Network Planning & Capacity Management

Frank Yeong-Sung Lin (林永松) Department of Information Management National Taiwan University Taipei, Taiwan, R.O.C.

Outline

- Introduction
- Network planning & capacity management
- Examples and demonstration
- Features for network planning & capacity management
- Summary
- Conclusion

Introduction

- Motivation
 - complexity of networks
 - needs of decision support systems (DSSs) and operation support systems (OSSs)
- Considerations
 - installation/operation/maintenance cost
 - network performance (sanity)
 - network integrity

Introduction (cont'd)

- Issues
 - efficiency and effectiveness
 - timeliness (development and response)
 - capacity
 - environment
 - user friendliness
 - integration and reliability
 - cost

Network Planning & Capacity Management

• System architecture



- Network planning
 - to design a network with the minimum installation and operation cost subject to performance (QoS), survivability/reliability and other constraints
- Network performance assurance/optimization
 - for an in-service traffic network, to assure prespecified QoS requirements and/or to optimize certain performance measures, e.g. to minimize the total system throughput/revenue or to minimize the average cross-network packet delay

- Network monitoring
 - for an in-service traffic network, by using traffic measurements or performance modeling techniques (or a combination of the two) to identify potential performance exceptions and to activate corrective actions
 - to collect traffic measurements for load forecasting purposes (to feed the servicing and the capacity expansion processes)

- Network servicing
 - using corrective actions to alleviate the performance exceptions identified by the monitoring process
 - three typical approaches
 - » traffic rerouting
 - » resource reallocation
 - » sizing (minimal-cost capacity augmentation to satisfy the current demand)

- Network capacity expansion
 - for an in-service traffic network, to determine the capacity augmentation strategy at each decision stage over a pre-specified time horizon such that the total cost, considering the effect of economies of scale and composite cost of money, is minimized

Performance Considerations

- Performance/service objectives/constraints
 - throughput
 - peak delay
 - mean delay
 - delay jitter
 - tail distribution of delay (percentile type)
 - call set-up delay
 - call blocking probability
 - packet/cell loss probability
 - interference
 - availability/reliability/survivability

- Performance evaluation
 - traffic measurements
 - » call/packet/cell counts
 - » packet/cell loss counts
 - » call blocked counts
 - » delay counts are usually not directly available
 - performance modeling
 - » to derive performance measures from available traffic measurements & appropriate queueing models
 - » optimization is used to derive performance bounds from imperfect information for engineering purposes

- Performance evaluation (cont'd)
 - introduction to queueing theories
 - » components of queueing systems
 - probability density function (pdf) of interarrival times
 - pdf of service times
 - the number of servers
 - the queueing disciplines
 - the amount of buffer

- Performance evaluation (cont'd)
 - introduction to queueing theories (cont'd)
 - » notation
 - *M*: exponential probability density
 - *D*: deterministic
 - G: general

e.g. *M*/*M*/1, *M*/*M*/*m*/*m*, *M*/*D*/1/*K*, *G*/*G*/*m*

- » Little's result $N = T\lambda$.
- » M/G/1 queues are fully solvable (P-K formula).
- » *GI/GI*/1 queues can be approximately analyzed by using the first two moments of the interarrival times and the service times.
- » *M/M/m/m* queueing models can be used to analyze the call blocking probability (Erlang B formula).

- Notion of equivalent bandwidth
 - Reference: R. Guerin et al, "Equivalent capacity and its application to bandwidth allocation in high-speed networks", IEEE Journal on Selected Areas in Communications, 9(7), Sep. 1991
 - The approximation for the equivalent capacity is based on a fluid-flow model, which focuses on the representation of traffic source.
 - A traffic source is modeled by a two-state Markov source, characterized by the connection metric vector (R_{peak}, ρ, b)



- » R_{peak} : the peak rate of the connection
- » *b*: the mean of burst period (the mean of times during which the source is active)
- » ρ : utilization (fraction of time the source is active)

- Notion of equivalent bandwidth (cont'd)
 - We wish to determine the bandwidth to allocate to the associated connection in isolation.
 - The distribution of the buffer contents, when such a source is feeding a buffer served by a constant rate server, can be derived using standard techniques.
 - From this distribution, it is then possible to determine the equivalent capacity \hat{c} , needed to achieve a given buffer overflow probability.
 - Assuming a finite buffer of size x and overflow probability ε (the PDU loss requirement), the equivalent capacity is obtained by

$$\hat{c} = \frac{\alpha b(1-\rho)R_{peak} - x + \sqrt{\left[\alpha b(1-\rho)R_{peak} - x\right]^2 + 4x\alpha b\rho(1-\rho)R_{peak}}}{2\alpha b(1-\rho)}$$

where $\alpha = \ln(1/\varepsilon)$

Cost Considerations

- Deployment cost
 - fixed cost
 - » real estate
 - » switches and interface cards
 - variable cost
 - » transmission capacity
 - » switching capability
- Operational cost
 - maintenance cost
 - personnel cost

Data in Support of Network Planning & Capacity Management

- Location data
 - candidate locations and corresponding real estate costs
- Traffic requirement
 - end-to-end (preferred) or network element demand
- Tariffs
 - charge policies with respect to services provided

Data in Support of Network Planning & Capacity Management (cont'd)

- Network element cost structure
 - cost of network elements considering pricing of available network element types and economies of scale
- Network element characteristics
 - load-service curves of each network element
- Performance objectives
 - user or system performance objectives specified by generic requirements or service contracts

Examples & Demo for Network Planning & Capacity Management

- Minimax OSPF (Open Shortest Path First) routing algorithms in networks supporting the SMDS service
- Access network design
- Headend interconnection and Internet access
- Integrated network design

Features for Network Planning & Capacity Management

- [1] Path Assignment and Admission Control to Satisfice Multiple Service Requirements in ATM Networks
- [2] Allocation of End-to-end Delay Objectives for CCSNs
- [3] A Real-time Distributed Routing and Admission Control Algorithm for ATM Networks
- [4] Minimax Open Shortest Path First Routing Algorithm in Networks Supporting the SMDS Service
- [5] Link Set Sizing for Networks Supporting OSPF Routing
- [6] Estimating IP Packet Loss Probabilities from ATM Cell Loss Probabilities
- [7] Allocation of End-to-end Delay Objectives for Networks Supporting SMDS
- [8] A Near-optimal Minimax Routing Algorithm in Networks with Single-path Routing
- [9] Virtual Path Assignment and Virtual Circuit Routing in ATM Networks
- [10] Link Set Capacity Augmentation Algorithms for Networks Supporting SMDS
- [11] On the Joint Virtual Path Assignment and Virtual Circuit Routing Problem in ATM Networks
- [12] Allocating End-to-end Delay Objectives for Common Channel Signaling Networks
- [13] An Optimal Algorithm for Threshold Determination in Network Monitoring Processes
- [14] Performance Analyses for SCPs in CCSNs
- [15] End-to-end QoS Assurance for Broadband Internet

Features for Network Planning & Capacity Management (cont'd)

- [16] Minimax End-to-end Delay Routing and Capacity Assignment for Virtual Circuit Networks
- [17] Optimal Real-time Admission Control Algorithms for the Video-On-Demand (VOD) Service
- [18] Admission Control and Routing Algorithms for the Permanent Virtual Connection (PVC) Service
- [19] Quasi-static Channel Assignment Algorithms for Wireless Communications Networks
- [20] Real-time Admission Control and Buffer Management Algorithms for the Video-On Demand (VOD) Service under Various Video Server Architectures
- [21] Performance Management of the Video-On-Demand Service over ATM Networks
- [22] End-to-end Packet Delay Analysis over ATM Networks
- [23] Channel Assignment, Transmission Power Control, Base Station Allocation and MTSO Allocation in Wireless Communications Networks
- [24] Near-optimal Design of CATV Networks
- [25] Near-optimal Design of Access Networks
- [26] Virtual Path Assignment and Virtual Connection Routing in ATM Networks with Multiple Types of Traffic
- [27] An Optimal Capacity Expansion Algorithm for STPs in CCSNs
- [28] HE Interconnection and Internet Access

Summary

- Architecture and functionality of network planning & capacity management (NPCM) are presented.
- Examples and demonstration are given.
- A number of features for network planning & capacity management are introduced.

Conclusion

- Information (technology) is power!
- Networking is the core of the information era.
- Network planning & capacity management is crucial for reliable and efficient information acquisition, exchange and distribution.
- Information over planned and well managed networks is even more powerful!

Q&A

