UNIVERSITEIT VAN AMSTERDAM
System and Network Engineering

Passive LAN Information Gathering

RP2 project report

Roy Duisters
{roy.duisters@os3.nl

August 16, 2011

Abstract

One can gather a lot of information about corporate Local Area Network (LAN) environments by passively listening on these networks. This information can be used during the reconnaissance phase of a penetration test. We started by researching the information generally gathered during the reconnaissance phase of a penetration test.

We selected a sample of six broadcast / multicast protocols to passively obtain information from: mDNS, SMB Browser, DHCP, NBNS, STP and CDP. These protocols were researched on their functionality, details and usefulness for network profiling. The effectiveness of this technique has been benchmarked by comparing the generally gathered information to the information possible to passively obtain.

The research concludes that one can passively create a profile of a corporate LAN environment, depending on the availability of the protocols. The researched technique is most effective in obtaining information about the structure/architecture of the LAN environment.

We created a proof of concept implementation of the researched technique. The proof of concept parses a traffic capture, stores "interesting" data in a database, creates relationships between the data and generates a report of the gathered information.

Acknowledgements

The research team gratefully acknowledges the contributions to this research. We greatly appreciate the support given by KPMG Advisory. In particular we would like to thank:

- Marc Smeets and Michiel van Veen for their supervision, insight, ideas and feedback.
- The System and Network Engineering master programme of the University of Amsterdam for the possibility to conduct this research.

Contents

Chapter 1

Introduction

Broadcast and multicast traffic from various protocols can be observed by passively listening on corporate Local Area Network (LAN) environments. Protocols such as the Server Message Block (SMB) and Spanning Tree Protocol (STP) use broadcast / multicast traffic for communication.

One can gather a lot of information about the setup and the main systems in the network by simply listening to these freely obtainable packets. This information can be used during the reconnaissance phase of a penetration test, as it could allow the penetration tester to obtain information about the internal corporate network.

Conventional network reconnaissance methods rely on active methods to obtain information about the network. By actively scanning and probing nodes on the network, chances increase that the penetration test is detected. By using a passive method to obtain information, the risk of detection decreases.

The research has been formulated into the following research- and sub-questions:

Which information can be obtained by listening passively in a corporate LAN environment and how can this information be combined, correlated and reported to create an "outline" of the network, to simplify and prevent detection of the reconnaissance phase of a penetration test?

The following sub-questions combined will answer the main research question:

- *Which protocols are the most interesting to gather (and combine) information from, to create a "map" of a corporate LAN environment that can be used for a penetration test?*

Chapter 2 (Information selection criteria) describes the information generally gathered during the reconnaissance phase of a penetration test. Chapter 3 (Empirical methods) describes the empirical methods used for this research, such as the methods used to select the protocol samples.

- *Which information can be passively gathered about the network segment and the systems in it by making use of the protocol samples?*

Chapter 4 (Protocol Analysis) describes the analysis of the protocol samples. The protocols in the sample are analysed on functionality, details and usefulness to obtain information required to perform a successful penetration test.

- *How can the gathered information be presented, to create an "outline" of the network that can be used for a penetration test?*

The data gathered from the protocol samples is combined into information related to a node, Internet Protocol (IP) subnetwork (subnet) and LAN environment. Chapter 4 (Protocol Analysis) describes the methods used to combine the gathered data into information. Chapter 5 (Proof of concept implementation) describes the implementation of the researched technique.

Chapter 2

Information selection criteria

To determine whether passive information gathering is useful for the reconnaissance phase of a penetration test, the first step is to determine the information generally gathered during the reconnaissance phase of a penetration test. Afterwards, a comparison can be made between the generally gathered information and the information possible to gather by passive information gathering.

This chapter describes the information selection criteria used in this research. The first section (2.1 - Determining the information to gather) describes the information generally gathered during the reconnaissance phase of a penetration test. The second section (2.2 - Passively gathering the information) describes the theoretical possibilities to passively gather information on an enterprise LAN environment.

2.1 Determining the information to gather

The penetration tester gathers information about the enterprise environment during the reconnaissance phase of a penetration test. The gathered information is used in the next phases of the penetration test.

We based the information generally gathered during the reconnaissance phase of a penetration test on multiple papers [1][2]. These multiple sources of information are combined into an independent view of the information generally gathered during the reconnaissance phase of a penetration test.

We divided the information generally gathered during the reconnaissance phase of a penetration test in three categories:

- Organisation and procedures
- Security of the enterprise IT environment
- Architecture of the enterprise IT environment

The following subsections describe the information gathered in these categories. The categories are divided in subcategories, to provide a more detailed view of the information generally gathered during the reconnaissance phase of a penetration test. The (sub)categories do not provide an exhaustive list of the information generally gathered during the reconnaissance phase of a penetration test. Appendix A (Information to gather) shows examples of information gathered in these (sub)categories.

2.1.1 Organisation and procedures

The penetration tester generally gathers information about the organisation and its procedures during the reconnaissance phase of the penetration test. This information may help the penetration tester in the next phases of the penetration test.

Naming conventions Naming conventions may reveal information interesting for the penetration test. For instance, usernames of employees may use a standard naming convention. Servers storing human resources data may use a different naming convention than end-user workstations. If the penetration tester can interpret (or create) these names, he/she can make more well-informed decisions about the systems or usersnames to target.

The penetration tester generally gathers information about the following naming conventions of the enterprise environment:

- Domain names
- System names
  - Servers/services
  - Network devices
  - End-user systems
- Usernames

Organisational structure The penetration tester generally gathers information about the organisational structure of the enterprise, in order to target the correct systems (and possibly the correct people). For instance, if the penetration tester wants to gain access to certain components, it may not be effective to crack the password of a random employee. This employee may not have the required access to these components. It may be more interesting to crack the password of the administrator of the components. Information about the organisational structure may help the penetration tester to target the correct employees or departments. This information allows the penetration tester to make more well-informed decisions in the next phases of the penetration test, increasing his / her chances of success.

The penetration tester generally gathers the following information about the organizational structure of the enterprise:

- Hierarchical structure
- Departments/workgroups
- Employees
  - Job information
  - Access levels

Physical locations The penetration tester generally gathers information about the physical locations of the enterprise. This information may be helpful in the next phases of the penetration test for multiple reasons. For instance, the physical security may differ per physical location, while each of these physical

locations has access to valuable information. Information about the physical locations of the enterprise allows the penetration tester to target the weakest link.

The penetration tester generally gathers the following information about the physical locations of the enterprise:

- Physical security levels
- Permitted access levels to internal resources

Partner organisations Partner organisations may have access to information valuable to obtain during the penetration test. Information about partner organisations and their permissions to access information is valuable for the penetration test, since it could allow the penetration tester to target the weakest link.

The penetration tester generally gathers the following information about partner organisations of the enterprise:

- Network connections to these organisations
- Permitted access levels to internal resources

2.1.2 Security of the enterprise IT environment

For penetration testing purposes, it is important to gather information about the security of the enterprise Information Technology (IT) environment. The penetration tester can use this information to determine the risk of detection associated with a certain test. The penetration tester can decide whether a certain test should be executed, avoided or adapted, using this information.

Security plans and policies If the penetration tester has knowledge about the security plans and policies he/she can make more well-informed decisions. For instance, the password policy is a common security policy in enterprise IT environments. The policy is used to configure rules passwords must comply with (for instance, a minimum of eight characters, including one number). This policy is interesting for the penetration test, since it gives an indication of the strength of the passwords used in the organisation.

The penetration tester generally gathers information about the following security plans and policies of the enterprise:

- IT security policy
  - Password policy
  - Additional security policies
- Physical security policy

Technical security measures The penetration tester can adapt his/her strategy to avoid technical security measures of the enterprise LAN environment, if information about these security measures is available.

For instance, an enterprise often uses Intrusion Detection System (IDS) or/and Intrusion Prevention System (IPS) systems to monitor its network for malicious activities or policy violations. These devices may detect

malicious activity on the network (possibly due to the penetration test). After detecting malicious activity, these systems can alert an administrator or possibly take preventive actions (in the case of an IPS system). If the penetration tester has information about these security measures, he/she can make the correct decisions to avoid these systems.

The penetration tester generally gathers information about the following technical security measures of the enterprise environment:

- IDS/IPS systems
- Firewalls
- Logging devices
- Other security devices
- Security domain segmentations
  - Native/current VLAN
  - Other used VLANs

2.1.3 Architecture of the enterprise IT environment

The penetration tester generally gathers gather information about the architecture of the enterprise IT environment. This information can be used to determine attack vectors and systems to focus on, in the next phases of the penetration test.

Infrastructure The infrastructure of the enterprise LAN environment can reveal information valuable for a penetration test. This information can allow the penetration tester to target the weakest part of the infrastructure. For instance, a Virtual Local Area Network (VLAN) may be used to limit access to certain services. The penetration tester may use information about the used VLANs in the enterprise to circumvent configured security policies in the next phases of the penetration test.

The penetration tester generally gathers the following information about the infrastructure of the enterprise LAN environment:

- OSI layer 2 network infrastructure
  - STP information
  - VLAN information
- OSI layer 3 network infrastructure
  - IP addressing information
- Upper layer network infrastructure
  - Workgroup/domain infrastructure
  - File shares
  - Proxy services
  - First hop redundancy services
  - Other network services

Hard- and software The penetration tester should gather information about the hard- and software used in the enterprise LAN environment. This information allows the penetration tester to decide upon the correct strategy and methods in the next phases of the penetration test.

The penetration tester generally gathers the following information about the physical hard- and software used in the enterprise LAN environment:

- Naming information
- Platform information
- OS information
- Service information
- Role information

Important IT components The penetration tester requires information about important components in the enterprise LAN environment, to perform a successful penetration test. Information needs to be gathered about each component, to determine which components can be considered important in the enterprise LAN environment. For instance, one can determine the amount of clients connecting to a certain component. Here we assume, if a large amount of clients connect to a certain component, the component is bound to have some importance in the enterprise LAN environment.

The penetration tester generally gathers the following information, to determine the importance of the IT components in the enterprise LAN environment:

- Most accessed components
- Services offered by components
- Purpose of the components
- Components storing valuable information

2.2 Passively gathering the information

This research focuses on passively gathering information from an enterprise LAN environment. Therefore, one needs to passively capture network traffic from the enterprise LAN environment. However, due to the nature of modern-day switched LAN environments, one cannot obtain all traffic sent on a switched LAN environment. One can only receive certain types of traffic.

After passively obtaining the traffic capture, the obtained data requires analysis and combination into a "profile" of the enterprise LAN environment. This profile contains information useful for the reconnaissance phase of a penetration test. The profile will be based on network traffic passively receivable on a switched LAN environment.

2.2.1 Passively obtainable network traffic

Originally, IP only supported unicast traffic [3]. The Internet Engineering Task Force (IETF) added broadcast and multicast to the existing types of IP network traffic afterwards [4][5]. One cannot capture all traffic

sent between nodes in a switched/bridged LAN environment [6]. In theory, only broadcasts and packets sent to the all-nodes link-local multicast address should be received by every node in the LAN environment.

The main advantage of multicast over broadcast communication is that multicast messages are only received by nodes, if these nodes are a member of the multicast group where the messages are sent to. Since the traffic captures are obtained passively, the capturing device does not subscribe itself to any multicast groups. However, if the switched Open Systems Interconnection (OSI) layer 2 LAN environment does not have support for IP multicast, these packets are sent to every host in the network (similar to a broadcast) [7].

Therefore, we decided not to focus on received unicast packets, since these packets cannot be passively observed in switched/bridged LAN environments. The following chapter (3 - Empirical methods) describes the methods used to select the protocol samples.

2.3 Chapter summary

We divided the information generally gathered during the reconnaissance phase of a penetration test in three categories: "Organisation and procedures", "Security of the enterprise IT environment" and "Architecture of the enterprise IT environment".

The first category (Organisation and procedures) contains information about naming conventions, the organisational structure, physical locations and partner organisations. The second category (Security of the enterprise IT environment) contains information about the security plans and policies and the technical security measures. The third category (Architecture of the enterprise IT environment) contains information about the physical layout/architecture, hard- and software used and important IT components.

One cannot receive all traffic sent between every connected host on switched/bridged LAN environments. However, one can receive broadcast and (in some cases) multicast traffic. Therefore, the research focuses on information that can be passively obtained in enterprise LAN environments (broadcast and multicast traffic).

Chapter 3

Empirical methods

This chapter describes the empirical methods used in the research. The first section (3.1 - Protocol samples) describes the methods used to select the protocol samples. The second section (3.2 - Protocol analysis overview) describes the analysis of the protocol samples, in order to obtain information usable for the creation of a "profile" of the enterprise LAN environment. The final section (3.3 - Research limitations) describes the limitations of this research that may have influenced its results.

The empirical methods are directly related to the information selection criteria. The empirical methods strive to obtain the information as described in the information selection criteria, by making use of passive LAN information gathering. After analysing and combining the gathered data, it forms the profile of the enterprise LAN environment.

3.1 Protocol samples

We selected a number of different sources of information to create a profile of the enterprise LAN environment. This section describes the methods used for the selection of the protocol samples. The protocol samples strive to passively obtain the information described in the information selection criteria.

3.1.1 Required information to create a network profile

We first determined the information required to create a profile of an enterprise LAN environment. Section 2.1 (Determining the information to gather) describes the generally gathered information during the reconnaissance phase of a penetration test. The profile strives to provide as much generally gathered information as possible, in order to benchmark the effectiveness of this technique to perform network reconnaissance.

3.1.2 Commonly used protocols in a corporate LAN environment

The research focuses on commonly used protocols in enterprise LAN environments, to increase its applicability in different enterprise LAN environments. The results of passive LAN information gathering are highly dependent on the protocols used in the tested LAN environment. We analysed several different (passively obtained) traffic captures, to determine which protocols are commonly used in these enterprise LAN environments.

Table 3.1 shows a list of protocols commonly used in enterprise LAN environments, based on a combination of (passively obtained) traffic captures from five different enterprise LAN environments. The table only

shows protocols passively observed in more than one enterprise LAN environment.

OSI layer	Protocols	Occurrence
Layer 2	ARP	●●●●●
	STP	●●●●○
	LLDP	●●●○○
	CDP	●●○○○
Layer 3	DHCP	●●●●●
	IGMP	●●●○○
	ICMPv6	●●●○○
Upper layer	mDNS	●●●●○
	NBNS	●●●○○
	SMB Browser	●●●○○
	HTTP	●●○○○

Table 3.1: Protocol occurrence

3.1.3 Selection criteria

To select the protocol samples where the research will focus on, the previously described commonly used protocols serve as input. However, selecting only commonly used protocols to create a profile of the network may not yield the best results. For instance, a protocol may be very common in enterprise LAN environments, but may not carry information usable to create a profile of the enterprise LAN network. On the other hand, a certain protocol may carry information useful to create a profile, but may be rare in enterprise LAN environments, making it less useful for this research as well.

A trade-off between the applicability in multiple enterprise LAN environments and gathering information directly useful for the creation of the network profile is used. The following criteria are used to select the protocol samples:

- The protocol must be commonly used and passively observable in enterprise LAN environments
- The protocol must contain information needed to create the network profile

We gave each protocol two scores, based on its applicability to both criteria. The first score describes the occurrence of the protocol in enterprise LAN environments, as described in section 3.1.2 (Commonly used protocols in a corporate LAN environment). The second score displays the probability whether the protocol contains information to create a profile of the network, based on an initial analysis of the functionality of the protocol (as described in section 3.2.1 - Functionality).

Five traffic captures are used to determine the commonly used protocols in enterprise LAN environments. If a certain protocol is observed in one of these traffic captures, its score is increased by one. Therefore, the occurrence of a protocol is directly mapped to its score, on a scale from zero to five. The second score (determining whether the protocol contains information required for the network profile) is graded on a scale from zero to five as well. Both scores are weighed equally in the total score. The scores can be viewed in table 3.2.

The two protocols with the highest total scores were researched (SMB Browser and DHCP), before deciding the total amount of protocols to research. We decided to research six protocols, based on the time needed to research these two protocols.

OSI layer	Protocols	Occurrence score	Network profiling score	Total score
Layer 2	ARP	●●●●●	●○○○○	●●●●●●○○○○
	STP	●●●●○	●●●○○	●●●●●●●○○○
	LLDP	●●●○○	●●●○○	●●●●●●○○○○
	CDP	●●○○○	●●●●●	●●●●●●●○○○
Layer 3	DHCP	●●●●●	●●●○○	●●●●●●●●○○
	IGMP	●●●○○	●●○○○	●●●●●○○○○○
	ICMPv6	●●●○○	●●○○○	●●●●●○○○○○
Upper layer	mDNS	●●●●○	●●●●○	●●●●●●●●○○
	NBNS	●●●○○	●●●●○	●●●●●●●○○○
	SMB Browser	●●●○○	●●●●●	●●●●●●●●○○
	HTTP	●●○○○	●●●○○	●●●●●○○○○○

Table 3.2: Protocol scores

For this purpose, the six highest scoring protocols are placed in the "Processed" category. The category contains the following protocols:

- Multicast DNS (mDNS)
- Server Message Block (SMB) Browser
- Dynamic Host Configuration Protocol (DHCP)
- NetBIOS Name Service (NBNS)
- Spanning Tree Protocol (STP)
- Cisco Discovery Protocol (CDP)

As described earlier, the statistics of the commonly used protocols are based on a sample of traffic captures, obtained from five different enterprise LAN environments. Therefore, the possibility is present that other protocols may be very common in other enterprise LAN environments. Therefore, protocols that may yield interesting results but are not present in the "Processed" class are added to the "Optional" protocol class. The protocols in this class are optional. These protocols are researched, depending on the time needed for the research of the protocols in the "Processed" class. The following protocols are in the "Optional" protocol class:

- Link Layer Discovery Protocol (LLDP)
- Address Resolution Protocol (ARP)
- Possibly others that are discovered during the research

3.2 Protocol analysis overview

The protocol samples are analysed in three steps. The first step is an analysis of the functionality of the protocols. The second step is to research the details of these protocols, such as the types of packets used by the protocol and the internal fields of these packet types. The third step is to determine the usefulness of this protocol for the creation of a profile of the enterprise LAN environment.

3.2.1 Functionality

We analysed the protocols based on their functionality. One can determine whether the protocol contains information usable for the creation of a network profile, based on an analysis of the functionality. For instance, certain protocols use broadcast/multicast communication to locate specific services in the LAN environment. Others may use broadcast/multicast to distribute network information. These different functionalities are analysed to determine whether the protocol can be used to gather information necessary for the creation of a profile of the enterprise LAN environment.

The information gathered from this analysis is used for the selection of the protocol samples.

3.2.2 Protocol details

The research analysed the details of the protocol samples in two steps. The first step in the analysis of the protocol details is the analysis of the protocol packets. These packets are analysed, to select the most interesting packets to (passively) gather the required information. For instance, certain packets only request information, rather than advertise information. These request packets often contain less information than packets that advertise information.

The next step in the analysis of the protocol details is to dissect the "interesting" packets. In this step, information interesting for the creation of the network profile of the enterprise LAN environment is extracted from the protocol.

3.2.3 Usability for the network profile

We analysed the usability of the protocol for the creation of a profile of an enterprise LAN environment in two steps. The first step is to determine the possibility to obtain information interesting for the creation of the network profile. For instance, SMB Browser HostAnnouncement packets are periodically transmitted (by default every 12 minutes). Since this information is periodically transmitted, the chance to passively obtain this information is relatively high. On the other hand, DHCP Acknowledgement packets are only sent when a host renews its DHCP lease or initiates the DHCP process (to request an IP address). Therefore, the chance to obtain a DHCP Acknowledgement packet is lower than the chance to capture a SMB Browser packet (if both protocols are present in the enterprise LAN environment).

The second step is to determine whether the protocol contains information useful for the reconnaissance phase of a penetration test (as described in section 2.1 - Determining the information to gather).

3.3 Research limitations

A number of limitations have influenced the research:

The results are dependent on the availability of the protocol samples - A sample of protocols has been selected, that will be analysed and implemented in the proof of concept. The effectiveness of this method to passively gather information is highly dependent on the availability of these protocols.

The protocol samples are highly dependent on the analysed traffic captures - To determine which protocols are commonly used, five different traffic captures have been analysed. However, it is possible

that these five captures contain protocols uncommon in other enterprise LAN environments, making the research less applicable to different environments.

3.4 Chapter summary

After determining the information generally gathered during the reconnaissance phase of a penetration test (the "required" information, as described in chapter 2 - Information selection criteria), the next step is to determine the methods to conduct the research. The empirical methods are directly related to the information selection criteria. The empirical methods strive to obtain the information generally gathered during the reconnaissance phase of a penetration test.

The first section (3.1 - Protocol samples) describes the selection of the protocol samples. We selected a sample of six protocols, based on two selection criteria, namely: "The protocol must be commonly used and passively observable in enterprise LAN environments" and "The protocol must contain information needed to create the network profile". The six protocols showing the highest affinity with these two criteria are selected as the protocol samples (mDNS, SMB Browser, DHCP, NBNS, STP and CDP).

The second section (3.2 - Protocol analysis overview) describes the methods used to analyse the protocol samples, in order to obtain information usable for the creation of a network profile. The protocols are researched on functionality, details and usefulness for network profiling.

The third section (3.3 - Research limitations) describes the limitations of this research that may have influenced the results of this research. The main limitation is that the effectiveness of this technique to obtain information is highly dependent on the availability of the selected protocols and obtained traffic captures.

Chapter 4

Protocol Analysis

This chapter describes the results of the protocol analysis. Section 3.1 (Protocol samples) describes the methods used for the selection of the protocol samples. Two classes of protocols were defined, namely the "Processed" and "Optional" classes. Only the protocols in the "Processed" class are described in this chapter.

The first section of this chapter, (4.1 - Analysis of the selected protocols) describes the analysis of the protocol samples. The second section (4.2 - Combining data) describes the process of combining the data, to create information usable for the network profile. The final section (4.3 - Protocol effectiveness) shows a comparison between the information generally gathered during the reconnaissance phase of a penetration test and the data passively obtainable using the selected protocols.

4.1 Analysis of the selected protocols

Section 3.1.3 (Selection criteria) describes the criteria used to select the protocols that will be researched. This section describes the analysis of the protocols in the "Processed" category (mDNS, SMB Browser, DHCP, NBNS, STP and CDP). The protocols are analysed on functionality, details and usability for network profiling.

4.1.1 mDNS

The Multicast Domain Naming System (mDNS) is selected based on its occurrence in enterprise LAN environments and the possibility that it contains data useful for the creation of a network profile (as described in section 3.1 - Protocol samples).

mDNS makes use of multicast communication. The protocol operates as an application-layer protocol. mDNS packets of other network nodes were observed in multiple traffic captures, indirectly causing the selection of mDNS as a part of the protocol samples. If the switched OSI layer 2 LAN environment does not have support for IP multicast, these packets are flooded to every host in the network (similar to a broadcast).

Functionality

The mDNS specification provides a standard to perform Domain Naming System (DNS) queries over IP Multicast [8].

Clients can use mDNS to convert a hostname to its associated IP address. In that way, it offers functionality

similar to DNS [9], without the need for a dedicated DNS server. In contradiction to the regular DNS system, mDNS can only be used to obtain IP addresses of systems in its own Local Area Network (LAN) environment. mDNS makes use of a link-local multicast address (224.0.0.251) for communication.

An extension to the DNS system, DNS-based Service Discovery (DNS-SD) allows a client to discover services on a network, using either the regular DNS system or mDNS [10].

Given a type of service that a client is looking for, and a domain in which the client is looking for that service, this convention allows clients to discover a list of named instances of that desired service, using only standard DNS queries. In short, this is referred to as DNS-based Service Discovery, or DNS-SD. [10]

The mDNS protocol (in combination with DNS-SD), forms the basis for Apple's service called Bonjour, used for zero-configuration networking [11].

Protocol details

mDNS uses flags to differentiate between query requests and query responses. Both the query requests and the query responses make use of multicast communication. The mDNS packet fields are the same in every mDNS packet. The fields of the packet are described in table 4.1.

Fields	Description
Transaction ID	The transaction ID of the packet
Flags	The flags indicate specific details of the packet (such as whether the packet is a response or query)
Questions	The amount of questions in the packet
Answer RRs	The amount of answer resource records (RRs) in the packet
Authority RRs	The amount of authority resource records (RRs) in the packet
Additional RRs	The amount of additional resource records (RRs) in the packet
Queries / Answers	The mDNS queries / answers of the packet

Table 4.1: mDNS packet fields [8]

Both the mDNS query request and query response packets are interesting for the creation of the network profile. Both packets give an insight into the DNS names clients request and the responses received. DNS-based Service Discovery (DNS-SD) allows a client to discover services on a network, using either the regular DNS system or mDNS [10]. DNS-SD makes use of the DNS TXT record to query services in the network.

Usability for network profiling

mDNS packets are not transmitted periodically. mDNS packets are sent when a client requests (or sends) a certain piece of data and uses mDNS to request or transmit this data.

The mDNS protocol can be used to identify systems in the network. The queries obtained from mDNS give an insight into the DNS names clients are requesting. These queries may give data about the services of the client and the services available on the network. For instance, suppose a client issues a mDNS query request for the A record of "WPAD". The "WPAD" query is sent by the proxy auto-discovery mechanism of

Microsoft Internet Explorer. If no query response is received (and the client possibly issues the same request multiple times), one can assume no proxy service is available on the internal network.

Since DNS-SD makes use of mDNS, these service-discovery queries can also be observed. The service discovery queries may be interesting to identify running services in the network, or to identify systems requesting a certain service. However, a detailed analysis of DNS-SD protocol is out of the scope of this research.

Section 2.1 (Determining the information to gather) describes the categories of information generally gathered during the reconnaissance phase of a penetration test. The following paragraphs describe the data obtainable from the mDNS protocol in these categories.

Organization and procedures The data obtained using mDNS about the organization and its procedures can be used as data about the naming conventions used in the enterprise environment. If clients request data using a Fully Qualified Domain Name (FQDN), the protocol may also reveal data about the physical locations and possible partner organizations of the enterprise.

Security of the enterprise IT environment mDNS can obtain minimal data about the security of the enterprise IT environment. One could obtain data about the technical security measures in place, if a client requests a name that can be mapped to a security device (such as IDS01 or Firewall3).

Architecture of the enterprise IT environment The data one can obtain using mDNS about the architecture of the enterprise IT environment is mainly knowledge about components in the enterprise LAN environment. This data can provide an indication of the level of importance of the component. For instance, one can gather data about the most-queried names, indicating a level of importance. Due to the data obtained about the naming of the components, one may determine where possibly company sensitive data is stored. If certain specific queries are sent (such as "WPAD", as described earlier) one may also obtain data about the services running on the clients and in the enterprise LAN environment.

4.1.2 SMB Browser

The Server Message Block (SMB) (also known as Common Internet File System (CIFS)) Browser protocol is selected, based on its occurrence in enterprise LAN environments and the possibility that it contains data useful for creating a profile of the LAN environment (as described in section 3.1 - Protocol samples).

The SMB protocol operates as an application-layer protocol. The protocol is mainly used by systems using the Microsoft Windows operating system. It is used to provide shared access to files, printers, serial ports, and miscellaneous communications between nodes on a network. The SMB protocol acts as a framework, consisting of multiple protocols (such as the Browser protocol) [12]. Most SMB protocols make use of unicast communication and are thus not observed when passively listening to the network.

Functionality

One of the protocols operating in the SMB framework is the Browser protocol. The Browser protocol has been specified in 1997 by the IETF [13]. The Browser protocol is used as a mechanism for discovering systems offering particular services on the local area network. To discover these services, the protocol makes

use of IP broadcast communication. Therefore, all nodes on the same broadcast domain will receive this communication.

Protocol details

As mentioned before, the SMB protocol consists of multiple layers. The Browser service is built upon a number of these layers. The Browser service makes use of the "Mailslot" layer. This "Mailslot" layer is carried within another layer, namely within SMB "transact" layer. The Browser protocol fields start with a field called "Transact data". The "Transact data" field contains an opcode, indicating the type of SMB Browser packet. The opcodes (and their associated description) are shown in table 4.2.

Opcode	Description
1	HostAnnouncement
2	AnnouncementRequest
8	RequestElection
9	GetBackupListReq
10	GetBackupListResp
11	BecomeBackup
12	DomainAnnouncment
13	MasterAnnouncement
15	LocalMasterAnnouncement

Table 4.2: SMB Browser Opcode description [13]

Only a number of these SMB Browser packet types contain data usable for the creation of a network profile. For instance, the "Request" packets, (such as GetBackupListReq and RequestElection) are merely sent to request data from other hosts. These packets do not contain data useful for the creation of a network profile (besides host discovery).

On the other hand, the "Announcement" packets contain more data about the connected LAN environment. The packet fields of the Host-, Domain- and LocalMasterAnnouncement packets can be viewed in table 4.3.

Announcement fields	Description
Opcode	Described in table 4.2
UpdateCount	Field must be set on 0 and ignored on receipt
Periodicity	Indicates the period (in seconds) after which a new Announcement is sent
ServerName	Used for multiple purposes. For instance, it can be used as the name of the domain (Domain Announcement) or server (Host Announcement)
VersionMajor	Indicates the major version of the OS sending the Announcement
VersionMinor	Indicates the minor version of the OS sending the Announcement
Type	Specifies the type of the server
Signature	The browser protocol minor version number
Comment	ASCII comment for the server

Table 4.3: SMB Browser fields of Host-, Domain- and LocalMasterAnnouncement packets [13]

The "Type" field in the SMB Browser "Announcement" packets indicates the capabilities (or services)

offered by the sending device. For instance, these flags indicate whether the device is a workstation, a domain controller, a Structured Query Language (SQL) server, a time source, has a print queue, etcetera. These flags allow the penetration tester to obtain data about the role of the node.

The MasterAnnouncement packets have a different structure than the Host-, Domain- and LocalMasterAnnouncement packets. The structure of the MasterAnnouncement packets can be viewed in table 4.4.

Master Announcement fields	Description
Opcode	Described in table 4.2
Master Browser Server Name	Specifies the name of the Master Browser server

Table 4.4: SMB Browser fields of MasterAnnouncement packets [13]

The "Announcement" packets have a default update periodicity. For instance, clients running Microsoft Windows 7 transmit the HostAnnouncement packet every 12 minutes by default. Every 12 minutes each SMB Browser-capable host advertises its existence on the local network to its SMB Browser-capable neighbours. Due to these regular timing intervals, one can use the regular HostAnnouncement packets as a host-discovery mechanism.

Usability for network profiling

By default, SMB-capable clients periodically advertise their existence on the network using the SMB Browser protocol. The SMB Browser protocol can be used to identify systems in the network. The SMB Browser service has the capability to provide data about the system itself, such as the name, operating system and type/purpose of the system.

Section 2.1 (Determining the information to gather) describes the categories of information generally gathered during the reconnaissance phase of a penetration test. The following paragraphs describe the data obtainable from the SMB Browser protocol in these categories.

Organization and procedures The data obtained by the SMB Browser protocol about the organization and its procedures is mainly knowledge about the naming conventions used in the enterprise environment. One can regularly gather data about the naming conventions in the network due to the regular update periodicity of the SMB Browser "Announcement" packets.

Security of the enterprise IT environment The data obtained by the SMB Browser protocol about the security of the enterprise IT environment is minimal. The only way one could obtain data about the technical security measures would be if security devices advertise their presence on the network through the SMB Browser protocol. The device could only be recognised if the hostname resembles the name of a security device, like IDS01 or Firewall2.

Architecture of the enterprise IT environment The main strength of the SMB Browser protocol lies in its ability to obtain data about the architecture of the enterprise IT environment. The SMB Browser protocol advertises the hostname, Operating System (OS), capabilities and other data regularly on the network. One can regularly obtain data about the components in the network and determine whether these

components are important or not, due to this combination. Due to the advertised OS version and capabilities, it also provides data about the software used in the enterprise LAN environment.

4.1.3 DHCP

The Dynamic Host Configuration Protocol (DHCP) is selected based on its occurrence in enterprise LAN environments and the possibility that it contains data useful for network profiling (as described in section 3.1 - Protocol samples).

DHCP is an automatic configuration protocol used on IP networks. The protocol operates as an application-layer protocol. DHCP allows a host to automatically configure its IP addressing.

Functionality

DHCP has been specified in 1997, in Request for Comments (RFC) 2131 [14]. DHCP provides a framework for passing configuration information to hosts on an IP network. DHCP is based on the Bootstrap Protocol (BOOTP). DHCP adds functionality to BOOTP, such as the automatic allocation of reusable network addresses and additional configuration options. These additional options are specified in RFC 2132 [15]. These options are used to send specific configuration options to the DHCP client.

Protocol details

Table 4.5 describes the packets used by the DHCP protocol. These packets are sent between the client and the server and used to sent or request information.

Depending on the state of the client, DHCP messages can be sent to a broadcast or unicast address [14]. For instance, a client may renew its IP address using DHCP. Since the client requested an IP address before, it knows the address of the DHCP server. This DHCP message would be sent as unicast, instead of broadcast. Whether the server or client broadcasts a message depends on the state of the client.

Message	Description/use
DHCPDISCOVER	Client broadcast to locate available servers
DHCPOFFER	Server to client in response to DHCPDISCOVER with offer of configuration parameters
DHCPREQUEST	Client message to servers, requesting the specified address
DHCPACK	Server to client with configuration parameters, including committed network address
DHCPNAK	Server to client indicating client's notion of network address is incorrect
DHCPDECLINE	Client to server indicating network address is already in use
DHCPRELEASE	Client to server relinquishing network address and cancelling remaining lease
DHCPINFORM	Client to server, asking only for local configuration parameters

Table 4.5: DHCP packet description [14]

Only a few of these different packets contain data usable to create a profile of the network. For instance,

certain "Request" packets are merely sent from the client to the DHCP server to indicate that the client requires certain information. These packets do not contain data useful for the creation of a network profile, other than host discovery.

Clients can automatically configure themselves with the correct domain, gateway, DNS servers or other information using DHCP. These configuration parameters and other control information are carried in data items, stored in the "options" field of the DHCP message. The data items themselves are also called "options" [15].

A number of commonly used options are described in table 4.6. These options are only contained in specific DHCP packets. For example, the DHCPACK sent by the DHCP server to the DHCP client contains these options. These options can be used to gain knowledge about services in the network and the clients configured to use these services.

Option	Description/use
Subnet Mask	The subnet mask specifies the subnet mask to be used by the client
Time Offset	The time offset field specifies the offset of the client's subnet in seconds from Coordinated Universal Time (UTC)
Router Option	The router option specifies a list of IP addresses for routers on the client's subnet
Time Server Option	The time server option specifies a list of time servers available to the client
Name Server Option	The name server option specifies a list of IEN 116 name servers available to the client
Domain Name Server Option	The domain name server option specifies a list of Domain Name System (STD 13, RFC 1035) name servers available to the client
Domain Name	This option specifies the domain name that client should use when resolving hostnames via the Domain Name System
Additional options	Additional options are described in [15]

Table 4.6: DHCP options [15]

Usability for network profiling

DHCP packets are not sent periodically. DHCP packets are sent when a client initiates the DHCP mechanism to obtain information. The DHCP protocol can allow one to passively obtain data about the services the DHCP clients are configured to use. For instance, DHCP can be used to obtain data about the configured gateway, subnet mask, name servers, domain names, time server, et cetera.

Section 2.1 (Determining the information to gather) describes the categories of information generally gathered during the reconnaissance phase of a penetration test. The following paragraphs describe the data obtainable from the DHCP protocol in these categories.

Organization and procedures The data obtained from the DHCP protocol about the organization and its procedures is mainly knowledge about the naming conventions used in the enterprise environment. However, one requires a traffic capture obtained over a long period of time before every host in the network has advertised its existence, since DHCP packets are not transmitted periodically.

Security of the enterprise IT environment The data obtained by the DHCP protocol about the security of the enterprise IT environment is minimal. The only way one could obtain data about the technical security measures would be if these devices use the DHCP protocol to request configuration information. Perhaps the device would be recognised if the hostname resembles the name of a security device, like IDS01 or Firewall2.

Architecture of the enterprise IT environment The main usefulness of the DHCP protocol lies in its ability to obtain data about the architecture of the enterprise IT environment. The DHCP options sent to DHCP clients can be used to determine which services the clients use and which component provides this service. Therefore, the DHCP protocol provides data useful to determine the important IT components of the enterprise LAN environment.

4.1.4 NBNS

The NetBIOS Name Service (NBNS) is selected based on its occurrence in enterprise LAN environments and the possibility that it contains data useful for network profiling (as described in section 3.1 - Protocol samples).

The NBNS protocol is used to discover the IP address(es) associated with a NetBIOS name. NBNS uses a name query (also known as "resolution" or "discovery") broadcast to obtain this information. The protocol operates as an application-layer protocol.

Functionality

NetBIOS Name Service (NBNS) is part of the NetBIOS service and was first specified in 1987, in RFC 1001 [16]. Whereas most NetBIOS services make use of unicast communication, NBNS makes use of broadcast communication. NetBIOS Name Service uses broadcast for the Name Query (discovery) packet.

The Name Query packet is used to resolve NetBIOS names to their associated IP address. However, in some cases the Name Query does not use broadcast communication to obtain the IP addresses associated to a specific NetBIOS name. One can implement a central repository, or Name Service, that records all NetBIOS name registrations. If the client has been configured to use a name service for NetBIOS names (such as Microsoft Windows Internet Name Service (WINS)), the NetBIOS Name Service will query this central repository (using unicast) instead of issuing a broadcast on the local LAN environment. Thus, when a central repository is in use, NetBIOS name queries cannot be passively observed.

Protocol details

Table 4.7 describes the packets used by the NBNS protocol. As indicated in the table, some of these packets are always broadcasted, while others are always unicasted.

Only a number of these packets can be used for the creation of the network profile. For instance, the packets sent as unicast cannot be passively observed. However, the "Name registration request" and "Name query request" can be passively observed on the LAN environment. The "Name registration request" is used by clients to advertise their own NetBIOS name on the network, to determine whether the NetBIOS name is free to use. The other packet, the "Name query request" is more interesting for the creation of the network

Message	Description/use	Packet type
Name registration request	A NetBIOS client attempt to register its NetBIOS name, to check if the name already exists	Broadcast
Negative Name registration response	A Response to the name registration request, notifying the client that the name already exists	Unicast
Name query request	Name query transaction to obtain the addresses associated with a NetBIOS name	Broadcast
Positive name query response	Response to the name query request, containing the addresses associated with a NetBIOS name	Unicast

Table 4.7: NBNS packet description [16]

profile, since it gives an insight into the NetBIOS names clients are requesting. Table 4.8 shows the fields a NBNS name query request packet consists of.

Fields	Description/use
Flags	The flags indicate specific details of the packet (such as whether the packet is a response or query)
Questions	The amount of questions in the packet
Answer RRs	The amount of answer resource records (RRs) in the packet
Authority RRs	The amount of authority resource records (RRs) in the packet
Additional RRs	The amount of additional resource records (RRs) in the packet
Queries	The NetBIOS name queries (the name that is being requested)

Table 4.8: NBNS name query request fields [16]

Usability for network profiling

NBNS packets are not transmitted periodically. NBNS packets are transmitted when a client requires (or offers) information and uses NBNS to request or transmit this information. The some cases, the NBNS protocol cannot be passively observed. When a central repository (or Name Service) is implemented, NetBIOS name queries are sent (as unicast) to this central repository, rather than broadcasted on the LAN environment.

The NBNS protocol can be used to identify systems in the network. The queries obtained from NBNS give an insight into the NetBIOS names clients are requesting. These queries may again provide data about the services of the client and the services available on the network.

Section 2.1 (Determining the information to gather) describes the categories of information generally gathered during the reconnaissance phase of a penetration test. The following paragraphs describe the data obtainable from the NBNS protocol in these categories.

Organization and procedures The data obtained about the organization and procedures is mainly knowledge about the naming conventions used in the enterprise environment.

Security of the enterprise IT environment The data one can obtain using NBNS about the security of the enterprise IT environment is minimal. One could obtain data about the technical security measures in place, if a client requests a name that can be mapped to a security device (such as IDS01 or Firewall3).

Architecture of the enterprise IT environment The data one can obtain using NBNS about the architecture of the enterprise IT environment is mainly knowledge about components in the enterprise LAN environment. Due to this data, one may consider certain components more important than others. For instance, one can gather data about the most-queried names, indicating a level of importance of these systems. Due to the data obtained about the naming of the components, one can determine where possibly company sensitive data is stored. If certain specific queries are sent (such as "WPAD", as described earlier in mDNS) one may also obtain data about the services running on the clients and in the enterprise LAN environment.

4.1.5 STP

The Spanning Tree Protocol (STP) has been selected, based on its occurrence in enterprise LAN environments and the possibility that it contains data useful for network profiling (as described in section 3.1 - Protocol samples).

STP is a network protocol that ensures a loop-free topology for any bridged Ethernet local area network. The basic function of STP is to prevent bridge loops. STP allows a layer 2 network design to include redundant links to provide a backup path if an active link fails, without the danger of bridge loops. STP provides an automated way to prevent bridge loops, without the need to manually enable/disable backup links.

Functionality

STP is a Data Link Layer protocol. It is standardized as Institute of Electrical and Electronics Engineers (IEEE) 802.1D [17]. STP is used to create a loop-free topology, to prevent bridge loops in a redundant topology. Since 1998, STP has been superseded by the Rapid Spanning Tree Protocol (RSTP):

> *In IEEE Std 802.1D, 1998 Edition, and prior editions of this standard, this clause specified the spanning tree algorithm and protocol (STP). STP has now been superseded by the Rapid Spanning Tree Protocol (RSTP) specified in Clause 17 of this standard.* [17]

Protocol details

Since STP operates on the Data Link Layer, it makes use of OSI layer 2 communication. STP uses Bridge Protocol Data Unit (BPDU) frames for communication. BPDU frames contain information on ports, addresses, priorities and costs. Bridges do not forward received BPDUs. Instead, they generate new BPDUs using the information in the received BPDU. BPDUs are (by default) flooded on every port of the bridge. The different BPDU types are shown in table 4.9.

The Topology Change Notification and Acknowledgement BPDUs are sent when the network topology changes. STP differentiates between the Topology Change Notifications and the Configuration BPDUs using flags. The Configuration BPDUs contain information similar to the Topology Change Notification BPDUs. The Configuration BPDUs are easily obtained since these are flooded every two seconds on every

BPDU type	Description/use
Configuration BPDU	BPDU used for the Spanning Tree computation
Topology Change Notification	BPDU used to announce changes in the network topology
Topology Change Notification Acknowledgment	BPDU used to acknowledge the Topology Change Notification BPDU

Table 4.9: STP BPDU types [17]

port, by default. Therefore, one can easily obtain Configuration BPDUs, while Topology Change Notification BPDUs are only sent when the STP configuration changes. The BPDU fields are displayed in table 4.10.

Fields	Description/use
Protocol ID	Identify different protocols supported by Spanning Tree Protocol entities
Version ID	Identifies the current version of the protocol being used
BPDU type	Field that identifies the type of BPDU being sent
BPDU flags	Field that contains flags used in response to a Topology Change Notification BPDU
Root ID	Field that contains the bridge identifier of the root for the spanning tree being deployed
Root path cost	Field used to indicate the cost of the path from the transmitting bridge to the root
Bridge ID	Field containing the bridge identifier of the transmitting bridge
Port ID	Field used to identify the port via which this BPDU was transmitted
Message ID	Field used to indicate the age of the current Configuration BPDU
Max age	Field used to indicate a timeout value to be used by all bridges in the STP domain
Hello time	Field defining the time interval between generation of Configuration BPDUs by the root
Forward delay	Field that defines the time a bridge port must wait in the listening state, and then again in the learning state, before entering the forwarding state

Table 4.10: STP BPDU fields [17]

An interesting detail is that Cisco's Per VLAN Spanning Tree (PVST) implementation adds a VLAN identifier (ID) to the Root and Bridge ID. This data could be used to determine the VLAN ID from which the traffic capture was obtained.

Usability for network profiling

STP Configuration BPDUs are flooded periodically on the network. Section 2.1 (Determining the information to gather) describes the categories of information generally gathered during the reconnaissance phase of a penetration test. The following paragraphs describe the data obtainable from the STP protocol in these categories.

Organization and procedures STP does not carry any data about the organization and its procedures.

Security of the enterprise IT environment The data one can obtain using STP about the security of the enterprise IT environment is minimal. One could obtain data about the technical security measures in place, due to Cisco's PVST implementation addition of the VLAN ID to the Root and Bridge IDs. By passively capturing network traffic on multiple points of the enterprise network, one could create a map of the VLANs in use.

Architecture of the enterprise IT environment The data one can obtain from STP about the architecture of the enterprise IT environment is mainly knowledge about the layer 2 infrastructure and components of the enterprise LAN environment. For instance, one can gather data about the spanning tree itself. Data such as the Root bridge, distance to the root bridge and possibly the associated VLAN of the spanning tree can be obtained. These components can be considered important IT components of the enterprise LAN environment.

4.1.6 CDP

The Cisco Discovery Protocol (CDP) is selected based on its occurrence in enterprise LAN environments and the possibility that it contains data useful for network profiling (as described in section 3.1 - Protocol samples).

CDP is a proprietary Data Link Layer network protocol developed by Cisco Systems. It is used to share information to directly connected network equipment, such as the operating system version and management IP address of the sending device. Since CDP is a proprietary protocol, it has not been specified in an open standard. The Link Layer Discovery Protocol (LLDP) is a standardised alternative for CDP.

Functionality

CDP is used to share information between directly connected network equipment. CDP announcements are sent every 60 seconds by default, using multicast communication.

The information contained in CDP announcements varies between devices, depending on the type and operating system version of the device [18]. The information contained in these announcements may include the operating system version, device ID (hostname), configured addresses, the port identifier from which the announcement was sent, device type, model, et cetera.

Protocol details

CDP uses one packet type: the CDP announcement. CDP announcement packets contain data about the sender of the CDP announcement. The packet consists of a header followed by a set of variable-length fields. One can only obtain data about the directly connected device, since CDP packets are not forwarded between devices.

Table 4.11 shows the fields used in the CDP announcement packets. Since CDP makes use of the extendible Type Length Value (TLV) frame format, the details of these announcements may differ. Certain devices may have more TLVs implemented than others.

Fields	Description/use
Version	The Version field indicates the version of CDP being used
Time-to-Live	The Time-to-Live field indicates the amount of time, in seconds, that a receiver should retain the information contained in the packet
Checksum	The Checksum field indicates the standard IP checksum
Type	The Type field indicates the type/length/value type
Length	The Length field indicates the total length, in bytes, of the type, length, and value fields
Value	The Value field contains the type/length/value value, depending on the type/length/value type

Table 4.11: CDP packet fields [18]

As mentioned before, the TLVs of the CDP announcements may differ. The most common TLV values are described in table 4.12. Often, additional TLVs can be observed in CDP announcement packets, such as the native VLAN, Voice over Internet Protocol (VoIP) VLAN, VLAN Trunking Protocol (VTP) management domain, duplex and available power.

TLV name	Description/use
Device ID	The device ID is the fully-qualified domain name
Address	The configured addresses of the device
Port ID	The port identifier of the port that sent the CDP announcement
Capabilities	Flags that indicate the capabilities of the device (for instance, whether it is capable of layer 3 routing)
Version	A character string that provides information about the software release version that the device is running
Platform	A character string that describes the hardware platform of the device
IP Prefix	The IP prefix string contains information about the configured IP prefix on the device
Additional TLVs	..

Table 4.12: CDP TLV types [18]

Usability for network profiling

The CDP announcement packets are sent periodically on regular timing intervals. By default, every CDP-capable network device advertises its existence and (part of) its configuration every 60 seconds.

Section 2.1 (Determining the information to gather) describes the categories of information generally gathered during the reconnaissance phase of a penetration test. The following paragraphs describe the data obtainable from the CDP protocol in these categories.

Organization and procedures Using CDP, one can obtain minimal data about the organization and its procedures. Since CDP frames are not forwarded between devices, one can only obtain data about the connected device. However, one can obtain naming data of the connected CDP device.

Security of the enterprise IT environment One can obtain some data about the security of the enterprise IT environment using CDP. CDP advertises data about the configured VLANs, such as the native VLAN and VoIP VLAN. This data can be used to obtain data about the segmentation of security domains in the enterprise LAN environment.

Architecture of the enterprise IT environment The main usefulness of the CDP protocol lies in its ability to obtain data about the architecture of the enterprise IT environment. The CDP protocol can be used to obtain data about the hard- and software used by CDP-capable devices. One can also obtain data about the layout of the enterprise LAN environment, since CDP advertises data about the configured VLANs.

4.2 Combining data

The data gathered from the protocol samples is fragmented. The data requires combination into information, to make it useful for the reconnaissance phase of a penetration test. For instance, a DHCP packet may indicate a node uses a certain DNS name server. Other protocols, such as the SMB Browser and NBNS protocol may have gathered information about the same DNS name server as well. The fragmented data gathered from these different protocols needs to be linked to a single system, to transform the data into information usable for the reconnaissance phase of a penetration test. Then, the system can be linked to a single network or infrastructure to create a profile of the enterprise LAN environment.

To combine the data and transform it into information usable for the reconnaissance phase of a penetration test, three steps are taken.

4.2.1 Mapping data gathered from multiple protocols to a single node

As mentioned before, the data gathered from the different protocols is fragmented. The first step to transform the data into usable information is to map the data to a single system.

The researched protocols operate on different layers of the OSI model. Protocols such as CDP and STP operate on the Data Link Layer of the OSI model (layer 2). Other protocols, such as NBNS and mDNS operate on upper layers. This complicates the task to map data to a single node.

One can gather the source addresses of the node, from the different OSI layers (such as the MAC and IP address). Data from the same source IP address can be mapped to a single node. However, protocols operating on the Data Link Layer (OSI layer 2) do not have a source IP address. Therefore, combining data obtained from these protocols (such as CDP and STP) requires another method. Source MAC address cannot be directly mapped to IP addresses, since source MAC address of frames are rewritten by intermediate upper layer devices (such as routers).

To create a correlation between an IP address and a MAC address, data from other protocols is required. Broadcast protocols (such as NBNS), transmit IP broadcast packets on a LAN environment. Since IP broadcast packets cannot traverse layer 3 boundaries, the source MAC address of these packets will not be overwritten. Therefore, data obtained from other protocols can be used to create a mapping between a layer 2 (MAC) and layer 3 (IP) address.

4.2.2 Mapping the node to an IP subnet

The nodes are mapped to a single IP subnet based on their IP address. Some nodes may not operate on OSI layer 3, such as OSI layer 2 network devices. These devices will only be mapped to their associated Ethernet LAN.

4.2.3 Mapping the node to an Ethernet LAN

Since multiple IP subnets can coexist in the same OSI layer 2 LAN environment, the nodes and IP subnets are mapped to the Ethernet LAN they operate on. The nodes and IP subnets are mapped to their associated Ethernet LAN based on the source traffic capture. Depending on the gathered data, the Ethernet LAN may be mapped to a VLAN identifier. The passively obtained traffic capture contains information about a single Ethernet LAN environment, making these traffic captures useful to differentiate between Ethernet LAN environments.

4.3 Protocol effectiveness

This section compares the required information (as described in section 2.1 - Determining the information to gather) to the data passively obtained from the protocol samples. The data passively obtainable from an enterprise LAN environment is highly dependent on a number of different factors:

The usage of these protocols in the enterprise LAN environment - We researched six protocol samples during this research. The effectiveness of this technique to passively obtain information is (highly) dependent on the available protocols in the enterprise LAN environment. For instance, the information one could gather passively is very limited, if only one of the six protocols is used in the enterprise LAN environment.

The information contained in the traffic capture(s) - Some protocols advertise their information infrequently, upon client request. For example, suppose one passively captures network traffic for a short period of time while the nodes in the network are idle. In that case, the data obtained from that traffic capture may be very limited, while the protocols are used the enterprise LAN environment.

We gave each protocol a score, indicating the amount of data the protocol can passively obtain in a specific category. The data passively obtainable from the protocol is compared to the information required to perform a penetration test. This "required" information has been categorized in three categories: "Organization and procedures", "Security of the enterprise IT environment" and "Architecture of the enterprise IT environment". Appendix A (Information to gather) shows examples of information gathered in these categories. The list shown in appendix A (Information to gather) is not exhaustive.

The given scores are explained in table 4.13. The comparison in table 4.14 is based on a optimal scenario, whereas one would passively obtain all "interesting" traffic. The traffic capture would contain all infrequently advertised data. The scores are based on the percentage of information "items" (as shown in Appendix A - Information to gather) possible to obtain by the protocol in a specific category. The scores provide an indication of the usefulness of a certain protocol to passively obtain information in a specific category.

Table 4.14 shows the association between the scores and the protocols. It is interesting to note that the obtainable information highly differs per category. For example, the table shows that the protocol samples

Score	Description
○ ○ ○ ○ ○	Indicates that no information can be obtained in the specified category by the protocol.
● ○ ○ ○ ○	Indicates that the information obtainable from the protocol for the specified category is minimal. This indicates that between 0 and 20% of the items of information can be obtained using the protocol.
● ● ○ ○ ○	Indicates that one can obtain some information from the protocol for the specified category. This indicates that between 20 and 30% of the items of information can be obtained using the protocol.
● ● ● ○ ○	Indicates that one can obtain a medium amount of information from the protocol for the specified category. This indicates that between 30 and 40% of the items of information can be obtained using the protocol.
● ● ● ● ○	Indicates that the protocol can provide a large amount of information from the protocol for the specified category. This indicates that between 40 and 50% of the items of information can be obtained using the protocol.
● ● ● ● ●	Indicates that the protocol can provide directly usable information for the specified category. This indicates that over 50% of the items of information can be obtained using the protocol.

Table 4.13: Protocol scores

are not very effective in obtaining information about the "Security of the enterprise IT environment". On the other hand, the technique to passively obtain information for the reconnaissance phase of a penetration test is most effective in gathering information about the "Architecture of the enterprise IT environment".

Category	mDNS	SMB Browser	DHCP	NBNS	STP	CDP
Organization and procedures	● ● ○ ○ ○	● ● ● ○ ○	● ● ○ ○ ○	● ● ○ ○ ○	○ ○ ○ ○ ○	● ● ○ ○ ○
Security	● ○ ○ ○ ○	● ○ ○ ○ ○	● ○ ○ ○ ○	● ○ ○ ○ ○	● ○ ○ ○ ○	● ● ○ ○ ○
Architecture	● ● ● ○ ○	● ● ● ● ●	● ● ● ● ○	● ● ● ○ ○	● ● ○ ○ ○	● ● ● ● ○

Table 4.14: Comparison of passively obtainable information per protocol

4.4 Chapter summary

This chapter describes the analysis of the protocol samples. The first section of this chapter, (4.1 - Analysis of the selected protocols) describes the results of the analysis of the protocol samples selected to research. A sample of six protocols is analysed (mDNS, SMB Browser, DHCP, NBNS, STP and CDP) on functionality, details and usability for network profiling.

The second section (4.2 - Combining data) describes the combination of gathered data, to create information usable for the reconnaissance phase of a penetration test. The data gathered from the protocol samples is fragmented. To make the data usable, it needs to be combined into information. The first step is to map the gathered data to a single node. The second step is to map this node to an IP subnet. The third step is to map the node to an Ethernet LAN.

The third section (4.3 - Protocol effectiveness) compares the information gathered during the reconnaissance phase of a penetration test to the data passively obtainable using the protocol samples. The obtainable

information highly differs per category of information. The technique to passively obtain information using the protocol samples is not effective in gathering information about the "Security of the enterprise IT environment". On the other hand, the technique to passively obtain information for the reconnaissance phase of a penetration test is most effective in gathering information about the "Architecture of the enterprise IT environment".

Chapter 5

Proof of concept implementation

We created a proof of concept implementation of the researched technique to passively obtain information useful for the reconnaissance phase of a penetration test. This chapter describes the implementation and architectural decisions of the proof of concept. Whereas the previous chapter (4 - Protocol Analysis) describes the theoretical analysis of the protocols, this chapter describes the implementation of the researched technique.

Section 5.1 (Architecture) gives a description of the architectural decisions made to create the architecture of the proof of concept.

Section 5.2 (Performance evaluation) evaluates the performance of the proof of concept implementation. The section describes the effectiveness and the throughput of the proof of concept implementation.

5.1 Architecture

This section describes the requirements and the architectural decisions of the proof of concept. The requirements are used as a base for the architectural decisions of the proof of concept implementation.

5.1.1 Requirements

We determined the requirements of the proof of concept, before deciding upon the architecture of the proof of concept. The proof of concept has the following two (main) requirements:

Extendibility - The proof of concept should be easily extendible. For instance, it should be easy to add additional protocols to the proof of concept, to increase the effectiveness of this method to obtain information.

Flexibility - The proof of concept should be flexible in reporting. For instance, one should be able to easily create new reports, showing other aspects of the gathered information.

5.1.2 Gathering data from the protocols

The proof of concept is created in Python. Python allows the application to be simple and easily extendible, due to the large amount of open source Python libraries useful for this project.

The Scapy library is used to parse a traffic capture. The Scapy library provides a simple way to parse traffic captures and has a relatively large amount of supported protocols [19]. One could easily add new protocols to the proof of concept to increase its effectiveness, without the need to program a protocol parser (if the Scapy library has support for the protocol).

5.1.3 Database architecture

The information gathered from the protocols is stored in a database, to allow for simple extraction of information. As described earlier, the proof of concept focuses on flexibility. One can extract information specific to a certain test, without rewriting parts of the proof of concept. The proof of concept achieves this, by storing the gathered data in a database.

Appendix B (Proof of concept database schema) shows the database schema.

5.1.4 Relationships between the information

The proof of concept creates relationships between the data, using the method described in section 4.2 (Combining data). One can easily extract and create new correlations between the data, due to these relationships. For instance, one can combine information gathered by the SMB Browser and NBNS protocols easily, allowing the creation of new correlations between the data. These relationships add flexibility to the proof of concept, since one can easily add custom reporting functionality or extract specific information.

5.1.5 Reporting

The proof of concept creates an example report, showing the collected information in a Portable Document Format (PDF) document. The report acts as an example, focusing on displaying node based information per layer 3 subnet. Since the "interesting" information is stored in a SQLite database, one could easily add reporting functionality showing different aspects of the gathered information. For instance, one could add reporting functionality showing known attack vectors of specific systems, based on the passively gathered information.

5.2 Performance evaluation

This section evaluates the performance of the proof of concept. It describes the effectiveness and throughput of the proof of concept.

5.2.1 Effectiveness

The effectiveness of this technique to passively gather information usable for the reconnaissance phase of a penetration test is highly dependent on the tested LAN environment.

For instance, the proof of concept gathers significantly less data from a Linux or Apple OSx environment, in comparison a Microsoft Windows environment. Due to the selection criteria of the protocol samples, the

protocol samples consist of protocols mainly used in Microsoft Windows environments. The technique is highly dependent on the protocols used in the enterprise LAN environment, complicating the measurement of the efficiency of the proof of concept.

An evaluation of the effectiveness of the proof of concept requires:

- Traffic captures obtained from multiple, selected enterprise LAN environments
- Measurements taken from the selected enterprise LAN environments, using other reconnaissance methods

The effectiveness of the researched technique is highly dependent on the tested network environment. Therefore, multiple enterprise LAN environments must be tested to obtain information usable to measure the effectiveness of the proof of concept implementation. The gathered data must be compared to data gathered by other reconnaissance methods, or network documentation to test the effectiveness of the proof of concept.

Due to time and policy constraints, the effectiveness of the proof of concept implementation cannot be measured in multiple enterprise LAN environments. A demo network environment was created, to demonstrate the proof of concept implementation. However, measurements obtained from this demo environment do not provide proper data to determine the effectiveness of this proof of concept to perform network reconnaissance in enterprise LAN environments.

5.2.2 Throughput

We measured the throughput of the proof of concept implementation, using six different traffic captures. The measurements are obtained using a dual-core 1.66Ghz Intel Centrino Duo system (2GB RAM). The results are shown in table 5.1.

Traffic capture	Amount of packets	Duration of traffic capture (in seconds)	Time to analyse (in seconds)	Analysed packets per second
1	675	12476	3.2	201
2	4044	834	6.4	631
3	46193	12640	43	1074
4	664	728	1.8	368
5	5895	1617	8.4	701
6	1070	406	1.9	563

Table 5.1: Throughput proof of concept

The table shows the throughput of the proof of concept implementation varies between 200 and 1075 packets per second. The throughput is highly dependent on the total amount of packets in the traffic capture. Traffic captures with a larger number of packets (such as traffic capture 3, containing 46193 packets) show the highest throughput, reaching approximately 1075 packets per second. Smaller traffic captures (such as traffic capture 1, containing 675 packets) show a significantly lower throughput of around 200 packets per second.

5.3 Chapter summary

This chapter describes the proof of concept implementation of the researched technique. The chapter describes requirements and the architectural decisions made for the proof of concept. The main requirements of the proof of concept are extendibility and flexibility.

The proof of concept is created in Python. The application makes use of several open source libraries, such as the Scapy library to parse traffic captures. The proof of concept parses between 200 and 1075 packets per second, depending on the total number of packets contained in the traffic capture. Obtained data is stored in a SQLite database. The proof of concept creates relationships between the data, to allow for easy (and custom) reporting functionality. As an example of this reporting functionality, the proof of concept generates a PDF report, displaying (part of) the collected information.

Chapter 6

Conclusion

6.1 Passive LAN information gathering

The research has been formulated into the following research- and sub-questions:

Which information can be obtained by listening passively in a corporate LAN environment and how can this information be combined, correlated and reported to create an "outline" of the network, to simplify and prevent detection of the reconnaissance phase of a penetration test?

The following sub-questions combined will answer the main research question:

- *Which protocols are the most interesting to gather (and combine) information from, to create a map of a corporate LAN environment that can be used for a penetration test?*

We decided to limit the protocol samples to six, due to the time schedule of this research. Two selection criteria are used to select the protocol samples most interesting to research, namely: "The protocol must be commonly used and passively observable in enterprise LAN environments" and "The protocol must contain information needed to create the network profile". Six protocols showing the highest affinity with these two criteria are selected as protocol samples. The following protocols are selected: mDNS, SMB Browser, DHCP, NBNS, STP and CDP.

- *Which information can be passively gathered about the network segment and the systems in it by making use of the protocol samples?*

During the reconnaissance phase of a penetration test, information is obtained about the enterprise LAN environment. This information is required for the next phases of the penetration test. We divided the generally gathered information in three categories: "Organisation and procedures", "Security of the enterprise IT environment" and "Architecture of the enterprise IT environment".

The research shows that passive information gathering can be effective to obtain information about the architecture of the IT environment. Passively gathering information about the organisation and its procedures is possible, but one cannot obtain all required information in this category. The research also showed that the researched protocols hardly contain information about the security of the enterprise IT environment.

- *How can the gathered information be presented, to create an "outline" of the network that can be used for a penetration test?*

The data gathered from the different protocols is fragmented. The first step is to combine this data into information. The data from the different systems is combined into information from a single node. After that, the nodes are coupled to their associated OSI layer 3 IP subnet. The third step is to couple these nodes and subnets to a OSI layer 2 network. By combining the information, one has the flexibility to extract information of interest for a specific penetration test.

The proof of concept stores this information in a database, allowing for a easy method to extract specific information. The proof of concept also generates an example (PDF) report, containing information about the enterprise LAN environment.

6.2 Future work

Research other methods to passively obtain information from - Other methods to passively obtain information can be researched. This research focuses on the gathering of information directly from specific protocols, such as the SMB Browser and CDP protocols. One could research another method to obtain information, that is not directly based on specific protocols. For instance, one could parse all traffic and use a mathematical method to extract "interesting" information.

Research more protocols to passively obtain information from - In this research, six protocols have been analysed to passively obtain information from. One could research more protocols, in order to increase the effectiveness of this technique to obtain information. For instance, one could research the "optional" protocols, described in section 3.1 (Protocol samples). The DNS-SD standard looks promising to passively obtain information from, making it interesting for future research as well.

Extending the proof of concept As described in chapter 5 (Proof of concept implementation), a proof of concept implementation of the researched technique has been created. One could improve this implementation, by (for instance):

- By adding more protocols to passively obtain more information, thereby increasing the effectiveness of the implementation.
- By adding more reporting functionality. For instance, one may add a reporting functionality that searches for "known attacks", based on the passively gathered information.

Acronyms

BOOTP Bootstrap Protocol. 19

BPDU Bridge Protocol Data Unit. 23, 24

CDP Cisco Discovery Protocol. 25–27

CIFS Common Internet File System. 16

DHCP Dynamic Host Configuration Protocol. 19–21, 27

DNS Domain Naming System. 14, 15, 20, 27

DNS-SD DNS-based Service Discovery. 15, 16, 36

FQDN Fully Qualified Domain Name. 16

ID Identifier. 24, 25

IDS Intrusion Detection System. 5

IEEE Institute of Electrical and Electronics Engineers. 23

IETF Internet Engineering Task Force. 7, 16

IP Internet Protocol. 2, 7, 8, 12, 14, 15, 17, 19, 21, 25, 27–29

IPS Intrusion Prevention System. 5, 6

IT Information Technology. 5–8, 16, 18, 21, 23, 25, 27–30, 35

LAN Local Area Network. 1–3, 5–14, 16, 17, 19, 21–23, 25, 27, 28, 32, 33, 35, 36

LLDP Link Layer Discovery Protocol. 25

mDNS Multicast Domain Naming System. 14–16, 23, 27

NBNS NetBIOS Name Service. 21–23, 27

OS Operating System. 18, 19

OSI Open Systems Interconnection. 8, 14, 23, 27, 28, 36

PDF Portable Document Format. 32, 34

PVST Per VLAN Spanning Tree. 24, 25

RFC Request for Comments. 19, 21

RSTP Rapid Spanning Tree Protocol. 23

SMB Server Message Block. 16–18, 27

SQL Structured Query Language. 18

STP Spanning Tree Protocol. 23–25, 27

subnet subnetwork. 2, 20, 28, 29, 32

TLV Type Length Value. 25, 26

VLAN Virtual Local Area Network. 6, 24–28

VoIP Voice over Internet Protocol. 26, 27

VTP VLAN Trunking Protocol. 26

WINS Windows Internet Name Service. 21

Bibliography

[1] Egil Andresen. Conducting a security audit of an oracle database. http://www.electroban.net/~mbauden/docs/auditing/GSEC-Conducting_a_Security_Audit_of_an_Oracle_Database.pdf, 2002. [Online; accessed 24-June-2011].

[2] National Institute of Standards and Technology. Technical guide to information security testing and assessment. http://www.itsecure.hu/library/file/Biztons%C3%A1gi%20%C3%BAtmutat%C3%B3k/Egy%C3%A9b%20biztons%C3%A1gi%20%C3%BAtmutat%C3%B3k/Technical%20Guide%20to%20Information%20Security%20Testing%20and%20Assessment.pdf, 2008. [Online; accessed 24-June-2011].

[3] J. Postel. Internet protocol. www.ietf.org/rfc/rfc791.txt, 1981. [Online; accessed 27-June-2011].

[4] S. Deering. Host extensions for ip multicasting. www.ietf.org/rfc/rfc1112.txt, 1989. [Online; accessed 27-June-2011].

[5] J. Mogul. Broadcasting internet diagrams. http://tools.ietf.org/html/rfc919, 1984. [Online; accessed 27-June-2011].

[6] IEEE Computer Society. Media access control (mac) bridges. http://www.dcs.gla.ac.uk/~lewis/teaching/802.1D-2004.pdf, 2004. [Online; accessed 27-June-2011].

[7] Internet Engineering Task Force. Considerations for internet group management protocol (igmp) and multicast listener discovery (mld) snooping switches. http://tools.ietf.org/html/rfc4541, 2006. [Online; accessed 27-June-2011].

[8] Internet Engineering Task Force. Multicast DNS. http://tools.ietf.org/html/draft-cheshire-dnsext-multicastdns-14, 2011. [Online; accessed 22-June-2011].

[9] P. Mockapetris. Domain names - Implementation and specification. http://www.ietf.org/rfc/rfc1035.txt, 1987. [Online; accessed 22-June-2011].

[10] Internet Engineering Task Force. DNS-Based Service Discovery. http://tools.ietf.org/html/draft-cheshire-dnsext-dns-sd-04, 2011. [Online; accessed 22-June-2011].

[11] Apple Inc. Introduction to Bonjour Overview. http://developer.apple.com/library/mac/#documentation/Cocoa/Conceptual/NetServices/Introduction.html#//apple_ref/doc/uid/10000119i, 2010. [Online; accessed 22-June-2011].

[12] Microsoft MSDN. Microsoft smb protocol and cifs protocol. http://msdn.microsoft.com/en-us/library/aa365233%28VS.85%29.aspx, 2011. [Online; accessed 27-June-2011].

[13] Internet Engineering Task Force. Cifs/e browser protocol - preliminary draft. http://tools.ietf.org/html/draft-leach-cifs-browser-spec-00, 1997. [Online; accessed 22-June-2011].

[14] Internet Engineering Task Force. Dynamic host configuration protocol. http://tools.ietf.org/html/rfc2131, 1997. [Online; accessed 4-July-2011].

[15] Internet Engineering Task Force. Dhcp options and bootp vendor extensions. http://tools.ietf.org/html/rfc2132, 1997. [Online; accessed 4-July-2011].

[16] Internet Engineering Task Force. Protocol standard for a netbios service on a tcp/udp transport: Concepts and methods. http://www.ietf.org/rfc/rfc1001.txt, 1987. [Online; accessed 5-July-2011].

[17] IEEE Computer Society. 802.1d - ieee standard for local and metropolitan area networks. http://standards.ieee.org/getieee802/download/802.1D-2004.pdf, 2004. [Online; accessed 6-July-2011].

[18] Inc. Cisco Systems. Frame formats - cdp. http://www.cisco.com/univercd/cc/td/doc/product/lan/trsrb/frames.htm#xtocid12, 2002. [Online; accessed 5-July-2011].

[19] IEEE Computer Society. Scapy. http://www.secdev.org/projects/scapy/, 2011. [Online; accessed 6-July-2011].

Appendix A

Information to gather

A.1 Organisation and its procedures

A.1.1 Naming conventions

- Domain names
- System names
  - Servers/services
  - Network devices
  - End-user systems
- Usernames

A.1.2 Organisational structure

- Hierarchical structure
- Departments/workgroups
- Employees
  - Job information
  - Access levels

A.1.3 Physical locations

- Physical security levels
- Permitted access levels to internal resources

A.1.4 Partner organisations

- Network connections to these organisations
- Permitted access levels to internal resources

A.2 Security

A.2.1 Security plans and policies

- IT security policy
  - Password policy
  - Additional security policies
- Physical security policy

A.2.2 Technical security measures

- IDS/IPS systems
- Firewalls
- Logging devices
- Other security devices
- Security domain segmentations
  - Native/current VLAN
  - Other used VLANs

A.3 Architecture

A.3.1 Infrastructure

- OSI layer 2 network infrastructure
  - STP information
  - VLAN information
- OSI layer 3 network infrastructure
  - IP addressing information
- Upper layer network infrastructure
  - Workgroup/domain infrastructure
  - File shares
  - Proxy services
  - First hop redundancy services
  - Other network services

A.3.2 Hard- and software

- Naming information
- Platform information
- OS information
- Service information
- Role information

A.3.3 Important IT components

- Most accessed components
- Services offered by components
- Purpose of the components
- Components storing valuable information

Appendix B

Proof of concept database schema

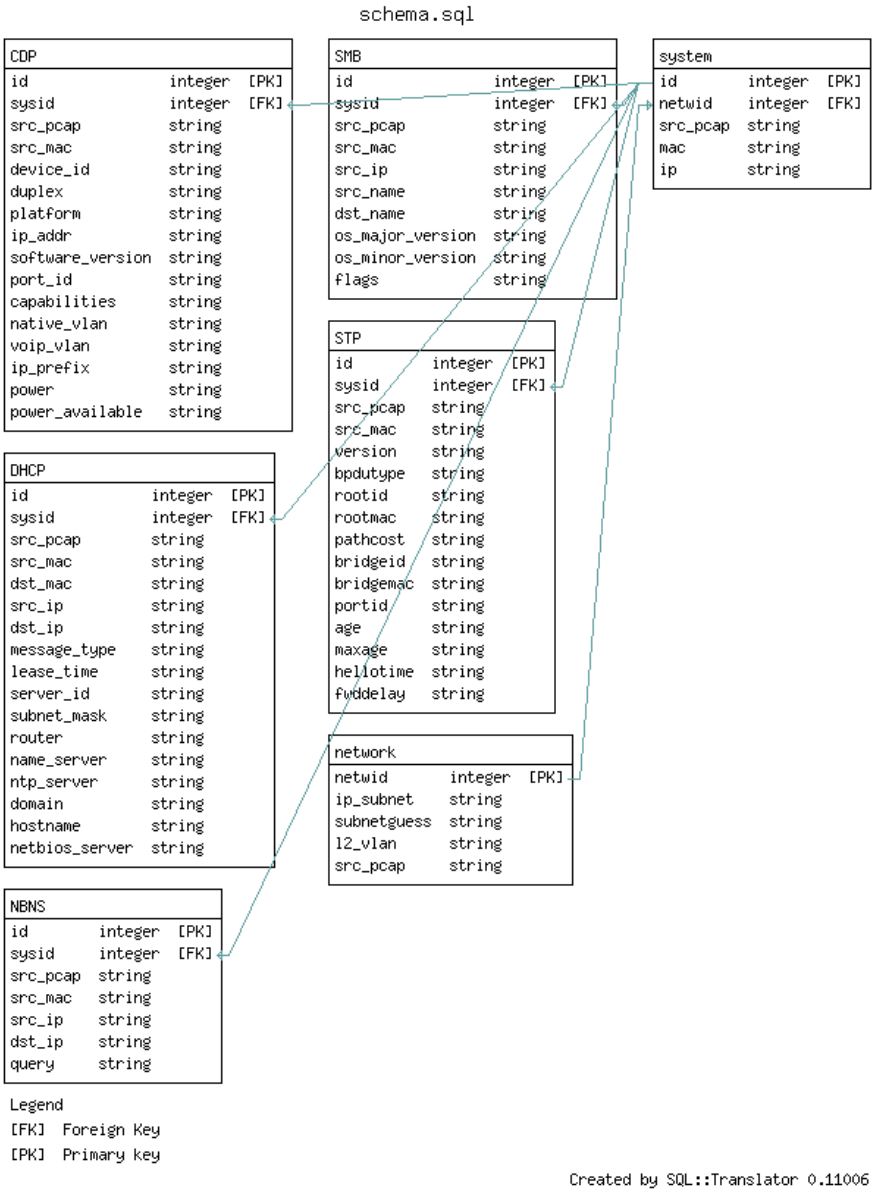

Figure B.1: Database schema