

# 支援 IP 群播樹區域快速重新繞送機制之設計與效能分析

蘇暉凱<sup>a\*</sup>、吳誌軒<sup>a</sup>、林信良<sup>b</sup>、徐茂馨<sup>a</sup>  
國立虎尾科技大學電機工程系<sup>a</sup>  
國立中正大學通訊工程所<sup>b</sup>

**摘要**— 近年來網路快速發展，即時性多媒體的應用越來越多，如：IPTV、MoD 與 Video Conference，網際網路群播服務可使用性 (Network Availability) 越來越受重視，因此也成為網路服務業者 (Internet Service Provider, ISP) 的重要服務品質參數。所以對於即時性的網路服務來說，因錯誤發生時而等待重新收斂的時間，對服務品質的影響會是很嚴重的，尤其是群播環境下，上游節點或是鏈結若出現損壞而遺失封包，會造成底下的接收者都遺失這個封包。因此，本論文以 IP Multicast Protection 為主題，在 Multicast 的環境下提出 Tree-Based 方法做單一鏈結的保護，並利用 ns2 模擬此機制與效能。

## 一、前言

由於現今網路快速的發展，讓即時性多媒體的應用快速增加，而大部分的多媒體傳出服務都利用群播方式 (IP Multicast) 傳送封包，IP Multicast 是一種結合 IGMP (Internet Group Management Protocol) 以及 Multicast Routing；IGMP [1] 是提供 End User 加入與離開群播服務，而 Multicast Routing 提供 Router 間交換 Multicast Group Information 與路由資訊，並建立 Multicast Tree 等協定，提供網路群播服務 (Multicast Service)。

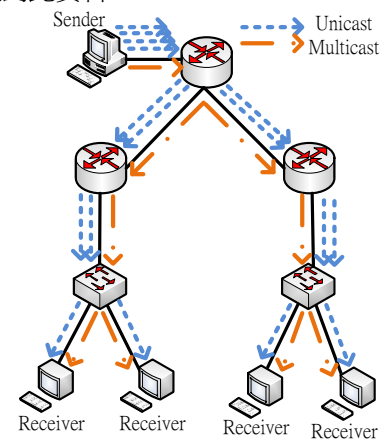
群播傳送一旦發生錯誤時可能會造成資料大量的流失，接而影響下游接收端的鏈結越多，所造成的影響也就越大，也因為如此，使得群播保護的議題越來越受到重視。網路服務中斷的原因可能是路由器之間的鏈結中斷或是網路設備損壞，對於網路服務中斷目前有兩種修復方式：IP 網路回復 (IP Restoration) 與 IP 網路保護 (IP Protection)[2]。IP 網路回復是當網路發生節點或線路中斷，網路觸發路由更新，透過 IGP (Interior gateway protocol) 交換新的網路資訊，最後讓每個路由器的路由資訊收斂到新的最佳狀態。IP 網路保護則是事先主動分析網路上的路由資訊，並計算出保護路徑，當服務中斷發生時，則將受影響的封包轉送至未受影響的路徑上來達到 IP 保護的作用。

我們可以發現，如果 IGP 收斂時間不夠快，往往無法立即反應網路最新狀態，對於 Real-Time Application 要求是不足的，例如：Video Conference 對網路中斷要求必須小於 200ms ~ 450 ms，Non-Real-Time Video Streaming Application 對網路中斷要求必須小於 1sec ~ 30 sec。目前 Multicast Tree 修復的標準作法，是當偵測出連線中斷後，立即觸發重建 Multicast Tree；因此，這種 IP Multicast Restoration 往往需要耗費許多 Multicast Tree 修復時間，並且造成 Multicast Transport Service 中斷。

IP Protection 與 IP Restoration (傳統 Multicast Routing Protocol 機制) 是不相衝突的，兩機制同時運作有加成效果。在群播路由協定收斂期間，由於 IP 網路無法保證每個路由器之路由表同步更新，因此網路會造成短暫 Micro-Loop 狀況。如果 IP Protection 與 IP Restoration 可以一起運作，當群播路由協定收斂完成後，也同時計算出每一 Multicast Tree 之 Backup Next Hop，一旦網路錯誤發生，在群播路由協定收斂與群播路由表更新期間，先暫時以 IP Protection 機制將受影響之封包繞送至相對應之 Backup Next Hop，一直到群播路由協定收斂與群播路由表更新完成後，路由器再切換回 IP Restoration 模式，根據新的群播路由表繞送封包，藉此縮短網路中斷時間與改善 Micro-Loop 問題。因此本篇論文提出 Tree-Based 方法在 IP Multicast 的環境下來對於群播樹做區域的快速路由，並利用 ns2 模擬此機制並觀察封包遺失和控制訊息量，來分析比較與原本的群播路由協定傳輸效能的差異性，最後模擬數據部分在第四章節會說明成果。

## 二、背景

在現行的網路環境中，傳送封包的方式大致上分為三種：單播 (Unicast)、廣播 (Broadcast) 與群播 (Multicast)。網路服務中大部分使用上都是以單播來進行封包傳送，單播就是以一對一的方式來傳送。如果同時要傳送相同的資料給不同的使用者時，傳送者就需要複製多份資料，分別傳送到不同的使用者，這樣會造成網路資源的浪費。因此群播 (IP Multicast) 技術提供了傳送者能透過網路同時向多個接受者傳送相同資料，如圖一，且傳送者不需要複製多份相同資料，只有特定的所需的群組會收到此資料。



圖一 SIP 的運作流程

## 2.1 IGMP (Internet Group Management Protocol)

網路群組管理協定 (IGMP)，在網路層中管理群組的成員加入/離開的協定，路由器利用此協定知道群組是在哪個網段上，目前最新的版本為 IGMPv3[3]。在第三版中主要特色是加入 Source Filter，當多個來源時，成員可以決定要接收哪個來源或不接收哪個來源，這樣可以讓 Multicast Source 傳送到有興趣的群組接收者上，不會傳送到不相關的使用者上以減少資源的浪費。

## 2.2 PIM-SM (Protocol Independent Multicast Sparse Mode)

PIM-SM [4]是一種不依賴特定單播路由協定的群播路由。PIM-SM 適用於群組成員較分散的網路上，主要特色是可依情況做群播樹的切換；群播樹的可分為二種：Source Based Tree (SBT) 與 Share Tree 兩種。Source Tree 是經由來源端到各個接收端所建立的最短路徑樹，如有多個來源則會產生多個群播樹；Share Tree 主要是將所有來源端的訊息集中於一個主要節點上，在由此節點將群播封包發送給底下的群組成員，好處是可節省許多路徑的維護，但缺點是流量都集中於一點，負擔會過重。PIM-SM 協定是由接收者來傳送 join 的訊息，中間所經過的路由器會記錄狀態，並將此訊息往集中點 (Rendezvous Point, RP) 傳送，而 RP 會將來源端所傳送的群播封包往剛剛所記錄的路徑傳送至接收端。而當上游的 DR (Designated Router) 察覺負擔過大時，則會切換成 Source Tree 的形式傳送。

## 2.3 群播保護機制

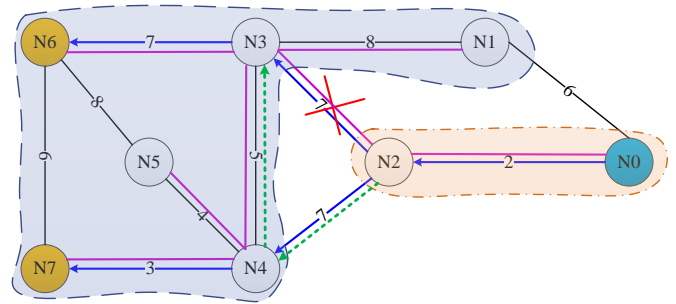
目前為止，從 IP 層與 Hop-by-Hop Routing (或 Destination Routing) 的特性來探討 IP Multicast Protection 研究議題之論文並不多，但主要可以分為兩類：備援群播路徑 (Backup Multicast Path) 與備援環 (Protection Cycle)。

1. 群播備援路徑 (Backup Multicast Path)：群播備援路徑最典型的作法是 A. Fei 等作者所提出之「Dual Tree」方法[5]，該想法是在 Multicast Network 上，找出多個 Link-Disjoint 或 Node-Disjoint Multicast Tree 與原本 Multicast Tree Disjoint，一旦發生連線或節點損壞，立即將受影響之封包轉送至 Link-Disjoint 或 Node-Disjoint Multicast Tree。此方法簡單，但有時受限於網路拓樸結構，在一個 Mesh Network 下，無法同時找到多個 Link-Disjoint 或 Node-Disjoint Multicast Tree；尤其在很多 Multicast Receiver 的狀況下，要找到多個 Node-Disjoint Multicast Tree 更不容易。
2. 備援環 (Protection Cycle)：備援環的想法是來自於 FDDI (Fiber Distributed DataInterface)，FDDI 使用 Dual-Attached, Counter-Rotating Token Ring Topology 來提供備援服務，但其缺點就是必須把整個網路所有節點連接成一個環狀。Pre-Configured Cycles (或 p-cycles) 的想法在[6]被提出；該方法避免實體拓樸必須連接成環狀，在 Mesh Network 中，可以事先設

定一個或多個 p-cycle，p-cycle 是一個邏輯上的 Ring，運作方式與 FDDI 類似，受影響的封包可以往 p-cycle 的反方向傳送，以繞過錯誤的連線，因此在 p-cycle 上的線路皆可以得到備援與保護；但是受影響的封包在 p-cycle 傳送，必須使用 Source Routing 的方式傳送，在純 IP 網路並不完全適用 (Shortest-Path Routing 之封包無法沿著 p-cycle 傳送)。

## 三、系統架構

一般 Multicast Routing 環境下，IP Multicast Source Tree 是由 Receiver 端 Trigger Join 所產生的 Reverse Shortest-Path Destination Tree。在此結構下，我們可以發現，該 Reverse Shortest-Path Destination Tree 與 Unicast 的 Destination tree 結構一樣，只有資料流方向相反。因此，參考 Hui-Kai Su [7][8] 等人的文獻，針對 Multicast 特性，我們提出 Tree-Based 的方式來做網路保護，以 PIM-SM 群播路由為例，以下將說明 Tree-Based 對鏈結保護之步驟。

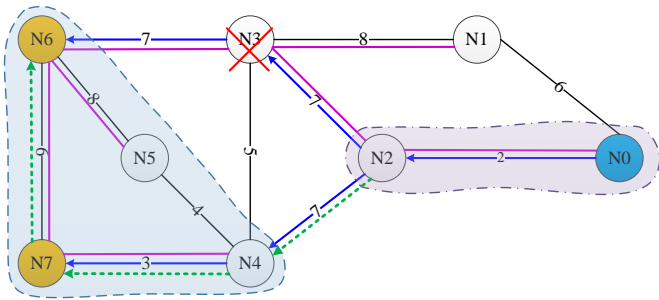


圖二 Multicast Cascaded Source Tree for Single Link Protection

圖二中，N0 是 Source node，N6 與 N7 是 Receiver；正常狀況下，藍色箭頭所形成的路徑為 Multicast source tree，亦為該群播資料流方向；此 Multicast source tree 是一種 reverse shortest-path destination tree。假設一旦鏈結 (N2,N3) 損壞，整棵 Multicast source tree 會被切成兩個 Sub-Tree，有 Sub-Tree(N2) 與 Sub-Tree(N4)。此結構下，在群播路由協定收斂前，只有 N2 會在第一時間知道鏈結損壞。因此，N2 將停止對 N4 Multicast 傳送封包；此時，只要 Sub-Tree(N0) 的 Leaf Node 可以將 N3 與 N4 串接起來，讓整個結構變成一棵不完整反向最短路徑目的地樹 (Non-Complete Reversed Shortest-Path Destination Tree)，那麼 (N2,N4) 鏈結損壞即可被保護。所以在群播樹建立後，因 pim join 訊息會往上游傳送，所以會知道下游節點的資訊，以圖二來看 N2 會知道下游有 N3 與 N4，所以 Tree-Based 機制首先會計算下一節點的 Destination Tree，N2 會計算 N3 的 Destination Tree (桃紅色線)，當鏈結 (N2,N4) 損壞時，N2 找尋鄰接節點並可利用計算好的目的地樹將受影響的封包來傳送到 N3，以圖二來看 N2 可以找鄰接節點 N4，而 N4 可利用目的樹傳送到 N3，此時 N2 會將群播封包外層加上 Unicast Header



並往 N4 傳送指定目的地為 N3(綠色虛線)，N3 接收到此封包後會將外層的 Unicast Header 解開，裡面的群播封包會再繼續往原本的群播路徑傳送，而鏈結就可被保護。



圖三 Multicast Cascaded Source Tree for Single Node Protection

上述是講 Tree-Based 對於單一鏈結的保護，接下來說明 Tree-Based 對單一節點的保護，節點保護與鏈結保護最大的不同在於節點若損壞則可能會有好幾條鏈結跟著受損，因此節點的保護會比鏈結的保護難度更高，要使網路存活性提高，節點保護的備援路徑會比鏈結保護來的難找，但如果節點可以被保護到代表鏈結一定也能被保護到。

環境與上述鏈結保護一樣，假設一旦 N3 損壞，整棵 Multicast source tree 會被切成二個 Sub-Tree，有 Sub-Tree(N0) 與 Sub-Tree(N4)。在 IGP 收斂前，只有 N2 鄰近的節點可以在第一時間偵測到節點損壞。因此，N2 的上游鄰近節點 (Aware Node) 必須負責將受影響的 Multicast 封包，快速轉送給備援下一節點；只要 Sub-Tree(N0) 的可以與 Sub-Tree(N4) 串接起來，讓整個結構變成一棵不完整反向最短路徑目的地樹 (Non-Complete Reversed Shortest-Path Destination Tree)，那麼 N2 節點損壞即可被保護。這邊作法與上述鏈結保護不太一樣，在節點保護上我們需增加 PIM join 的訊息，需讓 PIM join 訊息能夾帶下游接收者的資訊，以圖三為例，N3 會夾帶底下有 N6 的訊息並往 N2 傳送，這時 N2 會知道下一節點為 N3 與下一節點 N2，接下來 N2 會幫 N6 計算 Destination Tree (桃紅色線)，當 N3 損壞，N2 會開始找尋鄰近節點看是否能利用計算好的目的地樹將受影響的封包傳送到 N6，由圖三可知，N2 可以找鄰近節點 N4 且 N4 可以將封包利用目的地樹傳送到 N6，所以 N2 會將群播封包外層加上 Unicast Header 往 N4 傳送並指定目的地為 N6(綠色虛線)，N4 接收到後會依照目的地樹繼續往 N6 傳送，當 N6 接收到後會將外層的 Unicast Header 解開，裡面的群播封包會再繼續依原本的群播樹傳送，如此 N3 節點將可以被保護。

#### 四、 模擬與效能分析

##### 4.1 模擬環境

我們使用 ns2 網路模擬器來模擬，版本為 2.35，因 PIM-SM 在 ns2 中網路模擬行為不完全所以我們群播路由協定採用類似 PIM-SM 行為的 CtrMcast，其特點為集中管理時都會經由 RP 來傳送到 receiver。我們模擬時，從 0 秒預設每 0.5 秒就會有接收者加入群組，封包

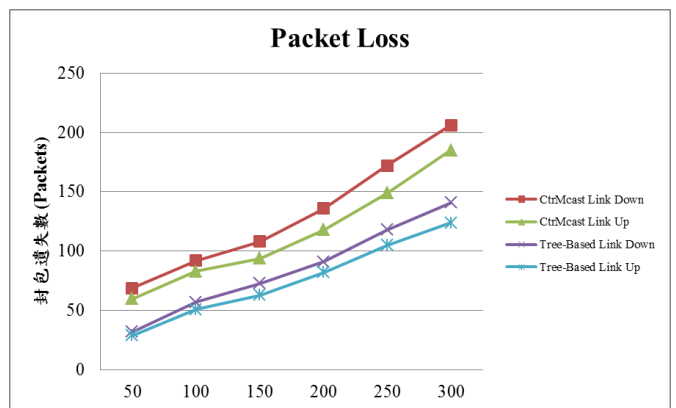
在第 3 秒時會開始傳送，並在第 8 秒時鏈結會中斷，第 25 秒時鏈結會回復，第 29.5 秒封包會停止傳送，而拓樸會從 50 個節點開始以每 50 個節點來增加直到 300 個節點，接收者的數量分為兩部分來看，一部分是依全部節點裡的 20% 為接收者，表示於網路中的低使用率；另一部是依全部節點裡的 60% 為接收者，表示網路中的高使用率，網路流量設定在平均 8 Mbps，這些不同的拓樸，我們將觀察針對封包遺失與控制訊息量，比較原本協定的修復與加入我們所提出的機制因路由更新所產生 miss route 的差別，以下是我們模擬時的參數。

- 模擬時間 (Simulation time)：30 秒
- 鏈結模式 (Link Connection Mode)：Duplex-link
- 頻寬 (Bandwidth)：100Mb
- 鏈結延遲時間 (Propagation delay time)：20ms
- 序列管理 (Queue management)：DropTail
- 群播封包型態 (Multicast packet type)：UDP
- 群播封包傳送型態 (Multicast packet traffic)：CBR
- 封包大小 (Packet Size)：1400 Byte
- 封包傳送間隔 (Rate)：0.0014 s

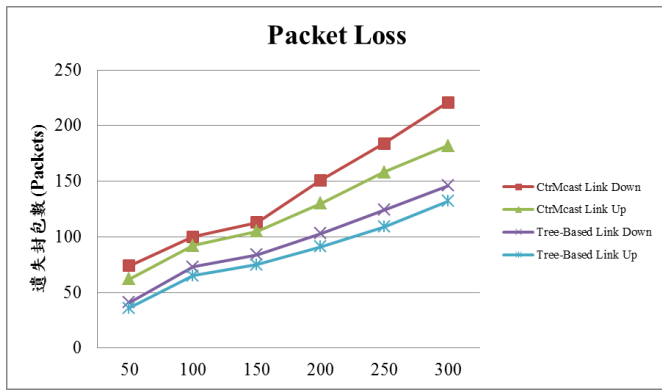
##### 4.2 封包遺失

封包遺失在群播環境中影響是會很大的，假設接收者有 5 位、RP 所傳送的封包有 10 個，在傳送的路徑當中有損壞，其中遺失了一個封包，這時所遺失的封包必須乘以 5，因為這些封包對於接收者是遺失的，因此我們定義這個叫 Effective Packet Loss。

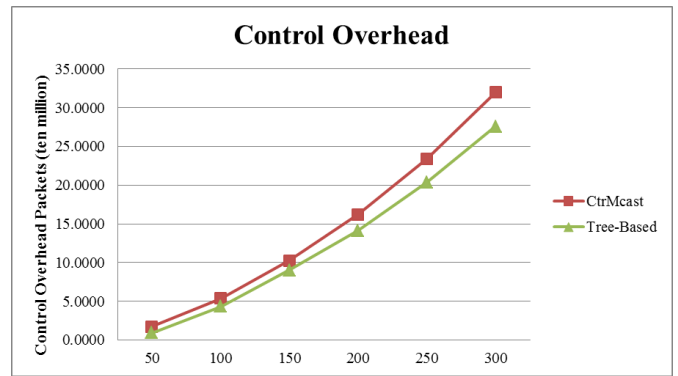
圖四與圖五 中，紅線方形線是代表是 CtrMcast 協定在鏈結損壞時所造成封包掉落數量的曲線，而綠色三角形線是代表那條損壞鏈結恢復時所影響的封包掉落數量的曲線；紫線是我們所提出的 Tree-Based 機制在鏈結損壞時所造成的封包遺失曲線，藍線是鏈結恢復時所影響的封包掉落數量曲線。因此可看出當網路越大鏈結的損壞也會造成越大的影響，當鏈結損壞時，ns2 會馬上更新路由資訊所以會造成一些封包因路由表更新的 miss route，所以我們的機制如有保護成功則會將這些受影響的 miss route 封包轉送到備援路徑上，所以原本的協定會造成這些封包找不到路而被丟棄，這些封包的遺失對於即時性的網路應用服務來說，是會很嚴重的。



圖四 接收者佔全部節點 20% 時的 Packet Loss



圖五 接收者佔全部節點 60% 時的 Packet Loss

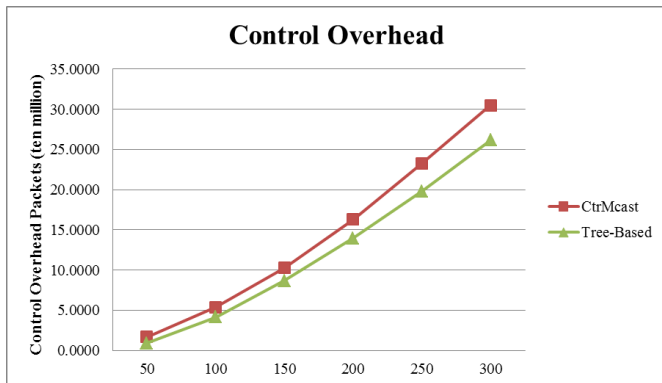


圖七接收者佔全部節點 60% 時的 Control Overhead

### 4.3 控制訊息

控制訊息封包是路由協定所發送的，除了定期的發送來維持網路的狀態，當網路有發生改變時，如節點或是鏈結損壞，則損壞地方旁的節點會最先知道，也會發送控制訊息封包來告知其他節點損壞之訊息之後再更新路由表資訊，所以我們在這觀察當鏈結或結點損壞時網路上所發送的控制訊息封包量，因為這些控制訊息封包如果數量多的話，有可能會造成網路的壅塞而影響路由協定的收斂時間，也會影響封包的傳送，可能會使封包遺失。

控制訊息封包是路由協定所發送的，除了定期的發送來維持網路的狀態，當網路有發生改變時，如節點或是鏈結損壞，則損壞地方旁的節點會最先知道，也會發送控制訊息封包來告知其他節點損壞之訊息之後再更新路由表資訊，所以我們在這觀察當鏈結或結點損壞時網路上所發送的控制訊息封包量，因為這些控制訊息封包如果數量多的話，有可能會造成網路的壅塞而影響路由協定的收斂時間，也會影響封包的傳送，可能會使封包遺失。



圖六 接收者佔全部節點 20% 時的 Control Overhead

## 五、 結論

未來網路只會更快速，應用也會越來越多，所以對於網路服務品質的保護會更顯得重要，群播協定所建立的群播樹若中間發生錯誤所造成的影響會比單播的環境來得大，尤其越靠近根節點影響會越大，因此群播的保護是非常重要的。

由第四章所模擬幾種環境中的圖來看，當網路越大假使鏈結損壞的剛好是在群播樹的路徑上，所掉落的封包會越多，而因鏈結或節點損壞則會觸發 Unicast 路由協定，讓 Unicast 的路由協定來發送訊息狀態封包並重新收斂，所以所傳送的訊息狀態封包也會越多，這些訊息的傳送每經過一個節點都會有個延遲時間，群播路由協定也須等待 Unicast 的路由協定收斂完成才會開始群播樹的收斂，而這樣則會導致網路回復的時間變長，影響網路服務的傳輸。也因此希望在錯誤發生前，能計算好備援路徑，在發生錯誤時能將受影響的封包轉送至備援路徑上，將網路服務的影響降到最低。

本篇論文我們所提出的機制是針對於 PIM-SM 的共享樹來做保護，未來可對於來源樹的部分來保護並結合共享樹保護的機制，除此之外還可分析在群播網路服務中對於網路整體的保護性並與其他保護機制來做比較。

## 參考文獻

- [1] W. Fenner, "Internet Group Management Protocol, Version 2," RFC 2236, Nov. 1997.
- [2] C. Metz, "IP protection and restoration," IEEE Internet Comput., vol. 4, no. 2, pp. 97-102, Mar-Apr 2000.
- [3] B. Cain, S. Deering, I. Kouvelas, B. Fenner and A. Thyagarajan, "Internet Group Management Protocol, Version 3," RFC 3376
- [4] B. Fenner, M. Handley, H. Holbrook, and I. Kouvelas, "Protocol independent multicast - sparse mode (PIM-SM): Protocol specification (revised)," RFC 4601, Aug. 2006.
- [5] A. Fei, J.-H. Cui, M. Gerla, and D. Cavendish, "A dual-tree scheme for fault-tolerant multicast," in Proc. ICC, 2001.
- [6] W. D. Grover and D. Stamatelakis, "Cycle-oriented distributed preconfiguration: Ring-like speed with mesh-like capacity for self-planning network restoration," in Proc. ICC, 1998.
- [7] H.-K. Su and C.-S. Wu, "Local link protection scheme in IP networks," LNCS 4490, May 2007.
- [8] H.-K. Su, C.-S. Wu, and Y.-S. Chu, "IP local node protection," in Proc. of the Second International Conference on Systems and Networks Communications (ICSNC 2007), 25-31 Aug 2007.