



MAILSNIFFER – ТЕХНИЧЕСКОЕ ОПИСАНИЕ

1. ВВЕДЕНИЕ	2
1.1. НАЗНАЧЕНИЕ	2
1.2. ОПРЕДЕЛЕНИЯ	2
2. ПРИНЦИПЫ РАБОТЫ СИСТЕМЫ	3
2.1. КОМПОНЕНТЫ СИСТЕМЫ	3
2.2. ПРИНЦИПЫ РАБОТЫ	4
2.3. УСТАНОВКА СЕРВИСА ПЕРЕХВАТА В ЛОКАЛЬНОЙ СЕТИ	5
3. УСТАНОВКА СИСТЕМЫ	11
3.1. УСТАНАВЛИВАЕМЫЕ МОДУЛИ NETWORKSNIFFER	12
4. ТЕХНИЧЕСКИЕ ХАРАКТЕРИСТИКИ	13
4.1. ИСПОЛЬЗОВАНИЕ ИНДЕКСОВ	13
4.2. МАСШТАБИРУЕМОСТЬ СИСТЕМЫ	14
5. СИСТЕМНЫЕ ТРЕБОВАНИЯ	14
5.1. РЕКОМЕНДУЕМЫЕ АППАРАТНЫЕ ТРЕБОВАНИЯ	14
5.2. ПОДДЕРЖИВАЕМЫЕ ОПЕРАЦИОННЫЕ СИСТЕМЫ	15
5.3. ПОДДЕРЖИВАЕМЫЕ БАЗЫ ДАННЫХ	15
6. КЛИЕНТСКАЯ ЧАСТЬ MAILSNIFFER	16
6.1. ПОИСК	16
6.2. ПРОСМОТР РЕЗУЛЬТАТОВ	17
7. ИНТЕГРАЦИЯ С ALERTCENTER	18



1. Введение

1.1. Назначение

Программный комплекс SearchInform MailSniffer (MailSniffer) предназначен для перехвата почтового трафика на уровне сетевых протоколов, индексирования перехваченных сообщений и осуществления поиска по ним. Вне зависимости от используемого почтового клиента, MailSniffer обрабатывает все сообщения, принятые или отправленные с компьютеров корпоративной сети по поддерживаемым протоколам.

1.2. Определения

- Система - комплекс программных средств, включающий в себя MailSniffer, развернутый в локальной сети организации или предприятия.
- Сервер - серверное приложение. Обеспечивает перехват почтового трафика, сохранение сообщений в базе данных и индексирование.
- Клиент - клиентское приложение. Используется для поиска документов по индексу на сервере и просмотра результатов поискового запроса.
- База данных (хранилище) – набор перехваченных документов, структурированный при помощи поддерживаемых систем управления (СУБД): Microsoft SQL Server, PostgreSQL, SQLite.
- Индекс - структура, предназначенная для хранения и быстрого поиска по базе документов. Индекс обеспечивает быстрый доступ к содержащимся в нем данным, поэтому поисковые решения на базе SearchInform отличаются высокой скоростью и стабильностью работы.
- Текстовый поиск - поиск по ключевым словам и выражениям.
- Фразовый поиск - частный случай текстового поиска, при котором фиксируется порядок следования слов и выражений запроса.
- Поиск похожих – запатентованный алгоритм, осуществляющий поиск документов, похожих по содержанию. При поиске похожих, запрашиваемый текст сопоставляется каждому документу индекса. Документы похожие по содержанию на запрос, фильтруются по степени сходства или "релевантности".



2. Принципы работы системы

2.1. Компоненты системы

Работа MailSniffer основана на архитектуре клиент-сервер.

Реализация клиент-серверной архитектуры обеспечивает:

- Централизованное хранение и целостность информации на сервере СУБД
- Возможность поиска по всей базе документов, проиндексированной при помощи сервера
- Доступ к серверу при помощи клиентского приложения
- Доступ к базе данных в удаленном режиме
- Возможность масштабирования системы.

Компоненты MailSniffer:

- Сервер **NetworkSniffer**
- **База данных** поддерживаемого типа: **Microsoft SQL Server, PostgreSQL, SQLite**
- **SearchInform Server**
- Клиент **MailSniffer**
- **AlertCenter**

Сервер NetworkSniffer - предназначен для перехвата почтового трафика по поддерживаемым протоколам (POP3, SMTP, IMAP, HTTP) и сохранения почтовых сообщений в базе данных.

Сервис перехвата нужно установить на компьютер, на который зеркалируется трафик с сетевого коммутатора (свитча) или сетевого концентратора (хаба). При использовании программного маршрутизатора, можно попытаться обойтись без зеркалирования трафика и установить сервер перехвата непосредственно на программный сетевой шлюз.

Сервис перехвата имеет модульную структуру. Возможна отдельная установка на несколько рабочих станций для перехвата по отдельным протоколам – POP3, SMTP, IMAP, HTTP (перехват сообщений, отправляемых с поддерживаемых сервисов веб-почты). Целесообразность отдельной установки сервиса следует рассматривать для сетей с 300 или более рабочих мест.

База данных - предназначена для хранения документов, перехваченных при помощи NetworkSniffer.

В качестве хранилища документов могут использоваться следующие СУБД:

- Microsoft SQL Server
- PostgreSQL
- SQLite



База данных SQLite может быть установлена только локально на рабочей станции с установленным сервисом перехвата. Эксплуатация баз данных SQLite не рекомендуется при числе рабочих станций от 200 и выше.

SearchInform Server - система, обеспечивающая возможность быстрого анализа перехваченных данных.

SearchInform Server индексирует базы почтовых сообщений. Благодаря использованию индексов, многократно ускоряется доступ к данным. Использование индексов и уникальных запатентованных поисковых алгоритмов обеспечивает высокую скорость и точность поиска.

Клиент **MailSniffer** - предназначен для формирования запроса к индексу, получения результатов поиска и просмотра найденных документов.

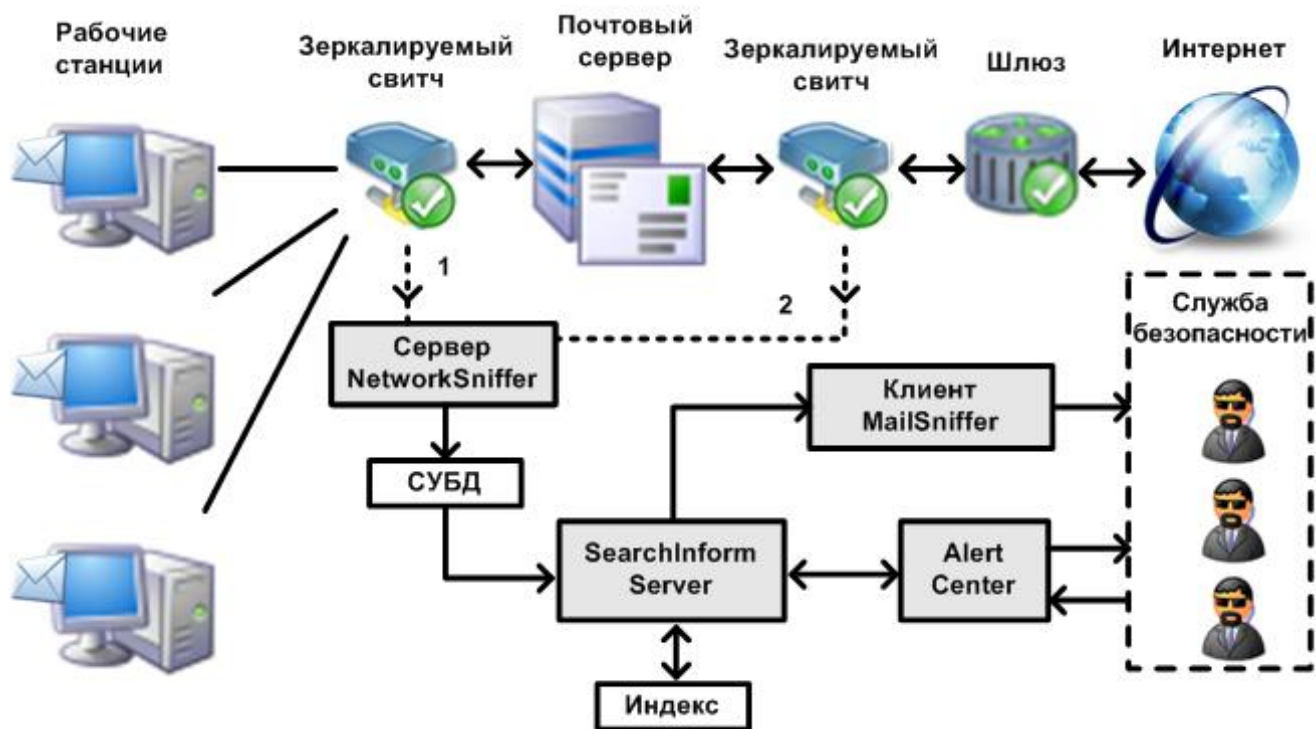
Эффективность проверок и точность идентификации обеспечиваются благодаря использованию трех типов поиска при формировании запросов: текстового поиска, фразового поиска и поиска похожих; а также целому ряду дополнительных функций и настроек, доступных в клиентской части.

AlertCenter - необходим для автоматизации процесса проверки: настройки расписания проверок и списка запросов, и для отсылки уведомлений в случае выявления подозрительных документов.

По заданному расписанию, AlertCenter может производить поиск по списку ключевых слов, выражений и текстовых фрагментов и автоматически отправляет администратору системы уведомления в случае обнаружения подозрительных совпадений.

2.2. Принципы работы

- Сервис перехвата помещает все сообщения, отсылаемые или получаемые по поддерживаемым протоколам в базу данных. Возможен импорт документов из баз Microsoft Exchange Server, почтовых клиентов и сохранённых на диске файлов MSG и EML.
- SearchInform Server подключается к СУБД, кэширует и индексирует данные.
- Индекс опрашивается по списку контрольных слов, выражений и текстовых фрагментов. Для настройки расписания проверок и редактирования списка запросов используется AlertCenter.
- В случае обнаружения совпадений, AlertCenter отправляет уведомление сотруднику службы безопасности.
- По ссылке в уведомлении открывается клиентское приложение MailSniffer, где можно просмотреть подозрительные письма и провести дополнительные проверки и анализ, исходя из ситуации.



2.3. Установка сервиса перехвата в локальной сети

Сетевой шлюз или отдельная рабочая станция?

Если планируется отслеживать почту с нескольких компьютеров локальной сети, то необходимо установить систему перехвата в том узле системы, через который проходит необходимый для перехвата входящий/исходящий трафик. Другими словами, необходимо выбрать «точку наблюдения». Существует множество конфигураций сети, в которые можно успешно встроить систему перехвата почтового трафика.

Не рекомендуется установка на сервер домена или файловый сервер.

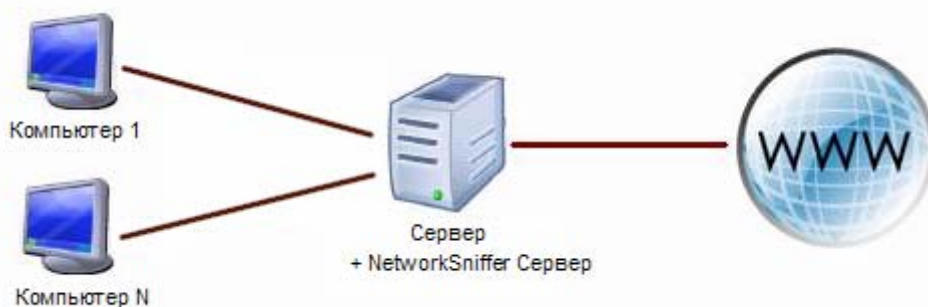
- **Работа на сетевом шлюзе**

Шлюз - компьютер, через который осуществляется обмен информацией между локальной сетью и Интернетом. Шлюзом может выступать аппаратный маршрутизатор или программный маршрутизатор.

- На аппаратный маршрутизатор невозможно установить дополнительное ПО для

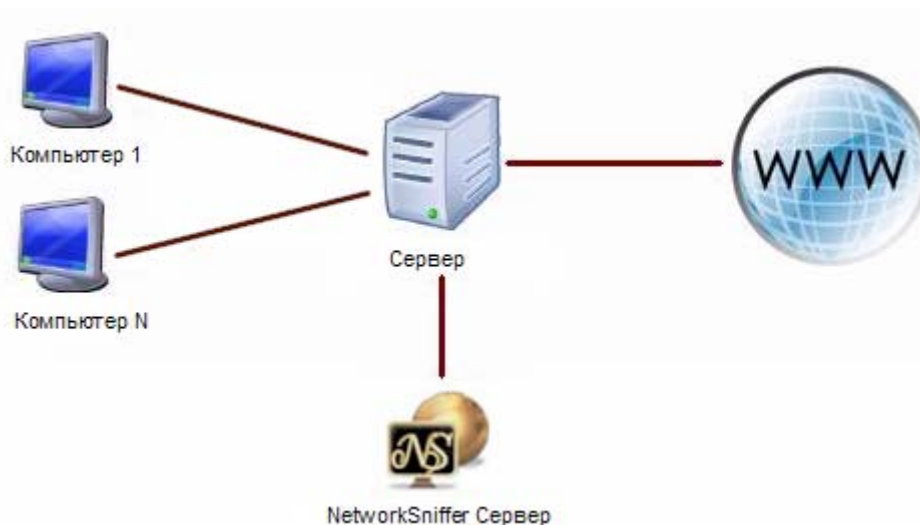
перехвата трафика. Поэтому трафик необходимо зеркалировать на отдельный сервер NetworkSniffer.

- Если используется программный маршрутизатор, трафик можно и перехватывать на этом же компьютере, и зеркалировать на отдельный сервер NetworkSniffer. При большом числе рабочих станций в сети, трафик рекомендуется зеркалировать.



• Работа на отдельной рабочей станции

Рекомендуемый вариант установки NetworkSniffer - установка дополнительной рабочей станции, которая не будет выполнять других задач, кроме перехвата почтового трафика и обеспечения поиска по базе перехваченных писем. Для того чтобы отдельная рабочая станция получила возможность перехвата почтового трафика, на эту станцию необходимо продублировать весь почтовый трафик локальной сети. В любом случае, станция перехвата подключается к локальной сети через какое-либо устройство, способное дублировать весь трафик на эту станцию.



Выбор точки установки (использование шлюза, хаба, свитча)

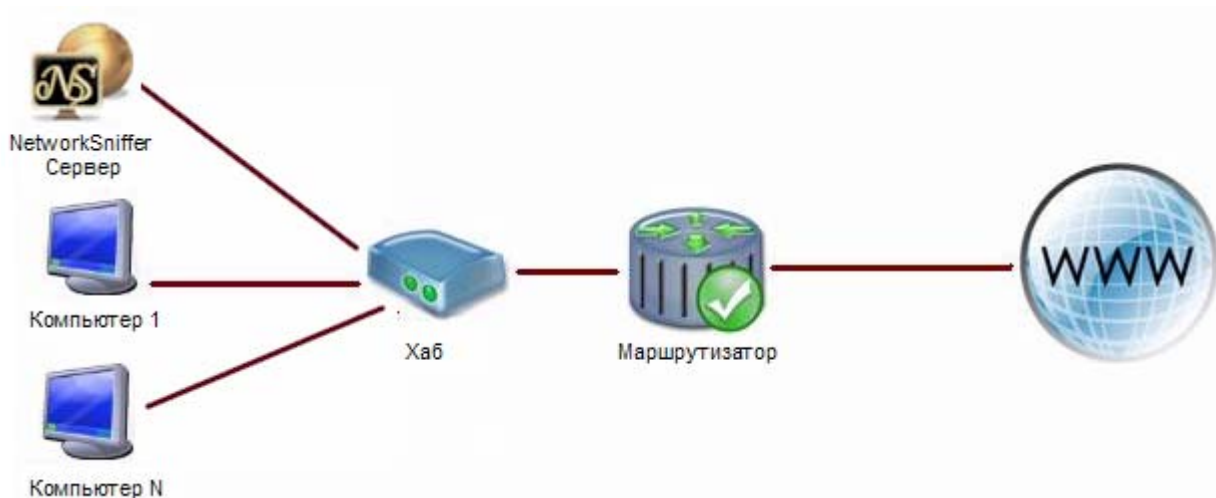
- **Шлюз**

Один из вариантов точки перехвата – сетевой шлюз. При этом через шлюз должен проходить весь входящий/исходящий трафик тех компьютеров локальной сети, которые требуется «прослушивать». При использовании аппаратного шлюза, трафик необходимо зеркалировать. При использовании программного шлюза, трафик можно и зеркалировать, и перехватывать прямо на аппаратном шлюзе.

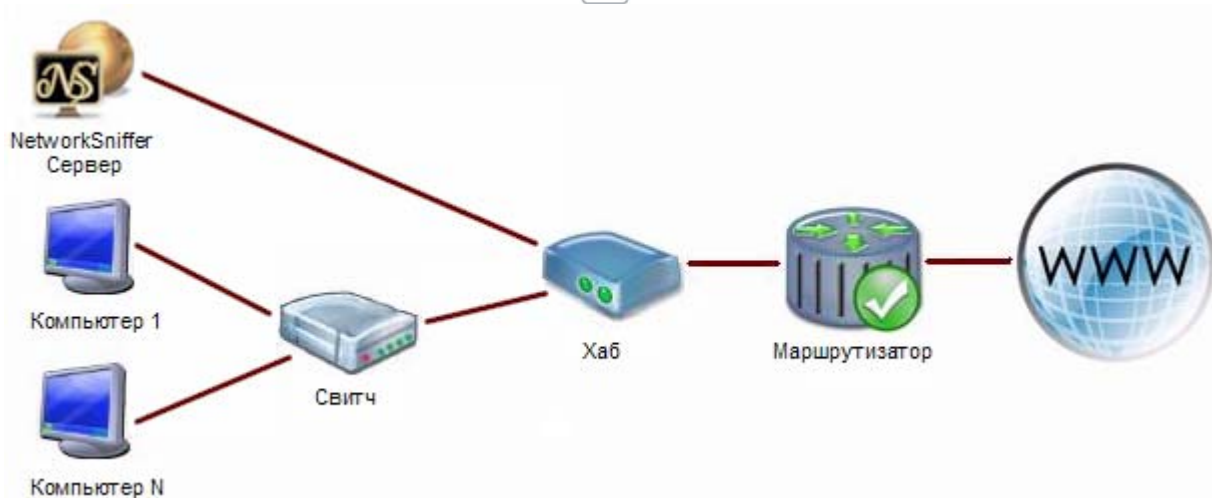
- **Хаб**

Хабы функционируют таким образом, что принимаемые на один порт сетевые пакеты рассылаются по всем остальным портам, независимо от их назначения. При подключении к одному из таких портов точки перехвата, появляется возможность перехвата всего сетевого трафика, проходящего через этот хаб.

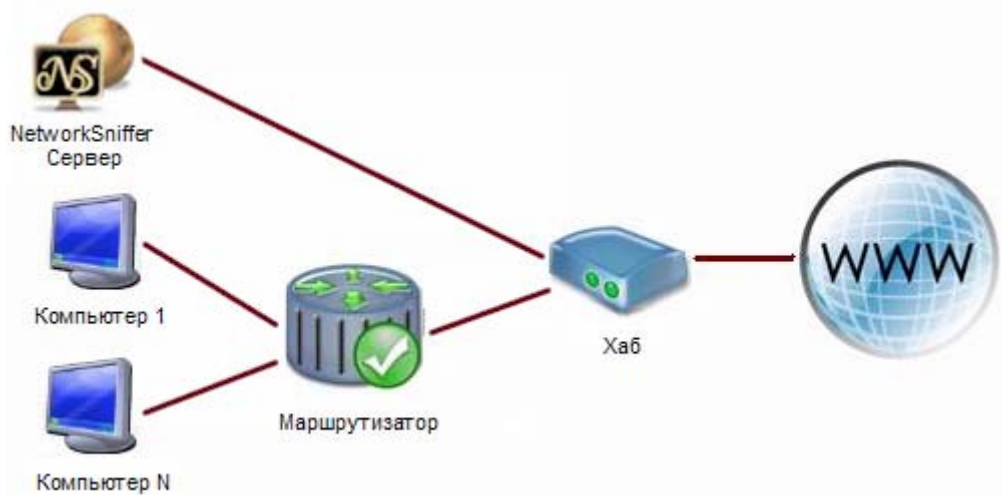
Следует заметить, что если хаб анализирует сетевые пакеты и пересылает их на соответствующие порты, то возможность перехвата всего трафика на таком хабе исчезает. Такие хабы, обладающие возможностью анализа пакетов, уже подходят под определение "свитч".



На рисунке приведён простейший способ перехвата трафика с использованием хаба. Любые компьютеры, подключенные к хабу, могут быть прослушаны, т.к. данные проходящие через маршрутизатор, направляются на все порты. Кроме того, можно перехватывать трафик, проходящий между компьютерами локальной сети.



Хаб может быть установлен между свитчем и маршрутизатором. Такая топология предоставляет возможность перехвата трафика входящего/исходящего в Интернет, но не позволяет отслеживать трафик между компьютерами локальной сети.



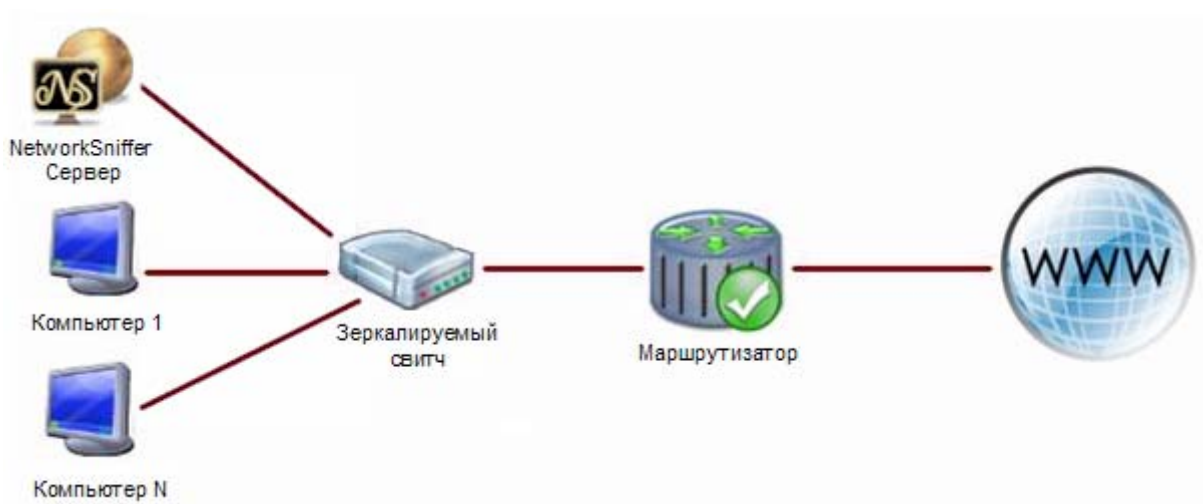
При установке хаба после маршрутизатора, существует возможность использования хаба для зеркалирования трафика. Данный способ часто используется в малых сетях с невысоким трафиком. Тем не менее, компания-разработчик не может гарантировать 100%-ную корректную работу сервиса перехвата при данной топологии сети.

- **Свитч**

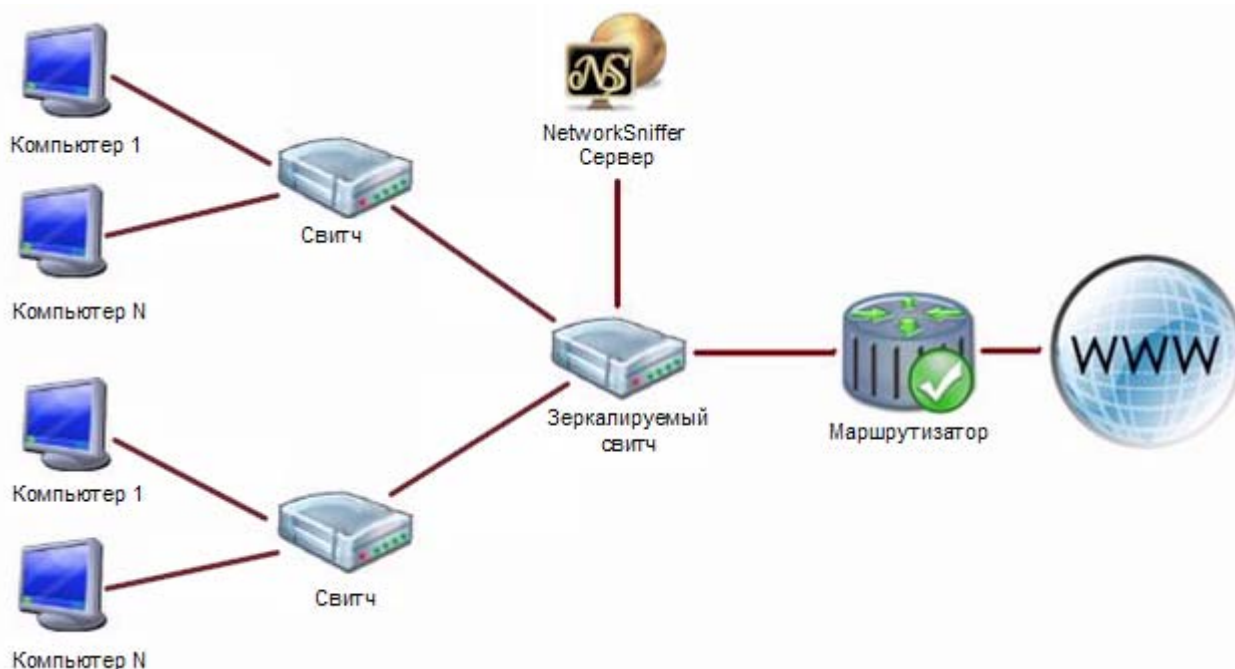
Свитчи представляют собой усовершенствованные хабы, дополненные функциями анализа сетевых пакетов для их пересылки на целевые компьютеры, что позволяет избежать дублирования трафика всей локальной сети.

Управляемые свитчи с «зеркалированием портов» позволяют дублировать сетевые пакеты с других портов на сервер перехвата. В таких свитчах можно настраивать порты,

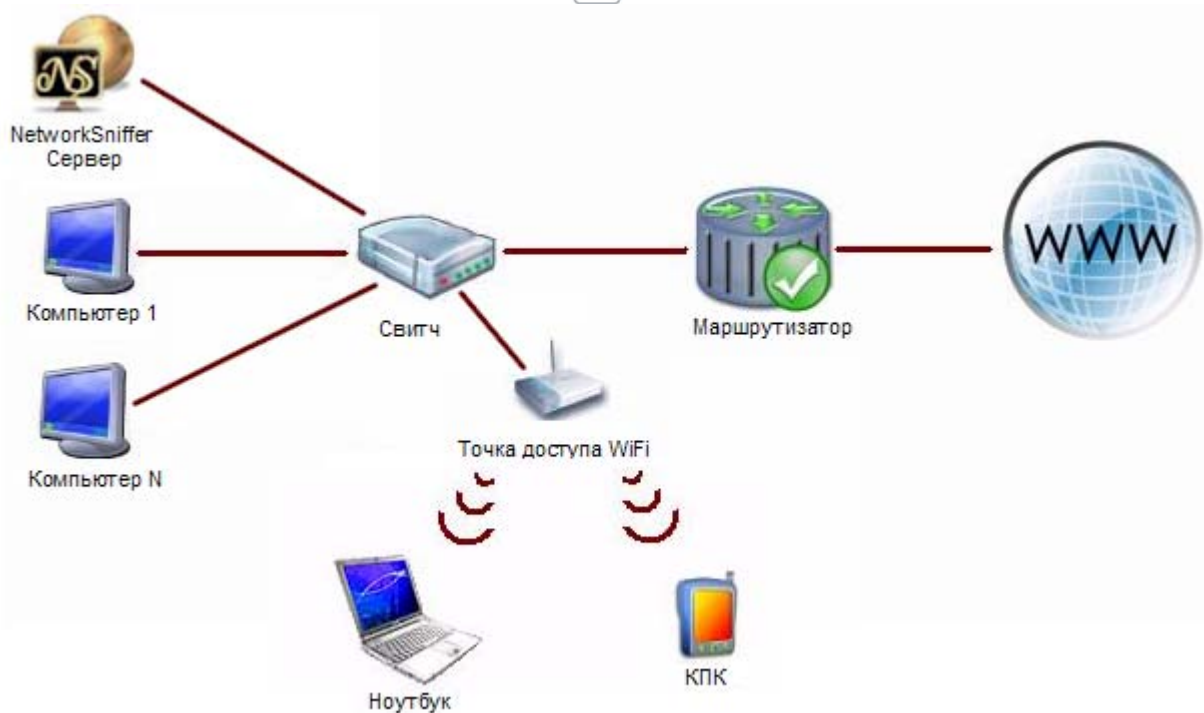
трафик с которых необходимо зеркалировать.



Сервер NetworkSniffer подключается к порту зеркалирования свитча для перехвата трафика, проходящего между локальными компьютерами и маршрутизатором. Свитч можно настроить таким образом, чтобы перехват информации производился как для всех, так и нескольких рабочих станций.



Если локальная сеть сконфигурирована с использованием свитчей, не имеющих портов зеркалирования, то можно добавить один управляемый свитч в такую сеть, который следует подключать между маршрутизатором и остальными свитчами локальной сети. Тем не менее, стоит учесть, что локальный трафик, проходящий между компьютерами в каждом сегменте, не будет перехвачен, т.к. не дойдёт до управляемого свитча.



В локальной сети может быть установлена точка доступа WiFi. В данном случае точка доступа подключается к одному из портов управляемого свитча (возможны и другие конфигурации подключения). В итоге, появляется возможность перехвата трафика компьютеров, получающих доступ в Интернет по беспроводным каналам связи.



3. Установка системы

Для установки MailSniffer необходимы три дистрибутива:

- Дистрибутив SearchInform Server
- Дистрибутив NetworkSniffer
- Дистрибутив AlertCenter

Дистрибутив SearchInform Server используется для установки сервиса индексации (SearchInform Server).

Дистрибутив NetworkSniffer позволяет установить следующие компоненты MailSniffer:

- Сервер NetworkSniffer
- Клиент MailSniffer.

Дистрибутив AlertCenter устанавливает AlertCenter.

База данных, Поддерживаются СУБД PostgreSQL или Microsoft SQL Server. В случае, если эти СУБД недоступны, из консоли NetworkSniffer можно создать базу SQLite.

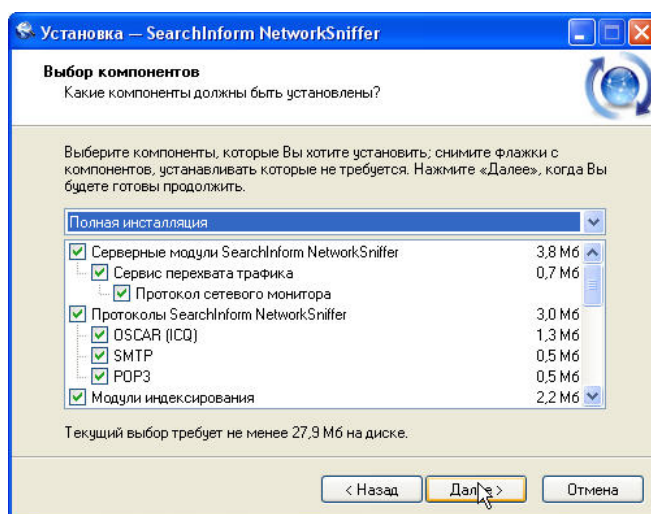
Установка MailSniffer производится в следующем порядке:

1. Установить сервис перехвата.
2. Установить сервис индексации.
3. В случае необходимости, создать базу SQLite.

Сообщения, перехваченные на одной машине, складываются в одну базу. Тем не менее, есть возможность работы с несколькими базами, например, MailSniffer поддерживает работу с несколькими хранилищами.

5. Установить клиентское приложение MailSniffer.
6. Установить AlertCenter.

3.1. Устанавливаемые модули NetworkSniffer



Дистрибутив NetworkSniffer используется при установке двух комплексов – MailSniffer и IMSniffer. Поэтому при установке MailSniffer, из дистрибутива NetworkSniffer нужно выбрать следующие модули:

Серверные модули SearchInform NetworkSniffer:

- Сервис перехвата трафика
- Протокол сетевого монитора

Протоколы SearchInform NetworkSniffer (модули поддержки перехвата по протоколам):

- SMTP
- POP3
- IMAP

Модули индексирования (подключаемые модули источника данных):

- SMTP
- POP3
- IMAP

Модули подключения к базам данных:

- Microsoft SQL Server
- PostgreSQL
- SQLite

Интерфейс пользователя:

- SearchInform NetworkSniffer Administrator Console (консоль администратора сервера)
- SearchInform NetworkSniffer Databases Control Console (консоль управления базами данных)
- SearchInform MailSniffer Client (клиент MailSniffer)

4. Технические характеристики

Для обеспечения эффективной работы системы в крупных корпоративных сетях особенно важны следующие два аспекта:

- Использование индексов
- Масштабируемость системы.

4.1. Использование индексов

Характеристика	Значение
Скорость индексирования текста	15-30 GB/час
Индексирование вложенных файлов	<p>Файлы MS Office: DOC, XLS, PPT, RTF, MDB, DOCX, XLSX, PPTX, DOT</p> <p>Файлы Internet: HTM, HTML, SHTML, CSS, JS, MAFF</p> <p>Электронные сообщения: MSG, EML</p> <p>Программные файлы MS Windows: TXT, PDF, XML, LST, CHM, BAT, LOG, INI, WRI, MHT, HLP, JPG, DJVU, NSF</p> <p>Архивы: RAR, ZIP, JAR, TAR, GZ, TGZ, ISO, CAB, 7Z, ARJ, GZIP, TGZ, TPZ, LZH, LHA, Z, TAZ, LZMA, BZ2, BZIP2, TBZ2, TBZ, HFS</p> <p>Форматы файлов языков программирования: JAVA, PAS, DFM, DPR, BAS, CPP, HPP, C, C++, H, CS, SQL, JSP, ASP, ASPX, PHP, WSDL, PY, PL, INC, VB, VBS, XLA, CMD</p> <p>Файлы PowerBuilder: SRA, SRJ, SRW, SRU, SRM, SRS, SRF, SRD, SRQ, SRP</p> <p>Аудио/видео файлы: MP3, AVI</p> <p>Файлы OpenOffice: SXW, STW, ODT, ODS</p> <p>Старые текстовые форматы: LEX, ASC</p>

Обновление индекса	Инкрементальное. Индексируются только документы, измененные или созданные с момента предыдущего индексирования.
Скорость поиска	< 1 с.
Объем индекса	10-25% размера проиндексированных файлов, но не менее 65 MB.

4.2. Масштабируемость системы

При очень высокой загрузке системы, вычислительных мощностей системы может оказаться недостаточно. Поэтому важной особенностью решений на базе SearchInform Server является *масштабируемость* - способность увеличения производительности пропорционально дополнительным ресурсам.

Для работы при высокой нагрузке сети (при большом количестве рабочих станций и большого трафика) можно использовать следующие возможности системы:

- Установка сервиса перехвата, сервиса индексации и базы данных на отдельные рабочие станции.
- Установка сервиса перехвата на отдельные рабочие станции. Например, на одной машине производить перехват по протоколу POP3, а на второй - по SMTP.

5. Системные требования

5.1. Рекомендуемые аппаратные требования

- **1...200 рабочих станций**

Для сервиса перехвата, базы данных и индексации базы (возможно размещение на одном компьютере):

- Процессор Intel Core 2 Duo E6400/ AMD 64 X2 4200+
- ОЗУ 2 ГБ

- **200...2000 рабочих станций**

Для сервиса перехвата, базы данных и индексации (возможно размещение на одном компьютере сервиса перехвата и индексации базы):

- Процессор Intel Core 2 Duo E6400/ AMD 64 X2 5200+
- ОЗУ 4 ГБ

Это ваша власть
над информацией



и должна присутствовать в сети поддерживаемая база данных:

- Microsoft SQL Server 2005+
- PostgreSQL 8.3+
- **2000 и более рабочих станций**

Сервис перехвата, индексация базы и база данных размещаются на отдельных компьютерах:

Сервис перехвата:

- 2 Процессора Phenom 9100e
- ОЗУ 8 ГБ

Индексирование базы данных:

- Процессор Intel Core 2 Duo E6400/ AMD 64 X2 4200+
- ОЗУ 4 ГБ

Базы данных:

- 2 Процессора Phenom 9100e
- ОЗУ 4 ГБ

5.2. Поддерживаемые операционные системы

- Microsoft Windows XP Professional SP1 x32 и выше
- Microsoft Windows Server 2003 x32 и выше

5.3. Поддерживаемые базы данных

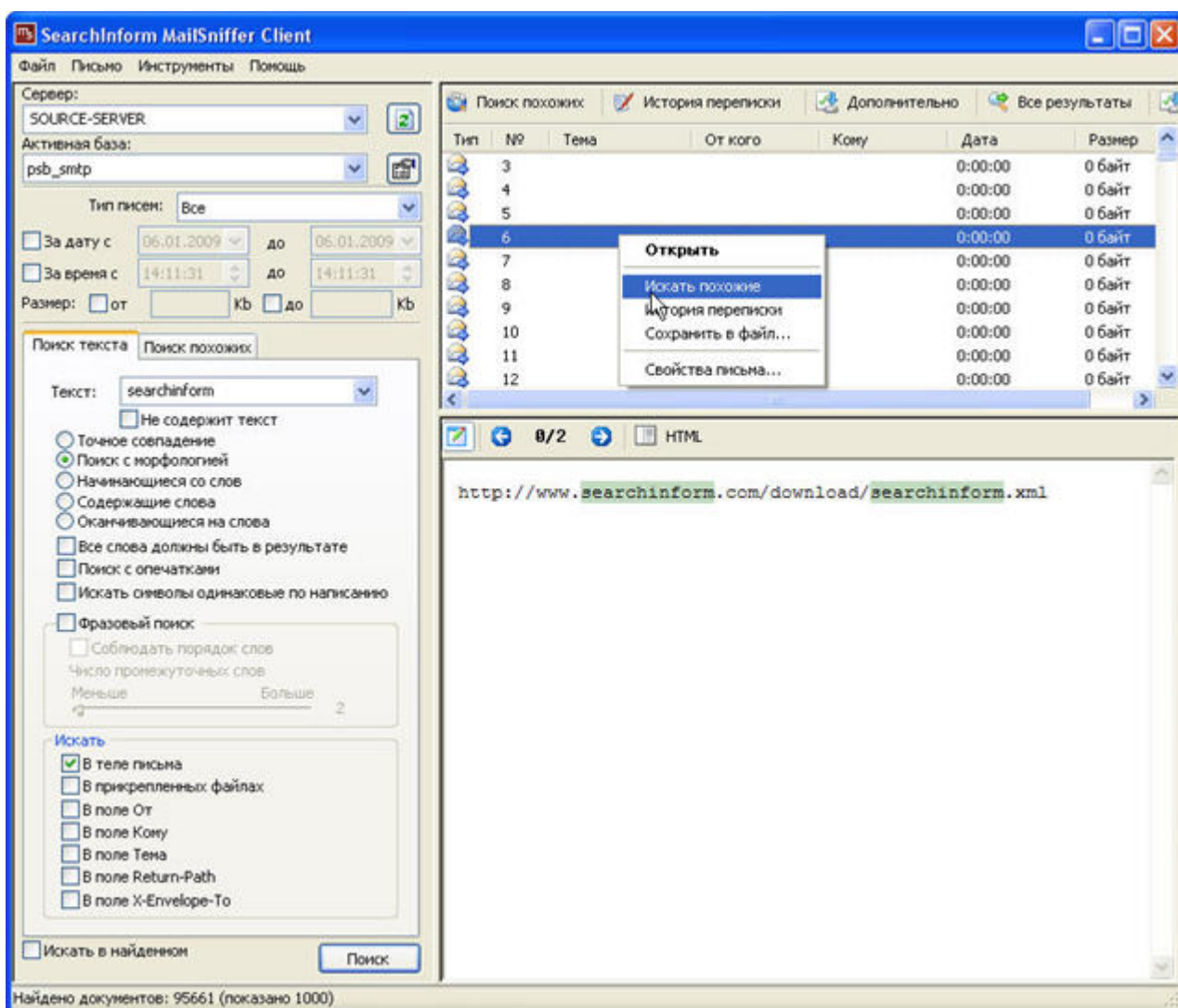
- Microsoft SQL Server 2005+
- PostgreSQL 8.3+
- SQLite (устанавливается только локально). Не рекомендуется эксплуатация SQLite в сетях с более 200 прослушиваемых рабочих станций.

6. Клиентская часть MailSniffer

Клиентская часть предназначена для выполнения поиска в базе данных электронной почты, хранящейся на сервере локальной сети, и просмотра результатов поисковых запросов.

На сервере может находиться несколько баз данных. Можно выбрать из списка необходимую базу. Если база всего одна, то программа автоматически к ней подключится.

После успешного подключения к серверу и базе, система будет готова к выполнению поисковых запросов.



6.1. Поиск



Можно выбрать тип писем, по которым будет вестись поиск нужной информации (Все, Полученные или Отправленные).

Поиск можно вести по письмам за определенный период времени. Причём временной интервал может определяться не только датами, но и (при надобности) временем с точностью до секунды. Это максимально сужает область поиска и делает выдачу результатов более релевантной.

Можно производить поиск уже среди найденных писем, которые попали в выдачу результатов по предыдущему запросу.

Поиск текста

Цель текстового поиска - обнаружение документов, содержащих ключевые слова и выражения.

Система обеспечивает гибкую настройку условий проверки по ключевым словам:

- Поиск по словам с учетом морфологии и синонимов. Простейший вид поиска, позволяющий находить документы, содержащие искомые слова, их словоформы и синонимы, вне зависимости от места их расположения в документе.
- Поиск по словам с учетом порядка слов и расстояния между ними (фразовый поиск). Для эффективного определения подозрительных документов, зачастую требуется анализировать документы не по отдельным словам, а по словосочетаниям, например, ФИО, или по устоявшимся определениям. Фразовый поиск имеет очевидные преимущества – возможность задать порядок слов и расстояние между ними. Данный вид поиска исключает документы, содержащие слова запроса, беспорядочно разбросанные по тексту документа.

Поиск похожих

Запатентованный алгоритм «Поиск похожих». Данный тип поиска позволяет отслеживать конфиденциальные данные даже в том случае, если текст документа был предварительно отредактирован. В качестве поискового запроса можно использовать как фрагменты текста, так и документы целиком. В результате поиска идентифицируются документы, содержащие запрос целиком или похожие на текст запроса.

Документы, обнаруженные в результате поиска, фильтруются по степени сходства или "релевантности". Пользователь может настроить процентное значение релевантности, при котором документ определяется как похожий.

6.2. Просмотр результатов



После выполнения поискового запроса, найденные письма отображаются в окне в виде списка. Письма можно отсортировать по любому доступному полю.

При выборе любого сообщения из списка, в просмотрщике будет отображено тело письма. Просмотрщик предназначен для быстрого ознакомления с содержимым документа, не открывая его во внешнем приложении. Зачастую это может сэкономить массу времени, так как, используя предварительный просмотр можно сразу же определить, искомый это документ или нет.

Удобство использования окна предварительного просмотра заключается еще и в том, что в нём автоматически подсвечиваются найденные ключевые слова и можно делать быстрый переход по найденным словам – это особенно актуально для объёмных документов, которые содержат небольшое количество найденных слов.

Благодаря функции “История переписки”, можно просмотреть всю переписку с адресом выбранного на данный момент письма. Причём отображаться будут только письма, отправленные на указанный адрес и полученные с него. В итоге будет показана переписка в хронологическом порядке только с данным адресатом – между двумя почтовыми ящиками.

7. Интеграция с AlertCenter

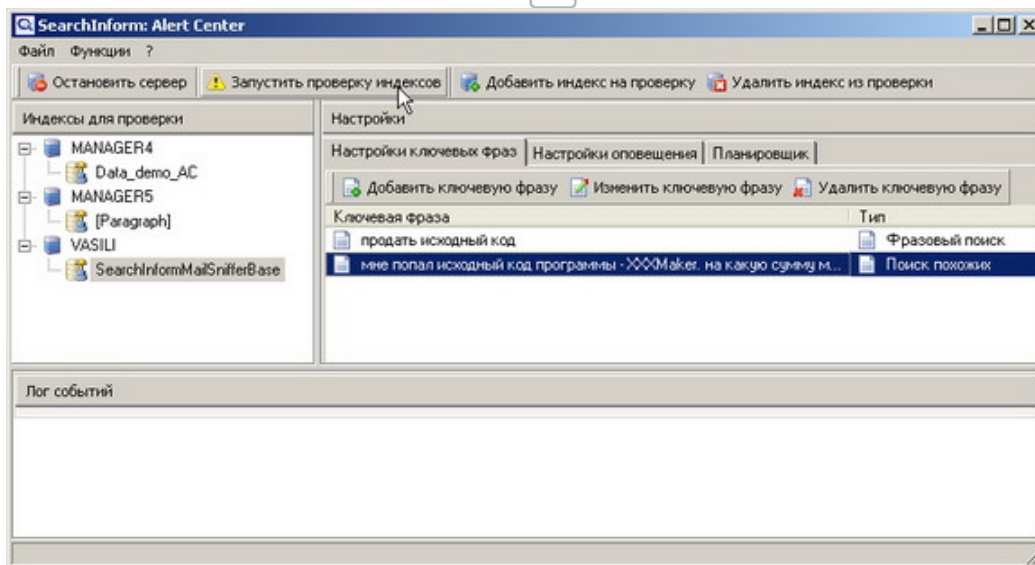
Программа **AlertCenter** предназначена для отсылки уведомлений о ключевых словах, обнаруженных в содержании документов, проиндексированных при помощи любых приложений линейки SearchInform, в том числе MailSniffer.

AlertCenter поддерживает не только фразовый поиск с богатыми возможностями настройки, но и уникальную функцию «поиска похожих». В случае обнаружения документов с критической информацией, по электронной почте высылаются сообщения, содержащие ссылку на выявленный документ и список ключевых фраз.

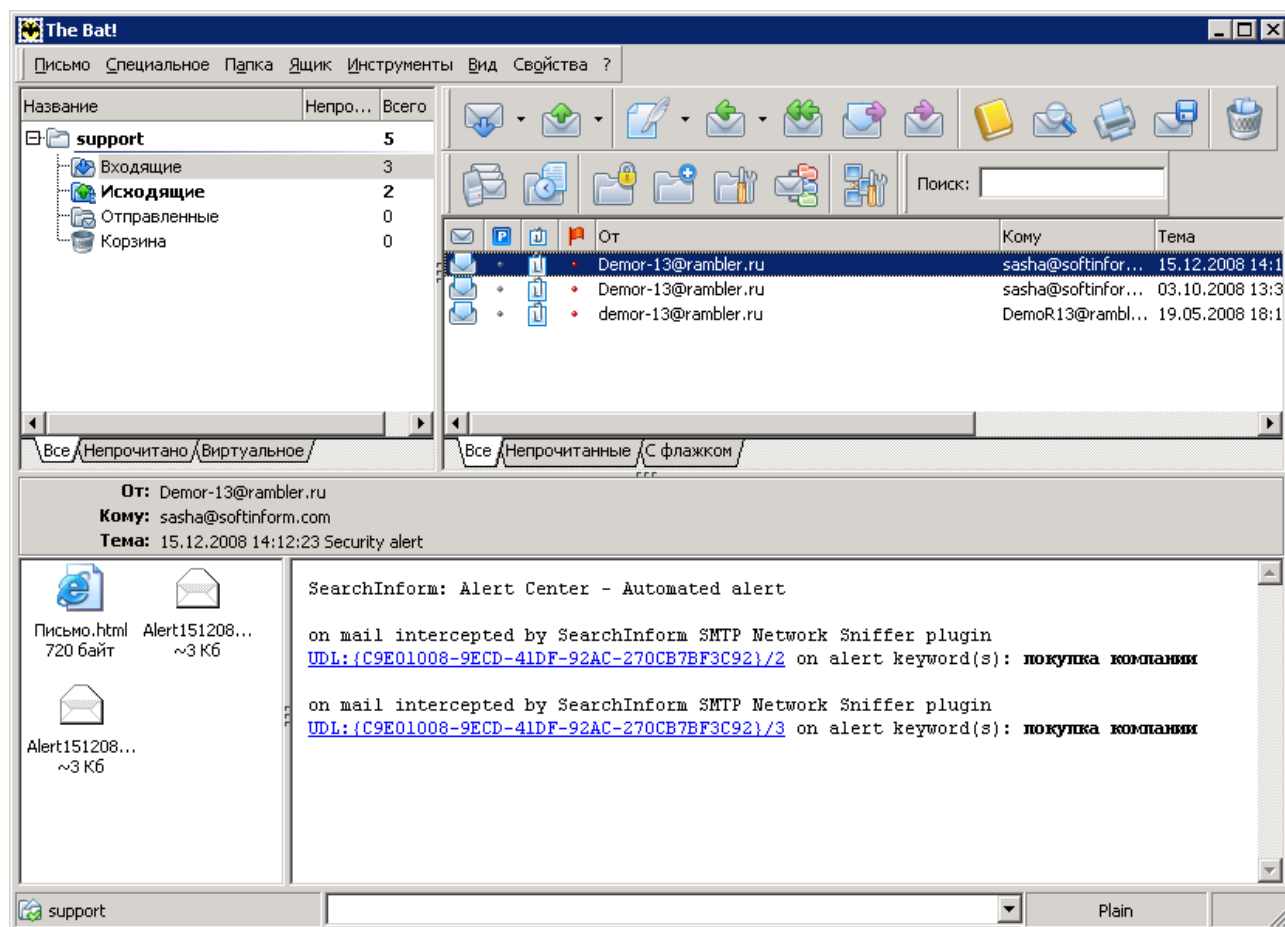
Функции AlertCenter:

- Настройка списка ключевых слов и документов с типовым содержанием
- Подключение индексов, созданных при помощи приложений семейства SearchInform
- Текстовый поиск критических документов по подключенным индексам, в том числе фразовый поиск с фиксированным порядком слов
- Поиск документов, похожих по содержанию
- Генерация извещений (алертов) при обнаружении критических документов
- Отсылка алертов по электронной почте
- Настройка расписания повторного поиска

Это ваша власть
над информацией



В случае обнаружения подозрительных документов на почту администратора приходит уведомление (алерт) со ссылкой для просмотра документа.



По ссылке открывается клиентское приложение MailSniffer непосредственно на письме, по которому произошло срабатывание.