

멀티서비스를 제공하는 IP 네트워크에서의 링크용량 산출 기법

A Capacity Planning Framework for a QoS-Guaranteed Multi-Service IP network

최 용 민*
(Yongmin Choi)

Abstract: This article discusses a capacity planning method in QoS-guaranteed IP networks such as BcN (Broadband convergence Network). Since IP based networks have been developed to transport best-effort data traffic, the introduction of multi-service component in BcN requires fundamental modifications in capacity planning and network dimensioning. In this article, we present the key issues of the capacity planning in multi-service IP networks. To provide a foundation for network dimensioning procedure, we describe a systematic approach for classification and modeling of BcN traffic based on the QoS requirements of BcN services. We propose a capacity planning framework considering data traffic and real-time streaming traffic separately. The multi-service Erlang model, an extension of the conventional Erlang B loss model, is introduced to determine required link capacity for the call based real-time streaming traffic. The application of multi-service Erlang model can provide significant improvement in network planning due to sharing of network bandwidth among the different services.

Keywords: IP network, Broadband convergence Network (BcN), capacity planning, network dimensioning, traffic modeling, multi-service Erlang model.

I 서론

BcN(Broadband convergence Network)은 IP를 기반으로 통신, 방송, 인터넷이 융합된 품질 보장형 멀티미디어 서비스를 언제 어디서나 끊김 없이 제공하고자 하는 통합 네트워크이다. 멀티서비스를 제공하기 위한 BcN은 단일 서비스를 제공하던 기존의 전화망이나 인터넷과는 다른 엔지니어링 기술이 필요한데 링크용량 산출이나 네트워크 계획 분야에서도 적절한 트래픽 모델링과 디멘저닝 방법이 필요하다.

BcN에서는 기존의 최선형 데이터 트래픽을 전송하는 인터넷 서비스뿐만 아니라 실시간 대화형 트래픽을 전송하는 IP 전화, 네트워크 게임 또는 비디오 스트리밍 서비스 등을 고려해야 한다. 인터넷 서비스는 그 품질이 전체 데이터를 전송하는데 걸리는 시간에 영향을 받으며 주로 TCP를 사용하여 전송되는 웹 또는 이메일과 같은 데이터 트래픽으로 구성된다. IP 전화 또는 비디오 전송과 같은 스트리밍 트래픽을 전송하는 서비스 품질은 개별적인 패킷의 시간 지연, 지터 또는 손실과 같은 전송 특성에 영향을 받는다. 따라서 BcN에서의 링크용량 설계나 네트워크 계획은 서로 다른 QoS 요구사항을 갖는 데이터 트래픽과 스트리밍 트래픽을 동시에 고려하는 기법이 필요하다.

또한 BcN에서 제공되는 통화형 서비스는 기존 전화망에서와 달리 가변적인 전송률을 갖는 코덱이 사용되고 다양한 모드의 서비스가 가능하게 된다. 예를 들면 IP 전화의 경우 영상의 유무 또는 사용하는 코덱의 종류에 따라 한 통화가 차지하는 대역폭이 달라진다. 이와 같이 BcN에서의 링크용량 설계 및 네트워크 계획은 전화망에서 사용하는 Erlang 모델 기반의 설계기법이나 인터넷 설계에 사용되는 overprovisioning과는 근본적으로 다른 접근 방법이 요구된다.

본 논문은 가입자로부터 백본 네트워크에 연결되는 링크에 대해 다양한 QoS 요구사항을 갖는 멀티서비스를 제공하는 BcN에서의 네트워크 계획 방법에 대해 고찰한다. 즉 네

트워크 구조가 결정되어 있고 노드에 수용되는 가입자 수가 주어졌을 때 개별적인 서비스가 요구하는 QoS 요구사항을 만족시키는 적정 링크용량을 산출하는 문제에 대해 논의한다.

이를 위해 먼저 BcN 멀티서비스가 요구하는 QoS 특성에 따른 트래픽 분류 및 모델링에 대해 검토한다. 멀티서비스 트래픽은 탄성적(elastic)과 스트리밍(streaming)으로 구분되는데 탄성적 트래픽은 목표연도의 가입자당 트래픽을 예측한 다음 주어진 패킷 손실률을 만족시키는 링크용량을 산출하고, 스트리밍 트래픽의 경우 통화 호 기반의 모델링을 기반으로 주어진 호 블로킹 확률을 만족하는 링크용량을 산출하게 된다.

이때 통화 호 기반의 서비스를 제공하는데 필요한 링크용량을 산출하는데 멀티서비스 Erlang 모델을 사용한다. 전통적으로 Erlang B 모델을 사용하여 서비스 수요로부터 호 블로킹 확률과 같은 QoS 요구사항을 만족하는 링크용량을 계산한다. 단일 서비스가 제공되고 한 통화가 차지하는 대역폭이 일정하다면 기존 Erlang B 모델이 적용될 수 있으나 BcN에서는 IP 전화, 비디오 등의 멀티서비스가 제공되고 서비스 모드도 다양하며 한 통화를 제공하는데 필요한 대역폭도 가변적이므로 기존 모델을 적용하기 어렵다. 멀티서비스 및 다양한 전송률을 갖는 서비스 수요로부터 소요 대역폭을 산출하기 위해서는 멀티서비스 Erlang 모델을 적용해야 한다.

본 논문에서는 멀티서비스 Erlang 모델을 사용하여 BcN 접속링크에 대해 IP 전화 및 비디오 스트리밍 서비스를 제공하는데 필요한 링크용량을 산출한다. BcN은 음성 및 영상 서비스에 대해 CAC(Call Admission Control)와 같은 자원제어를 적용하므로 IP 전화와 비디오 서비스에 대해 호 기반의 모델링을 적용하는 것이 가능하다. 멀티서비스 Erlang 모델에서 IP 전화와 비디오 서비스는 접속링크 대역폭을 공유하고 새로운 호가 도착했을 때 서비스를 제공하는데 필요한 대역폭이 존재하지 않으면 그 호를 거절한다. 멀티서비스 Erlang 모델은 단일 서비스에 대한 Erlang B 모델을 적용하는 방식보다 정확한 소요 대역폭을 산출할 수 있다는 장점이 있다.

* 책임저자(Corresponding Author)

최용민: KT BcN본부 (himrchoi@kt.co.kr)

본 논문의 구성은 먼저 2장에서 네트워크 계획시 고려하는 서비스에 대해 정의하고 QoS 요구사항을 고려한 트래픽 분류와 모델링에 대해 논의한다. 3장에서는 멀티서비스 Erlang 모델에 대해 논의하고 4장에서 BcN 환경에서의 네트워크 계획 방법에 대해 기술한다. 5장은 결론과 향후 과제에 대해 기술한다.

II. 트래픽 분류 및 모델링

적정한 멀티서비스 IP 네트워크 계획을 위해서 BcN 응용의 트래픽은 QoS 요구사항에 따라 탄성적(elastic) 트래픽과 스트리밍 트래픽의 두 가지로 분류된다 [4]. 이것은 각 응용의 모든 특성을 고려하는 것은 네트워크 계획을 지나치게 복잡하게 하고 현실적으로 불필요하기 때문이다.

탄성적 트래픽은 서비스 품질이 전체 데이터를 전송하는데 걸리는 시간에 영향을 받는 트래픽이다. 탄성적 트래픽을 발생시키는 응용은 웹, 이메일과 같이 데이터를 전송하는데 사용하는 응용으로서 기존 KORNET과 같은 인터넷에서 사용하는 응용과 큰 차이가 없다. 이 트래픽은 주로 TCP 프로토콜을 사용하여 전송되는데 TCP는 그 전송률이 가용한 네트워크 대역폭에 적응적으로 조절되는 특징이 있다. 탄성적 트래픽에 적합한 QoS 파라미터로는 데이터를 전송하는데 걸리는 시간으로부터 계산한 throughput이 있다.

스트리밍 트래픽은 서비스 품질이 개별 패킷들의 시간 지연, 지터 또는 손실과 같은 전송특성에 영향을 받는 트래픽이다. IP 전화, 네트워크 게임 또는 비디오 스트리밍과 같은 실시간 데이터를 전송하는 응용에 의해 발생하는 트래픽으로서 적합한 QoS 파라미터로는 시간 지연, 손실과 같은 패킷의 전송특성을 사용할 수 있다. BcN과 같이 호 기반의 연결 제어(CAC: Call Admission Control)가 적용된다면 앞서 언급한 QoS 척도는 블로킹 확률을 적용할 수도 있다. 실제로 네트워크 계획 단계에서는 시간지연이나 지터와 같이 예측하기 어려운 패킷 레벨의 트래픽 특성 보다는 연결 기반에서의 호 블로킹 확률과 같은 QoS 요구사항을 고려한다.

통화 호 또는 플로우 레벨에서의 스트리밍 트래픽은 호 도착률과 평균 지속시간(또는 이 두 파라미터의 곱인 Erlang으로 표현한 트래픽 볼륨)으로 모델링할 수 있다. 이때 호 도착간 시간과 평균 지속시간은 지수함수적으로 분포한다고 가정한다 (Poisson 가정). 연결이 지속되는 시간 동안 한 통화 호가 전송하는 데이터 전송률은 가변적이다. 이것은 BcN에서 제공하는 IP 전화 또는 VOD 서비스의 경우 가변 비트율을 지원하는 코덱을 사용하기 때문이므로 네트워크 계획시 한 통화가 차지하는 대역폭을 산정하는 경우 유효대역폭(effective or equivalent bandwidth) 개념을 적용한다.

BcN에서는 소프트웨어 스위치에 의해 IP 전화 또는 VOD와 같은 연결기반의 서비스에 대한 자원제어를 수행하므로 스트리밍 트래픽에 대한 QoS 파라미터로 호 블로킹 확률을 사용한다. 만약 CAC와 같은 연결제어가 적용되지 않는다면, QoS 파라미터로는 각 통화 호가 SLA (Service Level Agreement)에서 규정한 서비스 품질을 제공받지 못하는 시간의 비율로 계산하는 열화 확률(degradation probability)를 사용할 수 있다 [4].

본 논문에서 다루는 BcN 네트워크 계획 프레임워크는 기

존의 인터넷 서비스에 의한 탄성적 트래픽과 IP 기반의 통화형 서비스와 VOD 서비스에 의한 스트리밍 트래픽을 대상으로 한다. BcN은 멀티캐스트를 이용한 방송 서비스도 포함하고 있으나 본 논문에서는 네트워크 계획을 단순화하기 위해서 호 기반의 스트리밍 트래픽만을 고려한다. 향후 IPTV가 본격적으로 서비스되면 네트워크 계획시에도 멀티캐스트에 의한 스트리밍 트래픽을 고려할 필요가 있다.

III. 멀티서비스 Erlang 모델

멀티서비스 Erlang 모델은 두 가지 이상의 서비스가 동일한 링크를 사용하여 제공되는 경우 적용한다. 단일 서비스를 제공하는 경우 전통적인 Erlang B loss model에 의해 트래픽 수요로부터 주어진 QoS 요구조건을 만족하는 링크용량을 계산할 수가 있다. 멀티서비스의 경우에도 Erlang B loss model을 적용하면 각 서비스에 대한 소요 대역폭을 독립적으로 구할 수 있고, 전체 링크용량은 각 서비스들의 소요 대역폭의 합으로 구할 수 있다. 이와 같은 방법에 의할 경우 BcN 멀티서비스는 동일한 시스템 자원(즉, 링크 대역폭)을 공유함에도 불구하고 각 서비스가 시스템에서 독립적으로 제공된다고 가정하게 되므로 산출한 링크용량은 실제 필요한 대역폭보다 더 크게 된다. 따라서 멀티서비스의 joint blocking probability 계산에 기반하여 Erlang B loss model을 확장한 개념의 멀티서비스 Erlang model을 사용할 필요가 있다.

멀티서비스 Erlang 모델은 서로 다른 K개의 서비스가 시스템 자원 C를 공유하는 대기 시스템이다 [5]. 임의의 서비스 k에서 고객의 도착 프로세스는 도착률 λ_k 인 Poisson 분포를 따르고 서비스 시간 프로세스는 $1/\mu_k$ 인 지수분포를 따르며, 각 서비스가 대기 시스템에 부과하는 offered load는 $\rho_k = \lambda_k / \mu_k$ 이다. 하나의 k번째 서비스 요구는 b_k 의 대역폭을 요구하는데 서비스 블로킹은 시스템에 b_k 자원이 존재하지 않을 때 발생한다. 그림 1은 K개의 서비스가 존재하는 멀티서비스 Erlang 모델을 나타내었다.

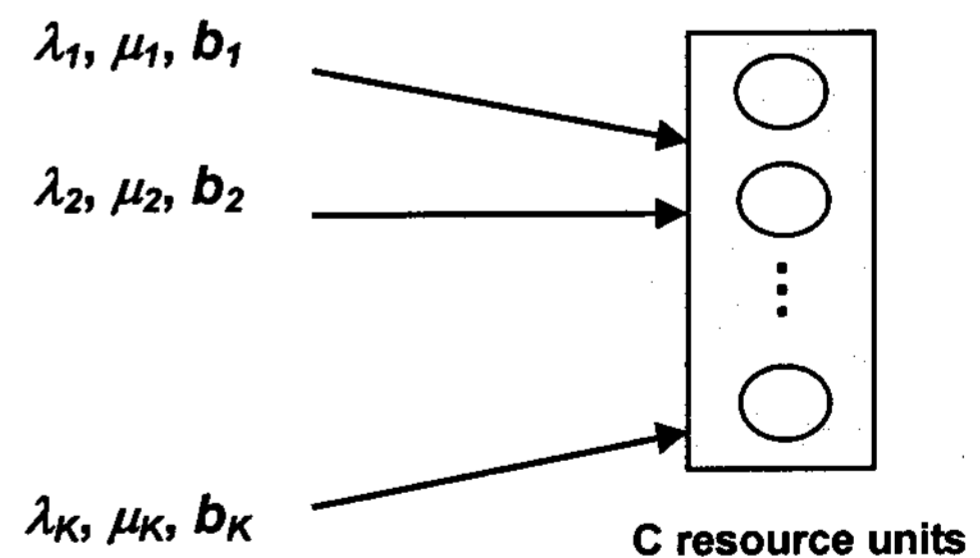


그림 1. 시스템 모델

멀티서비스 Erlang 모델에서 시스템 상태는 각 서비스의 accepted request의 수 (n_1, \dots, n_k) 로 정의하는데 여기서 $n_k=0, 1, \dots, \lfloor C/b_k \rfloor$ 의 값을 가진다. 따라서, 이 대기 시스템에 대한 상태 전이는 다차원 Markov chain으로 표현할 수 있는데 이 Markov chain은 제공하는 서비스의 개수만큼의 차원을 갖는다. 시스템 용량 C는 다음과 같이 가능한 시스템 상태의 집합 S를 결정한다.

$$S = \left\{ (n_1, \dots, n_K) \mid \sum_{k=1}^K n_k \cdot b_k \leq C \right\}$$

이 시스템에서 호 블로킹 확률은 product form 또는 recursive 알고리즘으로 구할 수 있다 [6]. 실제로 멀티서비스 Erlang B 모델에 대해서는 A. K. Erlang 이래로 많은 연구가 있었으며 효율적인 호 블로킹 확률을 계산하기 위한 알고리즘에 대해서도 많은 연구 결과가 있다.

각 서비스별 QoS 요구사항을 만족하는 멀티서비스 대기시스템의 용량은 멀티서비스 Erlang-B 공식에 의해 계산한다. 이때 b_k 는 각 서비스의 유효대역폭이며 QoS_k 는 호 블로킹 확률을 사용하는데, 그림 2와 같이 멀티서비스 Erlang-B 공식을 사용하여 각 서비스별 트래픽 수요로부터 시스템 용량(소요대역폭)을 계산한다.

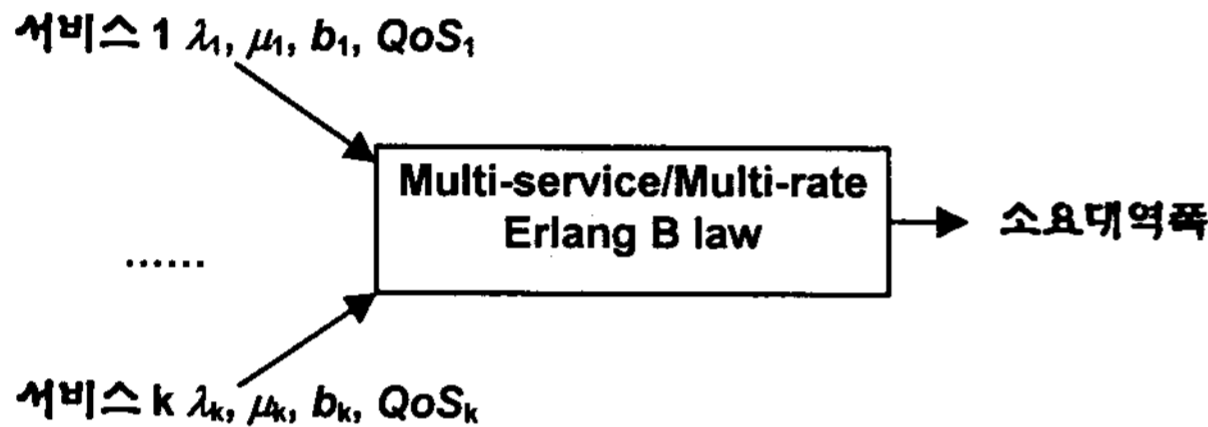


그림 2. 멀티서비스 Erlang 모델에 의한 소요대역폭 계산

IV. 네트워크 계획 방법

본 논문에서 제안하는 네트워크 계획 프레임워크는 탄성적 트래픽으로 인터넷 서비스를 고려하고 스트리밍 트래픽으로 IP 전화와 VOD 서비스의 회선 연결 기반의 서비스를 그 대상으로 한다. BcN에서는 이들 서비스 이외에도 다른 다양한 서비스가 제공될 것이나 본 논문에서는 제안한 프레임워크의 설명을 쉽게 하기 위해 예시한 세 가지 서비스만을 고려한다. 다른 서비스를 추가하는 경우 트래픽 속성에 따른 분류를 통해 스트리밍 트래픽은 멀티서비스 Erlang 모델에 포함하거나 탄성적 트래픽의 용량 산출에 반영하는 등 본 프레임워크를 확장할 수 있다.

본 프레임워크에서는 탄성적 트래픽에 대해 기존 인터넷에서와 같이 overprovisioning에 의한 소요대역폭을 산출하는데 이 방법은 기본적으로 인터넷 트래픽의 변동성을 수용하기 위해 산출하는 소요대역폭보다 충분한 링크용량을 공급하는 방법이다. 이 overprovisioning 방법에서는 가입자당 트래픽을 산출하고 목표시점에서의 트래픽을 예측한 후 가입자수를 곱하는 방식으로 소요대역폭을 산출한다. 이와 같이 산출한 소요대역폭으로부터 링크용량은 보정계수를 적용해 계산하게 되는데, 보정계수란 소요대역폭과 링크용량의 비로 정의된다. 보통 인터넷 백본 접속링크에서 관찰하는 인터넷 트래픽은 정규분포를 따른다고 가정하고 산출한 소요대역폭을 기준으로 주어진 패킷 손실률을 만족시키는 링크용량을 결정한다.

스트리밍 트래픽의 경우 앞서 소개한 멀티서비스 Erlang 모델에 의해 소요대역폭을 계산한다. 즉, 각각의 서비스에 대한 부과 트래픽(offered load)과 서비스별 QoS 요구사항인 호

블로킹 확률을 입력으로 받아 멀티서비스 Erlang-B 공식에 의해 서비스별 소요대역폭을 산출한다. 본 논문에서는 IP 전화의 경우 64, 128, 192, 384 및 1024 Kbps의 다섯 가지 서비스 모드를 가정하고 VOD는 WMT9, SD 및 HD의 세가지 모드를 가정하므로 멀티서비스 Erlang 모델은 8차원 Markov chain으로 표현된다. IP 전화의 경우 유효대역폭은 서비스 모드와 동일하며 VOD 서비스별 유효 대역폭은 사용하는 코덱에 따른 평균대역폭을 가정한다. 예를 들면 HD 모드의 경우 H.264의 경우 대략 8Mbps의 대역폭을 사용한다고 가정한다. 스트리밍 트래픽의 경우에도 서비스별 부과 트래픽을 기준으로 목표연도의 트래픽을 예측한 후 소요대역폭을 계산한다.

이와 같이 각각 구한 인터넷과 스트리밍 트래픽의 링크용량을 합산하여 목표시점에서의 링크용량을 결정한다. 이 네트워크 계획 프레임워크에 의한 경우 통화 호 기반의 스트리밍 트래픽 간의 대역폭 공유 효과는 고려할 수 있으나 탄성적 트래픽과 스트리밍 트래픽의 대역폭 공유 효과를 고려할 수는 없다는 한계가 있다. 그러나 BcN 멀티서비스 중 회선 연결 기반의 서비스간의 통계적 다중화를 고려하므로 단일 서비스 각각에 대해 Erlang-B loss model을 적용한 링크용량 산출 방법에 비해서 좀더 정확한 소요대역폭 산출이 가능하다.

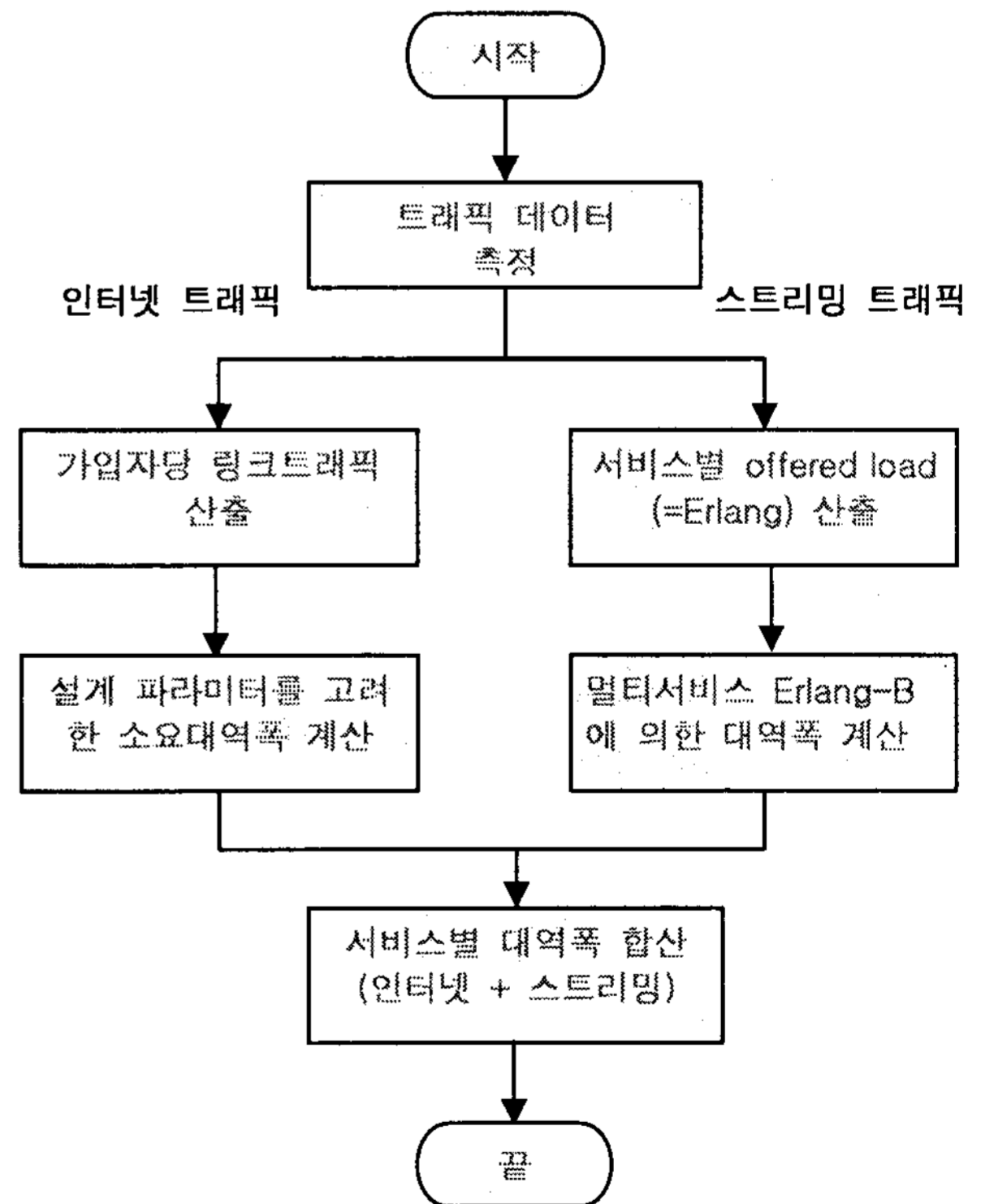


그림 3. 멀티서비스 Erlang 모델에 의한 소요대역폭 계산

V. 결론

본 논문은 BcN과 같이 멀티서비스를 제공하는 IP 네트워크에서의 네트워크 계획을 위한 프레임워크에 대해 논의하였다. 효과적인 링크용량 산출 및 네트워크 계획을 위해서 QoS 요구사항에 따라 탄성적 트래픽과 스트리밍 트래픽으로 분류한 다음 탄성적 트래픽은 가입자당 트래픽을 예측하여 주어진 패킷 손실률을 만족시키는 링크용량을 산출하고 스트리밍 트래픽의 경우 호 기반의 모델링을 기반으로 주어진 호 블로킹 확률을 만족하는 링크용량을 산출한다. 회선 연결 기반의 IP 전화와 VOD 서비스에 의해 전송되는 스트리밍 트래픽에 대해서는 기존 Erlang-B loss model을 확장한 멀티서비스 Erlang 모델을 사용하여 링크용량을 산출하는데 이 모델을 사용할 경우 멀티서비스간 대역폭 공유를 고려할 수 있기 때문에 보다 효과적인 네트워크 계획이 가능하게 된다. 향후 실제 네트워크에서 설계 파라미터를 산출하여 본 논문에서 제안한 프레임워크에 적용한 결과를 확인하는 과정이 필요하다.

참고문헌

- [1] M. Menth, R. Martin, and J. Charzinski, "Capacity overprovisioning for networks with resilience requirements," in *Proc. SIGCOMM '06*, pp. 87-98, Sep. 2006.
- [2] I. Atov and R. J. Harris, "Framework for capacity planning of multiservice IP networks," in *Proc. ICICS-PCM 2003*, pp. 1707-1711, Dec. 2003.
- [3] K. Wu and D. S. Reeves, "Capacity planning of DiffServ networks with best-effort and expedited forwarding traffic," in *Proc. ICC 2003*, pp. 1902-1906, May 2003.
- [4] A. Riedl, T. Bauschert, and J. Frings, "A framework for multi-service IP network planning," in *Proc. Networks 2002*, pp. 183-190, June, 2002.
- [5] D. Adiego and C. Cordier, "Multi-service radio dimensioning for UMTS circuit-switched services," in *Proc. VTC 2001*, pp. 2745-2749, Oct. 2001.
- [6] A. A. Nilsson, M. Perry, A. Gersht, and V. B. Iversen, "On multi-rate Erlang-B computations," in *Proc. 16th International Teletraffic Congress (ITC 16)*, pp. 1051-1060, 1999.
- [7] K. Lindberger, "Balancing quality of service, pricing and utilisation in multiservice networks with stream and elastic traffic," in *Proc. 16th International Teletraffic Congress (ITC 16)*, pp. 1127-1136, 1999.
- [8] K. Lindberger, "Dimensioning and design methods for integrated ATM networks," in *Proc. 14th International Teletraffic Congress (ITC 14)*, pp. 897-906, 1994.
- [9] J. S. Kaufman, "Blocking in a shared resource environment," *IEEE Trans. Communications*, vol. 29, no. 10, pp. 1474-1481, Oct. 1981.
- [10] L. Kleinrock, *Queueing Systems Volume II: Computer Applications*, Wiley-Interscience Publication, 1976.