



## Position Paper

# Internet Filtering & Reporting

### Problem

In an environment where access to a secure computing device is governed and controlled by inherently secure protocols, the problem still remains of how access to untrusted environments such as the Web is controlled.

When accessing the web there are three problems that exist;

1. Ensuring that where you browse is in line with the stated (corporate, country, personal or home) policy on web browsing
2. Ensuring that what a web server delivers back is free from malicious content
3. Ensuring that all end-devices, no matter where, or how they are connected, are protected

Existing solutions involve installing filtering solutions in a DMZ which generally cover only those users inside the Intranet. The same level of filtering is rarely available for SME or home users. Where a corporate policy exists for remote user, it involves either leaving mobile users unprotected or insisting that all web access required that the user first initiates an authenticated VPN tunnel back to the corporate environment.

### Why should I care

Browsing the web is a risky pastime. From users that are deliberately lured to web sites by e-mails, to URL's that are deliberately mis-spelt in the hope of luring the unsuspecting browser, to those users who stray to sites that are clearly inappropriate or visit a appropriate site that have been hacked and malicious code inserted.

There is a need to ensure that users are provided with a web browsing experience that both protects them from inadvertently straying to inappropriate<sup>1</sup> sites and ensures that wherever they browse the data returned is free from malicious content.

As end computing devices move into a de-perimeterised world then it is essential that all data feeds have adequate levels on integrity irrespective of their physical location / connection.

### Recommended solution/response

There are two separate problems to be solved, firstly an architecture that allows operation in a de-perimeterised environment, and secondly the provision of a distributed filtering service.

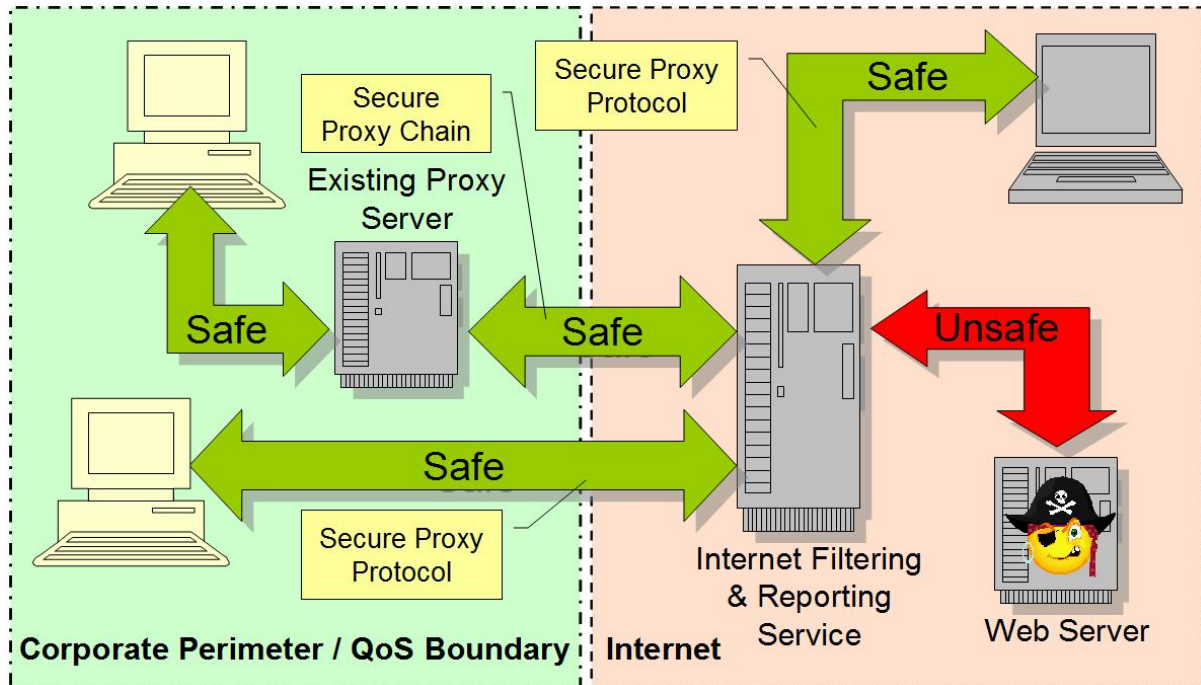
---

<sup>1</sup> An inappropriate site could be defined by corporate policy, local legal restrictions or age restriction

## Background & rationale

### Architecture – A service or internal solution?

In a truly de-perimeterised environment, whether this is provided as a service or as an internal solution should be irrelevant. In the interim, as we move to de-perimeterisation, then this does have relevance and will probably be decided by the company stance on how such services are provided.



For the company that will provide this internally, then this is simply a service that resides in the DMZ (or multiple DMZ's) capable of accepting connections from either roaming devices on the Intranet or internal devices on the Internet.

For the company that prefers to buy this as a service, then existing corporate proxies can proxy-chain to the service allowing connection to Intranet clients, while corporate devices on the Internet are capable of connecting to the service directly.

### Achievement of 100% web filtering

To ensure that 100% protection exists at all time all web traffic, regardless of where the device is physically connected (JFC#5<sup>2</sup>) must be filtered.

There is an issue when a remote user needs to make an initial local connection for authenticating / paying for access when typically using a local hot-spot or hotel. This issue is compounded by the plethora of web redirection methods for authentication / payment.

There is a need for a standard, agreed method/protocol for web charging to be agreed by the industry, thus allowing a standard secure interface to separately handle those access requests.

### Connection from the end-device to the IFR service

The proxy connection must be both mandatory and secure, allowing the passing of the credentials of both the device and the user (JFC#6 & 7), together with other essential user attributes to be passed thus enabling;

- Permission to use the service and possibly charging based on the end-device and/or authenticated user.

<sup>2</sup> The term JFC#n refers to the relevant Jericho Forum Commandment number. See [www.jerichoforum.org](http://www.jerichoforum.org)

This implies that a 3<sup>rd</sup> party, external service provider, would be able to allow access to their services, based on being able to establish a user as part of a company with whom they have a current contract to provide filtering services

- Access rules to web sites applied based on end-device and/or authenticated user and user attributes
- Full logging by device, user, and other group attributes

### Filtering features

Most of the filtering features are available in existing products available today, however for operating in a de-perimeterised environment will necessitate operating in a distributed environment, where multiple, replicated IFR environments allow users to connect to the local service or DMZ while common filtering rules are applied and reporting is consolidated into a central report interface irrespective of the actual hardware or route used to access the Internet;

- Full use of the credentials passed to service, allowing the implementation of rules, filtering, reporting and granular access by user-name, machine-name, users business and business groups

Passed user information and business hierarchy information must support a variety of browsers or be browser independent.

- The ability to define access, by website category and time of day (local time required), total access time and/or limit/throttle data throughput
- The ability to force the redirection to an Acceptable Usage Policy or Standard (AUP/AUS) page on first use by a user and/or at defined periods thereafter. The AUP page should cope with multi-lingual options.
- Should a site being accessed not be categorised then the options should be to fail open, to categorise using an AI engine or fail closed.
- All denial screens should be customisable and contain the reason for denial together with the option for users to request approved override and/or re-categorisation. These requests should be directed to the appropriate person in the IRF system (management) hierarchy or optionally should integrate with an external workflow program.

### Filtering capability

- Standard URL filtering by database lookup of known categorised sites with corporate-wide blocking of standard blocked categories (hate, criminal methods, racism etc.)
- A wildcard capability for black-list & white-list, for example `http://*.my-company.com`
- Sufficiently granular categorisation of sites (40 plus) + sub categories if required
- Intelligent handling and differentiation of port 80 tunnelling traffic, such as IM, Limewire, Skype / VoIP, Video and audio stream, with the ability to handle ports other than 80 & 443 (for video stream etc.)
- Inspection of HHTPS for malicious content, and/or a definable policy for how HTTPS is handled when scanning for potential malicious / inappropriate return traffic
- The ability to identify proxy-sites that deliberately mask the actual URL's as well as those that accidentally mask/bypass the URL's such as the Google cache.
- Blocking by computer name and/or individual user (typically for shared accounts or kiosk-type systems) or restriction to a list of URL's – in addition the ability to restrict web access to a paired computer and user account – for example only the user account “kiosk1” on computer “pc-kiosk1.mydomain.com”.

- Blocking filtering by attachment / file-type and the ability to identify and filter on content (such as streaming media) rather than URL
- The ability to allow access to a single site but block (for example) streaming video, or at a more generic level allow the category “sports sites” but disallow streaming media from those sites.
- Banner / advert replacement or blocking, to minimise page distortion or corruption. The option of replacing adverts with key messages.
- The ability to white-list sites or groups of sites based on individual users or groups of users (preferably using integration with existing groups – say from AD).

### Screening for malicious content

In a de-perimeterised environment an untrusted web site needs to have its trust level raised to ensure the user is presented with browsing that is 100% free of malicious content.

- The systems should bar on known sites with malicious code on that site
- Scanning of files and other non-HTML code for viruses and malicious code
- Use of heuristic detection to ensure a 100% malicious code browsing experience
- The guarantee of zero malicious content (ActiveX, Files, Downloads, ZIP, Java etc. Spyware, inc poisoned links)
- The option to block files (such as ZIP files) that are password protected such that their files contents cannot be inspected to ensure they are clean

### The service provision

In an externally provided service then this service should;

- Be tiered to the Internet at a suitable level
- Have access-points or POPs that are globally available, thus ensuring the mobile de-perimeterised workers take only a short hop to the nearest POP and from there directly to the Internet. Multiple global POPs also provide global load-balancing and resilience
- As a global service, there needs to be 24x7 support
- With a global external service, the only component should be the secure interface between the local Browser and the service, thus minimising upgrades or changes which are the responsibility of the service provider
- Once filtered; all protocols between systems must be inherently secure (JFC#4)

### Systems management

- The ability to allow management at a business, sub-business, group or user level, with the ability to allow groups within groups for ease of long-term management.
- Simple management interface, allowing either centralised or distributed business day-to-day management at appropriate levels within the business.
- The ability to define granular, hierarchical access with different access privileges and abilities for administrators of the system to configuration, rule-sets, management, logs, reports etc. (JFC#10)
- Fast replication of rule-set changes, across all global systems.

### Logging and reporting

The features required by such a service include:

- The ability to automate / schedule and run pre-defined reports and custom reports that can be distributed by e-mail.
- Hierarchical access, with restrictions on those people able to run reports that identify named users. Audit trails and alerting (via e-mail) should reports on identifiable users be requested to ensure the privacy of individuals is safeguard. (JFC#10)
- The facility to anonymise the usage so that the service providers, and or managers can see usage (and abuse), but are not able to identify the individuals.
- The ability to report the extent of Internet use, not just by individuals, but also aggregate to departments and locations, with suitable safeguards on individual privacy.
- Real Time reporting on filtered conditions to monitor for specific occurrences.
- Retention & archiving of log data, of a quality and integrity enabling it to be usable as evidence in potential court case
- Configurable data retention policies, to meet business, industry, regulatory and/or country specific requirements
- Browser type and version detection and reporting
- Encryption of all log data - if on a shared database
- The ability to backup of all logs, configuration, data, etc. (but in encrypted form, thus allowing both storage off-site and also proof that that data has not been tampered with)
- Tripping and alerting (via e-mail or even SMS) on;
- Categories and keywords
- Excess use / threshold trip
- Excess blocking / threshold trip
- Potential (bot) access to suspect sites (typically an infected computer “calling home”)
- XML for standardised reporting (with other security tools)
- Aggregation of data from global proxies into single interface or a single report
- Be able to give an accurate figure of unique users over a defined period (potentially to assist with charging)

## Challenges to the industry

The numbering here does not imply any priority.

1. There needs to be a “standard” for web-page redirection that allows a minimal subset of protocol exchange to allow access to be granted by either password, token, certificate, pre-authentication or payment card.
2. The industry need to agree a standard mechanism for secure proxy connectivity with credentials being passed.

## The way forward

The Jericho Forum believes that accelerating the use of inherently secure protocols for proxy connections, with the ability to use those protocols either within the corporation or outside will allow corporate to provide a simpler, yet more secure and holistic approach to web access.