Passive Information Gathering

The Analysis of Leaked Network Security Information

Gunter Ollmann, Professional Services Director

Abstract

Most organisations are familiar with Penetration Testing (often abbreviated to, "pentesting") and other ethical hacking techniques as a means to understanding the current security status of their information system assets. Consequently, much of the focus of research, discussion, and practice, has traditionally been placed upon active probing and exploitation of security vulnerabilities. Since this type of active probing involves interacting with the target, it is often easily identifiable with the analysis of firewall and intrusion detection/prevention device (IDS or IPS) log files.

However, too many organisations fail to identify the potential threats from information unintentionally leaked, freely available over the Internet, and not normally identifiable from standard log file analysis. Most critically, an attacker can passively gather this information without ever coming into direct contact with the organisations servers – thus being essentially undetectable.

Very little information has been publicly discussed about arguably one of the least understood, and most significant stages of penetration testing – the process of Passive Information Gathering. This technical paper reviews the processes and techniques related to the discovery of leaked information. It also includes details on both the significance of the leaked information, and steps organisations should take to halt or limit their exposure to this threat.

Information Leakage

Like it or not, every Internet-connected system unintentionally leaks internal information about their organisation which could be used to formulate a targeted attack. Depending upon the source of this leakage, the information may relate to the components used within the organisation's physical asset infrastructure, the management processes utilised, or the operational "personnel" hierarchy.

The significance of this leaked information may not be entirely obvious to some organisations. In the majority of cases, much of the leaked information relates to the topology of the organisations network and the types of services within. This enables an attacker to provisionally map out the network for coordinating more sophisticated attacks at a later date.

In addition, the harvesting of information relating to specific employees (in particular administrative staff) can be actively employed in social engineering attacks – perhaps through deception, bribery or blackmail. Past examples of this include staff blackmailed into providing confidential information due to racial and sexist remarks posted to public newsgroups via their corporate Internet account.

An important aspect of this information leakage is that most of it is publicly available using the Internet, and probably contained on systems unrelated to the organisation. Consequently, access to the information is independent of the organisations resources and thus can effectively be accessed "anonymously" by anyone.

The processes for discovering this leaked information tends to be very simple and a multitude of tools are readily accessible to make it as even easier. In fact, many of the tools are either already built in to the most popular operating systems, or can be freely accessed from various websites.

The techniques used to uncover this leaked information are commonly referred to as "Passive Information Gathering" – and they form a vital (and often overlooked) role in any quality penetration test or security assessment.

Definition of "Passive"

Before delving into the techniques of Passive Information Gathering, it is important to understand what is meant by the term "passive". A dictionary provides two relevant definitions:

pas·sive (adj.)
 1. Receiving or subjected to an action without responding or initiating an action in return.
 2. Accepting or submitting without objection or resistance.

From an ethical hacking perspective, the focus is upon identifying information about the organisation under investigation, without the organisation being aware that the information has been accessed.

In the context of this technical whitepaper, "passive" refers to techniques that either do not connect to a system owned or managed by the organisation (thus they would be unaware of any such access), access to information from the organisations systems which is commonly available and would not normally ever be associated as a precursor to future attacks, or via the increasingly numerous online security analysis websites.

This includes non-intrusive techniques such as searching generic Internet resources like www.google.com for information relating to the organisation, and encompasses analysis of data returned during normal interaction with the organisation – for example the banners and other system messages returned when connection to the web or mail server. However, it does not include intrusive network enumeration phases such as port-scanning.

Passive Information Gathering Techniques

There are a number of techniques and processes available when carrying out a Passive Information Gathering exercise. This technical whitepaper will detail the most relevant techniques and endeavour to elaborate both the thought processes necessary to identify the leaked information, and to evaluate the relative security risks associated with the leakage.

A lot of important information can be passively harvested and subsequently used in a direct attack or to reinforce other attacks targeted at an organisation. Depending upon the source, information such as current service patching levels, internal network architecture layout and account details can be easily obtained. Just as importantly, with a little insight as to where this information is obtained and the level of detail of information, an organisation can often rectify this information leakage simply and quickly.

The most critical phases or investigation processes revolve around the accessibility of various online resources such as:

- Internet Service Registration – The global registration and maintenance of IP address information
- Domain Name System – Local and global registration and maintenance of host naming
- Search Engines – The specialist retrieval of distributed material relating to an organisation or their employees
- Email Systems – The information contained within each email delivery process
- Naming Conventions – The way an organisation encodes or categorises the services their online hosts provide
- Website Analysis – The information intentionally made public, that may pose a risk to security

Internet Service Registration

In order to access networked resources over the Internet, every accessible host must have a unique and routable IP address. Whilst there is slow take-up of IPv6, IPv4 is still the de-facto standard and will be considered in this document. This IP address takes the form of xxx.xxx.xxx.xxx, where "xxx" may be some value between 0 and 255. In order to simplify host addressing, and to make hosts more memorable, services exist that associate IP addresses to a unique domain name (for instance www.abcexample.com).

Both the registration of IP addresses and domain names are coordinated at an international level. In order to administer these IP addresses (or address ranges) and domain names, organisations must supply administrative details including physical billing addresses and technical contact information. By necessity, this information is publicly available and may be requested by anyone over the Internet. Consequently, these international databases may are queried and form the first stages of the Passive Information Gathering exercise.

Dividing the responsibility between them, there are exist four Regional Internet Registries (RIR). These RIRs provide allocation and registration services and support the operation of the Internet globally. Responsibilities include the allocation of Internet (IP) address space, autonomous system numbers (ASNs) and the management of reverse domain name space. The four RIRs are:

- APNIC (Asia-Pacific Network Information Center)
- ARIN (American Registry for Internet Numbers)
- LACNIC (Latin American and Caribbean Internet Addresses Registry)
- RIPE NCC (Réseaux IP Européens Network Coordination Centre)

WHOIS

There are two primary WHOIS resources – Network service-based and Name service-based. As the names suggest, one focuses upon the registration and management of individual/blocks of IP addresses (note that a block or range of IP addresses is commonly referred to as a "netblock"), while the other focuses upon the registration and management of domain names.

The core RIR WHOIS services provide a mechanism for finding contact and registration information for resources registered with the individual registries. These databases contain IP addresses, autonomous system (AS) numbers, organisations or customers that are associated with these resources, and related Points of Contact (POC).

The RIR WHOIS services will typically not locate any domain-related information or any information relating to military networks. For domain name lookups, various online resources exist – however

www.internic.net/whois.html and www.uwhois.com are recommended for non-military, and whois.nic.mil for military network information.

Network Service-Based WHOIS

Network service-based WHOIS data provides details of network management data. The data is commonly used by network and security personnel to reach, or otherwise contact, an organisations technical staff should a problem arise. This registration and administration data include information such as the contact provider of the network numbers, and in some situations the company leasing the address space.

Worked Example – One

Let us consider the investigation of the example company ABC Example Ltd., and start with the primary web site www.example.com. A quick lookup of the host name (via nslookup or even ping) reveals the unique IP address to be 213.48.xx.45[1]. In this instance, we note that the IP address is part of a Class A address range (i.e. 213.xxx.xxx.xxx) normally associated with Western Europe[2].

As this is a European IP address range, we should direct any network service WHOIS queries to RIPE NCC, e.g. whois -h whois.ripe.net <query string>

```
1.      % This is the RIPE Whois server.
2.      % The objects are in RPSL format.
3.      %
4.      % Rights restricted by copyright.
5.      % See http://www.ripe.net/ripencc/pub-services/db/copyright.html

6.      inetnum:      213.48.xx.0 - 213.48.xx.63
7.      netname:      ABC-EXAMPLE
8.      descr:        INTERNET
9.      country:      GB
10.     admin-c:      TWIP1-RIPE
11.     tech-c:       TWIP2-RIPE
12.     status:       ASSIGNED PA
13.     mnt-by:       AS546x-MNT
14.     remarks:      report abuse to abuse@blueyonder.co.uk
15.     remarks:      All reports via other channels will be ignored.
16.     changed:      ripe-admin@blueyonder.co.uk 20020709
17.     source:       RIPE

18.     route:        213.48.0.0/16
19.     descr:        Telewest Broadband
20.     descr:        UK Broadband ISP
21.     origin:       AS5462x
22.     notify:       ripe@telewest.net
23.     mnt-by:       AS546x-MNT
24.     remarks:      report abuse to abuse@blueyonder.co.uk
25.     remarks:      All reports via other channels will be ignored.
26.     changed:      ripe-admin@blueyonder.co.uk 20020709
27.     source:       RIPE

28.     role:         Telewest Broadband IP Network Services
29.     address:      Genesis Business Park
30.     address:      Albert Drive
31.     address:      Woking
32.     address:      Surrey UK
```

---

[1] Note that "XX" has been used to replace potentially confidential information in this obscured example.

[2] This observation is based upon experience. However, many of the tools commonly used to query these registration databases will automatically select the most appropriate regional authority.

```
33.     address:        GU21 5RW
34.     e-mail:         ripe@telewest.net
35.     admin-c:        JH1542x-RIPE
36.     tech-c:         AH1530x-RIPE
37.     tech-c:         DR1307x-RIPE
38.     tech-c:         DS1550x-RIPE
39.     tech-c:         KJ2418-RIPE
40.     tech-c:         MG645-RIPE
41.     tech-c:         SA3620-RIPE
42.     tech-c:         SB5110-RIPE
43.     tech-c:         SL3595-RIPE
44.     nic-hdl:        TWIP1-RIPE
45.     notify:         ripe@telewest.net
46.     mnt-by:         AS546x-MNT
47.     changed:        ripe@telewest.net 20030820
48.     source:         RIPE

49.     role:           Telewest Broadband Business Internet
50.     address:        Genesis Business Park
51.     address:        Albert Drive
52.     address:        Woking
53.     address:        Surrey UK
54.     address:        GU21 5RW
55.     e-mail:         ripe@telewest.net
56.     admin-c:        JH1542x-RIPE
57.     tech-c:         LM5500x-RIPE
58.     tech-c:         MB3445x-RIPE
59.     nic-hdl:        TWIP2xx-RIPE
60.     notify:         ripe@telewest.net
61.     mnt-by:         as546x-mnt
62.     changed:        jim.haffey@telewest.net 20030103
63.     source:         RIPE
```

Example 1: WHOIS query on 213.48.14.45

Observations:
- Firstly, we note that the Class B IP address range (CIDR 213.48.0.0/16) is managed by Telewest Broadband – [lines 18 & 19]
- The Telewest Broadband ISP is UK based and appears to have two business roles; "Telewest Broadband IP Network Service" and "Telewest Broadband Business Internet" – [lines 19-20, 28 and 49]
- The ISP's real postal address details have been supplied – [details contained on lines 29-33 and 50-54]
- The web servers IP address (213.48.xx.45) is just one of 64 that have been allocated to the identifier "ABC-EXAMPLE" – [lines 6 & 7]
- The registrations for "ABC-EXAMPLE" and "Telewest Broadband" were last changed by "ripe-admin@blueyonder.co.uk" on the same date – [lines 16 & 26]
- The same MNTNER is responsible for all records (AS546x-MNT) – [lines 13, 23, 46 & 61]

From these observations we can conclude that ABC Example Ltd. has an IP address pool of 64 IP addresses (CIDR 213.48.14.0/26) that is allocated and managed by Telewest Broadband. Investigation of Telewest Broadband suggests that ABC Example Ltd.subscribes to their South-East England based business broadband offering, and must therefore host their web-site locally.

We also identified that the same maintainer (AS546x-MNT) is responsible for both Telewest Broadband and ABC Example's netblock registrations. By carrying out an appropriate WHOIS query on the maintainer (querying a MNTNER record), we get the following response from whois.ripe.net:

```
1.      % This is the RIPE Whois server.
2.      % The objects are in RPSL format.
3.      %
4.      % Rights restricted by copyright.
5.      % See http://www.ripe.net/ripencc/pub-services/db/copyright.html

6.      mntner:       AS546x-MNT
7.      descr:        Telewest Broadband;
8.      descr:        UK Broadband ISP
9.      descr:        report to abuse@blueyonder.co.uk
10.     descr:        All reports via other channels will be ignored
11.     admin-c:      JH1542x-RIPE
12.     tech-c:       MG645xx-RIPE
13.     tech-c:       SB511xx-RIPE
14.     upd-to:       ripe@telewest.net
15.     auth:         MD5-PW $1$p1Pjw.8d$yDpgtNz2KPpVoSArZaMsC/
16.     mnt-by:       AS546x-MNT
17.     mnt-nfy:      ripe@telewest.net
18.     referral-by:  RIPE-DBM-MNT
19.     changed:      xxx.haffey@telewest.net 20020613
20.     source:       RIPE

21.     person:       XXX Haffey
22.     address:      Telewest Broadband
23.     address:      Genesis Business Park
24.     address:      Albert Drive
25.     address:      Woking
26.     address:      UK
27.     phone:        +44 (0)1483 xx2542
28.     e-mail:       xxx.haffey@telewest.net
29.     notify:       xxx.haffey@telewest.net
30.     nic-hdl:      JH1542x-RIPE
31.     changed:      xxx.haffey@telewest.net 20020709
32.     source:       RIPE

33.     person:       XXX Brocklebank
34.     address:      Telewest
35.     address:      Genesis Business Park
36.     address:      Woking
37.     address:      Surrey
38.     phone:        +44 1483 750 900
39.     e-mail:       xxx@cableinet.net
40.     nic-hdl:      SB511x-RIPE
41.     notify:       xxx@cableinet.co.uk
42.     changed:      xxx@cableinet.net 19990908
43.     source:       RIPE

44.     person:       XXX Garrett
45.     address:      Telewest Communications (Cable Internet)
46.     address:      Genesis Busines Park
47.     address:      Woking, Surrey
48.     address:      GU21 5RW
49.     phone:        +44 1483 xx6796
50.     fax-no:       +44 1483 xx1 810
51.     e-mail:       xxx@cableinet.net
52.     nic-hdl:      MG64x-RIPE
53.     changed:      xxx@cableinet.net 20010426
54.     source:       RIPE
```

Example 2: WHOIS query on the MNTNR object AS5462-MNT

Observations:

- Firstly, we note that there are three individuals that are nominated maintainers within the RIPE database for the handle AS5462-MNT. These include "XXX Haffey", "XXX Brocklebank" and "XXX Garrett" – [lines 21, 33 & 44]
- We note that various email addresses for these individuals have been supplied belonging to the domains telewest.net, cableinet.co.uk and cableinet.net – [lines 19, 28, 29, 31, 39, 41, 42, 51 & 53]
- It appears that real, and possibly direct, phone numbers have been supplied for these individuals – [lines 27, 38, 49 & 50]

- In order to change any entries within the RIPE database, any communications from these individuals must include a password. This password validation is referenced on [line 15] and presented in a MD5 hashed format.

From a security perspective, a lot of unnecessary information has been leaked by the ISP. The inclusion of real names – combined with email addresses and contact phone numbers – may all be used with great effect in social engineering attacks. Wherever possible, generic role-based names and contact email addresses should be used. In addition, security best practices would recommend that a non-office postal address, and single "reception" phone number be used.

It is also important to note that, should anyone wish to impersonate one of the authorised netblock maintainers, they would need to supply a password. In this example, the use of an MD5 hash is a strong security choice. The following section briefly covers the significance of this maintenance authentication process.

NETBLOCK Registration Maintenance[3]

Netblock registration maintenance is normally carried out in a secure and controlled manner. Authorised maintainers are required to authenticate themselves before changes can be made. Authorisation is the determination of whether a transaction passing a specific authentication check is allowed to perform a given operation.

Different portions of the RIR databases require different levels of protection. Some applications require authentication based on strong encryption. In other cases strong encryption may not be necessary or may not be legally available. For this reason, multiple authentication methods are supported by the servers.

The MNTNER objects serve as a container to hold the authentication filters. A reference to a maintainer within another object defines authorisation to perform operations on the object or on a set of related objects. Such reference is provided by means of the "mnt-by:", "mnt-lower:", "mnt-routes:" and "mbrs-by-ref:" attributes.

The maintainer contains one or more "auth:" attributes. Each "auth:" attribute begins with a keyword identifying the authentication method followed by the authentication information needed to enforce that method.

When submitting an update that requires authorisation from a MNTNER, authentication information valid for that MNTNER must be supplied. Different methods require different authentication information, as shown below.

---

[3] This information was adapted from the FAQ located on the www.ripe.net site
(http://www.ripe.net/ripe/docs/databaseref-manual.html#3.6.1)

Method	Description
NONE	No authorisation checks are performed.
MAIL-FROM	This is a very weak authentication check and is discouraged. The authentication information is a regular expression over ASCII characters. The maintainer is authenticated if the "From:" field in RFC 822 mail headers are matched by this regular expression. Since mail header forgery is quite easy, this is a very weak form of authentication. (MAIL-FROM authentication scheme is not valid in the RIPE Database since 11 July 2002.)
CRYPT-PW	This is a relatively weak form of authentication. The authentication information is a fixed encrypted password in UNIX crypt format. The maintainer is authenticated if the transaction contains the clear text password of the maintainer. Since the password is in clear text in transactions, it can be captured by sniffing. Since the encrypted form of the password is exposed, it is subject to password guessing attacks. Authentication information is supplied using a "password:" pseudo-attribute. The value of this attribute is a clear-text password. It can appear anywhere in the body of the message, but not within mail headers. Line continuation is not allowed for this attribute.
MD5-PW	This scheme is based on the MD5 hash algorithm and provides stronger authentication than CRYP-PW. The authentication information stored in the database is a pass phrase encrypted using md5-crypt algorithm, which is a concatenation of the "1" string, the salt, and the 128-bit hash output. Because it uses 8-character salt and an almost unlimited pass phrase, this scheme is more stable against dictionary attacks. However, since the encrypted form is exposed it cannot be considered as a strong form of authentication. Authentication information is supplied using a "password:" pseudo-attribute. The value of this attribute is a clear-text pass phrase. It can appear anywhere in the body of the message, but not within mail headers. Line continuation is not allowed for this attribute, the attribute and the pass phrase should fit on one line. If the key gets split across multiple lines this will be treated as a syntax error.
PGPKEY	Strong form of authentication. The authentication information is a signature identity pointing to a public key certificate, which is stored in a separate object. The maintainer is authenticated if the transaction is signed by the corresponding private key. The RIPE NCC does not guarantee that a key belongs to any specific entity; it is not a certificate authority. Anyone can supply any public keys with any ownership information to the database and these keys can be used to protect other objects by checking that the update comes from someone who knows the corresponding secret key.

Table 1: MNTNER Authorisation Methods

Following is an example from a poorly secured MNTNER object. In this instance the maintainer has chosen to implement the less secure CRYPT encoding method to protect their administrative password, and also selected a very short password length (see line 7 below).

```
1.      mntner:       ABCEXAMPLE-MNT
2.      descr:        ABC Example Maintainer Object
3.      admin-c:      DC496x-RIPE
4.      tech-c:       DC496x-RIPE
5.      upd-to:       david@example.net.uk
6.      mnt-nfy:      david@example.net.uk
7.      auth:         CRYPT-PW haNlLfDHxARrM
8.      notify:       david@example.net.uk
9.      mnt-by:       ABCEXAMPLE-MNT
10.     referral-by:  RIPE-DBM-MNT
11.     changed:      hostmaster@ripe.net 20010706
12.     source:       RIPE
```

Example 3: WHOIS query on a poorly secure MNTNR object

It is important to note that CRYPT encoded passwords can be brute-forced. Thus organisations should be wary of recycling this type of disclosed password on other systems. In the past, organisations have been caught out when using the same password for managing domain registration details and also for accessing the management interface on their border routers.

Name Service-Based WHOIS

Name service-based WHOIS data provides a number of details about a domain. These details include the registrant of the domain, the street address the domain is registered to, and a contact number for the registrant. This data is supplied to facilitate the communication between domain owners in the event of a problem. This is the ideal method of handling problems that arise, although these days the trend seems to be whining to the upstream provider about a problem first (which is extremely bad netiquette).

Domain registration details tend to be more fragmented than Network service-based registration. Consequently there are many more authoritative systems managing domain names. The root domain servers (i.e. those associated with .com, .net, .org, .mil, .uk, .au, etc.) will often refer to other more authoritative WHOIS servers – which should be queried for full registration details. However, many online WHOIS services and downloadable tools will automatically locate and query the authoritative service. One such example is Network Solutions.
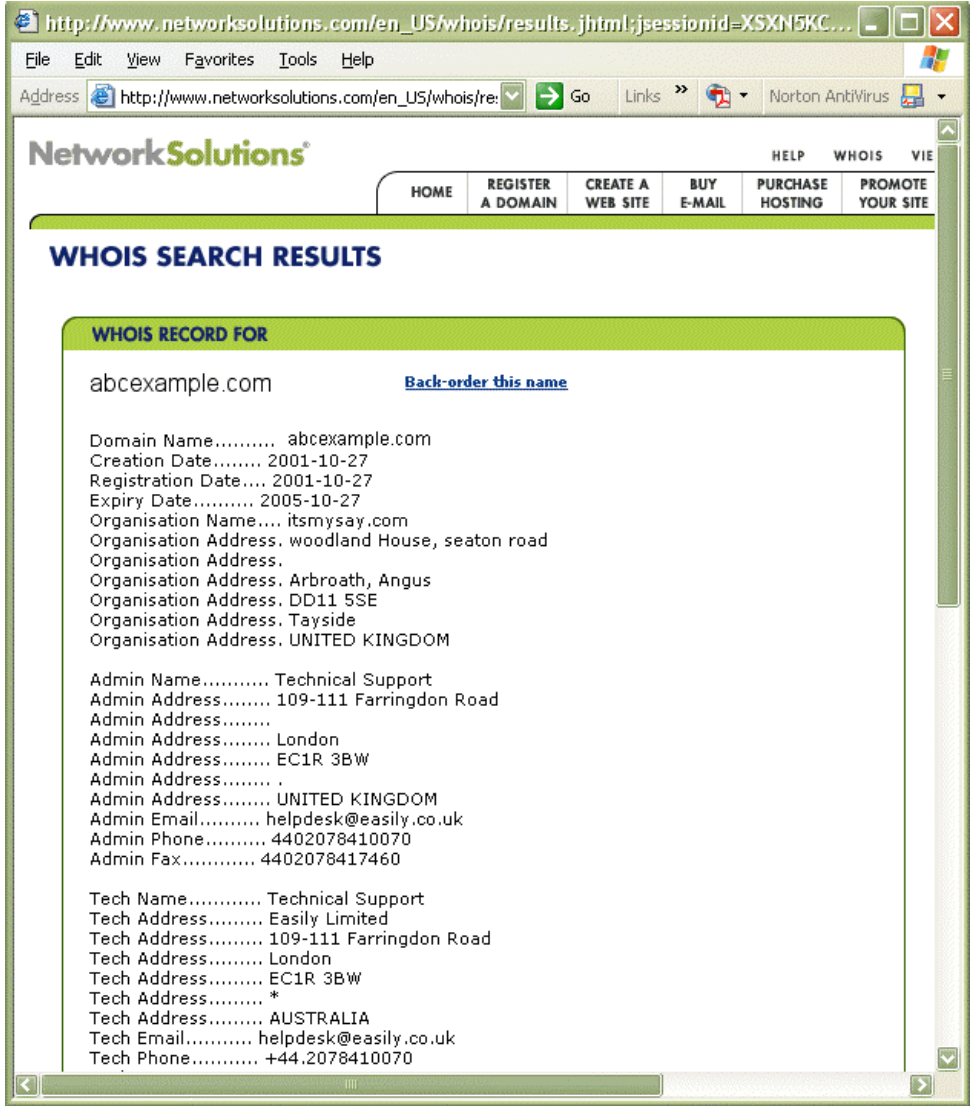
Worked Example – Two

Let us continue the investigation of ABC Example Ltd. Utilising a name service-based WHOIS query facility hosted by the main authority on .com name registrations (www.internic.net/whois.html), we query the domain "abcexample.com" and receive the following response:

```
1.     Domain names in the .com and .net domains can now be registered
2.     with many different competing registrars. Go to http://www.internic.net
3.     for detailed information.

4.     Domain Name: ABCEXAMPLE.COM
5.     Registrar: MELBOURNE IT, LTD. D/B/A INTERNET NAMES WORLDWIDE
6.     Whois Server: whois.melbourneit.com
7.     Referral URL: http://www.melbourneit.com
8.     Name Server: NS0.BYWORKWISE.COM
9.     Name Server: NS1.BYWORKWISE.COM
10.    Status: ACTIVE
11.    Updated Date: 22-sep-2003
12.    Creation Date: 26-oct-2001
13.    Expiration Date: 26-oct-2005


14.    >>> Last update of whois database: Fri, 7 Nov 2003 06:26:27 EST <<<
```

Example 4: WHOIS domain query response about abcexample.com from the root .com server

Observations:

- The domain "abcexample.com" has been registered with an entity called "Melbourne IT, Ltd" – [line 5]
- Melbourne IT maintains their own authoritative WHOIS service for their customers. Any queries should be conducted against "whois.melbourneit.com" – [line 6]
- The abcexample.com domain was initially registered (i.e. created) on the 26th October 2001, and must be renewed before the 26th October 2005.

In order to get more information about the domain registration of abcexample.com we must query the managing WHOIS server – whois.melbourneit.com. The response is as follows:

```
1.      Domain Name abcexample.com
2.      Creation Date 2001-10-27
3.      Registration Date 2001-10-27
4.      Expiry Date 2005-10-27
5.      Organisation Name itsmysay.com
6.      Organisation Address woodland House, seaton road
7.      Arbroath, Angus
8.      DD11 5SE
9.      Tayside
10.     UNITED KINGDOM
11.     Admin Name Technical Support
12.     Admin Address Farringdon Road
13.     London
14.     EC1R 3BW
15.     .
16.     UNITED KINGDOM
17.     Admin Email helpdesk@easily.co.uk
18.     Admin Phone 44020784100xx
19.     Admin Fax 44020784174xx
20.     Tech Name Technical Support
21.     Tech Address Easily Limited
22.     Farringdon Road
23.     London
24.     EC1R 3BW
25.     *
26.     AUSTRALIA
27.     Tech Email helpdesk@easily.co.uk
28.     Tech Phone +44.2078410070
29.     Tech Fax +44.2078417460
30.     Name Server     ns0.byworkwise.com
31.                     ns1.byworkwise.com
```

Example 5: WHOIS domain query response about abcexample.com from whois.melbourneit.com

Observations:

- Both the creation and expiry dates appear to be one day older that that indicated in the query against the root .com server. The managing registration server should have same or earlier dates. The discrepancy is most likely due to the fact that melbourneit.com is located in Australian and thus operating in a different time-zone to the root .com server – [lines 2 & 4]
- The organisation name refers to "itsmysay.com" with an address in Tayside UK – [lines 5-10]
- Technical administration for the domain registration is directed to generic addresses related to "Easily Limited" (easily.co.uk), with UK address details pointing to London.
- There are two name servers listed (ns0.byworkwise.com and ns1.byworkwise.com), both independent of both itsmysay.com and easily.co.uk – [lines 30-31]

Analysis of this name registration information reveals that the initial registration of abcexample.com was carried out in Australia, but is now managed by the third-party company, Easily Limited – located in the UK. We

---

also note that the name servers for the domain are managed by another entity – byworkwise.com.

A quick examination of the byworkwise.com web site reveals that the name servers belong to "Blueyonder Workwise". Relating this information to the network WHOIS registration information, we see that the name servers are related to the personnel who last changed the RIPE WHOIS details observed in the first worked example (ripe-admin@blueyonder.co.uk).

It is interesting to note that the ABC Example web site can be accessed using the www.abcexample.co.uk domain name. A query of the WHOIS server responsible for managing .co.uk domains (whois.nic.uk) reveals the following:

```
1.      Domain Name: ngssoftware.co.uk
2.      Registrant: itsmysay.com
3.      Registrant's Agent: Easily Limited t/a easily.co.uk [Tag = WEBCONSULTANCY]
4.      URL: http://www.easily.co.uk
5.      Relevant Dates:
6.      Registered on:  26-Oct-2001
7.      Renewal Date:   26-Oct-2003
8.      Registration Status: Renewal required.
9.      Name servers listed in order:
10.     dns0.easily.co.uk               213.161.76.87
11.     dns1.easily.co.uk               217.206.221.213
12.     WHOIS database last updated at 20:30:02 08-Nov-2003
```
Example 6: WHOIS domain query response about ngssoftware.co.uk from whois.nic.uk

Observations:

- The initial .co.uk registration date (26th October 2001) is the same as for the .com version – [line 6]
- The renewal data has expired and ABD Example must re-register the domain name to keep it, or another organisation could purchase it – [lines 7-8]
- There are two domain names associated with the .co.uk domain (dns0.easily.co.uk and dns1.easily.co.uk), both of which are controlled by easily.co.uk – the same organisation identified in the registration of the .com domain registration details – [lines 3, 10 & 11]

Security Issues and Advice

When focusing upon the "Internet Service Registration" analysis phase of a passive information gathering exercise, organisations should carefully review the detailed information returned. The primary security issues and advice include:

ISSUE	Description
ISP Selection	For most organisations, almost all network service-based information will be controlled and managed by the local ISP or Internet connectivity provider. It therefore extremely important that a good relationship is established. Before selecting or migrating to a new ISP, carefully review some of the registrations they manage and ascertain the level of information they disclose. It should be fairly clear whether they follow security best practices, or could represent a future security risk to your organisation. Lax security could result in social engineering attacks or loss of all IP-based Internet traffic (effectively causing a denial of service).

ISSUE	Description
	In addition, many organisations will rely upon their selected ISP to provide DNS name services. It is vital that these hosted services be secure from an attack & penetration perspective as well as not disclosing volumes of information about their clients through unregulated zone-transfers. Loss or alteration of DNS services can result in denials of service and possible hijacking of client connectivity through man-in-the-middle based attacks.
Address Details	Where possible use generic postal addresses for all registration details – preferably not the organisations main office address. When possible, the use of PO Box numbers are recommended. However, some ISPs or registrars will not accept PO Box numbers for addresses. Alternatively, with permission the organisation's accountant's, their address may be used.
Real Names	Never use real names and email addresses in registration details. Create and use role-specific names and email addresses such as "RIPE Admin" and RIPE@myorganisation.com. Real names can be used in social engineering attacks, or automated brute-forcing attacks against other online services.
Phone Numbers	Where possible use a single, generic phone number for the entire organisation. The phone receptionist should have available a list of appropriate contact names and numbers for any registration queries. Ideally, the receptionist would take the contact name and number of the caller and pass it on to the appropriate internal technical contact instead of transferring the call to the internal technician. To reduce the risk of war-dialling and other exchange based attacks, a national number (such as 0870 in the UK or 1-800 in America) could be used.
MNTNER Auth	It is extremely important that the highest level of authorisation method is used to manage MNTNER records. Ideally, PGP should be used. Failure to use strong MNTNER authorisation levels can lead to unauthorised alteration of registration details. The effect could include loss of corporate Internet IP address ranges, and consequently all connectivity. In addition, a careful attacker could cause subversion of all organisational clients, and present them with false services (such as a fake website to steal login information and credit card details). When using non-PGP level MNTNER authorisation methods, it is important that the registration maintainers DO NOT select or use guessable passwords (nor passwords they use for other administrative tasks).
Expiry Dates	It is important to keep a careful track of the expiry dates of domain names the organisation has registered. Failure to do so can easily result in someone else (an attacker or competitor) taking ownership of the domain name and using it themselves for whatever purpose. Ideally, most organisations should review this registration information as a monthly or bi-monthly process

Table 2: Internet Service Registration security advice

Domain Name System

Since host names are easier to remember than IP addresses, they are the preferred method of addressing hosts. The Domain Name System (DNS) is a service designed to provide a link between an IP address and a unique host name.

Several implementations of DNS are used on the Internet and for internal corporate name resolution. The most common DNS service is BIND (named after "The Berkeley Internet Name Daemon"), and most other DNS services provide BIND level functionality/compatibility by default.

Although multiple security vulnerabilities have been identified with the protocol and coded implementation of the service, exploitation of these vulnerabilities is not in any way "passive". Instead, in this section we will focus upon the methods available to query these services, and evaluate the significance of the information returned.

Given the nature of the service, querying DNS records can provide a wealth of information to an attacker in a few short moments. Most critical to the service is the disclosure of multiple IP and naming records for a single queried domain. Loosely implemented name services may also yield more information than expected.

Querying DNS

The most common (and popular) method of querying BIND services is through the use of the "dig" tool. This tool is freely distributed as part of BIND and is installed by default on most UNIX based operating systems.

Dig can be used to resolve the names of hosts into IP addresses, and reverse resolve IP addresses into names. In addition, dig can also be used to gather version information from name servers which may be used to aid exploitation of the host.

Utilities such as dig can perform other DNS services, such as a Zone Transfers. Authoritative name-servers for a domain retrieve zone files (complete records) from other name-servers using Zone Transfers. By manually conducting a zone transfer, an attacker can gain valuable information about all systems and addresses in the domain from the domain name server.

Another query tool in the arsenal of passive information gathering tools is nslookup (short for "Name Service Lookup"), and comes as standard on most operating systems. Nslookup is almost as flexible as dig, but provides a simpler default method of identifying primary hosts such as Mail and DNS servers.

Worked Example – Three

Let us continue the investigation of NGS by querying a DNS server for information about www.abcexample.com. The authoritative DNS server queried was identified from the previous WHOIS queries (ns0.byworkwise.com [217.28.130.50]).

The following response was received from the query using an online tool (http://network-tools.com):

```
1.      ns0.byworkwise.com [217.28.130.50] returned an authoritative response in 94 ms:
2.      Answer records
3.      name            class  type  data                                time to live
4.      www.abcexample.com IN   A    213.48.xx.45                        3600s (1h)
5.      abcexample.com  IN    NS    ns0.byworkwise.com                  3600s (1h)
6.      abcexample.com  IN    NS    ns1.byworkwise.com                  3600s (1h)
7.      abcexample.com  IN    SOA   server:    ns0.byworkwise.com       3600s (1h)
8.                                  email:     admin@byworkwise.com
9.                                  serial:    25
10.                                 refresh:   900
11.                                 retry:     600
12.                                 expire:    86400
13.                                 minimum ttl:3600
14.     abcexample.com  IN    MX    preference: 10                      3600s (1h)
15.                                 exchange:   mail.abcexample.com
16.     abcexample.com  IN    MX    preference: 20
17.                                 exchange:   thsmtpb1.byworkwise.com 3600s (1h)
18.     Authority records
19.     [none]
20.     Additional records
21.     name                   class  type  data                       time to live
22.     ns0.byworkwise.com      IN    A    217.28.130.50               3600s (1h)
23.     ns1.byworkwise.com      IN    A    217.28.130.51               3600s (1h)
24.     mail.abcexample.com     IN    A    213.48.xx.35                3600s (1h)
25.     thsmtpb1.byworkwise.com IN    A    213.166.14.22               3600s (1h)
26.     www.abcexample.com      IN    A    213.48.xx.45                3600s (1h)
27.     end --
```

Example 7: Nslookup DNS query on abcexample.com

Observations:

- The two name servers associated with abcexample.com (ns0.byworkwise.com and ns1.byworkwise.com) have been identified and are classified as TYPE=NS (Name Server) – [lines 5 & 6]
- The primary DNS server is ns0.byworkwise.com and the administrative email contact is admin@byworkwise.com. This server is classified as TYPE=SOA (Start of Authority) – [lines 7 & 8]
- Two email servers have been identified (mail.abcexample.com & thsmtpb1.byworkwise.com), of which mail.abcexample.com is the primary/preferred server (A lower "preference" value takes priority). Mail servers are normally indicated with the TYPE=MX identifier – [lines 14-17]
- It is clear that both the mail and www servers are within the same IP netblock (213.48.xx.45 & 213.48.xx.35) – [lines 24 & 26]
- The two name servers (ns0.byworkwise.com and ns1.byworkwise.com) are located on the same logical and physical network. Thus any internal networking problem (such as a flaw in routing tables) could make them both unavailable. – [lines 22-23]

Using this new information, we have been able to identify the name and address of the primary mail server, and have discovered that it is most likely connected to the Internet with the same connection as the "www" host. In addition, we see that a backup mail server (thsmtpb1) is hosted by an external provider – Blueyonder Workwise. We have also confirmed earlier findings related to the "live" status of the name servers.

A standard Dig query reveals less information than the nslookup query above. This is due to the fact that Dig defaults to an 'A' class lookup, while nslookup's default is 'Any'.

```
1.      ; <<>> DiG 2.2 <<>> @218.020.062.001 www.abcexample.com
2.      ; (1 server found)
3.      ;; res options: init recurs defnam dnsrch
4.      ;; got answer:
5.      ;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 10
6.      ;; flags: qr rd ra; Ques: 1, Ans: 1, Auth: 2, Addit: 1
```

```
7.      ;; QUESTIONS:
8.      ;;      www.abcexample.com, type = A, class = IN
9.      ;; ANSWERS:
10.     www.abcexample.com.    3600   A      213.48.14.45
11.     ;; AUTHORITY RECORDS:
12.     abcexample.com.        172800 NS     ns1.byworkwise.com.
13.     abcexample.com.        172800 NS     ns0.byworkwise.com.
14.     ;; ADDITIONAL RECORDS:
15.     ns0.byworkwise.com.    146290 A      217.28.130.50
16.     ;; Total query time: 47 msec
17.     ;; FROM: ADVENT to SERVER: 218.020.001.001
18.     ;; WHEN: Fri Nov 11 23:43:39 2004
19.     ;; MSG SIZE  sent: 37  rcvd: 116
```

Example 8: Dig DNS query on www.abcexample.com

Worked Example – Four

In a lot of instances, organisations may utilise multiple IP addresses and name aliases for a single service. This can make the process of querying DNS services a little more complex, such as the following example when querying about Microsoft's primary website www.microsoft.com.

Due to the size of Microsoft's client base, they have had to provide multiple hosts distributed globally to ensure a robust service. To manage this complex environment, a string of aliases are used. For instance, a Dig query on www.microsoft.com provides the following:

```
1.      ; <<>> DiG 2.2 <<>> @218.020.62.001 www.microsoft.com
2.      ; (1 server found)
3.      ;; res options: init recurs defnam dnsrch
4.      ;; got answer:
5.      ;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 10
6.      ;; flags: qr rd ra; Ques: 1, Ans: 12, Auth: 9, Addit: 3
7.      ;; QUESTIONS:
8.      ;;      www.microsoft.com, type = A, class = IN
9.      ;; ANSWERS:
10.     www.microsoft.com.               870  CNAME   www.microsoft.akadns.net.
11.     www.microsoft.akadns.net.        275  CNAME   www.microsoft.com.edgesuite.net.
12.     www.microsoft.com.edgesuite.net. 876  CNAME   a562.cd.akamai.net.
13.     a562.cd.akamai.net.    1     A       63.208.194.7
14.     a562.cd.akamai.net.    1     A       63.208.194.14
15.     a562.cd.akamai.net.    1     A       63.208.194.23
16.     a562.cd.akamai.net.    1     A       63.208.194.25
17.     a562.cd.akamai.net.    1     A       63.208.194.56
18.     a562.cd.akamai.net.    1     A       63.208.194.57
19.     a562.cd.akamai.net.    1     A       63.208.194.64
20.     a562.cd.akamai.net.    1     A       63.208.194.65
21.     a562.cd.akamai.net.    1     A       63.208.194.89
22.     ;; AUTHORITY RECORDS:
23.     cd.akamai.net.  881    NS     n1cd.akamai.net.
24.     cd.akamai.net.  881    NS     n2cd.akamai.net.
25.     cd.akamai.net.  881    NS     n3cd.akamai.net.
26.     cd.akamai.net.  881    NS     n4cd.akamai.net.
27.     cd.akamai.net.  881    NS     n5cd.akamai.net.
28.     cd.akamai.net.  881    NS     n6cd.akamai.net.
29.     cd.akamai.net.  881    NS     n7cd.akamai.net.
30.     cd.akamai.net.  881    NS     n8cd.akamai.net.
31.     cd.akamai.net.  881    NS     n0cd.akamai.net.
32.     ;; ADDITIONAL RECORDS:
33.     n0cd.akamai.net.       51    A       213.161.82.7
34.     n1cd.akamai.net.       2211  A       213.161.82.12
35.     n2cd.akamai.net.       1429  A       213.161.82.13
36.     ;; Total query time: 47 msec
37.     ;; FROM: ADVENT to SERVER: 218.020.001.001
38.
39.     ;; WHEN: Tue Dec 25 23:26:38 2003
40.     ;; MSG SIZE  sent: 35  rcvd: 507
```

Example 9: Dig DNS query on www.microsoft.com

Observations:

- A series of three linked aliases are used for www.microsoft.com. These include www.microsoft.akadns.net, www.microsoft.com.edgesuite.net and a562.cd.akamai.net – [lines 10-12]
- Nine authoritative (TYPE=A) records are returned with their IP addresses – all of which exist on the 62.208.194.xxx netblock – [lines 13-21]
- Multiple name servers are listed [lines 23-31], with the preferred name server being n0cd.akamai.net (having the lowest preference value of 51) – [line 33]

Using default nslookup settings on a Microsoft windows system reveals the following information for www.microsoft.com:

```
1.      nslookup www.microsoft.com
2.      Server:  inh2dns05.abcexample.com
3.      Address:  218.020.62.001
4.      DNS request timed out.
5.      timeout was 2 seconds.
6.      Non-authoritative answer:
7.      Name:    a562.cd.akamai.net
8.      Addresses:  63.208.194.65, 63.208.194.71, 63.208.194.73, 63.208.194.89,
        63.208.194.95, 63.208.194.9, 63.208.194.22, 63.208.194.46, 63.208.194.63
9.       Aliases:  www.microsoft.com, www.microsoft.akadns.net,
        www.microsoft.com.edgesuite.net
```

Example 10: Nslookup DNS query on www.microsoft.com

Observations:

- The nslookup tool has automatically identified the primary web server (a562.cd.akamai.net) for the alias www.microsoft.com – [line 7]
- The nine IP addresses corresponding to the web server have been identified and listed – [line 8]
- All known aliases for the primary web server address (a562.cd.akamai.net) have been identified and listed – [line 9]

It is important to note that, in this example, subsequent DNS queries about www.microsoft.com will list different IP addresses, or some of these IP addresses will be in a different order. This is because Microsoft has implemented a round-robin method of listing addresses as part of a load-balancing procedure.

Zone Transfers

A special method exists for a DNS server to exchange authoritative records for a domain between multiple servers. This method, referred to as a Zone Transfer, is the main method of transferring bulk lists of domain information between primary and secondary servers. However, any client system can query a DNS server and request a zone transfer.

If a DNS server has not been securely configured, it is likely to respond to the client query and provide a list of all the information about the queried domain. The net effect of a successful zone transfer is that an attacker can obtain a list of all named hosts, sub-zones and associated IP addresses.

A zone transfer is an effective method of obtaining a lot of information about an organisations network for very little effort. It is for this reason that security best practices recommend that zone transfers only be allowed between hosts that are recognised authoritative name-servers and have been specifically listed, and to not allow zone transfers to unknown or unauthorised hosts. This should be implemented in the name-server software and at perimeter security (DNS lookups are performed over UDP/53 and Zone-Transfers over TCP/53).

In the example below, a query such as "dig @ns1.example.com example.com. axfr" carries out a zone transfer and returns a listing of IP addresses and their corresponding host names. A typical listing may look something like this:

```
Domain: example.com.
Primary Nameserver: ns1.examplehosting.com E-mail Contact: admin@examplehosting.com

/www/cgi-bin/demon/external/bin/dig @ns1.example.com example.com. axfr

; <<>> DiG 2.1 <<>> @ns1.examples.com example.com. axfr ; (1 server found)
example.com.3600SOAns1.examplehosting.com. admin.example.com. (

        10; serial
        3600; refresh (1 hour)
        600; retry (10 mins)
        1209600; expire (14 days)
        3600 ); minimum (1 hour)

        example.com. 3600 A        10.2.3.4
        example.com. 3600 NS       ns1.examplehosting.com
        example.com. 3600 NS       ns2.examplehosting.com
        example.com. 3600 MX       10 smtp.example.com.

        webmail.example.com. 3600 CNAME  webmail.freemail.com.
        router.example.com.  3600 A      10.2.3.1
        fw1.example.com.     3600 A      10.2.3.2
        snort.example.com.   3600 A      10.2.3.3
        www.example.com.     3600 A      10.2.3.4
        ftp.example.com.     3600 A      10.2.3.5
        pdc.example.com.     3600 A      10.2.3.6
        mailsweeper          3600 A      10.2.3.10
        devserver            3600 A      10.2.3.10
        mimesweeper          3600 CNAME  mailsweeper.example.com.


        example.com.         3600 SOA    ns1.examplehosting.com
admin.examplehosting.com. (
                10; serial
                3600; refresh (1 hour)
                600; retry (10 mins)
                1209600; expire (14 days)
                3600 ); minimum (1 hour)

;; Received 10 records.
;; FROM: nu7www.noname.net to SERVER: 101.102.101.102 ;; WHEN: Mon Dec 11 23:21:49
2004
```

Example 11: A successful zone transfer for example.com

The example query above has identified 10 host names, corresponding to 8 unique hosts belonging to example.com. It is important to note that all the named hosts are too descriptive in their naming convention – thus an attacker can easily use this information to investigate potentially "soft" critical hosts such as pdc.example.com (most likely to be a primary domain controller for their internal and external networks).

It should be noted that requesting a Zone Transfer is not typically classed as a "passive" information technique. While domain lookup requests are usually conducted over UDP, Zone Transfers rely upon a TCP connection – consequently they are easier to detect and stop using perimeter defence systems. However, in the context of an organisation that does not host their own external DNS servers, attempting a Zone Transfer against their hosting providers DNS server is unlikely to alert the organisation and is "passive" in this context.

An extension of this "passive" information gathering is the use of third-party sites that will attempt to carry out a zone transfer for a domain against the DNS name server. Such sites allow attackers to anonymously retrieve zone transfer information. One such site is IP-Plus - part of the Swisscom Enterprise Solutions (http://www.ip-plus.net/tools/dns_check_set.en.html). The following screenshot displays a typical zone transfer analysis:
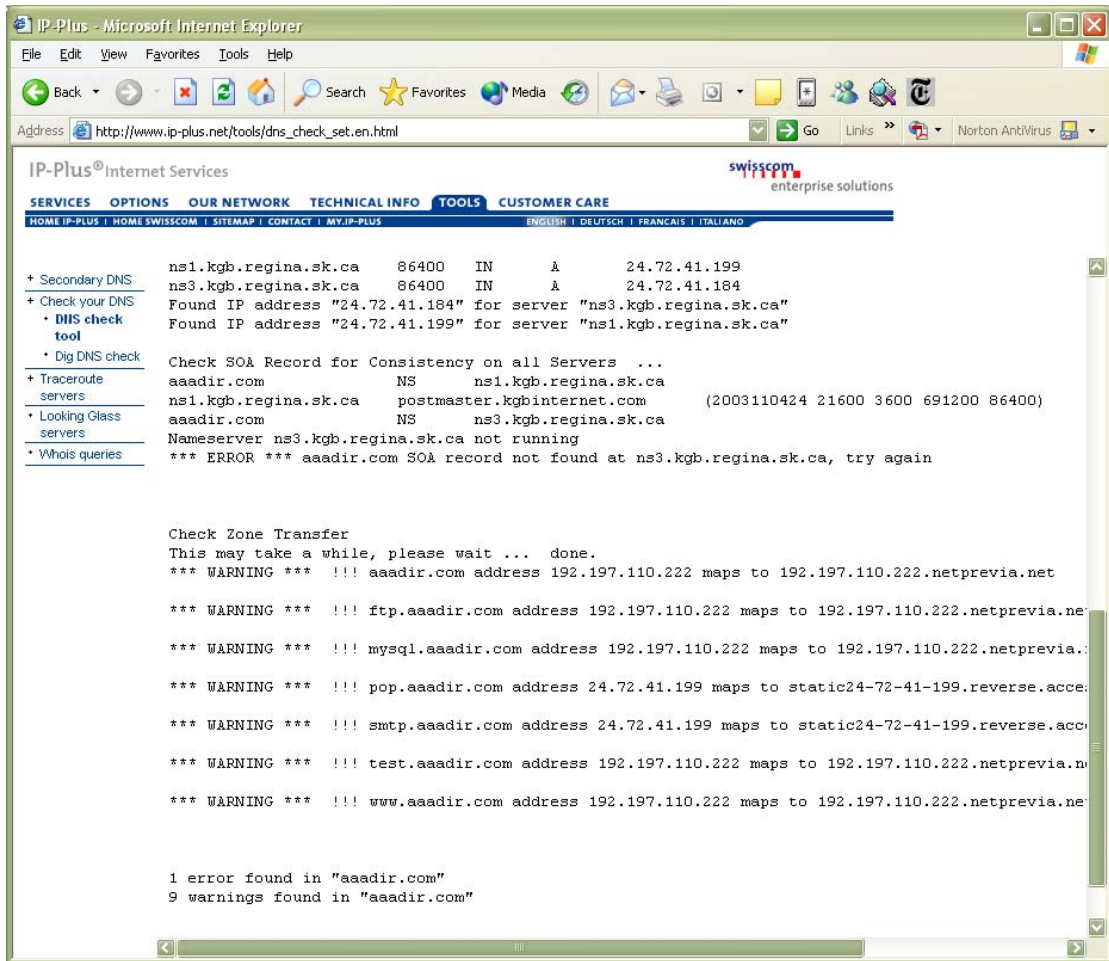
Reverse Resolution

Many of the queries conducted so far have been to take a named host or domain, and query for additional information including IP addresses. It is often possible to reverse this process and, by taking a known IP address, resolve it back to a host and domain name. Doing so can help identify other

hosts or network devices belonging to the organisation that did not show through previous query techniques.

A simple technique is to take a range of IP addresses above and below a known target. In the example below, a lookup on www.cisco.com revealed an IP address of 198.133.219.25. By creating a simple script to reverse lookup a range of IP addresses that included this IP, the following results were returned:

```
198.133.219.1     cco-net-hsrp.cisco.com
198.133.219.2     sjck-sdf-ciod-gw2.cisco.com
198.133.219.3     sjck-sdf-ciod-gw1.cisco.com
198.133.219.4     cco-wall-hsrp-dirty.cisco.com
198.133.219.5
198.133.219.6
198.133.219.7     sjck-sdf-ciod-sw1.cisco.com
198.133.219.8     sjck-sdf-ciod-sw2.cisco.com
198.133.219.9     test-garbage.cisco.com
198.133.219.10    ldir-ksdf-cionet.cisco.com
198.133.219.11    asp-web-sj-1.cisco.com
198.133.219.12    asp-web-sj-2.cisco.com
198.133.219.13    contact-sj-1-new.cisco.com
198.133.219.14    www.netimpactstudy.com
198.133.219.15    deployx-sj.cisco.com
198.133.219.16    contact-sj1.cisco.com
198.133.219.17    scc-sj-1.cisco.com
198.133.219.18    scc-sj-2.cisco.com
198.133.219.19    scc-sj-3.cisco.com
198.133.219.20    jmckerna-test.cisco.com
198.133.219.21    events.cisco.com
198.133.219.22    ldir-ksdf-cionet-failover.cisco.com
198.133.219.23    redirect.cisco.com
198.133.219.24    bill.cisco.com
198.133.219.25    www.cisco.com
198.133.219.26    partners.cisco.com
198.133.219.27    ftp-sj.cisco.com
198.133.219.28
198.133.219.29    deployx-sj-1.cisco.com
198.133.219.30    deployx-sj-2.cisco.com
198.133.219.31    ecommerce.cisco.com
198.133.219.32    www.stratumone.com
```

Example 12: A reverse lookup on the range 198.133.219.1 to 198.133.219.32 owned by Cisco

Observations:

- The IP addresses with blank rows are due to there not being a reverse lookup name.
- Out of the 32 IP addresses, 29 could be resolved to host names.
- The naming of many of the servers strongly suggests their organisational role.
- Two hosts not belonging to the cisco.net domain (www.netimpactstudy.com and www.stratumone.com) were identified, but exist within the same pool of IP addresses. This suggests that they are probably owned or managed by Cisco, and could be "softer" targets for gaining entry to the Cisco WAN.

It is important to note that such a simple type of query can yield a lot of information. Identification of non-domain related names within netblocks owned or managed by an organisation can prove useful to attackers. Investigation of these alternative domains may lead to other avenues of attack and potentially "softer" entry points into an organisation.

Reverse resolution can also reveal valuable information pertaining to the third-party hosting of web-based services. A popular method of providing cheap web hosting facilities has many ISP's hosting multiple organisations web-sites

on a single server (client browsers must support HTTP 1.1 to access these multiple sites). Thus, forward resolving a host name (e.g. www.example.com) would result in an IP address (e.g. 10.2.3.4), while reverse resolving this IP address would lead to the disclosure of the ISP's multiple-site host (e.g. webhost05.examplehosting.com).

DNS Brute Force

In cases where organisations have adequately controlled access to their DNS servers (e.g. Zone Transfers are refused) and reverse lookup is not available, it may still be possible to perform a dictionary-based attack against the DNS server to identify critical hosts and their primary function.

This type of investigation is typically automated and entails the use of a script or compiled application to forward resolve the IP address for a number of possible/probable named hosts. The script queries the DNS server for 'A' class records matching the guessed host name (e.g. firewall.domain.com, router.domain.com, dev.domain.com, ids.domain.com, owa.domain.com etc.), and reports the associated IP address.

```
firewall.domain.com    -    No Match
fire.domain.com        -    No Match
fw.domain.com          -    10.2.30.2
fw1.domain.com         -    No Match
fw2.domain.com         -    No Match
fw01.domain.com        -    No Match
fw02.domain.com        -    No Match
gateway.domain.com     -    No Match
gw.domain.com          -    10.2.30.1
gw1.domain.com         -    No Match
gw2.domain.com         -    No Match
gw01.domain.com        -    No Match
gw02.domain.com        -    No Match
webmail.domain.com     -    No Match
webm.domain.com        -    No Match
exchange.domain.com    -    No Match
outlook.domain.com     -    No Match
olk.domain.com         -    No Match
owa.domain.com         -    10.2.30.12
london.domain.com      -    No Match
lon.domain.com         -    No Match
ldn.domain.com         -    10.2.32.33
newyork.domain.com     -    No Match
ny.domain.com          -    No Match
nyk.domain.com         -    10.2.33.35
```

Example 13: An example DNS brute-force on the domain Domain.Com

Security Issues and Advice

When focusing upon the "Domain Name System" analysis phase of a passive information gathering exercise, organisations should carefully review the detailed information returned. The primary security issues and advice include:

ISSUE	Description
Zone Transfers	A Zone Transfer represents an easy method of extracting detailed information about an entire domain from a single DNS server. This information typically lists all named hosts belonging to the domain, and in many cases may also list internal systems – including their internal or

ISSUE	Description
	NAT IP addresses. Zone transfers should only be allowed between known and trusted systems, and not to unlisted client systems. It is a simple task to configure DNS servers to only allow Zone Transfers between listed hosts (IP addresses) and secure at perimeter by restricting access to tcp/53.
Reverse Lookup	Given a netblock of IP addresses associated to an organisation, it is often possible to enumerate host names with simple DNS reverse lookups. Reverse lookup functionality should only be granted to hosts or services that genuinely require this functionality. All other non-essential hosts should be scoured from reverse lookup DNS tables.

Table 3: Domain Name Service security advice

Search Engines

The use of search engines is vital for harvesting the often widely distributed cache of public material relating to the organisation under analysis. There is a popular saying in the computer underground when it comes to passive information gathering: "Google is your friend". It is surprising what can be unearthed using an advanced public search engine, particularly one as sophisticated as Google. Not only will Google allow you to search for specific text strings, it can also cache page content. Therefore, even after an offending or insecure page has been withdrawn from a web site, an attacker can still call up and analyze the cached page content.

Quite often other informational gems appear through conventional searching techniques. Past investigational queries have discovered client firewall configuration manuals, internal auditing manuals and confidential financial analysis documents when searching for different permutations of the organization's name, and restricting the search to .doc and .xls file extensions.

Searching newsgroups and other public posting areas often reveals infrastructure details as the organisation's administrators pose or answer questions relating to specific components of their network or software.

For example, one organisation had a public posting providing advice on getting a new security patch for AIX systems to work - telling the members of the newsgroup that the only way they managed to get a service functioning was by removing certain other "less likely to be exploited" security patches. Not only did this describe the type and patch level of their server, but also went on to explain what patches they had removed. In other cases, the details can be used for social engineering or extortion purposes.

Network Investigation Search Engines

There are a number of specialist search engines that focus upon information gathered about the health of the Internet and hosts that are frequently accessed. One popular site is Netcraft (http://www.netcraft.com), which specialises in the analysis of web hosts, the versions of the software they are running, and system uptime. One particular tool allows you to search for a particular string in the host's name (e.g. microsoft. or .cisco.com etc.) – and retrieve a list of all known web-enabled systems. An example screenshot is shown below:

Security Issues and Advice

When focusing upon the "Search Engines" analysis phase of a passive information gathering exercise, organisations should carefully review the detailed information returned. The primary security issues and advice include:

ISSUE	Description
Cached Content	Many popular search engines will allow users to search for, and request, cached page content in preference to current "live" content. Therefore, the possibility remains that, having removed specific content from an organisations web site, it may still be possible for others to retrieve a copy of the removed material.

ISSUE	Description
	Where ever possible, organisation must ensure that all pages to their websites contain appropriate information and meta-tags to limit third-party caching. However, organisations must also be aware that some search engines will ignore any caching limitations.
Error Messages	By restricting searches to a particular domain, it is sometimes possible to search for specific error messages. These error messages may have been generated by the website under investigation when the search engine requested content incorrectly. This error information can often be used to gain a better understanding of the type or supplier of the web-server technology (e.g. Broadvision and Microsoft SQL Server). If an organisations web services must provide error messages, they should be generic and not be indicative of the underlying application technologies. Preferably, any encountered error should result in a server-side redirect to a normal page (typically the home page).
Public Postings	It is important that searches of newsgroups and public messaging forums for content submitted using email addresses (i.e. domain names) belonging to the organisation are carried out on a regular basis, even if "acceptable use" policies are enforced internally. Message postings typically come from system administrators or internal development teams, and can thus hold a wealth of internal system information. Other postings to public message boards have, in the past, provided black-mail opportunities of internal staff after sexual or racial remarks have been made.
Public Documents	It is important that each document publicly released by an organisation be stripped of any internal editing references. For example, Microsoft Word documents (.doc) may contain internal information within the Document-Properties section, or may still contain undelete and tracked-changes content. With the Google search engine, the following search string would list all word documents on multiple web servers that are part of the example.com domain, and contain the words " top secret": `"top secret" site:example.com filetype:doc`
Robots.txt	Some web sites contain a file called "robots.txt" that is used by search engines to navigate parts of a website. The file typically lists the areas of a site that the search engine should or should not retrieve and catalogue. By manually reviewing this file, attackers may be able to discover sections of an organisations web site that they did not wish to be publicly disclosed.

Table 4: Search Engine security advice

Email Systems

Email hosts are probably the most important business critical systems organisations operate which are exposed to the Internet. While web sites present the public face of the organisation or services to their customers, their mail systems provide the essential business communications. In general, mail systems are often poorly secured and probably less understood by their administrators than web services.

A lot of information about an organisation can be gathered through passive analysis of the mail systems. In particular, enumeration of user accounts and mapping of the internal network.

SMTP Headers

During a Passive Information Gathering exercise, a lot of information may be obtained from the analysis of email headers. The SMTP protocol stipulates that email headers contain routing and address information for the safe delivery (and consequently reply) of the email message.

To manage email within a global organisation, multiple email servers are frequently utilised. As an email is routed internally, STMP headers are appended to the email message. Email headers are valuable for providing insight into internal server naming, IP numbering schemes, the type and version of content filter or anti-virus solution, service patch levels and even the version of the client's mail client.

Worked Example – Five

The following email header is taken from a real email that has passed through a large international organisation. Given the level of detail the original email headers contained, this example has been made anonymous.

```
1.      Return-path: <Elf@examplenetwork.net>
2.      Envelope-to: hosteddomain.net_bob@isphost.co.uk
3.      Delivery-date: Fri, 16 May 2004 15:57:03 +0100
4.      Received: from [127.0.0.1] (helo=localhost)
5.          by athena.isphost.co.uk with esmtp (Exim 4.14)
6.          id 19Ggdr-0004iN-Ln
7.          for hosteddomain.net_bob@isphost.co.uk; Fri, 16 May 2004 15:57:03 +0100
8.      Received: from athena.isphost.co.uk ([127.0.0.1])
9.          by localhost (athena.isphost.co.uk [127.0.0.1]) (amavisd-new, port 10024)
            with ESMTP id 18035-05
10.         for <hosteddomain.net_bob@isphost.co.uk>; Fri, 16 May 2004 15:57:03 +0100
            (BST)
11.     Received: from [106.253.xxx.132] (helo=mutex.netrex.com)
12.         by athena.isphost.co.uk with esmtp (Exim 4.14)
13.         id 19Ggdq-0004ht-BX
14.         for bob@hosteddomain.net; Fri, 16 May 2004 15:57:03 +0100
15.     Received: from USAMail.examplenetwork.local (plutonium [109.139.xxx.201])
16.         by mutex.netrex.com (8.12.9/8.12.9) with ESMTP id h4GEtRO9008774;
17.         Fri, 16 May 2004 10:56:34 -0400 (EDT)
18.     Received: from EuropeMail.examplenetwork.local ([10.2.1.20])
19.         by USAMail.examplenetwork.local with Microsoft SMTPSVC(5.0.2195.5329); Fri,
            16 May 2004 10:56:08 -0400
20.     Received: from EuropeMail.examplenetwork.local ([10.2.1.20])
            helo=mailhub.northpole.examplenetwork.net)
21.         by plutonium with esmtp (Exim 3.22 #23)
22.         id 19Mkov-0000j3-00
23.         for bob@hosteddomain.net; Fri, 16 May 2004 09:37:33 +0100
```

```
24.     Received: from mailhub.northpole.examplenetwork.net ([10.1.1.21)
25.        by mailhub.northpole.examplenetwork.net (Content Technologies SMTPRS
        4.2.10)
26.        with ESMTP id <T6294472d870a769a8c39c@ElfHost.northpole.examplenetwork.net>
        for <bob@hosteddomain.net>;  Fri, 16 May 2004 09:34:47 +0100
27.     content-class: urn:content-classes:message
28.     MIME-Version: 1.0
29.     Content-Type: text/plain;
30.     charset="us-ascii"
31.     Content-Transfer-Encoding: quoted-printable
32.     X-MimeOLE: Produced By Microsoft Exchange V6.0.6249.0
33.     Subject: FW: Xmas Time Deliveries
34.     Date: Fri, 16 May 2004 15:56:06 +0100
35.     Message-ID:
        <B9A2D6FDC131EA4C917246D65C69342B1A3BC3@ElfHost.northpole.examplenetwork.ne
        t>
36.     X-MS-Has-Attach:
37.     X-MS-TNEF-Correlator:
38.     Thread-Topic: Xmas Time Deliveries
39.     Thread-Index: AcMbuJ/6CTioicZzR+GzPie0ByrqiQAAmtYA
40.     From: "Helper Elf (North Pole)" <Elf@examplenetwork.net>
41.     To: <someone@somwhereelse.com>,
42.     "Annoying Elf" <Annoy@examplenetwork.net>
43.     Cc: <bob@hosteddomain.net>
44.     X-OriginalArrivalTime: 16 May 2004 14:56:08.0112 (UTC)
        FILETIME=[483A4B00:01C31BBB]
45.     X-Original-To: bob@hosteddomain.net
46.     X-Virus-Scanned: by isphost.co.uk
47.     X-UIDL: I+1!!U`e!!edg"!A]V!!
48.     Status: RO
49.     X-Status: U
50.     X-Keywords:
51.     X-UID: 267
52.     X-KMail-EncryptionState:
53.     X-KMail-SignatureState:
```

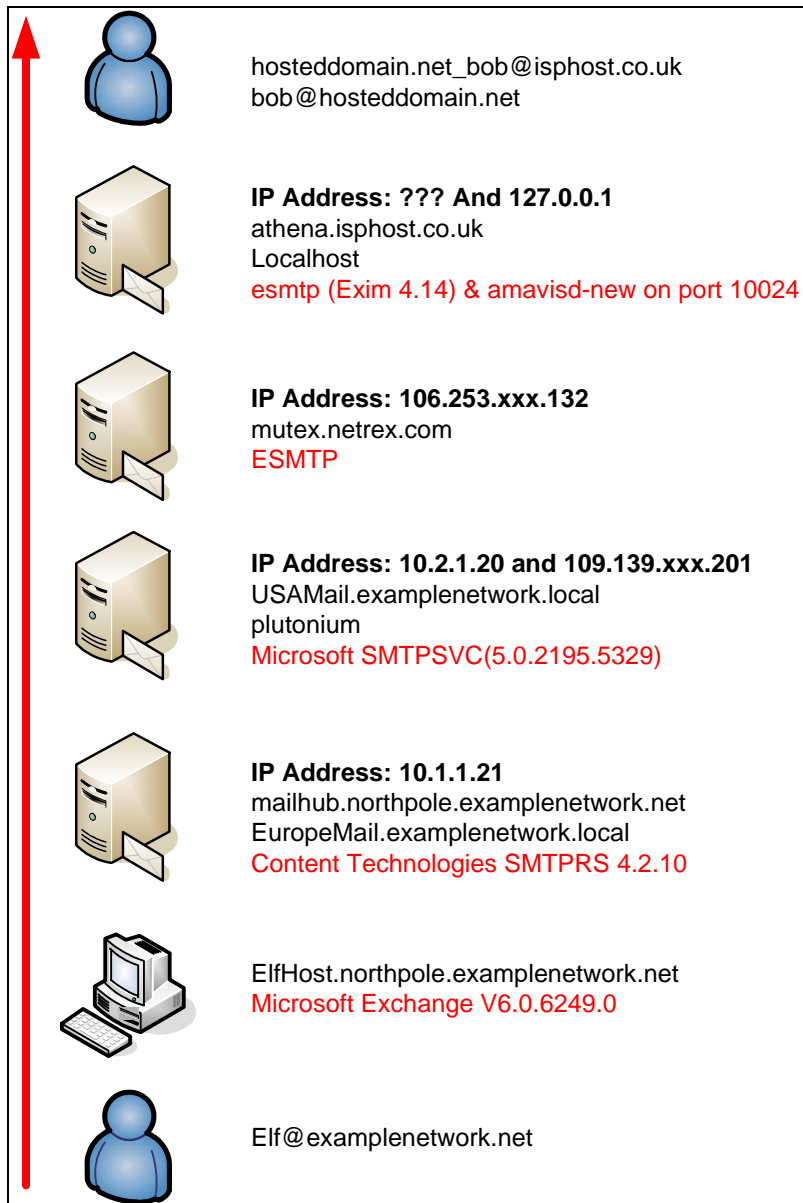Example 13: An example SMTP email header

Observations:

- The email was originally sent by Elf@examplenetwork.net – destined to be sent to two email addresses (someone@somwhereelse.com & Annoy@examplenetwork.net), and copied to one other (bob@hosteddomain.net). The email headers for this email are for the delivery to bob@hosteddomain.net. – [lines 40 to 43]

- The email was originally sent from a mail client connected with a Microsoft Exchange server (v6.0.6249.0) connector – [line 32]

- The email then passes to the next server. This server has two names (mailhub.northpole.examplenetwork.net and EuropeMail.examplenetwork.local) and, given the zone information within the domain name, exists within the main domain examplenetwork.net and the local domain "northpole". One internal IP address is revealed (10.1.1.21), and we know that the host is probably running some kind of mail filter/anti-virus system due to the reference to "Content Technologies SMTPRS 4.2.10".

- The next server in the chain goes by two names – "USAMail.examplenetwork.local" and "Plutonium". This servers internal IP address is 10.2.1.20 and transfers mail using Microsoft SMTPSVC(5.0.2195.5329) – typically associated with the Microsoft Exchange server application. Also referenced is the servers external IP address of 109.139.xxx.201.

- Given the internal mail server names of EuropeMail and USAMail, and the two different class B addresses (10.1.x.x and 10.2.x.x), a guess can be made at the internal IP numbering scheme of the organisation.

- The next server is the chain may be an email gateway, or anti-spam/anti-virus server (mutex.netrex.com) given that it does not appear to be affiliated with either organisation.

- The final server in the chain is athena.isphost.co.uk. This server probably runs additional services beyond SMTP due to the cyclic reference to itself (localhost – IP address 127.0.0.1). This additional service is probably related to the referenced "amavisd-new".

- A quick web search for "amavisd-new" reveals – "… is a high-performance interface between mailer (MTA) and content checkers: virus scanners, and/or SpamAssassin. It is written in Perl for maintainability, without paying a significant price for speed. It talks to

MTA via (E)SMTP or LMTP, or by using helper programs. Best with Postfix, fine with dual-sendmail setup and Exim v4, works with sendmail/milter, or with any MTA as a SMTP relay."

- Finally, it appears that the domain "hosteddomain.net" is managed (and probably hosted) by "isphost.co.uk" given the two linked email addresses (hosteddomain.net_bob@isphost.co.uk and bob@hosteddomain.net)

To help explain the linkages between the four different servers, the following illustration is provided:

hosteddomain.net_bob@isphost.co.uk
bob@hosteddomain.net

IP Address: ??? And 127.0.0.1
athena.isphost.co.uk
Localhost
esmtp (Exim 4.14) & amavisd-new on port 10024

IP Address: 106.253.xxx.132
mutex.netrex.com
ESMTP

IP Address: 10.2.1.20 and 109.139.xxx.201
USAMail.examplenetwork.local
plutonium
Microsoft SMTPSVC(5.0.2195.5329)

IP Address: 10.1.1.21
mailhub.northpole.examplenetwork.net
EuropeMail.examplenetwork.local
Content Technologies SMTPRS 4.2.10

ElfHost.northpole.examplenetwork.net
Microsoft Exchange V6.0.6249.0

Elf@examplenetwork.net

Example 14: The email passed through multiple servers on its way to bob@hosteddomain.net

These SMTP email headers can be extracted from a variety of sources including received emails and emails posted to public forums.

For many organisations, there are a number of opportunities to receive email and review headers. Common methods can include:

- Utilisation of online web-based customer service portals operated by the organisation under investigation,
- Sending emails to addresses within the organisation that could not possibly exist (e.g. randomnonsense@example.com and test@test.test.test.example.com), and reviewing the non delivery responses.

Email Address Information

Another important aspect of passive information gathering is the harvesting of email accounts. Most organisations follow one of two naming models for their users' email addresses: either the address contains the user's full name, or an abbreviated version that directly maps to their logon ID. Consequently, the full name is useful for social engineering attacks, and the abbreviated name forms half of the user-name/password pair needed to log into corporate resources. These addresses may be extracted from organisations' web sites or purchased from various spam mailing lists. Of most value are the names and email addresses of staff with technical administrative authority.

SMTP Server Banners

Most Internet mail servers depend upon SMTP to transport mail between hosts. Since a typical SMTP mail host is unlikely know in advance the mail servers that will connect to it and attempt to sent it email, the host will usually allow any remote host to connect to the service on TCP port 25. By default, most SMTP mail services provide an informative banner upon connection to the service. This banner may be used to positively identify the exact version or supplier of the SMTP mail gateway software. Using this information, an attacker may be better able to tune future attacks.

```
220 mail.example.com ESMTP Exim 4.21 Sat, 28 Dec 2003 18:14:37 +0000
```
Example 15: An example SMTP connection banner

In the example above, the mail server (mail.example.com) identifies itself as running the Exim SMTP service, version 4.21. A quick Internet search reveals details bout this service and security vulnerabilities associated with the version in use by mail.example.com

Security Issues and Advice

When focusing upon the "Email Systems" analysis phase of a passive information gathering exercise, organisations should carefully review the detailed information returned. The primary security issues and advice include:

ISSUE	Description
Internal Naming	It is important to be aware that host naming conventions and domain structures can be used to formulate a map of an organisations internal network structure. This map can be used by attackers to develop more sophisticated exploits and target key infrastructure components.
Internal IP Addresses	As with domain naming conventions (especially LDAP and Active Directory based services), IP addresses leaked through mail headers can help an attacker define the bounds and structure of an organisations internal routing structure.
Mail Services	The SMTP protocol is commonly configured to provide information about

ISSUE	Description
	the version being used. This information can help an attacker search for, or develop, exploit code specific to an identified version of the STMP service.
	Many common SMTP services allow administrators to alter or remove these banners. It is recommended that the necessary configuration changes be made to each mail service in order to remove any banner information.
SMTP Banners	As shown in the example, the SMTP banners can include an entire routing history of the email message. Organisations should implement processes that will strip all such routing information from outbound email messages at the last message gateway host. Many anti-virus mail gateways are capable of doing this.

Table 5: Email systems security advice

Naming Conventions

An important aspect of passive information gathering, and more subtle than many of the techniques described previously, is the observation and analysis of the actual names used to define each networked host or service. The naming convention used by an organisation can provide valuable insight to the use and position of hosts within an organisation. In extreme cases, poor naming conventions can even reveal the type of hardware used.

The most common mistakes include:

- The use of physical location information (e.g. London.example.com and HQOffice034.example.com) or common location shorthand (e.g. LONmail.example.com and ATLmail.example.com).
- The use of operations system information (e.g. the Microsoft Windows 2000 host w2k034.example.com and the HP Unix server HPUXserver.example.com).
- The use of functional information (e.g. FW.example.com for the firewall and OWA.example.com for the Outlook Web-mail Application server).
- The use of hardware manufacturer information (e.g. a Dell CPx laptop called DELLCPX002.example.com, or Cisco6100.example.com for a Cisco 6100 network manageable switch).
- The use of network location information (e.g. fwDMZ1.example.com and fwDMZ2.example.com)
- The use of common sequences – such as naming all servers after the planets (e.g. mars.example.com and venus.example.com) or ancient gods (e.g. titan.example.com and zeus.example.com)

For internal naming conventions, even smaller organisations should resist the temptation to use host names such as "johnscomputer" or "reception". Although such conventions make management easier in the short term, they can also raise the temptation for internal users to go snooping through Network Neighbourhood for the finance directors' computer or the personnel departments system holding salary details. Additionally, location based names are of limited value when office layouts change.

The naming conventions used and implemented by an organisation will always be a compromise between attacker obscurity and something internal users can remember. Most organisations will settle on a host naming convention that is memorable and easy to predict, thus information will always be leaked in some form. It is important however, to ensure that valuable information about the host's role or importance within the organisation is not readily discernable.

Trace route

The importance of choosing a good naming convention is most valuable when seen from an attacker's perspective, using tools that rely upon reverse lookups. One such tool is Traceroute.

Traceroute is supplied with almost all operating systems, although the actual name may differ depending on the particular flavour of operating system (e.g. most UNIX systems use the tool name "traceroute" while Microsoft operating systems use the shortened name of "tracert"). The tool is designed to show all the "hops" (intermediary network devices) that network traffic must go through to reach the final destination.

Various trace route implementations can be found online. Many of these tools include additional reporting functionality, as well as anonymous investigation. The following screenshot is from www.samspade.com, investigating the routing to www.cisco.com.
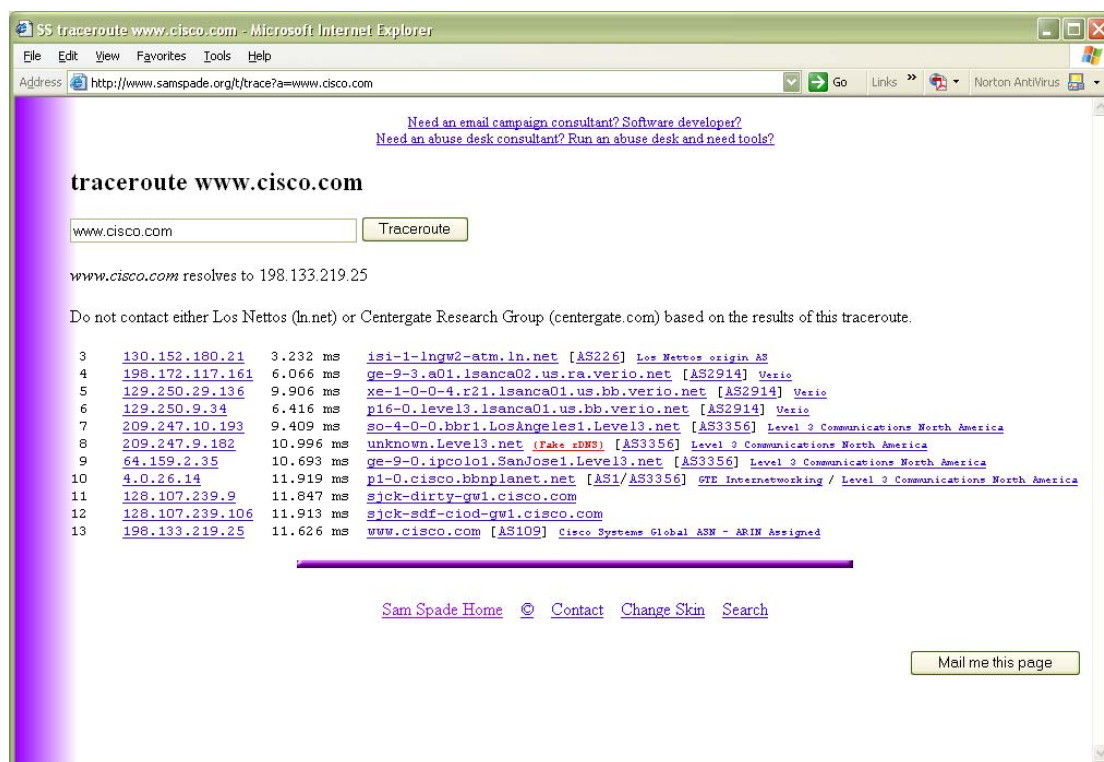
Worked Example – Six

The following example data was provided using the Microsoft Windows trace route tool Tracert.exe. The tool was launched from the local host (with an external IP address of 192.168.100.1), and the target was www.example.com (IP address 212.84.xx.6).

```
Tracing route to www.example.com
over a maximum of 30 hops:

  1  <10 ms  <10 ms  <10 ms   192.168.100.1
  2   40 ms   60 ms  160 ms   isi-1-lngw2-atm.ln.net [130.152.180.21]
  3   30ms    40ms    100ms   ge-9-3.a01.lsanca02.us.ra.verio.net [198.172.117.161]
  4  100 ms  120 ms  100 ms   xe-1-0-0-4.r21.lsanca01.us.bb.verio.net [129.250.29.136]
```

```
 5    70 ms   100 ms    70 ms    p16-0.level3.lsanca01.us.bb.verio.net [129.250.9.34]
 6    61 ms   140 ms    70 ms    so-4-3-0.bbr2.LosAngeles1.Level3.net [209.247.9.145]
 7    70 ms    71 ms   150 ms    so-1-2-0.bbr1.Washington1.Level3.net [64.159.0.138]
 8  1060 ms   960 ms  1091 ms    so-2-0-0.mp1.London2.Level3.net [212.187.128.137]
 9  1070 ms  1140 ms  1100 ms    so-2-0-0.mp1.London1.Level3.net [212.187.128.50]
10  1101 ms  1130 ms   900 ms    so-7-0-0.gar1.London1.Level3.net [212.113.3.2]
11  1180 ms  1190 ms   970 ms    so-2-0.metro1-londencyh00.London1.Level3.net [212.1.0.9]
12  1110 ms  1110 ms  1100 ms    80.253.125.11
13  1125 ms  1110 ms  1110 ms    h51-160.nolisp.net [212.84.160.51]
14  1145 ms  1125 ms  1130 ms    cisco-gw.example.com [212.84.xx.1]
15  1150 ms  1165 ms  1130 ms    cpfw1.example.com [212.84.xx.2]
16  1135 ms  1155 ms  1150 ms    www.example.com [212.84.xx.6]


Trace complete.
```

Example 16: An example Traceroute using Microsoft's tracert tool.

Observations:

- There exists a large time difference between hops 7 and 8. This is probably due to the network traffic being routed between Washington and London over a high latency device such as a satellite link.
- No host name is provided at hop number 12. This is probably due to the fact that no reverse lookup entry exists.
- The host name at hop number 14 (cisco-gw.example.com), combined with the fact that the IP address (212.84.xx.1) is probably the start of a netblock, suggests that this is the border router for example.com and that it is manufactured by Cisco.
- The host name at hop number 15 (cpfw1.example.com) is almost certainly a Checkpoint Firewall-1 firewall host.

Security Issues and Advice

When focusing upon the "Naming Conventions" analysis phase of a passive information gathering exercise, organisations should carefully review the detailed information returned. The primary security issues and advice include:

ISSUE	Description
Location Naming	The naming of key infrastructure devices to include shorthand notation of their physical location can be a boon to any network administrator when troubleshooting faults. However, the use of location information for all internal hosts should not be promoted as it is typically a redundant feature for internal users, while it adds detail useful to an attacker attempting to map an organisations internal network structure.
Service Naming	Unless the service is required to be obvious, or part of an accepted naming convention (e.g. "www" for web-servers), organisations should refrain from naming hosts after critical services they provide. While naming a key financial system "payroll.example.com" may make for an easy way of locating internal resources, this naming convention makes it easy for an attacker (or malicious internal user) to determine high priority targets.

Table 6: Naming Convention security advice

Web Site Analysis

As most organisations maintain large or complex Internet visible websites, the opportunity to inadvertently leak internal information is usually high. As such, detailed analysis of website content is valuable to an attacker.

The most effective way of analysing an organisations web site is to create a local mirror of the site content. This often requires the use of an automated tool to navigate the site and pull across a copy of every file referenced or linked to by the website. Obviously, such a task requires many consecutive connections and could be interpreted as intrusive. However, this type of activity is fairly common as both search engines and "offline web readers" will frequently perform this task. Therefore, from a passive information gathering perspective, such an activity is unlikely to be discovered or perceived as a threat or prelude to attack. In fact, for many organisations this type of activity may be perceived favourably as it increases the number of "hits" against the site and pleases internal PR staff.

The process of automatically retrieving web site content and analysing the content is commonly referred to as "web scraping". Web scraping, along with other manual investigation techniques can reveal a great deal of information about the organisation. Typical findings include:

- The harvesting of names and email addresses that can be later used for automated brute-force attacks or social engineering.
- The observation of internal developer's comments pertaining to the coding or operation of the sites content.
- Signatures of development tools contained within Meta-tags or other "hidden" fields.
- Commented or disabled code - linking to normally inaccessible site content or hosts.
- Links to data file URL's or other in-appropriately secured content (e.g. Form submission log files).
- References to internal development hosts or connectivity methods.
- The inclusion of badly coded site content that includes snippets of server interpreted code.
- The detail and nature of error pages in response to non-existent content requests and "dead" URL's.
- Links to external affiliated sites and hosts that may prove "softer" targets in later attacks.
- The existence of documents and other binary data that may contain internal information (e.g. auditing guides, network layout diagrams).

It is thus important that all content posted by an organisation be analysed for any unintentional disclosure. Any analysis is largely dependant upon the volume of information presented by the organisation.

Binary Download Data

Many organisations allow the download of binary data files. These files may range from trial software to whitepapers and press releases. In many cases, internal information located within the binary files is unintentionally leaked. Although there are many possible examples, some of the most common failures include the following:

- The inclusion of "created by", "last edited by", authors email address and other information contained within the document properties section. Documents commonly vulnerable to this include Microsoft Word files, Microsoft PowerPoint files, Microsoft Excel files and Adobe PDF files.
- The inclusion of "un-edit" information and "tracked changes" within Microsoft Word files that can be easily restored.
- The inclusion of internal host names within file-properties of compiled applications.
- The inclusion of third-party licensing information within most files (e.g. Microsoft Word files, compiled applications that rely upon third-party licensed components).
- The use of document passwords to protect against editing or copying (e.g. Adobe PDF and Microsoft Office suite documents). A poorly selected password may be indicative of internal password creation rules, or be recycled by the document creator for accessing other corporate resources.
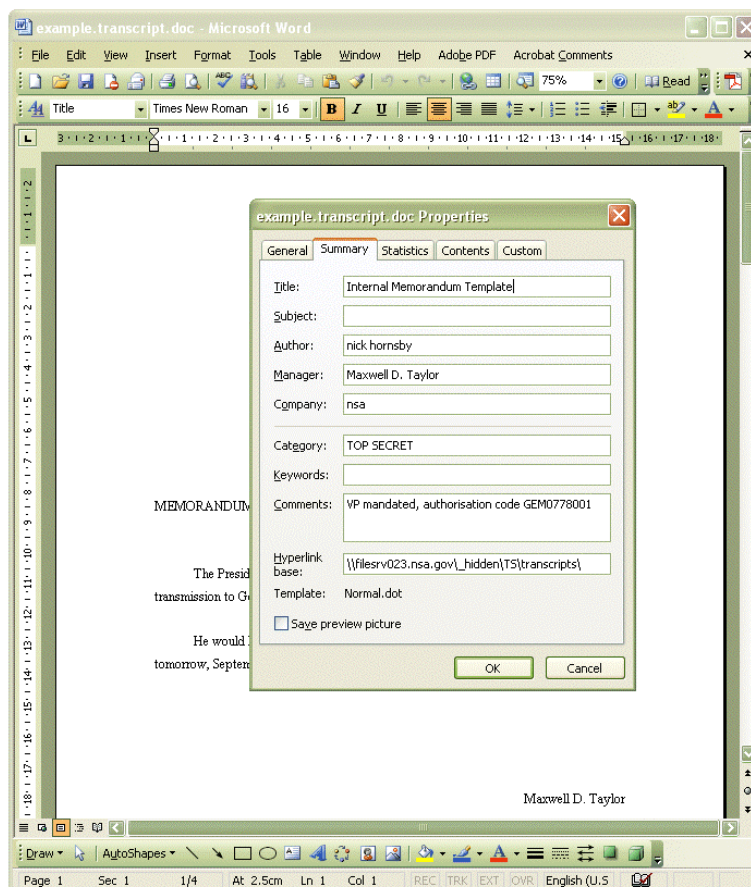
Web Server Banners

Similar to the issues encountered with SMTP mail headers, each request to a web server will result in a response containing information about the hosts web service. This information can be used to later target vulnerable web servers.

```
HTTP/1.1 404 Not found
Server: Zeus/4.2
Date: Sat, 28 Dec 2003 18:31:00 GMT
Connection: close
Content-Type: text/html
Set-Cookie: X-Zeus-Mapping-3ade68de=arion.isphost.com; path=/
```
Example 17: A sample banner from a Zeus web server.

```
HTTP/1.1 200 OK
Connection: close
Date: Sat, 28 Dec 2003 18:33:04 GMT
Server: Microsoft-IIS/6.0
P3P: CP='ALL IND DSP COR ADM CONo CUR CUSo IVAo IVDo PSA PSD TAI TELo OUR SAMo CNT COM
INT NAV ONL PHY PRE PUR UNI'
X-Powered-By: ASP.NET
Content-Length: 41032
Content-Type: text/html
Expires: Sat, 28 Dec 2004 18:33:04 GMT
Cache-control: private
```
Example 18: A sample banner from a Microsoft IIS web server.

```
HTTP/1.1 200 OK
Date: Sat, 28 Dec 2003 18:34:47 GMT
Server: Apache/2.0.48-dev (Unix)
Accept-Ranges: bytes
Cache-Control: max-age=86400
Expires: Sun, 29 Dec 2003 18:34:47 GMT
Content-Length: 8537
Connection: close
Content-Type: text/html; charset=ISO-8859-1
```
Example 19: A sample banner from an Apache web server.

Most web server software will enable system administrators to edit and change the banner of the web service. It recommended that web services are always renamed, and done in such a way to limit the amount of information disclosure.

However, it is interesting to note that although the web service banner may be changed, the request response (as indicated in the three examples above) have a layout that is common to the particular software vendor. Consequently, even after changing the service banner, it may still be able to identify the type or manufacturer of the web service software.

Security Issues and Advice

When focusing upon the "Web Site Analysis" phase of a passive information gathering exercise, organisations should carefully review the detailed information returned. The primary security issues and advice include:

ISSUE	Description
Addressing information	Care should be taken with organisational addressing information. Information such as email, telephone numbers and physical delivery locations are an important element in social engineering based attacks. Wherever possible, addressing details to should only refer to roles and/or

ISSUE	Description
	functions. For email, use marketing@example.com instead of the marketing staffs personal email account. In the case of telephone numbers, it is recommended that non-area code numbers such as 0800 be used to help prevent war-dialling attempts.
"Hidden" content	Organisations should ensure that the "hidden" content within their web pages does not hold any personal or revealing information. Typical sources of information leakage include meta-tags, broken links and commented code elements.
Error responses	It is important that non-standard or unexpected client requests to the website are dealt with a standard response that does not reveal data such as debug information, service-specific error messages or internal routing information (e.g. internal IP addresses or host names). Ideally, a sound session-management process should be used. Should a site visitor unintentionally or maliciously cause an internal error (or submit unexpected data), their session should be revoked and forcefully redirected to the first page of the web-site.
Binary Data	Before posting binary files to a corporate website or making it available through other electronic means, all binary data should be checked to ensure that it does not include any hidden information. Typical failures that have resulted in unintentional information leakage have been from Microsoft Word and Adobe PDF formatted documents. This leaked information is typically contained in hidden fields, editing, undelete and file-properties storage areas.
Service Banners	Wherever possible, the service banners of web servers should be changed so that they do not identify the vendor or version of the hosting software. Many automated tools exist for hacking web-sites and rely on the banner to tune their attacks. Thus, by removing the banner or replacing it with false information, many of these tools can be easily defeated. However, as discussed earlier, even with the removal of banner information, it may be possible to guess the vender/version of the software. This may be related to the layout of the HTTP HEAD request, or through typical coding structures and file extensions (e.g. PHP is associated with Apache while ASP is associated with IIS).

Table 7: Web Site information security advice

Conclusions

Passive information gathering is a vital stage in any black-box or zero-knowledge pentesting exercise, and consequently should form an important phase of any security assessment. The leaked information discovered during this analysis is often used by attackers to coordinate or plan more advanced attacks. Consequently, every effort should be taken by an organisation to ensure that the information leakage is limited as much as practically possible.

Whether an organisation conducts the passive information gathering exercise themselves or using a trusted third-party, they must ensure that the investigation is conducted thoroughly. Due to the diversity of the information, and the many opportunities for information to be leaked, organisations should ensure that a passive information gathering exercise be carried out multiple times a year.

Ideally a comprehensive passive information gathering exercise should be conducted twice per year in conjunction with a penetration test. Elements such as website and search-engine analysis should be conducted monthly due to the increased likelihood of content change.

The importance of passive information gathering techniques, both understanding the significance of the analysis techniques and the type of information available, is increasing yearly. With the substantial increases in the number of third-party hosted analysis tools capable of carrying out various degrees of analysis, organisations must be able to identify leaked information and take rapid steps in securing against future disclosure.

About Next Generation Security Software (NGS)
Established in 2001, NGS is the world's leading security vulnerability research company. In this capacity NGS act as adviser on vulnerability issues to the Communications-Electronics Security Group (CESG) the government department responsible for computer security in the UK. In addition to this NGS subscribe to the CESG IT Health CHECK Service. NGS Consulting provides a unique security auditing service based on the strengths of its Attack and Penetration team who lead the world in vulnerability research. The team members are all well known and widely respected professionals who have continually demonstrated to the security industry their technical competence and ability to innovate.

About NGS Insight Security Research (NISR)
The NGS Insight Security Research team are actively researching and helping to fix security flaws in popular off-the-shelf products. As the world leaders in vulnerability discovery, NISR release more security advisories than any other commercial security research group in the world.