1

Getting the Most Out of ORF

Preface

WHAT IS THIS GUIDE ABOUT?

This documentation provides ORF configuration suggestions on how to get the most out of ORF - the highest spam filtering performance with the lowest number of false positives.

WHO SHOULD READ THIS?

The guide targets readers who already spent a little time discovering ORF and are now looking for ways to fine-tune the configuration. The format is intentionally brief, so that you can quickly read it. Some sections are more comprehensive where deeper understanding of the content is required.

Getting the most out of ORF

GETTING THE INTERMEDIATE HOST LIST RIGHT

Among with DNS, the Intermediate Host List (IHL) is the most important setting of ORF which affects spam filtering performance and reliability. Having incorrect IHL settings may result in significantly lower spam filtering performance or losing legitimate email.

What does the IHL do that makes it so important?
As you might know, ORF has a bunch of tests which depend on the sender's actual IP address (e.g. DNS blacklists or the SPF test). When ORF is filtering on your system's perimeter, it is easy to tell the sender IP address: it is the IP address of the client which is sending the email. However, when the email arrives via your backup email exchanger (MX), this is clearly wrong: the client is your backup email exchanger and not the original, actual sender. It is the same problem with SMTP anti-virus proxies, active firewalls or SMTP front-ends.

So what can ORF do? Fortunately, email delivery hop history is recorded in the email header, so ORF can "look behind" of your backup email exchanger or other SMTP front-ends (referred as "Intermediate Hosts" in ORF) and find out the actual sender IP address: all you have to do is to tell ORF what your Intermediate Hosts are.

What should I add to my Intermediate Host List exactly?
Add the IP address of every host which relays for the ORF server to the list. These typically are:

- backup MXs
- front-end servers in the DMZ (when ORF runs on the back-end)
- non-transparent firewalls

Note that ORF's Intermediate Host List treats Class A, B and C private intranet address ranges as they were on the list, you do not have to add them manually. These address ranges are: *10.0.0.0 – 10.255.255.255* (Class A), *172.16.0.0 – 172.31.0.0* (Class B) and *192.168.0.0 – 192.168.255.0* (Class C).

Why should I care?
Let us assume that you have a backup MX, which is not listed on the IHL. It might sound quite innocent, but, in fact, it is quite dangerous. Remember: ORF thinks the actual sender is your backup MX. Now let us see the some scary situations.

SPF test enabled: Losing legitimate email
The email is coming from user@example.org via your backup MX. ORF looks up the SPF policy of *example.org* to check whether the sender IP is allowed to send in the name of *example.org*. Not surprisingly, *example.org* says that your backup MX has nothing to do with their email servers, so ORF takes action: you have lost a perfectly legitimate email.

Greylisting test enabled: Email delay
Spam is relayed via your backup MX (spammers often do this). The Greylisting test encounters an unknown "triplet", so the email gets temporarily rejected. Spammer software is not expected to retry the delivery, but legitimate servers will do, just like your backup MX, which retries in e.g. 15 minutes and this time passes the Greylisting test. Although other spam tests may stop the email, you potentially lose 15 minutes and cause extra headache for your backup MX.

DNS blacklists enabled: Missing spam
Spam is relayed via your backup MX again. ORF checks your backup MX's IP address in the DNS blacklists and finds that your backup MX is not a spam source. Although other spam tests may stop the email, you potentially let an unwanted email in.

GETTING THE DNS SETTINGS RIGHT

DNS is an often overlooked, yet very important setting of ORF. Problems with DNS settings may result in significantly degraded spam filtering effectiveness and/or in losing legitimate emails. The most effective spam filtering features of ORF's rely on DNS: DNS blacklists, SURBLs and the SPF Test all use DNS to check the email.

Getting the server list right
ORF requires the DNS server to be able to perform so-called *recursive lookups*. A DNS server which supports recursive lookups will do everything to get the requested DNS data for the client. The non-recursive server will tell the client (ORF in this case) to go and bother other servers, if the server is not authoritative for the requested zone. ISP primary DNS servers often support recursive lookups, but secondary DNS servers rarely do. What is worse, ISPs often change their mind and disable recursive lookups on their servers without prior notice.

Due to the above, it is *highly recommended* to use a local DNS server for ORF lookups. Most organizations already have a DNS server in place and have full control over that server, so you it is guaranteed that recursive lookups are supported and will be supported in the future.

You can check the health of your DNS servers using the *Test DNS* button on the *Configuration / Global / DNS and lookups* page of the ORF Administration Tool.

Typically, you do not need more than 2 servers
We recommend having only one or two DNS servers specified for ORF and at least one of them should be local (see above). Having too many servers may result in losing email, as described in the *"Why should I care?"* section below.

Be conservative with the DNS timeout
The default 12 seconds should be enough for any DNS lookups. Having a too high timeout, especially when combined with multiple DNS servers, may result in losing email: again, see the *"Why should I care?"* section.

Why should I care?
First, because if the DNS server does not support recursive lookups, ORF will think that the requested record was not found. While usually this will only cause lower spam filtering performance (because ORF actually never reaches the DNS blacklist or the SURBL), but when the Reverse DNS Test is turned on, the lack of recursion support will result in blacklisting of all incoming emails.

> **Reverse DNS Test turned on: Losing *all* email**
> The RDNS Test checks whether the sender's A or MX DNS records exist (in other words, whether an email response could be sent to the sender). If the email fails on this test, ORF will blacklist it. Consequently, you will lose all non-whitelisted email. Remember, if you are using the ISP DNS servers, you put yourself at risk of this.

> **High timeout problem: Losing emails**
> Assume that you use 2 of your ISP's DNS servers in ORF. Let us add that you got bored with the occasional DNS timeout errors and raised the timeout limit to *60* seconds from the default 12. A few days later, your ISP changes the DNS server IPs. What is going to happen? ORF will try each DNS lookup twice, because when the lookup using the first server times out, ORF switches to the second and repeats the lookup (fail-over). Each DNS operation will take 120 seconds. If you have 5 DNS blacklists and 2 SURBLs, checking the email might take 120 * 7 = 840 seconds = 14 minutes. ORF is a protocol-level filter, which holds up the email delivery until tests are done and 14 minutes is more than enough for the less patient SMTP clients to give up the delivery. This means that you will lose the most, if not all, incoming email.

SELECTING THE RIGHT FILTERING POINTS

Selecting the right filtering point for the ORF tests depends on both - your intentions, and your email delivery setup: what you want to do with blacklisted email and what you can do.

Intention questions
Do you want to keep the blacklisted email? If not, do you care if spam is not stopped at the earliest possible stage of the delivery? If your answer is 'yes' to any of these questions, the answer is simple: assign all possible tests to the On Arrival filtering point and you can skip reading this section. However, if you want to reject spam at the earliest stage, read on.

Email delivery setup questions
Basically, there are two email delivery setup elements that affect the filtering point selection: front-ends and backup MXs (also called "backup email servers").

The front-end is a host which receives the email before ORF would filter the email. The front-end is not necessarily a separate box, it may be e.g. an anti-virus proxy on the same server where ORF runs. You can easily tell whether you have a front-end: look at the ORF logs and if all emails seem to have arrived from the same host, then you have a front-end.

The backup MX is another server which receives emails for your domain when your server (the primary MX) is down. Spammers often send to the secondary MX directly to bypass the spam filtering software, which typically runs on the primary MX only and trusts the secondary MX.

The following table summarizes the reasonable filtering point selections for the various email delivery setups.

Case	Setup	Before Arrival	On Arrival	Both
A	Direct Delivery (no front-end or secondary MX)	Yes	Yes	Yes
B	No front-end, but there is a secondary MX	No	Yes	Yes
C	Front-end	No	Yes	No

- **Case A)** You receive emails directly from the Internet, so you can take advantage of ORF's full filtering capabilities.

- **Case B)** Your secondary MX address has to be on the Intermediate Host List. Due to this, emails from your secondary MX will be whitelisted at the Before Arrival filtering point. This would result in a decreased filtering effectiveness, but you can work this around by assigning all tests to the On Arrival filtering point. If you prefer to stop emails as soon as possible, assign tests to Both filtering points.

- **Case C)** The front-end address has to be on the Intermediate Host List and due to this, all email will be whitelisted at the Before Arrival filtering point. The only reasonable selection is the On Arrival filtering point in this case.

Sample Filtering Point Assignments I.
The table below shows an optimal filtering point selection for the following case:

- Setup: no front-end or secondary MX
- Intention: drop blacklisted emails

Test	Before Arrival	On Arrival	Both
DNS whitelist	•		
Auto Sender Whitelist	•		
Reverse DNS	•		
DNS blacklists	•		
HELO domain blacklist	•		
SPF test	•		

IP blacklist	•		6
Sender blacklist	•		
Recipient blacklist	•		
Recipient validation	•		
Tarpit Delay	•		

Sample Filtering Point Assignments II.
The table below shows an optimal filtering point selection for the following case:

- Setup: No front-end, but there is a secondary MX
- Intention: drop blacklisted emails

Test	Before Arrival	On Arrival	Both
DNS whitelist			•
Auto Sender Whitelist			•
Reverse DNS			•
DNS blacklists			•
HELO domain blacklist			•
SPF test			•
IP blacklist			•
Sender blacklist			•
Recipient blacklist			•
Recipient validation			•
Tarpit Delay			•

Sample Filtering Point Assignments III.
The table below shows an optimal filtering point selection for the following case:

- Setup: Front-end (and maybe a secondary MX in addition to the front-end)
- Intention: any

Test	Before Arrival	On Arrival	Both
DNS whitelist		•	
Auto Sender Whitelist		•	
Reverse DNS		•	
DNS blacklists		•	
HELO domain blacklist		•	

SPF test		•	
IP blacklist		•	
Sender blacklist		•	
Recipient blacklist		•	
Recipient validation		•	
~~Tarpit Delay~~ *			

* Enabling the tarpit delay makes no sense when you have a front-end, because this would "punish" your own front-end only.

SELECTING TESTS

ORF offers a number of tests which help to catch spam (e.g. DNS blacklists) and another bunch of tests which are not explicitly anti-spam, such as the Attachment Filtering and even not very test-like things such as Tarpit Delay. Of course, to get the highest spam performance out of ORF, you should have all tests enabled, however some tests may not work in your organization or delivery setup, e.g. you may not accept the delivery delay caused by Greylisting or you cannot use it, because you have a front-end server.

To help your test selection, the table below summarizes the benefits and risks of the ORF anti-spam tests.

Test	Effectiveness	Risk
Reverse DNS	Fair	None / Low
DNS Blacklists	Excellent	Low / Moderate
HELO Domain Blacklist	Fair	Low / Moderate
SPF Test	Fair	None / Low
Greylisting	Excellent	Moderate
Keyword Blacklist	Fair	None / Low
URL Domain Blacklist	Excellent	Low
Charset Blacklist	Fair	Low
Honeypot	Fair	Low
DHA protection	Fair	Low / Moderate

Risks detailed

- The **Reverse DNS Test** is virtually risk-free, it will never blacklist legitimate email, except if your DNS server stops performing recursive lookups (see the "Getting the DNS settings right" section). With the Sender IP validation turned on, you may get a higher false positive rate, as this subtest is more for policy and less for spam filtering.

- **DNS Blacklists** are operated by humans, so they always have the risk of false positives, however this risk is low if you select the blacklist carefully (see the "Selecting the Best DNS

blacklists" sections). Some blacklists, like BLARS, are known to be very aggressive.

- The **HELO Domain Blacklist** risk level depends on what features you turn on; if you turn on *"Blacklist if HELO/EHLO domain... is the same as the recipient domain"*, you can expect no false positives. The *"...is malformed"* check might catch a few legitimate emails, because some admins set up domains with underscore ("_") in the domain name, which is not legal in DNS host names. The *"...is not a Fully Qualified Domain Name (FQDN)"* check will surely catch a few legitimate emails, although FQDNs are required by the standards.

- With the **SPF Test** you can expect very few or no false positives if your Intermediate Host List settings are correct. However, if the IHL is not correct, you will get lots of false positives.

- **Greylisting** can cause legitimate email to arrive with much more delay than the expected 5-15 minutes, if the email was sent by a large organization (e.g. *gmail.com*), which has many outgoing servers and repeated email delivery attempts are made from different servers. You can reduce the chance for such delays by importing the Greylisting IP exception list from http://www.vamsoft.com/greyexcept.asp and by using the Automatic Sender Whitelist.

- The **Keyword Blacklist** has only as much risk as the keyword filters have: if they are carefully constructed, they will catch no legitimate emails.

- The **URL Domain Blacklist** and SURBLs are very reliable, because they list spamvertized domains that are collected manually.

- The **Charset Blacklist** can be used to detect spam written in specific languages/scripts. If your business focuses on domestic activities, the chance of false positives is is very low, otherwise you should carefully consider what languages/scripts you can afford blacklist.

- The effectiveness of the **Honeypot test** depends on many factors: if you receive spam sent to both non-existent and valid recipients often, you may find this test quite effective (if your published non-existent honeypot addresses are among these targeted recipient addresses: see our article regarding this at http://www.vamsoft.com/howto-publish-honeypots.asp)

- Spammers often perform Directory Harvest Attacks (DHAs) to discover valid email addresses in your domain. The **DHA Protection test** can be used to detect and stop some of these attacks.

Other tests of ORF are not explicitly anti-spam, these are the IP, Sender and Recipient Blacklists, the Recipient Validation, Tarpit Delay and the Attachment filtering. Basically, you can have all these enabled if you use them, there are virtually no risks.

SELECTING THE BEST DNS BLACKLISTS

Selecting DNS blacklists can be tricky, because what you consider to be a good DNS blacklist largely depends on your requirements: some might not tolerate false positives at all, but others may accept a few false positives for a higher spam catch rate.

It is worth checking out what others use at http://www.vamsoft.com/stats.asp, because this represents blacklist selection of users with various intentions and tolerance for false positives, thus this list is more or less reliable.

The *POP%* column tells you how popular a DNS blacklist is, i.e. how many ORF users use it. But do not rely solely on the popularity, check the spam/legitimate (*S/L%* column) ratio as well, which tells you how many percents of checks were positive (e.g. 25% means that every 4th check was positive). For instance, ORDB (before it was shut down in December 2006) had very high popularity, in spite of the fact that it caught virtually no emails.

At the time of writing this, we recommend the following DNS blacklists (in the following order):

- Barracuda Reputation Blocklist (BRBL)*
- Spamhaus ZEN
- Spamcop
- SORBS with the 127.0.0.6 action disabled
  (Select SORBS / Modify / SMTP Actions tab / Uncheck 127.0.0.6 / Click OK)
- Not Just Another Bogus List (NJABL Combined List)

* requires registration: http://barracudacentral.org/account/register

"CBL Composite Blocking List" could be a good choice if you do not use the Spamhaus zones (these already incorporate CBL data).

Before you select a DNS blacklist, read its description (double-click on the blacklist) – the DNS blacklist may incorporate data from another blacklists, like Spamhaus ZEN incorporates Spamhaus SBL, Spamhaus XBL and Spamhaus PBL data, which in turn incorporates the CBL Composite Blocking List data.

If any of the above DNS Blacklist definitions is missing from your ORF configuration (because you have an earlier version), you should download a more recent definition file from

http://vamsoft.com/dl.aspx?dnsbl-090416.xml

To import the new definition file:

1) Download the latest DNS Blacklist definition set from the link above (XML file)
2) Start the ORF Administration Tool
3) Select Configuration / Tests / DNS Blacklists from the navigation pane on the left
4) Right-click on the list and select Import Blacklist Definitions
5) Select the XML file that you downloaded in step 1)

Now you have two choices:

A) If you want to import only some of the new blacklist (e.g. Barracuda), select the new lists to be imported, uncheck "Delete current definitions not listed here (full overwrite)" and Click OK.

B) If you would like to update the entire definition set, select all lists, check "Delete current definitions not listed here (full overwrite)" and Click OK.

The latter choice is recommended (unless you have some custom definitions you added manually), because it will wipe out old DNS Blacklists which no longer operate (ORDB, DSBL, Blackholes.us etc) resulting in an up-to-date definition set.

You can also use our Country Blacklist Generator service at http://www.vamsoft.com/countries-nerd.asp to generate a DNS Blacklist definition for the countries.nerd.dk DNS blacklist, which allows blacklisting emails based on their source country. If you choose to enable this geographic blocklist, it should be the first in the list.

We do not recommended using more than 3-5 DNS blacklists at once.

SELECTING THE BEST URL DOMAIN BLACKLISTS

This is much simpler than DNS blacklist selection, because there are only two major URL domain blacklists providers, surbl.org and uribl.com. We recommend selecting the following URL domain blacklists, accordingly:

- SURBL: Combined SURBL list
- uribl.com Blacklist

Please consider that the *SURBL: Combined SURBL blacklist* combines data from the rest of the surbl.org zones, so if you decide to activate it, turn off the other URL blacklists starting with the *"SURBL:"* prefix.

USING THE MANUAL BLACKLISTS RIGHT

ORF offers a number of built-in automatic tests which do not require extra configuration once you set them up, such as the DNS Blacklists, URL Domain Blacklists or the Greylisting Test. There are also tests which allow you to specify various criteria for filtering: tests like the IP Blacklist or the Sender Blacklist. These later are called *manual blacklists*.

There appears to be a common misconception about the role of these manual blacklists. In many cases, we have seen extremely long lists of IP addresses on the IP Blacklist and the same long list of email addresses on the Sender Blacklist, which were supposed to stop spam. Well, they might help a little, but you would waste *much* more time and effort than it is worth.

Manual blacklists were designed to stop constant abuse. For instance, if you are getting a major virus hit from a specific IP or network, the IP Blacklist allows you to stop that IP or network from abusing you. If the virus or a long-running spam campaign has a specific sender domain or pattern, you can reject the domain or pattern by the Sender Blacklist. Similarly, you can ban specific words or phrases, such as offensive expressions by the Keyword Blacklist, or catch specific email worm patterns.

We suggest relying on the automated tests of ORF (like as DNSBLs or SURBLs) as much as possible and using the manual blacklists only when they have to be used. Adding spam sender addresses or domains to the Sender Blacklist takes a lot of time and you get very little results, because spammers use forged sender addresses. You will have the same little success with adding IPs to the

IP Blacklist – your time could be much appreciated, however, if you report the spam source to DNS blacklists instead of adding it to your own IP blacklist.

DNS blacklists, like Spamcop, usually get their spam feed from contributors, so you can register and improve the DNS blacklist databases by submitting spam samples. By submitting the spam sample instead of just blacklisting in your system, you not only get rid of spam, but you also help the thousands of blacklist users worldwide to fight the spammers. SURBLs also accept spam sample submissions.

You should also be conservative when adding entries to the manual whitelists: spammers often spoof legitimate domains, so they will avoid detection if they spoof a domain which you have whitelisted (including yours). You should add individual email addresses only, or you can rely on the Automatic Sender Whitelist Test instead of the manual whitelists.

Due to the above, whitelisting your own domain is not recommended. Moreover, you should consider blacklisting it instead: as ORF whitelists private local intranet addresses, Intermediate Hosts, and you can also configure it to whitelist authenticated clients (Configuration / Global / Miscellaneous), legitimate emails will not be affected by this, but will reject spammers who attempt to spoof your domain name. For more information regarding this, please read our article at http://www.vamsoft.com/howto-blacklist-self-spam.asp#solution-3

BOOSTING FILTERING WITH EXTERNAL AGENTS

External Agents allow you to attach external software to ORF to extend the filtering capabilities of ORF, typically by virus filtering and anti-spam.

At the time of writing this, we have 15 External Agent definitions available for download on our website (http://www.vamsoft.com/agentdefs.asp) for various anti-spam and anti-virus software, which includes a definition for ClamAV, a free and excellent anti-virus. Attaching these software to ORF could significantly boost ORF's filtering performance and allow you to set up a first line of defense against viruses.

You might think that free anti-spam software and anti-virus must have poor performance, but this is not the case with SpamAssassin and ClamAV, although they may need a little more effort to set up and maintain. They cost nothing, but time to set them up, so if you have a few spare minutes, they are worth the try. You may find useful tips and tricks from other users in our newsgroups (for ClamAV, see the "Native ClamAV for windows installation notes" thread).

Some External Agents focus on specific spam trends, which are harder to catch by regular test like DNS and URL blacklists. In response to the recent self-spam attacks (where the sender address seems to be the same as the recipient address) we developed an External Agent.

For detailed information regarding this threat and (and to download the agent definition), please visit

http://www.vamsoft.com/howto-blacklist-self-spam.asp

UNLEASHING THE POWER OF REGEXS

Another common misconception puts an equal sign between Perl and Regular Expressions (regexs), but it is not entirely true. Perl is a powerful scripting language, which offers sophisticated regex support, just as ORF does in many of its features. There are many flavors of regular expressions and ORF uses a flavor which is compatible with Perl 5's regular expression language, that is why ORF refers to it as "Perl-compatible regular expressions" (if you are interested, ORF uses the PCRE engine). Anyhow, although regexs might take a little time to learn, they are nearly not as complex as learning Perl.

If you ask why bother learning regular expressions, consider that regexs can be used in many ways from ORF to Microsoft Word's Search & Replace function and general text processing (grep/egrep).

Regexs come handy in ORF when you have to exceed the capabilities of simple wildcarded expressions, e.g. when you have to define complex email address patterns (e.g. email addresses which contain a sequence of 15 or more digits in the mailbox name, *.\d{15,}.*@.*) or keyword filters (e.g. the word "replica", followed by "for" or "4", followed by "you" or "u", *replica (for|4) (you|u)*).

Our Tips & Tricks newsgroup offers useful tips from user-contributed keyword and attachment filtering expressions to various regex-related resource links, which might help you in fine-tuning ORF according to your needs."

WARNING: It is strongly recommended to test the newly added Regular Expressions (and wildcarded simple text expressions) first to see whether they generate any false positives. We have seen lots of too "wide" regexs which caused legitimate email loss.

KEEP THE MAXIMUM HEADER CHECK DEPTH AT 1

Although you can increase the effectiveness of the DNS blacklist-based filtering by setting this value higher than 1 or by setting unlimited header check depth, you potentially block legitimate emails. (For more Information please read ORF Help).

CHECKLIST

Review the short checklist below to make sure that you configured ORF the best possible way.

ORF checklist

	Are Intermediate Host List Settings correct?
	Do I use local DNS?
	Are tests assigned to the right filtering point?
	Are all reasonable tests enabled?
	Do I use the best DNS blacklists and SURBLs?
	Are my manual filtering expressions safe?
	Do I have the Maximum Header Check Depth set to 1?

FEEDBACK

If you have anything to add to this guide, or you just think it is good or bad, contact us at orf@vamsoft.com and let us know what you think.