

Automated Takedown System for Peer-to-Peer File Sharing Networks

Submitted by:
Universal Music Group
Vivendi Universal Entertainment

In Response to the:
Joint Committee of the Higher Education and Entertainment Communities Technology Task Force
Request for Information
Technology Opportunities for Addressing Issues Associated with Peer-to-Peer File Sharing on the University and
College Campus

May 29, 2003

CONTACT INFORMATION:

Joseph Cates
Universal Music Group
2220 Colorado Avenue
Santa Monica, CA 90404
TEL: 310-865-8495
jcates@umusic.com

Aaron Markham
Vivendi Universal Entertainment
100 Universal City Plaza
University City, CA 91608
TEL: 818-777-3111
antipiracy@unistudios.com

TABLE OF CONTENTS:

1	Introduction.....	4
2	System Overview.....	4
3	System Elements.....	5
3.1	Electronic Notice.....	5
3.2	Control Station.....	5
3.3	Blocking Rule.....	5
3.4	Traffic Controller.....	5
3.5	Approval Point.....	5
3.6	Source Node.....	5
4	System Process.....	5
5	Source Identification.....	5
6	Response to Encrypted and Disguised Traffic.....	5
7	Requirements Development.....	5
8	Conclusion.....	5
	APPENDIX A: Response to Specific RFI Sections.....	5
A.1	Network Architecture.....	5
A.2	Scalability.....	5
A.3	Protocol identification.....	5
A.4	Granularity of protocols.....	5
A.5	Content identification.....	5
A.6	Examination of network packets or file content.....	5
A.7	Distribution systems.....	5
A.8	Resilience of the technology to countermeasures.....	5
A.9	Testing and installed base.....	5
A.10	Competitive approaches.....	5
A.11	Third-party components.....	5
A.12	Intellectual Property.....	5
A.13	Corporate Characteristics and Resources.....	5
A.14	Pilot Testing.....	5
A.15	Commercial Terms.....	5
A.16	Conflicts of Interest.....	5

1 Introduction

The Automated Takedown System (ATS) is designed to address illegitimate content distribution via peer-to-peer (P2P) file sharing systems on university networks. The system allows automation of the university's response to copyright takedown notices to block illegitimate distribution quickly and inexpensively, while being flexible enough to support each university's unique network usage policies. The ATS extends existing products and services to allow for rapid development and deployment, and the vast majority of needed features are already available in existing products. At the core of the system is the traffic controller which allows specific types of traffic to be limited or blocked for particular IP addresses. The system is resistant to encryption of P2P traffic and other countermeasures, and can be scaled to large institutions.

The preliminary goal of the ATS proposal is to establish a working group with universities to establish system requirements. Common requirements would allow existing vendors of traffic shapers and bandwidth management tools to extend their products to meet university needs. In addition, monitoring companies and content owners would develop a standardized electronic format for communication of notices.

Note: ATS is designed to integrate and extend existing technologies, and to be applicable to the widest possible range of vendors. This proposal is not intended to recommend or endorse any particular product, and the reader should refer to individual vendors for specific product details.

2 System Overview

Content companies regularly scan public P2P networks to locate P2P nodes illegitimately distributing copyrighted material. When an illegitimate source is located, a content owner may elect to send a copyright takedown notice to the individual university or ISP responsible for the identified IP address. Each university responds to the takedown notices according to the university's policy. Procedures vary widely, but the typical response includes identifying the user corresponding to the IP address, limiting or disabling access, notifying the user, initiating disciplinary action, and sending a compliance response. Almost unanimously, responding to takedown notices is a manual process. At some universities, the response even entails sending IT staff out to unplug cables in dormitories. The recent minor increase in the percentage of infractions for which content owners sent notices has led to a major increase in university workload. Universities have been forced to hire additional staff, and several have instituted fees to the user for the takedown response.

At the same time, a large number of universities are deploying bandwidth management tools such as traffic shapers to better manage their network and reduce bandwidth costs. Traffic shapers allow universities to identify the type of network

traffic and allocate bandwidth accordingly -- critical applications can be assigned the highest priority and allocated guaranteed bandwidth while less important applications can be assigned lower priority with fixed bandwidth limits. In addition, traffic shapers provide a broad range of rules which can control network traffic with great precision on an IP address basis.

The ATS seeks to extend the functionality of traffic shapers to allow universities to automate the response to copyright takedown notices. When a notice arrives, the ATS would process the notice and automatically enforce the university policy by limiting or blocking network traffic. Each university could choose the appropriate rule for their policy, but an example might be to block all outgoing P2P traffic from the source for a specified period of time.

Working with universities, a common set of requirements for ATS can be developed, which would allow existing traffic shaper vendors to extend their products and develop the tools needed by universities.

An automated system would greatly reduce the cost of response to takedown notices. The university's own network and acceptable use policies would be more efficiently enforced, and the system permits a more flexible, limited response than completely disconnecting a user. If content owners continue to increase the quantity of takedown notices, universities will not be overburdened. An automated system also allows for faster response, which helps limit further illegitimate distribution.

3 System Elements

The system elements are shown in Figure 1 and described in detail below:

3.1 Electronic Notice

When a content owner locates a source illegitimately distributing content, an Electronic Notice is sent which contains information in XML format specifying the details of the Source Node and the content being illegitimately distributed.

3.2 Control Station

The Control Station receives Electronic Notices, generates a Blocking Rule according to local university policy, and transmits the Blocking Rule to the Traffic Controller. In addition, the Control Station can notify the Source Node and the university administrators for appropriate follow-up actions.

3.3 Blocking Rule

The Blocking Rule specifies instructions for the Traffic Controller to prevent further illegitimate distribution from the Source Node. Each university will select blocking rules based on the university's policy.

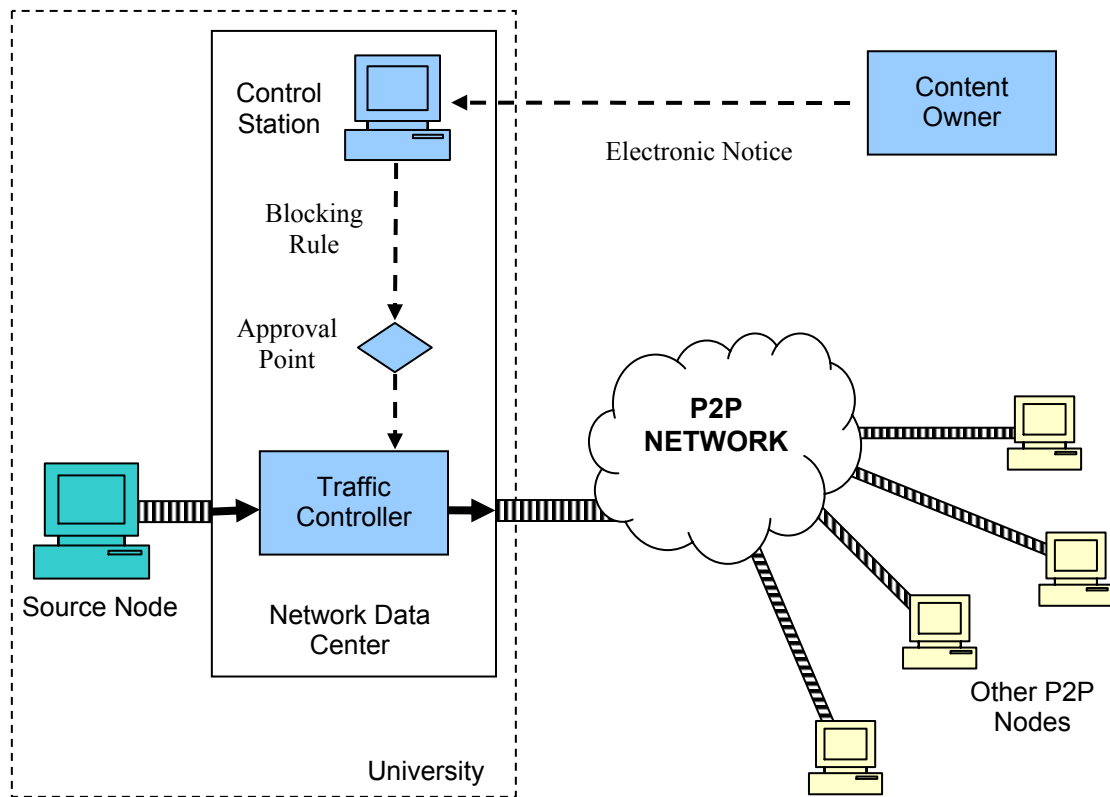


Figure 1: System Elements

Examples of possible rules for a Source Node illegitimately distributing content include:

- a) Block all network traffic from source except to university servers until administrative action
- b) Block all P2P traffic from source for specified time
- c) Block all *outgoing* P2P traffic from source for specified time
- d) Assign lower priority and bandwidth limits to all P2P traffic from source for specified time

Traffic shapers support a wide range of flexible rules for management of network traffic. Traffic can be categorized by source, destination, direction (incoming or outgoing), type of traffic, and other properties. Traffic can be limited to specific bandwidth or blocked altogether. Each of the rules can be implemented for a specific time periods.

3.4 Traffic Controller

The Traffic Controller enforces the Blocking Rule. Typically the Traffic Controller will consist of a traffic shaper, which identifies the type of network traffic based on packet information, and limits or blocks specific traffic as appropriate.

3.5 Approval Point

The Approval Point allows the option of manually approving each Blocking Rule before transmission to the Traffic Shaper.

3.6 Source Node

The Source Node is the P2P user identified by the Electronic Notice as illegitimately distributing content. The Source Node is typically identified by IP address, although in some cases a MAC address may also be used.

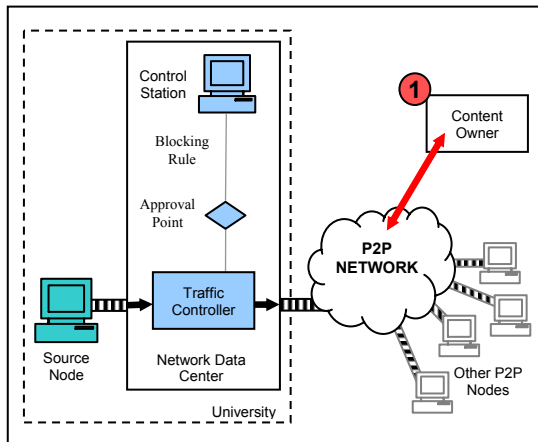
4 System Process

The system process is as follows:

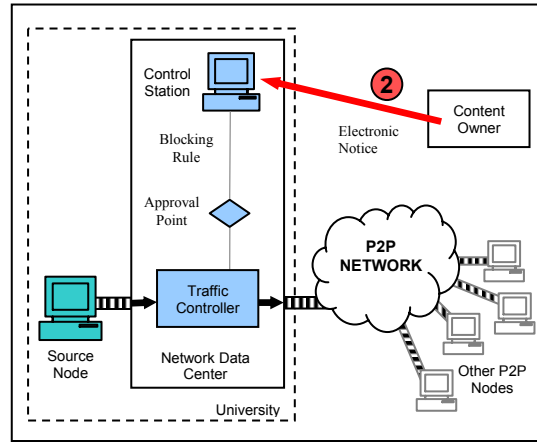
- i) A content owner connects to a P2P network and locates a Source Node which is illegitimately distributing content. The content owner may identify content based on a combination of file properties such as metadata, file size, P2P client file signatures (i.e., hashes), or potentially by downloading and identifying manually, with fingerprints, or with watermark detection. The Source Node is identified by IP address and other available information such as port.
- ii) From the IP address, the appropriate Control Station for the Source Node is determined and an Electronic Notice is dispatched.
- iii) The Control Station receives the Electronic Notice and determines the appropriate Blocking Rule according to the local university policy. The rule may depend on factors such as the type of content, amount of content, P2P protocol, and number of prior notices.
- iv) Optionally, the Control Station includes an Approval Point for administrators to review and approve the Blocking Rule.
- v) The Control Station updates the Traffic Controller to enforce the Blocking Rule.
- vi) The Traffic Controller limits or blocks network traffic to insure the rule is enforced for the Source Node.
- vii) Other P2P nodes within the university network, which may or may not be illegitimately distributing content, are not affected.

5 Source Identification

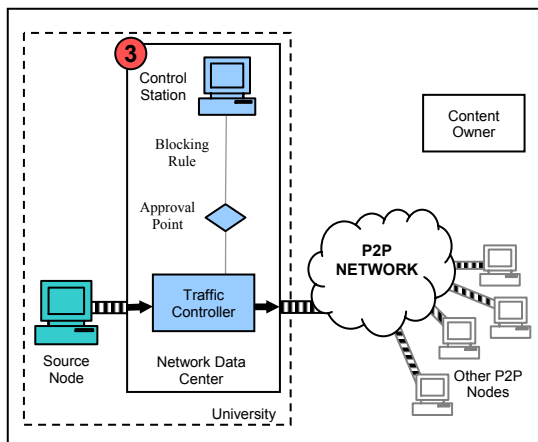
Typically the Electronic Notice will identify the Source Node by IP address. However, in many networks an IP address may not be a reliable means of identifying a source due to dynamic assigning of IP addresses (such as DHCP) and devices which perform network address translation (NATs).



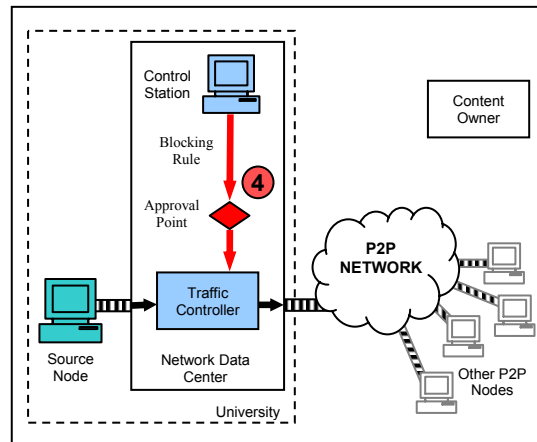
Step 1: A content owner connects to P2P network and locates a Source Node which is illegitimately distributing content.



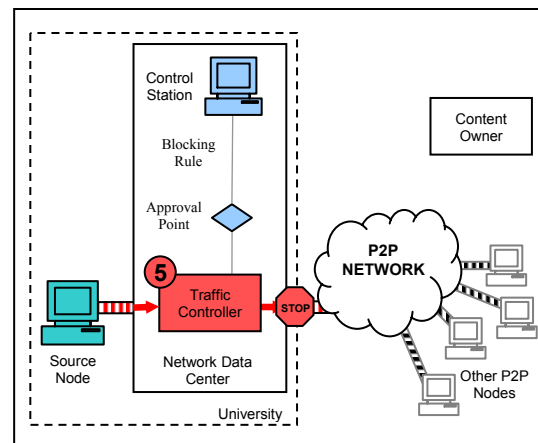
Step 2: The content owner determines the appropriate Control Station and dispatches an Electronic Notice.



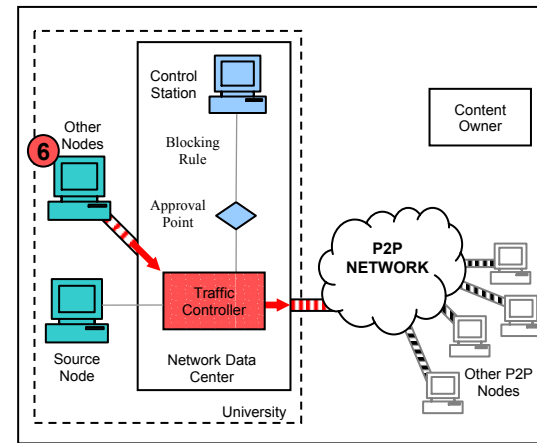
Step 3: The Control Station receives Electronic Notice and determines the appropriate Blocking Rule according to local university policy.



Step 4: The Control Station updates the Traffic Controller to enforce Blocking Rule. Optional Approval Point allows administrative approval.



Step 5: The Traffic Controller routes or blocks network traffic to enforce rule for Source Node.



Step 6: Other P2P nodes, which may or may not be infringing, are not affected.

Figure 2: System Process

Some existing traffic shapers are able to identify users by MAC address by communicating with DHCP servers and tracking the assignment of IP addresses. In addition, some existing traffic shapers are able to integrate with authentication systems such as RADIUS, so that a Blocking Rule can be linked to an account irrespective of IP address.

If these capabilities are not available, the Blocking Rules will have to be carefully selected according to the local network architecture and policies. For DHCP, Blocking Rules should balance any time periods for which a source is constrained against the likelihood of the IP address being reassigned. How often IP addresses are reassigned depends on a variety of factors unique to each university such as size of address pool, lease time, and turnover rate. For networks using network address translation (NAT), the IP address will typically only resolve to the NAT IP address. The Blocking Rules should take into account whether the NATs are allowed by university policy, whether all the devices behind the NAT are operated by the same user, and whether limiting traffic for all devices behind the NAT is appropriate. In addition, the traffic controller may also be able to utilize source and destination port information to identify the specific device behind a NAT.

The ATS requirements need to address the ability of the system to dynamically adapt to the local network, and specify capabilities to integrate with systems such as DHCP, RADIUS, and NATs.

6 Response to Encrypted and Disguised Traffic

The ATS requires the Traffic Controller to identify only the *type* of network traffic. Currently type identification for P2P traffic is straightforward. In the future, P2P networks may become stealthier, by encrypting traffic and disguising P2P traffic as other protocols such as HTTP or e-mail. However in most cases, type identification is still possible even for encrypted and disguised traffic. For example, several existing traffic shapers can distinguish fully encrypted Freenet traffic from other encrypted network traffic such as SSL.

Eventually, new protocols such as IPv6 (Internet Protocol Version 6) will allow all traffic to be encrypted at the transport layer. Encryption of traffic raises issues for the Internet as a whole; however, it is reasonable to assume that at some point even type identification may become impossible. If this occurs, the Blocking Rules will have to be formulated to constrain all traffic whose type cannot be determined. However traffic can still be distinguished by destination, so access to mail servers, university web sites, specific external web sites, etc. can remain unrestricted even for blocked nodes.

7 Requirements Development

The majority of the functionality required for ATS is already available via existing traffic shapers and monitoring services. However, additional work is needed to develop system requirements in several areas:

University Requirements:

- a) What support, if any, is needed for DHCP and other methods of dynamically assigning IP addresses?
- b) What support is needed for authentication methods such as RADIUS?
- c) What support is needed for messaging functions to communicate with users and administrators?
- d) What types of Blocking Rules are useful to universities?
- e) What information should be sent in Electronic Notices to identify Source Nodes?
- f) What information should be used to reply to the content owner that sent the initial notice?

Technical Work:

- a) Develop standardized message format for Electronic Notices.
- b) Define method for secure communication of Electronic Notices to Control Stations.

8 Conclusion

By extending existing traffic shaper products, ATS can be rapidly developed and deployed to address illegitimate P2P content distribution via university networks. The system allows universities to automate the response to copyright takedown notices to block illegitimate distribution quickly and inexpensively. The system is flexible enough to support the unique network policy of each university.

To proceed, ATS proposes to create a working group with universities to establish the requirements for the system. The requirements would allow existing vendors of traffic shapers and bandwidth management tools to extend their products to meet university needs. The result would be a wide range of products from which each university could choose.

In addition, trials of the ATS system are being established with traffic shaper vendors for summer 2003 which would include both university and ISP installations. The proposal seeks the assistance of the committee and interested universities in arranging other trials to allow the largest possible level of participation.

APPENDIX A: Response to Specific RFI Sections

A.1 Network Architecture

5.1.1: The vendor should provide descriptions of how its technology would be installed in typical networks including architecture diagrams.

The Traffic Controller is installed as normal for traffic shaper. The ATS proposal does not recommend any particular product; the reader should refer to the individual traffic shaper products for architecture and installation details.

The Control Station is installed within the university data center where it will be able to communicate with the university's Traffic Controllers and receive electronic notices from outside the university firewall.

A.2 Scalability

5.1.2: Due to the high capacity of many university networks, particular reference should be made to the scalability of the tool. Each vendor should clearly indicate the maximum throughput of its technology (using packet size or average packet size and number of packets per second.), and the effect of its technology on network latency.

The ATS architecture is designed to add additional functionality to existing traffic shapers, and introduces no additional performance penalties beyond normal operation of the traffic shaper. The ATS proposal does not recommend any particular product; the reader should refer to the individual traffic shaper products for performance details.

A.3 Protocol identification

5.1.3: The vendors should discuss how network protocols are identified (e.g. by port number, by layer 7 analysis, etc.) and the methods that the vendor uses to update its software to detect new and mutated protocols.

The Traffic Controller is required to identify the type of traffic. Although some protocols can be recognized by port number, in general, the Traffic Controller must inspect TCP packet data to distinguish protocols. Each Traffic Controller will use proprietary technologies to identify protocols and maintain and update the detection software. The ATS proposal does not recommend any particular product; the reader should refer to the individual traffic shaper products for performance details.

A.4 Granularity of protocols

5.1.4: Each vendor should discuss its technology (where applicable) in terms of addressing those P2P applications that employ multiple protocols (e.g. control, searching, file transfer, etc.). Descriptions should be provided as to: which protocols does the vendor's technology detect; can

the technology address each of these protocols independently; can different rate limits be set for “search” vs. “file download”.

The Traffic Controller is required to identify multiple protocols only to the extent that Blocking Rules distinguish between P2P operations. The ATS proposal does not recommend any particular product; the reader should refer to the individual traffic shaper products for performance details.

A.5 Content identification

5.1.5: If the technology is intended to selectively identify communications that represent potential infringements of copyright law, explain in detail how the technology will identify such content. The vendor should specify what categories of content can be identified, for example, compressed audio files, video, images, etc. The vendor should indicate any external content databases that are required.

The content owner joins P2P networks using the P2P client software and detects nodes distributing content (“sharing” or “uploading”). Other operations, such as searching or downloading, would not be identified.

A significant advantage of the ATS is that content identification is not constrained to real-time identification during transmission. Content can be identified based on file properties such as metadata, file size, and P2P client file signatures (i.e., hashes). Additionally, content can be downloaded and identified with manual inspection, fingerprinting, or watermark detection.

ATS would be able to identify any type of content desired. Existing monitoring companies regularly monitor for audio, video, and software content.

A.6 Examination of network packets or file content

5.1.6: Each vendor must specifically and clearly indicate any aspects of the use of its technology that requires the examination or “opening” of network packets or files of information in order to carry out the technology's work. The vendor should clearly indicate which fields of the Internet packets and which components of corresponding files are used by the technology and in what way. The vendor must include not only the degree of content inspection that is required in order to make the technology work, but also the degree of inspection and logging that the technology is capable of performing (even if “turned off” by the user or system administrator).

The Traffic Controller is required to examine network packets in order to identify and block traffic by type. The ATS is not necessarily required to “open” packets, depending on the complexity of the Blocking Rules implemented. In general, the more specific the blocking rule is, the greater the privacy implications. Each university deploying ATS must decide the appropriate balance between privacy and blocking for their application.

For example, consider the following three examples of Blocking Rules for an IP address:

- i) All Internet Access Blocked: In this case, the Traffic Controller only needs to inspect the IP header to identify the source and destination IP address.
- ii) Specific Port or Port Range Blocked: For some applications, blocking only a specific range of ports is sufficient. In this case, the Traffic Controller only needs to inspect the TCP header to identify the source and destination port.
- iii) Specific Application Blocked: For this more restrictive case, the Traffic Controller must open the TCP data. The number of bytes of TCP data which must be inspected and any potential requirement to maintain state are dependent on the Traffic Controller and the specific protocol.

Each traffic shaper uses proprietary technologies to identify protocols based on network data, and provides protections against unauthorized inspection or logging of network traffic. The ATS proposal does not recommend any particular product; the reader should refer to the individual traffic shaper products for performance details.

The content owner connects to public P2P systems to locate sources illegitimately distributing content by joining the P2P network using the P2P client software. The monitoring process does not involve “opening” network packets or monitoring communication.

A.7 Distribution systems

5.1.7: The response should specifically list all P2P networks for which the technology is effective. In addition, the vendor should list other distribution systems, such as FTP, IRC, Usenet, for which the technology is applicable.

Current traffic shapers recognize all the major distribution protocols including P2P networks, NNTP (Usenet), FTP, IRC, and HTTP. ATS could be applied to copyright notices for illegitimate sources found on any of these protocols.

A.8 Resilience of the technology to countermeasures

5.1.8: Each response should indicate the technology’s ability to resist:(i) Countermeasures by the P2P software, for example, file compression, data encryption, etc; (ii) Circumvention efforts by P2P users, for example, port tunneling, proxy servers, fragmented packets, etc.; (iii) Denial-of-service or other attacks against components of the technology.

The fundamental premise of P2P distribution is that users can connect anonymously to retrieve content. A content owner will always be able to connect to the network, instigate a download, and record the source of the data.

The ATS requires only identification of the type of network traffic and is highly resistant to countermeasures by P2P software, see Section 6.

Traffic shapers offer a variety of defenses against circumvention attempts and countermeasures. The ATS proposal does not recommend any particular product; the reader should refer to the individual traffic shaper products for performance details.

A.9 Testing and installed base

5.1.9: Each vendor should provide descriptions of the extent of testing of its technology, particularly the testing procedures employed prior to releasing an updated version (e.g. an updated detector) and specific examples (where available) of the application of the technology to real-life situations. The vendor should describe the maturity, status, availability, and installed base of its technology.

Existing traffic shaper products are widely deployed in universities for bandwidth management. The ATS proposal does not recommend any particular product; the reader should refer to the individual traffic shaper products for further testing and deployment details.

Universal Music and Universal Studios are in the process of establishing trials of the ATS system with traffic shaper vendors for summer 2003. The trials are anticipated to include both university and ISP installations.

A.10 Competitive approaches

5.1.10: Each vendor should provide clear descriptions of how its technology compares to other competitive approaches and the benefits of its technology over competitive approaches.

ATS offers a number of advantages over competitive approaches.

- All monitoring is performed outside the university by a monitoring service which joins public P2P networks. Remote monitoring avoids the privacy and performance issues of attempting to identify content “in-the-wire” during transmission. In addition, remote monitoring makes the system much more resistant to attempts to disguise content.
- All action is based on identification of sources illegitimately *distributing* content on public P2P network. This avoids any issues where a particular university might not consider objectionable the possession of content, downloading of content, or distribution within an academic context to a private group. By taking specific action against illegitimate sources, general bandwidth limits for all users are not required.
- All rules are customizable to support local university network policy and procedures.
- The system is based on the proven traffic shaper technology available from a variety of vendors, and already deployed at many universities.

A.11 Third-party components

5.1.11: Each response should describe any third-party components required by the technology which are not provided by the vendor, for example, content databases, etc.

The current proposal seeks to create an open system of requirements based on which multiple vendors can build compliant systems. As such, the ATS system would be dependent on individual technology companies providing components.

A.12 Intellectual Property

5.2: This section of the vendor's response to the RFI should provide information concerning intellectual property associated with the technology. The information should provide insight as to whether any patents have been applied for, when and in what territories, and what (if any) claims have been granted. Where a patent has been granted, the vendor should provide the patent number and attach an abstract of the patent.

As of May 29, 2003, Universal Music Group is considering filing a preliminary patent application in the United States for an automated monitoring and controlling system for networked communication, which would incorporate concepts from this proposal.

A.13 Corporate Characteristics and Resources

5.2: This section of the vendor's response to the RFI should provide information about the vendor's corporate capabilities so that the Task Force may evaluate the vendor's maturity, stability and future potential. Include information about the vendor in terms of general and specific corporate characteristics such as size, personnel, facilities, year of organization, corporate history, clients, revenues, capitalization, investors, etc. Information about corporate products or technologies other than that directly applicable to this RFI should also be included.

Not Applicable. The ATS proposal involves creating an open system of requirements based on which multiple vendors can build compliant systems. Each individual vendor would provide information on corporate capabilities.

A.14 Pilot Testing

5.4: It is anticipated that certain colleges or universities may elect to test some of the technologies that are considered to be exemplars of a particular class of tool (as described in the Technology section 5.1 above). The selection of those technologies will be done by the individual campuses based in part upon the results of the knowledge base to be created as a result of this RFI. Many factors will enter into the decision as to which technologies will be tested. One significant factor will be the level of interest and potential involvement of the vendor in the actual testing. Therefore, the vendor should present, in this section of the response, the vendor's concept for testing its technology in a real-life situation on campus.

It is currently anticipated that implementation of technology at such campuses would take place during the months of July and August so as to be in place for the fall 2003 semester. Testing and evaluation of technologies are anticipated to be conducted during the fall semester. The vendor's schedules for implementing its particular technology should be provided, allowing a period of sufficient duration to test and evaluate the results of that implementation.

The vendor should also provide information regarding the costs that might be incurred by a campus during pilot testing and to what extent the vendor would consider conducting the testing on a pro-bono basis.

Universal Music and Universal Studios are in the process of establishing trials with traffic shaper vendors for summer 2003. The trials are anticipated to include both university and ISP installations.

In addition, Universal Music and Universal Studios are working with monitoring companies to finalize the XML format for electronic notices. Universal Studios anticipates adding XML electronic notices to their current copyright takedown notices in the near future.

The proposal seeks the assistance of interested universities in arranging other trials to allow the largest possible level of participation with multiple vendors and universities. Specific costs will be dependent on vendors.

A.15 Commercial Terms

5.5: This section of the vendor's response to the RFI should discuss, in general, information about the commercial terms for its product (e.g. requirements for an annual license, requirements for annual maintenance, ability for one-time purchase). In those cases where the vendor has published prices for the subject technology, the vendor should provide those prices. If standard licenses exist they should be provided as well. In no instance should the vendor provide cost information that would not be considered public.

Not Applicable. The ATS proposal involves creating an open system of requirements based on which multiple vendors can build compliant systems. Individual vendors would each provide pricing and commercial terms for their products.

A.16 Conflicts of Interest

7: The vendor will disclose, within its response to this RFI, any potential or existing conflict of interest that it may have in either responding to this RFI or in the conduct of pilot testing at campuses that elect to participate in such tests. Conflicts of interest should also be noted with respect to any other products or services that may be required in order to deploy the vendor's technology for this project.

An employee of Universal Music Group, and employees of trade associations of which the Universal Music Group and Vivendi Universal Entertainment are members, are entertainment community representatives acting in a liaison capacity to the Joint Committee.