

Viewpoint™ ETA Suppression

Application Version: 7.0
August 19, 2026

© 2021 Conduent, Inc. All rights reserved. Conduent and Conduent Agile Star are trademarks of Conduent, Inc. and/or its subsidiaries in the United States and/or other countries.

Other company trademarks are also acknowledged.

Document Version: 7.0 (August 2021).

Document Conventions

Convention	Explanation
Bold	For file names, commands, fields, menus, options, and window names.
<code>Courier New</code>	Commands as you should type them.
<code>Lucida Console</code>	Example output generated by the system.
<i>Italics</i>	For configuration variables, including variable portions of file names and URLs. Also indicates a document name.
Note / Blue Callout	The Blue Callout text indicates information that is of special interest or importance, an idea that could be useful or additional information about a product or a feature.
	The Caution icon along with the text indicates actions that can lead to problems in system operation or configuration settings if the instructions are not followed properly.

Document Revision History

This section tracks the initial creation of the document after each major version thereafter.

Ver:	Date	Description	Reviewed / Approved By
1.0	Oct 16, 2012	Initial Version	Team
1.1	Feb 10, 2017	Added Rule 6 (BCC)	VP Dev
1.2	August 10, 2021	Rebranded adhering to the latest Conduent brand central documentation standards/guidelines.	Technical Writer

Table of Contents

1. ETA Suppression.....	1-1
1.1. Basic Example.....	1-1
1.2. Terminology.....	1-1
1.2.1. Email Cluster.....	1-1
1.2.2. Homogenous Email Cluster.....	1-1
1.2.3. Attachment Family.....	1-1
1.2.4. Message ID.....	1-1
1.2.5. Reference Chain.....	1-1
1.2.6. Missing Email Cluster.....	1-1
1.2.7. Leaf.....	1-2
1.2.8. Representative Document.....	1-2
1.3. Documents vs. Clusters.....	1-2
1.4. Views and Scopes.....	1-2
1.5. Rules.....	1-2
1.5.1. Rules 1 and 2 in Detail.....	1-2
1.5.2. Rule 3 in Detail.....	1-3
1.5.3. Rules 4 and 5 in Detail.....	1-3
1.6. Examples.....	1-3
1.7. Technical Implementation.....	1-5

1. ETA Suppression

ETA Suppression is a Viewpoint feature that culls the emails quoted by other emails, reducing the total number of documents required for review. When an email document is considered unique, its family (attachments, etc.) are also considered unique.

1.1. Basic Example

Alice sends an email to Bob asking “Do you want to go to lunch?”

Bob replies to Alice’s email as “Sure.”

This yields two emails – one from Alice to Bob, and one from Bob to Alice. The text in the second email entirely quotes the first email. As such, ETA suppression would consider the first email redundant and the second email unique. A View built from this tree with ETA Suppression in effect would contain only the second (unique) email, along with any other documents which were not ETA documents (not emails, or at least not emails that had ETA run against them).

1.2. Terminology

1.2.1. Email Cluster

A collection of documents that represent an email message. Often email clusters are made up of multiple copies of the same document.

For example, if an email is sent to two people, then there are three documents representing this cluster; the two that are in the inbox of the recipients and one in the outbox of the sender.

1.2.2. Homogenous Email Cluster

A cluster is homogenous if all of the documents in the cluster are the same or nearly the same. Nearly the same means that they have the same text, however, some flexibility is allowed here – some email systems add extra whitespace or line-breaks, and some systems add extra HTML tags around links. If an email cluster contains two (or more) documents that are not reasonably the same document, then the cluster is not homogenous.

1.2.3. Attachment Family

It consists of a set of attachments.

1.2.4. Message ID

Many, but not all email messages contain a unique Message-ID. The documents with the same Message-ID are generally considered as part of the same email cluster.

1.2.5. Reference Chain

Some email messages contain a reference chain, which is a catalog of the message IDs of each of the email’s parent emails.

1.2.6. Missing Email Cluster

An email cluster that does not contain any documents. These clusters are generated when an email cluster is inferred by a reference chain entry, but none of the actual documents exist.

Another way to think about a Missing Email Cluster is to consider it a placeholder for an email that is believed to exist, but no document has been processed to populate it.

1.2.7. Leaf

A leaf is an email cluster that has no children.

1.2.8. Representative Document

A document contains the information that represents a cluster. A cluster is often made up of multiple documents. The representative document for a cluster is any document in the cluster, provided that the cluster is homogenous.

Note: ETA Suppression implicitly suppresses the true duplicates as they should be a 100% match of other documents in their respective clusters.

1.3. Documents vs. Clusters

A Cluster can have any number of Documents, including zero. Generally, a cluster is made of two or more documents; a copy of the email from the sender's outbox and a copy of the email from the recipient's inbox. As mentioned in the terminology section, two documents in a cluster may not necessarily be identical. They may be nearly identical (homogenous) or not (non-homogenous). ETA Suppression suppresses the **documents**, not the **clusters**. The difference is subtle but pivotal.

1.4. Views and Scopes

ETA Suppression is performed over a source, such as a View, set of Views or a set of Scopes. It is not necessarily true that an entire email tree will fall within the source on which you are performing ETA suppression. Consequently, terms like 'leaf' and 'parent' are relevant only to the documents within the source. Taking the initial example of the emails between Alice and Bob, if ETA suppression runs over a view that contained Alice's initial email to Bob but not Bob's response, then Alice's email would be considered unique, because it is not quoted in any child email (because Bob's response is outside of the source and therefore 'invisible' to ETA suppression relative to this source).

1.5. Rules

A document is considered unique and therefore not suppressed based on the following rules:

- **Rule 1:** When the document represents a leaf cluster.
- **Rule 2:** When the document has child clusters but no child cluster fully quotes its text.
- **Rule 3:** When the document contains an attachment family which is not represented in any child cluster.
- **Rule 4:** When the document represents a distinct hash code within a non-homogenous cluster.
- **Rule 5:** When the document has children and none of the children are homogenous.

A document is either unique or not unique. Documents that are not unique will be suppressed. Suppression never breaks families, meaning, that if at least one document in a family is considered unique, then the entire family will be spared from suppression.

1.5.1. Rules 1 and 2 in Detail

The predominant rules in determining uniqueness are Rules 1 and 2. Rules 1 and 2 are concerned with establishing the endpoints of a thread of conversation. In the case of Rule 1, if no cluster quotes a particular cluster, then that cluster represents the end of a thread, and therefore unique

information (no other cluster quotes it). Similarly, Rule 2 looks for leaves of a different sort – Rule 2 seeks out cases where a quoted email body has been altered in some way. This means that an individual has redacted information from an email before forwarding it or manipulated the content in some other fashion. This cluster represents the start of a new thread and its parent (the last cluster in the parentage chain which contains the unadulterated content) is unique unless it is quoted by some other child which has not adulterated the parent's content.

Some clusters contain no documents at all (Missing Email Clusters). When considering Rules 1 and 2 these clusters are ignored as the documents that represent the cluster are necessarily outside of the documents that make up the source. That is to say that no document exists for the cluster, then no documents representing that cluster can be in the source.

1.5.2. Rule 3 in Detail

The presence of an attachment in an email will generally make the cluster's representative document unique; however, this is not always true. If in our Alice / Bob example, Alice attached an image file and when Bob responded he included the same image file, then Alice's initial email is not unique; the attachment is present in Bob's response email. However, if Bob did not quote the attachment, then both Alice's and Bob's emails are unique because both represent unique content. Alice's is unique because it is the only time that the attachment is represented in the thread, and the second email because of **Rule 1**.

An attachment family is a set of files attached to an email document. Using the Alice / Bob example again, if Alice attached two image files and Bob only quoted one of them in his response then both emails are considered unique, because although Bob's email represents some of what Alice sent it does not represent everything. Conversely, if Alice attaches an image file, and in Bob's response he includes Alice's attachment and also adds a file, then Alice's original email is not unique. It is not unique because all of her content is represented in a child email.

1.5.3. Rules 4 and 5 in Detail

When Rules 4 and 5 apply to a cluster, ETA has no way of knowing which of the texts representing the cluster is the "right" one. Consequently, ETA considers all representative documents in the cluster unique, along with the parent email cluster if the parent does not have any other non-homogenous children. This condition does not occur very often but has been known to happen.

1.6. Examples

The following examples show the results of applying the rules of ETA suppression on a series of email trees. For these examples, consider all documents in the tree to be in the source. Missing Email Clusters are not represented as they do not affect ETA suppression.

Example:

```
A
+-B(X)* Rule 1, 3
```

A is the parent cluster.

B is a child cluster (the response to, or forward of, cluster A).

B has an attachment – file "X".

'B' is considered unique because of Rules 1 and 3. (B would not be suppressed)

A is not unique. (A will be suppressed)

Case 1 – No attachments:

```

A
+-B
| +-C * Rule 1
| +-D * Rule 1
+-E
  +-F * Rule 1

```

Case 2 – Attachments are not unique:

```

A
+-B(X)
| +-C (X) * Rule 1, 3
| +-D * Rule 1
+-E
  +-F * Rule 1

```

Case 3 – Attachments are unique:

```

A
+-B (X) * Rule 3
| +-C * Rule 1
| +-D * Rule 1
+-E
  +-F * Rule 1

```

Case 4.1 – Attachments Family 1

```

A (X)
+-B (X,Y) * Rule 3
| +-C * Rule 1
| +-D * Rule 1
+-E
  +-F * Rule 1

```

Case 4.2 – Attachment Family 2

```

A (X, Y)
+-B (X,Y)
| +-C (X,Y) * Rule 1, 3
| +-D * Rule 1

```

```

+-E (X) * Rule 3
+-F * Rule 1

```

Case 4.3 – Attachment Family 3

```

A (X)
+-B (X,Y) * Rule 3
| +-C (X) * Rule 1
| +-D (Y) * Rule 1, 3
+-E
+-F(X) * Rule 1 ,3

```

Case 5 – Adulterated Content

```

A * Rule 2
+-B – the text of A was removed * Rule 1

```

Case 6 – Non-homogenous

```

A * Rule 5
+-B – two completely different docs make up this cluster. * Rule 4

```

In this case, both documents in the cluster are considered unique, not just one representative document.

1.7. Technical Implementation

To determine ETA suppression for a scope, perform the following steps:

1. Establish the set of clusters that contain the documents that are within the scope.
2. Apply Rule 4 – check each cluster in scope for non-homogeneity and note the representative documents as unique (even if only one document is in the source).
3. Apply Rule 5 – for every cluster that was not marked unique previously, find clusters that have only children that are eventually non-homogenous. The documents which are outside of the source are considered for this rule.
4. Apply Rules 1 and 2 - mark all remaining Leaf Clusters as unique.
5. Apply Rule 3 – mark as unique clusters which contain attachment families that are not already represented in a cluster that has been considered unique.