Server Load Balancing Architectures and Solutions

Agenda

- What is Server Load Balancing?
- Issues with Server Load Balancing
- Issues with Server Load Balancing Architectures
What Is Server Load Balancing?

In the beginning of the WWW one Web Server = domain.name

Problems quickly arose:

- More users on site, server performance suffered
- Maintenance on server disabled site
- Server HW/SW issues disabled site
Server Load Balancing (SLB) via DNS Round Robin

First Solution to the WWW Problem
- Add Servers and mirror content
- Use a DNS Server to Rotate Sequentially through Host List

Problems
- Issues with DNS Caching
- All servers receive equal traffic volume
- Off-line servers still receive requests

Evolution of Server Load Balancing

Limitations of DNS led to Creation of a new Server Load Balancing Device:
- Each server receives appropriate traffic volume
- No new requests to off-line servers
Overview of SLB: One Virtual and Many Real Servers

• Many real servers appear as one virtual IP address
• “Server farms” require only one DNS entry

Overview of SLB: Dispatch vs. Direct Mode

Servers can be attached two ways:

Direct Mode:
SLB uses NAT to translate Server addresses

Dispatch Mode:
Servers have two addresses. A ‘real’ address and the virtual address

Benefits
• Minimal Server Configuration
• SLB can work across Subnets
• No NAT Performance Penalty

Cons
• NAT Performance
• Server Configuration
• SLB on one Subnet
Agenda

- What is Server Load Balancing?
- Issues with Server Load Balancing
- Cisco’s Server Load Balancing Solutions
- Cisco’s Server Load Balancing Roadmap

New Server Load Balancing Issues

How can I Load Balance Other Services?
  - e.g. FTP servers, IBM based applications and Firewalls

Can I Base Server Load Balancing On Server information?
  - CPU, Memory, Disk, and Session Capacity
  - Server Content availability and Database Connectivity

Can I deliver URL content specific SLB decisions?
  - URL Load Balancing

SLB creates issues with my SSL and AOL Users?
  - SSL Sticky, Cookie Sticky, HTTP Redirection and other persistence features

Can I provide Load Balancing across my network?
  - Global Redirection based upon content availability and optimized network path to servers
**Firewall Scalability**

- All traffic must pass through firewall
- Firewall becomes the key bottleneck for network scalability and reliability

  Number of Connections

  Number of Connections-Per-Second (PIX 6000+CPS vs. Checkpoint 600+)

  Throughput

  Scheduled and Unscheduled outage

**Firewall Load Balancing Benefits**

- Transparent load balancing
- Automatic fail-over when a firewall is down or added-no service disruption
- Must support both Proxy and Transparent Firewalls
Server Feedback

Probing Server Content

SLB Device probes for content
- Can enable and disable servers based upon probe
- Notify managers on unavailable web pages
Benefits
- Specific Redirection of URL’s to servers
- Segregation of Content to specific servers

Disadvantages can outweigh Benefits
- Static Configuration
- Performance
- HTML coding provides the same service

Persistence Problem

"Dropped shopping carts—resulting not only from proxy server issues but from poor application response time, caching issues and impatient customers—has become a $6 billion drain on e-commerce revenues"

Zona Research Inc., in Redwood City, Calif.
http://www.zdnet.com/eweek/stories/general/0,11011,2504237,00.html
**Persistence Problem**

Hey!! Where did my Grocery List Go?

Joelle
ISP Proxy Server changes IP Address to:
IP: 10.1.1.01
ISP Proxy Server does not change IP Address for 'new' SSL session
IP: 171.68.1.02
SLB Device Load Balances IP: 171.68.1.02 to Server: 10.0.0.3

**Solving the Persistence Problem**

Redesign the Site

Joelle
ISP Proxy Server changes IP Address to:
IP: 10.1.1.01
ISP Proxy Server does not change IP Address for 'new' SSL session
IP: 171.68.1.02
SLB Device Load Balances IP: 171.68.1.02 to Server: 10.0.0.3

Backend Database
Holds state information
Solving the Persistence Problem - SSL Sticky

I love Shopping on the Web!

Joelle

ISP Proxy Server does not change IP Address for ‘new’ SSL session
IP: 171.68.1.02
SLB Device Load Balances Identifies SSL Session ID of 4
IP: 171.68.1.02+ SSL:4 to Server: 10.0.0.4

SSL Session ID: 4
SLB Tracks SSL ID #

Issues:
- IE 5.0 SSL Timeout
- Server Performance
- Rainbow Card helps

Solving the Persistence Problem - Cookie Sticky

I love Shopping on the Web!

Joelle

ISP Proxy Server changes IP Address to:
IP: 10.1.1.01
- Server or SLB device sends ‘cookie’ back to client
- ISP Proxy Server changes IP Address for ‘new’ session
IP: 10.1.1.02 Browser ‘cookie’ set to “10.0.0.4”
SLB Load Balancer Device Identifies Cookie server info
IP: 10.1.1.02 + cookie info Server: 10.0.0.4

Cookie Server 10.0.0.4
SLB Tracks cookie info

Issues:
- Minor Performance
- Maintenance
- SSL(cookie monster)
## Solving the Persistence Problem - HTTP Redirection

**I love Shopping on the Web!**

Joelle
ISP Proxy Server changes IP Address to: 10.1.1.01

- Server or SLB device redirects user to specific server
- User web browser selects www.grocery1.com
- Client always requests that server for further connections and persistence is established
- High Performance Solution

---

### Issues:

- Maintenance
  - DNS Server must have all domain names
  - Bookmarks on Web Browser

---

## Persistence Issue Review

<table>
<thead>
<tr>
<th>Persistence Feature</th>
<th>ISP Proxy Support</th>
<th>SSL</th>
<th>Performance</th>
<th>Maintenance</th>
<th>Browser Independence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Database</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SSL Sticky</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Cookie Sticky</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>HTTP Redirection</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Source IP Sticky</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Agenda

- What is Server Load Balancing?
- Issues with Server Load Balancing
- Issues with Server Load Balancing Architectures

E-Commerce Building Blocks

Global Load Balancers

- Distributed Director
- Global Content Delivery Software

Local Load Balancers

- Local Director
- Catalyst 4840G SLB Switch
- Catalyst 6000

- 16 10/100 Ports
- Ideal for Small 10/100 Server Farms
- 40 10/100 Ports, 2 Gigabit in a 2 RU chassis
- Integrated Server Load Balancing
- Ideal for Web Hosting in POPS or Midsise Server Farms
- Scales up to 384 10/100 Ports or 130 Gigabit Ports
- Integrated Server Load Balancing
- Ideal for high density Web Hosting and 10/100/1000 Server Farm Aggregation
Layer 2, 3, 4, 7 Switching Depth

MAC  IP  TCP/UDP  Application

128→1518+ Layer 7

48 bytes Layer 4
Layer 7 Load Balancing Requires TCP Termination HTTP 1.0(Multi) VS HTTP 1.1

32 bytes Layer 3
Layer 7 QoS & Switching Does not Require Packet Termination Maybe Multi-Packet

12 bytes Layer 2
Layer 2, 3, 4, 7 QoS & Switching and Load Balancing Do Not Require Packet Termination

Layer 4 Load Balancing

MAC  IP  TCP/UDP

48 bytes Layer 4

SYN
SYN ACK
ACK
HTTP Get
HTTP Res.1
...
FIN

CPU
L4 Lookup

SYN
SYN ACK
ACK
HTTP Get
HTTP Res.1
...
FIN
Layer 7 Load Balancing

HTTP 1.0 Single HTTP Get Per-Session

HTTP GET WWW.CISCO

TCP/UDP

IP

MAC

... HTTP Get1...

HTTP Res1...

FIN

Seq. 200 SYN

Seq. 200 SYN ACK

Seq. 300 TCP Termination

CPU

URL Lookup

Seq. 400 SYN

Seq. 400 SYN ACK

Seq. 400 ACK

HTTP Get1

HTTP Res1

FIN

HTTP 1.0 Single HTTP Get Per-Session

Layer 7 Load Balancing

HTTP 1.1 Multi HTTP Get Per-Session

HTTP GET WWW.CISCO

TCP/UDP

IP

MAC

... HTTP Get1...

HTTP Res1...

FIN

HTTP Get2

HTTP Res2

FIN

HTTP 1.1 Multi HTTP Get Per-Session
Sites On HTTP 1.0 VS 1.1

Centralized L4 Switching

- More memory = More connections
- Centralized State
- Centralized Persistence
- Physical port independent!
Centralized L7 Switching

- One place to keep persistent information
  - Cookies
  - URL
  - SSL
- Works with and without interface caching
- Interface “punts” to CPU
- Cache examples: Vito, Catalyst 6000 (Dispatch), IOSSLB and ASLB, Arrowpoint

Other Sessions of Interest

2605—Content Networking Product Update
25XX—Network Security and Components