Service Affinity Solution for TCP based Application

[draft-wang-tcpm-tcp-service-affinity-option]

Wei Wang (China Telecom)
Aijun Wang (China Telecom)
IETF 119, Mar. 2024
• Background and motivation of this draft
• Considerations on the existing solutions
• Proposed Solutions
• Further Actions
Background and motivation of this draft

Anycast Services Scheduling scenario

- Service nodes provide the same service function may be deployed different resource pools, and use the same anycast IP address.
- When customer A accesses to the service:
  - Customer A sends packet to R1
  - R1 determine the optimal service node for customer A based on the network status.
- The network status is constantly changing, different customers may be scheduled to different service nodes. For customers who have established connections, the service node providing services must remain unchanged.
Background and motivation of this draft

• The current solutions need to maintain the customer-based connection status table in ingress and egress routers.

• We propose a solution for the service affinity between client and server based on one newly defined TCP Option, which can realize the comprehensive scheduling based on real-time network status. This solution eliminates the need to maintain customer-based connection status tables for network devices, and improves the flexibility and scalability of large-scale deployment of anycast services scheduling.
Considerations on the existing solutions

- **Load-balancer**
  - This deployment may lead to the load balancer becoming the bottleneck when the traffic increases.
  - Direct traffic redirection and traffic scheduling between the client and server can avoid the bottleneck of load balancer.

- **MPTCP**
  - It enables hosts to send packets belonging to one connection over different paths, but it is confined to the MPTCP framework.
  - We want to find one solution that can meet such requirements in a more general manner for TCP-based applications.

- **HTTP redirection**
  - It mainly involves the communication between client and server.
  - Both client and server do not perceive changes in network status and cannot achieve comprehensive optimization based on network status and computing resource status.

- **DNS redirection**
  - The caching time of DNS records will take some time for the modification to take effect.
  - Customer experience will be affected.
Proposed Solution

Anycast Services Scheduling scenario

Procedures for the service affinity solution

A new Flag (“SAF”) is requested for identify the sender supports TCP Service Affinity Option.

1. Customer A accesses to the service.
2. R1 schedules the request and determines service node behind R4 will provide the service.
3. Service node returns its IP address and port information
4. Customer A re-establishes the connection to service node behind R4.
Proposed Solution

- **Type (1 octet):** identifies the newly defined TCP Option, which is allocated by IANA.
- **Length (1 octet):** identifies the length of the TCP Option.
- **(IPv4 Address, Port) (6 octets):** identifies the IPv4 address and port owned by the service node that provides the service.

- **Type (1 octet):** identifies the newly defined TCP Option, which is allocated by IANA.
- **Length (1 octet):** identifies the length of the TCP Option.
- **(IPv6 Address, Port) (18 octets):** identifies the IPv6 address and port owned by the service node that provides the service.

This Option is carried in the TCP FIN packet sending by the service node, and the address carried must be the address owned by the service node. After receiving the TCP FIN packet, if this TCP Option is included in the packet, the customer will establish the connection to the address specified in this Option.
Proposed Solution

Security considerations
The attack source may send TCP packets with SAF to a service node to obtain its unicast IP address, thereby illegally obtaining information on the server, or launching DDoS attacks on the service node.

• **DDoS attacks**
  – To avoid DDoS attacks, traffic accessing service nodes can first pass through the firewall.

• **Traffic hijacking**
  – To avoid information theft on the server, users and sites accessing the network can be authenticated and verified.
  – CATS solutions for various network attacks: [draft-li-cats-attack-detection]
Further Action

• Comments?
• Request for Adoption Call

wangw36@chinatelecom.cn
wangaj3@chinatelecom.cn
IETF119