DOIC Status and Issues

IETF #89

Summary

- 36 Issues Opened
- 6 closed in issue tracker
- 18 resolved on list
- 12 open in various stages of resolution

Issues for discussion

- Realm-Routed-Request and Realm report types
- Behavior of Agent acting as reacting node
- Host report scope (All, single)
- OC-Supported-Features AVP handling
- OC-OLR AVP handling
- Definition of Overload Control Endpoints

# Realm-Routed-Request and Realm report types

- Proposal to support three report types:
  - Host – Throttle requests targeted to a specific Destination-Host
  - Realm-Routed-Request – Throttle requests targeted to a realm with no Destination-Host AVP specified
  - Realm – Throttle requests targeted to a realm

Use Case for Realm scoped reports Partitioned servers

Applies for both Realm-Routed-Request and proposed Realm reports
Question: Who knows the Realm wide scope?
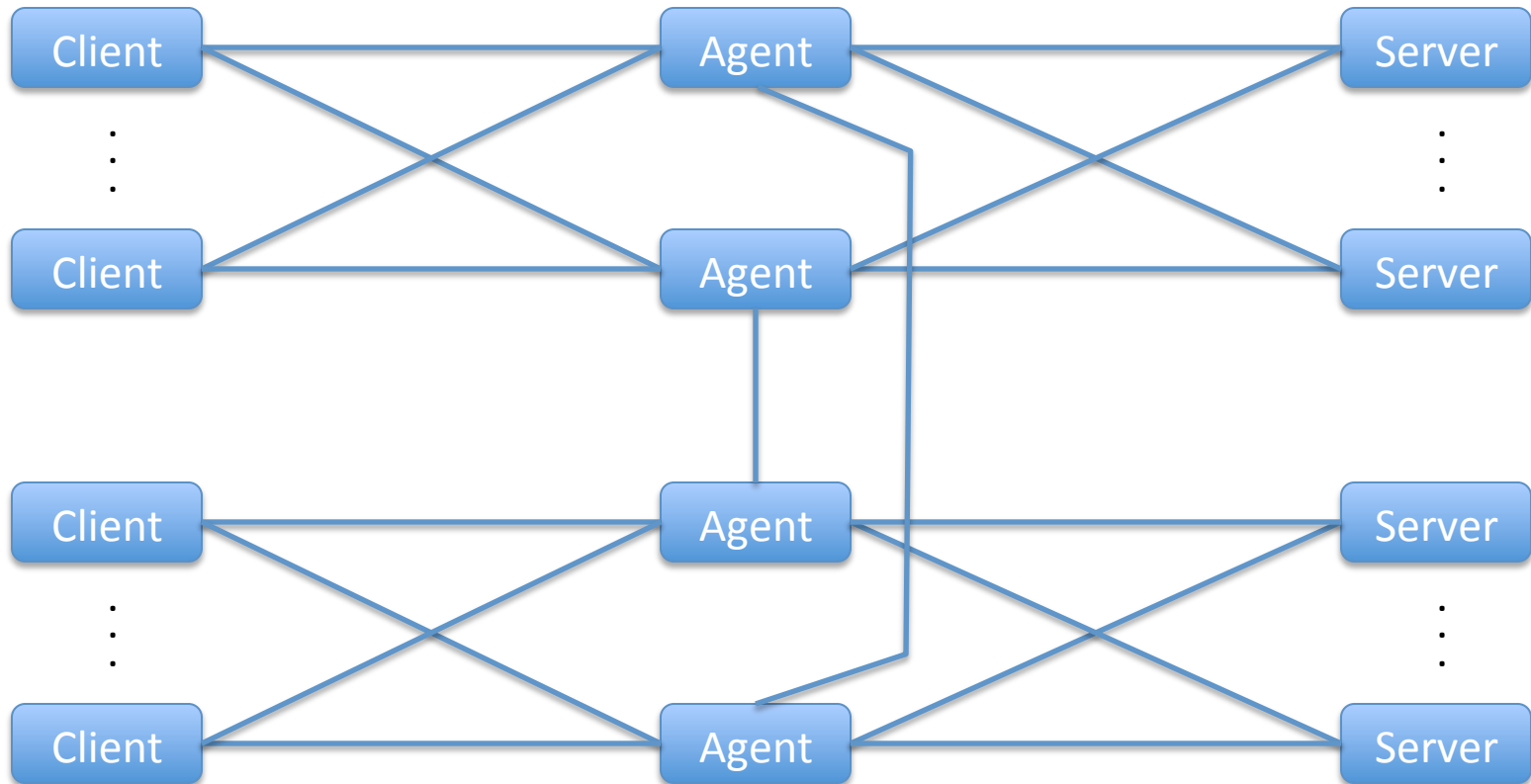
Issue for Realm scoped reports

- No one Diameter entity can determine realm wide scope based on observing DOIC OLRs

Behavior proposal for realm-scoped reports

- Reporting nodes
  - Must have knowledge of realm overload state (method is out of scope)
  - Send realm scoped reports when appropiate
  - Other behavior is the same as for host reports
  - Multiple reporting nodes do not require synchronization of sequence numbers
- Reacting nodes
  - Apply first received Realm-scoped report, remembering the originator of the report
  - For as long as there is a valid Realm-scoped report, ignore reports from other sources
  - Other behavior around managing the status of the report is the same as for host reports

Behavior of Agent acting as reacting node

- Issues
  - What error response does agent send when throttling requests
  - 3002 Too-Busy doesn't appear to work for Realm Routed Requests
  - Too-Busy behavior can be updated in a new RFC but that doesn't help for nodes that only support 6733

Host report scope (All, single)

- General agreement that server should be able to send an overload report that is global and be able to send overload reports that are specific to a single reacting node

- Proposal – Add AVP to OC-OLR that indicates the scope of the report
  - Type 0 Overload report applies to single reacting node
  - Type 1 Overload report applies to all reacting nodes
  - Reacting node uses the most specific report

OC-Supported-Features Handling

- Question – Does reporting node include single algorithm or all supported algorithms
- Proposal –
  - Reacting node advertises all supported algorithms
  - Reporting node responds with the single algorithm it will be using
  - Handling of other feature bits is defined in the extension drafts

OC-OLR Handling

- Options:
  - Separate overload report defined per abatement algorithm.
  - Single OC-OLR AVP used by all abatement algorithms with the meaning of the overload report indicated in the OC-Supported-Features AVP.

Definition of Overload Control Endpoints

- Alternatives
  - Hop-by-hop
  - End-to-end with ends defined by context of transaction

- Principles and Assumptions
  - Throttling should be done as close to the source of the Diameter transaction as possible
  - Capability negotiation lifetime is single transaction

End-to-end Definition

- DOIC Endpoints defined by node that inserts the OC-Supported-Features AVP
  - Reacting node is node that inserts it in request
  - Reporting node is node that inserts it in answer
- Questions/Issues:
  - How are end-points identified when agents are one of the end-points?
  - Extensibility - What happens when multiple nodes insert OC-Supported-Features in a request as likely will be required for Agent Overload.

Closed

#	Summary	Status	Notes
#25	Section 3.1.5 Diameter Agent Behavior	Closed	Resolution of #24 addresses this issue.
#28	Report type support in capabilities?	Closed	No change required.
#33	Overload Mitigation Differentiation per Client	Closed	Duplicate
#34	Semantics of OC-Report-Type AVP	Closed	Agreed – Replace text in section 4.6 with text proposed in the ticket.
#47	reduction percentages greater than 100 should be ignored	Closed	Agreed – In Section 4.7, change "treated as is the OC-Validity-Duration AVP was not present" to "Invalid validity duration values are given the default value".
#48	Setting M-Bit gives wrong semantics	Closed	Closed as invalid and no longer an issue based on better understanding of M-bit handling.

Resolved - 1

#	Summary	Status	Notes
#24	Terminology and Abbreviations	Resolved	Agreed – Remove definition of Diameter layer load balancing. Agreed – Remove definition of Topology Hiding. Agreed – Remove discussion of agents acting as topology hiders or server front ends. Agreed – Add wording in agent behavior section to cover the case when an agent is acting as a reporting node for host or realm-routed-request reports. This will be new wording in the draft that will need to be reviewed before is it finalized. This issue cannot be closed until the wording of the agent behavior section is reviewed and agreed to.
#29	OC-Sequence-Number in OC-Supported-Features	Resolved	Agreed – Removed OC-Sequence-Number from OC-Supported-Features AVP. Agreed – The scope of an OC-Supported-Features AVP is a single transaction. Agreed – Diameter nodes that support DOIC must include the OC-Supported-Features AVP in all requests. Update: 140226 – Removed OC-Sequence-Number from OC-Supported Features AVP section.

Resolved - 2

#	Summary	Status	Notes
#31	Sending OC-Ongoing-Throttling-Info AVP in request messages that survived a throttling	Resolved	Agreed – No consensus to add OC-Ongoing-Throttling AVP in request messages surviving throttling. Agreed - Missing OLR in answer means "no change"; it does not mean "no overload/no throttling requested" Agreed - Reporting nodes SHOULD include OLR in every answer that it sends in response to a request message which indicated OLR_DEFAULT_ALGO support unless the reporting node has very good reasons not to do so. Exact wording is not yet agreed.

Resolved - 2

#	Summary	Status	Notes
#32	Sequence-Number Time-Stamp values within OC-OLR	Resolved	Agreed - Sequence-Number is of type Unsigned64. Agreed - When generated, a new sequence number must be greater than the sequence number contained in the active overload report to which it applies (including over reboot of that node). When received, a sequence number is used to detect outdates/replays/freshness. Sequence numbers of expired OLRs MUST NOT be remembered by reacting nodes.
#36	OC-Validity-Duration AVP	Resolved	Agreed – Change based on Lionel's wording in email dated Feb 7, 2014 timestamped at 4:19am US central standard time, as shown in the proposed wording section for #36 below.
#37	Inconsistent name and abbreviation	Resolved	Agreed – Update the document to use the DOIC name and abbreviation. Updated – Change made in abstract, introduction and in specification to change DOC to DOIC where appropriate.

Resolved - 3

#	Summary	Status	Notes
#38	Server Farm Definition Issue	Resolved	Addressed by resolution of issue #24.
#39	Definition of Diameter Routing	Resolved	Addressed by resolution of issue #24.
#42	IETF defined example of a stateless application.	Resolved	Agreed – Add reference to RFC4740 as an IETF defined stateless application.
#43	Overstated guidance on session-ending requests.	Resolved	Agreed – No changes required.
#44	Incorrect sequence number behavior	Resolved	Proposed – Change the last sentence of section 4.3, paragraph 3 to "The reacting node SHOULD discard an OC-OLR AVP with a sequence number that is less than or equal to the previously received sequence number."
#45	Why is a validity duration of 0 disallowed?	Resolved	Proposed – A reporting node communicates that an overload report is no longer valid by sending an OLR with a Validity-Period AVP with a value of zero. This is the only way for a reporting node to indicate that an overload report is no longer valid. For instance, setting the reduction-percentage to zero does not make the overload report invalid.

Resolved - 4

#	Summary	Status	Notes
#46	Bad normative advice on not letting overload reports expire	Resolved	Agreed – Update the language to indicate that it is acceptable to let overload reports time out if the change in overload state at the reporting node expires close to the time that the report would have expired anyway.
#50	OC-OLR AVP implicit info	Resolved	Agreed – Change the wording in section 4.3 as captured on the list.
#51	OC-Supported-Features in requests	Resolved	Agreed – OC-Supported-Feature AVP MUST be included in all request messages sent by a DOIC supporting node.
#52	Throttling not needed to be based on previous history	Resolved	Agreed – Change wording as captured on the list.

Resolved - 5

#	Summary	Status	Notes
#52	Throttling not needed to be based on previous history	Resolved	Agreed – Change wording as captured on the list.
#57	Handling of "Realm-Routed" Overload report type	Resolved	Proposal – No change in meaning of the Realm-Routed-Requests report. Some wording on interaction between host and this report might be needed.

Open - 1

#	Summary	Status	Notes
#23	DOIC behavior for realm overload	Open	Proposal – Change the name of realm report. Proposed name - "Realm-Routed-Request" report. Proposal – Update all discussion on "Realm-Routed-Request" reports to indicate that they apply only to requests targeted to a realm when there is no destination-host AVP in the request.
#26	Overload Control Endpoints under specified	Open	Not much discussion but this overlaps with discussions in other threads. Issue: Definition of Overload Control Endpoints and how they are used.
#27	Behavior of agent acting on behalf of Client that does not support DOIC	Open	Proposal – Agent sends too busy response to throttled requests. Proposal – Agent behavior is captured in new section outlined in by issue #24. This issue cannot be closed until the wording of the agent behavior section is reviewed and agreed to.

Open - 2

#	Summary	Status	Notes
#30	OC-Supported-Features AVP in answer messages	Open	Agreed - Absence of Supported-Features-AVP from an answer message MUST not result in reacting nodes to cease sending of any DOIC related AVPs in subsequent requests. Proposal – Is the OC-Supported-Features AVP must be included in all answer messages originated by a supporting node.
#35	additional report types are proposed	Open	Agreed – There is benefit to allowing a reporting node to specify per Origin-Host OLRs. Proposal – Add AVP to loss report that indicates if report applies to all reacting nodes or to just the reacting node involved in the transaction.
#40	Need defintions for Overload Report and Abatement Algorithm	Open	Need proposed wording
#41	Need better overview	Open	Need proposed wording. Ben has indicated he will provide wording.

Open - 3

#	Summary	Status	Notes
#49	capabilities announcement mechanism needs to be rethought	Open	Agreed: Lifetime of a capabilities announcement is a single transaction. Proposal: OC-Supported-Features AVP must be included in all requests sent by Diameter nodes supporting DOIC. Proposal: OC-Supported-Features AVP must be included in all answers sent by Diameter nodes supporting DOIC. Open: Need behavior definition of responding and reacting nodes.
#53	5.5.2 chapter typo?	Open	Proposal – Remove "If the OC-Supported-Features AVP is received for the first time for the reporting node or the OC- Sequence-Number" from section 5.5.2, paragraph 2, sentence 2.
#54	OC-Report-Type as mandatory AVP	Open	Proposal – Make OC-Report-Type mandatory in all OC-OLR reports.

Open - 4

#	Summary	Status	Notes
#55	Lack of overload control for realm overload condition	Open	Proposal – Add new report type and subsequent behavior definition associated with that report type. Must capture interaction between the realm report type and other report types.
#56	Bad Description of Overload Control State	Open	New issue in early stages of discussion. Interacts with the resolution of issue #35.