Automated management of operational activities in forensic computer laboratories

Hernán Horacio Herrera[1, 2], Leopoldo Sebastián Gómez [1, 2]
[1]Poder Judicial de Neuquén - [2] Universidad Nacional de Río Negro
{hernanhoracio.herrera, sebastian.gomez}@jusneuquen.gov.ar
{hhherrera,sgomez}@unrn.edu.ar

Abstract. The paper shows the design of an extensible computer solution for computer forensic laboratories that will allow the automated management and monitoring of a set of tasks related to data processing to improve daily duties on digital evidence. The software works on an infrastructure composed of a local high-speed network and a group of workstations on which various forensic computer tools are executed. The forensic software tool is able to coordinate operational activities and the transfer of digital information on a set of network storage devices in which the sources of digital evidence and the results obtained from data processing are safeguarded. The framework has been designed for the automated management of operative activities and will enable the simultaneous and autonomous processing of multiple sources of digital evidence corresponding to different cases being processed in a computer forensic laboratory. Furthermore, it can also be managed through a web interface and will allow programming, controlling and reporting the progress of automated tasks that are executed on digital evidence. The results of those finalized jobs are stored in a database and after being validated they will be available and accessible through an online review system, so that judicial operators have without delays a fully set of forensic analysis reports and other potentially relevant findings that allow them an early evaluation of the digital evidence submitted to expertise. The solution proposed seeks to contribute to the automated management of operative activities in the laboratory as a first step towards the so-called forensic computer systems of second generation.

Gestión automatizada de actividades operativas en laboratorios de informática forense

Resumen. El trabajo presenta el diseño de una solución informática extensible para laboratorios de informática forense que posibilita la gestión automatizada y el monitoreo de un conjunto de tareas vinculadas con el procesamiento de datos en el marco de una pericia informática. El software funciona sobre una infraestructura compuesta por una red local de alta velocidad y un grupo de estaciones de trabajo sobre las que se ejecutan diversas herramientas de informática forense. Esta herramienta coordina las actividades operativas y la transferencia de información digital sobre un conjunto de dispositivos de almacenamiento en red en los que se resguardan las fuentes de evidencia digital

y los resultados que vayan obteniéndose del procesamiento de datos. El software para gestión automatizada de actividades operativas permite el procesamiento en forma simultánea y autónoma de múltiples fuentes de evidencia digital correspondientes a diferentes casos en trámite en un laboratorio de informática forense. Asimismo se puede administrar a través de una interfaz web y permite programar, controlar y notificar los avances tareas automatizadas que se ejecutan sobre el material probatorio. Los resultados de aquellos trabajos finalizados se almacenan en una base de datos y luego de ser validados quedan disponibles y accesibles a través de un sistema de consulta online para que los operadores judiciales dispongan en plazos menores de informes de análisis forense y otros hallazgos potencialmente relevantes que les permitan una evaluación temprana de la evidencia digital sometida a peritaje. Esta herramienta procura contribuir en la gestión automatizada de actividades operativas, avanzando hacia los llamados sistemas de informática forense de segunda generación.

Palabras clave: automatización, SOPs, laboratorios de informática forense.

1 Introducción a los frameworks para informática forense

El crecimiento en la capacidad de almacenamiento de información que presentan las computadoras y los dispositivos de telefonía celular ha ocasionado que la preservación de las fuentes de evidencia digital y principalmente la extracción de artefactos potencialmente relevantes para la investigación insuma tiempos elevados que dilatan la finalización de una pericia informática. El tiempo que demanda la adquisición y el procesamiento de información digital extraída de los dispositivos investigados se hace prohibitivo y conlleva al aumento de las listas de espera de los laboratorios. Los investigadores ocupan más tiempo en adquirir y preparar la evidencia que en analizar y presentar la misma [In de Braekt et al., 2016].

El procesamiento de las fuentes de evidencia digital llevado a cabo en un laboratorio de informática forense debe canalizarse a través de diferentes procedimientos operativos estandarizados -en adelante SOPs- que garantizan que los resultados de la actividad pericial sean verificables y repetibles, es decir, científicos [Gómez, 2015]. El método tradicional de trabajo con fuentes de evidencia digital consiste en crear una imagen forense del dispositivo de almacenamiento investigado, verificarla, copiar la misma a un repositorio digital de gran capacidad y luego procesarla con herramientas de informática forense para extraer los contenidos y artefactos digitales potencialmente relevantes, los que luego serán analizados por los responsables de la investigación. Esto se repite por cada dispositivo involucrado en el caso. Teniendo presente la considerable variedad de dispositivos móviles disponibles en la actualidad, con sistemas operativos propietarios, sistemas de archivos embebidos y la complejidad que conlleva la extracción de datos ante los nuevos métodos de protección de acceso y cifrado, cada caso presenta un nuevo desafío resultando muchas veces imposible obtener una imagen física de la memoria interna

en procura de evidencia digital. Usualmente se prepara la herramienta de informática forense especificando los parámetros de ejecución, se selecciona la fuente de evidencia digital a ser procesada y el destino en donde serán guardados los artefactos extraídos. En particular, cuando se trata de dispositivos de telefonía celular, el procedimiento forense es sutilmente diferente al tratamiento que se aplica con otros elementos tecnológicos y se dificulta aplicar la automatización en las primeras actividades operativas prescriptas en el SOP, ya que en general se requiere una considerable interacción manual con el dispositivo. En este caso existen tiempos de espera considerables ya que las actividades operativas son en su mayor parte secuenciales y sólo se inician dentro del horario laboral del laboratorio. Más allá de las particularidades propias de los dispositivos móviles, dentro de los SOPs definidos en el laboratorio de informática forense hay una gran cantidad de actividades operativas que pueden automatizarse y muchas de ellas pueden ejecutarse en paralelo, lo cual permite optimizar los recursos del laboratorio y mejorar el tiempo de respuesta a los requerimientos jurisdiccionales.

El diseño de sistemas de informática forense de segunda generación debe satisfacer requisitos esenciales como la velocidad de procesamiento, la precisión y completitud en los resultados, la confiabilidad de la herramienta, la posibilidad de ejecutar actividades operativas auditables y repetibles, y finalmente contemplar el manejo de abstracciones de alto nivel sobre los datos para permitir a los investigadores hacer revisiones sobre la evidencia digital sin la intervención constante de los peritos informáticos. Asimismo, el software debe ser: a) escalable: para poder procesar grandes volúmenes de información digital, b) extensible: con capacidades para incorporar desarrollos de terceros en forma modular y c) abierto: las salidas pueden ser utilizadas como entradas para ser vinculadas con otros sistemas.

En esta línea de trabajo, los investigadores [Vermaas et al., 2010] desarrollaron una arquitectura denominada OCFA[1] (Open Computer Forensic Architecture) que permitió automatizar ciertos procesos de informática forense. La herramienta fue construida por la policía nacional de Holanda y está compuesta por un back-end que funciona en una plataforma GNU Linux, utiliza una base de datos PostgreSQL y un indexador implementado con la tecnología Apache Lucene. En [Karabiyik & Aggarwal, 2014] se ha presentado como prueba de concepto una herramienta novel para analizar discos rígidos llamada Audit (Automated Disk Investigation Toolkit), la que fue diseñada para soportar la integración de herramientas de informática forense open source dentro de un sistema experto. Audit consta de tres componentes: una base de datos de tareas y herramientas, una base de conocimiento en la que se definen reglas y hechos y el sistema experto mencionado que asiste al usuario.

En [Van Baar et al., 2014] se expone una tecnología denominada DFaaS (Digital Forensics as a Service), la que se implementó en la infraestructura de servicios de informática forense llamada Hansken[2], actualmente utilizada por el Netherlands Forensics Institute, y que permite procesar e investigar grandes volúmenes de evidencia digital. Esta herramienta es el sucesor de Xiraf, un producto no comercial

[1] OCFA: http://ocfa.sourceforge.net/ (Último acceso: 29/04/2018).

[2] Hansken: https://www.forensicinstitute.nl/research-and-innovation/latest-forensic-innovations/hansken (Último Acceso: 29/04/2018).

de código cerrado que comenzó como un proyecto de investigación [Alink et al., 2006] en el año 2006 en el Netherland Forensic Institute y fue transferido a la Policía Nacional de Holanda en el 2010. Xiraf fue el primer software de informática forense holandés que permitió el procesamiento computarizado de la información digital recolectada en procedimientos judiciales. Hansken posibilita la interacción simultánea durante el proceso forense digital entre los investigadores digitales, los detectives y los analistas del caso. Al igual que otras metodologías tradicionales de informática forense el flujo de trabajo atraviesa una etapa de preservación en la cual se crean imágenes forenses de los dispositivos a analizar y se las almacena en un repositorio central, donde quedan disponibles para que a través de diferentes herramientas y de manera automática se extraiga información digital del sistema de archivos, archivos borrados, archivos recuperados a través de procesos de carving, archivos de logs, historiales de navegación en Internet y bandejas de correo electrónico. Los resultados mencionados son transferidos a una base de datos centralizada que puede ser consultada usando múltiples métodos. Tantos los detectives encargados de llevar adelante la investigación, como los analistas del caso, pueden acceder a los resultados a través de una interfaz web y ejecutar consultas sobre la base de datos. Por otro lado los investigadores digitales que efectúan el análisis forense pueden trabajar sobre los resultados obtenidos, los cuales son almacenados en formato XML. La administración de Hansken es llevada a cabo por un equipo de administradores de sistemas sin experticia en informática forense. Esta infraestructura centraliza el almacenamiento de las fuentes de evidencia digital y el software necesario para procesarlas, permitiendo que los usuarios puedan acceder al servicio desde diferentes lugares del país. Usando el modelo descrito, los responsables de la investigación pueden trabajar sobre los casos desde sus lugares de trabajo vía Internet, lo cual reduce el tiempo de respuesta para la obtención de prueba digital.

En [Chabot et al., 2014] se detalla una arquitectura multicapa llamada SADFC (Semantic Analysis of Digital Forensic Cases) diseñada para extraer información de fuentes heterogéneas de evidencia digital. A través de esta arquitectura se construye un modelo de conocimiento centrado en la implementación de una ontología a partir del procesamiento de registros de eventos obtenidos de los dispositivos analizados. Un evento puede ser la ejecución de una aplicación, la descarga de un archivo, una entrada en un archivo de log, etc. La arquitectura está conformada por una capa de extracción que adquiere automáticamente los registros mencionados, los filtra y los normaliza para luego entregarlos a la capa de semántica que se encarga de almacenar y extraer el conocimiento contenido en ellos a través de la implementación de una ontología. Luego la capa de razonamiento analiza el conocimiento vertido por la capa de semántica y genera nuevo conocimiento. Esta capa también provee una herramienta de análisis temporal de la información obtenida. Finalmente la capa de interfaz permite que los investigadores puedan acceder al conocimiento generado a través de una herramienta de visualización de eventos y una interfaz de consultas.

En [In de Braekt et al., 2016], los autores proponen un framework para gestionar la automatización de tareas de informática forense, con el objetivo de reducir el costo del procesamiento inicial de las fuentes de evidencia digital y darle más tiempo al investigador para el análisis y presentación de la misma. El framework utiliza herramientas forenses de terceros que pueden ejecutarse desde la línea de comandos del sistema operativo, lo cual permite integrarlas a scripts que se ejecutan sin la

intervención del usuario. Consta de tres componentes que funcionan de manera independiente: un generador de imágenes forenses, un gestor de cola de procesos -servidor de cola-, que gestiona las herramientas mencionadas, y un gestor de archivado y limpieza que se ejecuta periódicamente para archivar y limpiar los casos cerrados. Usando el framework propuesto por estos investigadores, la preparación de las herramientas se efectúa automáticamente después de que la imagen forense es creada para que se ejecuten en paralelo sobre las fuentes de evidencia digital sin intervención humana. Mediante este método, el tiempo entre los procesos secuenciales es nulo y se pueden iniciar fuera del horario laboral del personal. Las pruebas efectuadas por los autores revelan que el proceso forense ejecutado con el framework es un 65% más rápido que el proceso tradicional.

DFORC2 es un proyecto open source [Gonzales, D. et al, 2017] que extiende las funcionalidades de Autopsy. Este framework utiliza cómputo distribuido para procesar evidencia digital y permite paralelizar la ingestión de evidencia y el procesamiento de la misma. Puede funcionar sobre un clúster local o bien utilizar cómputo en la nube a través de los servicios Web de la empresa Amazon (AWS). Esta aplicación no integra herramientas forenses de terceros y utiliza una versión headless de Autopsy que funciona en los nodos de procesamiento (workers).

Algunos proveedores de software comercial para informática forense están incursionando en el camino hacia una segunda generación de herramientas para informática forense. Access Data ha comenzado a ofrecer Quin-C[3], una plataforma para informática forense con mayores capacidades colaborativas entre investigadores y peritos informáticos mediante una refinada interfaz web desarrollada en HTML5. En igual sentido empiezan a presentarse frameworks de código abierto como Turbinia[4] que permiten la distribución y paralelización de trabajos, procurando la automatización de tareas usuales de informática forense.

2 Evaluación de herramientas de informática forense

En la actualidad no se dispone de un software open source que facilite la gestión automatizada de actividades operativas para un laboratorio de informática forense, integrando múltiples herramientas forenses desarrolladas por terceros y que contemple la posibilidad del uso de aplicaciones específicas que permitan procesar evidencia digital en modo distribuido. Salvo contadas excepciones, las soluciones de automatización que provienen de la academia o bien generadas por especialistas en informática forense han sido siempre ad-hoc, brindando soluciones puntuales a los tópicos de mayor demanda en la comunidad forense informática. Se han desarrollado herramientas específicas de informática forense que procesan artefactos digitales del sistema operativo en un esfuerzo por lograr una correlación automática de datos que facilite el análisis de eventos digitales. A título de ejemplo, Log2timeline permite procesar diferentes archivos de logs y otros artefactos digitales que se pueden encontrar en una computadora y produce un archivo de salida que puede ser usado

[3] Quin-C: https://accessdata.com/products-services/quin-c (Último acceso: 29/04/2018).
[4] Turbinia: https://github.com/google/turbinia (Último acceso: 06/11/2018).

para presentar cronológicamente una serie de eventos usando herramientas tales como mactime, incluida en la suite TSK. Es posible citar un sinnúmero de utilidades de informática forense para el procesamiento de artefactos digitales, sin embargo han sido pocos los intentos de crear plataformas de gestión de laboratorio para informática forense que contemplen la automatización y el procesamiento distribuido con un enfoque más amplio y orientado al uso concurrente de múltiples herramientas de informática forense.

Una revisión de varias herramientas de informática forense permite comprobar que la mayor parte del software comercial permite que los profesionales extiendan sus funcionalidades, pero mantienen sus propios lenguajes de scripting lo que limita su portabilidad o bien no son suficientemente flexibles para integrar herramientas desarrolladas por terceros. Si bien existen productos como ADLab[5] y EnCase[6] que procesan grandes volúmenes de datos utilizando procesamiento distribuido, y sin contar con que el costo de estas herramientas resulta en muchos casos prohibitivo para muchos laboratorios de informática forense, en todos los casos la gestión de actividades operativas automatizadas se realiza siempre sobre las funcionalidades disponibles para cada uno de ellos. Por otra parte, ADLab sólo admite la integración parcial con algunas herramientas de terceros (v.gr. Belkasoft add-on, ProjectVIC hashes) y Encase restringe su portabilidad al utilizar un lenguaje propietario llamado Enscript para permitir a terceros el desarrollo de nuevas funcionalidades que se integran a la herramienta. La evaluación efectuada (ver Tabla 1) contempla otras soluciones informáticas de código cerrado, herramientas open source y frameworks.

Autopsy es un software para informática forense open source que brinda una interfaz gráfica de la biblioteca The Sleuth Kit[7] (v.gr. TSK), siendo este último una colección de herramientas para el análisis de imágenes forenses. De forma análoga a los productos comerciales líderes del mercado, Autopsy permite procesar y analizar múltiples fuentes de evidencia digital. Esta herramienta tiene una arquitectura que admite la incorporación de módulos desarrollados en Java o Python pero no posee capacidades de cómputo distribuido. Se han desarrollado algunos prototipos como Sleuth Kit Hadoop Framework, pero estas iniciativas no están suficientemente maduras para ser utilizadas en laboratorios de informática forense.

Herramienta	Procesamiento distribuido	Integra herramientas forenses de terceros	Existe una implementación disponible	Open source	Comercial
OCFA	No	No	Sí	Sí	No
Audit	No	Sí	Sí	Si	No
Hansken	Sí	Sí	No	No	No

[5] ADLab: https://accessdata.com/products-services/ad-lab (Último acceso: 29/04/2018).
[6] EnCase: https://www.guidancesoftware.com/encase-forensic (Último acceso: 29/04/2018).
[7] Sleuth-Kit: https://www.sleuthkit.org (Último acceso: 29/04/2018).

SADFC	No	No	No	No	No
Workflow Management Automation Framework	No	Sí	No	No	No
DFORC2	Sí	No	No	Sí	No
AD Lab AccessData	Sí	No[8]	Si	No	Sí
EnCase Forensic	Si	No[9]	Sí	No	Sí
Autopsy	No	Si	Sí	Si	No
Tracks Inspector	Sí	No	Sí	No	Sí

Tabla 1. Evaluación de herramientas de informática forense

Tracks Inspector[10] es un producto comercial orientado al descubrimiento de evidencia digital resultante de un trabajo de investigación [Henseler et al., 2014] que anuncia estar diseñado para que investigadores no técnicos puedan analizar evidencia digital desde un web browser. La arquitectura de este producto presenta un sistema con varios procesos que se ejecutan en múltiples servidores y que se conectan entre sí a través de llamados a procedimientos remotos (RPC). Entre los hosts que componen la arquitectura se destacan el monitor de evidencia, al cual se conectan lo dispositivos analizados, y el controlador de evidencia, que asigna las tareas de análisis a las denominadas unidades de procesamiento. Existe también una base de datos en la cual se almacenan los resultados obtenidos. El sistema cuenta también con un front-end que ofrece una interfaz web intuitiva desarrollada en HTML5. Tracks Inspector procesa fuentes de evidencia digital que han sido identificadas por el sistema y han sido asignadas a un caso por un administrador. Durante el procesamiento todos los metadatos de los artefactos forenses son extraídos y almacenados en la base de datos de evidencia, a la cual acceden los usuarios.

3 Gestión automatizada de actividades operativas

La solución informática propuesta está compuesta por un conjunto de herramientas de informática forense que se ejecutan de manera autónoma y procesan fuentes de

[8] Parcialmente con algunas herramientas comerciales.
[9] Parcialmente mediante un lenguaje propio de scripting.
[10] Tracks Inspector: https://tracksinspector.com (Último acceso: 07/04/2018).

evidencia digital almacenadas en un conjunto de repositorios compartidos a través de una red datos de alta velocidad. Los resultados obtenidos son transferidos a una base de datos, para luego ser publicados en la intranet de la organización a través de un servidor web al cual tienen acceso los operadores judiciales.

Las herramientas de informática forense que se integran en el sistema están gestionadas por una aplicación de coordinación llamada FJP -Forensic Job Processor-, la cual se ejecuta en un servidor de aplicaciones y es la encargada de ofrecer el servicio de procesamiento de las fuentes de evidencia digital. A través de una interfaz web los analistas forenses del laboratorio pueden conformar una lista de actividades operativas que se ejecutan sobre las imágenes forenses de los diferentes dispositivos a analizar o bien a posteriori sobre el corpus digital resultante de otras tareas previas de procesamiento de información digital. Las tareas se ejecutan en los denominados hosts de procesamiento (ver Figura 1) utilizando herramientas de informática forense de terceros allí instaladas, las que son gestionadas por una aplicación local denominada Worker, subordinada al coordinador FJP bajo el modelo de comunicación Master/Slave.

Para efectuar las tareas programadas se utilizan principalmente aquellas herramientas de informática forense que admitan su ejecución desde la línea de comandos de un sistema operativo. Esto permite que se integren a un script de procesamiento por lotes o bien que se invoquen desde una aplicación Worker, como ocurre en este caso. Asimismo puede considerarse el desarrollo de otros scripts para automatización de actividades operativas mediante lenguajes de automatización como AutoIt para aquellas herramientas de informática forense que requieren la utilización de una interfaz gráfica para el procesamiento de datos.

Los servicios mínimos que ofrece el sistema de gestión automatizada de actividades operativas son los siguientes:

- Carving de archivos de imagen/video digital y documentos de ofimática
- Extracción de fotogramas desde archivos de video digital
- Extracción de rostros desde archivos de imagen digital
- Detección de pornografía en archivos de imagen y video digital
- Indexación de contenidos digitales
- Búsqueda de contenidos digitales mediante palabras clave
- Extracción de información relacionada con la actividad del usuario en la Web
- Archivado y relocalización de evidencia digital

A las herramientas tradicionales y bien conocidas por la comunidad de informática forense para el procesamiento de fuentes de evidencia digital se irán sumando otras aplicaciones desarrolladas ad-hoc en el laboratorio de informática forense con el objeto de aprovechar los beneficios de la automatización de actividades operativas[11]. Todas ellas se ejecutan desde la línea de comandos del sistema operativo y resultan de utilidad en diversas etapas del proceso forense digital.

A continuación se detallan las principales herramientas que se utilizan:

[11] GetFrames, Faces, Mover, Indexer y PornDetectionCluster son herramientas de informática forense desarrolladas ad-hoc.

- PhotoRec[12]: recupera archivos de imagen y video digital, así como también documentos desde un disco rígido o bien desde una imagen forense.
- Bulk Extractor[13]: permite escanear una imagen de un disco rígido, un archivo o un directorio. Este software extrae información útil sin procesar el sistema de archivos o las estructuras del mismo. Los resultados pueden ser fácilmente inspeccionados, analizados o procesados con herramientas automatizadas. Esta herramienta también crea un histograma de las características halladas en la fuente analizada. Bulk extractor puede también ser utilizado para procesar otros dispositivos de almacenamiento tales como memorias flash, medios ópticos y capturas de paquetes de red.
- GetFrames extrae un número determinado de fotogramas de un conjunto de archivos de video digital.
- Faces: identifica rostros en un archivo de imagen digital. Puede ser configurada para que desde un conjunto de imágenes seleccione aquellas en las que existe un número determinado de rostros.
- Mover: permite relocalizar un conjunto de archivos en diversos repositorios del laboratorio asegurando la integridad de los mismos mediante la aplicación de algoritmos de hashing.
- Indexer: genera índices sobre grandes volúmenes de datos haciendo uso de la herramienta DtSearch[14] para luego poder efectuar búsquedas sobre ellos.
- PornDetectionCluster: detecta contenidos pornográficos en archivos de imagen digital. Una de sus principales características es que utiliza el poder de cómputo de las estaciones de trabajo vinculadas mediante una red local.

[12] PhotoRec: http://www.cgsecurity.org/wiki/PhotoRec (Último acceso: 29/04/2018).
[13] Bulk Extractor: https://github.com/simsong/bulk_extractor/wiki (Último acceso: 29/04/2018).
[14] DTSearch: https://www.dtsearch.com (Último acceso: 09/04/2018).

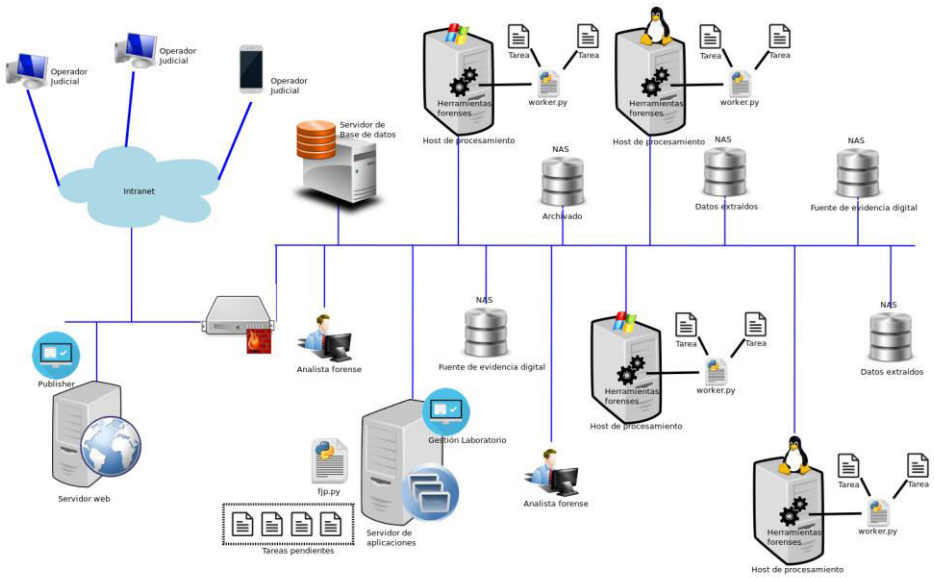

Figura 1: Arquitectura para la gestión automatizada de actividades operativas

4 Descripción del flujo de trabajo de FJP

A partir del montaje de las imágenes forenses de los dispositivos correspondientes a los casos que integran la lista de espera del laboratorio se obtienen los datos de entrada para las diferentes actividades operativas (ver Figura 2).

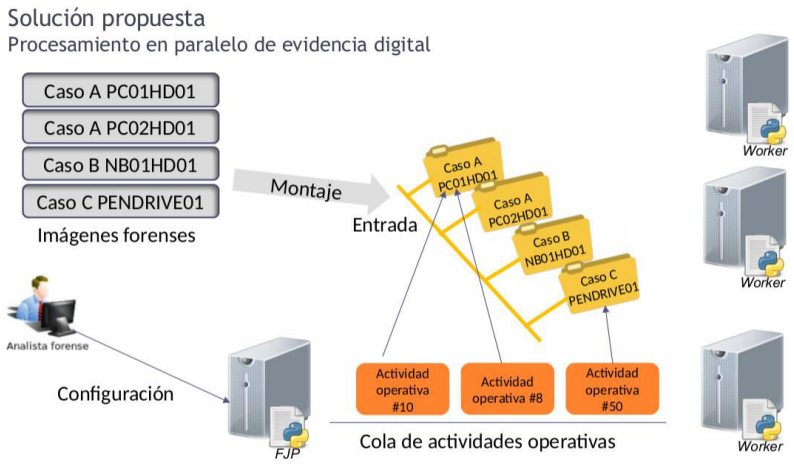

Figura 2: Registro de actividades operativas

Las herramientas ejecutadas por los workers dejan sus resultados en repositorios destinados para tal fin (ver Figura 3). Usualmente se trata de dispositivos de almacenamiento en red que son parte de la infraestructura del laboratorio.

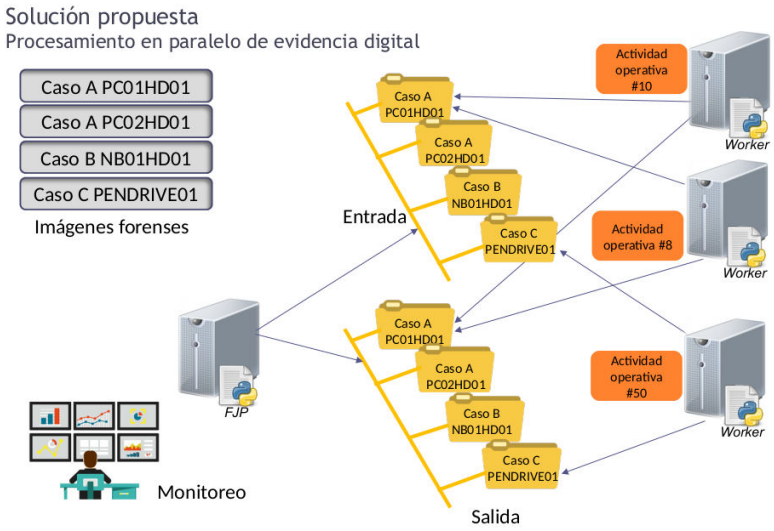

Figura 3: Procesamiento distribuido de actividades operativas

Cuando el proceso correspondiente a la aplicación FJP recibe una notificación de finalización de tarea proveniente de un worker, procede a publicar los resultados en la base de datos del sistema y a notificar a los analistas forenses sobre los nuevos resultados disponibles para revisión (ver Figura 4). Dichos usuarios inician una sesión en la aplicación Publisher, la cual reside en el servidor web de la infraestructura y es la encargada de publicar los resultados obtenidos previa revisión y autorización de los analistas forenses.

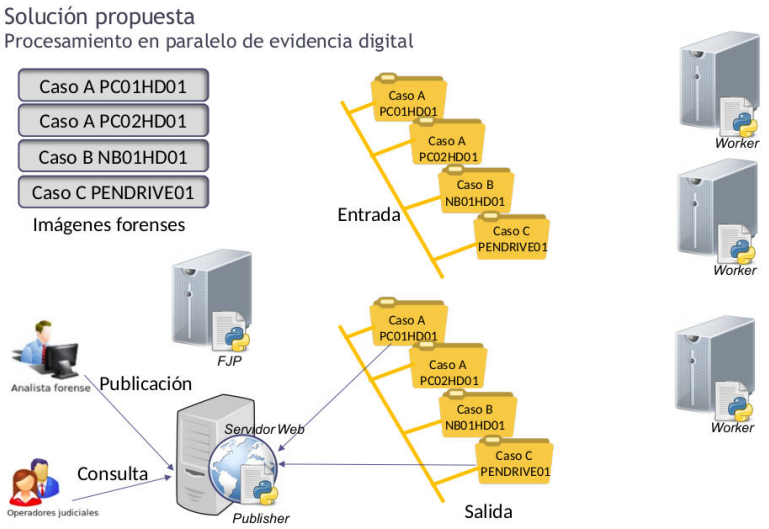

Figura 4: Publicación de resultados para revisión por operadores judiciales

El framework cuenta también con un módulo de archivado que se encarga de transferir la evidencia digital de los casos cerrados a un repositorio en la red utilizado para este propósito. Este proceso se activa cuando se cierra un caso en el sistema de gestión del laboratorio. Luego de un tiempo determinado, el cual depende de las políticas de gestión definidas en el laboratorio, la evidencia digital archivada es borrada definitivamente para liberar recursos de almacenamiento.

5 Conclusiones

Las arquitecturas empleadas por las herramientas de informática forense de primera generación han alcanzado su límite operativo y no están preparadas para manejar el aumento del volumen digital y la complejidad de las investigaciones actuales. El Big Data plantea dos importantes desafíos: el almacenamiento y el procesamiento de datos en aquellos casos en los que se deben superar las limitaciones los sistemas de archivos y computadoras tradicionales. Las herramientas actuales de informática forense no están preparadas para procesar en tiempos aceptables grandes volúmenes de datos contenidos en imágenes forenses. La mayor parte de ellas están pensadas para ser utilizadas un equipo informático de altas prestaciones, y si bien son capaces de utilizar múltiples núcleos, recaen en la idea de "llevar los datos al código". Por otra parte, la oferta de herramientas de informática forense con capacidad de ofrecer procesamiento distribuido está sumamente acotada.

Existen nuevas soluciones informáticas que intentan abordar estos problemas mediante el uso de sistemas de archivos distribuidos como Google File System y Hadoop Distributed File System. El surgimiento de frameworks como Hadoop junto el modelo de programación MapReduce ha sido diseñado para enfrentar los desafíos planteados por el Big Data utilizando la idea de "llevar el código a los datos". Mientras que el cómputo de altas prestaciones tiene por objeto el abordaje de problemas que exceden las capacidades de un CPU, el modelo MapReduce se enfoca en problemas centrados en la entrada/salida de datos. Estas tecnologías están siendo estudiadas y se han desarrollado pruebas de concepto para su aplicación en informática forense, sin embargo plantean otras complejidades como la reescritura de las herramientas de informática forense para poder adaptarlas al nuevo paradigma.

La evaluación de las principales herramientas de informática forense ha señalado que no existe un producto maduro capaz de aprovechar simultáneamente los beneficios de la extensibilidad de funcionalidades mediante la integración de herramientas desarrolladas por terceros y la productividad que brinda el cómputo distribuido sumada a la automatización de actividades operativas. La próxima generación de herramientas de informática forense necesitará de métodos distribuidos para el almacenamiento y procesamiento de evidencia digital.

Se han presentado los aspectos centrales y las funcionalidades requeridas durante el diseño de una solución informática orientada a la gestión automatizada de actividades operativas para laboratorios de informática forense. El uso del sistema de gestión automatizada de actividades operativas permite programar tareas de

procesamiento de datos sobre múltiples fuentes de evidencia digital y reducirá el tiempo total de la pericia informática mediante la ejecución en paralelo de herramientas de informática forense, sin perder de vista todos aquellos requisitos esenciales de confiabilidad e integridad que garantizan la solidez de la prueba digital producida.

La solución propuesta mejora la interacción entre el laboratorio pericial y los investigadores, posibilitando la revisión de informes periciales y contenidos digitales en línea a medida de que se van concluyendo determinados flujos de trabajo, los que son pautados mediante SOPs en el marco de la pericia informática. FJP puede ser categorizado como un sistema de gestión de flujo de trabajo forense (Forensic Workflow Management System) y principia una nueva etapa hacia el desarrollo de sistemas de informática forense de segunda generación.

Referencias

1. Alink, W., Bhoedjang R., Boncz P., De Vries A. (2006), XIRAF - Ultimate Forensic Querying, Digital Forensic Research Conference USA, Agosto de 2006.
2. Ayers, D. (2009), A second generation computer forensic analysis system, Digital Forensic Research Workshop.
3. Chabot, Y., Bertaux, A., Nicolle, C., Kechadi, T. (2014), Automatic Timeline Construction and Analysis for Computer Forensics Purposes. 10.13140/2.1.3595.1040.
4. Gómez, L. (2015), Análisis forense de dispositivos de telefonía celular mediante procedimientos operativos estandarizados, 15° Simposio Argentino de Informática y Derecho, 44 JAIIO - SID 2015 - ISSN: 2451-7526.
5. Gonzales, D., Winkelman, Z., Tran, T., Sanchez, R., Woods, D., Hollywood, J. (2017), Digital Forensics Compute Cluster (DFORC2): A High Speed Distributed Computing Capability for Digital Forensics, International Science Index, Computer and Information Engineering Vol:11, No:8, 2017 waset.org/Publication/10007817
6. Henseler, J., Hofse, J., Post, A. (2014), Tracks Inspector: Putting Digital Investigations in the Hands of Investigators, ISBN: 978-0-9891305-7-8 ©2014 SDIWC978-0-9891305-7-8 ©2014 SDIWC978-0-9891305-7-8 ©2014 SDIWC
7. In de Braekt, R., Le-Khac, N., Farina, J., Scanlon, M., Kechadi, T. (2016), "Increasing Digital Investigator Availability Through Efficient Workflow Managment And Automation", 4th IEEE International Symposium on Digital Forensics and Security (ISDFS).
8. Karabiyik, U., Aggarwal, S. (2014), Audit: Automated Disk Investigation Toolkit, Journal of Digital Forensics, Security and Law, Volume 9|Number 2, Article 11.
9. Van Baar, R., Van Beek, H., Van Eijk, E. (2014), Digital Forensics as a Service: A game changer, Digital Investigation 11, S54-S62, Elsevier.
10. Vermaas, O., Simons, J., Meijer, R. (2010), Open Computer Forensic Architecture a Way to Process Terabytes of Forensic Disk Images. Open Source

Software for Digital Forensics, Huebner, E., Zanero, S. (Eds.). Springer US, 2010, 45-67.