

# HTTP Message Steering



Lori MacVittie, 2013-01-04

*It's like content steering, only different, and it's part of the application layer #SDN.*

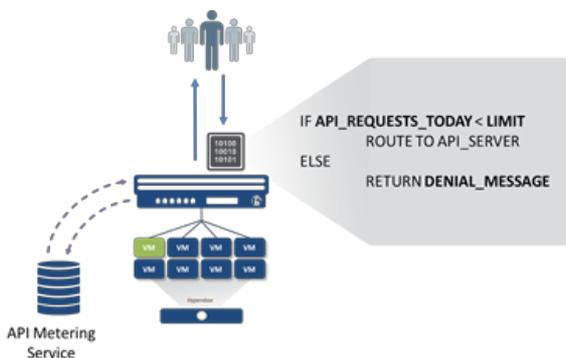


If you're immersed in the world of service providers (sometimes called carriers, telecoms, or carrier service providers) then you've probably heard the term "traffic steering".

Traffic steering is the ability to control and direct data and voice traffic to the best suitable cell layer and radio technology within any network. It becomes a key feature in LTE networks deployed over multiple layers, such as frequency layers and/or hierarchy layers as macro, pico or femtocells. Those networks make available resources from multiple layers to an end-user in a certain geographical area. Therefore, traffic steering could optimise network capacity and user experience via an efficient utilisation of the whole pool of available resources (e.g. radio, transport, backhaul).

-- LTE SELF-ORGANISING NETWORKS (SON)

A similar concept, HTTP Message Steering, is used in data center networks to provide this same type of capabilities with respect to routing HTTP requests and responses to services as determined to be appropriate based on operational and business requirements.



Now, more astute readers will immediately ask the question, "How is this different from content switching (or content steering, if we want to adopt modern nomenclature)?"

It's a good question, and one that deserves an answer. Content steering executes very similar logic in that it is able to determine how to route or handle HTTP requests based on the context of the request.

The thing with content steering (or page routing or whatever the nom de jour might be this year) is that it's self-contained.

Whatever variables are used to determine how to route the request had better be either in the content, on the system, or able to be deduced from the connections between client and intermediary, intermediary and client.

HTTP Message Steering, like its cousin traffic steering, implies the use of an external, off-the-box service to augment the data upon which the routing decision is made. If we take (again, because I like this use case) API metering, it is fairly obvious that the intermediate proxy is *not* going to be storing the data necessary to determine whether a given request will exceed a user's quota. Proxies store session-related data, but do not weight themselves down with the mechanisms and memory needed to store long term data. A failure, remember, means anything in memory necessarily disappears, and something like API usage metrics is not something you want lost. It's longer term, it exists outside any given session. It's persistent, in developer-speak, and thus requires a more permanent means of storage to ensure it continues to exist across sessions.

Thus, if an intermediate proxy requires persistent data to make a routing decision, we call it message steering to differentiate it from the more near-term, session-constrained content steering. The differentiation is necessary because it has architectural ramifications that need to be considered. If you want to be able to execute HTTP message **steering** then you are going to need a proxy. Period.

That's because only a true (full) proxy terminates (acts as the client destination) the underlying TCP connection. Because there is a clear separation between client-side and server-side, a proxy can steer a message in any direction - including to a third-party persistent data service. A solution that is not a true proxy can only perform content steering, because it doesn't have the mechanisms required to initiate a connection on its own; it's purely a seamstress, stitching up a connection between the client and the application service and maintaining the flow between them. A proxy intercedes, it intercepts, it inspects and redirects if necessary. It can take other action because technically it *is* the destination endpoint.

HTTP Message Steering is a component of an application layer SDN; programmatically enabling the routing of requests and responses to services in the data center fabric based on policy. It is one of the core technologies comprising an application layer SDN and the underlying nature of a proxy-based application network service is the foundation upon which such solutions are built.



---

F5 Networks, Inc. | 401 Elliot Avenue West, Seattle, WA 98119 | 888-882-4447 | [f5.com](http://f5.com)

F5 Networks, Inc.  
Corporate Headquarters  
[info@f5.com](mailto:info@f5.com)

F5 Networks  
Asia-Pacific  
[apacinfo@f5.com](mailto:apacinfo@f5.com)

F5 Networks Ltd.  
Europe/Middle-East/Africa  
[emeainfo@f5.com](mailto:emeainfo@f5.com)

F5 Networks  
Japan K.K.  
[f5j-info@f5.com](mailto:f5j-info@f5.com)