Internet Area Working Group
Internet-Draft
Updates: 791, 2003, 2780, 4301, 4727, ietf-intarea-ipv4-id-update (if approved)
Intended status: Standards Track
Expires: September 13, 2012

B. Briscoe
BT

March 12, 2012

Reusing the IPv4 Identification Field in Atomic Packets
draft-briscoe-intarea-ipv4-id-reuse-01

Abstract

This specification takes a new approach to extensibility that is both principled and a hack. It builds on recent moves to formalise the increasingly common practice where fragmentation in IPv4 more closely matches that of IPv6. The large majority of IPv4 packets are now 'atomic', meaning indivisible. In such packets, the 16 bits of the IPv4 Identification (IPv4 ID) field are redundant and could be freed up for the Internet community to put to other uses, at least within the constraints imposed by their original use for reassembly. This specification defines the process for redefining the semantics of these bits. It uses the previously reserved control flag in the IPv4 header to indicate that these 16 bits have new semantics. Great care is taken throughout to ease incremental deployment, even in the presence of middleboxes that incorrectly discard or normalise packets that have the reserved control flag set.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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

This Internet-Draft will expire on September 13, 2012.

Copyright Notice

Table of Contents

Intended status: Standards Track? (to be removed before publication)

This draft defines a process and a protocol for enabling new protocols, including their progression from experimental track to standards track. A process specification cannot have lesser status than the protocols it enables. So if this specification were to start on the experimental track, it would not initially have sufficient status to enable standards track protocols.

In order for the IETF to consider whether this draft itself should be experimental or standards track, it has been written as if it is intended for the standards track. Otherwise the parts of the process for enabling standards track protocols would have had to have been written hypothetically, which would have been highly confusing. If the IETF decides this specification ought to start out on the experimental track, the standards track parts of the process will have to be edited out.

Appendix B discusses whether this draft itself would be better to start as experimental or standards track.

1. Introduction

The Problem: The extensibility provisions in IP (v4 and v6) have turned out not to be usable in practice. Hardware has been optimised for the common case, so packets using extensibility mechanisms (e.g. IPv4 options or IPv6 hop-by-hop options) are very likely to be punted to the software slow-path and consequently likely to be dropped whenever the software processor is busy [Fransson04, Cisco.IPv6Ext].

This specification takes a different approach to extensibility. Rather than flagging protocol extensions as 'extensions', it places extension headers where they will be ignored by pre-existing hardware. As code is added to routers to handle newly added extensions, the code can tell the machine where to look for the relevant header.

This approach recognises that extensions added after a protocol suite was first defined are different to options defined as a coherent part of the original protocol suite. Machines that have no code to understand a protocol extension that was added later do not need to punt a packet to the software processor merely to scan through chains of headers that it will not know how to process.

Having only settled on this approach long after the TCP/IP suite has been defined, it becomes necessary to find places in the existing protocol headers that are already ignored by existing machines. In an 'atomic' IPv4 packet, the Identification (IPv4 ID) field is one

such place that is redundant. This specification defines the process through which the 16 bits in this field can be returned to the IETF for use in future standards actions, at least within the constraints imposed by their original use for reassembly.

Background: [ipv4-id-update] proposes to update IPv4 to more closely match the approach taken to fragmentation in IPv6. It recommends that IPv4 sources send 'atomic' packets whenever possible. An atomic packet is one that has not yet been fragmented (MF=0 and fragment offset=0) and for which further fragmentation is inhibited (DF=1) [ipv4-id-update]. If fragmentation is necessary, it is only permitted at devices that control the uniqueness of the IP ID field, e.g., sources, tunnel ingresses (for the outer header), and the public side of NATs.

In practical scenarios, the IPv4 ID field is too small to guarantee uniqueness during the lifetime of a packet anyway [RFC4963]. Therefore it has become safer to disable fragmentation altogether and instead use an approach such as packetization layer path MTU discovery [RFC4821]. The large majority of IPv4 packets are now atomic.

Approach: This specification defines the IPv4 control flag that was previously reserved [RFC0791] as the Recycled flag (RC). An implementation can set RC=1 in an atomic packet to unambiguously flag that the IPv4 ID field is not to be interpreted as IP Identification, but instead it has the alternative semantics of an ID-Reuse field. By setting RC=1, IPv4 implementations can distinguish a value deliberately written into the ID-Reuse field from the same value that just happened to be written into the IP ID field of an atomic packet by a pre-existing implementation.

Thus, this specification effectively uses up the last bit in the IPv4 header in order to free up 16 other bits. However, there are some constraints on the use of these 16 bits due to their original use as the IP ID field (enumerated in Section 5.1). Of course the main constraint is that the bits are not available in non-atomic packets. But fragmentation is now used only rarely anyway, so it makes sense to see if the the Internet community can invent ways to use the 16 bits in the IPv4 ID field despite the constraints.

Frequently Asked Questions:

1. There are many cases where a non-compliant machine ignores Don't Fragment (DF=1) and fragments a packet anyway.

   One answer is that we cannot allow non-complaint behaviour to always block progress. Another answer is that we may be able to

detect and circumvent such non-compliant behaviour. For instance, if a non-compliant router fragments packets with DF=1, it may be possible to enhance path maximum transmission unit discovery (PMTUD) to find a lower segment size small enough to prevent the offending box from fragmenting packets.

2. {ToDo}

Document Roadmap: Section 3 defines the semantics of the updated IPv4 wire protocol and Section 4 defines intermediate node behaviour. Section 5 defines the process to be used for reassigning sub-fields of the IPv4 ID-Reuse field. Then Section 6 describes a way to circumvent problems likely to arise when deploying this new protocol. Finally, Section 7 enumerates the updates to pre-existing RFCs, before the tailpiece sections considering IANA, Security and draw conclusions.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Further terminology used within this document:

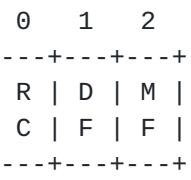
Atomic packet: A packet not yet having been fragmented (MF=0 and fragment offset=0) and for which further fragmentation has been inhibited (DF=1), or in the syntax of the C programming language ((DF==1) && (MF==0) && (Offset==0)) [ipv4-id-update].

Recycled (RC) flag: The control flag that was 'reserved' in [RFC0791] (Figure 1). The flag positioned at bit 48 of the IPv4 header (counting from 0). Alternatively, some would call this bit 0 (counting from 0) of octet 7 (counting from 1) of the IPv4 header.

ID-Reuse field: Octets 5 and 6 (counting from 1) of the IPv4 header of an atomic packet (Figure 3). The field that would have been the IP Identification field if the packet were not atomic.

3. IPv4 Wire Protocol Semantics for Reusing the Identification Field

This specification defines the control flag that was defined as 'reserved' in [RFC0791] as the Recycled (RC) flag (Figure 1).



The Recycled (RC) Flag was previously reserved.

Figure 1: The Control Flags at the Start of Byte 7 of the IPv4 Header

Figure 2 recaps the definitions of octets 5 to 8 (counting from 1) of the IPv4 header [RFC0791].

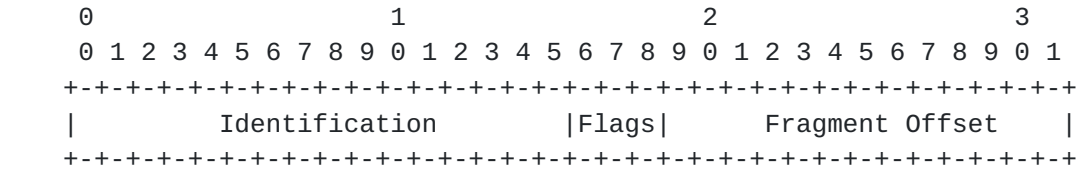
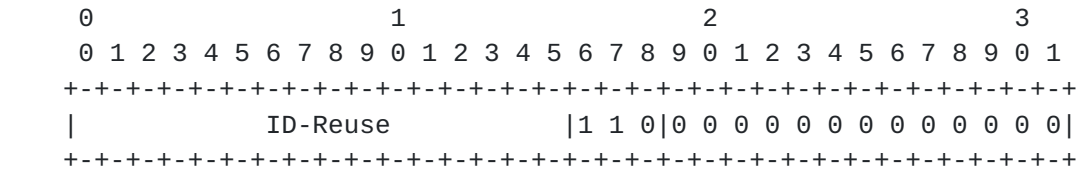


Figure 2: Recap of RFC791 Definition of Octets 5 to 8 of the IPv4 Header.

If an IPv4 implementation sets RC=1 on an atomic packet, octets 5 & 6 of the IPv4 header MUST be interpreted with the semantics of the ID-Reuse field, and MUST NOT be interpreted as the Identification field. Figure 3 shows how octets 5 & 6 are redefined as the ID-Reuse field when the packet is atomic, in the case where RC=1.



The Identification Field is redefined as the ID-Reuse Field when the Packet is Atomic and specifically when RC=1

Figure 3: Octets 5 to 8 of the IPv4 Header.

If the Recycled flag is cleared to RC=0 on an atomic packet, some sub-fields of octets 5 & 6 of the IPv4 header MAY be interpreted with the semantics of the ID-Reuse field, but only in the highly constrained circumstances defined in Section 6.2.

For the avoidance of doubt, the Recycled flag alone MUST NOT be assumed to indicate that the packet is atomic. Only the combination of ((DF==1) && (MF==0) && (Offset==0)) indicates that a packet is atomic. Then if the Recycled flag is also set, the ID field unambiguously has the semantics of the ID-Reuse field. If the Recycled flag of an atomic packet is cleared, its ID field only has

the semantics of the ID-Reuse field in specific limited circumstances.

It is expected that proposals to use the ID-Reuse field will each need a few bits, not the whole 16 bit field. Therefore this specification establishes a new IANA registry (Section 8) to record assignments of sub-divisions of the ID-Reuse field. In this way, it will be possible for new uses of different sub-divisions to be orthogonal to each other. The process for incrementally defining new sub-divisions is specified in Section 5.

If an IPv4 packet header has RC=1 but it is not atomic ((DF==0) || (MF==1) || (Offset !=0)), then all the fields of the IPv4 header are undefined and reserved for future use. If an implementation receives such a packet, it could imply:

- that some currently unknown attack is being attempted
- or that some future standards action has defined a meaning for this reserved combination of header values

Therefore, if an implementation receives a non-atomic packets with RC=1, it MUST treat the packet as if the Recycled flag were cleared to 0, but it MUST NOT change the Recycled flag to zero. It MAY log the arrival of such packets and/or raise an alarm. It MUST NOT always drop such packets, but it MAY drop them under a policy that can be revoked if it is established that the appearance of such packets is the result of a future standards action.

For convenience only, the above rules are summarised in Table 1. The semantics of octets 5 & 6 of the IPv4 header are tabulated for each value of the RC flag (rows) and for whether the packet is atomic or not (columns).

RC flag	Non-Atomic	Atomic
0	Identification	ID-Reuse (Limited)
1	Undefined	ID-Reuse

Table 1: The Dependence of the Semantics of Octets 5 & 6 of the IPv4 Header on whether the Packet is Atomic and on the RC Flag

4. Behaviour of Intermediate Nodes

4.1. End-to-End Preservation of ID-Reuse Semantics

If the source sets the RC flag to 1 on an atomic packet, another node MUST NOT clear the RC flag to zero. Otherwise the semantics of the ID-Reuse field would change (see the Security Considerations in Section 9 for discussion of the integrity of the ID-Reuse field). Note that intermediate nodes are already not expected to change an atomic packet to non-atomic, which otherwise would also risk changing the semantics of the ID-Reuse field.

If the source zeros the RC flag on an atomic packet, an intermediate node MAY change the RC flag to 1. At this time, no case is envisaged where an intermediate node would need to do this. However, as this behaviour preserves ID-Reuse semantics safely, it is not precluded in case it will prove useful (e.g. for sender proxies).

4.2. Tunnel Behaviour

This specification does not need to change the following aspects of IPv4-in-IPv4 tunnelling, which already provide the most useful semantics for the ID-Reuse field:

- For some time, it has been mandated that an atomic packet "MUST" be encapsulated by an atomic outer header [RFC2003] (although some implementations are broken in this respect).

- On decapsulation the outgoing header will naturally propagate the ID-Reuse field of the inner header.

However, compliant IPv4 encapsulation implementations SHOULD copy the ID-Reuse field when encapsulating an atomic IPv4 packet in another atomic IPv4 header, irrespective of the setting of the Recycled flag. It would be ideal but impractical to assert 'MUST' in this last clause, given it cannot be assumed that pre-existing IPv4-in-IPv4 encapsulators will propagate the ID-Reuse field to the outer header (see Section 5.1).

IPv6 packets without a fragmentation extension header are inherently atomic. Therefore, if an IPv4 header encapsulates an IPv6 packet, the encapsulator is already required to set the outer as atomic.

There is no direct mapping between the IPv4 ID-Reuse field as a whole and any IPv6 header field (main or extension), because the ID-Reuse field is merely a container for yet-to-be-defined sub-fields. However, sub-fields of the ID-Reuse field might be defined to provide a mapping for IPv6 extension headers that need to be visible in the

outer IPv4 header of a tunnel. The present specification cannot say anything in general about any such mappings or any associated tunnel behaviour. Any such behaviour will have to be defined when individual ID-Reuse sub-fields are specified.

5. Process for Defining Subdivisions of the ID-Reuse Field

When IPv4 was designed, then later IPv6, all the fields in the main IP header were initially defined together in a coordinated fashion. In contrast, the only practical way to define new uses for the bits in the ID-Reuse field will be to adopt a gradual addition approach, in which subsets of the bits or codepoints will have to be assigned on the merits of each request at the time.

Each new scheme will need to submit an RFC that requests a subdivision of the ID-Reuse field and assigns behaviours to the codepoints within this subdivision. A specification defining a new use of a subdivision of the ID-Reuse field MUST register this use with the IANA, which will maintain a registry for this purpose (Section 8).

Proposals to reuse the IP ID field could relate to other parts of the IPv4 header in the following different ways {ToDo: this list is not exhaustive}:

Orthogonal: Some new protocol proposals will need to apply whatever is in the rest of the packet, e.g. whether unicast or multicast, whatever the Diffserv codepoint and whatever else might have been added in the rest of the IP-Reuse field. Schemes that need to be orthogonal to other elements of the IPv4 protocol will require assignment of a number of bits as a dedicated sub-field of the ID-Reuse field.

Mutually exclusive: It might be impossible for two uses of the ID-Reuse field to both apply to the same packet. Such mutually exclusive schemes will only each require a range of codepoints within a sub-field.

Conditional: Some protocol proposals might only apply when other parts of the header satisfy certain conditions, e.g. only for multicast packets. The IANA will need to register these conditions so that the bits can still be assigned for other uses when the conditions do not apply.

To allow interworking between sub-fields that are being defined incrementally, every new protocol MUST assign the all-zeros codepoint of its sub-field to mean the new protocol is 'turned off'. This means that implementations of the new protocol will treat such

packets as they would have been treated before the new protocol was defined.

Implementations MUST also clear to zero any bits in the ID-Reuse field that are not defined at the time the implementation is coded.

Proposals to use sub-fields of ID-Reuse will have to be assessed in the order they arrive without knowing what future proposals they might preclude. To judge each proposal, at least the following criteria will be used:

Constraint satisfaction: Each proposal MUST either satisfy all the constraints in Section 5.1 below, or include measures to circumvent them.

General usefulness: Proposals that are not applicable to a broad set of services that can be built over the internetwork protocol SHOULD NOT warrant consuming the newly freed up IPv4 header space.

Parsimony: Burning up a large proportion of the remaining bits will count against a proposal.

Backward compatibility with prior uses of ID-Reuse: As more sub-fields of the ID-Reuse field become defined, each new proposal SHOULD ensure that it takes into account potential interactions with earlier standards actions or experiments defining other sub-fields.

Forward compatibility with potential uses of ID-Reuse: In addition, proposals that demonstrate sensitivity to potential future uses of the remaining sub-fields of the ID-Reuse field will be more likely to progress through the IETF's approval process.

Do no harm: Proposals that do no harm to existing uses of the Internet will be favoured over those that do more harm.

5.1. Constraints on Uses of the ID-Reuse Field

Atomic packets: The IPv4 ID field cannot be reused if the packet is not atomic, because then the IP ID field will need to be used for its original purpose: fragment reassembly.

IPsec interaction: The IP Authentication Header (AH) [RFC4302] assumes and requires the IPv4 ID field to be immutable, otherwise verification of authentication and integrity checks will fail. Any new use of bits in the ID-Reuse field MUST ensure the bits are immutable, at least between IPsec endpoints (whether transport or tunnel mode). It cannot be assumed that pre-existing IPsec

implementations will check the setting of the Recycled flag.

Note that the Recycled flag itself is considered mutable and masked out before calculating an authentication header [RFC4302] (see Section 9).

Tunnelling: Any new use of the ID-Reuse field in atomic packets cannot reliably assume that the ID-Reuse field will propagate unchanged into the outer header of an IPv4-in-IPv4 tunnel [RFC2003, RFC4301]. It is likely that an IPv4 tunnel ingress will encapsulate an atomic packet with another atomic outer header, as this behaviour was mandated in [RFC2003]. However it is known that some implementations are broken in this respect. It is possible that an IPv4 encapsulator might copy the IP ID field of an arriving atomic packet into the outer header. However this behaviour has never been required and therefore cannot be guaranteed for pre-existing tunnels.

Nonetheless, it can be assumed that the IPv4 ID field will be preserved through the inner header into the outgoing packet at the other end of the tunnel (even though this behaviour would not strictly have been necessary for an atomic packet).

Incremental deployment: Each new proposal will need to consider any detrimental effects from pre-existing IPv4 implementations, assuming that they are likely to act on atomic packets without first checking on the setting of the Recycled flag.

5.2. Process Example

For illustration purposes, imagine two RFCs have been published: an experimental track RFC called Experiment A (ExA) and a standards track RFC called Standard B (StB) and . Imagine they define respectively a use for bits bits 14 to 15 and 11 to 13 of the ID-Reuse field. Figure 4 shows example IANA registry entries for these imaginary sub-fields.

```
Protocol name:         StB
RFC:                   BBBB
Leftmost bit:          11
No. of bits allocated: 3
Sub-field defined if:  Atomic packet and RC=1

Protocol name:         ExA
RFC:                   AAAA
Leftmost bit:          14
No. of bits allocated: 2
Sub-field defined if:  Atomic packet and RC=1
```

Figure 4: Example IANA Registry of Sub-fields of the ID-Reuse Field

Figure 5 shows an example of how incremental specification of subdivisions of ID-Reuse would work.

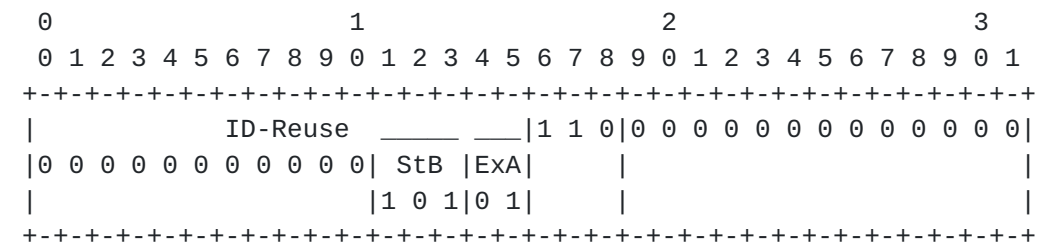


Figure 5: Example of Reuse of Octets 5 & 6 using RC=1

The bits shown in each row of Figure 5 define the semantics of the bits shown in the next row down, as follows:

- The top row identifies that the packet is atomic and the RC flag is 1. Therefore octets 5 & 6 of the IPv4 header are redefined as the ID-Reuse field.

- The middle row shows the bits assigned to Standard B and Experiment A by IANA. An implementer has to ensure that all the bits of the ID-Reuse field that are yet to be defined (bits 0-10) are cleared to zero.

- The bottom row shows that an implementation of ExA has set its 2-bit sub-field to codepoint 01 and an implementation of StB has set its 3-bit sub-field to codepoint 101. The meaning of each would be defined in the RFCs for ExA and StB respectively.

Imagine now that Experiment C (ExC) is defined later to use bits 0-7 of the ID-Reuse field. If the packet in Figure 5 is received by an implementation of ExC, then it will see only zeros in the ExC sub-field. Therefore the implementation of ExC will treat the packet as if ExC is turned off (as mandated in Section 5).

Similarly, the implementation of protocol StB can rely on being able to turn off Experiment A by setting bits 14 & 15 to zero.

6. Incremental Deployment of New Uses of the IPv4 ID Field

When implementations first set the Recycled flag to 1, they are likely to be blocked by certain middleboxes, either deliberately (e.g. firewalls that assume anomalies are attacks) or erroneously (e.g. having misunderstood the phrase "reserved, must be zero" in RFC791). It is also possible that broken 'normalisers' might clear RC to zero if it is 1, although so far no tests have found such broken behaviour.

To address this problem, Section 6.2 introduces a way to use a sub-field of ID-Reuse without having to set RC=1. In this approach, packet headers using the new protocol will be indistinguishable from an IPv4 header not using the new protocol. Therefore it will be possible to guarantee that middleboxes will not treat packets using the new protocol any differently from other IPv4 packets.

Many pre-existing IPv4 hosts cycle through all the values in the IP ID field even when sending atomic packets in which the IP ID field has no function. Therefore, these pre-existing IPv4 hosts will occasionally issue a packet that happens to look as if it is using a codepoint of a new protocol using the IP ID field. Without RC=1, there will be no way to distinguish the two.

	middlebox traversal	new protocol recognition
RC=0	Assured	Uncertain
RC=1	Uncertain	Assured

Table 2: Tradeoff between deterministic middlebox traversal and deterministic protocol recognition

Table 2 shows the tradeoff between using RC=0 or RC=1:

RC=0: If an implementation of a new protocol uses RC=0, its packets will traverse middleboxes, but it will suffer a small fraction of false positives when recognising which packets using the new protocol -- occasionally it will mistakenly assume a packet is using the new protocol when it is actually just random noise in the IP ID field from a pre-existing implementation.

RC=1: If an implementation of a new protocol uses RC=1, its packets may be black-holed by some middleboxes, but it will be certain which packets use the new protocol and which don't.

Nonetheless, a probabilistic protocol that can be deployed may be more useful than a deterministic protocol that cannot.

6.1. Process Example with RC=0

Figure 6 shows an example of how this approach would work with RC=0. For illustration purposes imagine, as in the previous example in Section 5.2, that an experimental track RFC has been published called Experiment A (ExA) that defines bits 14 to 15 of the ID-Reuse field for atomic packets with RC=1. Now imagine another experimental track RFC has been published called Experiment B (ExB) that defines a use for bits 11 to 13 of the ID-Reuse field, but does not require RC=1. In fact a packet is defined as complying with ExB whether RC=1 or RC=0 (i.e., RC=X, where 'X' means don't care). Figure 7 shows the IANA registry entries for these imaginary sub-fields.

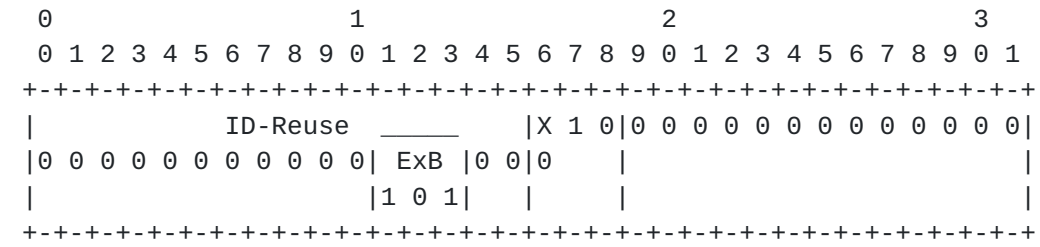


Figure 6: Example of Experimental Reuse of Octets 5 & 6 Without Requiring RC=1

The bits shown in each row of Figure 6 define the semantics of the bits shown in the next row down, as follows:

- The top row identifies that the packet is atomic. The RC flag is don't care ('X'), so RC does not have to be 1. Implementations can clear RC to 0 to traverse awkward middleboxes, but RC can be set to 1 otherwise.

- The middle row shows that an implementation of Experiment B (ExB) has set RC=0. It is also using the ID-Reuse field, so it clears all the bits to zero except those in its own sub-field (bits 11-13). It will have registered this experimental use with the IANA as shown in the top example of Figure 7.

- The bottom row shows that an implementation of ExB has set its 3-bit sub-field to codepoint 101, the meaning of which will have been defined in the RFC specifying the ExB protocol.

Note that, the process for using protocol ExB without RC=1 (Section 6.2) precludes an implementation from using the ExA protocol in the same packet -- any one packet can only be part of one RC=0 protocol at a time.

6.2. Process for Using the ID-Reuse Field Without Requiring RC=1

This approach SHOULD NOT be used unless the preferred approach (Section 5) is impractical due to middleboxes blocking packets with RC set to 1.

To follow this non-preferred approach, the registration with the IANA MUST specify that the sub-field of ID-Reuse is defined for 'RC=X', meaning "don't care", that is RC may be either set or cleared (for an example, see the final bullet of the imaginary registration details in Section 8). The RFC defining the relevant ID-Reuse sub-field MUST also make it clear that the sub-field is defined for either value of the Recycled flag (RC=X) in an atomic IPv4 packet.

This approach will not be feasible for all protocols; only those that satisfy the severe constraints laid down below. Otherwise, for protocols that cannot satisfy these prerequisite constraints, the preferred approach in Section 5 wth RC=1 will be the only option.

Once a sub-field of the ID-Reuse field has been registered with the IANA, implementations of the protocol can use any of the available codepoints within that sub-field in atomic packets without having to set RC=1, if and only if the following constraints can be satisfied:

1. New protocol implementations MUST NOT use RC=0 unless the treatments associated with all the new codepoints are generally benign to packets not taking part in the protocol. 'Benign' means the new protocol SHOULD do no more harm to other packets than previous implementations did. Using the term 'SHOULD' rather than 'MUST' does not completely rule out new protocol proposals that might sometimes introduce slightly more harm, but such proposals will need to give strong justifications

2. Implementations MUST clear all the other bits of the ID-Reuse field (except those in the new protocol's sub-field) to zero. Note that this is different to the approach with RC=1, where more than one sub-field at once can be non-zero

3. In addition the constraints in Section 5.1 must also be satisfied.

Constraint #1 is severe but necessary in order to ensure that a new protocol (e.g. ExB) does not harm atomic packets from pre-existing

IPv4 implementations. For example, a receiving implementation of ExB can assume that most packets with all zeros in bits 0-10 and 14-15 were deliberately set by another implementation of ExB. But many pre-existing implementations of IPv4 will be cycling (sequentially or randomly) through all the IPID values as they send out packets. Occasionally they will send out a packet that happens to look like it complies with protocol ExB. For the case of ExB with a 3-bit sub-field, such false positives will occur with probability 1 in 2^13 (~0.01%). We term this the misrecognition probability.

If the new protocol were designed to do harm (e.g. to deprioritise certain packets against others) that would be fine for those packets intended to take part in the new protocol. But it would not be acceptable to harm even a small proportion of packets misrecognised as using the new protocol. This is why the RC=0 approach can only be allowed for a new protocol that is generally benign.

Constraint #2 is necessary in order to ensure the misrecognition probability remains low. If only one sub-field is allowed at one time, all the other bits in the ID-Reuse field will have to be zero. This ensures that a pre-existing IPv4 implementation cycling through all the IP ID values will collide less frequently with values used for each new protocol.

As already stated (Section 5), each new protocol MUST define the all-zeros codepoint of its sub-field to mean that the new protocol is 'turned off'.

This arrangement ensures that packets with an IPv4 ID of zero will never collide with a codepoint used by any ID-Reuse scheme, whether RC=0 or RC=1. All zeros was deliberately chosen as the common 'turned off' codepoint because some pre-existing implementations have used zero as the default IP ID for atomic packets.

In either case, whether the Recycled flag is set or not, a sub-field of the ID-Reuse field MUST be registered with the IANA, initially for experimental use, by referencing the relevant experimental track RFC. This will ensure that experiments with different sub-fields of the ID-Reuse field can proceed in parallel on the public Internet without colliding with each other. The referenced RFC MUST define a coherent process for returning the bits for other uses if the experimental approach does not progress to the standards track.

The same sub-field can be used with the same semantics as the experiment progresses, initially with the Recycled flag cleared to 0 and later set to 1. And the same protocol semantics can be used whether the proposal is experimental or standards track. Thus, the whole process is designed to:

1. allow initial experiments to use RC=0 to traverse non-compliant middleboxes (Section 6);

2. then, once sufficient middleboxes forward RC=1 packets, the experiment can either be continued with RC=1 (Section 5);

3. or the experiment can progress cleanly to the standards track, while still using the same sub-field but with RC=1;

4. or the experiment can be terminated without having wasted any header bits.

(Step 1 is only feasible if the extra constraints in Section 6.2 can be satisfied. If not, Step 2 will be the only feasible first step.)

For the avoidance of doubt, any use of ID-Reuse, whether experimental or not, is also subject to the general constraints already enumerated in Section 5.1.

7. Updates to Existing RFCs

Great care has been taken to ensure all the updates defined in this specifications are incrementally deployable.

The definition of the RC flag in Section 3 updates the status of this flag that was "reserved, must be zero" in [RFC0791]. The redefinition of the IP Identification field as the ID-Reuse field when an IPv4 packet is atomic also updates RFC791.

Updates to existing RFC791 implementations are only REQUIRED if they discard IPv4 packets with RC=1, or change RC from 1 to 0, both of which are misinterpretations of RFC791 anyway. Otherwise, there will be no need to update an RFC791-compliant IPv4 stack until new use(s) for the ID-Reuse field are also specified.

The recommendation in Section 4.2 to copy the ID-Reuse field when encapsulating an atomic IPv4 packet with another atomic IPv4 header updates IPv4-in-IPv4 encapsulation specifications [RFC2003] [RFC4301]. These updates to tunnels are likely to be recommended rather than essential for interworking, so they can be implemented as part of routine code maintenance.

The ability to redefine the IPv4 ID field of an atomic packet updates [ipv4-id-update], which states "The IP ID is not defined if the packet (datagram) is atomic". Nonetheless, octets 5 & 6 of an atomic packet still MUST NOT be interpreted with the semantics of the Identification field.

[RFC2780] provides the IANA with guidelines on allocating values in IP and related headers. The process defined in Section 5 and Section 6 update RFC2780, given ID-Reuse is effectively a new field in the IPv4 header.

[RFC4727] defines the processes for experimental use of values in fields in the IP header that are managed by the IANA. The processes defined in Section 5 and Section 6 update RFC4727 to include the new alternative use of the IPv4 ID field as an ID-Reuse field.

8. IANA Considerations

The IANA is requested to establish a new registry to record allocation of sub-divisions of the ID-Reuse field and to avoid duplicate allocations. The ID-Reuse field is an alternative use of the Identification field of the IPv4 header in atomic packets (Section 3). All 16 bits are available for assignment, either as sub-fields of bits or as sets of codepoints within a sub-field of bits. Each sub-division of the ID-Reuse field MUST be allocated through an IETF Consensus action. The registry MUST then record:

Protocol name: the name for the protocol, as used in the RFC defining it

RFC: the RFC that defines the semantics of the codepoints used by the protocol

Leftmost bit: the leftmost bit allocated, counting from bit 0 at the most significant bit (which is bit 32 of the IPv4 header, counting from 0)

No. of bits allocated: the width in bits of the allocated sub-field

Codepoint range (optional): The range of codepoints within the assigned sub-field of bits that the protocol uses

Sub-field defined if: the precondition for the sub-field to be defined (Section 5). Valid entries MUST include the condition that the packet is atomic and MUST specifiy valid values of the Recycled (RC) flag, either 'RC=1' or 'RC=X', where 'X' means don't care (Section 6).

Two example registrations are shown in Figure 7.

```
Protocol name:         ExB
RFC:                   BBBB
Most significant bit:  11
No. of bits allocated: 3
Codepoint range:       all
Sub-field defined if:  Atomic packet and RC=X

Protocol name:         ExA
RFC:                   AAAA
Most significant  bit: 14
No. of bits allocated: 2
Codepoint range:       all
Sub-field defined if:  Atomic packet and RC=1
```

Figure 7: Example IANA Registry of Sub-fields of the ID-Reuse Field

9. Security Considerations

Integrity Checking: This specification make the semantics of octets 5 & 6 of the IPv4 header (IP ID or ID-Reuse) depend on the setting of octets 7 & 8 (all the Control Flags and the Fragment Offset field). The IP Authentication Header (AH) [RFC4302] covers octets 5 & 6 but not octets 7 & 8. Therefore AH can assure the integrity of the bits in the ID-Reuse field, but it cannot verify whether or not the sender intended those bits to have the semantics of an ID-Reuse field.

Any security-sensitive application of the ID-Reuse field will therefore need to provide its own integrity checking of the status of the Control Flags and Fragment Offset. Such a facility would need to take into account that the present specification allows an intermediate node to set the Recycled flag, but not to clear it (Section 4.1).

Covert channels: It has always been possible to use bit 48 of the IPv4 header for a 1 bit per packet covert channel, for instance between a network protected by IPsec and an unprotected network. Bit 48 could be covertly toggled to pass messages because it had no function (so no-one would notice any affect on the main communication channel) and it was not covered by IPsec authentication. On the other hand, once alerted to the vulnerability, it has always been easy for an IPsec gateway to spot bit 48 being used as a covert channel, given bit 48 was meant to always be zero.

Now that bit 48 has been given a function, it will often no longer be possible for an attacker to toggle it without affecting the main data communication. However, whenever the main communication

does not depend on bit 48, it will be easier to for an attacker to toggle it covertly given it will no longer stand out as anomalous behaviour.

10. Conclusions

This specification builds on recent moves to make the approach to fragmentation in IPv4 more closely match that of IPv6. Already the fields that support fragmentation in the IPv4 header are usually redundant, but unfortunately they are non-optional.

This specification makes it possible to reuse the 16 bits of the IPv4 ID field when they are not needed for reassembly. The last unused bit in the IPv4 header is used in order to unambiguously flag that the IP ID field has new semantics. The bit is called the Recycled flag, because it allows the IP ID field to be recycled for new purposes when it would otherwise be redundant. Whenever the IP ID field has new semantics, it is termed the ID-Reuse field.

The process for redefining the semantics of sub-fields of this ID-Reuse field has been laid down, both for experimental and standards actions. Great care has been taken throughout to ease incremental deployment. The same sub-field can be used with the same semantics as an experiment evolves into a standards action. Initially it is even possible for certain experiments to leave the Recycled flag cleared to zero, in order to traverse any awkward middleboxes that incorrectly discard or normalise packets if the Recycled flag is set.

11. Acknowledgements

Rob Hancock originally pointed out that code to handle new protocols can tell the machine where to look for the relevant header. Dan Wing pointed out that codepoints, not just whole bits, could be assigned for protocols that are mutually exclusive.

Bob Briscoe is partly funded by Trilogy, a research project (ICT-216372) supported by the European Community under its Seventh Framework Programme.

Comments Solicited (to be removed by the RFC Editor):

Comments and questions are encouraged and very welcome. They can be addressed to the IETF Internet Area working group mailing list <int-area@ietf.org>, and/or to the author(s).

12. Outstanding Issues (to be removed when all resolved)

1. ...

13. References

13.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

[RFC2003] Perkins, C., "IP Encapsulation within IP", RFC 2003, October 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2780] Bradner, S. and V. Paxson, "IANA Allocation Guidelines For Values In the Internet Protocol and Related Headers", BCP 37, RFC 2780, March 2000.

[RFC4301] Kent, S. and K. Seo, "Security Architecture for the Internet Protocol", RFC 4301, December 2005.

[RFC4302] Kent, S., "IP Authentication Header", RFC 4302, December 2005.

[RFC4727] Fenner, B., "Experimental Values In IPv4, IPv6, ICMPv4, ICMPv6, UDP, and TCP Headers", RFC 4727, November 2006.

[ipv4-id-update] Touch, J., "Updated Specification of the IPv4 ID Field", draft-ietf-intarea-ipv4-id-update-04 (work in progress), September 2011.

13.2. Informative References

[Cisco.IPv6Ext] Cisco, "IPv6 Extension Headers Review and Considerations", Cisco Technology White Paper , October 2006, <http://www.cisco.com/en/US/technologies/tk648/tk872/technologies_white_paper0900aecd8054d37d.html>.

[Fransson04] Fransson, P. and A. Jonsson, "End-to-end measurements on performance penalties of IPv4 options", Luleae University of Technology, Technical Report 2004:03, 2004, <http://pure.ltu.se/portal/files/1299598/

LTU-TR-0403-SE.pdf>.

[RFC4821] Mathis, M. and J. Heffner, "Packetization Layer Path MTU Discovery", RFC 4821, March 2007.

[RFC4963] Heffner, J., Mathis, M., and B. Chandler, "IPv4 Reassembly Errors at High Data Rates", RFC 4963, July 2007.

Appendix A. Why More Bits Cannot be Freed (To be Removed by RFC Editor)

Given this specification uses the last unassigned bit in the IPv4 header, it is worth checking whether it can be used to flag a new use for more than the 16 bits in the IP ID field of atomic packets.

IHL: Ideally, the Internet header length field (4 bits) could be made redundant if the length of those IPv4 headers with bit 48 set were redefined to be fixed at 20 octets. Then a similar approach to IPv6 could be taken with the Protocol field redefined as a Next Header field and each extension header specifying its own length.

Unfortunately, although IPv4 options are rarely used and generally ignored, this idea would not be incrementally deployable. There are probably billions of pre-existing implementations of the IPv4 stack that will use the IHL field to find the transport protocol header, without ever looking at bit 48. If the IHL field were given any other semantics conditional on bit 48 being set, all these pre-existing stacks would break.

Header Checksum: Ideally, the Header Checksum (16 bits) could be made redundant in those IPv4 headers with bit 48 set. Then a similar approach to IPv6 could be taken where the integrity of the IP header relies on the end-to-end checksum of the transport protocol, which includes the main fields in the IP header.

Unfortunately, again, this idea would not be incrementally deployable. Pre-existing implementations of the IPv4 stack might verify the header checksum without ever looking at bit 48. And anyway IPv4 stacks on probably every pre-existing router implementation would update the checksum field without knowing to check whether bit 48 was set. Therefore if the field were used for any other purpose than a checksum, it would be impossible to predict how its value might be changed by a combination of pre-existing and new stacks.

It is clear that reusing fields other than the IPv4 ID would be fraught with incremental deployment problems. The reason the IPv4 ID field can be reused, is that an atomic packet already does not need

an Identification field, whether bit 48 is set or not. Setting bit 48 merely allows new implementations that understand ID-Reuse semantics to be certain the value in the ID-Reuse field was not written by an implementation that intended it to have Identification semantics.

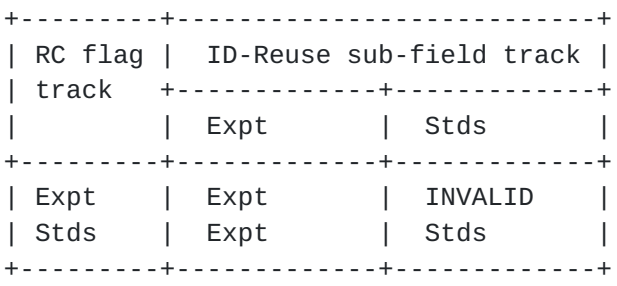
Appendix B. Experimental or Standards Track? (To Be Removed Before Publication)

This document defines a protocol (using the Recycled flag) to enable other protocols (using the ID-Reuse field). The Recycled flag protocol is currently written as if it is on the IETF standards track. Nonetheless it might be feasible to write it for the experimental track. This appendix discusses the pros and cons of each.

The Recycled flag uses up the last unused bit in the IPv4 header. The present specification defines a use for this last bit in the expectation that the Internet community will find ingeneous new use(s) for sub-fields of the ID-Reuse field, because then the Recycled flag will be needed to unambiguously indicate the new semantics. However, there is a risk that the last IPv4 header bit could be wasted, if no new uses for the IP ID field can be found within the constraints of its previous use for fragment reassembly, or if new experimental uses are proposed but none successfully proceed through to standards actions.

The risk of wasting the last bit would be mitigated if the definition of the Recycled flag itself was initially on the experimental track. Then, if some experimental use(s) of the ID-Reuse field did see widespread adoption, the RC flag protocol could progress to the standards track. On the other hand, if no ID-Reuse experiments happened, the RC flag could possibly be reclaimed for another use in the future. This would require all experiments with the RC flag to be confined in time, so that stray implementations of old experiments would not conflict with future uses of the flag.

Eventually, each specification for each sub-field of ID-Reuse might either progress on the experimental track or standards track. However, an enabler for standards track specifications cannot itself only be experimental. Therefore the RC flag protocol would have to be on the standards track, to enable standards track protocols as well as experimental. Figure 8 illustrates this need for the RC flag protocol to have sufficient rank for any protocols it enables.



The IETF track of the RC flag protocol in the present document (rows) and of any particular RFC specifying a sub-field of the ID-Reuse field (columns). The combination determines the status of any particular sub-field as shown at the intersection of the relevant row and column.

Figure 8: Validity of Combinations of IETF tracks for the RC flag and an ID-Reuse Subfield

One purpose of the present draft is to outline how new uses of ID-Reuse sub-fields can progress seamlessly from experimental track to standards track. Therefore, this draft is written as if it were on the standards track. Otherwise the processes for enabling standards track documents would have had to be written hypothetically, which would have been highly confusing. Nonetheless, no intent to prejudge that this document should be or will be on the standards track is implied.

If it were decided that the present draft should start on the experimental track, all the text about enabling standards track protocols would have to be edited out, or perhaps moved to a non-normative appendix.

Alternatively, the IETF might see some obvious new uses for sub-fields of the ID-Reuse field that would make it reasonable to fast-track the RC flag straight onto the standards track.

Appendix C. Changes in This Version (to be removed by RFC Editor)

From briscoe-00 to 01:

* Updated to preserve liveness.

* No changes other than updates to refs and minor corrections.

Author's Address

Bob Briscoe
BT
B54/77, Adastral Park
Martlesham Heath
Ipswich IP5 3RE
UK

Phone: +44 1473 645196
EMail: bob.briscoe@bt.com
URI: http://bobbriscoe.net/