

Domain Name Forensics: A Systematic Approach to Investigating an Internet Presence

by Bruce J. Nikkel
nikkel@digitalforensics.ch

Originally published by Elsevier in Digital Investigation
The International Journal of Digital Forensics and Incident Response
Vol. 1, No. 4 (doi:10.1016/j.diin.2004.10.001)

August 1, 2005

Abstract

Over the last few years the typical Internet presence has become a crowded outsourcing arrangement of multiple organizations dividing up the complexity of maintaining various parts of an infrastructure. Finding the parties responsible for the different infrastructure areas has become time consuming and error prone. This paper presents a systematic approach to investigating a complex Internet presence, including collecting, time-stamping, packaging, preserving, and presenting evidence. It is geared towards the network forensics practitioner.

Keywords: Digital Forensics, Network Forensics, Domain Name Investigation, Domain Name Forensics, DNS Investigation, Website Investigation

Contents

1	Introduction	3
2	Advantages of Complexity	3
3	Identifying Points of Responsibility	4
3.1	Domain Name Registrars	4
3.2	Domain Name Registrants	4
3.3	DNS Server Owners	4
3.4	Regional Internet Registries	5
3.5	Network Owners	5
3.6	Web Server Owners	5
3.7	Email Server Owners	6
3.8	Upstream ISP	6
3.9	Telecommunications Carriers	6
3.10	Routes and AS owners	6
3.11	Other Responsible Parties	7
3.12	The next generation, IPv6	7
4	Collecting and Preserving the Evidence	7
4.1	Preparing for the Investigation	8
4.2	Investigating the Domain Registry and Registrant	9
4.3	Investigating the DNS Owners	10
4.4	Investigating the IP Network Owners	10
4.5	Investigating the Reverse DNS	11
4.6	Investigating the Webserver Owner	12
4.7	Investigating the Upstream ISPs	13
4.8	Investigating the Routing Information	13
4.9	Investigating the Physical Location	14
4.10	Investigating the Email Owners	15
4.11	Finding Additional Information	15
5	Packaging and Preserving the Evidence	16
6	Presenting the Evidence	17
7	Conclusion and Future Work	17

1 Introduction

Tools such as whois or nslookup have traditionally provided a quick and simple method of investigating who is behind a particular Internet site. Unfortunately, the number of involved parties playing a role in maintaining an Internet presence has increased dramatically over the past few years making it more difficult for an investigator to identify those responsible for a site or its infrastructure.

A typical modern Internet presence has become a crowded outsourcing arrangement of multiple organizations dividing up the complexity of maintaining various parts of an infrastructure. Today there are often separate organizations managing the DNS, the IP network, and the various other server system platforms (email, web, application, and database servers, etc.). Servers are physically co-located, websites are virtually hosted, and other critical infrastructure components are sub-contracted out. Even the previously centralized role of Network Solutions Inc.¹ has turned into a large competitive market of registrars.

What used to be a trivial investigative task has now become time consuming and error prone. A more systematic approach is needed which identifies the various responsible parties in an orderly manner and treats the information gathered as evidence. A method for collecting, time-stamping, packaging, preserving, and presenting this evidence is needed. This paper outlines some simple procedures for achieving this goal.

2 Advantages of Complexity

In spite of the issues just mentioned, the additional complexity brings some interesting benefits for investigators. Having critical infrastructure spread across multiple parties can, in some cases, help investigators overcome legal jurisdiction hurdles, as well as solve issues regarding anonymity.

Illegal activity done using Internet infrastructure residing outside a local jurisdiction has always been difficult to bring under control. But with more parties involved, the chances of having a piece of critical infrastructure residing within a region's legal jurisdiction are increased. This could provide investigators with additional sources of evidence, or even opportunities to disable a site.

¹Once the sole registrar for .com, .net, and .edu

The privacy and protection offered by anonymity is continuously misused by criminals to hide their identities from law enforcement and other investigative bodies. The more parties involved in the existence of an Internet presence, the more difficult it becomes for an entity to remain completely anonymous. Each outsourced party may have a certain amount of information about the anonymous entity (billing, registration, physical location, etc.). Piecing this information together could assist in the identification of the anonymous entity.

3 Identifying Points of Responsibility

In order to systematically proceed with an investigation, we need to identify the major parties responsible for maintaining an Internet presence. This section outlines the major responsible parties involved, identifies information each party may be able to provide, and identifies the capability they have to disable an Internet presence (if legally or ethically compelled to do so).

3.1 Domain Name Registrars

Before a domain name is recognized, it must be registered with the registrar of a Top Level Domain (TLD). Most countries centrally manage registrations to their own Country Code TLDs (for example .uk or .ch). Generic TLDs (for example .com or .org) are managed by independent registrars (for example Network Solutions Inc, or Register.com). Among other things, the registrar is responsible for maintaining contact information and name server information for registered domains under its control. A registrar also has the ability to deactivate or delete domain names under its responsibility.

3.2 Domain Name Registrants

The domain name registrants are those parties responsible for registering and maintaining a domain name. This typically includes the registrant, an administrative contact, a technical contact, and possibly a billing contact. These domain owners have the ability to deactivate their domain or to modify the information specifying the DNS servers and contacts.

3.3 DNS Server Owners

Name server owners control the DNS zones which resolve IP addresses and domain names. Maintaining DNS servers for a domain can be done by anyone, anywhere on the Internet (for example, a remote ISP, the registrar, dynamic DNS hosters, etc.). The DNS server owners can provide information about other hosts within the same domain or IP range. The DNS server

owners also have the ability to modify or disable DNS resolution for domains under their control.

3.4 Regional Internet Registries

Control of the Internet Protocol (IP) address space has been delegated to four Regional Internet Registries (RIRs). In general, regional delegation is as follows[4]:

- ARIN for North America and sub-Saharan Africa
- LACNIC for South America and the Caribbean Islands
- APNIC for the Asia/Pacific region
- RIPE for Europe, the middle east, northern Africa and central Asia

A newcomer, AfriNIC, may manage address space for the African continent in the future. These RIRs further delegate IP ranges to Local Internet Registries (LIRs) and National Internet Registries (NIRs) within a region. Internet Registries provide information about who is responsible for a particular IP range.

3.5 Network Owners

Network owners are those parties responsible for managing a particular IP network. They typically handle things like routing, link-layer connectivity and possibly reverse DNS lookups for a network. They can provide information about network topologies, physical links (including geographic location), and contact information for further delegated subnets. They have the ability to monitor Internet traffic and the possibility to provide access logs for dial-up connections. The owner of a network has the ability to disconnect subnets from the Internet either physically or by removing the relevant routing information.

3.6 Web Server Owners

Web server owners are responsible for providing a platform for storing and serving web content. They are responsible for the system administration of the server operating system and the web serving software. They may also be responsible for the server hardware (but not necessarily). These machines may be located at a remote site, or co-located at a central location together with other servers. The web server owner may be able to provide information about the people responsible for the web content, including file transfer logs or content management access. They may be able to provide HTTP access logs of the site including downloaded files. They even have the ability to provide investigative access to restricted areas of a site. The

web server owner is able to shutdown a website, redirect it elsewhere, or modify/remove content contained on the site.

3.7 Email Server Owners

Email service can be broken down into two basic parts, which may or may not be owned by the same party. Sending email servers² forward or deliver email to other servers or mailboxes. Servers which store delivered email provide access to mailboxes via the POP or IMAP protocols, and may provide a web front-end as well. Email server owners can provide email logs of both incoming and outgoing email traffic, and may provide investigative access to mailboxes. Email server owners are also in a position to monitor message activity for investigative purposes. The email server owners also have the ability to reject email to/from a particular domain.

3.8 Upstream ISP

Large ISPs and commercial networks are connected to each other through peering agreements at Internet Exchanges (IXs), Network Access Points (NAPs), or Metropolitan Area Exchanges (MAEs). These larger entities often provide bandwidth to smaller ISPs. Sometimes information about a site can be found by examining the surrounding ISPs. These upstream ISPs may also have the ability to monitor or block traffic to or from an attached site.

3.9 Telecommunications Carriers

The local telecommunications carriers are responsible for the “last mile” of physical connectivity. They maintain the physical link between a site and the ISP’s nearest point of presence. This could be a leased line, frame relay connection, or some other link layer technology. Home and small business connections may also have ADSL or cable connections. The carrier will be able to provide information about the physical location of a site. It may also have information about other network links at that physical location. The telecommunications carrier is able to disconnect a site from an ISP.

3.10 Routes and AS owners

In the distant past a single routing table was maintained for every network on the entire Internet. Internet routing has evolved considerably since then and today[2] uses Autonomous Systems (AS) which exchange routing information with peers using the Border Gateway Protocol (BGP³). Each AS

²properly referred to as Message Transfer Agents or MTAs

³Currently BGP version 4

has an AS Number (ASN) with the associated contact information available by querying whois databases. An AS has the ability to filter routing information entering or leaving the AS.

3.11 Other Responsible Parties

There are many other possible places to investigate an Internet presence. So far we have identified the most common parties involved at the infrastructure level. Other potential investigative points include such areas as:

- content management or publishing services
- web design companies
- distributed web caching services
- SSL Certificate Authorities
- application and database back-end providers
- other hosts on the same network segment or same domain name
- other parties mentioned or linked in the website's content

All of these areas could provide an investigator with useful information or assistance in disabling a site or finding the identities of people behind it.

3.12 The next generation, IPv6

The original IP protocol (version 4) has worked well for a number of years, but it is not without its problems. The next generation, IP version 6 was designed to fix a number of shortcomings in scalability and security which haunt the current Internet. The IPv6 Internet already exists and is growing (it is already common to see IPv6 name servers appearing in DNS lookups). In the future this will also be territory which will need to be part of every investigation. Fortunately, the methods outlined here will transfer easily to this new network layer, since it is conceptually similar to IPv4.

4 Collecting and Preserving the Evidence

This next section shows a practical investigation of an Internet presence. We start with only one piece of information: the website name. From this, we systematically do whois and DNS lookups to uncover the responsible parties. Investigating a modern Internet presence can generate a lot of query data, so it is imperative to save everything in an organized manner and to keep a record of the work that was done.

4.1 Preparing for the Investigation

The data we collect is evidence. The record of all our activity shows the procedures used to acquire the evidence. From this evidence, we will generate a report outlining the various points of responsibility together with contact information. It should be possible for an independent investigator to verify the contents of the report from the evidence collected.

For our investigation we will use a command line shell to systematically pipe collected evidence into an organized set of files. The example procedure shown is from a generic Unix shell, but it will also work on Linux, Mac OSX, and Windows shells (for Windows, additional tools may need to be installed separately). There are certain forensic advantages to collecting evidence with a command line instead of GUI or online Web interfaces:

- Each file containing evidence has a system generated time-stamp showing the exact time of evidence collection.
- Collected evidence is transferred from the collection tools directly to the files without human intervention⁴
- The Whois and DNS server names are explicitly defined and logged, showing that the evidence was collected from authoritative sources.
- A complete transcript log of the evidence collection procedure is available for scrutiny.

The example site we will be investigating is “www.example.com” [3] We begin by creating and entering our evidence directory and recording our session:

```
$ mkdir evidence
$ cd evidence
$ script record.txt
```

The Unix script command will keep a record of everything we see or type (including control characters). The file record.txt is the raw transcript of our investigation and will be included together with the rest of the evidence we will collect.

Throughout this document we will be using Unix redirection (“>”) to pipe evidential data into files. The first character of the file name denotes the content (W=Whois, N=Nlookup, T=Traceroute). These files may be read with standard Unix tools such as ‘cat’ or ‘more’, etc. Do not edit or modify these files. Doing so will update the modification time-stamp, causing the time of evidence collection to be lost.

⁴Human errors from graphical interactions such as copying and pasting are eliminated.

The Network Time Protocol[8] (NTP) can keep the system time accurate to a few tens of milliseconds. To show that our investigative machine is properly time synchronized[7], we use the `ntpq` command to save the time synchronization status⁵:

```
$ ntpq -p > timesync.txt
```

Every file we create will have system time-stamps which we will later preserve when we package the evidence. Since we are collecting evidence “snapshots” of things which will change over time, accurately time-stamping the moment of evidence collection is important.

At this point we can start writing our report which will be updated as we proceed with the investigation. To begin the report, state the time and date as seen with the `date` command:

```
$ date
```

Also state in the report that the investigative machine had the correct time and date when the investigation took place.

If website content is of evidential value, now would be a good time to acquire it. Proper acquisition of remote websites is outside the scope of this paper, but may be addressed in future work. We now move on to getting the information about the various parties involved.

4.2 Investigating the Domain Registry and Registrant

We will need to do a number of whois lookups, but first, we need to know which whois servers to use. The initial whois information can be found at the Internet Assigned Numbers Authority (IANA). A summary is provided here:

- For `.com`, `.net`, and `.edu` domains, queries to `whois.internic.net` will refer us to the appropriate independent registrar’s whois server.
- For `.org` domains, we use `whois.pir.org`
- Databases for other generic TLDs (gTLDs) such as `.biz` or `.info`, can be found at <http://www.iana.org/gtld/gtld.htm>
- Databases for country code TLDs (ccTLDs) such as `.ch` or `.uk` can be found at <http://www.iana.org/cctld/cctld-whois.htm>

⁵If NTP synchronization is not an option, the `'rdate'` command could be used

Using whois⁶, we look up example.com using whois.internic.net. Since this is a .com domain, it is not managed by a central registrar and the query will point us to the managing registrar’s whois server. We save the registrar information as follows:

```
$ whois -h whois.internic.net example.com > W_DomainRegistry.txt
```

From this, we will get the name of the registrar, a link to their website, and their whois server. The registrar is one of our points of responsibility. We note the registrar’s name and website in our report under the heading “Domain Name Registry”.

For example.com, the registrar’s whois server is whois.iana.org. Using this server, we do another whois lookup of the domain name:

```
$ whois -h whois.iana.org example.com > W_DomainOwner.txt
```

Here we find the traditional whois database info about the registrant, technical, admin, and billing contacts. If they are shown, these contacts are all potential points of responsibility. We note these contacts in our report under the heading “Domain Owners”.

A word of caution: The registrars are under no obligation to verify the quality of this information. It could be falsified, anonymized, or registered by a proxy service.

4.3 Investigating the DNS Owners

The last whois query also gave us a list of the name servers. Anyone controlling the DNS servers is a potential point of responsibility. We therefore do whois lookups on the name server’s domain name(s) and save them as evidence. For example.com the name server’s domain name was iana-servers.net:

```
$ whois -h whois.internic.net iana-servers.net # find registrar
$ whois -h whois.register.com iana-servers.net > W_DNSserver.txt
```

We note the names of the name server contacts in our report under the heading “DNS Server Owners”

4.4 Investigating the IP Network Owners

Now that we know who controls the domain’s whois data and name servers, we move on to finding out more about the owners of the IP address. For this we use the traditional nslookup tool⁷.

⁶The whois client used in this paper is a simple traditional whois client without any special query features

⁷‘dig’ is a more modern replacement, but not all systems include it by default

```
$ nslookup www.example.com a.iana-servers.net > N_Website.txt
```

Note how we used the domain’s name server (a.iana-servers.net) which was found in previous whois lookups. This ensures the data we receive is authoritative[1]. A thorough investigation would include additional nslookups against each name server to reveal any inconsistencies.

We now have the IP address which will help us find information about the network owners. The whois server used depends on the geographic region of the IP range, and will be one of the four Regional Internet Registries:

- whois.arin.net (North America/Sub-Sahara Africa)
- whois.ripe.net (Europe/Mid-East/Central Asia/North Africa)
- whois.apnic.net (Asia/Pacific)
- whois.lacnic.net (South America/Caribbean)

If you don’t know the region, then try all four. Here we guess that it is in North America:

```
$ whois -h whois.arin.net 192.0.34.166 > W_IPaddress.txt
```

Local Internet Registries (LIRs) and National Internet Registries (NIRs) may also have additional whois servers with more information specific to smaller IP networks[4].

Any contact information found from the Internet Registry whois lookups, we record in our report under the heading “Network Owners”. We also include the names of the Internet Registries (RIR, LIR or NIR) where the IP range exists.

4.5 Investigating the Reverse DNS

There is still more info we can find about potential network owners. This is done by querying for the Start Of Authority (SOA) resource records of the in-addr.arpa address.

The SOA Resource Record (RR) provides us with the authoritative name server for a zone, that is, the name server with the “best” data out of all the name servers listed. The SOA RR also gives us the email address of the person responsible for that data (this email address is somewhat obscured, the first dot in the “mail addr” field is the “@”).

The .arpa (Address and Routing Parameter Area) domain is a special domain used for Internet infrastructure[10]. The few sub-domains are under the responsibility of the Internet Architecture Board (IAB). The in-addr.arpa domain is of interest to us, because it provides us with information about Internet addresses (ip6.arpa is the IPv6 equivalent). Specifically, the in-addr.arpa domain tells us which name servers will provide us with reverse lookups of an IP network.

Performing an in-addr.arpa query can be a bit tricky. It involves reversing the order of the IP network address and adding it to the in-addr.arpa domain. A thorough explanation of this can be found in “DNS and BIND”[1]. To illustrate, we perform this query on the network containing www.example.com[6]:

```
$ nslookup -type=SOA 34.0.192.in-addr.arpa > N_inaddrarpa.txt
```

Note how 34.0.192 is the first three octets of the webserver’s IP address (192.0.34), but reversed. This is intentional and part of how DNS works[1].

A thorough investigation would include additional nslookups against each name server to reveal inconsistencies. Record the email address in the report and do a whois lookup of the server domain in the “origin” field, as well as any other name server domains. These are all potential points of responsibility and should be noted under the heading “Network Owners”. Since example.com’s name server domain (for the in-addr.arpa zone) is icann.org, we use whois.pir.org to perform the query:

```
$ whois -h whois.pir.org icann.org > W_IPDNSserver.txt
```

4.6 Investigating the Webserver Owner

If the website is being virtually hosted[5], the IP address may resolve to something other than itself, possibly revealing the owner of the machine. We can find the Fully Qualified Domain Name (FQDN) of the machine by doing a reverse lookup on the IP address.

A word of caution: be careful when trusting DNS lookups of IP addresses. Anyone controlling address resolution for a particular IP range can arbitrarily assign any host or domain name to those IP addresses. This is an easy method for malicious network owners to frame other people or mislead investigators.

Using the a.iana-servers.net name server found in previous queries, we lookup the FQDN behind the IP address:

```
$ nslookup 192.0.34.166 a.iana-servers.net > N_IPaddress.txt
```

If it is not the same as the website, we do a whois lookup on the FQDN and save it as evidence, it may be the web server owner:

```
$ whois -h whois.internic.net examplehoster.com # find registrar
$ whois -h whois.godaddy.com examplehoster.com > W_Webserver.txt
```

Here examplehoster.com and whois.godaddy.com were just used to show the procedure. The site www.example.com's IP address actually resolves back to itself correctly, indicating that the web server owner is probably already identified in a previous whois query. Either way, we make a note of the results in our report under the heading "Web Server Owners" since this is a potential point of responsibility.

4.7 Investigating the Upstream ISPs

We now try to find some information about adjacent networks using traceroute. This may take some time since some firewalls will silently drop certain ICMP traffic along the way, causing timeouts and gaps in the trace:

```
$ traceroute 192.0.34.166 > T_FromMyISP.txt
```

The data collected here will give us some addresses of upstream ISP routers to which the site is attached. It could also give us some clues as to the geographic region of the site. The last hops in the traceroute could be a potential points of responsibility. We do a whois lookup of alter.net (seen in our traceroute) and make a note in our report under "Upstream ISPs".

```
$ whois -h whois.internic.net alter.net # find registrar
$ whois -h whois.networksolutions.com alter.net > W_UpstreamISP.txt
```

4.8 Investigating the Routing Information

Another useful way to find information about the upstream providers is to analyze an IP's routing information. This will tell us the Autonomous System Number to which the site belongs, as well as contact information for that ASN. To find this, we look to the Internet Routing Registry (<http://www.irr.net>). The IRR maintains a list of routing registries which provide information about routes and routing policies[11][12] which can be queried with the standard whois command. While the IRR has a large list of routing registries, Merit Network Inc. (<http://www.radb.net>) has attempted to mirror most of them. We query their whois server with www.example.com's IP address as follows:

```
$ whois -h whois.radb.net 192.0.34.166 > W_IRR.txt
```

Here we see various bits of information, including the size of the block and the ASN to which the IP address belongs. Once we know the AS number, we can find out more by querying the Regional Internet Registry servers. The AS Numbers allocated to the RIRs can be found at IANA: <http://www.iana.org/assignments/as-numbers>. Not all whois servers use the same syntax for AS queries, but here are some examples:

- `whois -h whois.arin.net "a 11"`
- `whois -h whois.ripe.net AS559`
- `whois -h whois.lacnic.net 27670`
- `whois -h whois.apnic.net AS23552`

Using `whois.arin.net` (identified at IANA), we query for AS 20144 where `www.example.com`'s IP address belongs:

```
$ whois -h whois.arin.net 'a 20144' > W_AS.txt
```

We record any contact information in our report under the heading “AS Owners”

For `www.example.com` this procedure was fairly simple and straight forward. Other IP addresses may be more difficult to investigate, possibly requiring additional whois lookups at local routing registries. A detailed analysis of Internet routing investigation is outside the scope of this paper, but may be covered in future work.

4.9 Investigating the Physical Location

The physical location of a site is also of importance. If the ISP and general location of a site is known, it may be possible to determine the telecommunications carrier responsible for the lower layer network connections (leased lines, frame-relay, etc.). In smaller countries there is often a single carrier which manages the “last mile” communications links for the entire country. If the country of origin is known, the carrier may also be known, and should be added to the report as a potential point of responsibility under the heading “Telecommunications Carrier”.

Another possibility of determining the physical location of a site is to examine the type of ISP. If the site is using a co-location provider, then the server will be physically located at the ISP. In this case, the investigated site's geographical address is the same as the ISP.

4.10 Investigating the Email Owners

Email servers can also be interesting information for a case. Investigating the owners of a domain's email servers is just applying more of the same techniques as we've previously used. First, we'll query for the domain's 'mx' (Mail Exchanger) resource records using an authoritative server:

```
$ nslookup -type=MX example.com a.iana-servers.net > N_MXServer.txt
```

This will typically include a list of mail servers and backup mail relays which are accepting email on behalf of a domain⁸. The owners of these machines are potential points of responsibility for the suspect domain's email traffic. A whois lookup of each mail server's domain should be saved as evidence and the contacts noted in the report under the heading "Email Server Owners".

4.11 Finding Additional Information

Whois databases can be used for more than just domain name and IP address queries. As we saw with ASNs, searching can also be done with other keywords. Depending on the whois database used, searches for Names, Addresses, NIC handles, etc. can be queried. Check with the whois database for more information on creating queries. For example:

```
$ whois -h whois.networksolutions.com help
$ whois -h whois.ripe.net help
```

Other DNS Resource Records could also be queried for information. For example, RR's which may be of interest (if they exist) could be the Canonical Name (CNAME), the Responsible Person (RP) or Host Info (HINFO). For more information on other RR's used, see [13].

The investigation that we have just completed is somewhat simplified. There may be large numbers of servers providing redundant DNS, Web, or email services for a site. These are all potential points of responsibility and each one should be analyzed for inclusion in the report. In such cases, we simply number the files as follows:

```
$ whois -h networksolutions.com domainA.com > W_MXserver1.txt
$ whois -h networksolutions.com domainB.com > W_MXserver2.txt
$ whois -h networksolutions.com domainC.com > W_MXserver3.txt
$ whois -h networksolutions.com domainD.com > W_MXserver4.txt
and so on...
```

⁸example.com is a poor example here, since it has no mx records (probably for good reason)

Exhaustively collecting all this information will generate a lot of data, but the procedure outlined here will keep the process organized and manageable.

5 Packaging and Preserving the Evidence

Once we are satisfied that our collection of evidence is complete, we need to package it in a simple format that can be distributed to people who may need it for reference. We also need to include a method to ensure the evidence has not been altered. Before we do this, we exit the script program and leave the directory. Then we package the collected evidence using the Unix tar command:

```
$ exit
$ cd ..
$ tar cvf evidence.tar evidence
```

The tar command will create a single file which can be opened with most modern unpacking programs (like winzip). This command will also package any sub-directories that may have been created (like recursive downloads of websites). Tar will also preserve the file modification time-stamps of our evidence files (the time of collection in this procedure), and preserve the Unix user-ID of the investigator who created the files⁹

The final step in packaging our evidence is to make a cryptographic hash of the tar file. This hash needs to be kept separately for the purpose of verifying that the evidence integrity has been maintained. This is done as follows¹⁰:

```
$ md5sum evidence.tar > evidence.md5
```

if you have no 'md5sum' command, but you have openssl:

```
$ openssl md5 evidence.tar > evidence.md5
```

The evidence.md5 file should be included with the report, and also stored separately in a safe place. If a particular investigator needs to “sign off” on the evidence collection, a PGP signature of the hash or the tar file could also be taken.

⁹This entire investigation can be done without root privileges.

¹⁰md5sum and the MD5 algorithm are used in these examples but there are many other tools and algorithms available.

6 Presenting the Evidence

Our report so far has consisted almost entirely of contact information. For every potential point of responsibility we have identified names, phone numbers, fax numbers, email and physical addresses, and website names. Without going into too much technical detail, we have created a report during the course of the investigation that non-technical staff can use within the context of their roles (legal staff, for example).

The information in the report can be independently verified based on the data in the evidence.tar file. The integrity of the evidence.tar file can be verified with the evidence.md5 file (a copy of which should be kept separately in a safe place).

7 Conclusion and Future Work

This paper organizes the investigation of an Internet presence. It outlines the forensic analysis of domain names and IP networks and has completed the following:

- Defined the points of responsibility related to an Internet presence.
- Systematically collected and time-stamped the evidence which identifies these parties.
- Saved and packaged the evidence in an organized manner.
- Created a cryptographic hash of the evidence to ensure integrity is preserved.
- Created a verifiable report presenting the contact information found in the evidence.

Several areas were touched upon in this paper where future work could be done. The area of remote website acquisition for investigative purposes needs to be addressed. The forensic analysis of IP routing (including the core routing protocols) would be useful. Network forensics relating to IPv6 also needs to be analyzed in more depth.

References

- [1] Paul Albitz and Cricket Liu. DNS and BIND, 3rd Ed. 1998, O'Reilly & Associates

- [2] Bassam Halabi. Internet Routing Architectures, Cisco Press, 1997, New Riders Publishing
- [3] RFC2606 allows example.com to be used for documentation and testing. RFC 2606 - Reserved Top Level DNS Names, 1999
- [4] RIRs are listed at <http://www.iana.org/ipaddress/ip-addresses.htm>. Follow the links for more info about NIRs and LIRs
- [5] Virtual hosting allows multiple websites to be served from a single IP address. See section 14.23 of RFC 2616 - Hypertext Transfer Protocol – HTTP/1.1, 1999
- [6] This was a simple example using a class C address. For information on Classless networks see RFC 1519 - Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy, 1993
- [7] For information on time synchronizing a machine using NTP servers see <http://www.ntp.org>
- [8] RFC 1305 - Network Time Protocol (Version 3) Specification, Implementation, 1992
- [9] RFC 1771 - A Border Gateway Protocol 4 (BGP-4), 1995
- [10] RFC 3173 - Management Guidelines & Operational Requirements for the Address and Routing Parameter Area Domain (“arpa”), 2001
- [11] RFC-2622: Routing Policy Specification Language, 1999
- [12] RFC-2650: Using RPSL in Practice, 1999
- [13] RFC 1183 - New DNS RR Definitions, 1990

Document History

Aug-Oct 2004: Created original article

Oct 14, 2004: Submitted to Elsevier

Dec 14, 2004: Published in Digital Investigation Journal

April 27,2005: Minor corrections. Changed examples to use md5sum instead of md5

August 1,2005 Added table of contents and document history