A Resource Allocation Framework for the Predictable Continuity of Mission-Critical Network Services

Rajiv Kumar
Jaypee University of Information Technology, Waknaghat, INDIA
Email:rjv.ece@gmail.com

Krishna Gopal[1], G.L. Pahuja[2]
[1,2]National Institute of Technology, Kurukshetra, INDIA
Emails: [1]krishna46@gmail.com, [2] pahuja_gl@yahoo.co.in

***Abstract* – Users of mission-critical (MC) network services such as telesurgery demand high degree of predictability as the seamless transportation of information in real time is necessitated. This paper introduces a resource allocation framework considering network reliability as a measure of the system's ability to provide deterministic and accurate delivery of information as per McCabe [2]. Proposed framework is able to support uninterrupted delivery of information from the source to the destination end.**

1 Introduction

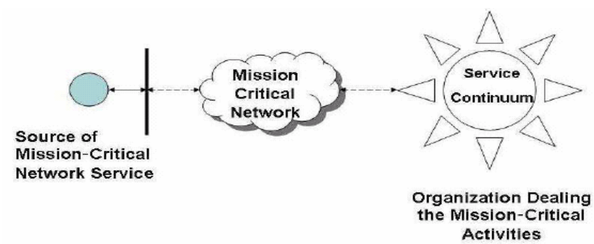

Figure 1. Mission Critical Network Service

Most MC services require seamless transportation of information in real time and must maintain continuity in service i.e. uninterrupted delivery of service transaction over the entire mission time (Fig. 1). High quality of service delivery is of paramount importance for such MC systems as interruption/reduction in QoS delivery for MC systems may involve huge loss and therefore may be unacceptable. However, it has been felt that seamless delivery of critical services is not possible through only QoS routing without taking into account the robustness of service delivery infrastructure that is frequency of failure and mean restoration time (MRT) for the service delivery infrastructure [1,7].

This paper focuses on the continuity of MC services considering the robustness of service delivery infrastructure along with service quality. Robustness of service delivery infrastructure has been considered, first for developing the routing algorithm and second for developing fault-tolerance during path failure in MC services delivery system.

2 Problem Formulation

Basic problem solved in this paper is – "How to allocate the network resources incorporating infrastructure robustness within timeliness constraints and overcoming network largeness problems in terms of large link constraints and interconnectivity." To translate the service requirements into the service offering from the network via network layer a methodology of resource allocation for the provisioning of end-to-end connectivity has been adopted.

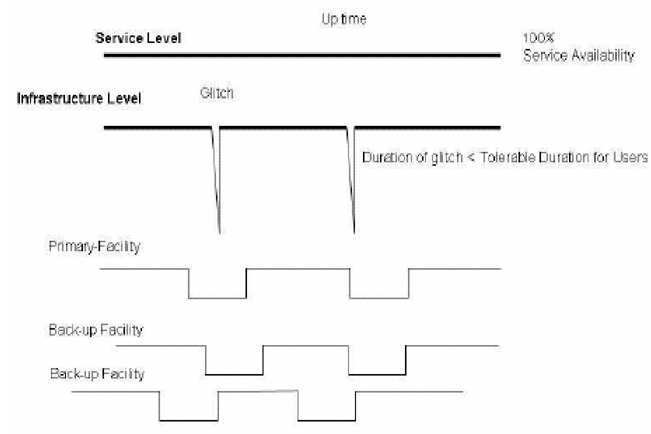

Figure 2. Timing Diagram for Mission-Critical Service Availability

Tortorella [3] characterizes the continuity of services as the ability of network to deliver the service without any interruption and it also defines continuity in terms of conditional probability, g(t; h):

$g(t; h) = P_r$\{transaction of duration h continues uninterrupted until completion at time t + h | transaction established at time t\}, where h-duration of transaction is a random variable.

Requirement of 100% service availability level as indicated in Fig. 2 can be achieved by two approaches – either by maintaining very high Mean Time to Failure (MTTF) or minimizing Mean Restoration Time (MRT) to zero. This paper exploits both the approaches.

3 Methodology

This paper proposes two separate heuristic schemes for predictable routing and fault recovery. With the help of these schemes this paper proposes an algorithm for the self-continuity based on the self-management principle.

3.1 Scheme – 1

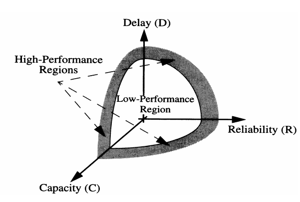

Figure 3. 3-D Service Performance Envelope [4]

Most real-time routing algorithms consider the criterion of path selection that is related to the service quality whereas robustness of the medium has been ignored [4,5,7]. However,

Copyright ISSRE 2009

robustness of service delivery medium is important as the mission-time may range from few seconds to hours or even days. Fig. 3 shows the 3-D service performance envelop. Following heuristic scheme has been proposed for the path selection according to the 3-D service performance requirements.

Computer Communication Network (CCN) considered as a multi-constrained distributed network can be easily modeled by a connected graph $G(V, E)$; where V is the set of service centers, routers, switches etc. and E is the set of communicating links. Each communication link has been assigned three different constraints: reliability, capacity and delay. These three constraints represent three different capabilities of communicating links explained as follows: *capacity* represents the rate of transfer of data, *delay* represents transportation-lag time required for delivery from the source to the destination node and *reliability* has been considered as the operating probability of the link. A specified amount of flow is to be transmitted from a source node'sʼ to a target node 'tʼ. This scheme proposes the concept of Best-Performing Path (BPP) by defining a unified link weight, w(u,v). Weight w(u,v) is function of amount of data flow, capacity, delay and reliability of link. Monte Carlo simulation of functional relationship for weight w(u,v) reveals the trend as shown in Fig. 4.

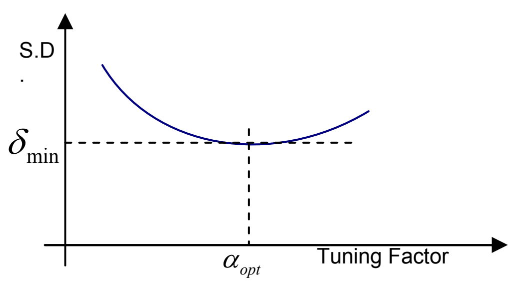

Figure 4. Standard Deviation of Path Weight vs. Tuning Factor

3.2 Scheme – 2

Second heuristic scheme proposes a predictable routing scheme for the service delivery and a real-time fault-recovery scheme so as to provide failure free transparent services. Proposed fault-tolerant scheme is a self-healing way of failure recovery that is based upon the self-management principle i.e. service continuity over the distributed network is maintained without any external support. Scheme introduces the concept of persistent states for providing failure resiliency and is able to provide up to 100% availability of the service [6]. *Persistent states* are pre-computed but are unreserved path-sets connected between'sʼ and'tʼ nodes. Methodology has been developed to allocate the resources from the persistent state which protects the temporal requirements of fault-recovery.

A complete self-continuity scheme has been shown in Fig. 5 with the help of flow chart.

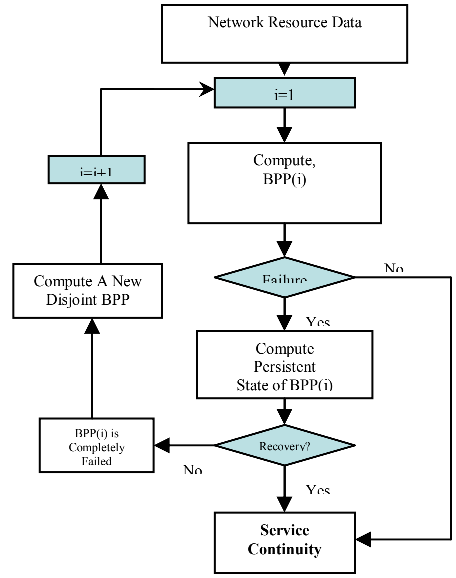

Figure 5. Flow Chart for Self-Continuity of Services

4 Conclusion

For increasing the predictability of service, this paper considers reliability in the QoS metrics. Framework proposed is able to reduce the frequency of failures. A fast method of path computation has been proposed. A fault-recovery methodology has been proposed for making the failure tolerant service delivery. Framework is able to handle the largeness during resource allocation those are related with link constraints and connectivity. Algorithms developed for the path computing and fault-recovery are useful for real-time and online applications.

References

[1] International Telecommunication Union Recommendations No. E.800, Terms and Definitions related to the Quality of Telecommunication Services, 1994.

[2] J.D. McCabe, "Network Analysis, Architecture, and Design", Morgan Kaufmann, 2003.

[3] M. Tortorella, "Service Reliability Theory and Engineering", http://www.soe.rutgers.edu/~ie/research/working_paper/paper%2004-017.pdf

[4] S. Tragdous, "The Most Reliable Data-Path Transmission", IEEE Trans. Reliability, Vol. 50, No. 3, pp 281-285, 2001.

[5] G. Xue, "End-to-End Data Paths: Quickest or Most Reliable?", IEEE Communications Letters, Vol. 2, No.6, pp156-158, 1998.

[6] David Patterson, Aaron Brown, et al, "Recovery-Oriented Computing (ROC): Motivation, Definition, Techniques, and Case Studies", Computer Science Technical Report CSD-02-1175, University of California at Berkeley, March 2002.

[7] X. Masip-Bruina, M. Yannuzzib, et al, "Research challenges in QoS routing", Computer Communications, Vol. 29, Issue 5, pp 563-581, 2006.