Independent Expert Assessment of MarkMonitor AntiPiracy Methodologies

[REDACTED]

November 1, 2012

STROZ FRIEDBERG

Table of Contents

Executive Summary

1. Background

Stroz Friedberg, LLC was retained by the Center for Copyright Information ("CCI") as an Independent Expert under the July 6, 2011 CCI Memorandum of Understanding ("MOU") executed between the Content Owner Representatives and the Participating ISPs, both as defined by the MOU. Under the MOU, the Content Owner Representatives are required to develop and maintain Methodologies (as defined by the MOU) for identifying instances of Peer-to-Peer ("P2P") online infringement and gathering reliable evidence that the identified content was uploaded, downloaded, copied, or offered on a P2P network. The Independent Expert, in turn, is required by the MOU to review on a periodic and ongoing basis the Methodologies and any modifications thereto, and recommend enhancements as appropriate, with the goal of ensuring and maintaining confidence on the part of the Content Owner Representatives, the Participating ISPs, and the public in the accuracy and security of the Methodologies.

To this end, the Content Owner Representatives have retained the company MarkMonitor to implement its P2P antipiracy services to identify instances of online copyright infringement. Stroz Friedberg conducted an assessment of the MarkMonitor Methodologies for monitoring, verifying, and enforcing online copyright infringement on P2P file sharing networks. Stroz Friedberg assessed the efficacy of MarkMonitor's Methodologies to monitor, identify, collect evidence, and generate notices to P2P infringers by conducting a series of in-person and remote interviews, reviewing documentation, and conducting technical analysis. This report details Stroz Friedberg's findings from this assessment. Stroz Friedberg reserves the right to amend, modify, or supplement this report as necessary or based on newly discovered information.

2. Summary of Findings and Recommendations

Based on our analysis and review of MarkMonitor's Methodologies, Stroz Friedberg found that:

- MarkMonitor's Methodologies effectively identify P2P online copyright infringers.
- MarkMonitor Methodologies are well-developed and matured.
- MarkMonitor's evidence collection in connection with P2P infringement is robust, defensible, and will withstand adverse party scrutiny or evidentiary challenges.
- The Methodologies include appropriate checks and balances at key points in the work flow to ensure accuracy.

- The reporting and notice-generating abilities allow MarkMonitor to accurately report on identified infringers.
- The Methodologies have a number of inherent and added system redundancies designed to ensure that MarkMonitor can provide continuous and consistent scanning.

Though on the whole we found MarkMonitor's present Methodologies to be well developed and robust, we offer the following recommendations as potential enhancements:

- Consider adding additional checks and auditing to the file identification and verification protocols.
- Increase the percentage of verified infringing works deployed to the remote [collection mechanisms].
- Consider modifications to the infringement file identification process to include works with obfuscated names, or that are sourced from private trackers.
- Create and maintain a cryptographic hash value for each collected case file.
- Regarding ISP notifications, to the extent possible, move away from reliance on email/SMTP and towards the use of other secure protocols such as HTTPS.
- Develop an audit framework that allows for assessment and evaluation of the effectiveness of the Methodologies on an ongoing periodic basis.

These recommendations have been communicated to MarkMonitor, and we understand that many have been implemented or are being considered for implementation in the future.

Independent Assessment Methodology

MarkMonitor has been extremely helpful in preparing detailed documentation, making appropriate resources available, responding to our specific inquiries, and overall cooperating with the Independent Expert assessment. Stroz Friedberg's independent assessment involved: (1) a review of MarkMonitor's network, application components, and antipiracy workflow; (2) in-person and remote walk-throughs of the antipiracy work flow from beginning to end; (3) identification and review of in-process verification checks and controls and potential breakpoints; (4) limited live testing of the antipiracy system and technical analysis of captured P2P content; and (5) multiple live/real-time reviews of the MarkMonitor antipiracy applications and systems. Sequentially, our assessment began with an overview of MarkMonitor's network and systems and progressed through the following specific topics, organized loosely with discrete steps in MarkMonitor's own antipiracy work flow:

- Identification of infringing files and file verification;
- Search for and identification of infringing files using P2P data [collection mechanisms];
- Infringement verification and evidence collection; and
- Notice generation and enforcement.

Outlined immediately below is a list of the documentation that Stroz Friedberg reviewed in connection with our assessment.

- MarkMonitor [REDACTED] Report, dated [REDACTED];
- MarkMonitor Presentation entitled [REDACTED];
- Screenshots of various components of the MarkMonitor antipiracy system;
- Test torrent files prepared by MarkMonitor;
- Case file packages generated from Stroz Friedberg's test of P2P content hosting; and
- Infringement Notice generated as a result of Stroz Friedberg's test P2P content hosting.

Assessment of MarkMonitor AntiPiracy Platform

This section provides an overview of the MarkMonitor AntiPiracy Methodologies and outlines the steps it takes to monitor, verify, collect evidence regarding, and enforce against online copyright infringement on P2P file sharing networks. For purposes of our assessment, Stroz Friedberg focused specifically on the extent to which these steps accurately identify copyright infringers; the checks and balances in place to ensure accuracy; and where any potential process break points exist.

1. Platform Overview

The MarkMonitor AntiPiracy platform consists of an enterprise client-server environment designed to scan P2P networks and websites, collect and preserve evidence, and send infringement notices. The system has been in use and evolving for over eight years, and in this time has developed into a reliable, dynamic, and highly extensible system. In connection with MarkMonitor's MOU-related work, the system currently monitors [REDACTED].

At the highest level, the MarkMonitor platform consists of [scanning systems], multiple [databases], [collection mechanisms], and [configuration systems]. Figure 1 below

visually depicts the antipiracy workflow and the interaction of these systems and databases.

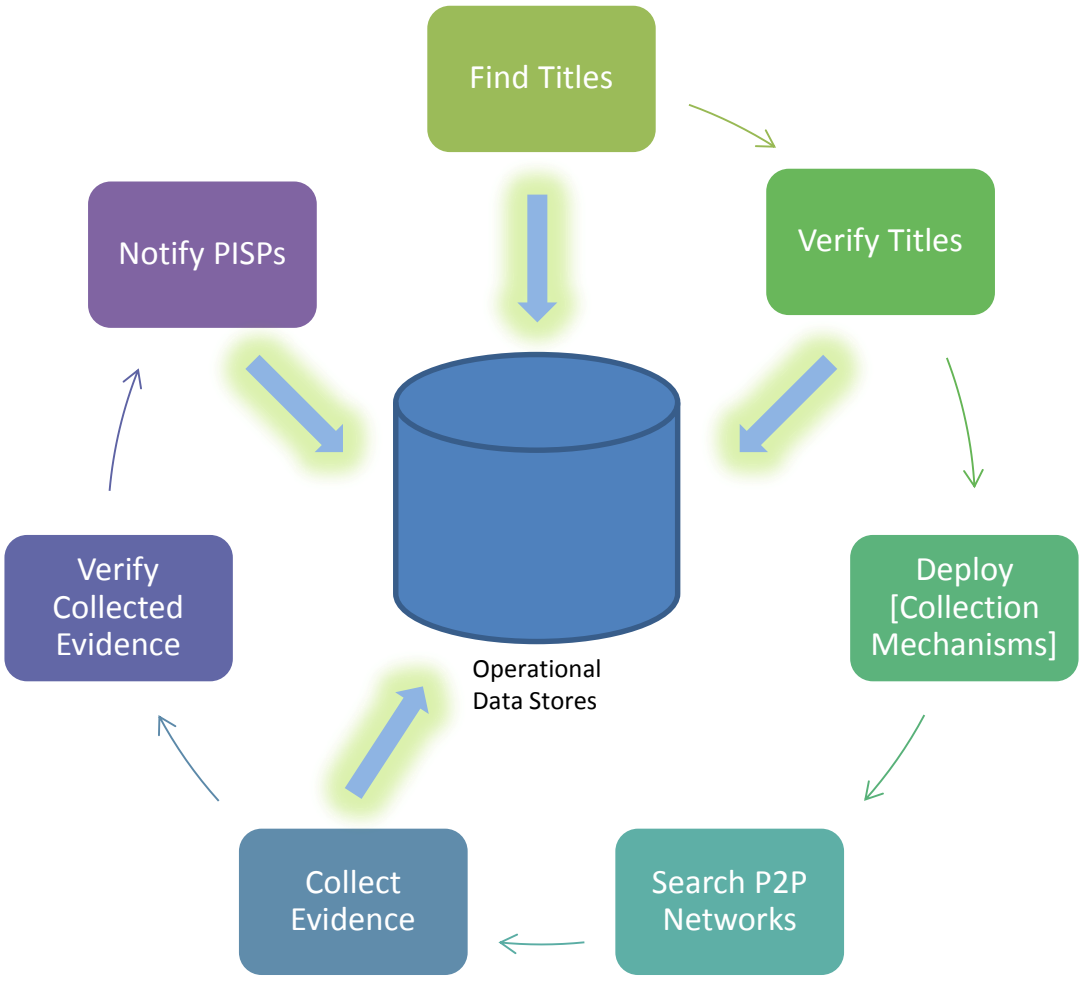

Figure 1: MarkMonitor AntiPiracy Workflow

To identify infringing works, MarkMonitor personnel search for [potentially offending files] and add the results to a [database] that captures relevant metadata about each file, including its name, hash values, and size. This identified content is next reviewed manually or with automated fingerprinting technology to determine if it is an actual infringing copy of the protected work. Once the work has been reviewed, its status is updated in a [database] to indicate that it has been confirmed as an actual infringing work.

Concurrently, identified infringing searches and torrents are deployed to [collection mechanisms]. The [collection mechanism] is a custom-built software application that runs on servers deployed in datacenters geographically spread [REDACTED]. MarkMonitor has designed the [collection mechanisms] to specifically target [REDACTED]. The [collection mechanisms] search for, download portions of, and create evidence packages or "cases" of infringing works including (among other data points) IP address, port, time/date, size, PeerID, and hash values.

[REDACTED] [S]cripts run to verify that the collected evidence is in fact a verified infringing work and meets all program requirements.

After the collection has been verified, in the final step, infringement notices are generated and sent to the appropriate ISP based on specifications specific to each provider.

2. Identification of Infringing Files and File Verification

Once a month the Content Owner Representatives provide MarkMonitor lists of titles of copyrighted works they would like monitored. MarkMonitor identifies infringing online versions of these titles through [comprehensive scanning across multiple sites].

Each new instance of an identified work is fully downloaded and reviewed by MarkMonitor personnel to verify that it is in fact an actual infringing copy of the title. The purpose of this step is to verify that the file being targeted is the asset intended for monitoring by the content owners. This verification step is crucial to ensure that the content owners are not enforcing antipiracy measures on non-protected works and is a key part of MarkMonitor's ongoing success in the proper identification of P2P copyright infringement.

In order to conduct the verification, for the MPAA, MarkMonitor analysts are required to:

- Manually review at least [REDACTED] minutes of a video file at the beginning, middle and end of the file, including reviewing all titles and credits; and
- Manually review at least [REDACTED] full minutes of content from each [REDACTED] minutes of the remaining content of the file.

For the RIAA, MarkMonitor analysts are required to:

- Use industry standard audio fingerprinting technology provided by [REDACTED] to identify the first instance of an identified file, and match each subsequent version of the same file by hash value. The [REDACTED] settings require a "Type 3" match.[1] With this setting, the entire suspected infringing file is fingerprinted, and that fingerprint is sent to a lookup server for identification. An arbitrary [REDACTED] segment is used to initially match the suspected infringing file against the reference file (i.e. reference sound recording). Once initially identified in this manner, the remainder of the suspected infringing file is matched to the reference file to confirm the match across all or substantially all of the sound recording.

[1] For the Type 3, the media sample up to [REDACTED] minutes is fingerprinted is sent to a lookup server for identification.

The status of each reviewed instance of each file and its hash value is recorded in [a database]. The database will reflect the status of each reviewed file, indicating whether it is incomplete, fake, or a verified infringing work. Figure 2 below [REDACTED]:

Figure 2: [REDACTED]

Searches for new versions of an in-scope title are conducted and added to the database continuously throughout the duration of a project/program. However, only the first unique instance of each identified file is subjected to the full review process; each subsequent version can be verified as identical through SHA1 cryptographic hash matching, which confirms that the file is identical to the previously reviewed version. Though all identified versions of a file are deployed to the P2P data [collection mechanisms], as described in the next section, only verified infringing works are eligible to generate infringement notices.

3. Search for and Identification of Infringing files with P2P [Data Collection Mechanisms]

For purposes of its work pursuant to the MOU, MarkMonitor uses a distributed network of servers [REDACTED] all with custom designed data scanning and [collection mechanism] software. This software, which consists of [REDACTED], is responsible for collecting information from P2P networks and websites. The [collection mechanisms] search process and functionality vary by P2P network, but are all designed to integrate and function on the P2P network as a standard peer and communicate and download data from other peers, just like a standard P2P client.

Unlike a standard P2P client, however, the P2P [collection mechanisms] are designed to document users' activities and generate evidentiary collections of shared content with relevant supporting information. Also unlike standard P2P clients, the [collection mechanisms] request and download only a portion of a shared file from each peer, typically around [REDACTED] kilobytes. These individual pieces are verified by SHA1 cryptographic hash values to be part of the original targeted work and after the content is confirmed to part of the original targeted work, the download is stopped. Downloading only a limited portion of the file allows the [collection mechanisms] to minimize required storage space and target as many peers sharing a particular work as possible. The figure below contains a screenshot of a P2P [collection mechanism]:

Figure 3: [P2P Collection Mechanism] [REDACTED]

[REDACTED] Figures 4, 5, and 6 below contain screenshots of [MarkMonitor proprietary tools]:

Figure 4: [REDACTED]

Figure 5: [REDACTED]

Figure 6: [REDACTED]

4. Infringement File Collection and Verification

When a P2P [collection mechanism] connects with a peer and identifies an in-scope file, [REDACTED], it attempts to download a piece of that file and generate an infringement "case" package.

Stroz Friedberg found that the MarkMonitor case file generation follows practices consistent with industry standard forensic data collection, including generation of audit trails, robust documentation, hash verification, and repeatability. As described below, this infringement case package is an exhaustive and defensible collection of evidence about the infringement.

In addition to collecting actual portions of the infringing work, each infringement case contains a series of XML-formatted log files with case specific information, a packet capture file ("PCAP"), and a presentation layer that renders the case information cleanly in a web browser. Timestamps included in the log files are synchronized with Internet time servers and are configured to Coordinated Universal Time (UTC). The evidence file includes the following key information:

- IP Address
- Country
- Capture Initiated time
- Capture Completed time
- P2P Protocol
- Target ISP
- Target Port
- Target Hostname
- Server ISP
- Server IP
- Server Port
- Server Hostname
- SHA1 hash value
- Connection type
- Peer Client Info
- Peer ID

Among the XML files in the case file are an activity log and a communication log. These files provide a timeline for the entire collection process and detail the P2P protocol related communications between the [collection mechanism] and end user's peer sharing or downloading the infringing content.

Another of the XML files, [REDACTED], includes the name, SHA1 hash value, and size of the target file, as well as how much of the infringing work was shared by that particular user, and the portion downloaded and hash verified by the [collection mechanism]. The

figure below contains an example of the [REDACTED] file as displayed through the HTML presentation layer:

Figure 7: Content Info from Case File [REDACTED]

The evidence package also includes a traceroute of the user's IP address in connection with a reverse DNS lookup of all devices encountered, to determine the IP address, ISP, and geographic location of the peer.

MarkMonitor relies on the industry standard WinPcap library to preserve network packets exchanged between the [collection mechanism] and the target peer. The packet captures preserve and allow for review of the entire communication session between the [collection mechanism] and the investigation subject.

[A] second level of verification [] is conducted at this point in the process. This [second level of verification] consists of a check to ensure that the individual file found shared by a user (and captured in part in the case file) is a true copy of the asset being monitored. Stroz Friedberg confirmed that for each instance of infringement, the MarkMonitor system verifies that:

- The IP address is sharing a file [REDACTED];
- The IP address is verified as active [REDACTED];
- The IP address is detected as having a P2P client; and
- The hash reported by the peer matches a hash value in the database of confirmed infringing files.

[REDACTED]

Stroz Friedberg finds that these [second level verification] steps are adequate and sufficient to confirm that files identified by the [P2P collection mechanism] are in fact copies of known infringing works.

5. P2P Online Infringement Enforcement

MarkMonitor's enforcement-specific Methodologies consist of identifying the ISP associated with an infringement case, generating infringement notices, sending a notice to the ISP, and (at least in the case of RIAA-related works) logging notice responses received back from the ISPs and/or customer. After a [P2P collection mechanism] has identified and downloaded a shared file, generated a case file, and that work has been verified as infringing (through both the File Verification and Infringement Verification process), notices are programmatically generated and sent by email.

Stroz Friedberg took steps to ensure that MarkMonitor's approach included safeguards to: (1) only send notices for verified infringement cases that have been captured within the past 48 hours; (2) ensure notices are not duplicative of notices sent to the same IP address within the past 24 hours; (3) create notices using the proper template and

format; (4) send notices to the proper ISP and correct contact address using the appropriate method; (5) make sure notices are delivered properly; and (6) for RIAA works, document responses from ISPs and users. A mistake or breakdown in any of these areas could potentially result in a failure to properly report on or track infringement cases. The specific steps MarkMonitor takes to address each of these six areas is discussed below:

- To ensure that notices are only sent on verified cases, MarkMonitor relies on scripts that generate notices only on cases that have been verified within the [REDACTED] database.

- To ensure that notice timeframe limitation rules are met, settings within the [MarkMonitor's proprietary application] allow MarkMonitor to specify the frequency of collection for a particular identified IP address.

- To identify the appropriate ISP for each infringement case, MarkMonitor maintains its own database of IP-block to ISP correlation. To ensure that this database is current, MarkMonitor uses the IP addresses identified in the reverse DNS lookups performed in the collection stage and compares them [REDACTED]. MarkMonitor also proactively maintains and updates a list of ISP contact information as well as a report on each ISP to ensure its notices are directed to the appropriate location.

- To make sure its notices are delivered properly, MarkMonitor runs and reviews a weekly report on the SMTP server used to send notices. Any address message blocking errors are addressed on a timely basis.

- [REDACTED]

Stroz Friedberg found these steps to be adequate and sufficient to ensure the proper creation and delivery of infringement notices.

6. AntiPiracy System Redundancies

The MarkMonitor Methodologies have a number of inherent and added system redundancies designed to ensure that it can provide continuous consistent scanning.

[REDACTED]

Controlled System Test of MarkMonitor Platform

Stroz Friedberg worked with MarkMonitor to conduct a controlled test of its antipiracy Methodologies and to further understand its workflow and evaluate the efficacy of its processes. As described below, this test further established the effectiveness of the MarkMonitor's Methodologies.

Our test environment consisted of a VMWare virtual machine running Windows XP and the BitTorrent client application, version 7.7.0. This virtual computer was connected to a DSL line provided by one of the Participating ISPs, AT&T. Our test machine connected to the Internet with IP address [REDACTED].

To test the operation of the Methodologies, Stroz Friedberg downloaded and became a seeder of a video file called, "House season 4.avi".[2] Figure 8 below contains a screenshot of the BitTorrent client running on Stroz Friedberg's test computer.

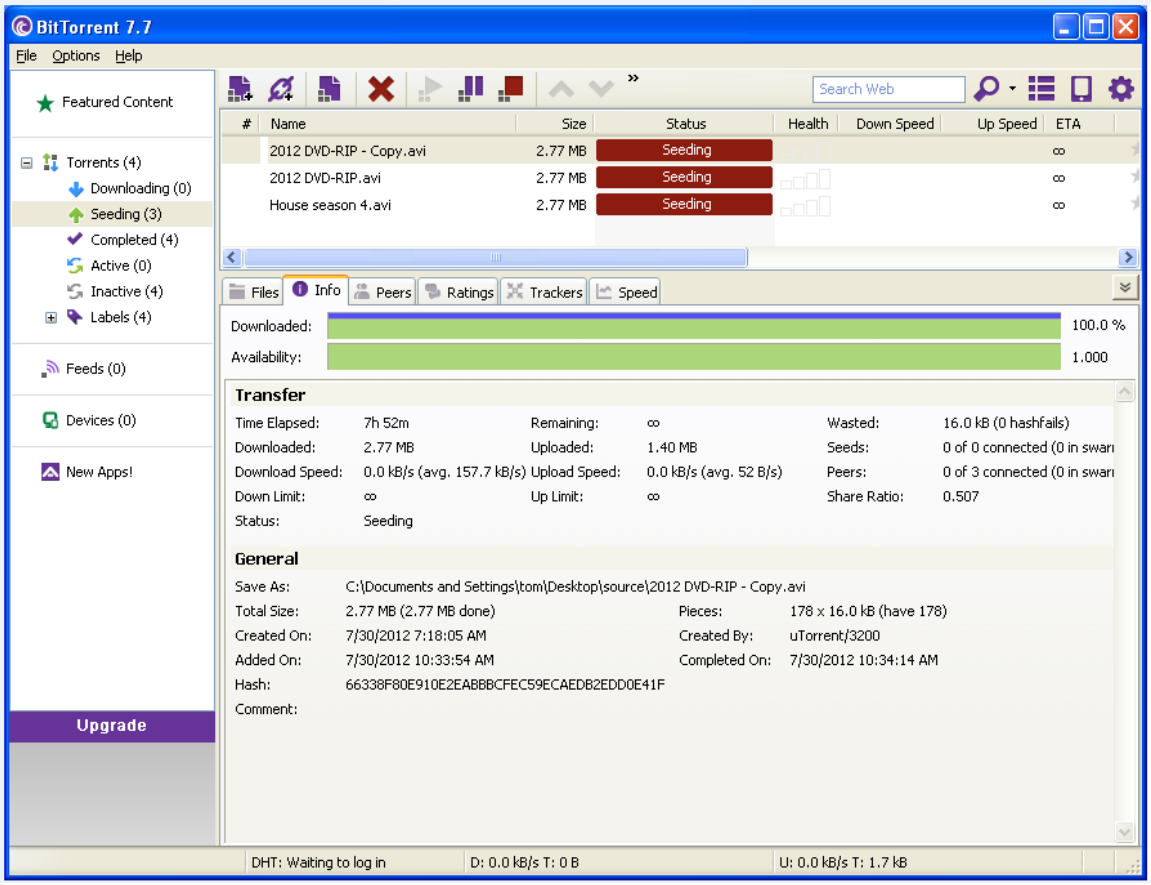

Figure 8: BitTorrent running on Stroz Friedberg test computer

The "House season 4.avi" content file was one that MarkMonitor had identified as a file to search for by the [collection mechanisms]. In our live test, the MarkMonitor [collection mechanism] identified and began collecting data about our test system almost immediately after our BitTorrent client began seeding the file.

[REDACTED]

Figure 9: [REDACTED]

[2] The "House season 4.avi" was not, in reality, a pirated copy of the FOX television series.

Our inspection of [the generated] case file confirmed that the XML files contained within maintained the key information required to document our test computer's download and sharing of the "House season 4.avi" file, as well as information necessary to authenticate the collection. The evidence collection package is included as **Exhibit 1** to this report.

MarkMonitor also took steps to generate a sample notice based on this test case using the template in place for a current MPAA program. And though this manual notice generation did not fully duplicate the automated methods used to create notices in the MarkMonitor production environment, it demonstrates the workings of the overall process. The notice email is included as **Exhibit 2** to this report.

Enhancement Recommendations

Stroz Friedberg offers the following recommendations focused on improving the accuracy and reliability of MarkMonitor's already robust processes to ensure P2P infringers are properly identified in a defensible and well documented manner. These recommendations are designed to further mitigate any potential break points and lower the probability of exposure as a result of inaccurate identification.

These recommendations have been communicated to MarkMonitor, and we understand that many have been implemented or are being considered for implementation in the future.

1. **Augment the File Identification Audit Trail**

[REDACTED]

2. **Deploy a Higher Percentage of "Verified" Infringing Works to the [Collection Mechanisms]**

[REDACTED]

3. **File Identification of Obfuscated Infringing Works**

[REDACTED]

4. **Create and Maintain a Hash Value with the Case File**

[REDACTED]

5. **Notice Generation and Sending Should Move Away from Email/SMTP**

[REDACTED]

6. **Develop a Periodic Audit Framework**

[REDACTED]

Recommendations for Ongoing Periodic Review

The MOU establishes that the Independent Expert shall evaluate the Content Owner Representatives' Methodologies on a periodic basis. Stroz Friedberg proposes to accomplish this through a quarterly review of MarkMonitor's Methodologies. [REDACTED]. The substance of the periodic review should include at least the following events:

- A briefing on protocols that have been changed or revised.
- Discussion of any P2P platform, network changes, or file sharing trends that may implicate the MarkMonitor Methodologies.
- A review of metrics related to notice generation sending and responses and trends compared to the prior quarter.
- A limited audit of a statistically relevant sample of reviewed/verified titles.
- A limited review of a statistically relevant sample of notices generated to ensure they relate to verified infringing works.
- Review a sample of collected cases to ensure they contain all required components.

BOSTON, MA

CHICAGO, IL

DALLAS, TX

LOS ANGELES, CA

MINNEAPOLIS, MN

NEW YORK, NY

SAN FRANCISCO, CA

SEATTLE, WA

WASHINGTON, DC

LONDON, UK

STROZ FRIEDBERG

www.strozfriedberg.com

© 2012 Stroz Friedberg. All rights reserved.