



用恐龍書剖析 AWS Legacy Infra 效能瓶頸

3ric@DevOpsDays

2025/06/05

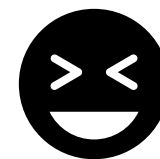


王仁宏 (3ric)

Trend Micro

~~Senior Site Reliability Engineer~~

Senior Reboot Engineer



趨勢科技 DevOps 人才招聘中！

■ (Sr.) DevOps Engineer

Help build and maintain the cloud infrastructure for Trend Vision One!

- 1+ years of experience as a DevOps Engineer
- Familiarity with cloud and containerization technologies such as AWS, Azure, and Kubernetes would be a big plus



\ 更多資訊請見 104 /

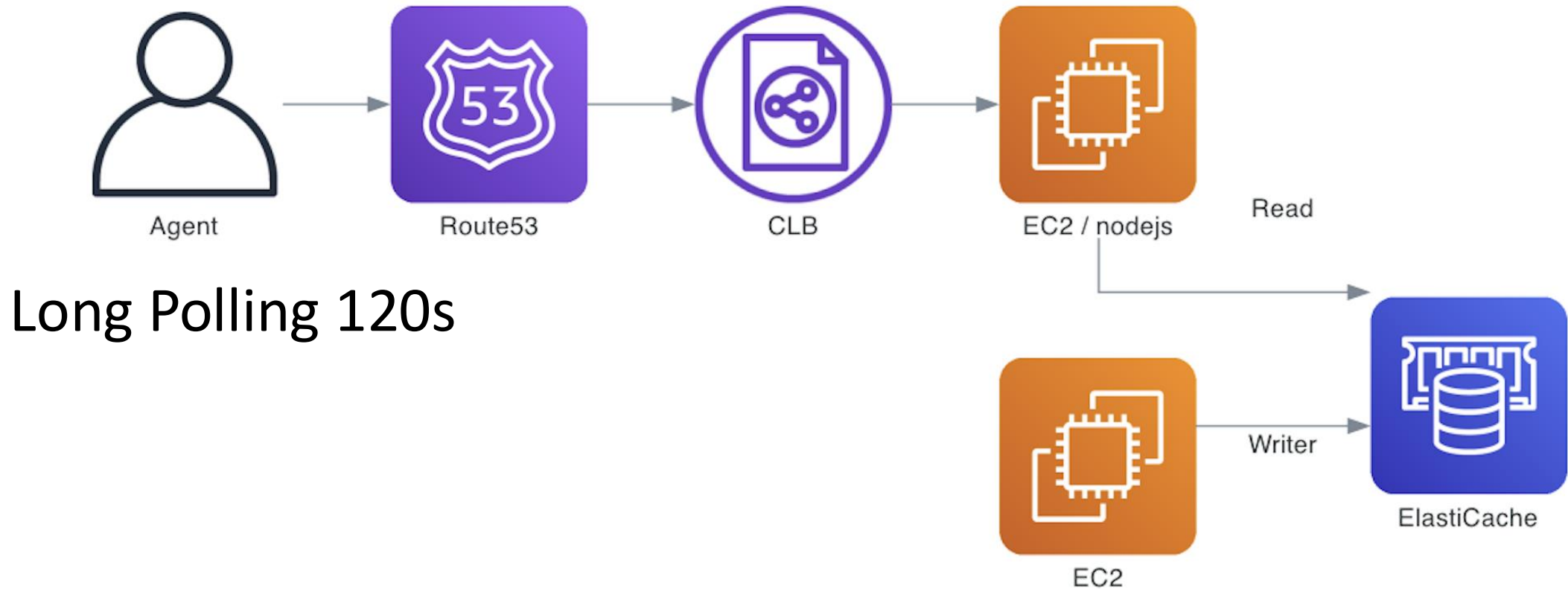
■ (Sr.) Staff DevSecOps Engineer

Drive security and efficiency across Trend Micro's entire software development lifecycle!

- 7+ years of experience in software development (including 3+ years in DevOps)
- Familiarity with cloud and containerization technologies such as AWS, Azure, and Kubernetes



Legacy Infra





🏛️ 工程師聖經守則 🏛️

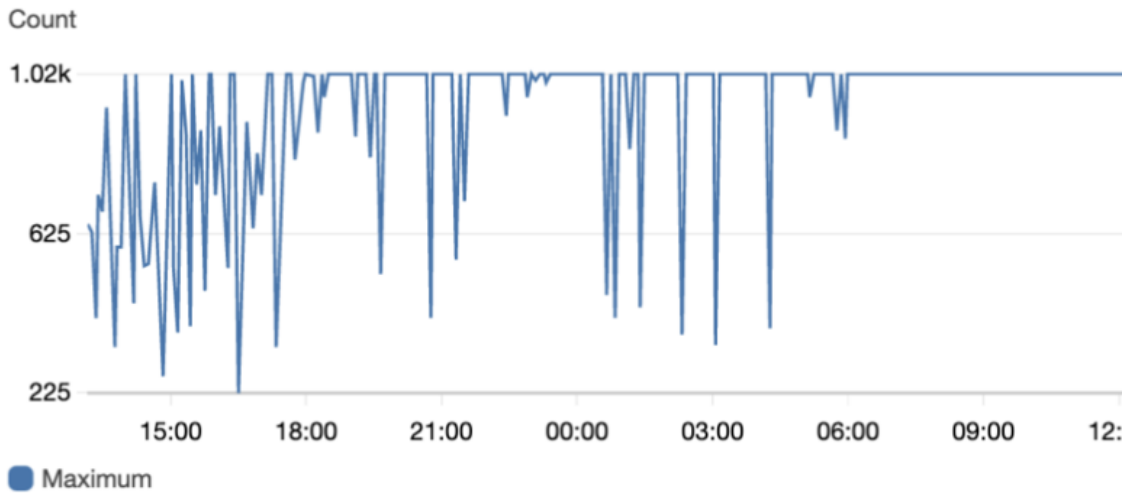


第一條



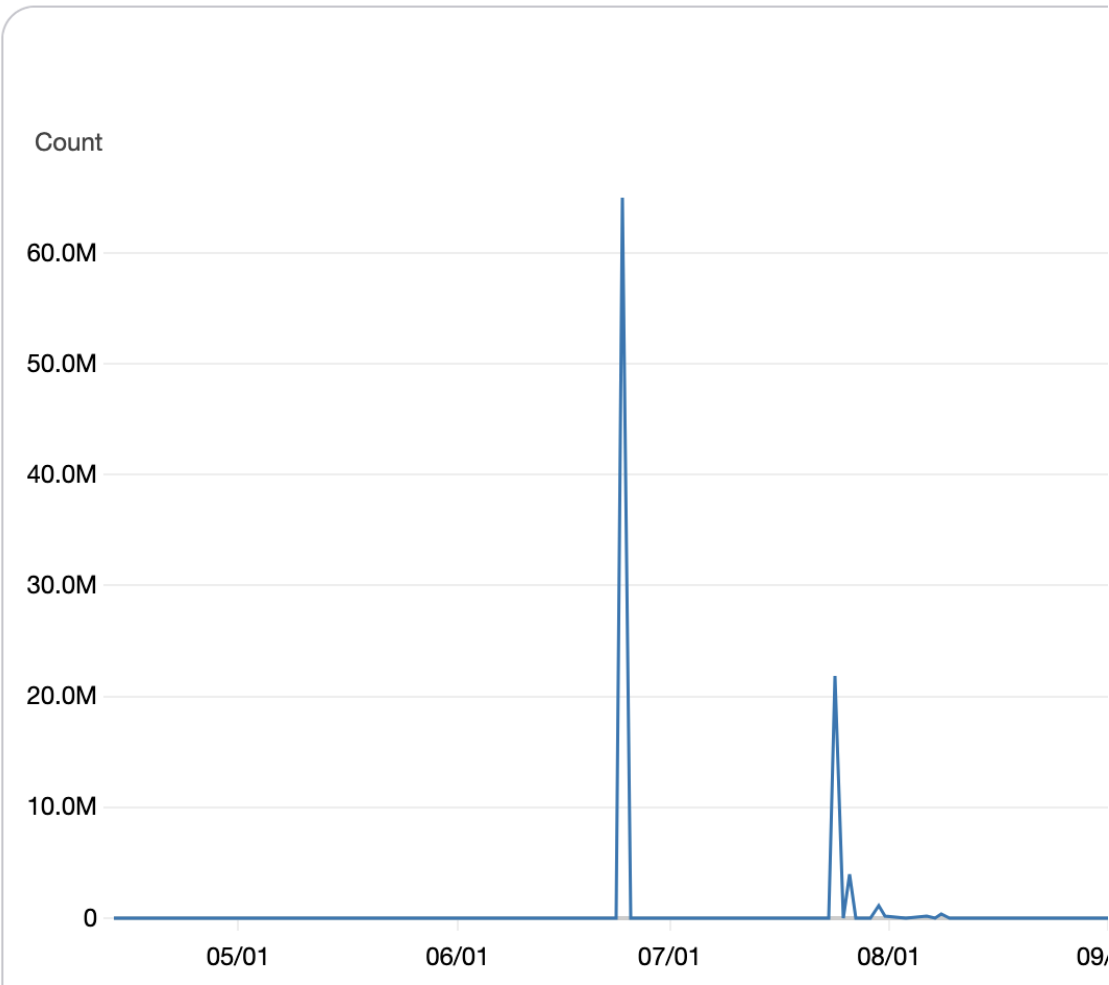
服務可以跑
就不要去動它

SurgeQueueLength



ELB 5XXs

1 day ▼



Zone ▼	Status ▼
us-east-1b	⊗ unhealthy
us-east-1d	⊗ unhealthy
us-east-1a	⊗ unhealthy
us-east-1c	⊗ unhealthy

UnHealthyHostCount

Zone ▾	Status ▾
us-east-1b	⊗ unhealthy
us-east-1d	⊗ unhealthy
us-east-1a	⊗ unhealthy
us-east-1c	⊗ unhealthy

The number of unhealthy instances registered with your load balancer. An instance is considered unhealthy after it exceeds the unhealthy threshold configured for health checks. An unhealthy instance is considered healthy again after it meets the healthy threshold configured for health checks.

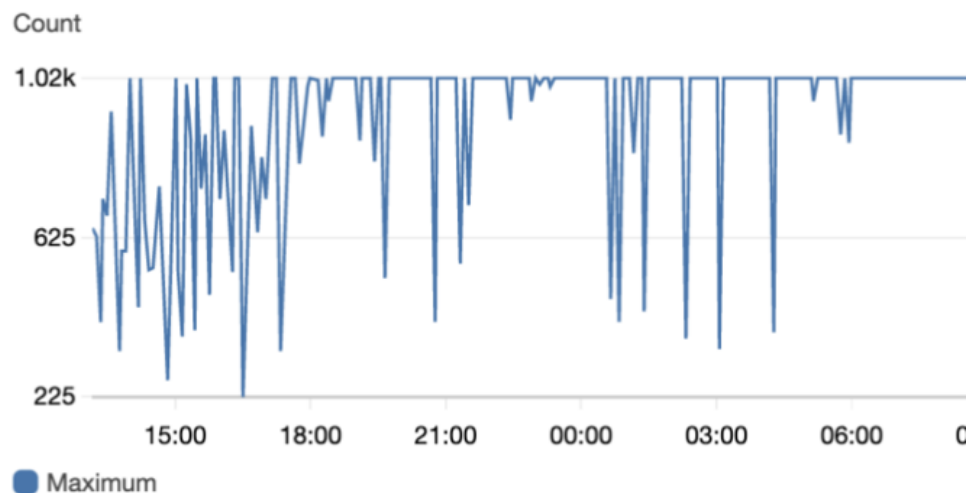
註冊到您的負載平衡器的不健康實例數量。當實例超過為健康檢查配置的不健康閾值後，該實例被認為是不健康的。當不健康實例達到為健康檢查配置的健康閾值後，該實例再次被認為是健康的。

Reporting criteria: There are registered instances

報告標準：有註冊的實例

SurgeQueueLength

SurgeQueueLength



The total number of requests (HTTP listener) or connections (TCP listener) that are pending routing to a healthy instance. The maximum size of the queue is 1,024. Additional requests or connections are rejected when the queue is full. For more information, see `SpilloverCount`.

正在等待路由到健康實例的請求（HTTP 監聽器）或連接（TCP 監聽器）的總數。隊列的最大大小為 1,024。當隊列已滿時，額外的請求或連接將被拒絕。更多信息，請參見 `SpilloverCount`。

Reporting criteria: There is a nonzero value.

報告標準：存在非零值。

SpilloverCount

The total number of requests that were rejected because the surge queue is full.
因為突增隊列已滿而被拒絕的請求總數。

[HTTP listener] The load balancer returns an HTTP 503 error code.

[HTTP listener] 負載均衡器返回 HTTP 503 錯誤代碼。

[TCP listener] The load balancer closes the connection.

[TCP 監聽器] 負載均衡器關閉連接。

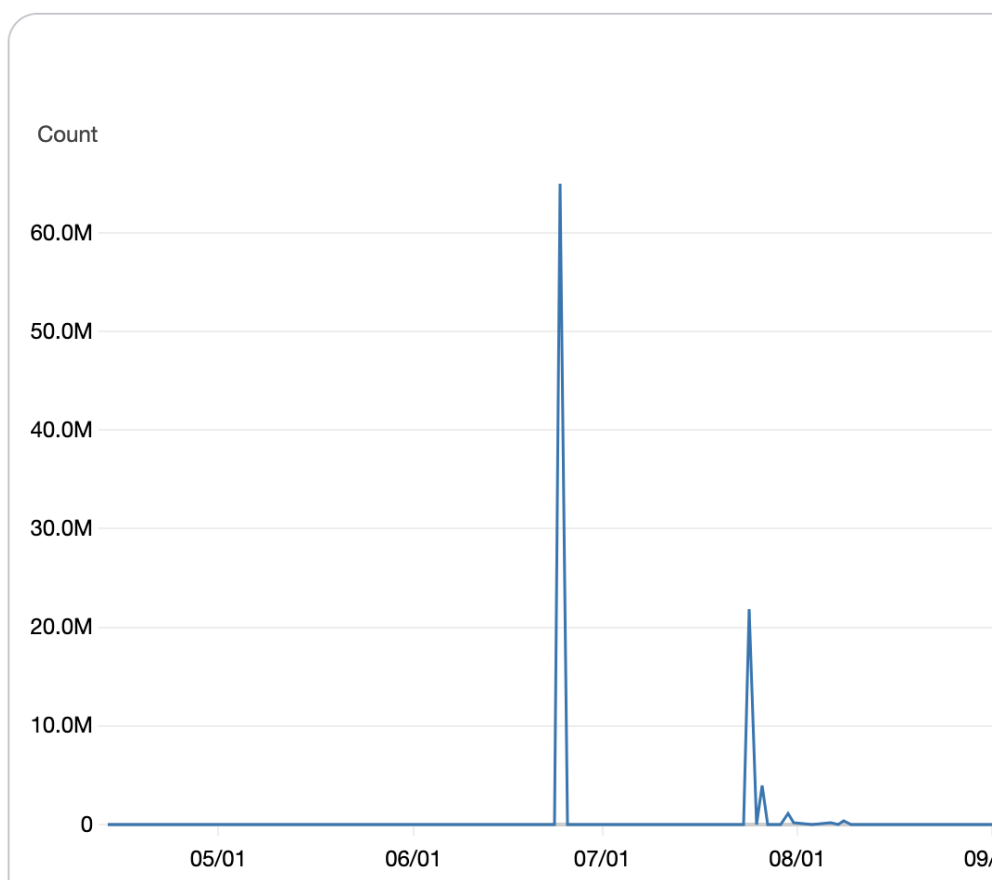
Reporting criteria: There is a nonzero value

報告標準：存在非零值

HTTPCode_ELB_5XX

ELB 5XXs

1 day ▼



[HTTP listener] The number of HTTP 5XX server error codes generated by the load balancer. This count does not include any response codes generated by the registered instances. The metric is reported if there are no healthy instances registered to the load balancer, or if the request rate exceeds the capacity of the instances (spillover) or the load balancer.

[HTTP 監聽器] 負載均衡器生成的 HTTP 5XX 伺服器錯誤代碼數量。此計數不包括註冊實例生成的任何響應代碼。如果沒有健康的實例註冊到負載均衡器，或者請求率超過實例（溢出）或負載均衡器的容量，則會報告該指標。

Reporting criteria: There is a nonzero value

報告標準：存在非零值

HTTPCode_Backend_2XX, HTTPCode_Backend_3XX,
HTTPCode_Backend_4XX, HTTPCode_Backend_5XX

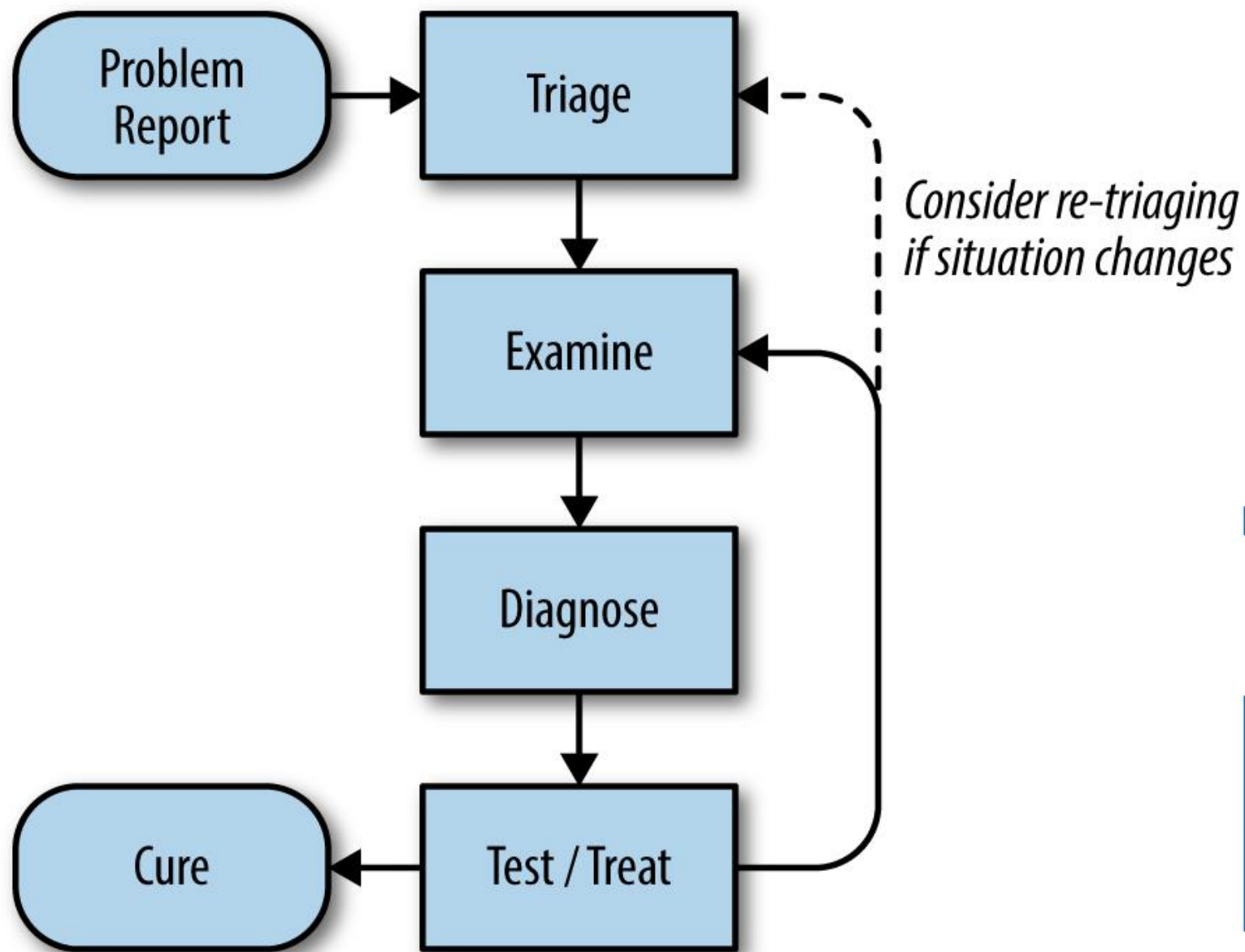
[HTTP listener] The number of HTTP response codes generated by registered instances. This count does not include any response codes generated by the load balancer.

[HTTP 監聽器] 註冊實例生成的 HTTP 響應代碼數量。此計數不包括負載平衡器生成的任何響應代碼。

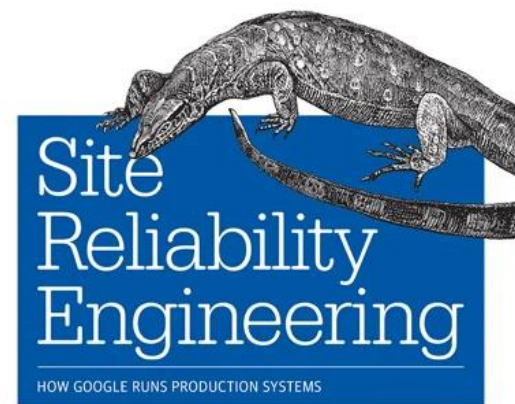
Reporting criteria: There is a nonzero value

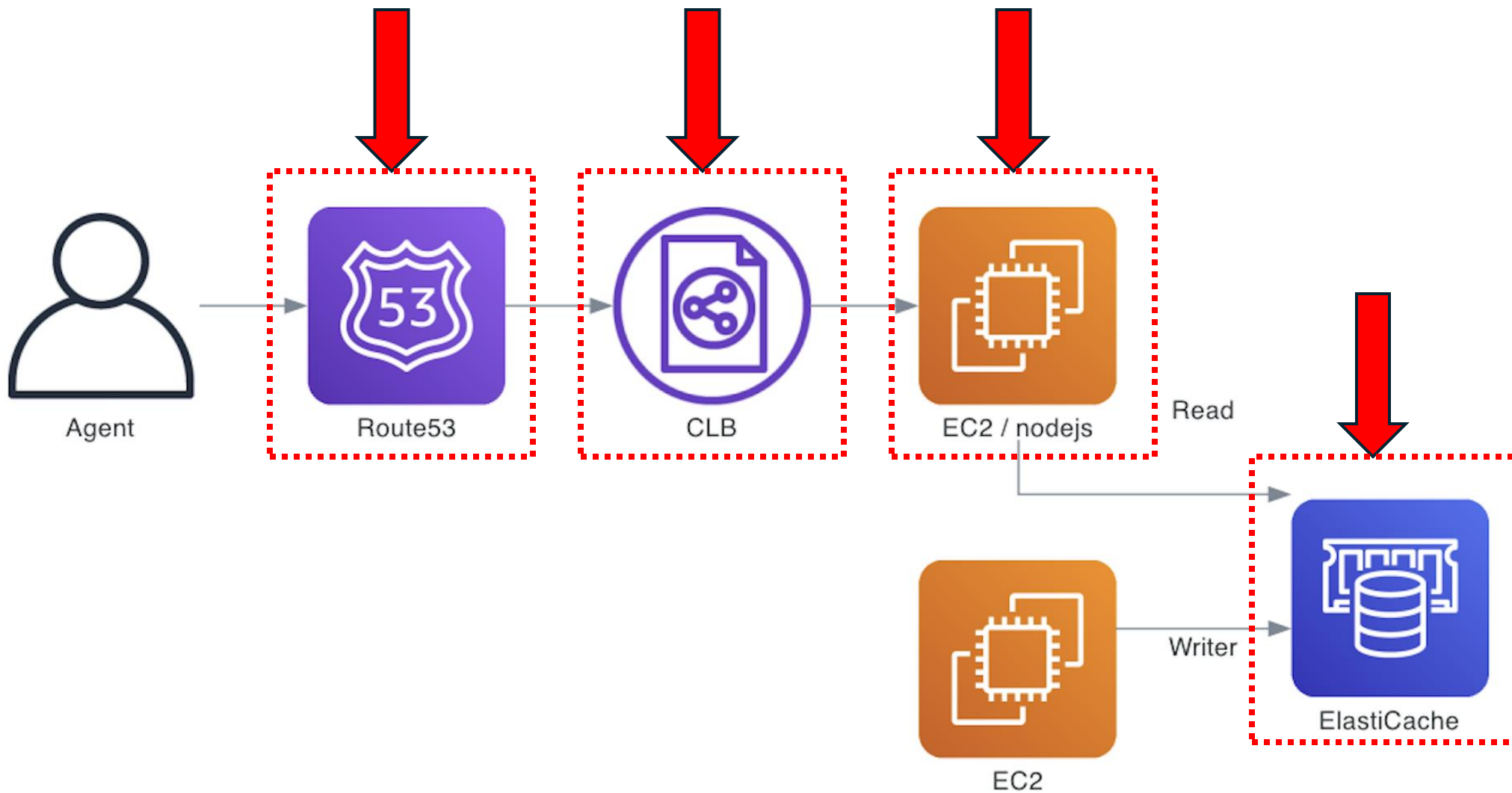
報告標準：存在非零值



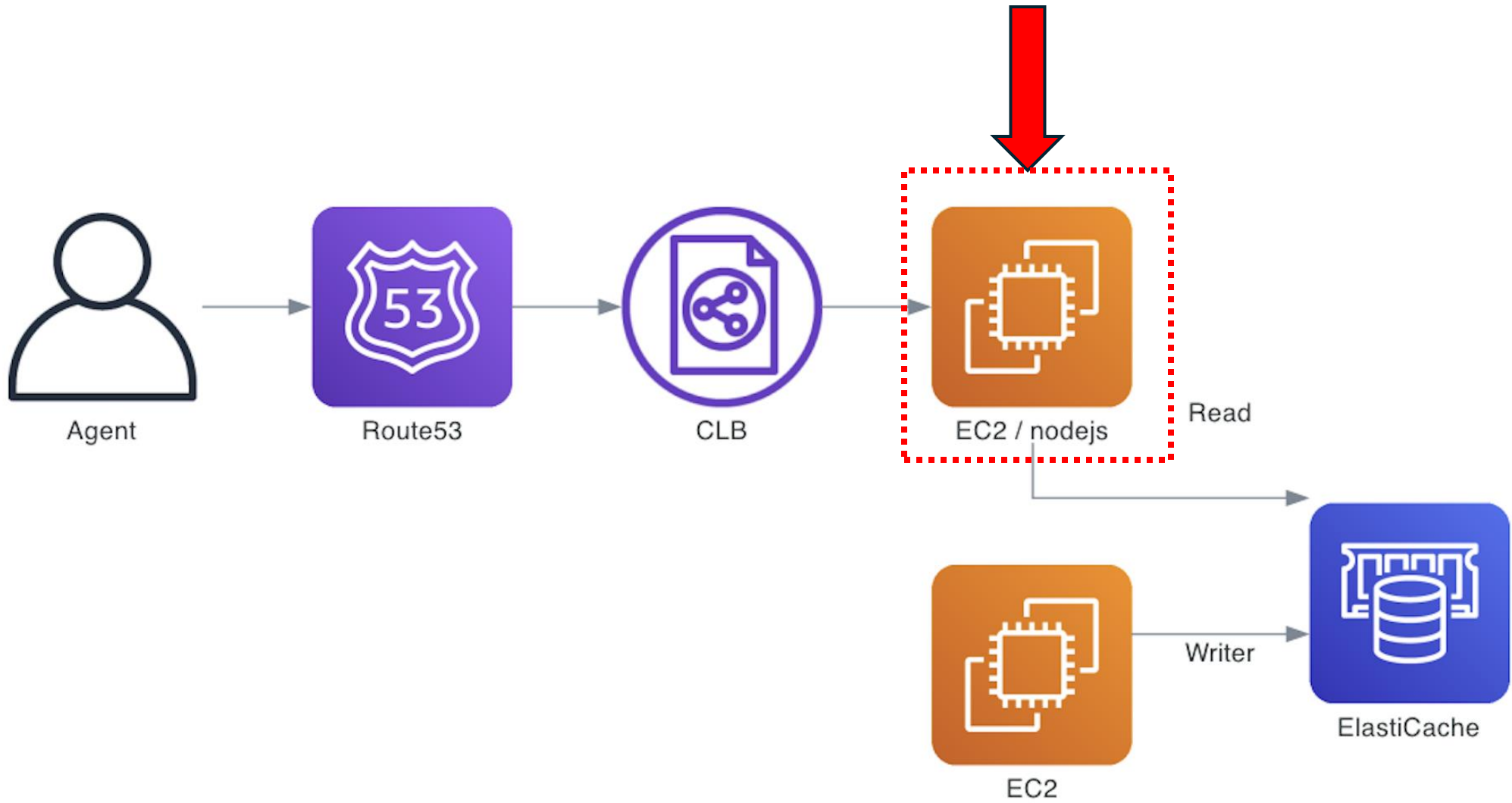


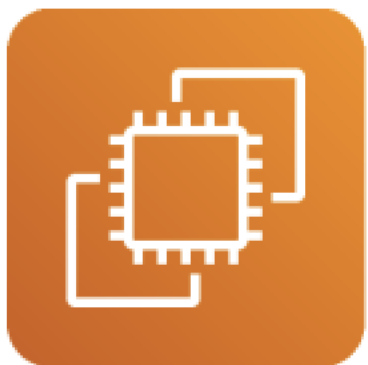
O'REILLY





	UnHealthy HostCount	CLB 5XX	Backend 5XX	SurgeQueue Length	Spillover Count
Route53	X	X	X	X	X
CLB	O	O	X	O	O
EC2	O	O	O	O	O
Elasticache	X	X	O	X	X





AWS EC2



Application / NodeJS

CPU %

MEM %

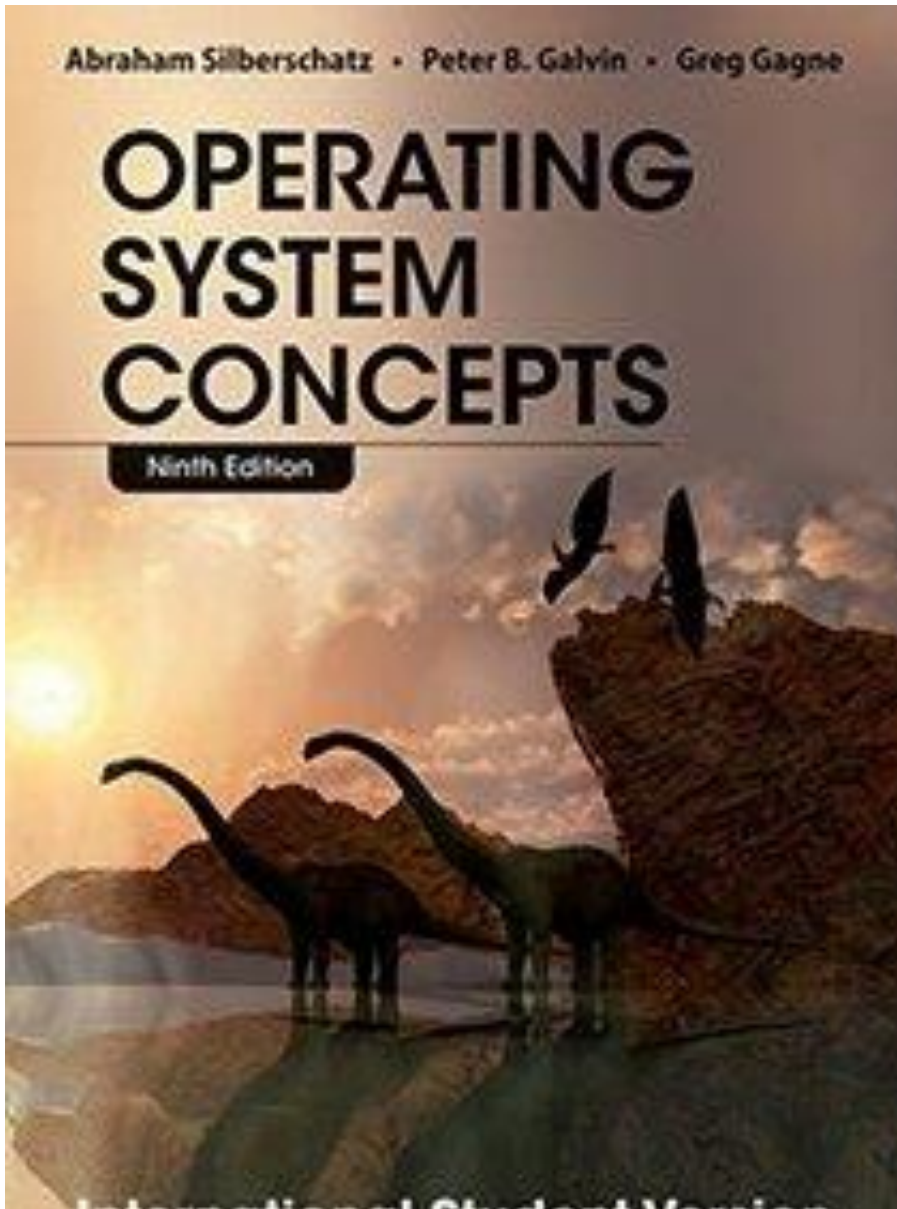
Disk %

Disk I/O

Network I/O

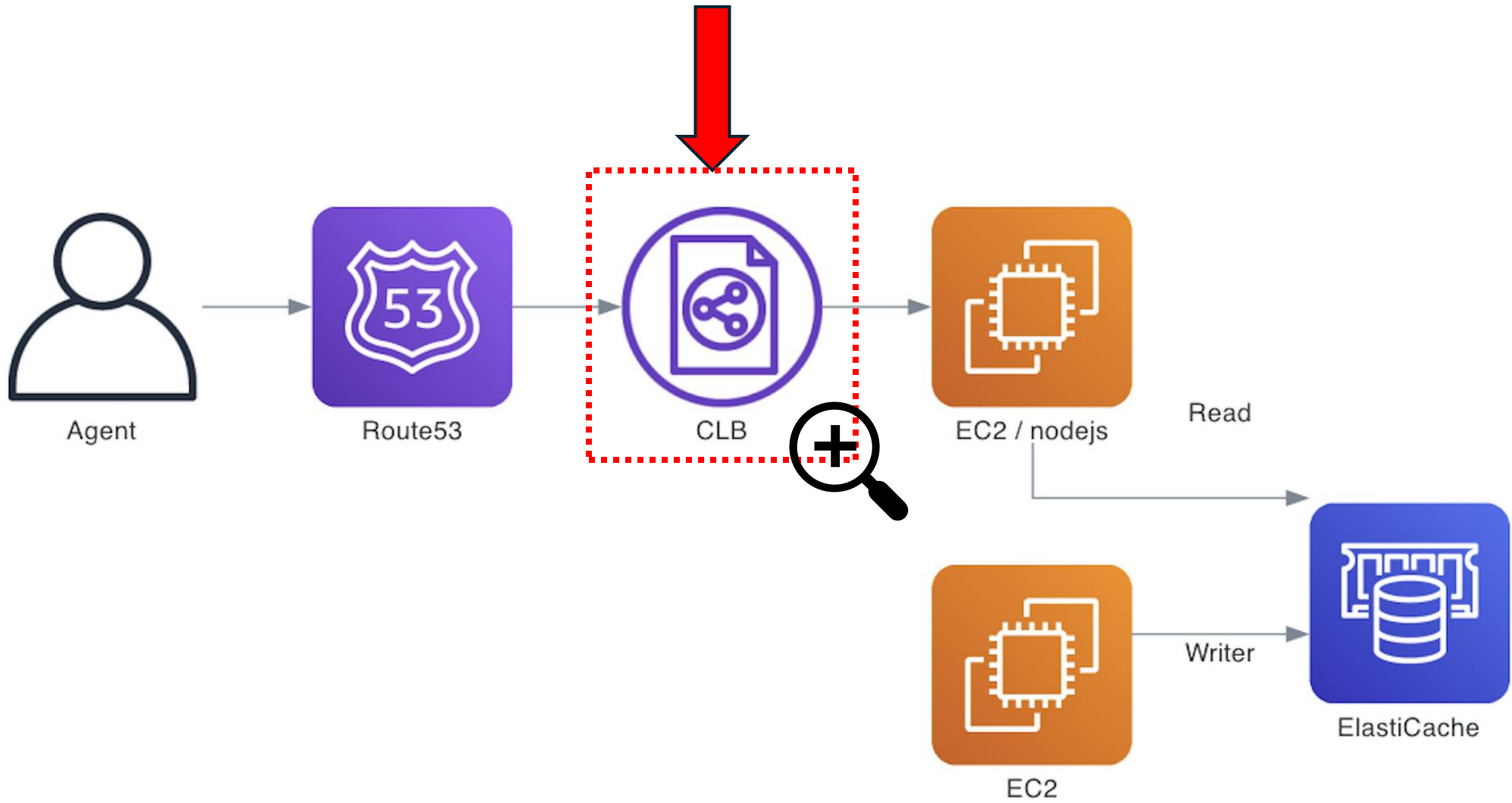
Network Connections

Open Files



```
ulimit -a

core file size          (blocks, -c) X
data seg size          (kbytes, -d) X
scheduling priority      (-e) X
file size              (blocks, -f) X
pending signals         (-i) X
max locked memory      (kbytes, -l) X
max memory size        (kbytes, -m) X
open files              (-n) 100000
pipe size              (512 bytes, -p) X
POSIX message queues   (bytes, -q) X
real-time priority      (-r) X
```





Request routing 請求路由

Before a client sends a request to your load balancer, it resolves the load balancer's domain name using a Domain Name System (DNS) server. The DNS entry is controlled by Amazon, because your load balancers are in the `amazonaws.com` domain. The Amazon DNS servers return one or more IP addresses to the client. These are the IP addresses of the load balancer nodes for your load balancer. With Network Load Balancers, Elastic Load Balancing creates a network interface for each Availability Zone that you enable, and uses it to get a static IP address. You can optionally associate one Elastic IP address with each network interface when you create the Network Load Balancer.

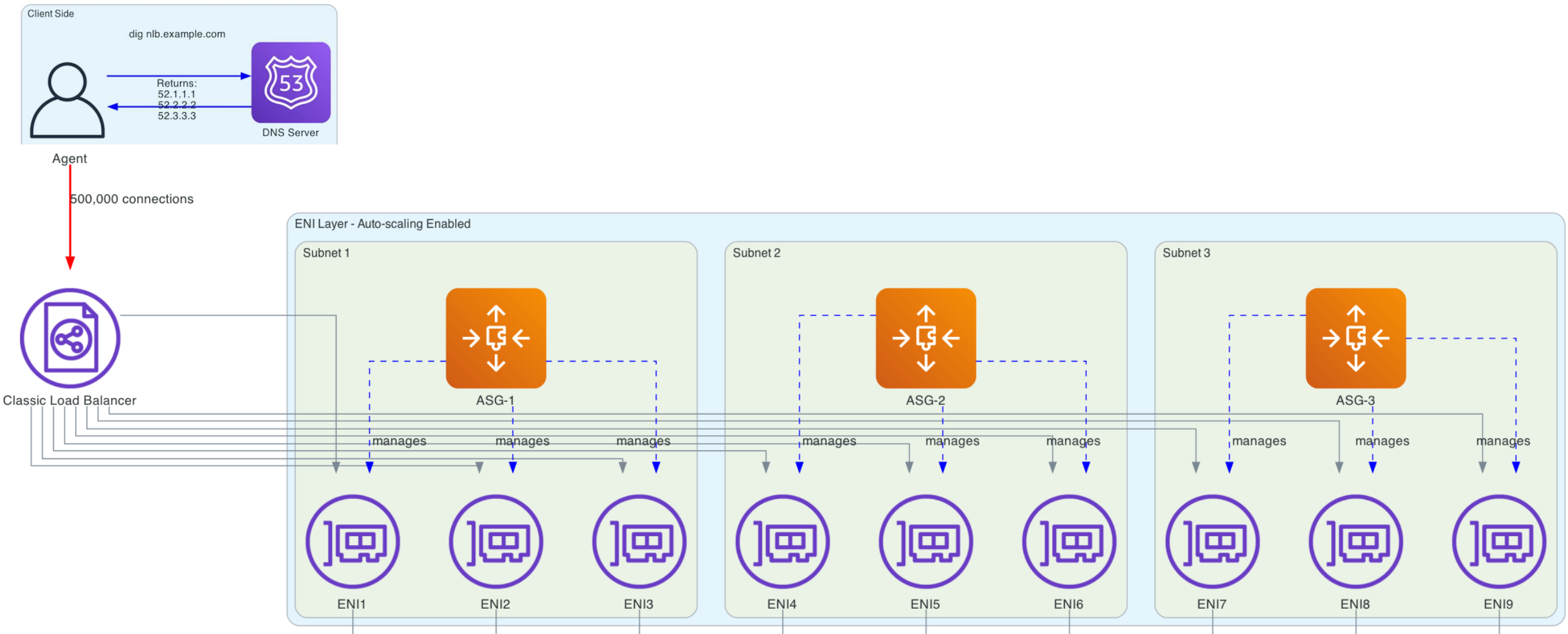
在客戶端向您的負載平衡器發送請求之前，它會使用域名系統 (DNS) 伺服器解析負載平衡器的域名。由於您的負載平衡器位於 `amazonaws.com` 域，因此 DNS 條目由 Amazon 控制。Amazon 的 DNS 伺服器會向客戶端返回一個或多個 IP 地址。這些是您的負載平衡器節點的 IP 地址。對於網路負載平衡器，Elastic Load Balancing 會為您啟用的每個可用區域創建一個網路介面，並使用它來獲取靜態 IP 地址。您可以選擇在創建網路負載平衡器時將一個彈性 IP 地址與每個網路介面關聯。

As traffic to your application changes over time, Elastic Load Balancing scales your load balancer and updates the DNS entry. The DNS entry also specifies the time-to-live (TTL) of 60 seconds. This helps ensure that the IP addresses can be remapped quickly in response to changing traffic.

隨著應用程式流量隨時間變化，Elastic Load Balancing 會擴展您的負載平衡器並更新 DNS 條目。DNS 條目還指定了 60 秒的生存時間 (TTL)。這有助於確保 IP 地址能夠快速重新映射以應對變化的流量。

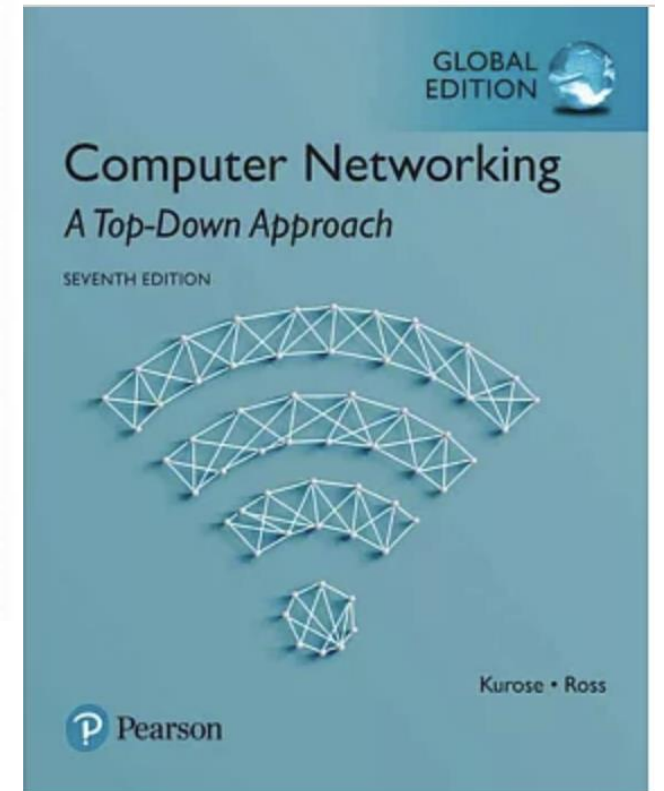
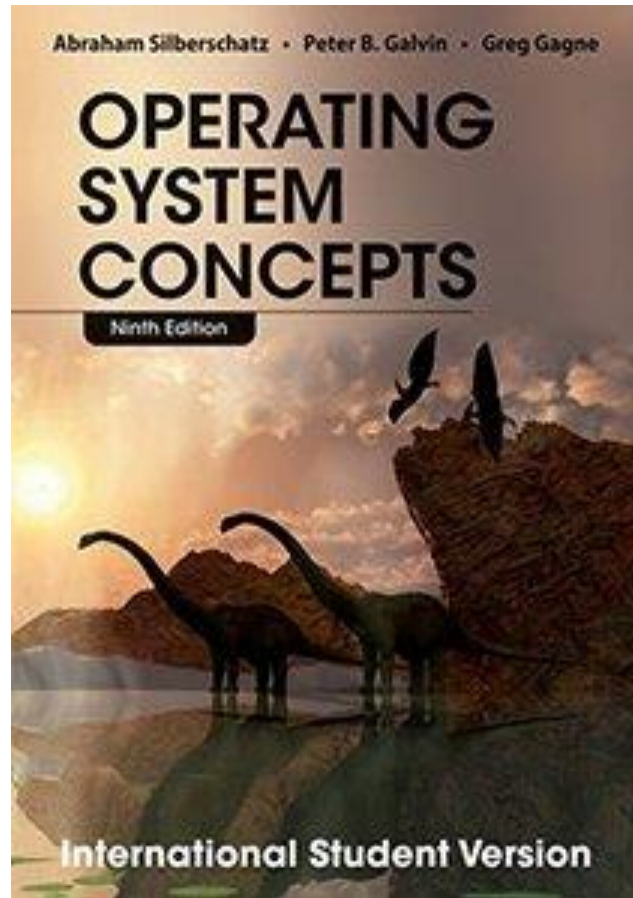
The client determines which IP address to use to send requests to the load balancer. The load balancer node that receives the request selects a healthy registered target and sends the request to the target using its private IP address.

客戶端決定使用哪個 IP 地址來向負載均衡器發送請求。接收到請求的負載均衡器節點選擇一個健康的註冊目標，並使用其私有 IP 地址將請求發送到目標。



Network...	Subnet ID	VPC ID	Availabi...	Security...	Security...	Interface Type	Descr	Public IPv4 address	Primary private IPv4 address
eni-0			...	us-east-1a		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1a		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1a		Elastic network interface	ELB c	4	1!
eni-0			...	us-east-1a		Elastic network interface	ELB c	3	1!
eni-0			...	us-east-1a		Elastic network interface	ELB c	3	1!
eni-0			...	us-east-1c		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1c		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1c		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1c		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1c		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1d		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1d		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1d		Elastic network interface	ELB c	-	1!
eni-0			...	us-east-1d		Elastic network interface	ELB c	-	1!

考驗



https://gd-hbimg.huaban.com/007904a643021ad3c8a8c228d4fa21c55726d74125284-XCYoCi_fw658

<https://down-tw.img.susercontent.com/file/8059c4cc16fd85bde4680472519616e5>

<https://im1.book.com.tw/image/getImage?i=https://www.books.com.tw/img/001/082/45/0010824592.jpg&v=5cf64854k&w=375&h=375>

理論上一台Server可以接收多少TCP連線?

C(lient)10k/100k問題

假設Server提供服務的IP和Port只有一個

192.168.1.1:80

理論上一台Client可以發起多少TCP連線？

假設Client的IP只有一個

TCP Connection 4 Tuple

Source / Client		Destination / Server		Protocol
Source IP Client IP	Source Port Client Port	Destination IP Server IP	Destination Port Server Port	TCP
1	2	3	4	

理論上一台Server可以接收多少TCP連線?

假設Server提供服務的IP和Port只有一個

192.168.1.1:80

Source / Client		Destination / Server		Protocol
Source IP Client IP	Source Port Client Port	Destination IP Server IP	Destination Port Server Port	TCP
2 ³² 42.9億	2 ¹⁶ 65536	192.168.1.1 固定	80 固定	TCP 固定

	Source / Client	Source Port	Destination / Server	Destination Port	Protocol
1	1.2.3.4	2^16 65536	10.20.30.40(固定)	443(固定)	TCP(固定)
2	5.6.7.8	2^16 65536	10.20.30.40(固定)	443(固定)	TCP(固定)
3	9.10.11.12	2^16 65536	10.20.30.40(固定)	443(固定)	TCP(固定)
4	13.14.15.16	2^16 65536	10.20.30.40(固定)	443(固定)	TCP(固定)
5	17.18.19.20	2^16 65536	10.20.30.40(固定)	443(固定)	TCP(固定)
N	IPv4 2^32 42.9億	2^16 65536	10.20.30.40(固定)	443(固定)	TCP(固定)

理論上一台Client可以發起多少TCP連線?

假設Client的IP只有一個

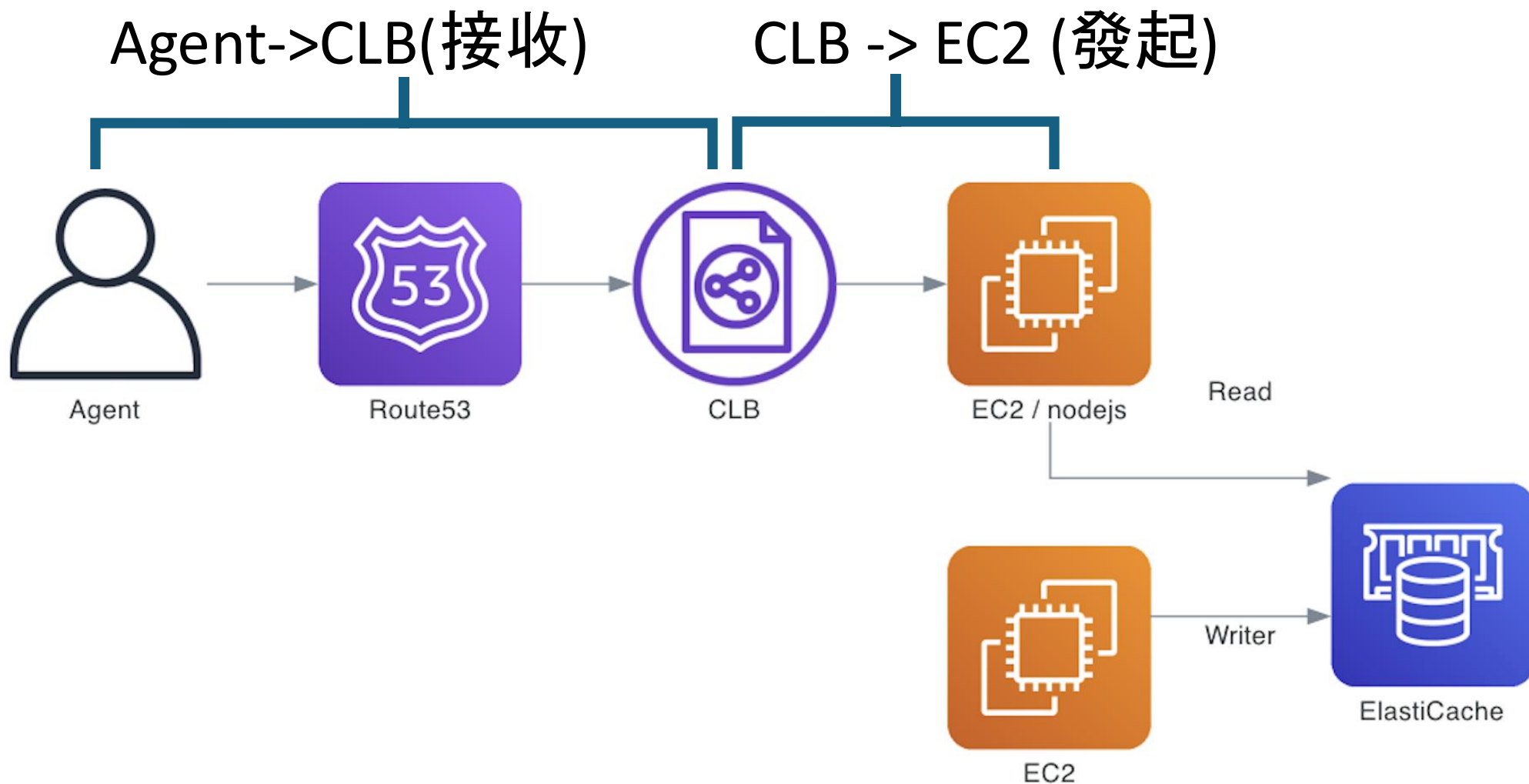
Source / Client		Destination / Server		Protocol
Source IP Client IP	Source Port Client Port	Destination IP Server IP	Destination Port Server Port	TCP
1.2.3.4 固定	2^{16} 65536	2^{32} 42.9億	2^{16} 65536	TCP 固定

Source / Client	Source Port	Destination / Server	Destination Port	Protocol
1.2.3.4(固定)	2^16	10.20.30.40	80(固定)	TCP(固定)
1.2.3.4(固定)	2^16	10.20.30.40	443(固定)	TCP(固定)
1.2.3.4(固定)	2^16	10.20.30.60	443(固定)	TCP(固定)
1.2.3.4(固定)	2^16	10.20.30.70	443(固定)	TCP(固定)
1.2.3.4(固定)	2^16	10.20.30.80	443(固定)	TCP(固定)
1.2.3.4(固定)	2^16	10.20.30.90	443(固定)	TCP(固定)

實際上呢?

- 可用記憶體
- 作業系統可開啟的最大檔案數量
- 檔案系統的inode數量
- `net.ipv4.ip_local_port_range` (32768 60999)
- etc

CLB ENI可以接收/發起多少連線?

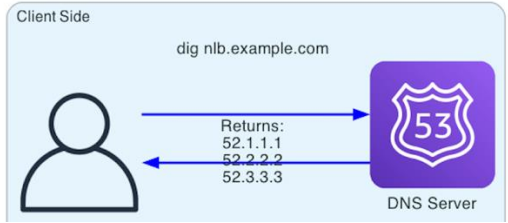


Agents -> CLB ENI Public

Source / Agent IP		Destination / CLB Public		Protocol
Source IP Client IP	Source Port Client Port	Destination IP CLB IP	Destination Port CLB Port	TCP
看有多少 agent	預設 32768~60999 約28232個	看有幾張 CLB ENI	443	TCP

CLB ENI Private -> EC2 Private

Source / Client		Destination / Server		Protocol
Source IP Client IP	Source Port Client Port	Destination IP Server IP	Destination Port Server Port	TCP
看有幾張 CLB ENI	預設 32768~60999 約28232個	看有幾張 EC2 ENI	8888	TCP



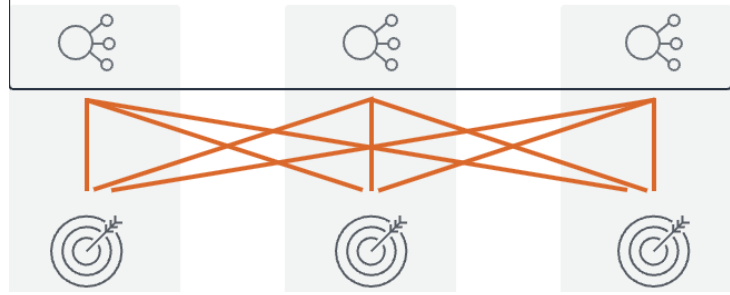
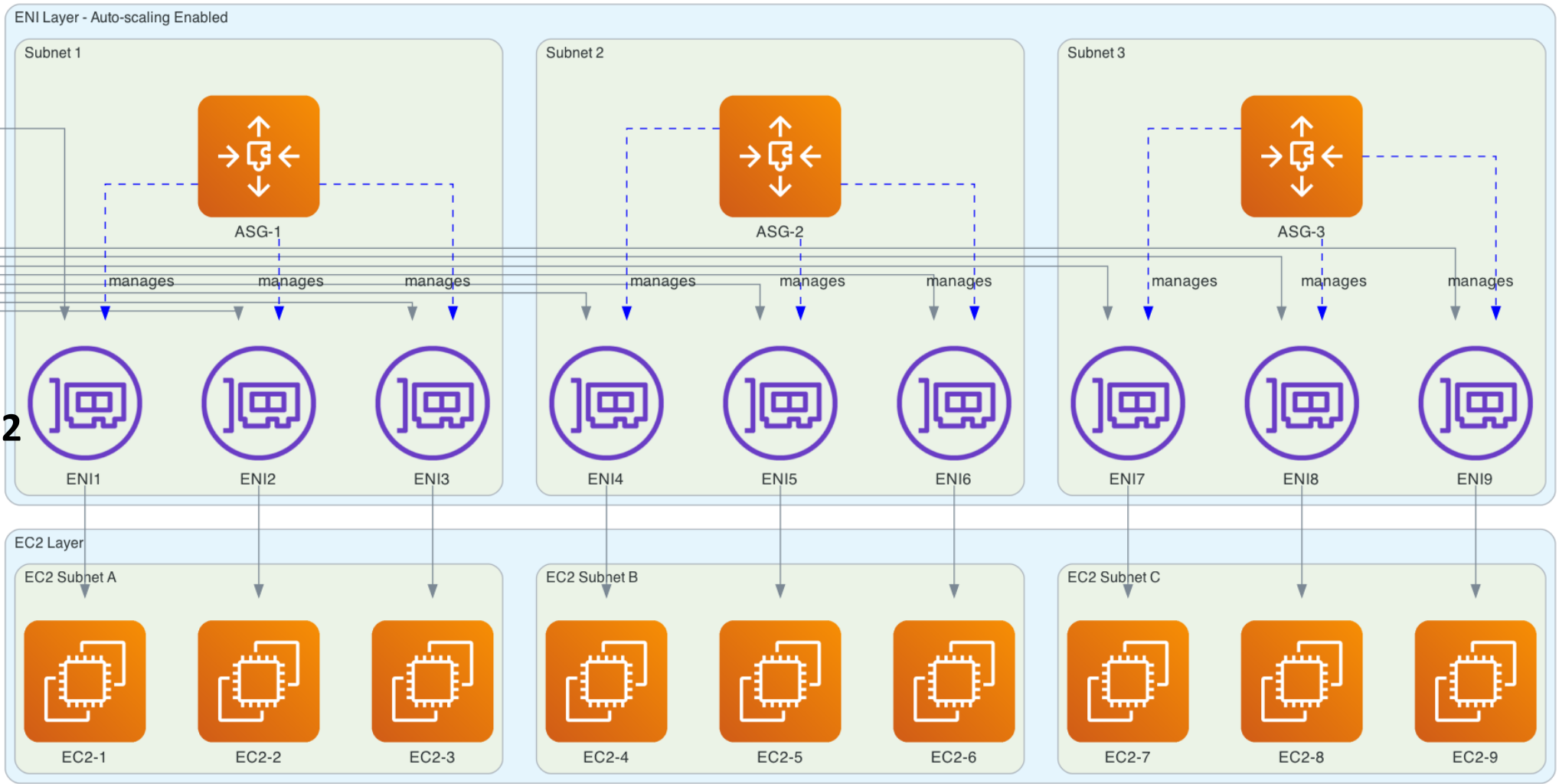
1M Connections



Classic Load Balancer

111k Connections x2
??? Connections

111k Connections
150k Connections





⌕ HTTP 503: Service unavailable

HTTP 503：服務無法使用

Description: Indicates that either the load balancer or the registered instances are causing the error.

說明：表示錯誤可能是由負載平衡器或已註冊的執行個體所引起。

Cause 1: Insufficient capacity in the load balancer to handle the request.

原因 1：負載平衡器的容量不足以處理請求。

Solution 1: This should be a transient issue and should not last more than a few minutes. If it persists, go to the [AWS Support Center](#) for assistance.

解決方法 1：這應該是暫時性的問題，通常不會持續超過幾分鐘。如果問題持續存在，請前往 AWS 支援中心尋求協助。

猜測是不是這個問題

- 不管CLB ENI背後原理，依照連線組成定義角度
- 假設問題是CLB來不及擴展ENI
- 或是 ENI可接收連線、可發起連線到上限，
- 導致CLB出現5XX
- HealthyCheck也是透過TCP，造成UnHealthy等連鎖問題
- AWS沒有公開CLB擴展的邏輯

觀察到的有趣現象

- Agent數量相同
- A Region CLB ENI 數量到 N 就上不去了
- B Region CLB ENI 數量會超過 N
- 為什麼?





Benefits of migrating from a Classic Load Balancer

從經典負載平衡器遷移的好處

Using a Network Load Balancer instead of a Classic Load Balancer has the following benefits:

使用網路型負載平衡器取代經典負載平衡器具有以下優點：

- Ability to handle volatile workloads and scale to millions of requests per second.
能夠處理波動性工作負載並擴展至每秒數百萬個請求。
- Support for static IP addresses for the load balancer. You can also assign one Elastic IP address per subnet enabled for the load balancer.
支援為負載平衡器設定靜態 IP 位址。您也可以為啟用的每個子網路分配一個彈性 IP 位址給負載平衡器使用。
- Support for registering targets by IP address, including targets outside the VPC for the load balancer.
支援以 IP 位址註冊目標，包括位於負載平衡器 VPC 之外的目標。
- Support for routing requests to multiple applications on a single EC2 instance. You can register each instance or IP address with the same target group using multiple ports.
支援將請求導向單一 EC2 執行個體上的多個應用程式。您可以使用多個連接埠，將每個執行個體或 IP 位址註冊到相同的目標群組。

Elastic Load Balancing 彈性負載平衡

Network Load Balancers 網路負載平衡器

What is a Network Load Balancer?

什麼是網路負載平衡器？

[Documentation 文件](#) > [Elastic Load Balancing 彈性負載平衡](#) > [Network Load Balancers 網路負載平衡器](#)

What is a Network Load Balancer? 什麼是網路負載平衡器？

PDF

RSS

Focus mode 專注模式

For TCP traffic, the load balancer selects a target using a flow hash algorithm based on the protocol, source IP address, source port, destination IP address, destination port, and TCP sequence number. The TCP connections from a client have different source ports and sequence numbers, and can be routed to different targets. Each individual TCP connection is routed to a single target for the life of the connection.

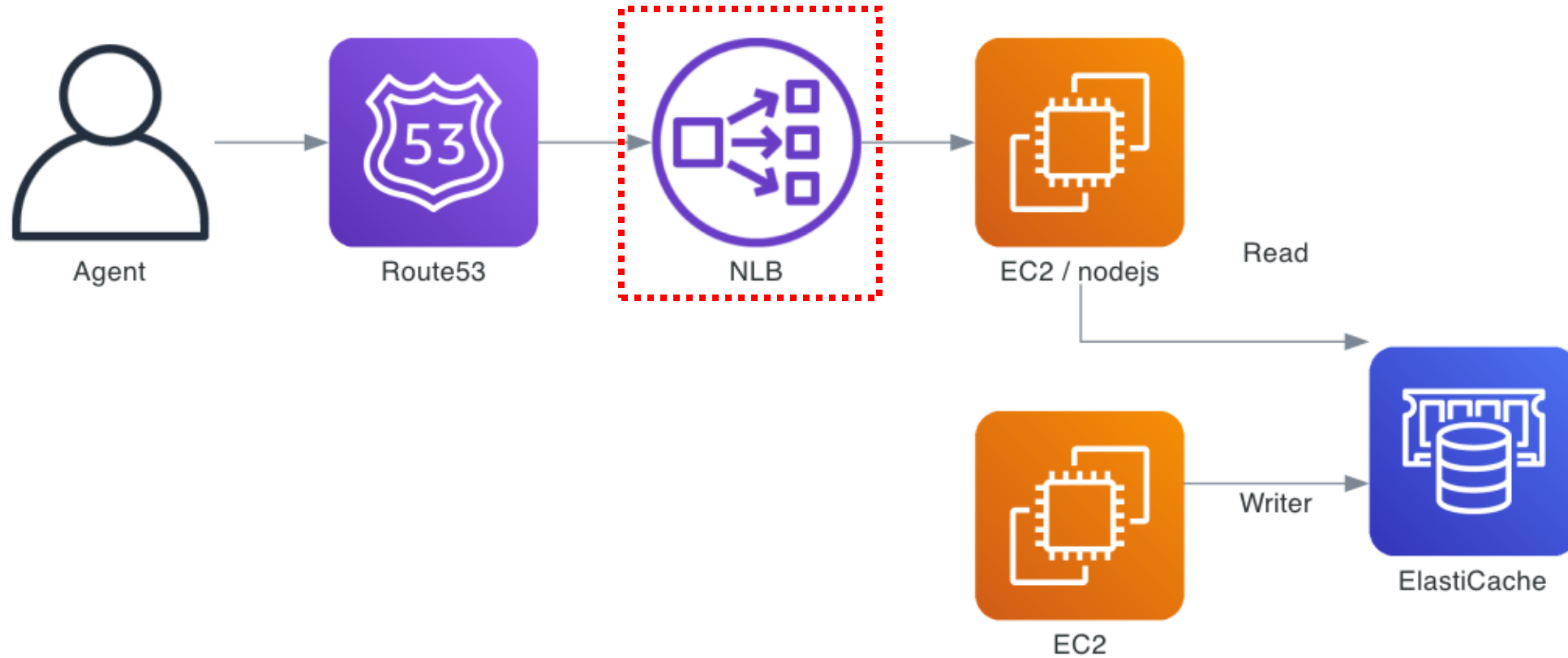
TCP 4-Tuple

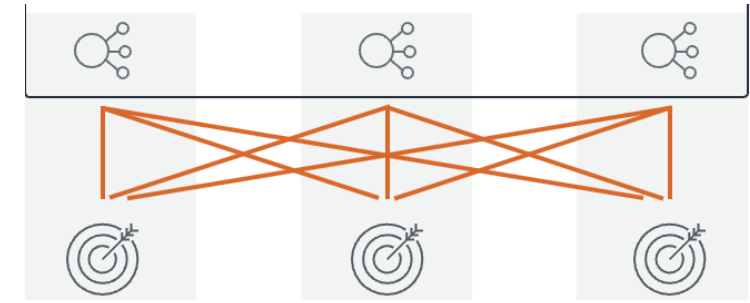
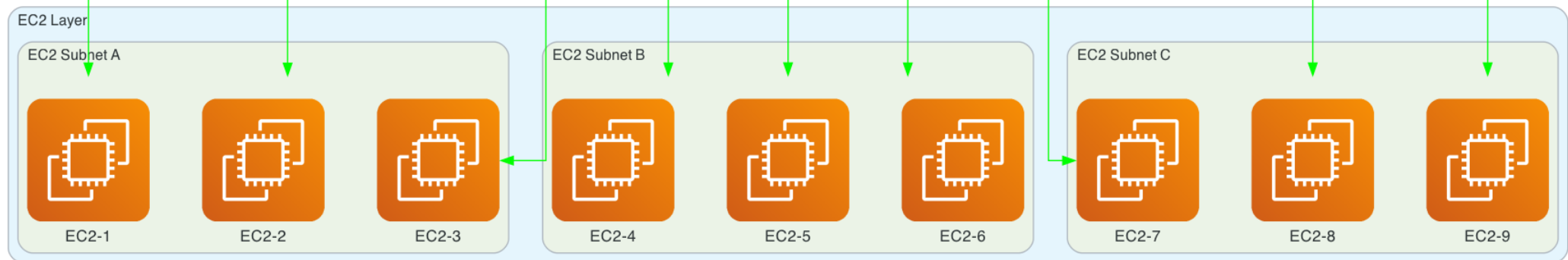
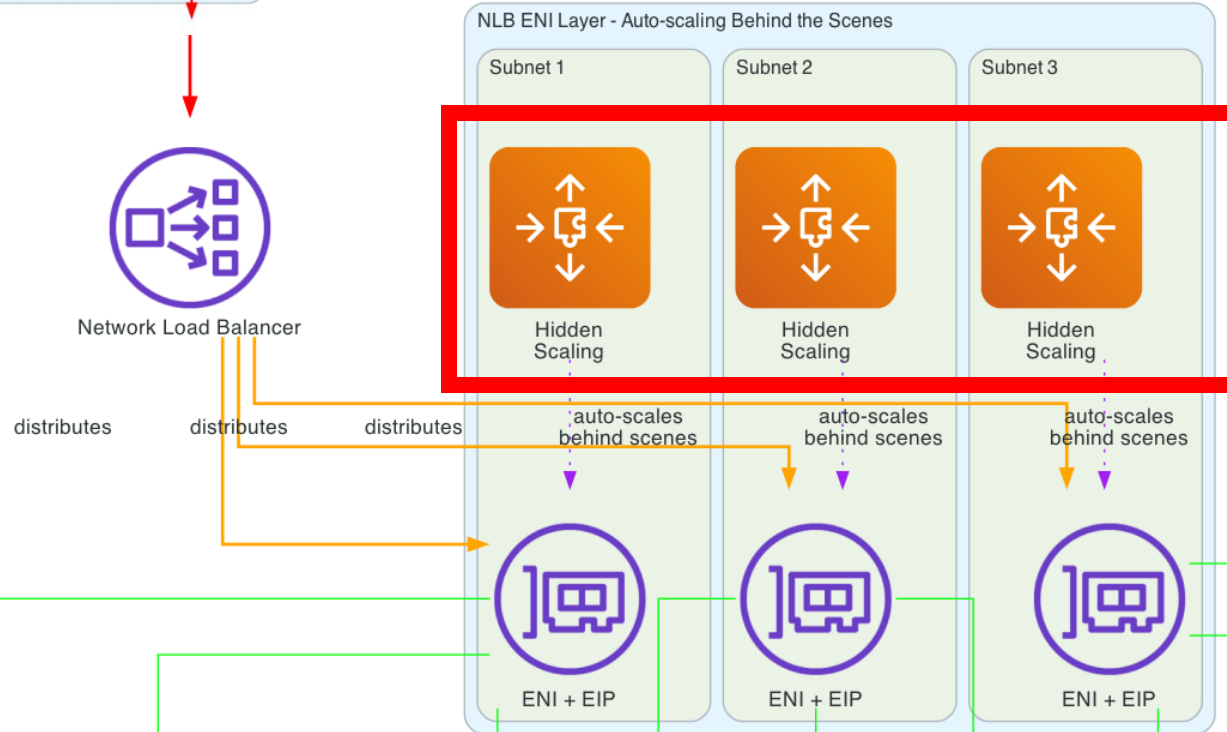
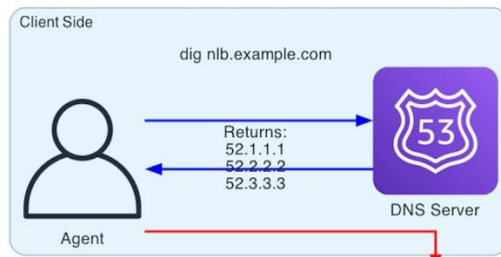
對於 TCP 流量，負載平衡器會根據協定、來源 IP 位址、來源埠號、目的 IP 位址、目的埠號以及 TCP 序列號，使用流量雜湊演算法選擇目標。來自同一用戶端的 TCP 連線會有不同的來源埠號和序列號，因此可能會被導向不同的目標。每一個獨立的 TCP 連線在其存續期間都會被導向單一目標。

You can add and remove targets from your load balancer as your needs change, without disrupting the overall flow of requests to your application. Elastic Load Balancing scales your load balancer as traffic to your application changes over time. Elastic Load Balancing can scale to the vast majority of workloads automatically.

您可以根據需求變化隨時新增或移除負載平衡器的目標，而不會中斷應用程式的整體請求流程。Elastic Load Balancing 會隨著應用程式流量的變化自動調整負載平衡器的規模。Elastic Load Balancing 能自動擴展以應對絕大多數的工作負載。

NLB







AWS文件並沒有說明NLB架構和背後技術

不過Google有公開他們的NLB論文

Google Research

Who we are ▾

Research areas ▾

Our work ▾

Programs & events ▾

Careers

Blog



[Home](#) > [Publications](#) >

Maglev: A Fast and Reliable Software Network Load Balancer

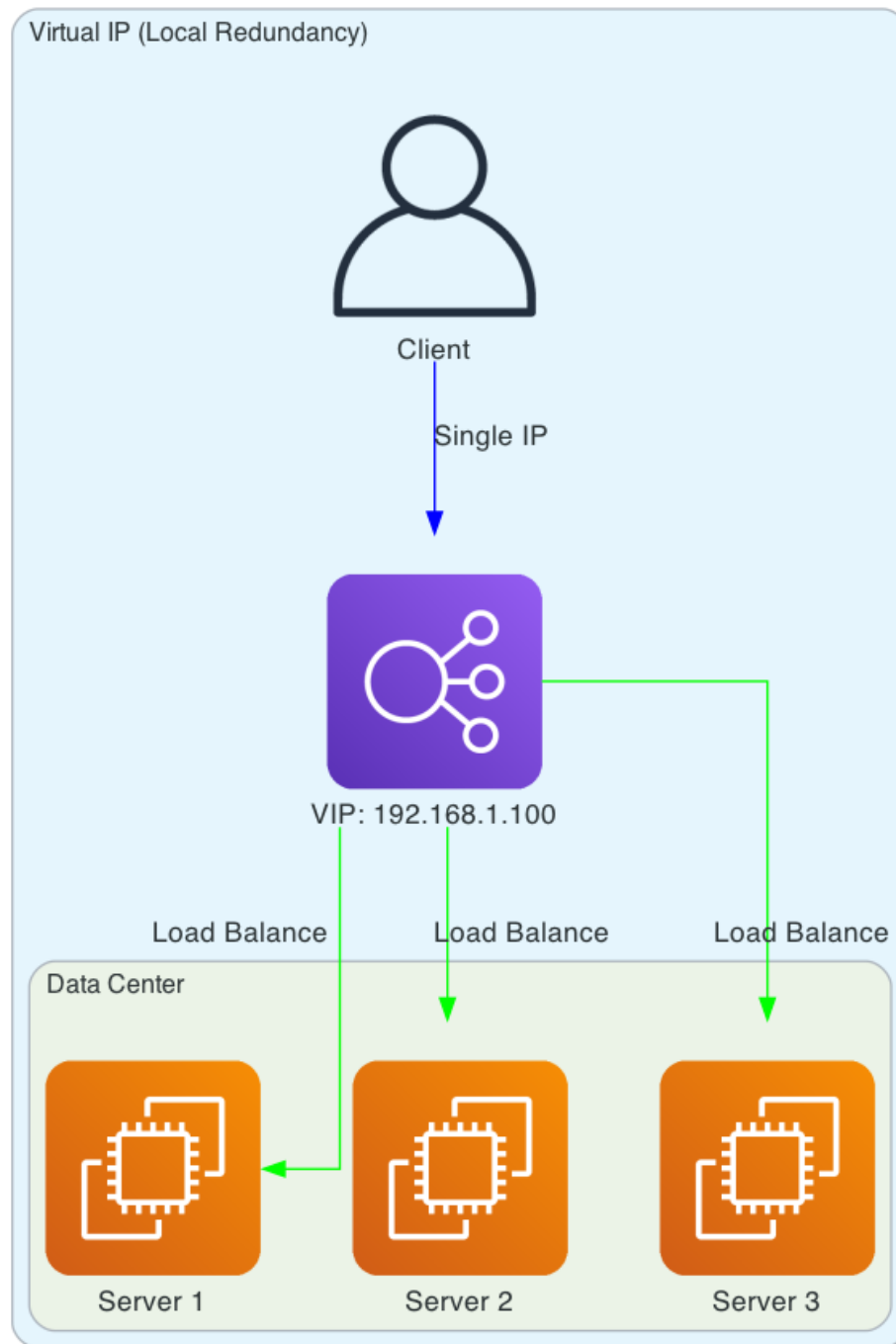
[Danielle E. Eisenbud](#) · [Cheng Yi](#) · Carlo Contavalli · Cody Smith · Roman Kononov · Eric Mann-Hielscher · Ardas Cilingiroglu · Bin Cheyney · Wentao Shang · Jinnah Dylan Hosein ·
13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16), USENIX Association, Santa Clara, CA (2016), pp. 523-535

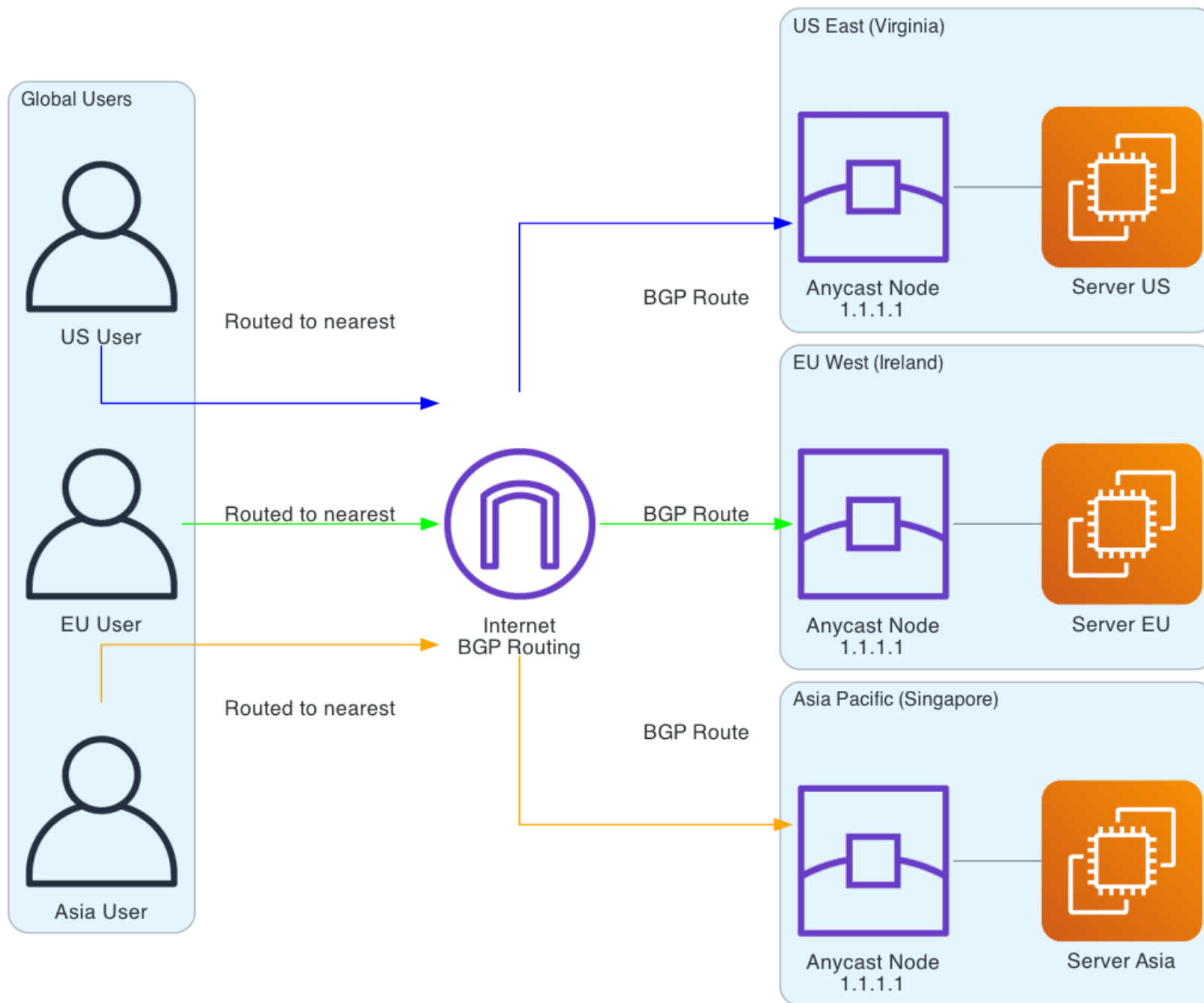
⬇ Download

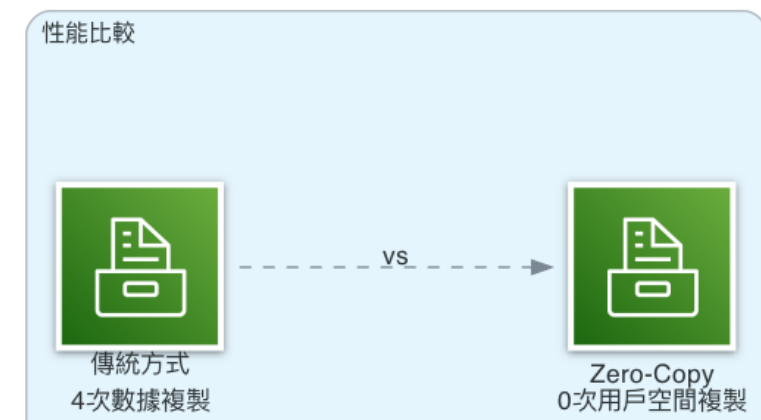
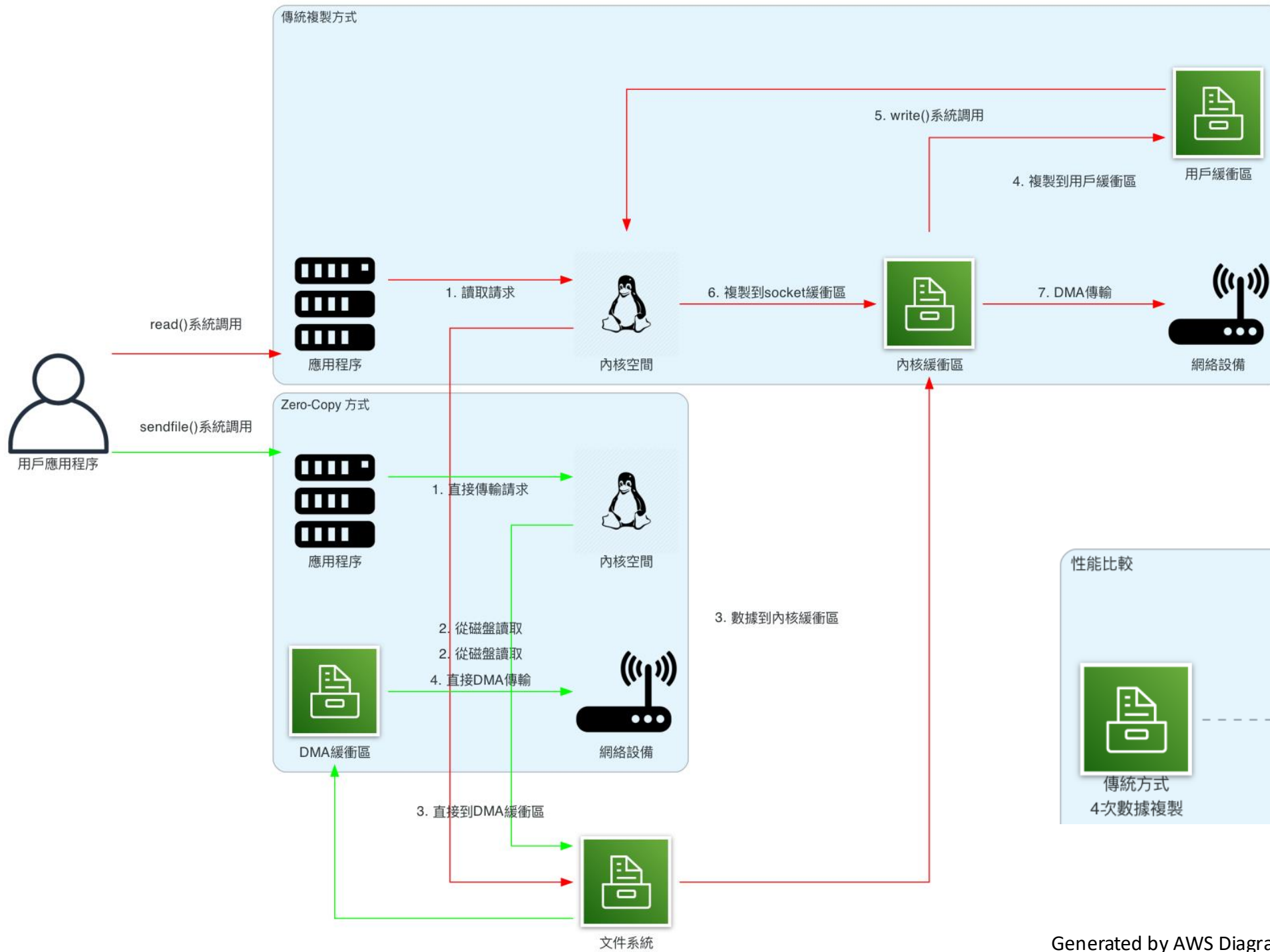
Google Scholar

Copy Bibtex

	Google NLB	AWS NLB
OSI Layer	L4	L4
靜態IP	Virtual IP (VIP)	AWS Elastic IP (EIP)
延遲	Anycast	
效能	Kernel bypass (類似 Zero Copy)	
黑魔法	Y	Y

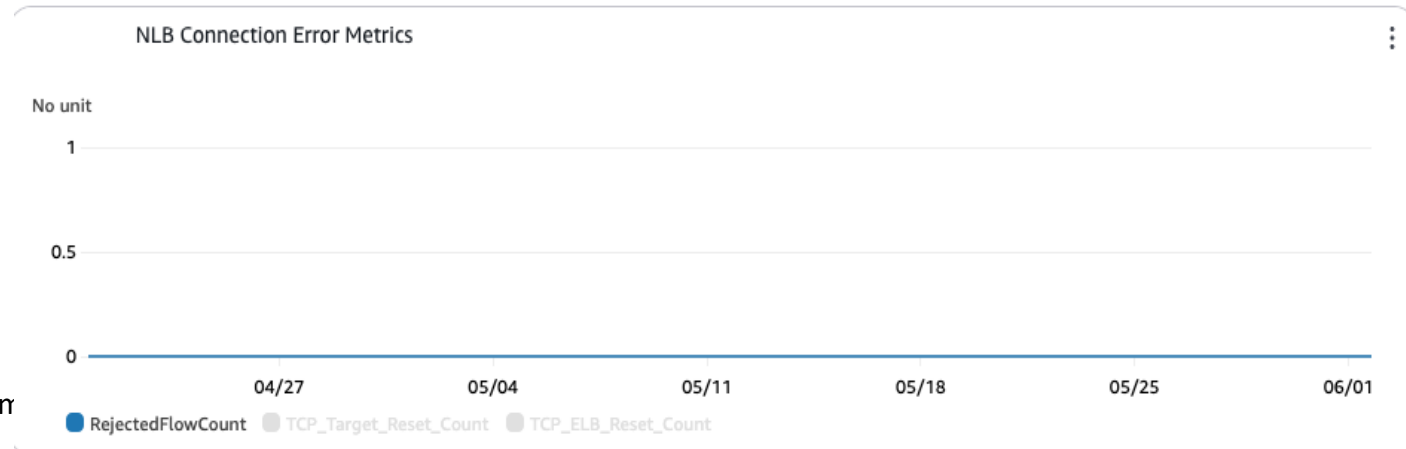
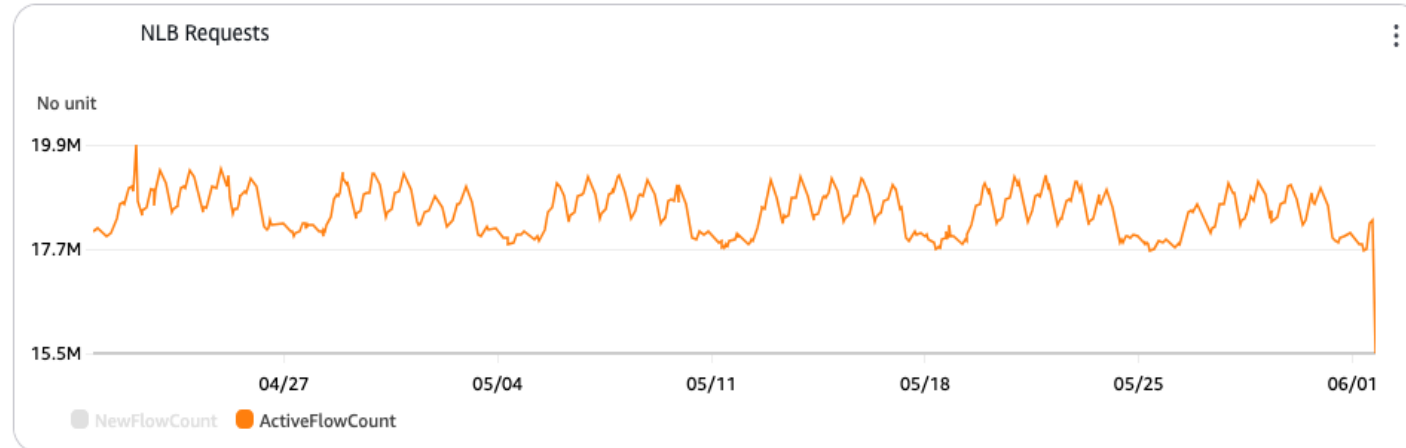
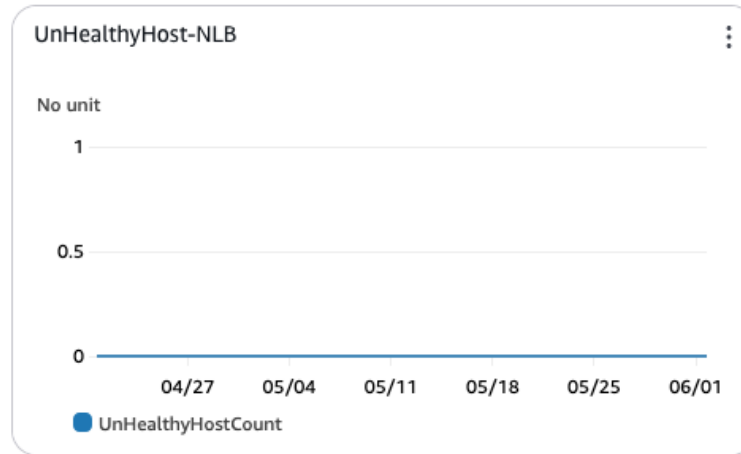








CLB -> NLB



Key Takeaways

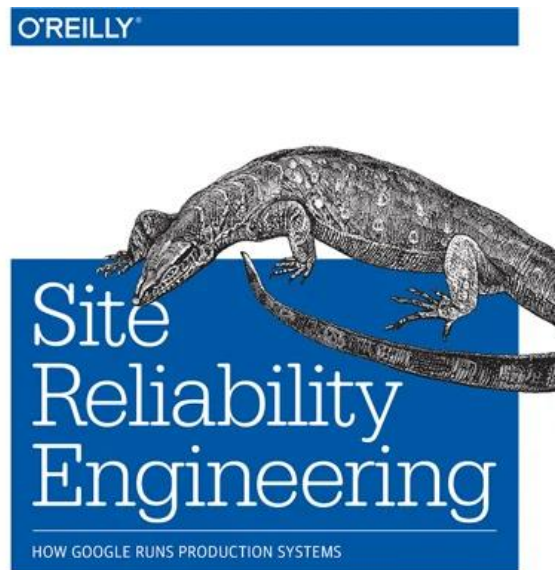
Google Research Who we are ▾ Research areas ▾ Our work ▾ Programs & events ▾ Careers Blog

[Home](#) > [Publications](#) >

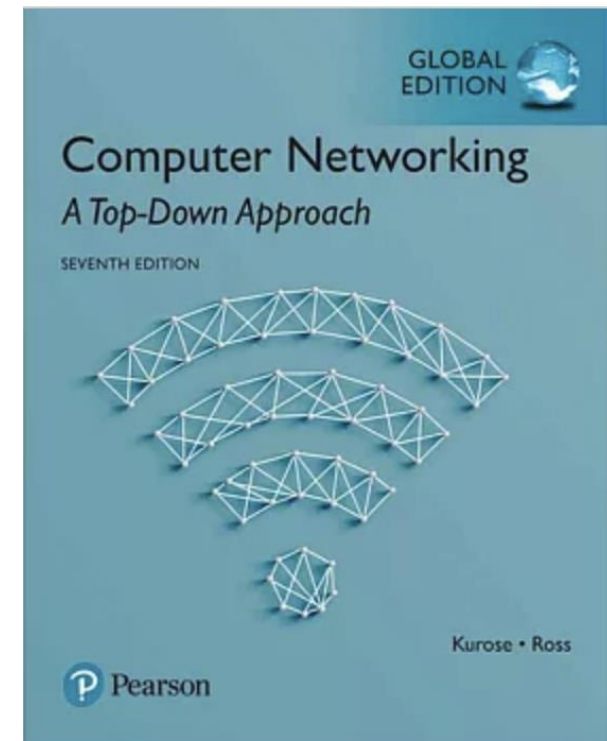
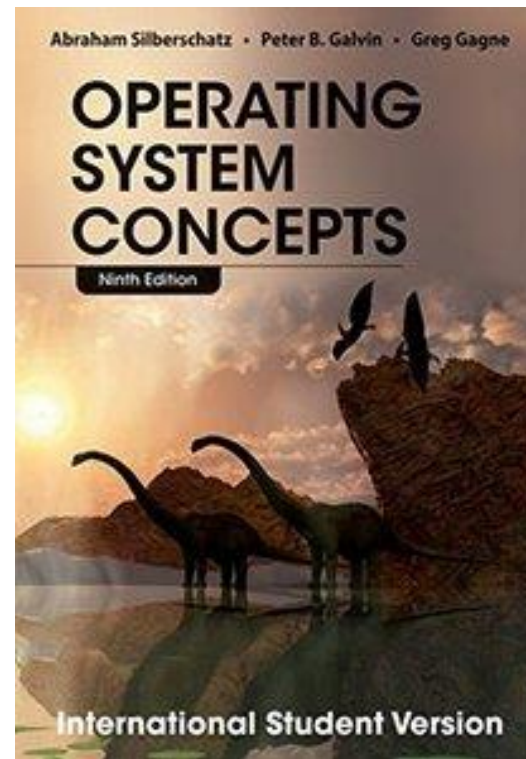
Maglev: A Fast and Reliable Software Network Load Balancer

[Danielle E. Eisenbud](#) · [Cheng Yi](#) · Carlo Contavalli · Cody Smith · Roman Kononov · Eric Mann-Hielscher · Ardas Cililingiroglu · Bin Cheyney · Wentao Shang · Jinnah Dylan Hosein ·
13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16), USENIX Association, Santa Clara, CA (2016), pp. 523-535

[Download](#) [Google Scholar](#) [Copy Bibtex](#)



Edited by Betsy Beyer, Chris Jones,
Jennifer Petoff & Niall Murphy





Thank You!