LESSON 23:
ULTRA HIGH CAPACITY NETWORKS

Objective

The objective here is to learn ultrahigh capacity networks.

Introduction

Ultra high capacity optical networks that are able to provide bandwidth on demand cannot be modeled with high fidelity using today's methods and tools. Discrete event and hybrid simulators, and (near) real time network emulators capable of 'processing' several hundred thousand packets per second cannot cope with aggregate traffic volumes four or five orders of magnitude larger. Analytical techniques provide first-order steady state approximations, but cannot capture dynamic and transient behaviors. All relevant information about the network's *data plane* is contained in the signaling messages that transit the network's *control plane*, so it is not necessary to precisely

simulate data plane traffic, nor to use parallel or concurrent simulation techniques to synchronize the operation of the data plane's components.

Modeling these networks in (near) real time requires a combination of techniques: emulators to mimic the network's control plane; an inference engine to deduce data plane behavior and performance from traffic observed in the control plane; analytics to model core network behaviors; a fast hybrid simulator to inject traffic and network impairments; and a supervisory kernel to interconnect and manage the components. This approach allows one to model ultra high capacity networks at full operational load in (near) real time with very high fidelity.

Ultra High Capacity Networks

Many optical networks have link rates of 2.5 Gigabits per second, and research teams are deploying 10 and 40Gbit/sec optical links in experimental network test beds (*e.g.* the Internet2 Abilene backbone, and the NSF's TeraGrid) [Dr00]. If the optical links are conditioned to carry multiple channels via wavelength division multiplexing, then a network's capacity can exceed tens of Petabits per second. An important feature in some ultra high capacity networks is the ability to provision wavelength connections between endpoints ("light paths"). This can done in a number of ways (*e.g.*, IETF GMPLS, and the User-Controlled Lightpath Provisioning protocol [Wu03]). Several experimental architectures are capable of provisioning end-to-end light paths in a few milliseconds [*e.g.*, Ch99, Ch00, Ve00, Ya00, Du02, Ma02, Ra02a, Vu02, Yu02a], and several applications actually utilize user- or application initiated light path provisioning. *E.g.*, the high energy physics community uses "lambda grids" for Terabyte file transfers; others are experimenting with visualization applications in which data is speculatively pre-fetched using a set of ultra high bandwidth "Terapipes", and stored in close proximity to endpoints so that the latency seen by the application is quite small. Many believe user- and application-initiated provisioning applications will migrate to carrier and ISP optical networks within five years.

The Global Grid Forum has identified basic requirements for applications and services that use optical network resources; these include: **(1)** a scalable, flexible, and rapidly reconfigurable optical network infrastructure; **(2)** ultra high bandwidth on demand between arbitrary endpoints; and **(3)** user/application provisioning and control of bandwidth with sub-wavelength granularity [GRID].

Provisionable optical networks have been identified by the U.S. National Science Foundation and various Federal agencies as an essential and critical part of the Nation's information infrastructure. Unfortunately, large-scale, ultra high capacity optical networks cannot be realistically modeled and analyzed using today's methods and tools. A new approach is required to provide insight into how applications and services that rapidly provision and release network resources might behave in these networks.

Limitations Of Today's Methods

Networks with Petabit per second capacities (and billions of packets in transit) supporting services that opportunistically 'scavenge' network bandwidth cannot be realistically *analyzed* with today's performance tools. In some networks, the behavior of packets transiting a single ultra high capacity optical link cannot even be *measured* – optical monitoring equipment is prohibitively expensive, there may be tens of channels sharing the link, and the sheer volume of data crossing the link can incur enormous sampling, storage and analysis costs.

Furthermore, most ultra high capacity networks cannot be realistically *modeled* with today's methods and tools; *viz*., discrete event simulation, analytical techniques, hybrid techniques, and network emulation.

(1) Discrete event simulation is widely used to analyze the performance of low capacity communications networks, but is severely taxed when applied to higher capacity networks. Discrete event simulators typically process 10 to 100 thousand events per second per simulation instance, and a single packet transfer can spawn tens to hundreds of events (depending on the granularity of the simulation). A simulator's resources are usually exhausted when its event list exceeds several million entries. Sophisticated parallel and distributed simulation techniques can extend the total event processing rate to several tens of millions of events per second at high fidelity [Fu90, Fu00], but are unlikely to scale to the traffic volume found in ultra high capacity networks.

(2) Analytical techniques have been successfully used to model networks at an abstract level. They typically describe steady-state network behavior using exact or approximate mathematical formulae derived from queueing theory, Markov processes, and numerous extensions [*e.g.*, Ja57,

Ba75, Ge76, Ke79, Br80, To80, Ch83, Di83], and/or from operational analysis and its extensions [e.g., Bu76, De78, Bu82]. These techniques provide coarse first-order approximations of network performance without the run time and memory requirements and the scalability ceiling of discrete event simulators. However, analytical techniques cannot model a dynamic system in great detail, nor can they model feedback-based algorithms. Ultra high capacity networks are extremely difficult to model analytically due to behaviors induced by: widely disparate traffic sources; closed loop controls (e.g., TCP rate controls, resource provisioning); network elements (e.g., routers with active queue management schemes, switches with preemptive blocking); cross-layer protocol effects (e.g., wavelength routing and IP routing at different layers of the same network); asymmetries in bandwidth requirements and session lengths, etc. The confounding effects caused by interactions among these factors can also have a profound impact on the dynamics of the network.

(3) Hybrid techniques use discrete event simulation to model parts of the network in great detail. The remaining sub-systems are represented by analytic expressions that have been derived by studying these sub-systems in isolation. Crafting a hybrid model requires expert knowledge, careful tuning, and a precise balance between model fidelity and the resources available to the hybrid modeling tool (e.g., CPU cycles, simulator memory, simulation run time) [e.g., Je95, Lu97, Bo02, Ko02].

(4) Network emulation is an effective technique for some types of networks, especially those using IP as the network layer protocol. Emulators run (nearly) unmodified protocol stacks on a few tens to a few thousands of ‘instances’ – commodity PCs, or processes executing on multi-tasking hosts, or both – in (near) real time. The instances are interconnected over a virtual network core that routes traffic, simulates network bandwidth and other resources, and models congestion, latency, loss, and error [*e.g.*, Ri97, Ri98, Fa99, NSF02, Va02, Wh02, Yo02]. Some

implementations emulate transmission of several hundred thousand packets per second per instance. It is not clear whether these data-intensive emulators might realistically scale to the aggregate volume of traffic found in ultra high capacity optical networks (which may be four or five orders of magnitude larger).

Notes