Thursday, March 14, 2013
Douglas Otis

DOMAINS AS A BASIS FOR MANAGING TRAFFIC

A move to domains instead of IP addresses to manage traffic can be an attractive method to cope with the larger IPv6 address space. Assessments by IPv6 addresses is less practical, even when limited to just the prefix space that has an announced size of 18,205,630,682,693,634 not counting the /64 network identifiers. A manageable basis for assessments can leverage a smaller number of related domains, compared to IPv6 or even IPv4 addresses. Although technically the domain name space can be larger than the massively large IPv6 address space, in practice it is not. One hundred thousand domains control 90% of Internet traffic out of approximately 100 million domains active each month. The top 150 domains control 50% of the traffic, and the top 2,500 domains control 75%. This level of domain consolidation permits effective fast-path white-listing.

In general, abusive sources of traffic is determined by undesired reception of commercial solicitations, inclusion of compromising malware, or deceptive spoofing of sources. DKIM signatures play a limited role in preventing erroneous detection of deceptive content, but are silent on actual sources by design. DMARC may request a From header field be affirmed by either a matching DKIM signature or SPF record set, but even these limited requirements are unsuitable for much of today's traffic. SPF and to a lesser extent DKIM suffer from a significant failure rate ameliorated by allowing either to fail in acceptance policies. Examples on the impact of this approach will be given later.

It is essential that reputations be defendable for both domain owners and those qualifying sources. To ensure fairness, assessments MUST authenticate domains initiating traffic qualified as abusive. Therefore, reputation defensibility precludes use of DKIM domain signatures, From or Sender header field domains, and even opportunistic associations of IP addresses with message content unrelated to source/destination tuples. By design, DKIM signature validity is independent of source/destination tuples, and even permit injection of third-party pre-pended header fields. Often pre-pended headers are displayed in lieu of subsequent signed header fields. DKIM source/destination independence means DKIM validation can not prevent trivial spoofing or poisoning of a domain's reputation since it also permits circumvention of rate limiting.

Several methods can offer a fair basis for qualifying abusive sources. These methods may include use of DANE public keys published by specific domains, or address ranges referenced in APL Resource Records. Examples of some changes that could be introduced would be cryptographically authenticated traffic sources using

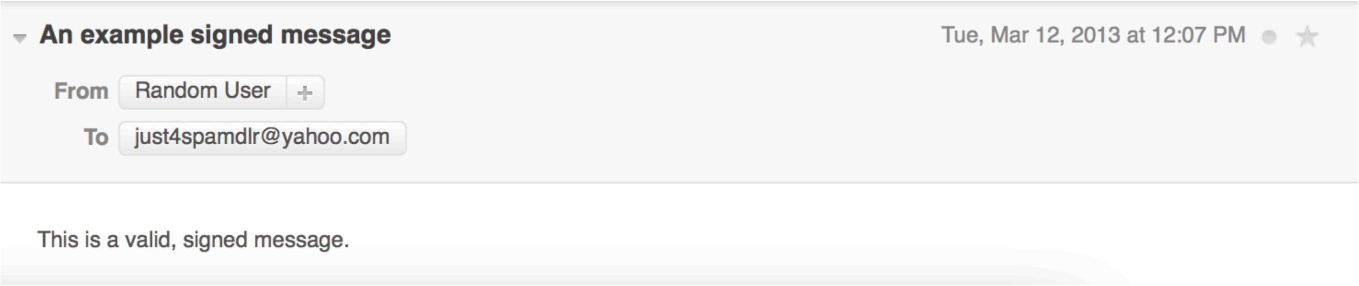
StartTLS, or forward reference confirmation of ELHO domains by referencing APL RRs. Algorithmic blocking of abusive actors can not be based on assumed relationships between message initiators based on signed message content. Email does not permit initiators or intended recipients to be associated with message signatures. Assumed associations with either traffic initiators or recipients can not serve as evidence of abuse by a domain offering message signatures as seen in the following examples.

Example of a properly signed DKIM message:

```
From Random User Tue Mar 12 12:07:37 2013
X-Apparently-To: just4spamdlr@yahoo.com via 72.30.237.8; Tue, 12 Mar 2013 12:08:37 -0700
Return-Path: <Fake.user@gmail.com>
Received-SPF: neutral (192.83.249.65 is neither permitted nor denied by domain of gmail.com)
 A3RleHQvcGxhaW4DAzACA3RleHQvaHRtbAMDMQ–
X-YMailISG: Po8J_9cWLDuz5QIo_tChc7OagZYPBIscsK7APx8FMj835hEX
 clyJxoQr6Ojy40ccEugqmkym_ayJu65fKm.KJY73k6aprxb9s7Bj6P32lpml
 6yGzxWFYdNXCwcxHtFGdhKe3v7Tjh8x051jkxjIqfuS0vo8J5rZOr.Z__6vD
 4wiGFDUwFHNUWAwuz_pwp7pZ5HCivuuuyszYVvH0eIFsrQ9crR.rrk_3EQU2
 Xkv_fInlGDFR8fafFPMOgQ7QOrHhy0zQUbptDEFGdh1QVOyLwIpjwEC7264k
 4MqxUH7zz_M5JOQzj6dJslH0.iz5y9Sgp6y6kTUHAVP2f_t1hMeRvf3F7WJ6
 1yY2rZJALIME1CtiNKQJoDctzgGFRnh_5mo415MvUcEIH7qqS5RFgWtXEQpd
 JIpyYlECDXVUcuASoLmzbuGSiCEVLq7f4EiBTAsaMwXJ07OgXBR.QYDw3VfA
 Z0AcfnFrUVHNLZtLaFukQKzdk9c6SpHFHSuCAsvLPuZeRy4Ij5ndXd7viyCS
 IkAHsnhG_u3.nZr3zUDFOrqw8sEKphobj6ZJ8KEXtuhr_tx.94abE1JRJYi5
 fukj2h8y9s.K10ZxoTClaw41_DD8fxESbyfyTRPytiEXUdK1WEjgS3rAZ0TA
 WPJPDr063xLYk20UY0V.N5J15lBCtqZcde_9pdXwxVySyXo1KEQOaH3TNRBZ
 AKMFuCC7NF56aklkiUgk2EWm8iYoHsFez5_HtOz1zmc1dv4mNFOPTaNrXF2X
 qjFiwfdUipupIlAEc6pIdv0_le.xvz1jnaewEOyxo4dKd2XLVvybLfsLY16U
 FzLS9MJJ1wC0Cmf3G2SbOmT4ZiAvPjyv8QnHzbSDDDy3hqg8F0uEE03sJ5dm
 on5FxOHZZ1wCH7DL1QAXpZYxYWKV.h3q69dKQMl6HbnmfT_WZQY4X8uKXqkZ
 o34v.YmvJxHSRCSmhFpug1EstpJ4gHVitl_eJzT_n6xYQwhNAuMZ9uRjN2xE
 1Lf7NpgzRf9bFvOpJAlyLoK5Xvxbx711cMgEUfGIha_JtL1P7hyfncRszHDv
 txgUYzcsVvRyAyVvwDAM.TEBsFhAtqqwOibqo2l5xCBj2yXRbKJ0EOC1JDMs
 HA–
X-Originating-IP: [192.83.249.65]
Authentication-Results: mta1225.mail.bf1.yahoo.com  from=gmail.com; domainkeys=neutral (no sig);
from=gmail.com; dkim=pass (ok)
Received: from 127.0.0.1  (EHLO rdaver.bungi.com) (192.83.249.65)
  by mta1225.mail.bf1.yahoo.com with SMTP; Tue, 12 Mar 2013 12:08:36 -0700
Received: by rdaver.bungi.com
        via smail with stdio
        id <m1UFUYr-00KeXPC@rdaver.bungi.com>
        for Just4spamdlr@yahoo.com; Tue, 12 Mar 2013 12:08:33 -0700 (PDT)
        (Smail-3.2.0.94 1997-Apr-22 #591 built 2011-Feb-5)
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;
     d=gmail.com; s=20120113;
     h=mime-version:x-received:date:message-id:subject:from:to
      :content-type;
     bh=PS9xMxYwwTGwWXbCd8bjBBm2rwb79wVOSDLhmp+k4b4=;
     b=qnYVUccLSAi2DGJdUgDDIP9A3uPk3PaxgqhYLBn6xU382MsCi/ICFgKAoFPuwM7BvL
      AuSuqL6P54cIJ3Pn36h2xmXy+ucNr5r5OqIY63rtvj6Apjr4uW1PzG47J7BGEiP9iwDZ
      PLTzl9ZLpZXvZZpTCJOXUQP2HF8q6aivCblYZIQcCdVRCftG+A4z0+dEyTHbxoAMx9U3
      GFISRRHcZ7k7GAyYmLrSr3fUTjvpa1YWoNK+IcSALC2tKVSW5FP1IQAT07f1e8+bOgHh
      JleaQIw8b1VjIzhs4hFKLdedmjQqjDJXVP/K3J+t/ggfYn4H547fu6Pb5syKZIiuPf1e
      yJqA==
MIME-Version: 1.0
X-Received: by 10.220.221.143 with SMTP id ic15mr6773333vcb.32.1363115257152;
 Tue, 12 Mar 2013 12:07:37 -0700 (PDT)
Received: by 10.52.70.169 with HTTP; Tue, 12 Mar 2013 12:07:37 -0700 (PDT)
Date: Tue, 12 Mar 2013 09:07:37 -1000
Message-ID: <CA+VnpPKv0s-p2nKkAkNHS4V2SxZehw_6S9QF5p1p2ji+FMof=Q@mail.gmail.com>
Subject: An example signed message
From: Random User <random.j.user.994@gmail.com>
To: just4spamdlr@yahoo.com
Content-Type: multipart/alternative; boundary=14dae9cdc33bb0ff5204d7bf00ff
Content-Length: 280
```

Example of a faked DKIM email:

```
From Fake User Tue Mar 12 12:07:37 2013
X-Apparently-To: just4spamdlr@yahoo.com via 72.30.237.8; Tue, 12 Mar 2013 12:09:01 -0700
Return-Path: <Fake.user@gmail.com>
Received-SPF: neutral (192.83.249.65 is neither permitted nor denied by domain of gmail.com)
 A3RleHQvcGxhaW4DAzACA3RleHQvaHRtbAMDMQ–
X-YMailISG: gFqc.ysWLDtqkdjDpSCH39uGWhgFfnsGdWobzNb5os6sP0We
 _L38eAdX.VKZWQ2F75gFwoipcPyj4g0uKMm_vSayLjrnps9lBxMGLvtTE8kT
 XYxIw6vZb4aFZ_jEcpoRntvJDkZQl4XSGWGakfmJ5G2blTWZ_i1BVkBvj0Sv
 jEymvhoIXZTb_l8C0Jh69ot3MgrNBvjhrBmhCK3sziUtDPpKQPJb_lxCnYKN
 O0SiArQ_TUXrCRFRNsyEiJxzVfSgJWIdsCV5BN3cp..NZ17X8fguB.YxNQjt
 qjVcGMd4IjQioY.a4f1luQxuiCN1yWvYqiLpP6eOCQhMrHt9XOdk32HAXNuJ
 GBraVtjrySTl9Db7PpRC46wlMs3iIUHl3z0d4o6293sMA5qFmnbczGoLRGFs
 RUVlBJuRoJCSYZh5LOwbj0RPQNX2Nmw.LHwF7SY3XcZWFUjvUQQ2sdx63m_J
 Mgy7JHAwBTVH6ytULsbXvu38a5GIYHccfNnDKVjtsrIg9qBDpVASHrRkncL0
 MFLy5FHLb_XBW1TPztCFtlRViKr_HFxMob6aZIte6T57AMqlV2YAHwVNObwx
 WE8ZWTkKNWbXqJYytd3vyuyAHfuseBFP_Jfmj0zVtg52EXpIlDiTANEOTamP
 zeu23QbeRWJd_Gpz9bbGw_OorPdcV.WJOQ29DHpiYAQRgWjJNLjkd8dI.vuM
 vs1Fr7LOiE3wRpSU5AW_hrR4anvGrnwSPOQaFmpNE0pl8n.Vomrp.5NU8cgU
 QYI1UCSPoE_HK5Som2HMPYZFQv0pJSu1NeitXlRM3DHkIMvW4aVYqrHSNVjl
 gGCFFx77c25QW.XAGtySBYWcTzcUlHP4fMa7Wli4u06C4N3pDPiQoXKOC10U
 koXUMKFYmedaZYvEeQRPO3_8xHwKyZ.QInDsnQRwPFWYKvcWCJu4c5zxDMG4
 h1AsyT3CM80nZXk8.ZGhzfTgo810Xjn_OJVgUfkG1z3..ReN990deaWJY8F5
 _j6lRWLZZRzCMwOGpJ6I.jgaN5mNk38Kj6.NYLFCpMTEIt28jIRHD85cfpa3
 iOL3drg1TIKQWrEhS9u3H29niQ_hjHbk7ys6uSJvowilRwO8eB2s.Wz0
X-Originating-IP: [192.83.249.65]
Authentication-Results: mta1266.mail.bf1.yahoo.com  from=gmail.com; domainkeys=neutral (no sig);
from=gmail.com; dkim=pass (ok)
Received: from 127.0.0.1  (EHLO rdaver.bungi.com) (192.83.249.65)
  by mta1266.mail.bf1.yahoo.com with SMTP; Tue, 12 Mar 2013 12:09:00 -0700
Received: by rdaver.bungi.com
            via smail with stdio
            id <m1UFUZI-00KeXRC@rdaver.bungi.com>
            for Just4spamdlr@yahoo.com; Tue, 12 Mar 2013 12:09:00 -0700 (PDT)
            (Smail-3.2.0.94 1997-Apr-22 #591 built 2011-Feb-5)
From: Fake User <fake.user@gmail.com>
DKIM-Signature: v=1; a=rsa-sha256; c=relaxed/relaxed;
       d=gmail.com; s=20120113;
       h=mime-version:x-received:date:message-id:subject:from:to
        :content-type;
       bh=PS9xMxYwwTGwWXbCd8bjBBm2rwb79wVOSDLhmp+k4b4=;
       b=qnYVUccLSAi2DGJdUgDDIP9A3uPk3PaxgqhYLBn6xU382MsCi/ICFgKAoFPuwM7BvL
        AuSuqL6P54cIJ3Pn36h2xmXy+ucNr5r5OqIY63rtvj6Apjr4uW1PzG47J7BGEiP9iwDZ
        PLTzl9ZLpZXvZZpTCJOXUQP2HF8q6aivCblYZIQcCdVRCftG+A4z0+dEyTHbxoAMx9U3
        GFISRRHcZ7k7GAyYmLrSr3fUTjvpa1YWoNK+IcSALC2tKVSW5FP1IQAT07f1e8+bOgHh
        JleaQIw8b1Vjlzhs4hFKLdedmjQqjDJXVP/K3J+t/ggfYn4H547fu6Pb5syKZIiuPf1e
        yJqA==
MIME-Version: 1.0
X-Received: by 10.220.221.143 with SMTP id ic15mr6773333vcb.32.1363115257152;
 Tue, 12 Mar 2013 12:07:37 -0700 (PDT)
Received: by 10.52.70.169 with HTTP; Tue, 12 Mar 2013 12:07:37 -0700 (PDT)
Date: Tue, 12 Mar 2013 09:07:37 -1000
Message-ID: <CA+VnpPKv0s-p2nKkAkNHS4V2SxZehw_6S9QF5p1p2ji+FMof=Q@mail.gmail.com>
Subject: An example signed message
From: Random User <random.j.user.994@gmail.com>
To: just4spamdlr@yahoo.com
Content-Type: multipart/alternative; boundary=14dae9cdc33bb0ff5204d7bf00ff
Content-Length: 280
```

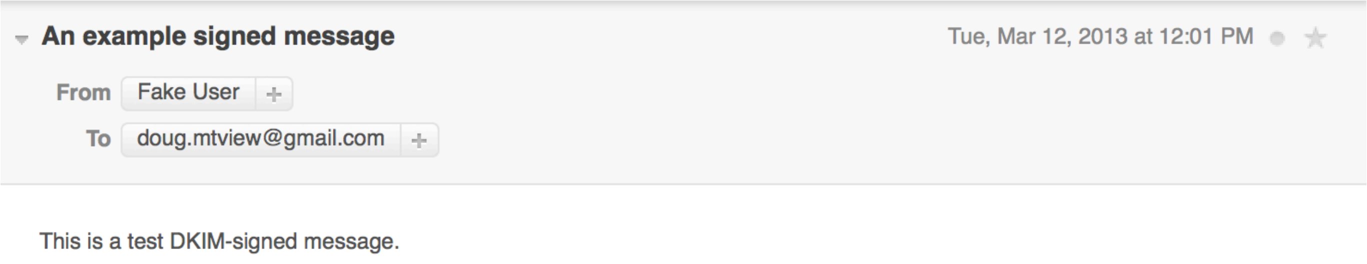


An example signed message Tue, Mar 12, 2013 at 12:07 PM

From Random User

To just4spamdlr@yahoo.com

This is a valid, signed message.



An example signed message Tue, Mar 12, 2013 at 12:01 PM

From Fake User

To doug.mtview@gmail.com

This is a test DKIM-signed message.

Screen shots taken from the Yahoo! web client:
A convincing, albeit fake, header field can be prepended onto DKIM messages displayed to users instead of the signed header fields. This problem exists with Yahoo!, Comcast, Microsoft, and other email providers supporting DKIM. It is possible for DKIM to be corrected to ensure against messages with deceptive header fields being marked as having a valid DKIM signature. SMTP is not to enforce message formats as specified in the second to the last paragraph in RFC5321 Section 3.3. Message enforcement by the transport would thwart message structure evolution.

A valid DKIM signed message as shown, could be issued from 1, 1,000 or 1,000,000 different IP addresses. Each message would have a valid DKIM signature according to the current standard, and allow users to see any imaginable pre-pended From header field. DKIM, as designed, allows any message to be distributed any number of times from any number of sources. As a result, reputations based upon DKIM signatures have highly questionable worth.

If providers had offered users a button to indicate “This Is Spam”, which domain (the signed From:, the forged From:, or the DKIM signing domain) would be identified as having sent the abuse? Clearly, DKIM does not offer requisite domain specific protections for either users or senders when used to establish domain reputation.

Applying reputations against sources making use of StartTLS extended with OCSP (Online Certificates Status Protocol Extensions to IKEv2) could offer the same scalability as that of HTTP while ensuring retention of a domain’s reputation. Otherwise, opportunistic techniques many suggest that might work will be gamed and spoofed extensively.

Looking at a few minutes of spam...

Total spams:	9438
DKIM pass:	688 (about 25% relayed from large ESPs)
DKIM fail:	189
DKIM pass w/multiple from:	28 (about 2% on average)
Unsigned:	8561

In summary, moving to reputations based on domains and away from IP addresses will be essential in coming years. Use of Reverse DNS PTR records is likely to consume excessive resources due to DNS timeouts and caching loads. The use of ARPA name servers (Reverse DNS) as a means to obtain a responsible domain has been problematic with IPv4, and will be even more so with IPv6. IPv6 is designed to accommodate rapid reassignment, where network providers are ill equipped at offering meaningful PTR records in the ARPA zone.

DKIM does not offer a suitable basis by which traffic can be fairly managed, so change to email will be required. To be useful, domain identifiers MUST represent those accountable for traffic between source and destination. As such, it is likely hop-by-hop assessments will be required. Such assessments are best done using cryptographic techniques like DANE, but could make use of a lighter weight APL Resource Records able to define the entirety of the address space involved within a single DNS transaction.