Distance-Agnostic, Application- and Resource-Aware Transport for Next-Generation Networks

NAME: Chase Wu, Raj Kettimuthu
AFFILIATION: New Jersey Institute of Technology, Argonne National Laboratory
EMAIL: chase.wu@njit.edu, kettimut@anl.gov
EXPERTISE: high-performance networks, transport control, big data, security

KEY DOE NETWORK/TRANSPORT CHALLENGE

Extreme-scale e-science applications within DOE are generating colossal amounts of data, now frequently termed as "big data", which must be transferred over long distances for remote operations. There are several collaborations that span from coast-to-coast within the U.S. and also across continents. The need to move data rapidly at such distances will become far greater ten years from now. We list a few key network/transport challenges below:

i) Can we build distance-agnostic transport protocols?

The data transfer throughputs achieved with current transport protocols are inversely proportional to the round-trip time between the end hosts, which makes it hard to achieve satisfactory performance over long distances. Can we design transport protocols that are distance-agnostic? Should we make each hop responsible for delivering data to the next hop? Would it change the notion of end-to-end? How do we handle failures?

ii) How to make network and transport solutions science application-aware?

E-science activities have specific performance requirements or metrics, e.g. a large file transfer may only care about the total turnaround time and the rate of progress; a scientific workflow in a data streaming application may need a data transfer rate within a certain range; a computational steering operation may require a smooth and stable data flow for parameter control on a remote high-performance computing or experimental facility. So far, we have left it to applications to identify appropriate transport methods to cater to their needs. Why don't we move that intelligence into the network, as the network has more knowledge about its state than the application? If so, how do we build network and transport protocols that are aware of application requirements and make intelligent decisions while making routing decisions and establishing connections? Should we augment existing protocols or build new protocols?

iii) How to design transport protocols for multiple dedicated connections?

The extremely high capacity and low (or zero) congestion within high-performance networks (HPNs) have posed unprecedented challenges to the design of transport protocols, particularly when a pool of them are needed, for example, data and control connections between geographically dispersed users controlling a complex computation on a supercomputer. In such cases, data transfers may not be bottlenecked by the network but rather by the host and storage systems and the performance achieved may be far less than the minimum performance of the individual components measured individually. But still, the current models for transport protocols predict infinite or maximum throughput in these cases as there is zero packet loss. Transient load variations on host and storage systems may make it very difficult to maintain host data rates that match connection bandwidths. Should transport protocols treat network bottlenecks and host/storage bottlenecks differently? If so, how? Is there a need for interactions between hosts, storage systems and networks?

DOE RESEARCH CHALLENGE DIRECTION

Maximizing the application-level transport performance over complex high-speed connections is (and will be) challenging because: i) the optimal operational zone is affected by the configurations and dynamics of the network, host/storage, and the transport protocol and ii) application users typically do not have the necessary knowledge to choose the appropriate transport protocol and parameters. It seems that the future network and transport protocols have to be application- and resource-aware. However, there are several challenges: How do we design extensible network and transport protocol APIs to enable applications to express performance requirements? What kinds of interactions are required among network devices to achieve this? How do transport protocols interact with hosts and storage systems? Can we use profiling techniques? How do we minimize overhead? How do we determine appropriate transport parameters based on application requirements and end-system capabilities? Can we dynamically adapt transport methods according to time-varying load conditions?