

DeepSweep™

CBIS Surveillance Module

User's Manual

December 2007

Copyright © IP Fabrics, Inc. 2007

IP Fabrics Corporate Headquarters
14964 NW Greenbrier Parkway
Beaverton, OR 97006
Telephone (main line): 503 444-2400
Telephone (FAX line): 503 444-2401
Website: <http://www.ipfabrics.com>

Information in this document is furnished in connection with IP Fabrics products. No license, express or implied, to any intellectual property rights is granted by this document. This document and the software described in it are furnished under license and may only be used or copied in accordance with the terms of the license.

Copyright © 2007, IP Fabrics, Inc. All rights reserved.

Packet Processing Language™, PPL™ and PPL-VM™ are owned and copyrighted by IP Fabrics, Inc.

Microsoft®, Windows® and Windows® XP are registered trademarks of Microsoft Corporation.

Linux® is a registered trademark of Linus Torvalds.

Red Hat® is a registered trademark of Red Hat, Inc.

RedBoot™ is a trademark of Red Hat, Inc.

MontaVista® is a registered trademark of MontaVista Software Inc.

Intel®, XScale® and Pentium® are registered trademarks of Intel Corporation.

Java™ is a trademark Sun Microsystems, Inc.

Table of Contents

1 Introduction.....4
 1.1 Implementation note.....4
 1.2 Overview4
 1.3 CBIS SM (Surveillance Module) Overview.....5
 2 Browser Pages7
 2.1 Case Information8
 2.1.1 CBIS SM-Wide Information.....10
 2.1.2 Changing an Active Case.....10
 2.2 Changes on Other Pages.....11
 2.3 SM Statistics.....11
 3 The Mediation Function Interface (MFI).....12
 3.1 Full (Content) Intercept Files13
 3.2 Limited (No-Content) Intercept Files14
 3.3 DHCP Processing and Files.....14
 3.3.1 DHCP State Tracking and Processing.....15
 3.3.2 IP Address Releases and Subsequent Traffic.....16
 3.4 Surveillance Status Reports.....16
 3.5 XXXXX and YYYYY17
 4 MF-BIF Communications.....18
 4.1 BIF Connection18
 4.2 File Copy via SFTP18
 4.3 Errors and Retry18

Table of Figures

Figure 1. DeepSweep as CBIS AF/MF.....5
 Figure 2. DeepSweep buffering (DSB) “Secure Buffered Delivery” as CBIS BIF5
 Figure 3. "CBI1" - CBIS SM definition screen.....7
 Figure 4. "CBI2" - New SubjectID definitions9
 Figure 5. CBIS MF-CF interface12

Table of Tables

Table 1. "SubjectID" type definitions8
 Table 2. CBIS file descriptions.....13
 Table 3. CBIS messages and related information15
 Table 4. DHCP tracking logic.....16
 Table 5. CBIS Surveillance Status Reports17

1 Introduction

This document describes a Surveillance Module for CALEA “broadband” use by the cable industry. Familiarity with the surveillance module for T1.IAS is helpful but not required to understand and use the CBIS surveillance module. Detailed information on how to create SMs, SM actions, and creating surveillance assemblies (SAs) can be found in the DeepSweep™ User's manual. This document assumes those concepts are understood by the reader.

1.1 Implementation note

This document may occasionally refer to items that are not supported in the current release of the product. These will be shown with a gray background. None of these restrictions are expected to limit the adherence of DeepSweep to the required standard but one should consider these aspects prior to initial deployment and work with IP Fabrics to ensure compliance.

1.2 Overview

This defines the architecture and external user interface of the of DeepSweep functions that provides support for lawfully authorized electronic surveillance of cable broadband access per the CableLabs Cable Broadband Intercept Specification CM-SP-CBI2.0-I01-070611 and the subsequent ECN CBI2.0-N-070517-1. This product draws heavily on existing IP Fabrics' products for ATIS “T1.IAS” and for the ATIS buffering specification.

As is the case for the “T1.IAS” buffering solution, there are several possible roles for DeepSweep products and technology in the network. One is that shown in Figure 1. Here DeepSweep is the access function and mediation function and does not rely on any intercept support in the CMTS or elsewhere; it serves as the complete solution up to the CBIS MFI interface. That is, the DeepSweep provides for the intercept administration and provisioning, filtering for all the intercepts, construction of the intercept information according to the CBIS specification, the first level of buffering per the CBIS specification, and proactive delivery of the intercept files to the BIF. The DeepSweep talks to the BIF as an SFTP client.

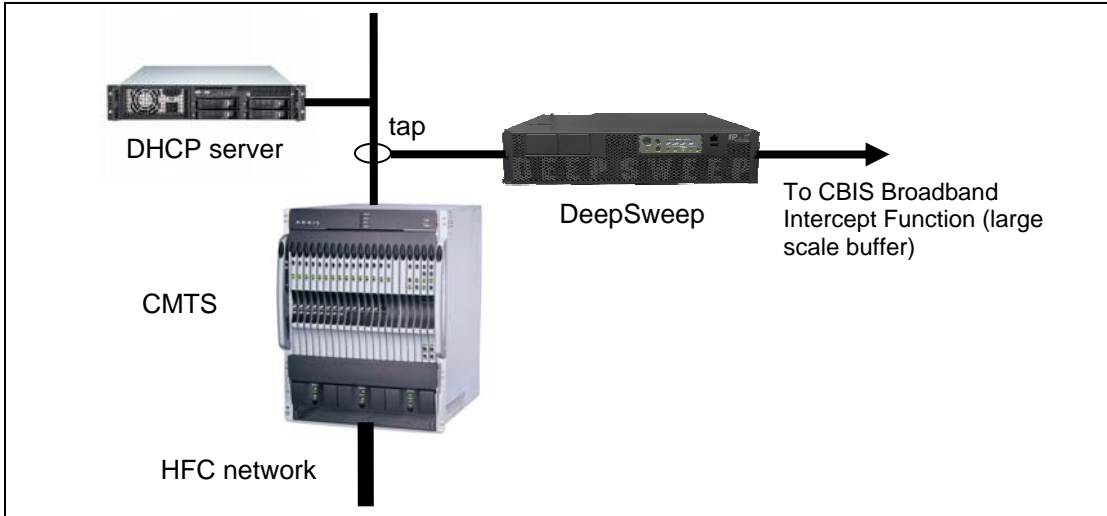


Figure 1. DeepSweep as CBIS AF/MF

Another possibility is use of the buffering-only DeepSweep (DSB) “Secured Buffered Delivery” product as only the BIF, as shown in Figure 2. Here the DSB receives intercept files from some MF (which could be the DeepSweep of Figure 1, or some other MF implementing the CBIS standard), and provides files to the LEAs. A DSB that implements both CBIS buffering and ATIS buffering is also available from IP Fabrics as a separate product offering.

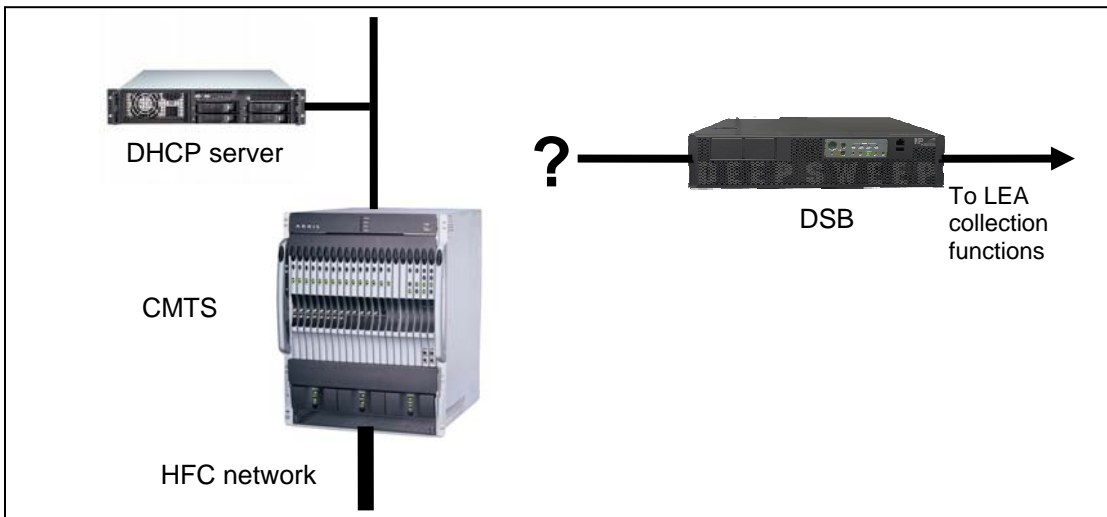


Figure 2. DeepSweep buffering (DSB) “Secure Buffered Delivery” as CBIS BIF

1.3 CBIS SM (Surveillance Module) Overview

For DeepSweep to function in the model of Figure 1,oe needs to have a CBIS surveillance module. Unlike the implementation for T1.IAS, CBIS has been implemented as a single SM (rather than the split used for T1.IAS).

The design allows multiple intercepts to be active simultaneously; multiple intercepts on behalf of different LEAs (law enforcement agencies), including multiple intercepts on the same subject (aka subscriber or target); and adding or deleting intercept cases without interrupting ongoing intercepts.

It is important first to understand the relationships among *cases*, *subjects*, *subject ids*, and *access sessions*.

Case	Typically a court order authorizing surveillance, typically of a single subject.
Subject	A term used loosely herein. Typically a person to which a case applies. The terms subject and subscriber are often used interchangeably. There can be more than one case that involves the same subject.
Subject ID	A specific network identification of a subject. A subject can be known by multiple IDs, and thus a case can typically define multiple subject IDs. In CBIS, where there is no authentication to the network, subject IDs are IP addresses and the various forms of DHCP identification of the cable modem or CPE. Because subjects can be in multiple cases, so can subject IDs.
Access session	The CBIS <i>access session</i> is <u>not</u> the same as the “access session” in T1.IAS; it is more like the “packet data session” in T1.IAS. View it as the set of packets related to a subject’s use of one IP address. This surveillance module equates access session to the possession of a specific IP address, and actually uses the IP address as the access-session ID.

2 Browser Pages

Configuration page CBI1 (Figure 3) shows the primary configuration parameters for the CBIS SM. There are three sections to the page:

- List of defined cases, in upper left.
- Definition of the selected case, in center. Here one can
 - Fill in information describing a new case
 - See the current definition about a defined case
 - Edit information about a defined case
- Definition and state of this CBIS SM. As for a case, one can use these fields to define attributes of the SM, see the attributes, and change the attributes.

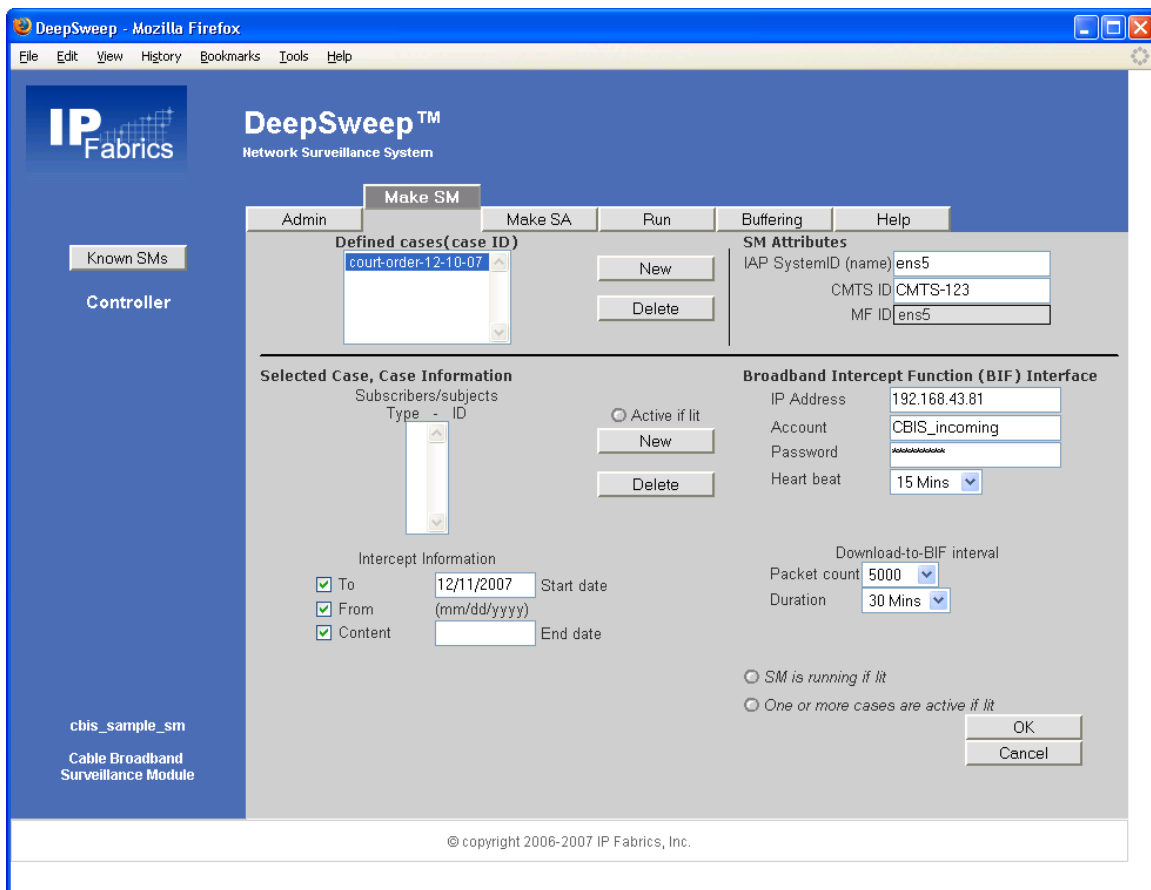


Figure 3. "CBI1" - CBIS SM definition screen

The upper left has a scroll box that shows the name(s) of all cases that have been defined to this surveillance module. A case is identified by a case ID, which is a 1-25 character string. Depressing the NEW button brings up a box that asks the user for a 1-25 character case ID. Providing that the case ID is different from all existing case ID's, depressing OK in that box returns to CBI1, where the user can then describe the case beneath. Depressing DELETE next to the case scroll box causes the entire case to be deleted. If the case is not

active, the system will prompt the user for confirmation. If the case is active, the system will prompt the user with stronger wording, because deleting a case while it is active is unusual and serious.¹

2.1 Case Information

The center of the screen shows the definition of the selected case, or in the case of a newly created case, is blank. An indicator shows whether the case is active, meaning that surveillance is active. By definition, if this indicator is lit, the two indicators on the lower right are also lit.

A scroll box shows the subject IDs that are part of the case. We provide for any number of subject IDs per case because a subject may be known to the network in multiple ways. Subject ID can be a number of things, and the scroll box shows the type that was chosen when the subject ID was entered. The following types are provided:

Type	Meaning	Notes and Examples
MAC	MAC address	12 hex digits. Hyphens or colons may be used as visual separators. E.g., 00-5B-78-A4-00-9E. This is found in field chaddr in the DHCP message. Depending on circumstances, this could be the CM MAC, a CPE MAC, or in CableHome a WAN-Man MAC or WAN-Data MAC. As for all of these, one can describe multiple ID's per subject.
CLIN	MAC address and DHCP option 61 client ID	Data entry boxes will appear for two items: <ul style="list-style-type: none"> MAC Address - 12 hex digits (hyphens or colons may be used as visual separators) This is matched to the chaddr field in the DHCP message. Client ID - N hex digits (where again hyphens or colons may be used as visual separators). This field is matched in an opaque way to the DHCP option 61 client ID. (No option 61 means no match).
SUBS	Subscriber ID	An ASCII string that is matched to DHCP option 82 suboption 6 (RFC 3993). No DHCP option 82 suboption 6 means no match.
RMCT	DHCP Option 82 remote ID and circuit ID	Data entry boxes appear for two items: <ul style="list-style-type: none"> Remote ID – "R" hex digits, and Circuit ID – "C" hex digits. Hyphens or colons may be used as visual separators. The R-part is compared opaquely with the option 82 remote ID and the C-part is compared opaquely with the option 82 circuit ID. Both fields must be non-blank. Use "O82R" or "O82C" to match only Remote ID or Circuit ID.
O82R	DHCP Option 82 remote ID	"R" hex digits compared opaquely with the option 82 remote ID. Hyphens or colons may be used as visual separators
O82C	DHCP Option 82 circuit ID	"C" compared opaquely with the option 82 circuit ID. Hyphens or colons may be used as visual separators.
IPV4	IPv4 address	Permanent IPv4 address. ² Must be standard dot notation, e.g., 68.100.1.1.

Table 1. "SubjectID" type definitions

Note that one can use as many subject IDs as wanted to define a specific subject.

Note that it is possible for a subject to be in possession of zero, one, or multiple IP addresses (by doing a variety of different DHCPDISCOVER's). In the situation of multiple IP addresses, each IP address must be

¹ An active case is one for which surveillance is currently underway. To be active, the case must have at least one subject ID, must have its BIF interface defined, and the current date must be greater than the start date (if the start date is not blank) and not greater than the end date (if the end date is not blank).

² Allowing for a static IP address is not mentioned in the CBIS spec, but it seems useful to have.

intercepted, and each different address is the access-session ID of a different access session. The DHCP processing for this is discussed in a later section.

Depressing the NEW button next to the subjects scroll box brings up a box that asks for the subject ID and its type. The type is selected from a set of choices in this box. Providing that the subject ID is different from all existing subject ID's of this type in this case and that its format matches the selected type, depressing OK in that box returns to CBI1. Adding a subject ID to an active case will cause that subject ID to be active immediately.

For ID types MAC, CLIN, SUBS, and RMCT (including O82R and O82C), the user fills in the identifier according to the rules in the table above. One can also optionally enter an IP address if it is known that an access session is already in progress by the subject at the start of the intercept. This IP address is treated as if it were leased in a DHCP transaction prior to the intercept. Also, if one wants (one of) the subject ID to be a permanent IP address, he can choose this option and enter the permanent IP address in the identifier field.

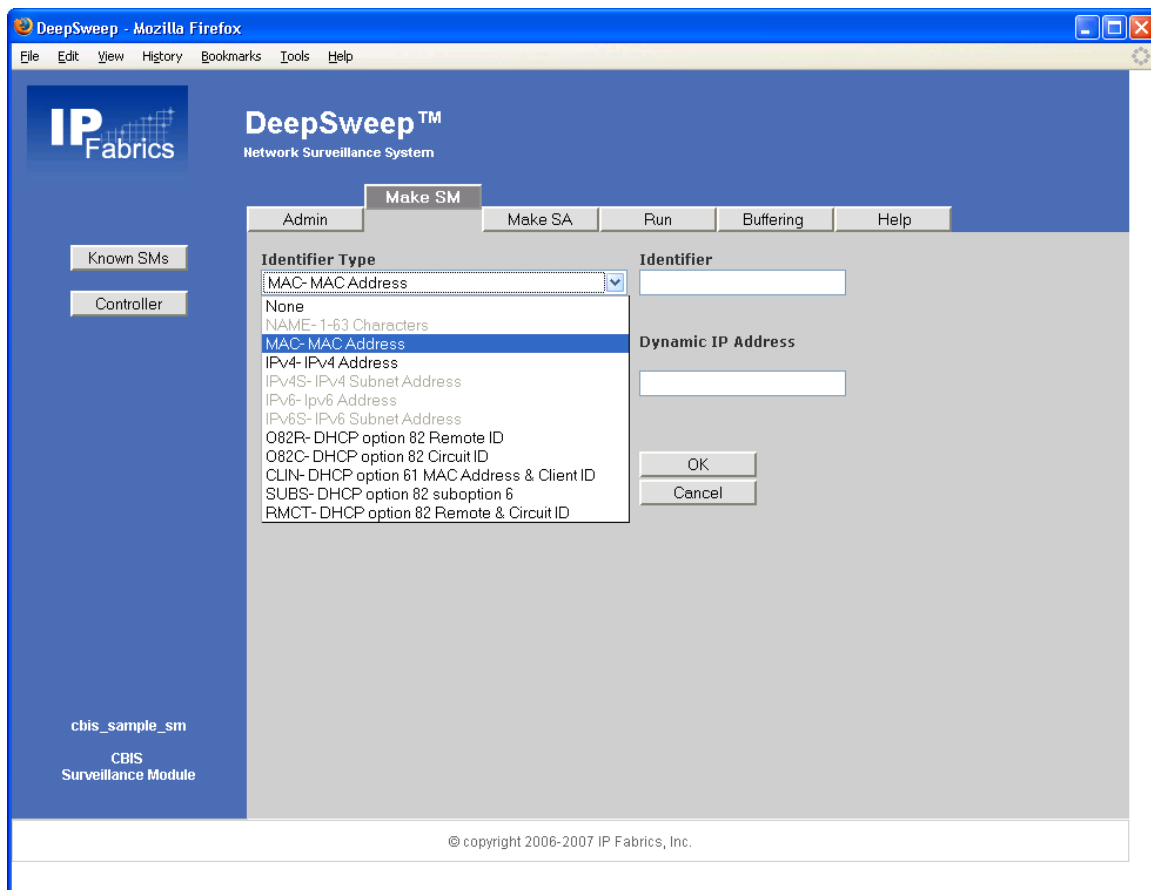


Figure 4. "CBI2" - New SubjectID definitions

Returning to page CBI1, depressing DELETE next to the subject scroll box deletes the subject. If the case is active, the system will prompt the user for confirmation. Deleting a subject ID from an active case causes intercept related to that specific subject ID to stop for this case.

The next three check boxes define the type of intercept authorized. At least one of TO and FROM needs to be checked. .

The next two fields are the dates, in the time zone³ of the DeepSweep system, on which intercept is to start and be completed. If start is left blank, it means “immediately.” If end is left blank, there is no automatic cessation of the intercept.

The last pieces of information for a specific case are information about the BIF system. The BIF need not be co-located because a reliable and secure communication mechanism is used. Defining the BIF as part of the case means that different cases can potentially go to different BIFs, thus providing for a configuration where there are multiple BIFs. In addition to needing the IP address of the BIF, the system also needs the name of the incoming directory within which the intercept files are transferred. This field is initialized with the value “CBIS_incoming.” This is the name used in the IP Fabrics DeepSweep BIF product.

The password is the password to allow SSH/SFTP access to the directory on the BIF.

CBIS defines a heart beat event; its timer is configurable here. The menu provides the following values: 30 sec, 1 min, 5 min, 15 min.

The next item is a specification of the interval between downloads of the intercept information to the BIF. It is expressed in terms of packet (capture) count and time; the first one to occur causes a download. It has the same meaning for both full content intercepts and non-content (limited) intercepts. The selections for packet count include 100, 500, 1000, 5000, 10,000, and 50,000. The selections for time include 1 min, 5 min, 10 min, 30 min, 1 hour, 4 hours, 8 hours.

2.1.1 CBIS SM-Wide Information

At the bottom right, note that this SM page is unusual in that it indicates if this SM is actually running as the page is viewed. Another indicator shows if any of the cases are currently active (intercept is active). The primary purpose of these are to help the user understand what adding or deleting an intercept will mean.

At the top right are also some fields representing additional information, independent of specific intercepts, that is communicated when hits occur. There are three names or IDs communicated in CBIS messages; two are defined here, and the third – the MF system identity – uses the DeepSweep system host name defined on page A2. This “MF ID” is displayed on this page for information only.

2.1.2 Changing an Active Case

It is important to understand what it means to change an active case. The situations are the following:

- The user deletes a subject ID. This was discussed earlier; intercept relative to this subject ID will stop.
- The user adds a subject ID. This was discussed this earlier also. The subject ID will become live immediately.
- The CBIS SM contains information about communication with the BIF that might warrant changing mid stream. The system allows the following to be changed (by changing the value and clicking OK):

³ To be precise, in the time zone as provisioned in the DeepSweep. That is, a DeepSweep could be physically in the Pacific time zone but provisioned to use UTC+0 time.

- Heart beat
 - Download-to-BIF intervals
- Changing of all other information will be disallowed for an active case – either the fields will be unchangeable or the OK click will report an error.

2.2 Changes on Other Pages

This section considers impacts or changes on aspects of the DeepSweep system user interface.

On the “SA Actions” page (“Make SA” tab - see “DeepSweep User’s Manual”) the only standard action allowed for the CBIS SM will be monitor. Monitor is very useful in determining that an installation or intercept is working properly, but it also represents a potential privacy exposure, so it should only be used under proper circumstances.

The “Statistics” page (“Run” tab) has a statistic for external messages sent. Unlike the IAS and VoIP SMs, the CBIS SM sends no messages, so nothing from the CBIS SM will be counted here.

It is unlikely that the CBIS SMs will be used in chains with other SMs but the system can handle such a circumstance. Consistent with the IAS and VoIP SMs, the CBIS SM always pass both hits and misses on to the next SM.

2.3 SM Statistics

In the DeepSweep architecture, every SM can collect up to four statistical values that are represented on the RUN/STATISTICS page in a 2x2 matrix. The lower right corner is always the number of packets examined by the SM. For the CBIS SM, the statistics should be

Pertinent DHCP packets processed	Packets captured
Downloads performed	Packets examined

3 The Mediation Function Interface (MFI)

The MFI is the interface from the MF to the BIF. It is shown below in a diagram from the CBIS specification.

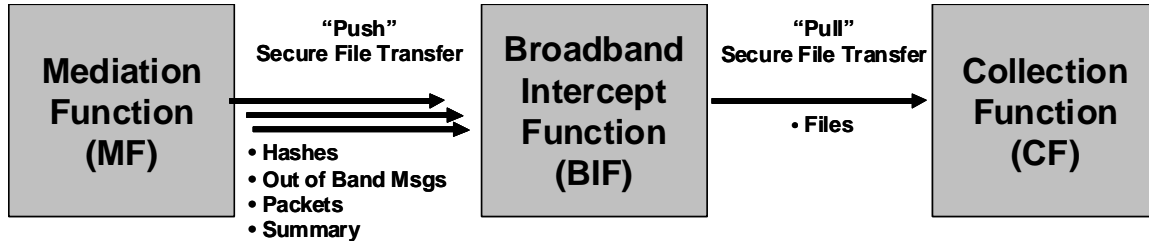


Figure 5. CBIS MF-CF interface

All communication between the MF and the BIF is initiated and driven by the MF, all using SFTP over SSH2. Thus all communication takes the form of the MF writing files to the BIF and the MF examining the file directory.

There are three file types used:

- PCAP. DHCP packets (only those relevant to an active case) are always captured in a PCAP file. If a content intercept is enabled, the subject's data packets are captured in a PCAP file.
- XML. There are three types of XML files:
 - If a content intercept is not enabled, a packet data summary report (similar to T1.IAS) is built in an XML file.
 - Each pertinent DHCP packet is also decoded into an XML file.
 - Surveillance status reports
- Hash. The full content PCAP capture file is always hashed and a digest put in an accompanying hash file. The packet data summary report is similarly hashed, with a digest put in an accompanying hash file. (Also the DHCP PCAP file is hashed, but its digest is placed in the accompanying XML file).
- Flag. This is a zero-length file used to signal to the BIF that the other files of same sequence number have been completely transferred to the BIF.

Almost all of these files are transient; that is, they contain just a single record, so they are created and disposed of in one motion. Only the content PCAP capture file or its alternative, the packet data summary report, have more than a fleeting lifetime.

Below is a table of information about all of the files. The xxxxx and yyyy parts of the file names are sequential decimal numbers starting at 1 and with no leading zeros. When a new file needs to be created, except for a hash file, the counter value is used and then incremented.

The full... and limited... files are mutually exclusive for a specific case, so a specific case will have two groups of files: full or limited, and oob. However, keep in mind that there might be multiple cases tracking the same subject.

Description	Form of name ⁴	When created	When downloaded
Content PCAP capture	full/xxxx.dmp	On a content capture, if this file doesn't exist	When packet count or time interval expires (or surveillance period ends)
Content PCAP hash	full/xxxx.hash	Right before a content PCAP capture file is closed for downloading	With content PCAP capture file
Packet data summary	limited/xxxx.xml	As for content PCAP capture, but when content intercept not enabled	When packet count or time interval expires (or surveillance period ends)
Packet data summary hash	limited/xxxx.hash	Right before a packet data summary file is closed for downloading	With packet data summary file
DHCP PCAP	oob/yyyy.dmp	Discovery of a pertinent DHCP packet	Immediately (always one record)
DHCP event	oob/AccessAttempt-yyyyy.xml oob/AccessAccepted-yyyyy.xml oob/AccessFailed-yyyyy.xml oob/AccessSessionEnd-yyyyy.xml	Discovery of a pertinent DHCP packet, or of an access-session-end for other reasons	Immediately (always one record)
Surveillance status report	oob/SurveillanceStatusReport-yyyyy.xml	Depends on the status type	Immediately (always one record)
Flag	full/xxxx.flag limited/xxxx.flag oob/xxxx.flag		When the file or file pair in the same subdirectory and same sequence number has been downloaded and verified

Table 2. CBIS file descriptions

3.1 Full (Content) Intercept Files

When content intercept is enabled, intercepted packets are captured in an xxxx.dmp PCAP file. The file starts with a PCAP header record defining the standard magic number and version number, the time zone offset, and other standard stuff (sigfigs=0, snaplen=65535, network = Ethernet). Each captured packet (in the DeepSweep case, Ethernet frame) is placed in the file, along with a timestamp denoting the time of capture.

If the file does not yet exist, it is created and counter xxxx is incremented.

There is just one such file open per case. The file will be named caseidentity/full/xxxx.dmp when transferred to the BIF.

The DeepSweep keeps track of the number of packets written to each open file and the duration since the file was actually created (which also is the time the first packet was recorded in it). If the packet count or duration specified (page CBI1) is hit, or if “end of case” occurs,⁵ the following happens:⁶

- The file is closed. (If a subsequent captured packet arrives, it causes a new file to be created.)
- A SHA-256 hash is computed of the file, and this 32-byte value is written into new file xxxx.hash, where xxxx is the same as the closed file.
- Both files are copied to the BIF and verified (see section 4 “MF-BIF Communications” for details on this)
- An empty file xxxx.flag is written into the same subdirectory on the BIF.

⁴ All are prefixed with *caseidentity/*

⁵ The term “end of case” is used to describe various ways of reaching the termination of an intercept case, such as the case’s authorized time limit occurring or an administrator explicitly deleting a case.

⁶ Packet count or duration expiration can never occur if there is no current (open) non-empty xxxx.dmp file, but it could occur on the “end of case” situation. The system never attempts to transfer an empty file (CBIS requirement R-310).

- The values caseidentity, xxxxx, timestamp, and the 32-byte digest are written to the system log. This is requirement R-70 in CBIS, which requires that this be maintained by the MSO as a business record.
- The .dmp and .hash files in the DeepSweep are deleted.

3.2 Limited (No-Content) Intercept Files

Alternatively, if content intercept is not enabled for a specific case, the following is done. The core of this is called the packet data summary report in CBIS.

A five-tuple of information is extracted from each intercepted packets and used to build a table. The five pieces of information are source IP address, destination IP address, protocol (from the IP header), source transport port, and destination transport port.⁷ If the transport protocol is other than UDP, TCP, or SCTP, it is treated it as having null-value ports. If the tuple is already in the table, the system increments a count associated with it; otherwise the the tuple is inserted in the table and set its count to 1.

If the packet counter has reached the packet count⁸ or has reached the time duration (both specified on page CBI1 for the case) or “end of case” arises, the system creates an xxxxx.xml file. This document will not specify the detailed XML format here, because it is specified along with an example in the CBIS document. Basically the file contains XML elements containing case identify, IAP system ID, timestamp (time that the system determined it necessary to create the file), and access session ID, and then a sequence of packet signatures (five-tuples and counts).

Similar to the case above, once the XML file is complete, a SHA-256 hash is computed of it and the value written to a new file xxxxx.hash. Both files are copied to the BIF (see section 4 “MF-BIF Communications”). The flag file is written to the BIF. A write occurs to the system log as in the above section, and the .xml and .hash files on the DeepSweep are deleted.

3.3 DHCP Processing and Files

The main DHCP processing is watching all DHCP messages for a match on any of the specified subject IDs. Any matching DHCP message is one that is to be captured. The actions taken with this DHCP message are

Creation and SFTP transmission of an oob/yyyy.dmp file containing exactly one PCAP record, namely the DHCP packet.⁹

- Creation and SFTP transmission of an oob/Accessmmmm-yyyy.xml file, which essentially is a formatted description of the DHCP event.
- Conditionally, some internal state changes

The creation of the PCAP file is straightforward; it follows the process described earlier, except that it always contains exactly one packet. A hash is computed of the file, but the hash is not placed in a .hash file; the hash value will be used within the xml file.

⁷ If the protocol is one that doesn't have transport ports, it is treated this as having null-valued ports.

⁸ CBIS doesn't include a packet count limit for the packet data summary report, just a “summary timer.” However, for uniformity, DeepSweep includes both. Having an extra limit available doesn't violate the standard.

⁹ Note the use of sequence number yyyy for this type of file.

The DHCP message is also decoded and as a result generates an XML file. The XML file contains an access message. Standard things in every access message are

- Case identity
- MF system identity
- Timestamp
- Subject ID
- Name of the PCAP file (containing the DHCP packet)
- 32-byte SHA-256 hash value of the PCAP file

The specific messages and message-dependent information are shown below.

DHCP message	Resultant XML file	Message/file specific parameters
Discover Offer Request Inform	oob/AccessAttempt-yyyyy.xml	CMTS ID from page CBI1. Access device ID would appear to be always the chaddr field
Ack	oob/AccessAccepted-yyyyy.xml	Above two things, plus IP address, lease duration, access session ID. However, for a DHCPACK with no IP address, no message is created.
Release	oob/AccessSessionEnd-yyyyy.xml	IP address and access session ID.
Nak Decline	oob/AccessFailed-yyyyy.xml	IP address if available, failure reason = "DHCPNAK" or :DHCPDECLINE"

Table 3. CBIS messages and related information

As in the previous sections, once the DeepSweep copies both files to the BIF, and creates, on the BIF, a yyyy.flag file in the same subdirectory.

3.3.1 DHCP State Tracking and Processing

The important internal processing with DHCP messages is determining which IP addresses should be used for filtering. The system gets an IP address to consider from any of the following situations:

- A DHCPACK (access accepted event)
- A starting temporary IP address assigned as a subject ID
- A permanent IP address assigned as a subject ID

The logic needs to consider three questions:

1. Is the new IP address incremental to those already associated with the subject?
2. Or does the new IP address replace one already associated with the subject?
3. When is an IP address no longer in use by a subject?

To find new assignments, the system primarily watches the DCHPACK, but watching it alone (i.e., stateless) is insufficient for two reasons: (1) it can't distinguish a lease renewal from a new assignment and (2) all of the subject IDs do not appear in the DHCPACK (specifically option 61 and 82 parameters don't appear back in the DHCPACK).

The overall logic is summarized in the table below.

An ID matches something in a DHCPDISCOVER and eventually there is an associated DHCPACK.	Associate the IP address from the DHCPACK with the ID
An ID matches something in a DHCPREQUEST, there is no associated DHCPDISCOVER, and eventually there is an associated DHCPACK with an IP address matching that currently associated with the ID.	No state change.
An ID matches something in a DHCPREQUEST, there is no associated DHCPDISCOVER, eventually there is an associated DHCPACK with an IP address that <u>doesn't</u> match that currently associated with the ID.	Disassociate the current IP address with its ID (i.e., discard the IP address) and associate the new one with the ID
A DHCPRELEASE has as its IP source address an IP address being monitored.	Disassociate the IP address with its ID
One of the IP addresses being monitored (i.e., is associated with a user ID) appears in a DHCPACK, and this DHCPACK is not associated with a DHCPDISCOVER or DHCPREQUEST that had an ID being monitored.	Disassociate the IP address with its ID

Table 4. DHCP tracking logic

3.3.2 IP Address Releases and Subsequent Traffic

There are two situations of IP address releases – those that are explicit (e.g., a DHCPRELEASE message) and those that are implicit (e.g., a tracked IP address being assigned to other than its associated ID). The DeepSweep design objective is to do this “promptly.” Promptly means before the address could be feasibly reallocated by a DHCP server, and the router(s) discover the new IP address location.

3.4 Surveillance Status Reports

The last message and file type is surveillance status. The file is named oob/SurveillanceStatusReport-
yyyy.xml. This is an instance where the system creates and sends a single file instead of a pair of files¹⁰. There is no hash involved. A flag file is used so that the BIF is not fooled by a created file that is about to be written to.

There are five types of reports. In addition to the report number, the XML file contains the case id, MF system id, time stamp, and access session ID(s)¹¹. The table below shows the circumstances under which the system creates each report. Once created, the normal process of file transfer and deletion as described in the previous sections occurs.

Report	When created
Up (surveillance activated)	One per active case when system is booted For newly active case (because of creation during operation or because start date arrives)
Down (surveillance deactivated)	One per active case when the Surveillance Assembly is terminated Active case is deleted Active case hits its end date
Error	Currently know of no circumstances where this would be emitted.
Unknown	CBIS says this denotes “indeterminate status.” DeepSweep will not create this report.
Heartbeat	One per active case every heartbeat duration (defined on page CB11)

¹⁰ There is one other instance – certain events leading to an access-session-end.

¹¹ Note that where the CBIS spec indicates there is one access session ID in the report, a recent decision changed this to zero to N ID’s per report. I.e., each time a report is sent, it will list the current access sessions for the case, or indicate that there is none.

Table 5. CBIS Surveillance Status Reports

3.5 XXXXX and YYYYY

For xxxxx, each case has a counter starting at 1. When creating a file needing xxxxx in its name (packet capture PCAP file, packet data summary PCAP file), except for .hash files, DeepSweep uses the counter value and then increments it. There are no leading zeros, and the counter has no size limit.

For yyyyy, each case has a counter starting at 1. When creating a file needing yyyyy in its name (all files in the oob subdirectory), DeepSweep uses the counter value and then increments it. No leading zeros and no size limit.

Thus on the receiving end (BIF and CF), one should expect to see two sequences of numbered files with no gaps. A gap would denote a lost file.

As noted earlier, this system uses yyyyy instead of xxxxx in the DHCP PCAP file. This simplifies life, has the benefit of easy correlation of the DHCP PCAP and XML files, and avoids ambiguity in the numbering of flag files.

4 MF-BIF Communications

4.1 BIF Connection

The CBIS document doesn't specify the SSH2 login (to the BIF(s)) and the duration of the login. DeepSweep will take the approach of doing an SSH2 login in following circumstances:

- When a case becomes active.¹² This occurs in the following:
 - When a case becomes active during system operation (dynamically created or start time arrives)
 - When the surveillance system is started (technically when an SA containing a CBIS SM is started)
- When a file needs to be transferred and the system finds that it is not currently logged into the BIF
- Retries after a delay when one of the above fails

If DeepSweep cannot connect to the BIF, it generates an inability-to-connect alert (causing an SNMP trap if enabled) and puts a descriptive entry in the DeepSweep system log (/ftp/systemlog that is available to the "ens_administrator" administrative account). After a delay (suggested to be about 1 minute), the system will retry connecting.

4.2 File Copy via SFTP

The BIF is responsible for creating the incoming directory and its subdirectories, so if the incoming directory is the working directory and the caseid is "case123," then one should expect to see case123/full, case123/limited, and case123/oob as the subdirectories.

In situations where DeepSweep transfers a pair of files (.dmp/.hash, .xml/.hash, .dmp/.xml), after the transfer the system reads the directory to verify the presence of the two files and that their sizes are correct. If there are no problems, the .flag file is generated.

In situations where DeepSweep transfers one file (just the XML file for the surveillance status report), after the transfer it reads the directory to verify the presence of the file and that its size is correct. If there are no problems, the .flag file is generated.

4.3 Errors and Retry

CBIS defines, but does not require, a method for testing for successful SFTP transfer and retrying in the event of an error.

For every file transfer via SFTP, there are attempts to detect a transfer error in the following circumstances:

- The system receives an SFTP error code of 300 or greater

¹² Note that this is done case by case because the DeepSweep design allows each case to be associated with a specific BIF.

- A read of the directory after the transfer shows that the transferred file name is not present
- A read of the directory after the transfer shows a file size different than that of the file on the sending system

In the second and third case, the system retries a second file transmission, adding “.retry” to the end of the file name (e.g., “5.dmp.retry”). If on this it receives an error code of 300 or greater, or if the read of the directory again shows a problem (missing file or wrong size), it discards the local file in the MF) and generates a system error. Otherwise the retry has succeeded and the local file is discarded and processing proceeds as if no error occurred.

If the original transfer resulted in an error code of 300 or greater, a system error is generated.

In the case of a system error, the system puts a descriptive entry in the DeepSweep “systemlog,” generates an unrecoverable-error alert (causing an SNMP trap if enabled), and then proceeds as if there were no error. In both the log entry and the alert, a text string will identify the problem and the name of the file that could not be transferred.