radware
availability | performance | security

A Superior Global Redirection Solution For Multiple Data Center Environments:

Using a Three-Tier Architectural Approach

December 2009

North America

Radware Inc.

575 Corporate Dr., Lobby 1
Mahwah, NJ 07430
Tel: (888) 234-5763

International

Radware Ltd.

22 Raoul Wallenberg St.
Tel Aviv 69710, Israel
Tel: 972 3 766 8666

www.radware.com

Radware Superior Global Redirection Solution

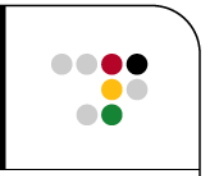
Abstract

For multi-site application environments, Radware's AppDirector Global Redirection Solution delivers highest level of service and optimum user experience. Radware's advanced technology ensures that users are always served by the best performing site and server by using patented site selection methods.

Radware's Global Redirection Solution capabilities provide:

- The highest level of redundancy in all levels: servers, devices and sites.
- Guaranteed application persistency by the widest set of application redirection methods so that applications and transactions are not interrupted.
- The most flexible and versatile mechanism to address any global application delivery requirement.
- Fully supports DNS based redirection, and in addition provides a solution for the DNS shortcomings, ensuring global failover and session completeness.

Table of Contents



Introduction

For multiple data center environments, Radware offers superior *Global Redirection* technology as part of its core application delivery services to ensure uptime and transaction completion, and to deliver the fastest application response times. Radware's approach to optimizing the delivery of applications over IP networks utilizes a host of integrated application delivery services including health monitoring, traffic redirection, load balancing, traffic shaping, bandwidth management, intrusion prevention and denial of service protection. Radware's multi-tier solution is a key differentiator that fully optimizes globally distributed server resources to maximize return on IT infrastructure investment and drive higher business productivity.

The Challenge

Ensuring that users' transactions are executed on through to completion is a key data center requirement. This means that throughout the "life" of the transaction, users must always be directed to an available site that has knowledge of the user information and the status of the transaction. With multiple data center environments there is the additional challenge of ensuring that users are directed to the site delivering the best user experience in terms of fast response time and transaction continuity. This is a fundamental application delivery service need and must be supported by Application Delivery Controllers (ADC) for multi-site configurations all while dealing with business fluctuations (bursts and peaks) and potential application or network failures.

Limitations of DNS-Only Global Load Balancing Solutions

DNS redirection is a key construct used by global server load balancing solutions. However, DNS redirection presents unique challenges in global data center infrastructures. The DNS system has limitations and a DNS-only solution does not guarantee accurate user persistency or transaction completion. Therefore, DNS redirection may cause inconsistencies in user and session persistency. This results in aborted transactions and from a user perspective, is surprising, and often times infuriating.

The nature of today's web-applications is such that it is very likely that a user will issue multiple DNS resolution requests for the same domain name during a single transaction. Unfortunately, regular DNS-only global load balancing solutions can not guarantee transaction completion because they do not "understand" that multiple DNS requests may be originating from the same user. In this case, the new request is treated as if it is coming from a new user and the user will then be redirected to the wrong data center in the middle of a transaction which would cause the transaction to fail.

DNS caching is another problematic issue for multi-site environments. When a user receives a DNS reply it is often cached and used in the future – thus a "live" domain name to IP address translation does not occur. Therefore if a data center goes down in the middle of a transaction, a portion of users will continue to use their cached DNS information in order to access the site, resulting in application downtime for those users. In addition, DNS is limited in identifying the applications and application content and thus a user can be forwarded to a sub-par server for the user's request.

There are also situations where users request a DNS resolution mid-session. This may cause an application delivery controller to resolve the request to a different site than the user's current site. Final outcome – transactions are dropped mid-session.

Radware's Global Redirection Solution

Radware's Global Redirection is the only solution in the market to **guarantee transaction completion and fastest response times** by fully optimizing globally distributed server resources across multiple data centers based on application/transaction persistency, content availability, load and proximity. Radware's superior Global Redirection capabilities enable transparent failover, complete disaster recovery among multiple sites, and optimal service delivery thru best site selection. This allows customers to create a true global network for delivering a wide range of applications while ensuring maximum uptime and optimized service delivery of all hosted applications for globally dispersed users.

Radware's Global Redirection is offered as a core application delivery service. It is available as a software module as part of Radware's APSolute™ architecture for application-smart networking, and runs on Radware's AppDirector™ intelligent application delivery controller. AppDirector's global redirection solution has three main attributes:

1. **Fastest site failover**: AppDirector's unique ability of announcing VIP (Virtual IP) reachability using common routing protocols such as BGP, RIP and OSPF, provides fast and transparent service continuity (without requiring additional DNS query) in case of site failure.

2. **Most accurate load & proximity reporting**: Both the end user's proximity to the data centers, and the data centers load and availability are calculated in real time to achieve the industry's most accurate measurements and decision on the best site to serve a user. AppDirectors installed in different sites frequently exchange load and proximity information to ensure that site selection is based on up-to-date and accurate information.

3. **Widest range of redirection methods**: The wide array of redirection methods which includes any combination of: DNS, Global Triangulation®, HTTP redirection, RTSP redirection and Client NAT provide globally distributed data centers with the only solution that can ensure high availability and optimized delivery for any IP-based service/application. Moreover, this array of redirection methods is also used to correct inferior site selection decisions, which due to DNS caching or due to lack of persistency identifiers in DNS request, may interrupt transaction completion.

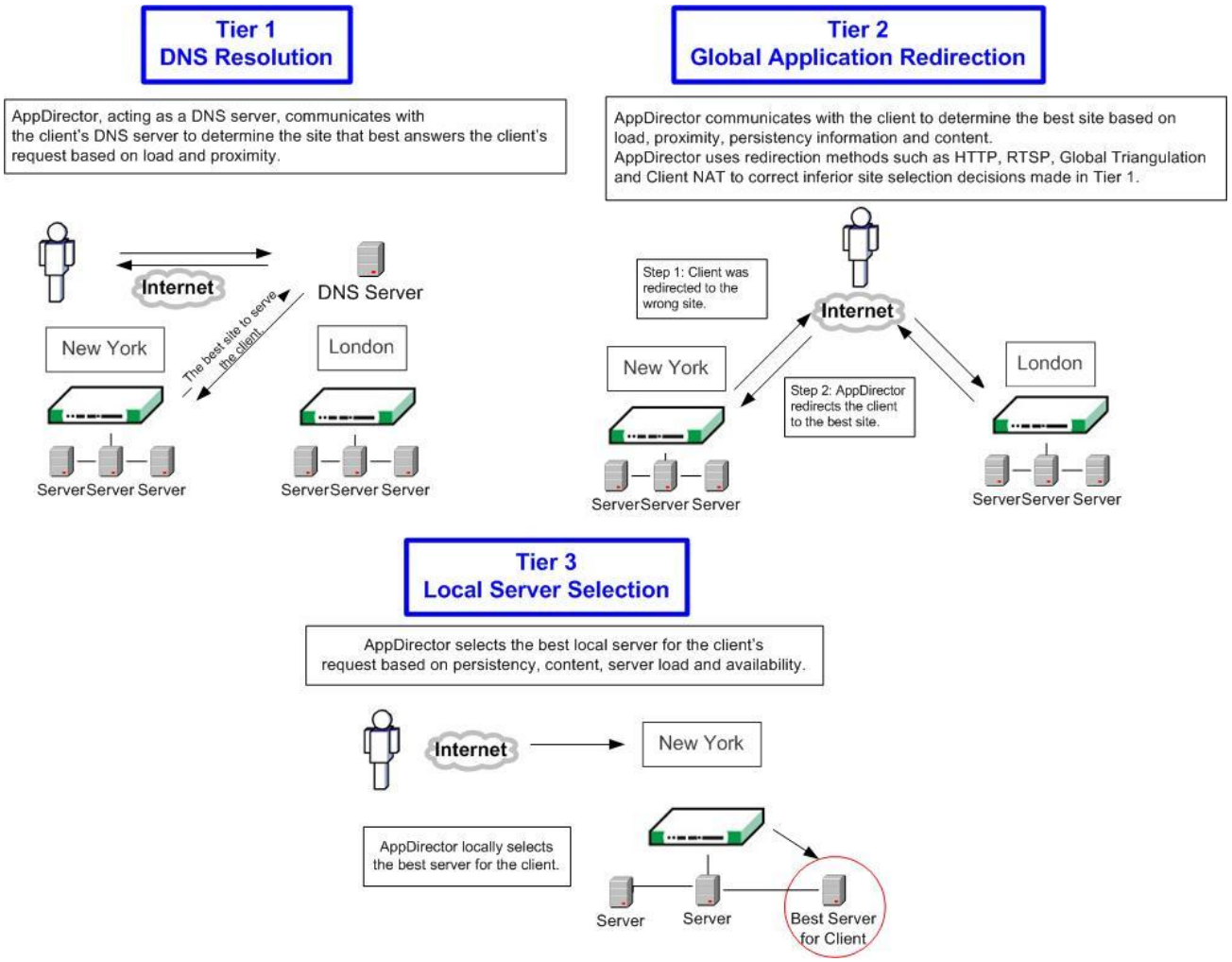
Radware Differentiators

AppDirector overcomes known limitations of DNS-only global solutions to ensure transaction completeness, business continuity and highest quality of service for the users for widest range of applications. Additionally, unique resilience mechanisms, such as Anycast, ensure the industry's fastest failover for DNS resolution services, as well as any type of application service.

Radware's advanced global redirection technology supports both application and DNS layer redirection between sites. This innovative solution injects session identification into the process, so that even if DNS resolution produces a result that would redirect a user to the wrong site, the global application redirection layer forwards the user to the correct data center.

Multi-tier Global Redirection

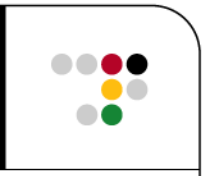
Radware's multi-tier global redirection architecture delivers fast user response time and application/transaction persistency to ensure transaction completeness, session consistency, and end-user satisfaction. This three tier approach overcomes DNS system failures using three redirection approaches with AppDirector communicating and sharing site information to ensure that each step of the global redirection is achieved. The three tiers of Radware's global redirection architecture are (1) DNS Resolution Redirection, (2) Global Application Redirection and (3) Local Server Selection.



AppDirector's multi-tier global redirection architecture locates the best DNS server, site and local server to optimize the user's transactions.

Tier 1 – DNS Resolution Redirection

Radware's DNS redirection uses DNS resolution for site selection, which is based on the site availability, load and proximity. When a DNS query is made, AppDirector responds with the virtual IP address of the best site for the user.



Global load balancing is set up with back up, or redundancy, in mind. Activating DNS resolution service on globally dispersed AppDirector units provides DNS redundancy to ensure users have continuous access to websites and mission-critical applications. If one of the sites goes down, Radware's AppDirector is there to serve user requests.

To further ensure availability, the virtual address used for DNS resolution may also be shared by multiple sites, either concurrently or in case of failure. This is achieved by "Anycast," which announces the IP address to the routing domain in multiple network locations.

Tier 2 – Global Application Redirection

The DNS resolution of a global system is limited and must be backed up to ensure transaction completion and session consistency. If the user's session was redirected by DNS to the wrong site, Radware's unique global application redirection mechanisms, with built-in persistency, overcomes DNS inconsistencies and prevents session disconnections.

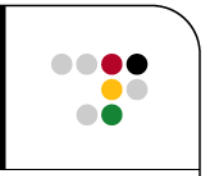
Application persistency ensures that the user is served by the same server throughout a transaction and during future transactions. AppDirector uses application server identifiers that are stored in either in the URL, HTTP cookie, or SIP Call-ID to recognize the correct site/server for handling all subsequent user requests.

Application Redirection methods include: HTTP redirection, Global Triangulation, RTSP redirection, and Client NAT.

HTTP Redirection
For HTTP traffic, AppDirector uses the standard HTTP redirection mechanism (code 302/moved temporarily) to redirect the client to an alternate site. After being redirected, the client will directly approach the alternate site for the rest of the transaction. Redirection to an alternate site, with optimal proximity to the client, reduces the response time that the client experiences. As the redirection is "temporary," a client's new transaction will be resolved again, and the AppDirector reconsiders application persistency, availability, load and proximity.

Global Triangulation®
Global Triangulation is Radware's patented redirection method which is used with multiple AppDirectors to redirect traffic in a Global Application Redirection tier.

In Global Triangulation a data "triangle" is formed where the client is the first vertex, the redirecting AppDirector (Global) is the second vertex and the third vertex is a local or global AppDirector that actually handles the client request. Clients approach one AppDirector, but are transparently redirected to another remote AppDirector unit in a different site. The remote AppDirector responds with the VIP of the original AppDirector that the client initially contacted. Radware's Global Triangulation can be used for all applications, web based and legacy, and is done transparently to the client, overcoming all caching issues at the client side. Triangulation can also be used to "globalize" legacy applications that are destined at specific IP addresses and not use DNS mechanisms.



As most of the application traffic goes from the server to the client, selecting a server with an optimal path for the client results in accelerating response time and enhancing application client experiences.

RTSP Redirection
For RTSP traffic, AppDirector uses the standard RTSP redirection mechanism to redirect the client to an alternate site. After being redirected, the client will directly approach the alternate site for the rest of the transaction. Redirection to an alternate site, with optimal proximity to the client, reduces the response time that the client experiences. As the redirection is "temporary," a client's new transaction will be resolved again, and the AppDirector reconsiders application persistency, availability, load and proximity.

Client NAT
Client NAT is primarily used to redirect client requests to a remote server, in scenarios where Global Triangulation method (see above) cannot be used. In Client NAT mode, AppDirector acts as a proxy at the IP level of the client communication, i.e. retrieves the content from the server on behalf of the client and then responds to the client.

To ensure that the server's responses are always routed via AppDirector and not sent directly to the client, AppDirector replaces the client IP with its own IP address before redirecting the traffic to the server.

Client NAT redirection method, like Global Triangulation, can be used for redirecting any IP based application.

Tier 3 – Local Server Selection

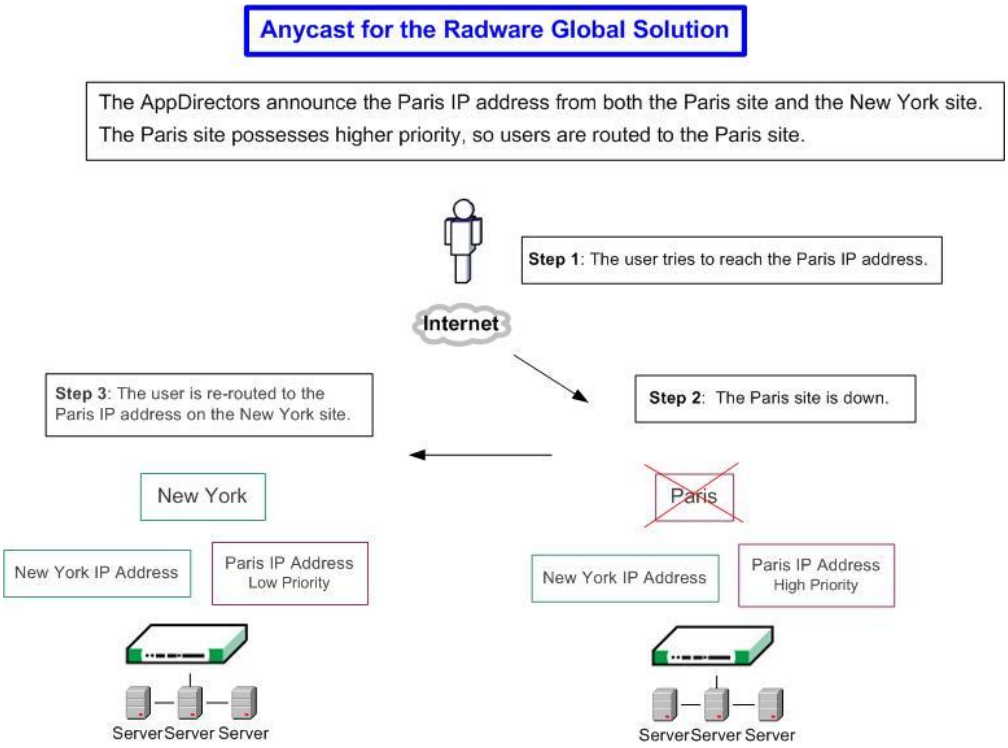
Local Server Redirection
After determining the site, the local AppDirector selects one of the servers in the site to serve each client's request. This selection is based on the availability and load of the servers of the local site, while maintaining application persistency to a single application server throughout the transaction.

Transparent Site Failover

Anycast Failover
Radware's global failover Anycast mechanism ensures application availability in case of a site failure. The failover mechanism provides transparent high availability of service without the need for new DNS resolution. The Anycast mechanism offers two failover methodologies:

(i) "Equal Anycast" uses standard routing protocols (BGP, OSPF, RIP) to advertise service availability from all global AppDirector sites. It relies on the routing logic to govern the delivery of messages to one of these locations. Using "Equal Anycast" optimizes the user response time, as each request is routed to the closest site.
(ii) "Prioritized Anycast" allows the selective failover mechanism by advertising the service with different priorities to different locations. As a result, once a site fails, the routing system will immediately redirect its traffic to the backup sites. Using "Prioritized Anycast" allows transparent application continuity in case of site failures.

Using Anycast optimizes the user response time, as each request is routed to the closest site that can serve the user.



Anycast broadcasts the same IP address from multiple sites.

Summary

Radware's innovative multi-tier Global Redirection architecture is a superior and unique solution for ensuring transaction completion and fastest application response time in multiple data center environments. Radware Global Redirection differentiators include:

- Overcoming all DNS shortcomings including DNS caching by the end user, and mid-session DNS resolutions - to ensure transaction completion.
- Largest and most flexible application redirection capabilities - to guarantee best response time to any application.
- Most advanced patented proximity detection to identify the best site - best user experience.
- Support for global operation for legacy application that lack DNS support - thus all applications can run across global data centers.

For additional reading, please refer to the AppDirector Global Solution Technology Overview white paper.

Glossary

ADC – Application Delivery Controller – device for optimizing response time of applications on a network. Typically use load balancing technologies on layer 4-7.

BGP – Border Gateway Protocol - core routing protocol of the Internet. It works by maintaining a table of IP networks or 'prefixes' which designate network reachability among autonomous systems.

DNS – Domain Name System - stores and associates many types of information with domain names, but most importantly, it translates domain names (computer hostnames) to IP addresses.

HTTP – HyperText Transfer Protocol - is a method used to transfer or convey information on the World Wide Web.

IP – Internet Protocol - is a data-oriented protocol used for communicating data across a packet-switched internetwork.

NAT – Network Address Translation - involves re-writing the source and/or destination addresses of IP packets as they pass through a router or firewall.

OSPF – Open Shortest Path First - protocol is a link-state, hierarchical interior gateway protocol (IGP) for network routing.

RIP – Routing Information Protocol - is one of the most commonly used interior gateway protocol (IGP) routing protocols on internal networks (and to a lesser extent, networks connected to the Internet), which helps routers dynamically adapt to changes of network connections by communicating information about which networks each router can reach and how far away those networks are.

RTP – Real-time Transport Protocol - defines a standardized packet format for delivering audio and video over the Internet.

SIP – Session Initiation Protocol - is an application-layer control (signaling) protocol for creating, modifying and terminating sessions with one or more participants.

URL – Uniform Resource Locator - is a widespread synonym for Uniform Resource Identifier (URI) — many popular and technical texts will use the term "URL" when referring to URI.

VIP – virtual IP address - is an IP address that is not connected to a specific computer or network interface card (NIC) on a computer. Incoming packets are sent to the VIP address, but all packets travel through real network interfaces.