

*UDP-based Application Layer Protocol Recipes to the Rescue*

Mark Nottingham <mnotting@akamai.com>, Miroslav Ponec <mponec@akamai.com>

Akamai Technologies

It is entirely appropriate to design UDP-based protocols when TCP is inadequate to meet the needs of the application. This is not a huge leap for protocol design, but what is different is that more applications are encountering what they perceive as limitations of TCP. This is especially true for HTTP/2, where connection setup latency, congestion control reaction to loss, and in-order delivery pose serious problems.

It is possible to address many of the issues in TCP itself by extending the protocol (e.g., with TCP Fast Start, new congestion control algorithms, TCP Minion, respectively). However, as is well-known, deploying these measures into existing networks and onto existing platforms is glacially slow. Thus, protocols built on top of UDP provide immediate solutions for deployment in today's Internet.

Perhaps the biggest question facing the community is how to facilitate this transition (where appropriate). We strongly believe that defining a new, generic transport layer(s) on top of UDP is destined to fail. A one-size-fits-all solution presumes we have good understanding of the problem space, and that we can balance concerns from different applications (many of them unknown). This is boiling the ocean, or perhaps several.

Instead, we believe that we have a much better chance of success if individual protocols make the choice to change their underlying transport independently, and when performing the design work, are able to rely on resources from the Transport Area regarding issues like fairness, congestion control, reliability methods, integrity verification, signaling, and so on. In other words, the specification deliverables spurred by this workshop should at most be "small pieces loosely joined" specs that application protocols can choose to use if they are suitable, along with human resources to advise on their use. Even then, an application might not be required to use them, if its community believes that another approach is better. It can at least serve as an inspiration, aid learning from common pitfalls, and support best practices, especially in security considerations area.

This approach does not deny that there are significant issues inherent in this area; moving away from TCP's rich body of knowledge and capabilities means that some ground will need to be re-tread (e.g., fairness). However, with proper support from the Transport Area, this should be manageable.