The Authenticated Received Chain (ARC) Protocol

Abstract

The Authenticated Received Chain (ARC) protocol provides an authenticated "chain of custody" for a message, allowing each entity that handles the message to see what entities handled it before and what the message’s authentication assessment was at each step in the handling.

ARC allows Internet Mail Handlers to attach assertions of message authentication assessment to individual messages. As messages traverse ARC-enabled Internet Mail Handlers, additional ARC assertions can be attached to messages to form ordered sets of ARC assertions that represent the authentication assessment at each step of the message-handling paths.

ARC-enabled Internet Mail Handlers can process sets of ARC assertions to inform message disposition decisions, identify Internet Mail Handlers that might break existing authentication mechanisms, and convey original authentication assessments across trust boundaries.
Status of This Memo

This document is not an Internet Standards Track specification; it is published for examination, experimental implementation, and evaluation.

This document defines an Experimental Protocol for the Internet community. This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are candidates for any level of Internet Standard; see Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc8617.

Copyright Notice

Copyright (c) 2019 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 (https://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.
Table of Contents

1. Introduction ................................................. 4
2. General Concepts ............................................. 5
   2.1. Evidence ............................................. 5
   2.2. Custody .............................................. 5
   2.3. Chain of Custody ..................................... 6
   2.4. Validation of Chain of Custody ......................... 6
3. Terminology and Definitions ................................. 6
   3.1. ARC Set ............................................. 7
   3.2. Authenticated Received Chain (ARC) ..................... 7
   3.3. Internet Mail Handlers / Intermediaries ................. 7
   3.4. Authentication Assessment ............................. 7
   3.5. Signing vs. Sealing ................................... 8
   3.6. Sealer ............................................... 8
   3.7. Validator ............................................ 8
   3.8. Imported ABNF Tokens .................................. 8
   3.9. Common ABNF Tokens ................................... 8
4. Protocol Elements ............................................ 9
   4.1. ARC Header Fields ..................................... 9
      4.1.1. ARC-Authentication-Results (AAR) .................. 9
      4.1.2. ARC-Message-Signature (AMS) ..................... 9
      4.1.3. ARC-Seal (AS) ................................... 11
      4.1.4. Internationalized Email (EAI) .................... 12
   4.2. ARC Set .............................................. 12
      4.2.1. Instance Tags .................................... 12
   4.3. Authenticated Received Chain .......................... 13
   4.4. Chain Validation Status ............................... 13
5. Protocol Actions ............................................. 14
   5.1. Sealer Actions ....................................... 14
      5.1.1. Header Fields to Include in ARC-Seal Signatures 15
      5.1.2. Marking and Sealing "cv=fail" (Invalid) Chains .. 15
      5.1.3. Only One Authenticated Received Chain per Message 16
      5.1.4. Broad Ability to Seal ............................. 16
      5.1.5. Sealing Is Always Safe ........................... 16
   5.2. Validator Actions ...................................... 17
      5.2.1. All Failures Are Permanent ....................... 18
      5.2.2. Responding to ARC Validation Failures during the SMTP Transaction .................. 19
6. Communication of Validation Results ....................... 19
7. Use Cases .................................................. 19
   7.1. Communicate Authentication Assessment across Trust Boundaries ........................... 19
      7.1.1. Message-Scanning Services ....................... 20
      7.1.2. Multi-tier MTA Processing ....................... 20
      7.1.3. Mailing Lists ................................... 20
   7.2. Inform Message Disposition Decisions .................. 21
      7.2.1. DMARC Local Policy Overrides ..................... 21
1. Introduction

The utility of widely deployed email authentication technologies such as Sender Policy Framework (SPF) [RFC7208] and DomainKeys Identified Mail (DKIM) [RFC6376] is impacted by the processing of Internet Mail by intermediate handlers. This impact is thoroughly documented in the defining documents for SPF and DKIM and further discussed in [RFC6377] and [RFC7960].

Domain-based Message Authentication, Reporting, and Conformance (DMARC) [RFC7489] also relies upon SPF and DKIM authentication mechanisms. Failures of authentication caused by the actions of intermediate handlers can cause legitimate mail to be incorrectly rejected or misdirected.

Authenticated Received Chain (ARC) creates a mechanism for individual Internet Mail Handlers to add their authentication assessment to a message’s ordered set of handling results. ARC encapsulates the authentication assessment in a DKIM signature derivative to grant other handlers the ability to verify the authenticity of the individual assessment assertion as well as the aggregate set and sequence of results.

Ordered sets of authentication assessments can be used by ARC-enabled Internet Mail Handlers to inform message-handling disposition, identify where alteration of message content might have occurred, and provide additional trace information for use in understanding message-handling paths.

2. General Concepts

ARC is loosely based on concepts from evidence collection. Evidence is usually collected, labeled, stored, and transported in specific ways to preserve the state of evidence and to document all processing steps.

2.1. Evidence

In ARC’s situation, the "evidence" is a message’s authentication assessment at any point along the delivery path between origination and final delivery. Determination of message authentication can be affected when intermediate handlers modify message content (header fields and/or body content), route messages through unforeseen paths, or change envelope information.

The authentication assessment for a message is determined upon receipt of a message and documented in the Authentication-Results header field(s). ARC extends this mechanism to survive transit through intermediary Administrative Management Domains (ADMDs).

Because the first-hand determination of an authentication assessment can never be reproduced by other handlers, the assertion of the authentication assessment is more akin to testimony by a verifiable party than to hard evidence, which can be independently evaluated.

2.2. Custody

"Custody" refers to when an Internet Mail Handler processes a message. When a handler takes custody of a message, the handler becomes a custodian and attaches its own evidence (authentication assessment upon receipt) to the message if it is ARC enabled. Evidence is added in such a way that future handlers can verify the authenticity of both evidence and custody.
2.3. Chain of Custody

The "chain of custody" of ARC is the entire set of evidence and custody that travels with a message.

2.4. Validation of Chain of Custody

Any ARC-enabled Internet Mail Handler can validate the entire set of custody and the authentication assessments asserted by each party to yield a valid chain of custody. If the evidence-supplying custodians can be trusted, then the validated chain of custody describes the (possibly changing) authentication assessment as the message traveled through various custodians.

Even though a message’s authentication assessment might have changed, the validated chain of custody can be used to determine if the changes (and the custodians responsible for the changes) can be tolerated.

3. Terminology and Definitions

This section defines terms used in the rest of the document.

Readers should to be familiar with the contents, core concepts, and definitions found in [RFC5598]. The potential roles of transit services in the delivery of email are directly relevant.

Language, syntax (including some ABNF constructs), and concepts are imported from DKIM [RFC6376]. Specific references to DKIM are made throughout this document. The following terms are imported from [RFC5598]:

- Administrative Management Domain (ADMD), Section 2.3
- Message Transfer Agent (MTA), Section 4.3.2
- Message Submission Agent (MSA), Section 4.3.1
- Message Delivery Agent (MDA), Section 4.3.3

Syntax descriptions use ABNF [RFC5234] [RFC7405].

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "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.
3.1. ARC Set

Section 4.1 introduces three (3) ARC header fields that are added to a message by an ARC-enabled Internet Mail Handler. Together, these three header fields compose a single "ARC Set". An ARC Set provides the means for an Internet Mail Handler to attach an authentication assessment to a message in a manner that can be verified by future handlers. A single message can contain multiple ARC Sets.

In general concept terms, an ARC Set represents Evidence and Custody.

3.2. Authenticated Received Chain (ARC)

The sequence of ARC Sets attached to a message at a given time is called the "Authenticated Received Chain" or "ARC". An Authenticated Received Chain is the record of individual authentication assessments as a message traverses through ARC-participating ADMDs.

The first attachment of an ARC Set to a message causes an Authenticated Received Chain to be created. Additional attachments of ARC Sets cause the Authenticated Received Chain to be extended.

In general concept terms, an Authenticated Received Chain represents a chain of custody.

3.3. Internet Mail Handlers / Intermediaries

Internet Mail Handlers process and deliver messages across the Internet and include MSAs, MTAs, MDAs, gateways, and mailing lists as defined in [RFC5598].

Throughout this document, the term "intermediaries" refers to both regular MTAs as well as delivery/reposting agents such as mailing lists covered within the scope of transit services per [RFC5598].

"Intermediaries" and "Internet Mail Handlers" are used synonymously throughout this document.

3.4. Authentication Assessment

The authentication assessment that is affixed to a message as part of each ARC Set consists of the "authres-payload" [RFC8601]. For the integrity of an ARC Set, the authentication assessment only needs to be properly encapsulated within the ARC Set as defined in Section 4.1. The accuracy or syntax of the authres-payload field does not affect the validity of the ARC Chain itself.
3.5. Signing vs. Sealing

Signing is the process of affixing a digital signature to a message as a header field, such as when a DKIM-Signature (as in [RFC6376], Section 2.1), an AMS, or an AS is added. Sealing is when an ADMD affixes a complete and valid ARC Set to a message to create or continue an Authenticated Received Chain.

3.6. Sealer

A Sealer is an Internet Mail Handler that attaches a complete and valid ARC Set to a message.

In general concept terms, a Sealer adds its testimony (assertion of authentication assessment) and proof of custody to the chain of custody.

3.7. Validator

A Validator is an ARC-enabled Internet Mail Handler that evaluates an Authenticated Received Chain for validity and content. The process of evaluation of the individual ARC Sets that compose an Authenticated Received Chain is described in Section 5.2.

In general concept terms, a Validator inspects the chain of custody to determine the content and validity of individual evidence supplied by custodians.

3.8. Imported ABNF Tokens

The following ABNF tokens are imported:

- tag-list ([RFC6376], Section 3.2)
- authres-payload ([RFC8601], Section 2.2)
- CFWS ([RFC5322], Section 3.2.2)

3.9. Common ABNF Tokens

The following ABNF tokens are used elsewhere in this document:

```
position     = 1*2DIGIT  ; 1 - 50
instance     = [CFWS] %s"i" [CFWS] "="
                [CFWS] position
chain-status = ("none" / "fail" / "pass")
seal-cv-tag  = %s"cv" [CFWS] "="
                [CFWS] chain-status
```

Andersen, et al.  Experimental  [Page 8]
4.  Protocol Elements

4.1.  ARC Header Fields

ARC introduces three new header fields. The syntax for new header fields adapts existing specifications. This document only describes where ARC-specific changes in syntax and semantics differ from existing specifications.

4.1.1.  ARC-Authentication-Results (AAR)

The ARC-Authentication-Results (AAR) header field records the message authentication assessment as processed by an ARC-participating ADMD at message arrival time.

In general concept terms, the AAR header field is where evidence is recorded by a custodian.

The AAR header field is similar in syntax and semantics to an Authentication-Results field [RFC8601], with two (2) differences:

- the name of the header field itself and
- the presence of the instance tag. Additional information on the instance tag can be found in Section 4.2.1.

The formal ABNF for the AAR header field is:

```
arc-info = instance [CFWS] ";" authres-payload
arc-authres-header = "ARC-Authentication-Results:" [CFWS] arc-info
```

Because there is only one AAR allowed per ARC Set, the AAR MUST contain the combined authres-payload with all of the authentication results from within the participating ADMD, regardless of how many Authentication-Results header fields are attached to the message.

4.1.2.  ARC-Message-Signature (AMS)

The ARC-Message-Signature (AMS) header field allows an ARC-participating ADMD to convey some responsibility (custodianship) for a message and possible message modifications to future ARC-participating custodians.

In general concept terms, the AMS header field identifies a custodian.
The AMS header field has the same syntax and semantics as the DKIM-Signature field [RFC6376], with three (3) differences:

- the name of the header field itself;
- no version tag ("v") is defined for the AMS header field. As required for undefined tags (in [RFC6376]), if seen, a version tag MUST be ignored; and
- the "i" (Agent or User Identifier (AUID)) tag is not imported from DKIM; instead, this tag is replaced by the instance tag as defined in Section 4.2.1.

ARC places no requirements on the selectors and/or domains used for the AMS header field signatures.

The formal ABNF for the AMS header field is:

```
arc-ams-info = instance [CFWS] ";" tag-list
arc-message-signature = "ARC-Message-Signature:" [CFWS] arc-ams-info
```

To reduce the chances of accidental invalidation of AMS signatures:

- AMS header fields are added by ARC-participating ADMDs as messages exit the ADMD. AMS header fields SHOULD be attached so that any modifications made by the ADMD are included in the signature of the AMS header field.
- Authentication-Results header fields MUST NOT be included in AMS signatures as they are likely to be deleted by downstream ADMDs (per [RFC8601], Section 5).
- ARC-related header fields (ARC-Authentication-Results, ARC-Message-Signature, and ARC-Seal) MUST NOT be included in the list of header fields covered by the signature of the AMS header field.

To preserve the ability to verify the integrity of a message, the signature of the AMS header field SHOULD include any DKIM-Signature header fields already present in the message.
4.1.3. ARC-Seal (AS)

The AS header field permits ARC-participating ADMDs to verify the integrity of AAR header fields and corresponding AMS header fields.

In general concept terms, the AS header field is how custodians bind their authentication assessments (testimonials) into a chain of custody so that Validators can inspect individual evidence and custodians.

The AS header field is similar in syntax and semantics to DKIM-Signature header fields [RFC6376], with the following differences:

- the "i" (AUID) tag is not imported from DKIM; instead, this tag is replaced by the instance tag as defined in Section 4.2.1;
- the signature of the AS header field does not cover the body of the message; therefore, there is no "bh" tag. The signature of the AS header field only covers specific header fields as defined in Section 5.1.1;
- no body canonicalization is performed as the AS signature does not cover the body of a message;
- only "relaxed" header field canonicalization ([RFC6376], Section 3.4.2) is used;
- the only supported tags are "i" (from Section 4.2.1 of this document), and "a", "h", "d", "s", and "t" from [RFC6376], Section 3.5. Note especially that the DKIM "h" tag is NOT allowed and, if found, MUST result in a cv status of "fail" (for more information, see Section 5.1.1); and
- an additional tag, "cv" ("seal-cv-tag" in the ARC-Seal ABNF definition), is used to communicate the Chain Validation Status to subsequent ADMDs.

ARC places no requirements on the selectors and/or domains used for the AS header field signatures.

The formal ABNF for the AS header field is:

```
arc-as-info = instance [CFWS] "" tag-list
arc-seal = "ARC-Seal:" [CFWS] arc-as-info
```
4.1.4. Internationalized Email (EAI)

In internationalized messages [RFC6532], many header fields can contain UTF-8 as well as ASCII text. The changes for EAI are all inherited from DKIM as updated by [RFC8616] and Authentication-Results (A-R) as updated in [RFC8601], but they are called out here for emphasis.

In all ARC header fields, the d= and s= tags can contain U-labels. In all tags, non-ASCII characters need not be quoted in dkim-quoted-printable.

The AAR header allows UTF-8 in the same places that Authentication-Results does, as described in [RFC8601].

4.2. ARC Set

An "ARC Set" is a single collection of three ARC header fields (AAR, AMS, and AS). ARC header fields of an ARC Set share the same "instance" value.

By adding all ARC header fields to a message, an ARC Sealer adds an ARC Set to a message. A description of how Sealers add an ARC Set to a message is found in Section 5.1.

4.2.1. Instance Tags

Instance tags describe which ARC header fields belong to an ARC Set. Each ARC header field of an ARC Set shares the same instance tag value.

Instance tag values are integers that begin at 1 and are incremented by each addition of an ARC Set. Through the incremental values of instance tags, an ARC Validator can determine the order in which ARC Sets were added to a message.

Instance tag values can range from 1-50 (inclusive).

_INFORMATIONAL_: The upper limit of 50 was picked based on some initial observations reported by early working group members. The value was chosen to balance the risk of excessive header field growth (see Section 9.1) against expert opinion regarding the probability of long-tail, but non-looping, multiple-intermediary mail flows. Longer ARC Chains will also impose a load on Validators and DNS to support additional verification steps. Observed quantities of "Received" header fields were also considered in establishing this as an experimental initial value.
Valid ARC Sets MUST have exactly one instance of each ARC header field (AAR, AMS, and AS) for a given instance value and signing algorithm.

For handling multiple signing algorithms, see [ARC-MULTI].

4.3. Authenticated Received Chain

An Authenticated Received Chain is an ordered collection of ARC Sets. As ARC Sets are enumerated sets of ARC header fields, an Authenticated Received Chain represents the output of message authentication assessments along the handling path of ARC-enabled processors.

Authentication assessments determined at each step of the ARC-enabled handling path are present in an Authenticated Received Chain in the form of AAR header fields. The ability to verify the identity of message handlers and the integrity of message content is provided by AMS header fields. AS header fields allow message handlers to validate the assertions, order, and sequence of the Authenticated Received Chain itself.

In general concept terms, an Authenticated Received Chain represents a message’s chain of custody. Validators can consult a message’s chain of custody to gain insight regarding each custodian of a message and the evidence collected by each custodian.

4.4. Chain Validation Status

The state of the Authenticated Received Chain at a specific processing step is called the "Chain Validation Status". Chain Validation Status information is communicated in several ways:

- as the AS header field in the "cv" tag and
- as part of the Authentication-Results and AAR header field(s).

Chain Validation Status has one of three possible values:

- none: There was no Authenticated Received Chain on the message when it arrived for validation. Typically, this occurs when a message is received directly from a message’s original Message Transfer Agent (MTA) or Message Submission Agent (MSA), or from an upstream Internet Mail Handler that is not participating in ARC handling.
- fail: The message contains an Authenticated Received Chain whose validation failed.
pass: The message contains an Authenticated Received Chain whose validation succeeded.

5. Protocol Actions

ARC-enabled Internet Mail Handlers generally act as both ARC Validators (when receiving messages) and ARC Sealers (when sending messages onward, not originated locally).

An Authenticated Received Chain with a Chain Validation Status of "pass" (or "none") allows Internet Mail Handlers to ascertain:

- all ARC-participating ADMDs that claim responsibility for handling (and possibly modifying) the message in transit and
- the authentication assessments of the message as determined by each ADMD (from AAR header fields).

With this information, Internet Mail Handlers MAY inform local policy decisions regarding disposition of messages that experience authentication failure due to intermediate processing.

5.1. Sealer Actions

To "seal" a message, an ARC Sealer adds an ARC Set (the three ARC header fields AAR, AMS, and AS) to a message. All ARC header fields in an ARC Set share the same instance tag value.

To perform sealing (aka to build and attach a new ARC Set), the following actions must be taken by an ARC Sealer when presented with a message:

1. All message modifications (including adding a DKIM-Signature header field(s)) MUST be performed before sealing.

2. If the message already contains an Authenticated Received Chain with the most recent AS reporting "cv=fail", there is no need to proceed and the algorithm stops here.

3. Calculate the instance value. If the message already contains an Authenticated Received Chain, the instance value is 1 more than the highest instance number found in the Authenticated Received Chain. If no Authenticated Received Chain exists, the instance value is 1.
4. Using the calculated instance value, generate and attach a complete ARC Set to the message as follows:

   A. Generate and attach an ARC-Authentication-Results header field as defined in Section 4.1.1.

   B. Generate and attach an ARC-Message-Signature header field as defined in Section 4.1.2.

   C. Generate and attach an ARC-Seal header field using the AS definition found in Section 4.1.3, the prescribed headers defined in Section 5.1.1, and the Chain Validation Status as determined during ARC validation.

5.1.1. Header Fields to Include in ARC-Seal Signatures

   The ARC-Seal is generated in a manner similar to how DKIM-Signature header fields are added to messages ([RFC6376], Section 3.7), with explicit requirements on the header fields and ordering of those fields.

   The signature of an AS header field signs a canonicalized form of the ARC Set header field values. The ARC Set header field values are supplied to the hash function in increasing instance order, starting at 1, and include the ARC Set being added at the time of sealing the message.

   Within an ARC Set, header fields are supplied to the hash function in the following order:

   1. ARC-Authentication-Results

   2. ARC-Message-Signature

   3. ARC-Seal

   Note that when an Authenticated Received Chain has failed validation, the signing scope for the ARC-Seal is modified as specified in Section 5.1.2.

5.1.2. Marking and Sealing "cv=fail" (Invalid) Chains

   In the case of a failed Authenticated Received Chain, the header fields included in the signature scope of the AS header field b= value MUST only include the ARC Set header fields created by the MTA that detected the malformed chain, as if this newest ARC Set was the only set present.
5.1.3. Only One Authenticated Received Chain per Message

A message can have only one Authenticated Received Chain on it at a time. Once broken, the chain cannot be continued, as the chain of custody is no longer valid, and responsibility for the message has been lost. For further discussion of this topic and the design restriction that prevents chain continuation or re-establishment, see [ARC-USAGE].

5.1.4. Broad Ability to Seal

ARC is not solely intended for perimeter MTAs. Any Internet Mail Handler MAY seal a message by adding a complete ARC Set, whether or not they have modified or are aware of having modified the message. For additional information, see Section 7.1.

5.1.5. Sealing Is Always Safe

The utility of an Authenticated Received Chain is limited to very specific cases. Authenticated Received Chains are designed to provide additional information to an Internet Mail Handler when evaluating messages for delivery in the context of authentication failures. Specifically:

- Properly adding an ARC Set to a message does not damage or invalidate an existing Authenticated Received Chain.
- Sealing an Authenticated Received Chain when a message has not been modified does not negatively affect the chain.
- Validating a message exposes no new threat vectors (see Section 9).
- An ADMD may choose to seal all inbound messages whether or not a message has been modified or will be retransmitted.
5.2. Validator Actions

A Validator performs the following steps, in sequence, to process an Authenticated Received Chain. Canonicalization, hash functions, and signature validation methods are imported from [RFC6376], Section 5.

1. Collect all ARC Sets currently attached to the message.
   * If there are none, the Chain Validation Status is "none", and the algorithm stops here.
   * The maximum number of ARC Sets that can be attached to a message is 50. If more than the maximum number exist, the Chain Validation Status is "fail", and the algorithm stops here.
   * In the following algorithm, the maximum discovered ARC instance value is referred to as "N".

2. If the Chain Validation Status of the highest instance value ARC Set is "fail", then the Chain Validation Status is "fail", and the algorithm stops here.

3. Validate the structure of the Authenticated Received Chain. A valid ARC has the following conditions:
   A. Each ARC Set MUST contain exactly one each of the three ARC header fields (AAR, AMS, and AS).
   B. The instance values of the ARC Sets MUST form a continuous sequence from 1..N with no gaps or repetition.
   C. The "cv" value for all ARC-Seal header fields MUST NOT be "fail". For ARC Sets with instance values > 1, the values MUST be "pass". For the ARC Set with instance value = 1, the value MUST be "none".
   * If any of these conditions are not met, the Chain Validation Status is "fail", and the algorithm stops here.

4. Validate the AMS with the greatest instance value (most recent). If validation fails, then the Chain Validation Status is "fail", and the algorithm stops here.
5. _OPTIONAL_: Determine the "oldest-pass" value from the ARC Set by validating each prior AMS beginning with N-1 and proceeding in decreasing order to the AMS with the instance value of 1:

A. If an AMS fails to validate (for instance value "M"), then set the oldest-pass value to the lowest AMS instance value that passed (M+1), and go to the next step (there is no need to check any other (older) AMS header fields). This does not affect the validity of the Authenticated Received Chain.

B. If all AMS header fields verify, set the oldest-pass value to zero (0).

6. Validate each AS beginning with the greatest instance value and proceeding in decreasing order to the AS with the instance value of 1. If any AS fails to validate, the Chain Validation Status is "fail", and the algorithm stops here.

7. If the algorithm reaches this step, then the Chain Validation Status is "pass", and the algorithm is complete.

The end result of this validation algorithm SHOULD be included within the Authentication-Results header field for the ADMD.

As with a DKIM signature ([RFC6376], Section 6.3) that fails verification, a message with an Authenticated Received Chain with a Chain Validation Status of "fail" MUST be treated the same as a message with no Authenticated Received Chain.

_INFORMATIONAL_: Recipients of an invalid or failing Authenticated Received Chain can use that information as part of a wider handling context. ARC adoption cannot be assumed by intermediaries; many intermediaries will continue to modify messages without adding ARC seals.

5.2.1. All Failures Are Permanent

Authenticated Received Chains represent the traversal of messages through one or more intermediaries. All errors, including DNS failures, become unrecoverable and are considered permanent.

Any error validating an Authenticated Received Chain results in a Chain Validation Status of "fail". For further discussion of this topic and the design restriction that prevents chain continuation or re-establishment, see [ARC-USAGE].
5.2.2. Responding to ARC Validation Failures during the SMTP Transaction

If an ARC Validator determines that the incoming message fails ARC validation, the Validator MAY signal the breakage through the extended SMTP response code 5.7.29 ("ARC validation failure") and the corresponding SMTP basic response code. Because ARC failures are likely only to be detected in the context of other underlying authentication mechanism failures, Validators MAY use the more general 5.7.26 ("Multiple authentication checks failed") instead of the ARC-specific code.

6. Communication of Validation Results

Chain Validation Status (described in Section 4.4) is communicated via Authentication-Results (and AAR) header fields using the authentication method "arc". This authentication method is described in Section 10.1.

If necessary data is available, the ptypes and properties defined in Section 10.2 SHOULD be recorded in an Authentication-Results header field:

- smtp.remote-ip - The address of the connection-initiating SMTP server, from which the message is being relayed.
- header.oldest-pass - The instance number of the oldest AMS that still validates, or 0 if all pass.

7. Use Cases

This section explores several message handling use cases that are addressed by ARC.

7.1. Communicate Authentication Assessment across Trust Boundaries

When an intermediary ADMD adds an ARC Set to a message’s Authenticated Received Chain (or creates the initial ARC Set), the ADMD communicates its authentication assessment to the next ARC-participating ADMD in the message-handling path.

If ARC-enabled ADMDs are trusted, Authenticated Received Chains can be used to bridge administrative boundaries.
7.1.1. Message-Scanning Services

Message services are available to perform anti-spam, anti-malware, and anti-phishing scanning. Such services typically remove malicious content, replace HTTP links in messages with sanitized links, and/or attach footers to messages advertising the abilities of the message-scanning service. These modifications almost always break signature-based authentication (such as DKIM).

Scanning services typically require clients to point MX records of an Internet domain to the scanning service. Messages destined for the Internet domain are initially delivered to the scanning service. Once scanning is performed, messages are then routed to the client’s own mail-handling infrastructure. Rerouting messages in this way almost always breaks path-based authentication (such as SPF).

Message-scanning services can attach Authenticated Received Chains to messages to communicate authentication assessment into client ADMDs. Clients can then benefit from the message-scanning service while processing messages as if the client’s infrastructure were the original destination of the Internet domain’s MX record.

7.1.2. Multi-tier MTA Processing

A large message-processing infrastructure is often divided into several processing tiers that can break authentication information between tiers. For example, a large site may maintain a cluster of MTAs dedicated to connection handling and enforcement of IP-based reputation filtering. A secondary cluster of MTAs may be dedicated and optimized for content-based processing of messages.

Authenticated Received Chains can be used to communicate authentication assessment between processing tiers.

7.1.3. Mailing Lists

Mailing lists take delivery of messages and repost them to subscribers. A full description of authentication-related mailing list issues can be found in [RFC7960], Section 3.2.3.

Mailing list services can implement ARC to convey the authentication assessment of posted messages sent to the list’s subscriber base. The ADMDs of the mailing list subscribers can then use the Authenticated Received Chain to determine the authentication assessment of the original message before mailing list handling.
7.2. Inform Message Disposition Decisions

Intermediaries often break authentication through content modification, interfere with path-based authentication (such as SPF), and strip authentication results (if an MTA removes Authentication-Results header fields).

Authenticated Received Chains allow ARC Validators to:

1. identify ARC-enabled ADMDs that break authentication while processing messages and

2. gain extended visibility into the authentication-preserving abilities of ADMDs that relay messages into ARC-enabled ADMDs.

Through the collection of ARC-related data, an ADMD can identify handling paths that have broken authentication.

An Authenticated Received Chain allows an Internet Mail Handler to potentially base decisions of message disposition on authentication assessments provided by different ADMDs.

7.2.1. DMARC Local Policy Overrides

DMARC introduces a policy model where Domain Owners can request email receivers to reject or quarantine messages that fail DMARC alignment. Interoperability issues between DMARC and indirect email flows are documented in [RFC7960].

Authenticated Received Chains allow DMARC processors to consider authentication assessments provided by other ADMDs. As a matter of local policy, a DMARC processor MAY choose to accept the authentication assessments provided by an Authenticated Received Chain when determining if a message is DMARC compliant.

When an Authenticated Received Chain is used to determine message disposition, the DMARC processor can communicate this local policy decision to Domain Owners as described in Section 7.2.2.
7.2.2.  DMARC Reporting

DMARC-enabled receivers indicate when ARC validation influences
DMARC-related local policy decisions.  When an ARC-enabled handler
generates a DMARC report, it MAY indicate the influence of ARC on
their local policy decision(s) by adding a reason of "local_policy"
with a comment string (per [RFC7489], Appendix C) containing a list
of data discovered during ARC validation, which at a minimum
includes:

- the Chain Validation Status,
- the domain and selector for each AS, and
- the originating IP address from the first ARC Set.

EXAMPLE:

```xml
<policy_evaluated>
  <disposition>none</disposition>
  <dkim>fail</dkim>
  <spf>fail</spf>
  <reason>
    <type>local_policy</type>
    <comment>arc=pass as[2].d=d2.example as[2].s=s2
                 as[1].d=d1.example as[1].s=s3
                 remote-ip[1]=2001:DB8::1A</comment>
  </reason>
</policy_evaluated>
```

In the example DMARC XML reporting fragment above, data relating to
specific validated ARC Sets are enumerated using array syntax (e.g.,
"as[2]" means an AS header field with an instance value of 2).
d2.example is the sealing domain for ARC Set #2 (i=2), and d1.example
is the sealing domain for ARC Set #1 (i=1).

Depending on the reporting practices of intermediate message
handlers, Domain Owners may receive multiple DMARC reports for a
single message.  Receivers of DMARC reports should be aware of this
behavior and make the necessary accommodations.

8.  Privacy Considerations

The Authenticated Received Chain provides a verifiable record of the
handlers for a message.  This record may include personally
identifiable information such as an IP address(es) and domain names.
Such information is also included in existing non-ARC-related header
fields such as the "Received" header fields.
9. Security Considerations

The Security Considerations of [RFC6376] and [RFC8601] apply directly to this specification.

As with other domain-based authentication technologies (such as SPF, DKIM, and DMARC), ARC makes no claims about the semantic content of messages. A received message with a validated ARC Chain provides evidence (at instance N) that:

1. the sealing domain (ARC-Seal[N] d=) emitted the message with this body,
2. the authentication assessment reported in the ARC-Authentication-Results was determined upon receipt of the corresponding message at the sealing domain, and
3. the preceding ARC Chain (1..N-1) (with the validation status as reported in the cv field) existed on the message that was received and assessed.

9.1. Increased Header Field Size

Inclusion of Authenticated Received Chains into messages may cause issues for older or constrained MTAs due to increased total header field size. Large header field blocks, in general, may cause failures to deliver or other outage scenarios for such MTAs. ARC itself would not cause problems.

9.2. DNS Operations

The validation of an Authenticated Received Chain composed of N ARC Sets can require up to 2*N DNS queries (not including any DNS redirection mechanisms that can increase the total number of queries). This leads to two considerations:

1. An attacker can send a message to an ARC participant with a concocted sequence of ARC Sets bearing the domains of intended victims, and all of them will be queried by the participant until a failure is discovered. DNS caching and the difficulty of forging the signature values should limit the extent of this load to domains under control of the attacker. Query traffic pattern analysis may expose information about a downstream validating ADMD infrastructure.
2. DKIM only performs one DNS query per signature, while ARC can introduce many (per chain). Absent caching, slow DNS responses can cause SMTP timeouts and backlogged delivery queues on validating systems. This could be exploited as a DoS attack.

9.3. Message Content Suspicion

Recipients are cautioned to treat messages bearing Authenticated Received Chains with the same suspicion applied to all other messages. This includes appropriate content scanning and other checks for potentially malicious content.

ARC authenticates the identity of some email-handling actors. It does not make any assessment of their trustworthiness.

Just as passing message authentication is not an indication of message safety, forwarding that information through the mechanism of ARC is also not an indication of message safety. Even if all ARC-enabled ADMDs are trusted, ADMDs may have become compromised, may miss unsafe content, or may not properly authenticate messages.

9.4. Message Sealer Suspicion

Recipients are cautioned to treat every Sealer of the ARC Chain with suspicion. Just as with a validated DKIM signature, responsibility for message handling is attributed to the sealing domain, but whether or not that Sealer is a malicious actor is out of scope of the authentication mechanism. Since ARC aids message delivery in the event of an authentication failure, ARC Sealers should be treated with suspicion, so that a malicious actor cannot seal spam or other fraudulent messages to aid their delivery, too.

9.5. Replay Attacks

Since ARC inherits heavily from DKIM, it has similar attack vectors. In particular, the replay attack described in [RFC6376], Section 8.6 is potentially amplified by ARC’s chained statuses. In an ARC replay attack, a malicious actor would take an intact and passing ARC Chain and resend it to many recipients without making any modifications that invalidate the latest AMS or AS. The impact to a receiver would be more DNS lookups and signature evaluations. The scope of this attack can be limited by caching DNS queries and following the same signing scope guidance from [RFC6376], Section 5.4.1.
10. IANA Considerations

This document defines one new authentication method and several status codes (Section 10.1), new ptypes and properties (Section 10.2), three new headers fields (Section 10.3), and a new enumerated status code (Section 10.4).

10.1. Update to Email Authentication Result Names Registry

Per this document, IANA has added one authentication method with three codes to the IANA "Email Authentication Result Names" registry:

- **Auth Method**: arc
  - **Code**: "none", "pass", "fail"
  - **Specification**: RFC 8617, Section 4.4
  - **Status**: active

10.2. Update to Email Authentication Methods Registry

Per this document, IANA has added the following to the "Email Authentication Methods" registry, which is defined in [RFC8601]:

- **Method**: arc
  - **Definition**: RFC 8617, Section 6
  - **ptype**: smtp
  - **Property**: remote-ip
  - **Value**: IP address (v4 or v6) of originating SMTP connection
  - **Status**: active
  - **Version**: 1

- **Method**: arc
  - **Definition**: RFC 8617, Section 6
  - **ptype**: header
  - **Property**: oldest-pass
  - **Value**: The instance id of the oldest validating AMS or 0 if they all pass (see Section 5.2)
  - **Status**: active
  - **Version**: 1
10.3. New Header Fields in Permanent Message Header Field Registry

Per this document, IANA has added the following three new header fields to the "Permanent Message Header Field Names" registry:

- Header field name: ARC-Seal
  Applicable protocol: mail
  Status: experimental
  Author/Change controller: IETF
  Specification document(s): RFC 8617
  Related information: RFC 6376

- Header field name: ARC-Message-Signature
  Applicable protocol: mail
  Status: experimental
  Author/Change controller: IETF
  Specification document(s): RFC 8617
  Related information: RFC 6376

- Header field name: ARC-Authentication-Results
  Applicable protocol: mail
  Status: experimental
  Author/Change controller: IETF
  Specification document(s): RFC 8617
  Related information: RFC 8601

10.4. New Status Code in Enumerated Status Codes Registry

Per this document, IANA has added the following value to the "Enumerated Status Codes" registry:

- Code: X.7.29
  Sample Text: ARC validation failure
  Associated basic status code: 550
  Description: This status code may be returned when a message fails ARC validation.
  Reference: RFC 8617
  Submitter: K. Andersen
  Change controller: IESG
11. Experimental Considerations

The ARC protocol is designed to address common interoperability issues introduced by intermediate message handlers. Interoperability issues are described in [RFC6377] and [RFC7960].

As the ARC protocol is implemented by Internet Mail Handlers over time, the following should be evaluated in order to determine the success of the protocol in accomplishing the intended benefits.

11.1. Success Consideration

In an attempt to deliver legitimate messages that users desire, many receivers use heuristic-based methods to identify messages that arrive via indirect delivery paths.

ARC will be a success if the presence of Authenticated Received Chains allows for improved decision making when processing legitimate messages, specifically resulting in equal or better delivery rates than achieved through the use of heuristic approaches.

11.2. Failure Considerations

ARC should function without introducing significant new vectors for abuse (see Section 9). If unforeseen vectors are enabled by ARC, this protocol will be a failure. Note that the weaknesses inherent in the mail protocols ARC is built upon (such as DKIM replay attacks and other known issues) are not new vectors that can be attributed to this specification.

11.3. Open Questions

The following open questions are academic and have no clear answer at the time this document was published. However, additional deployments should be able to gather the necessary data to answer some or all of them.

11.3.1. Value of the ARC-Seal (AS) Header Field

Data should be collected to show if the AS provides value beyond the AMS for either making delivery decisions or catching malicious actors trying to craft or replay malicious chains.
11.3.2. Usage and/or Signals from Multiple Selectors and/or Domains in ARC Sets

Any selectors and/or (sub)domains (under the control of the sealing ADMD) may be used for ARC header field signatures.

While implementers may choose to use various selectors and/or domains for ARC Set header fields, no compelling argument for or against such usage has been made within the working group. As such, we have chosen to allow maximum freedom for the experimental definition of this protocol.

Wider deployment experience and higher volumes of traffic may show whether this is useful.

11.3.3. DNS Overhead

Longer Authenticated Received Chains will require more queries to retrieve the keys for validating the chain. While this is not believed to be a security issue (see Section 9.2), it is unclear how much overhead will truly be added. This is similar to some of the initial processing and query load concerns that were debated at the time of the DKIM specification development.

Data should be collected to better understand usable length and distribution of lengths found in valid Authenticated Received Chains along with the DNS impact of processing Authenticated Received Chains.

An effective operational maximum will have to be developed through deployment experience in the field.

11.3.4. What Trace Information Is Valuable?

There are several edge cases where the information in the AAR can make the difference between message delivery or rejection. For example, if there is a well-known mailing list that seals with ARC but doesn’t do its own initial DMARC enforcement, an Internet Mail Handler with this knowledge could make a delivery decision based upon the authentication information it sees in the corresponding AAR header field.

Certain trace information in the AAR is useful/necessary in the construction of DMARC reports.
Further, certain receivers believe the entire set of trace information would be valuable to feed into machine learning systems to identify fraud and/or provide other signals related to message delivery.

At this point, however, it is unclear what trace information will be valuable for all receivers, regardless of size.

Data should be collected on what trace information receivers are using that provides useful signals that affect deliverability and what portions of the trace data are left untouched or provide no useful information.

Since many such systems are intentionally proprietary or confidential to prevent gaming by abusers, it may not be viable to reliably answer this particular question. The evolving nature of attacks can also shift the landscape of "useful" information over time.

12. References

12.1. Normative References


12.2. Informative References


[ARC-USAGE] Jones, S., Ed. and K. Andersen, "Recommended Usage of the Authenticated Received Chain (ARC)", Work in Progress, draft-ietf-dmarc-arc-usage-07, April 2019.

Appendix A. Design Requirements

The specification of the ARC framework is driven by the following high-level goals, security considerations, and practical operational requirements.

A.1. Primary Design Criteria

- Provide a verifiable "chain of custody" for email messages;
- Not require changes for originators of email;
- Support the verification of the ARC header field set by each hop in the handling chain;
- Work at Internet scale; and
- Provide a trustable mechanism for the communication of Authentication-Results across trust boundaries.

A.2. Out of Scope

ARC is not a trust framework. Users of the ARC header fields are cautioned against making unsubstantiated conclusions when encountering a "broken" ARC sequence.
Appendix B. Example Usage

The following message is an example of one that has passed through several intermediary handlers, some of which have modified the message and others which have not:

Return-Path: <jqd@d1.example>
Received: from example.org (example.org [208.69.40.157])
for <fmartin@example.com>; Thu, 14 Jan 2015 15:02:40 -0800 (PST)
Received: from segv.d1.example (segv.d1.example [72.52.75.15])
by lists.example.org (8.14.5/8.14.5) with ESMTP id tOEkaNU9010123
for <arc@example.org>; Thu, 14 Jan 2015 15:01:30 -0800 (PST)
(envelope-from jqd@d1.example)
Received: from [2001:DB8::1A] (w-x-y-z.dsl.static.isp.example [w.x.y.z])
(authenticated bits=0)
by segv.d1.example with ESMTP id t0FN4a8084569;
Thu, 14 Jan 2015 15:00:01 -0800 (PST)
(envelope-from jqd@d1.example)
Received: from mail-ob0-f188.google.example
(mail-ob0-f188.google.example [208.69.40.157]) by
clochette.example.org with ESMTP id d200mr22663000ykbb.93.1421363268
for <fmartin@example.org>; Thu, 14 Jan 2015 15:03:15 -0800 (PST)
ARC-Seal: i=3; a=rsa-sha256; cv=pass; d=clochette.example.org; s=
clochette; t=12345; b=CU87XzX1NhK5X/yW4173UvPUco9ivwYxWyWBVrRo7
+HPx3K0snJuy2fvyRfReAmOA9GTH/y+k9keC59hAKVg==
ARC-Message-Signature: i=3; a=rsa-sha256; c=relaxed/relaxed; d=
clochette.example.org; h=message-id:date:from:to:subject; s=
clochette; t=12345; bh=KWSe46TZKCCDH4k1Jpof+tj5LWJnVR1p5pvjXFZY
LQs; b=071wvysLoK+wM4cOSlixr0XOWzEv1vqIjd/2/GkYFY1SD/GOFkZkAgPqxf
K7ccBMP7Zjb/pegswphjEMS8x5NQ==
ARC-Authentication-Results: i=3; clochette.example.org; spf=fail
smtp.from=jqd@d1.example; dkim=fail (512-bit key)
header.i=@d1.example; dmarc=fail; arc=pass (as.2.gmail.example=pass,
ams.2.gmail.example=pass, as.1.lists.example.org=pass,
ams.1.lists.example.org=pass (message has been altered))
Authentication-Results: clochette.example.org; spf=fail
smtp.from=jqd@d1.example; dkim=fail (512-bit key)
header.i=@d1.example; dmarc=fail; arc=pass (as.2.gmail.example=pass,
ams.2.gmail.example=pass, as.1.lists.example.org=pass,
ams.1.lists.example.org=fail (message has been altered))
ARC-Seal: i=2; a=rsa-sha256; cv=pass; d=gmail.example; s=20120806; t=
12345; b=Zpukh/kJLq47Kv391FkwTepgS56dgHicdhhZjOsalh3qFtQQAJ4T9BE
8ujLWXpRnHuh8lyqo81/TjHN8RweyzGw==
ARC-Message-Signature: i=2; a=rsa-sha256; c=relaxed/relaxed; d=
gmail.example; h=message-id:date; s=20120806; t=
12345; bh=KWSe46TZKCCDH4k1Jpof+tj5LWJnVR1p5pvjXFZYLQs; b=CVoG44
cVz0o2z2mMig2zwqPaJ40ZSS5XGMcgwQwQslwvRZJS894tJM0x01RJgLCPs0x1a5

Andersen, et al.             Experimental         [Page 33]
Hey gang,
This is a test message.
--J.

Andersen, et al.  Experimental  [Page 34]
Acknowledgments

This document originated with the work of OAR-Dev Group.

The authors thank all of the OAR-Dev and the subsequent DMARC WG for the ongoing help and thought-provoking discussions from all the participants, especially J. Trent Adams, Marc Bradshaw, Alex Brotman, Greg Colburn, Dave Crocker, Tim Draegen, Mark Eissler, Peter Goldstein, Bron Gondwana, Mike Hammer, Mike Jones, Steve Jones, Scott Kitterman, Barry Leiba, Franck Martin, John Rae-Grant, Paul Rock, Gene Shuman, Terry Zink, and Elizabeth Zwicky.

Grateful appreciation is extended to the people who provided feedback through the arc-discuss mailing list.

Authors’ Addresses

Kurt Andersen
LinkedIn
1000 West Maude Ave
Sunnyvale, California 94085
United States of America

Email: kurt+ietf@drkurt.com

Brandon Long (editor)
Google

Email: blong@google.com

Seth Blank (editor)
Valimail

Email: seth@valimail.com

Murray Kucherawy (editor)
TDP

Email: superuser@gmail.com