1550 Larimer Street, Suite 168
Denver, CO. 80202

Contact: Kaleb A. Sieh, Deputy Director, ksieh@bitag.org

BITAG Publishes Report on Congestion Management

Denver, CO (October 22, 2013): Today, the Broadband Internet Technical Advisory Group ("BITAG") announced the publication of its technical report on the subject of Congestion Management. The executive summary and recommendations of the report are attached here, and the full report itself can be found at: http://www.bitag.org/documents/BITAG_-_Congestion_Management_Report.pdf.

Every link and router in the various networks that make up the Internet has a limit on its capacity to handle data. When aggregate user demand at any point in time exceeds the capacity of a link or router, the result is congestion, which can degrade performance. Choices made by a variety parties, including ISPs, ASPs, and CDNs may alleviate or exacerbate congestion on both their own networks and others carrying their traffic.

The report describes how network resources are allocated on a short time scale in order to, among other objectives, manage congestion on the network, and how such congestion management impacts applications and users. The report also recommends best practices regarding congestion management and network resource management. Among other things, the report recommends that:

- Internet Service Providers (ISPs) and Application Service Providers (ASPs) should disclose information about their user- or application- based network management and congestion management practices for Internet services in a manner that is readily accessible to the general public.
- Network operators should use accepted industry "Best Practices," standardized practices, or seek industry review of practices.
- When engaging in a congestion management practice that could have a detrimental impact on the traffic of certain users or certain applications, the practice should be designed to minimize that impact.
- If application-based congestion management practices are used, those based on a user's expressed preferences are preferred over those that are not.
- If application-based criteria are used by a network operator, they should be tested prior to deployment and on an ongoing basis.
- ASPs and Content Delivery Networks (CDNs) should implement efficient and adaptive network resource management practices.

Scott Jordan, Professor of Computer Science at the University of California-Irvine, and Fred Baker, Cisco Fellow, were the lead editors of the report. Douglas Sicker, Executive Director of BITAG, Chair of BITAG's Technical Working Group and Endowed Professor of Computer Science at the University of Colorado Boulder, chaired the review.

About BITAG. BITAG is a non-profit, multi-stakeholder organization focused on bringing together engineers and technologists in a Technical Working Group (TWG) to develop consensus on broadband network management practices and other related technical issues that can affect users' Internet experience, including the impact to and from applications, content and devices that utilize the Internet.

Questions or Comments? BITAG welcomes any questions, comments or suggestions. Please contact our Executive Director, Douglas Sicker, at dsicker@bitag.org or our Deputy Director, Kaleb Sieh, at ksieh@bitag.org.

– ATTACHMENT –

Executive Summary and Recommendations of BITAG Report on Congestion Management

The full report is available at:
http://www.bitag.org/documents/BITAG_-_Congestion_Management_Report.pdf

Executive Summary

The Internet, as is the case with many other networks such as highways and electricity grids, operates under the assumption that capacity will be set to a level such that total peak demand will occasionally exceed capacity. Further, the Internet is designed so that multiple users may dynamically share capacity and multiple services may share the same network links and routers, which is more efficient than offering individual users dedicated capacity or different services using separate links and routers.

Every link and router in the various networks that make up the Internet has a limit on its capacity to handle data. The capacity of each link and router in individual networks is determined by the equipment installed by the entity that runs each network in an attempt to optimize performance and cost; the lower the capacity relative to expected demand, the greater the probability that demand upon that link or router at times may exceed its capacity.

Significantly, a user's instantaneous demand for broadband Internet is bursty, meaning that it changes rapidly in time – and when aggregate instantaneous demand exceeds capacity on a network it causes congestion, which can degrade performance.

Network operators typically estimate demand months to years in advance, and use such demand estimates to plan a schedule for capacity upgrades. Since it may take months to implement a capacity upgrade, the time scale for managing congestion in this manner is months to years. Thus, although capacity planning can greatly affect how much congestion occurs on a network over time, it cannot react to congestion as it occurs.

The impact of congestion upon applications depends on the duration of congestion – which can vary from thousandths of a second up to hours or more – and the nature and design of the application. If the duration of congestion is short enough or the application is tolerant enough of congestion, a user will not notice any degradation in performance. Congestion is thus a problem only when its duration is long enough to be disruptive to applications. Congestion in a network can occur for a wide variety of reasons, some of which can be anticipated and some of which cannot.

This report describes how network resources are allocated on a short time scale in order to, among other objectives, manage congestion on the network, and how such congestion management impacts applications and users.

Congestion management practices are an important subset of network management practices implemented by a variety of parties or organizations, including Internet Service Providers (ISPs) and Application Service Providers (ASPs). Policymakers have expressed great interest in learning what congestion management practices are used in the Internet and how these practices impact users and the broader Internet ecosystem. Furthermore, an understanding of congestion management techniques and practices is crucial in discussions about reasonable network management.

One of the key design questions about any congestion management practice concerns the subset of network traffic to which the practice is applied, and its impact upon users and applications. Network operators apply some practices to all traffic on their networks, whereas in other cases practices are applied only to the traffic of specific users or to the

traffic associated with specific applications. Application- or user- based congestion management practices may achieve better performance for selected applications. They also may enable service providers to offer connectivity products that cater to particular customer's tastes or needs. However, they add complexity, which may result in added costs that each network operator will evaluate. In some cases application- or user- based congestion management practices may be harmful to applications.

Congestion management practices are composed of generic technical building blocks, described in this report as traffic management "techniques". This report discusses a range of user- and application- based congestion management techniques, including classification of packets, reservation of resources for particular network flows, storage of content in multiple locations, rate control, routing and traffic engineering, packet dropping, and packet scheduling.

Congestion management techniques may be combined to offer a collection of capabilities in various network architectures, and can create services with differentiated performance either within an operator's network or end-to-end. There are also architecture-specific implementations of congestion management techniques for broadband Internet access over cable, telephone, and cellular networks and for Content Delivery Networks. The offerings of a service provider often include multiple services that may utilize the same network links and routers. While there are benefits and efficiencies to sharing capacity between multiple services, such sharing of capacity also requires the use of congestion management practices.

Congestion management "practices" are the uses of particular techniques by particular network operators to avoid, limit, or manage congestion. This report illustrates a range of congestion management practices that show how providers may combine user- or application- based congestion management techniques, including traffic shaping, prioritization, transcoding, resource reservation, and preferential treatment.

The report begins in Sections 1 and 2 by giving an overview of congestion and BITAG's interest in the issue. Section 3 defines congestion and describes instances in which congestion can occur, the locations in the network where congestion can occur, the indicators of congestion, and the impact congestion can have on applications.

In Section 4, the report articulates the differences between congestion management techniques and congestion management practices, and describes the different time scales at which congestion can be seen to occur in the network. This section also describes the parties that implement congestion management practices and on what basis.

Although all congestion management is important, in order to limit scope and length Sections 5-7 focus on congestion management techniques and practices that: (1) are implemented or potentially implemented in a network that supports consumer broadband Internet access services; (2) act on a time scale of minutes or less; (3) are used for purposes of congestion management; and (4) are based on user or application.

In Section 5, the report focuses on specific congestion management techniques. Section 6 gives specific examples of congestion management practices that are based on user or application. Finally, Section 7 gives the Technical Working Group's recommendations.

At a high level, the recommendations of BITAG's Technical Working Group user are:

- **ISPs and ASPs should disclose information about their user- or application-based network management and congestion management practices for Internet services in a manner that is readily accessible to the general public**. This information should be made available on network operators' public web sites and through other typically used communications and channels, including mobile

apps, contract language, or email. ISPs and ASPs may choose to use a layered notice approach, using a simple, concise disclosure that includes key details of interest to consumers complemented by a more thorough and detailed disclosure for use by more sophisticated users, application developers, and other interested parties. The detailed disclosure should include: descriptions of the practices; the purposes served by the practices; the types of traffic subject to the practices; the practices' likely effects on end users' experiences; the triggers that activate the use of the practices; the approximate times at which the practices are used; and which subset of users may be affected. The disclosures should also include the predictable impact, if any, of a user's other subscribed network services on the performance and capacity of that user's broadband Internet access services during times of congestion, where applicable.

- **Network operators should use accepted industry "Best Practices," standardized practices, or seek industry review of practices.** Network standards setting organizations and technical industry bodies produce considered recommendations of Best Practices and standard practices for a variety of operational issues including congestion and congestion management. Where network operators see the need for an innovative solution that has not been standardized or documented as a Best Practice, these network operators should consider bringing their unique network or congestion management practices to such groups for discussion and documentation.

- **When engaging in a congestion management practice that could have a detrimental impact on the traffic of certain users or certain applications, the practice should be designed to minimize that impact.** Some congestion management practices may cause certain users or certain applications to experience performance degradation. ISPs and ASPs should seek to minimize such degradation to the extent possible while still managing the effects of the congestion that originally triggered the use of the practice.

- **If application-based congestion management practices are used, those based on a user's expressed preferences are preferred over those that are not.** User- and application- agnostic congestion management practices are useful in a wide variety of situations, and may be sufficient to accommodate the congestion management needs of network operators in the majority of situations. However, at times network operators may choose to use application-based congestion management practices, in which case those that prioritize application traffic according to a user's expressed preferences are preferred over those that do not.

- **If application-based criteria are used by a network operator, they should be tested prior to deployment and on an ongoing basis.** Application-based classification by network operators (e.g., using deep packet inspection) can sometimes be erroneous. If network operators choose to use application-based criteria for congestion management, the accuracy of the classifier should be tested before deployment.

- **ASPs and CDNs should implement efficient and adaptive network resource management practices.** ASPs and CDNs should match use of network resources to the performance requirements of the application. Applications should be designed to efficiently and adaptively use network resources, to the extent feasible given the application's requirements.