International Journal of Computer Trends and Technology (IJCTT) – volume 15 number 3 – Sep 2014

Forwarding Web Service Requests To A Single Service Instance in Service Oriented Networking

A Srilatha[1], S Tirupathi Rao[2]

[1]*M.Tech (CSE) Student, Geethanjali college of Engineering & Technology, Hyderabad, Telangana, India.*
[2] *CSE, Professor, Geethanjali college of Engineering & Technology, Hyderabad, Telangana, India.*

Abstract: Service providers within an enterprise network are often governed by client service contracts (CSC) that specify, among other constraints, the rate at which a particular service instance may be accessed. The service can be accessed via multiple points in a proxy tier configuration. The CSC and thus the rate specified have to be collectively respected by all the middleware appliances. The appliances locally shape the service requests to respect the global contract. Investigation is done in the case where the CSC limits the rate to a service to X requests with an enforcement/observation interval of T seconds. This paper extends, and investigates the Credit-based Algorithm in a production level enterprise network setting. CBA is a decentralized algorithm for service traffic shaping in middleware appliances.

Keywords: Client service contracts,Traffic shaping, Credit-based Algorithm,Service-oriented architecture, Service-oriented networking.

1.INTRODUCTION:

In order to ensure quality of service additional mechanisms besides the stream reservation rotocol (SRP)are necessary. IEEE Std. 802.1Q-2005 only described the strict priority transmission selection algorithm for the prioritization of frames. This mechanism follows the basic idea that highest priority traffic goes first. Such a concept works well as long as there is only a small mount of high priority traffic and no need to fulfill hard latency guarantees. This mechanism does not provide a deterministic low latency; hence the number of interfering higher and same priority frames is not limited. This type of prioritization scheme does not fit to environments in which audio and video streams are the predominant type of traffic, i.e. occupy a big part of the bandwidth. In the past this problem was solved with big buffers in the end stations, which guaranteed, that enough samples are buffered. This solves the problem as long as the buffers in the devices (end stations and bridges) are big enough and the applications do not require low latency. But many audio and video applications have very high requirements regarding latency (i.e. very low latency) and as the latency of the network is only one part of the total latency, it needs to be in the rage of few milliseconds. In any case the worst case latency needs to be known in order to know how many bytes a device needs to buffer to allow a reliable playback. Not only applications require low latency, but also the network itself. Latency in a network is also a measure of the memory requirements in bridges. This results of the simple fact, that a frame which is not in transmission has to be stored somewhere (accumulating latency). As the memory in bridges is limited, it is necessary to transmit traffic without undue delay through the network. This especially applies to bandwidth intensive applications like audio and video streams.

SOA networking is the use of the service-oriented architecture (SOA) model to enhance the capabilities of networks that use Web services. In SOA networking, events originating from diverse computers and communications devices are linked immediately and seamlessly to relevant business processes. The ultimate goal is the distribution of intelligence so the network functions as if it were a gigantic, self-contained computer.
One of the most important features of SOA networking is the consolidation of privacy and security services such as authentication, authorization, firewalls, anti-malware programs and encryption. Such consolidation reduces the complexity of network administration, minimizes the risk of vulnerabilities and lowers operational costs. It can allow for a more robust and reliable network than would otherwise be possible. SOA networking also facilitates streamlined testing for compliance with standards and regulations. Therefore, breaches become less likely and can be corrected in the shortest possible time when they do occur.

An SOA network functions in three layers:

The application layer includes all the software used by businesses and subscribers.
The interactive services layer ensures constant and reliable communication among all users, devices and applications. The systems layer maintains the

physical integrity of the network and ensures hardware interconnectivity and compatibility.

1.1. SERVICE-ORIENTED NETWORKING

Network infrastructure layer: Contains the enterprise network architecture, which includes switches, routers, communication links, and so on. This layer has redundancy built into it and contains network layer security to enforce business policies as needed.

Integrated service layer: Virtualizes services (or unties them from specific pieces of hardware) to allow them to be provided over a dispersed or centralized network environment. The following services are provided at this layer:

Identity: Authentication services for user or device credentials, which can play a role for network or application access.

Mobility: Allowing access to network resources from any location. This may rely on wireless technologies or a Virtual Private Network (VPN).

Storage: Storage of important network data and replication or duplication of that data, over the network, to remote locations for disaster recovery.

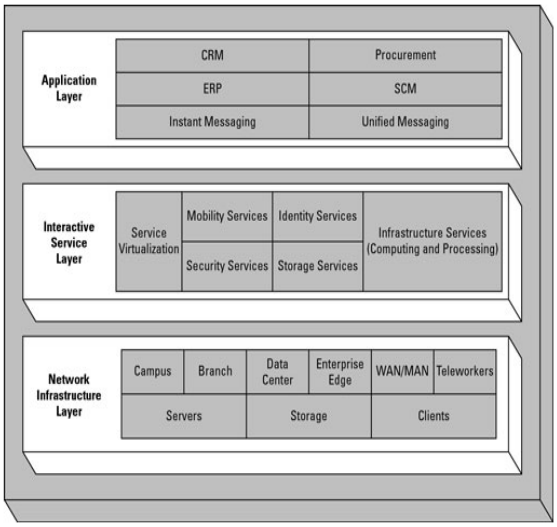

Fig. 1. System architecture showing the different elements and parameters

Computing or processing: Servers represent the main element of this component, while virtual servers allow for scaling and betting utilization of server processing power.

Security: Security for your business is crucial, and the security level makes use of security features at the network level, such as intrusion detection and prevention systems (IDS and IPS).

Voice and collaboration: Voice services now run over the main corporate data network, and have allowed for more options for users to communicate. These communication methods include the traditional telephone, but also include instant messaging and collaboration through websites, such as Microsoft's SharePoint.

Application layer: Carries the responsibility for providing the applications that users rely on. These applications include the following product areas:

Customer relationship management (CRM): Communication with clients, as well as all of their pertinent data, can be found in CRM applications.

Enterprise resource planning (ERP): Business data for your organization is found in your ERP system. This is everything that would have been in a traditional accounting system, plus information on business processes and business logic, thereby allowing you to derive more planning and statistical information from the accounting system.

Procurement: Purchasing can sometimes be tracked as part of the overall corporate ERP system, or can be a standalone system to manage purchasing from the request for a quote through to the deployment of the purchased product to the end user.

Supply chain management (SCM): Procurement systems can purchase items, but SCM systems tell procurement what parts need to be purchased and when. In manufacturing and service organizations, good SCM systems will provide you with "just in time" inventory items right before you need those items.

Instant messaging (IM): Instant messaging has come into businesses who now expect to be able to instantly communicate within their network infrastructure. This assists in users on your network in their collaboration goals.

Unified messaging (UM): Unified messaging talks all of the forms in which users can communicate and ties them together, allowing for unique situations, such as where an e-mail can be relayed to office voicemail, and then forwarded to a cell phone as a text message. Unified messaging takes control and integrates all communication and messaging formats within an organization, either partially or completely.

1.2. THE PURPOSE OF TRAFFIC SHAPING

Traffic shaping, or traffic management, controls the bandwidth available and sets the priority of traffic processed by the policy to control the volume of

traffic for a specific period (bandwidth throttling) or rate the traffic is sent (rate limiting).Traffic shaping attempts to normalize traffic peaks and bursts to prioritize certain flows over others. But there is a physical limitation to the amount of data which can be buffered and to the length of time. Once these thresholds have been surpassed, frames and packets will be dropped, and sessions will be affected in other ways. A basic traffic shaping approach is to prioritize certain traffic flows over other traffic whose potential loss is less disadvantageous. This would mean that you accept certain sacrifices in performance and stability on low-priority traffic, to increase or guarantee performance and stability to high-priority traffic. If, for example, you are applying bandwidth limitations to certain flows, you must accept the fact that these sessions can be limited and therefore negatively impacted. Note that traffic shaping is effective for normal IP traffic at normal traffic rates. To ensure that traffic shaping is working at its best, make sure that the interface Ethernet statistics show no errors, collisions or buffer overruns.

2. RELATED WORK

Many recent works in the literature addressed the real-time performance of IEEE AVB in multiple automation domains, namely, automotive, aeronautics, and industrial automation. The work [5] indicates AVB as one of the possible candidates for automotive communications. Encouraging simulation results obtained with the IEEE 802.1AS standard are provided in [6] and [7]. The IEEE AVB suitability for supporting the traffic flows of both Advanced Driver assistance Systems (ADAS) and multimedia/infotainment systems was proven in [8], [9] and [10]. IEEE AVB has the potential to be used not only as a common networking technology within a single functional domain, but also as an in-car backbone network for inter domain communication [11] through suitable gateways interconnecting the heterogeneous networks used in the different functional domains [12]. The work in [13] addressed the performance of AVB in aeronautic networks and showed that, when using time synchronization, the timing requirements are met.A number of works focused on the Ethernet AVB ability to cope with the requirements of the real-time traffic typically found in industrial automation [14]. In [15] a performance comparison between AVB and standard Ethernet is presented. The outcome of the study is that, although AVB allows for determining the worst case latency for all real-time message classes, further improvements are still needed for use in industrial automation. In fact, as the CBFQ used in AVB adopts non-preemptive scheduling, in the worst case a real-time frame might be delayed in every bridge by the ongoing transmission of a maximum sized frame not belonging to the real-time class. Approaches to mitigate this interference, such as packet preemption, fragmentation, and synchronous scheduling [16], are discussed in [17].

3. EXISTING SYSTEM

They are many approaches are available for client services contract enforcement. Among them one of the approaches is Credit based approach. It is for assigning the credits for each appliance. From this to calculate the weight for each appliance. It monitors the traffic based on regular basis and based on the credits for each appliance and queue sizes. The credit based approach has some drawbacks such as fast start, starvation, flooring effect.

A. Flooring effect

It is the existing credit-based solutions require the use of a flooring function to approximate the results to the integer immediately below. In some cases, when the number of appliances is not a divisor of the available credits, the use of a flooring function leads to under-utilization of the system.

B. Fast start

When the system operates under high input rates, all the available credits are rapidly consumed early in the enforcement period. This may result in overwhelming the service host, because a large number of requests are being sent during a time period substantially smaller than the specified enforcement period.

C. Starvation

The weighted strategies used for dynamic credit allocation are based on queue sizes. As a consequence, the appliances with at least one queued event may be allocated all the credits, thus depriving the appliances with empty queues from credits.

4. PROPOSED SYSTEM

In the existing process, traffic is shaped by using the algorithm called DOWSS. In that, there is no policies has been used. It is processed in the centralized approach. In Enhancement, we have to provide some policies for traffic shaping. Instead of DOWSS Credit-based Algorithm has been used for traffic shaping.

4.1. CREDIT-BASED ALGORITHM

Concepts in this section:
– Time sensitive Streams
– Credit based shaper for RC Traffic
– Stream reservations
AVB Terminology:
– ECU or Node = End System
– Sending Node = Talker, Receiving Node = Listener

– Stream:
Unidirectional flow of data from a Talker to one or more Listener.
– Time sensitive stream: Guaranteed bounded latency.

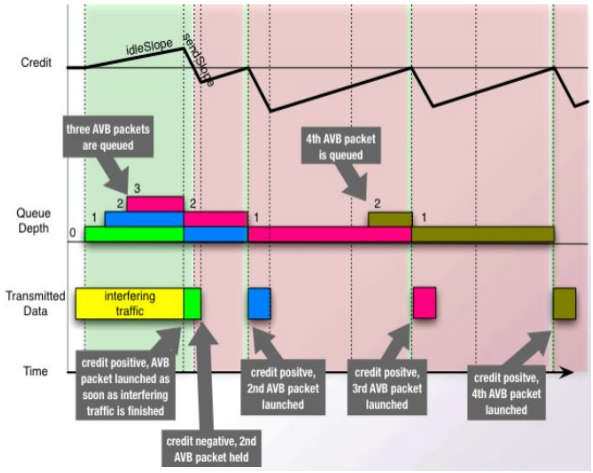

FIG.2.CBA Algorithm overview

➢ Devices in AVB network must "shape traffic"
➢ Schedule transmission of packets to prevent bunching, which causes overloading of network resources.

Time Aware Shaper

➢ The credit-based shaper reduces "bunching"
– Smooths out the traffic flow to greatly reduce the possibility of dropped packets due to congestion
➢ Average delay is actually increased
– Only the worst case is better
– Control traffic needs small-as-possible delays

Credit Based Shaper

The CBS spaces out the high priority AVB stream frames as far as possible. For this the shaper uses the information about the reserved amount of bandwidth for AVB streams, which is calculated by SRP. The spaced out traffic prevents the formation of long bursts of high priority traffic, which typically arise in traffic environments with high bandwidth streams. These bursts are responsible for significant QoS reductions of lower priority traffic classes in such traffic environments, as they completely block the transmission of the lower priority traffic for the transmission time of the high priority burst. This strongly increases the maximum latency of this traffic and thereby also the memory demands in the bridges. On the other hand long bursts also increase the interference time between high priority stream frames from different streams (which arrive from different ports) inside a bridge. This interference increases the maximum latency of this high priority stream frames and again the memory requirements in bridges. Another task of the shaper is to enforce the bandwidth reservation. Hence the shaping is performed on a per stream per class basis in the talker and on a per class per port basis in the bridges. This enforces on the one hand that every AVB stream is limited to its reserved bandwidth in the talker, and on the other hand that the overall AVB stream bandwidth of each port (in talker and bridges) is limited to the reserved one.

AVB stream frames are sent with a specific frequency. For SR class A the minimum packet frequency is 8 kHz and for SR class B 4 kHz. These frequencies are used for the bandwidth reservation. It is possible to use multiple of this frequencies and it is not required that a stream frame is sent in every transmission period, i.e. if a stream with an 8 kHz packet frequency is reserved it is also allowed to send less than 8000 stream frames in a second (e.g. necessary for rate adaptive codecs). The unused bandwidth is not lost and is used for best effort traffic (i.e. non AVB stream traffic).These frequencies also define the observation interval in which the reserved bandwidth can be measured if there is no interference with non AVB stream traffic. Hence this interval is also called class measurement interval. On the basis of the reserved amount of bandwidth and the class measurement interval it is possible to calculate two parameters which define the accumulation and reduction rate for the credit. The shaper algorithm is similar to the leaky bucket algorithm. AVB stream frames are sorted in two queues, one for SR class A stream frames and one for SR class B. The two AVB stream queues have the highest priority (SR class A is above SR class B).Frames of a specific SR class are only transmitted as long as there is positive or zero credit for this class. When the credit of a class is negative no frame of this AVB queue is transmitted, even though AVB stream frames have the highest priority. The calculation of the credit is based on the two already mentioned parameters. The idle slope, which defines the rate with which credit is accumulated, is defined as:

idleSlope = reservedBytes / classMeasurementInterval = reservedBandwidth (1)

The send slope defines the rate with which the credit is reduced and can be calculated as:

sendSlope = idleSlope portTransmitRate (2)

The credit is calculated according to the following rules:

• If there is positive credit but no AVB stream frame to transmit, the credit is set to zero.

• During the transmission of an AVB stream frame the credit is reduced with the send slope
• If the credit is negative and no AVB stream frame is in transmission, credit is accumulated with the idle slope until zero credit is reached.
• If there is an AVB stream frame in the queue but cannot be transmitted as a non AVB stream frame is in transmission, credit is accumulated with the idle slope.

In this case the credit accumulation is not limited to zero, also positive credit can be accumulated. Credit Based Shaper spaces out the frames based on the idleSlope and sendSlope. Interfering traffic which blocks the transmission of an AVB stream frame leads to an accumulation of positive credit which allows for a limited burst of stream frames to catch up. Thus the Credit Based Shaper allows for a converged network with Best Effort and Reserved Traffic (AVB stream traffic) in one network with controlled small latency.

5.CONCLUSION

CBA is a decentralized algorithm, The next exciting (and growing) application areas are automotive infotainment and home networks where LAN heterogeneity is an obvious requirement where product capabilities naturally expand from wired Ethernet to Wi-Fi and other coordinated shared networks like MoCA, HomePlug/IEEE 1901, and HomeGrid all of which are supported by the AVB architecture and standards. With strong industry support through the AVnu Alliance, multiple certification programs for these and other markets are expected to ensure interoperability of devices that implement the AVB capabilities on a diversity of IEEE 802-compatible networks. The detailed study of the modes and the influence of the dynamism of the environment, various environmental factors involved, on the behavior of the modes and thus for respecting the contract is the subject of future work.

6.REFERENCES:

[1] “Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems”, IEEE Std 1588-2009

[2] “IEEE Standard for local and metropolitan area networks—Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks”, IEEE Std 802.1AS-2011

[3] “IEEE Standard for local and metropolitan area networks — Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks”, IEEE Std 802.1Q-2011

[4] “IEEE Standard for Local and metropolitan area networks--Audio Video Bridging (AVB) Systems”, IEEE Std 802.1BA-2011

[5] “IEEE Standard for Layer 2 Transport Protocol for Time Sensitive Applications in a Bridged Local Area Network”,IEEE Std 1722-2011

[6] “Draft Standard for Device Discovery, Connection Management and Control Protocol for IEEE 1722 Based Devices”, IEEE P1722.1 draft 21, August 2012

[7] http://tools.ietf.org/html/draft-ietf-avtcore-clksrc

[8] AVnu Alliance, http://www.avnu.org

[9] Johas Teener, M., Huotari, A., Kim, Y., Kreifeldt, R., and Stanton, K., “No-excuses Audio/Video Networking: the Technology Behind AVnu,” AVnu Alliance White Paper,2009

[10] M. Head, M. Govindaraju, R. Engelen, and W. Zhang, “Benchmarking XML processors for applications in grid web services,” in ACM/IEEE Conference on Supercomputing, Tampa, FL, USA, November 2006.

[11] R. D. Callaway, A. Rodriguez, M. Devetsikiotis, and G. Cuomo,“Challenges in service-oriented networking,” in IEEE Globecom, San Francisco, CA, USA, November 2006.

[12] Gennaro (Jerry) Cuomo, “IBM SOA "on the edge"” , SIGMOD '05: Proceedings of the 2005 ACM SIGMOD international conference on Management of data, June 2005.

[13] G. Cuomo, “IBM SOA "on the edge",” in ACM Sigmod, Baltimore, MD,USA, June 2005.

[14] M. Zukerman, T. Neame, and R. Addie, “Internet traffic modeling and future technology implications,” in IEEE Infocom, San Francisco, CA,USA, January 2003.

[15] A. Varga, “The OMNeT++ Discrete Event Simulation System,” in European Simulation Multiconference, Prague, Czech Republic, June 2001.

[16] A. Elwalid and D. Mitra, “Traffic shaping at a network node: theory,optimum design, admission control,” in IEEE Infocom, Kobe, Japan,March 1997.

[17] A. Elwalid and D. Mitra, “Traffic shaping at a network node: Theory, optimum design, admission control,” in IEEE Infocom, Kobe, Japan, Apr. 1997.

[18] J. Rexford, F. Bonomi, A. Greenberg, and A. Wong, “Scalable architectures for integrated traffic shaping and link scheduling in high-speed ATM switches,” IEEE Journal on Selected Areas in Communications, vol. 15, no. 5, June 1997

[19] A. Parekh and R. Gallager, “A generalized processor sharing approach to flow control in integrated services networks: the single-node case,”IEEE/ACM Transactions on Networking, vol. 1, no. 3, pp. 344 – 357,June 1993.

[20] A. K. Parekh and R. G. Gallager, “A generalized processor sharing approach to flow control in integrated services networks – the single node case,” IEEE/ACM Transactions on Networking, vol. 1, no. 3, pp. 344–357, June 1993.