unencrypted headers in the associated data of AES-GCM. A malicious SFU could thus modify these headers undetected.
Of particular concern are the H.264/H.265 and AV1 frames, which contain variable-length unencrypted header data. Because the unencrypted header frame size value is not authenticated, a malicious SFU can replace it with a larger value. The SFU can then include additional unencrypted video-coding layers, which will be copied directly into the output frame. Additionally, if nonce replay protections are not present on the decryptor side, the SFU could use the nonce and authentication tag from a fully unencrypted header-only frame to forge arbitrary plaintext frames. 66 if (codecUnencryptedHeaderBytes) { 67 // copy the unencrypted header to the output frame 68 memcpy(frame.data(), encryptedFrame.data(), codecUnencryptedHeaderBytes); 69 }
Figure 11.1: Unauthenticated data is directly included in the visible frame. (discord/discord_common/native/secure_frames/decryptor.cpp#66-69)
Allowing the SFU to include or replace content in user media frames would allow impersonation of users by the Discord server, even when ML Send-to-end persistent identity verification is in use. Additionally, unauthenticated code headers present a large surface area for exploitation of implementation flaws in the underlying codecs. Furthermore, lack of header authentication may open clients to adaptive attacks against confidentiality. For example, forged video-coding layers may reference elements defined in encrypted video-coding layers, disclosing information about the encrypted layers by observing differential failures in the decoder. Exploit Scenario
A malicious Discord employee modifies the SFU to add video codec NALs that replace a particular user's broadcast stream with a pre-recorded video of a fake to look like a legitimate stream. Advanced users check the persistent identity fingerprint and are convinced that the stream is legitimate.
Recommendations Short term, include all unencrypted components of media frames in the authenticated data of the underlying AEAD. Long term, keep abreast of new extensions to the WebRTC API that would allow Discord to switch to a packet-based transport encryption mechanism. This would allow a future version of the protocol to encrypt media content and headers at the RTP packet level, which would allow the protocol to encrypt and authenticate the entire encoded frame in a way that would be invisible to the RTP packetizer and depacketizer. hashey 25 Discord DAVE Protocol Cryptographic Design Review PUBLIC

A. Vulnerability Categories

The following tables describe the vulnerability categories, severity levels, and difficulty levels used in this document. Vulnerability Categories Category Description
Access Controls Insufficient authorization or assessment of rights
Auditing and Logging Insufficient auditing of actions or logging of problems
Authentication Improper identification of users
Configuration Misconfigured servers, devices, or software components
Cryptography A breach of system confidentiality or integrity Data Exposure Exposure of sensitive information
Data Validation Improper reliance on the structure or values of data
Denial of Service A system failure with an availability impact
Error Reporting Insecure or insufficient reporting of error conditions
Patching Use of an outdated software package or library
Session Management Improper identification of authenticated users
Testing Insufficient test methodology or test coverage Timing Race conditions or other order-of-operations flaws Undefined Behavior Undefined behavior triggered within the system hashey 26 Discord DAVE Protocol Cryptographic Design Review PUBLIC

SeverityLevels SeverityDescription
InformationalTheissuedoesnotposean immediateriskbutisrelevanttosecuritybest practices.
UndeterminedTheextentoftheriskwasnotdeterminedduringthisengagement.
LowTheriskissmallorisnotonetheclienthasindicatedisimportant.
MediumUserinformationisatrisk;exploitationcouldposereputational,legal,or moderatefinancialrisks.
HighTheflawcouldaffectnumeroususersandhaveseriousreputational,legal, orfinancialimplications.
DifficultyLevels DifficultyDescription
UndeterminedThedifficultyofexploitationwasnotdeterminedduringthisengagement.
LowTheflawiswellknown;publictoolsforitsexploitationexistorcanbe scripted.
MediumAnattackermustwriteanexploitorwillneedin-depthknowledgeofthe system.
HighAnattackermusthaveprivilegedaccesstothesystem,mayneedtoknow
complextechnicaldetails,ormustdiscoverotherweaknessestoexploitthis issue. hashey
28DiscordDAVEProtocolCryptographicDesignReview PUBLIC

B.GuidanceonUsingShortAES-GCMtags

ThisappendixlistsconsiderationsthatmustbeobservedwhenusingtruncatedAES-GCM tags.WhileNISTSP800-38Dpermitstheuseof32-and64-bitauthenticationtagsinsome real-timeapplications,certainpropertiesofAES-GCMmakeshorttagsdangerousinmany situations.
WerecommendGCMtaglengthsof64bitsataminimumandatleast96bitsifpossible.
OtherAEADsbasedonHMACsratherthanpolynomialMACsaremorerobusttousewith truncatedtags.
TwopropertiesofAES-GCM,detailedinAuthenticationweaknessesinGCM,impactthe securityofschemasusingtruncatedGCMtags:

1.Longmessagesdecreasetheeffortrequiredtoconstructaforgery.If an adversaryobservesaproperlyauthenticatedmessageofbitlength $2k$,withtag length t , thentheadversarycanforgeadifferentmessagewithaprobabilityof roughly $2^{-(t-k)}$.
Asaconcreteexample,ifanhonestusersendsa128-KBpacketusing a32-bit tag,thentheadversarywillneedtosendroughly 2^{22} messagesbefore findingasuccessfulforgery.

2.Forgeriesagainsttruncatedtagsdiscloseinformationaboutthe authenticationkey.If anadversarycandeterminewhichforgeryattemptsare successful,theycanusethatinformationtoincreaseprobabilityoffuture forgeries,eventuallyrevealingthefullauthenticationkey. NISTSP800-38D:AppendixCspecifiesrequirementsandguidelinesforusingshorttags; wesummarizethoserelevanttotheDiscordDAVEprotocol:

1.Donotrevealinformationaboutdecryptionssuccessforindividualpackets.In particular,donotsendretryrequestsforframes thatfailauthentication.Consider mitigatingtimingsidechannelssothatexternallyvisibletimingbehaviorissimilar fordecryptionssuccessorfailure.
2.Includeaslittleencrypted/authenticateddataaspossibleineachframe.Beawareof opportunitiesforadversaries,suchastheSFU,torequesttheencryptionor authenticationoflargeframes.
3.Designthesystemtobero bustagainstindividualframeforgeries.Forgeryofa singleframeshouldresultinonlyaminorchangeintheenduser'smediastream andshouldnotcausesystem-wideeffects. 4.Rotatesenderkeys frequently,preferablyviasymmetricderivation.Asymmetrickey rotationviaMLSCommitsmaybeselectivelydelayedbyamaliciousDS. hashey
29DiscordDAVEProtocolCryptographicDesignReview PUBLIC

C.FixReviewResults Whenundertakingafixreview,hasheyreviewsthefixesimplementedforissues identifiedintheoriginalreport.Thisworkinvolvesareviewofspecificareasoftheupdated protocol designdocumentation,notcomprehensiveanalysisoftheentiresystem.
OnNovember9,2023,hasheyreviewedthefixesandmitigationsimplementedbythe Discordteamfortheissuesidentifiedinthisreport.Wereviewedeachfixtodetermineits effectivenessinresolvingtheassociatedissue.
Insummary,ofthe11issuesdescribedinthisreport,Discordhasresolved10issuesand haspartiallyresolvedtheonere remainingissue.Foradditionalinformation,pleaseseethe DetailedFixReviewResultsbelow. IDTitleStatus 1CommitspamcouldpreventgroupupdatesResolved
2CommitleafnodesarenotvalidatedbythemlspllibraryResolved
3ThemlspllibrarydoesnotvalidatekeypackagelifetimesResolved
4Webclientmayallowamaliciousservertoconducta machine-in-the-middleattackonthesession Resolved
5CommittingmembersmayfailtosendWelcomemessagesResolved
6UserscoulddecryptmessagesafterbeingremovedfromthecallUIResolved
7KeyfingerprintverificationisvulnerabletopartialpreimageattacksResolved
8AbusereportingmechanismisvulnerabletomessageforgeryPartially Resolved
9Secureframesencryptionprovidesweakenedforwardsecrecy guarantees Resolved hashey
30DiscordDAVEProtocolCryptographicDesignReview PUBLIC

10SecureframesmayfailtoencryptAV1-encodedmediaframesResolved
11VideoframeheaderlengthandheaderdataarenotauthenticatedResolved hashey

DetailedFixReviewResults TOB-DISCE2EE-1:Commitspamcouldpreventgroupupdates Resolved.TheDShasbeenaddedasanexternalsenderandistheonlyoneallowedto makeproposalstothe group.CommitmessagesarefilteredbytheDS,whichacceptsonly commitscontainingallthecurrentlyopenAddandRemoveproposalsmadebytheDS. TOB-DISCE2EE-2:Commitleafnodesarenotvalidatedbythemlspplibrary Resolved.BoththeDSandtheindividualmembersofthegrouppnowvalidateallcommits toensurethatmembercredentialandsigningkeysarenoupdatedbythecommit. TheDiscordteamhasalsodecidedtoperformaninternalreviewofthe mlspplibraryto identifyanypotentialsecurityissuesinit.Anyissuesidentifiedwillbeaddressedandthefix willbemergedupstreamto mlsppl . TOB-DISCE2EE-3:Themlspplibrarydoesnotvalidatekeypackage lifetimes Resolved.TheDSnowensuresthatthekeypackagefields not_before and not_after are settothemaximumallowedlifespan. TOB-DISCE2EE-4:Webclientmayallowamaliciousservertoconducta machine-in-the-middleattackonthesession Resolved.TheDiscordteamwillinformusersthattheDiscordwebclientcannotofferthe same security guarantees as native clients against machine-in-the-middle attacks. TOB-DISCE2EE-5:CommittingmembersmayfailtosendWelcome messages Resolved.TheDSnowrequiresthecommittingmembertoincludethe Welcome message withthecorresponding Commit.Ifamemberfailstotransmit Welcome messages,theymay beremovedfromthegroupofcommittingmembers.If a Welcome messageisreportedas invalidbythejoiningmember,theDSmayrequestanew Welcome messagefroma differentmemberofthegroup. TOB-DISCE2EE-6:UserscoulddecryptmessagesafterbeingremovedfromthecallUI Resolved.OfficialDiscordclientswillbeabletoreviewthecurrentMLSgroupstatetoseeif leavingmemberscanstilldecryptmessagestothe group.AccordingtothesectiononUI considerations,thisshouldalsobereflectedbythemediasessionmodalUI. Additionally,theDiscordteamisexploringmorelong-term solutionswheretheMLSgroup membershipstatewouldbedirectlyreflectedbytheclientUI. TOB-DISCE2EE-7:Keyfingerprintverificationisvulnerabletopartialpreimageattacks Resolved.ThefingerprintverificationUIwilluseaQRcodeforautomaticverificationand word-basedfingerprintsasamanualfallback. hashey 32DiscordDAVEProtocolCryptographicDesignReview PUBLIC

TOB-DISCE2EE-8:Abusereportingmechanismisvulnerabletomessageforgery Partiallyresolved.Thesectiononabusereportinginthedesigndocumentationnowclearly states that a design based on message frankings should be preferred as long as the impacts on bandwidth and storage are negligible. However, the abusereporting mechanism will not be part of the initial release of the protocol, according to the Discord team. Since the details around the design of this feature were not determined at the time of the fix review, this issue is considered partially resolved. TOB-DISCE2EE-9:Secureframesencryptionprovidesweakenedforwardsecrecy guarantees Resolved.These secureframes protocol now uses an MLS secret tree initialized using the exporter_secret key to derive symmetric encryption keys. TOB-DISCE2EE-10:SecureframesmayfailtoencryptAV1-encoded media frames Resolved. All OBU payloads are now encrypted independently of their position within the frame. TOB-DISCE2EE-11:Video frame header length and header data are not authenticated Resolved. All unencrypted data from the frame will now be included as additional data for the AEAD. hashey 33DiscordDAVEProtocolCryptographicDesignReview PUBLIC

D. Fix Review Status Categories

The following table describes the statuses used to indicate whether an issue has been sufficiently addressed. FixStatus StatusDescription UndeterminedThe status of the issue was not determined during this engagement. UnresolvedThe issue persists and has not been resolved. PartiallyResolvedThe issue persists but has been partially resolved. ResolvedThe issue has been sufficiently resolved. hashey 34DiscordDAVEProtocolCryptographicDesignReview PUBLIC