Network Design Problem

- **Goal**
  - **Given**
    - QoS metric, e.g.,
      - Average delay
      - Loss probability
    - Characterization of the traffic, e.g.,
      - Average interarrival time (arrival rate)
      - Average holding time (message length)
  - **Design the system**
  - Three systems will be studied:
    - Circuit switch, e.g., determine the # lines
      \( \rightarrow \) System 1 \( \rightarrow M/M/S/S \ (M/M/S/S/\infty) \)
    - Ideal router output port, e.g., determine link capacity
      \( \rightarrow \) System 2 \( \rightarrow M/M/1 \ (M/M/1/\infty/\infty) \)
    - Real router output port, e.g., determine link capacity and buffer size
      \( \rightarrow \) System 3 \( \rightarrow M/M/1/N \ (M/M/1/N/\infty) \)
Network Performance Evaluation

- Solution methodologies:
  - Mathematical analysis
    - Model this type of process as a Queueing System\(\rightarrow\) good for initial design
  - Simulation techniques\(\rightarrow\) good for more detailed analysis

Network Simulation

- Define network simulation
- Discuss attributes and application of simulation
- Present implementation of simulation systems
- Discuss analysis of simulation results
- Discuss selection of simulation tools
- Provide an overview of ExtendSim8. On the start up ExtendSim window there is:
  - A button for tutorials and a video showing how to build models
  - A link to “ExtendSim for DESS Textbook”, a tutorial on the tool.
  - Other useful tools.
  - There is a link to getting the whole user manual on the class web page. (It is long DO NOT PRINT the whole pdf file.)
A Definition of Communication Network Simulation

Communication network simulation involves generating *pseudo-random sequences* of message lengths and interarrival times (or other input processes, e.g. time varying link quality) then using these sequences to *exercise an algorithmic description of the network operation*.

Attributes of Simulation

- **Simulation Is a Very Flexible Evaluation Tool**
  - General Network Characteristics (Sources, Topology, Protocols, Etc.)
  - Minute Detail
- **Simulation Models Can Be Expensive to Construct**
  - Human Effort
- **Simulation Models Can Be Expensive to Run**
  - Computer Effort
- **Statistical Analysis of the Results Can Be Difficult**
  - Requires Careful Interpretation
- **Difficult to Gain Insight Into System Behavior**
  - Simulate Only a Set of Specific Scenarios
When to Use Simulation

- Whenever Mathematical **Analysis Is Difficult or Impossible**
  - For Studying Transient Behavior of Networks
  - For Systems With Adaptive Routing
  - For Systems With Adaptive Flow Control
  - For Systems With Blocking (Finite Buffers)
  - For Systems With General Message Interarrival Statistics
- For **Validating Analytic Models** and Approximations
  - How Accurate Is the Model?
  - Do Approximations Distort the Results?
- For **Experimentation Without Disturbing** an Operational System
  - Test Possible Modifications and Adjustments

Modeling Elements for Communication Networks

- **Traffic and Input Processes**
  - Message Arrival Process
    - Often Interarrival Times
  - Message Lengths
  - Other Message Attributes
    - Service Class
    - Error models
- **Algorithmic Descriptions of Network Processing**
  - Protocols
  - Links and Queues
  - Routing
Sample Realization of an Input Process

<table>
<thead>
<tr>
<th>Message number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interarrival time between i+1 and i message (seconds)</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>Length of i-th message (seconds)</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

Time Step Approach to Network Simulation

- Approaches to Discrete Event Simulation
  - Time Step Approach (Fixed Increment Time Advance)
  - Event-Scheduling Approach
- Fixed Increment Time Advance
  - Choice of Increment Important
  - Too Large: Multiple Events Happen In Single Step
  - Too Small: Wasted Processing Time
  - Update System States at End of Each Fixed Time Interval
Event Scheduling Approach to Network Simulation

- Variable Time Advance
  - Advance Time To Next Occurring Event
- Update System State Only When Events Occur
  - For Example, Arrivals or Departures
- Event Calendar
  - Events: Instantaneous Occurrences That Change the State of the System
  - An Event is Described by
    - The Time the Event is to Occur
    - The Activity to Take Place at the Event Time
  - The Calendar is a Time-Ordered List of Events

Event Scheduling Approach: Simplified Flow Control

An Executive (or Mainline) Controls the Selection of Next Event

- Use Event List to determine next event to process
- Advance simulation clock to event time
- Update system state using event routines
- Update event list using event routines
Event Scheduling for Simple Statistical Multiplexer

Verification and Validation of Simulation Models

- **Model**
  - Mathematical (Algorithmic) Description of Behaviour of “Real Thing”

- **Verification**
  - Determining Whether the Simulation Model Performs As Intended
  - In Programming Terminology, “Debugging”
  - Example: Is $M/M/1$ Model Producing Exponential Message Lengths?

- **Validation**
  - Determining Whether the Simulation Model Itself Is an “Accurate” Representation of the Communication Network Under Study (the “Real Thing”)
  - Example: Is the Assumption of Exponential Message Lengths Accurate?
Verification Methods

- Modular Development and Verification
  - Break Large System Into Smaller Components
  - Verify Component-by-Component
- Structured Walk-Through
  - Step-by-Step Analysis of Behavior for Simple Case

Verification Methods

- Event Trace
  - Detailed Analysis of Model Behavior
  - Compare to Walk-Through Analysis
- Model Simplification and Comparison to Analytic Results
- Graphical Display of Network Status As the Model Progresses
  - To “See” What Is Happening As It Happens
Some Comments on Validation

- Simulation Models Are Always Approximations
- A Simulation Model Developed for One Application May Not Be Valid for Others
- Model Development and Validation Should Be Done Simultaneously
- Specific Modeling Assumptions Should Be Tested
- Sensitivity Analysis Should Be Performed
- Attempt to Establish That the Model Results Resemble the Expected Performance of the Actual System
- Generally, Validation Is More Difficult Than Verification

Analysis of Results: Statistical Considerations

- Starting Rules
  - Overcoming Initial Transients
  - An Initial Transient Period Is Present Which Can Bias the Results
  - Achieving Steady State
    - Use a Run-in Period:
      - Determine $T_b$ Such That the Long-Run Distribution Adequately Describes the System for $t > T_b$
    - Use a “Typical” Starting Condition (State) to Initialize the Model
- Quality of Performance Estimates
  - Variance of Estimated Performance Measures
Quality of Performance Estimates

- Simulation results are like laboratory measurements, they can be modeled as random variables.
- Performance estimates should have acceptable variance.
- The more observation reduces the variance.
- HOWEVER→ Observations taken from network simulation will be correlated.
  - Cannot directly apply standard statistical approaches based on iid (Independent, Identically Distributed) observations.

Dealing with Lack of Independence

- Simple Replication: Multiple Simulation Runs
  - Assume Results for Each Replication Are Independent
  - Can be Inefficient Because of Multiple Startup Periods
Criteria for Selecting a Network Simulation Tool

- Availability
- Cost
- Usage
- Documentation
- Ease of Learning
- Computation Efficiency
- Flexibility
- Portability
- User Interface
- Extendibility

Common Tools

- ns-3
  (http://www.nsnam.org/)
- Opnet
  (http://www.opnet.com/)
- QualNet
  (http://www.scalable-networks.com)
- ExtendSim

Guidelines to Network Modeling and Simulation

- Things to Know
  - Know the Customer
  - Know the Network
  - Know the Important Performance Metrics

- Things to Do
  - Establish a Credible Model
  - Expect the Model to Evolve → Plan for success
  - Apply Good Software Management Techniques
Conclusions

- Simulation Can Be an Important Tool for Communication Network Design and Analysis
- Care and Thought Must Go Into Construction of Communication Network Models
- Care and Thought Must Go Into Interpretation of Model Output

Extend® Overview

- Allows Graphical Description of Networks
  - Sources, Links, Nodes, Etc.
- Data Flow Block Diagrams
- Hierarchical Structure to Control Complexity
- Be sure and create libraries when creating complex models
Network Performance Evaluation: Elements of a Queueing System

System

Queue

Blocked customers

Server

Server

Server

Departing customers

Number in system

Delay

Number in Queue

Number in Servers

Network Performance Evaluation: Elements of a Queueing System

System

Queue

Blocked customers

Server

Server

Server

Departing customers

Number in system

Delay

Number in Queue

Number in Servers
Network Performance Evaluation: Specific cases for theoretical analysis

- Assumptions:
  - Interarrival times are exponentially distributed
  - Service times are exponentially distributed
    - Holding time
    - Packet length
  - Types of systems
    - One server (Stat Mux)
      - Infinite memory
      - Finite memory
    - S servers and a system size of S (Circuit Switch)

Network Performance Evaluation: Approach

- Analysis of a pure birth process to characterize arrival processes
- Extension to general birth/death processes to model arrivals and departures
- Specialization to the specific cases to find:
  - Probability of system occupancy,
  - Average buffer size,
  - Delay,
  - Blocking probability

- Goal: Design and analyze statistical multiplexers and circuit switching systems
Network Performance Evaluation: Analysis of a Pure Birth Process

Arrivals and no departures

$\lambda$ = Arrival rate

Only Births (Arrivals) Allowed

$K = \text{System State (number in system)}$
- number of arrivals for 0 to t sec
- number in system at time t

Goal: Find Prob \[k \text{ arrivals in a t sec interval}\]
Network Performance Evaluation:
Analysis of a Pure Birth Process

- The number represents the *State* of the system. In networks this is usually the number in the buffer plus the number in service. *The system includes the server.*
- The time to clock the message bits onto the transmission facility is the service time. The server is the model for the transmission facility.
- Goal: Find Prob [k arrivals in a t sec interval]

---

Network Performance Evaluation:
Analysis of a Pure Birth Process: Assumptions

- Prob[ 1 arrivals in $\Delta t$ sec ] = $\lambda \Delta t$
- Prob[ 0 arrivals in $\Delta t$ sec ]
  = $1 - \lambda \Delta t$
- Number of arrivals in non-overlapping intervals of times are statistically independent random variables, i.e.,
  Prob [ N arrivals in $t$, $t+T$ AND M arrivals in $t+T$, $t+T+\tau$] =
  Prob [ N arrivals in $t$, $t+T$]*[M arrivals in $t+T$, $t+T+\tau$]
**Network Performance Evaluation:**

How to get to state $k$ at $t + \Delta t$?

- Define probability of $k$ in the system at time $t$
  
  $= \text{Prob}[k, t]$

- Probability of $k$ in the system at time $t + \Delta t$
  
  $= \text{Prob}[k, t + \Delta t] = \text{Prob}[(k \text{ in the system at time } t \text{ and 0 arrivals in } \Delta t) \text{ or (k-1 in the system at time } t \text{ and 1 arrival in } \Delta t)]$

  $= (1 - \lambda \Delta t) \text{Prob}[k, t] + \lambda \Delta t \text{Prob}[k-1, t]$
Network Performance Evaluation: Analysis

- Rearranging terms
  \[
  \frac{(\text{Prob}[k, t+ \Delta t] - \text{Prob}[k,t])}{\Delta t} + \lambda \text{Prob}[k,t] = \lambda \text{Prob}[k-1,t]
  \]

- Letting \( \Delta t \to 0 \) results in the following differential equation:
  \[
  \frac{d\text{Prob}[k, t]}{dt} + \lambda \text{Prob}[k, t] = \lambda \text{Prob}[k - 1, t]
  \]

Network Performance Evaluation: Analysis

- For \( k = 0 \) the solution is:
  \[
  \text{Prob}[0, t] = e^{-\lambda t}
  \]

- For \( k = 1 \) the solution is:
  \[
  \text{Prob}[1, t] = \lambda t e^{-\lambda t}
  \]

- For \( k = 2 \) the solution is:
  \[
  \text{Prob}[2, t] = \frac{(\lambda t)^2 e^{-\lambda t}}{2}
  \]
Network Performance Evaluation: Analysis

- In general the solution is a Poisson probability mass function of the form:

\[
\text{Prob} [k, t] = \frac{(\lambda t)^k e^{-\lambda t}}{k!}
\]

**A Poisson pmf of this from has the following moments:**

\[
E[k] = \lambda t
\]

\[
Var[k] = \lambda t
\]

**Poisson Arrival Process**

The number of arrivals in any T second interval follows a Poisson probability mass function.
Network Performance Evaluation: Interarrival Time Analysis

The interarrival time is denoted as $T_a$. The probability of $t < T_a < t + \Delta t$ can be calculated as follows:

\[
Prob[t<T_a<t+\Delta t] = Prob[0 \text{ arrivals in } t \text{ sec and } 1 \text{ arrival in } \Delta t]
\]

\[
Prob[t<T_a<t+\Delta t] = Prob[k=0,t]Prob[k=1, \Delta t]
\]

\[
Prob[t<T_a<t+\Delta t] = (e^{-\lambda t})\lambda \Delta t e^{-\lambda \Delta t}
\]

Let $\Delta t \to 0$ results in the following:

\[
Prob [ t < T_a < t + dt ] = \int_{T_a} f(t) dt = \lambda e^{-\lambda t} dt
\]

So:

\[
f_{T_a} (t) = \lambda e^{-\lambda t} \text{ for } t > 0 \quad f_{T_a} (t) = 0 \text{ for } t < 0
\]
Network Performance Evaluation: Interarrival Time Analysis

MAIN RESULT:
The interarrival time for a Poisson arrival process follows an exponential probability density function.

\[ E[T_a] = \frac{1}{\lambda} \quad \text{Var}[T_a] = \frac{1}{\lambda^2} \]

Network Performance Evaluation: Birth/Death Process Analysis

Now allow arrivals and departures.
The Model for the Birth/Death Process

Note that the arrival and service rates are now state dependent.
Network Performance Evaluation: Birth/Death Process Analysis

- The departure process is Poisson--
- $\text{Prob}[1 \text{ departure in } \Delta t \text{ sec when the system is in state } k] = \mu_k \Delta t$
- $\text{Prob}[0 \text{ departure in } \Delta t \text{ sec when the system is in state } k] = 1 - \mu_k \Delta t$
- Number of departures in non-overlapping intervals of times are statistically independent random variables
- $\text{Probability}[\text{arrival AND departure in } \Delta t] = 0$
Network Performance Evaluation: Birth/Death Process Analysis

Poisson service process implies an exponential probability density function for the message length.

To solve for the state probabilities:
Follow the procedure used for the pure birth process and use the transitions shown.
Network Performance Evaluation:
Birth/Death Process Analysis

- Specific queueing systems are modeled by

  - Setting state dependent arrival rates, $\lambda_k$
  - Setting the state dependent service rates, $\mu_k$
  - Solving for the steady state probabilities

  (or any queueing theory book)

Network Performance Evaluation:
Queueing System Notation (Kendall's notation)

- A / b / m / K / L
  - A = type of arrival process
  - b = type of service process
  - m = number of servers
  - K = maximum number of elements allowed in the system = system size
  - L = population size (if L missing then $\infty$)
Network Performance Evaluation:
Special cases: A / b / m / K / L

- A = M => the arrival process is Poisson and the interarrival times are independent, identically distributed exponential random variables. (M = Markov)
- b = M => the service process is Poisson and the interdeparture times are independent, identically distributed exponential random variables.
- A or b = G => times are independent, identically distributed general random variables.
- A or b = D => times are deterministic, i.e., fixed times

Examples:
- M/M/1/∞/∞ (Ideal router output port)
- M/M/1/N/∞ (Real-finite-buffer router output port)
- M/M/S/S/∞ (Circuit Switch)

Network Performance Evaluation:
M/M/1

- No limitation on buffer size means that the arrival rate is independent of state or \( \lambda_k = \lambda \)
- Only one server means that the service rate is independent of state or \( \mu_k = \mu \)
Network Performance Evaluation: M/M/1

Solving for the state occupancy probabilities

\[ P[k] = \rho^k (1 - \rho) \]

where

\[ \rho = \frac{\lambda}{\mu} = \frac{\lambda T_x}{\mu} = \text{load} = \text{traffic intensity} \]

where

\[ T_x = \text{Average message length in sec} = \frac{L}{C} \]

where

\[ L = \text{Average message length in bits} \]
\[ C = \text{Link capacity in b/s} \]

Note \( \rho \) is the traffic intensity and \( \rho < 1 \).

Note \( \Rightarrow \rho = \frac{\lambda L}{C} = r_g/C \)

Network Performance Evaluation: M/M/1

If the load is greater than 1 then the systems is not stable and the buffer occupancy grows without bound.

The expected number in the systems is

\[ E[K] = \frac{\rho}{1 - \rho} \]

and the variance is

\[ \text{Var}[K] = \frac{\rho}{(1 - \rho)^2} \]
Network Performance Evaluation: M/M/1

Load = 0.5  \( E[K] = 1 \)

Network Performance Evaluation: M/M/1

Load = 0.95  \( E[K] = 19 \)  Final simulated value = 12.05
Network Performance Evaluation: M/M/1

Network Performance Evaluation: M/M/1/N

- Only one server means that the service rate is independent of state or $\mu_k = \mu$
- The limitation on system size means that the arrival rate is dependent of
  $\lambda_k = \lambda$ for $k < N$
  $\lambda_k = 0$ for $k \geq N$

Arrivals to a full system are blocked so there can be no arrivals to a full system.
Network Performance Evaluation: M/M/1/N

Solving for the state occupancy probabilities

\[ P[k] = \frac{(1 - \rho) \rho^k}{1 - \rho^{N+1}} \quad \text{for } k \leq N \]

\[ P[k] = 0 \quad \text{for } k > N \]

Network Performance Evaluation: M/M/1/N

- The Quality of Service (QoS) metric in this case is the probability of blocking.
- For a M/M/1/N queue the blocking probability is given by

**Probability of Blocking**  \( = P_B = \)

\[ P_{k=N} = \frac{(1 - \rho) \rho^N}{1 - \rho^{N+1}} \]

Design Problem: Given \( P_B \) and \( \rho \) find \( N \)
(recommend constructing a spread sheet)
Network Performance Evaluation: M/M/1/N

Average Holding Time = 0.95 Arrival rate = 1
(load = 0.95) → Theory $P_B = 0.23$
Simulated $P_B = 0.219$

What is going on during this time?

Network Performance Evaluation: M/M/S/N

The limitation on system size means that the arrival rate is dependent of state or

$\lambda_k = \lambda$ for $k < N$

$\lambda_k = 0$ for $k \geq N$

Arrivals to a full system are blocked so there can be no arrivals to a full system.
Multiple servers means that

\[ \mu_k = k\mu \text{ for } k \leq S \]

\[ \mu_k = S\mu \text{ for } k > S \]
Network Performance Evaluation: Erlang B formula

Solving for the state occupancy probabilities

\[ P[k] = \frac{\rho^k}{N \sum_{n=0}^{k} \frac{\rho^n}{n!}} \quad k=0\ldotsN \]

Relationship among \( P_B \), \( N \), \( \rho \) found using provided table or web Erlang B calculator

\[ P_B = P[k = N] = \frac{\rho^N}{N! \sum_{n=0}^{N} \frac{\rho^n}{n!}} \]

Network Performance Evaluation: M/M/S/N

Holding time=3min, Arrival rate=0.833 calls/min

\[ N = 4 \]

\[ \text{Load} = 3 \times 0.833 = 2.5 \text{ Erlangs} \]

\( \Rightarrow \) Theory \( P_B = 0.15 \) (From Erlang B table)

Simulated \( P_B = 0.198 \)
Network Performance Evaluation: Example

- Design of a building phone system. The design goal is to minimize the number of lines needed between the building and the phone company. The blocking QoS is specified to be 5%.
- A building has four floors, on each floor is a separate department. Each department has 22 phones, each busy 10% of the time during the busy hour.

Network Performance Evaluation: Example-Case A

- Acquire one telephone switch for each floor.
- 2.2 Erlangs/floor & B=5% gives:
- 5 lines/floor or 20 lines for the building.
Network Performance Evaluation: Example-Case B

- Acquire one telephone switch for the building.
- 88 phones @ .1 Erlang/phone = 8.8 Erlangs
- 8.8 Erlangs & B=5% gives:
- 13 lines for the building
- Select Case B

Network Performance Evaluation: Delay Analysis for M/M/1

- Average delay = average transmission time + queueing time

- Queueing time =
  
  (average number of messages in system ahead)*
  (average time to serve each message)
Network Performance Evaluation: Delay Analysis for M/M/1

- Average number ahead: \( \frac{\rho}{1 - \rho} \)
- Average time/mess age: \( \frac{1}{\mu} = \frac{L}{C} = \bar{T}_x \)
- \( E[T] \) = Average delay: \( \frac{1}{\mu} + \frac{1}{\mu} \left( \frac{\rho}{1 - \rho} \right) \)
- \( E[T] = \frac{1}{\mu(1 - \rho)} = \frac{\bar{T}_x}{(1 - \rho)} = \frac{1}{\mu - \lambda} \)
Network Performance Evaluation: Example

- Which is better?

Assume you have 10 PC's each with a 100 kb/s line to a server.
- Assume 1000 bits/packet
- Assume 50 packets/sec/PC
- Service time =
  \[1000 \text{bits}/100,000 \text{ b/s} = 10\text{ms}\]
- Traffic intensity \(=\frac{50 \times 1000}{100,000} = 0.5\)
- Average delay \(=\frac{10\text{ms}}{.5} = 20\text{ms}\)
Network Performance Evaluation: Example-continued

- Assume the 10 PC’s are connected to a statistical multiplexer at 1 Mb/s
- Traffic intensity = 0.5
- Service time = 1 ms
- Average delay = 2 ms
- Show that traffic aggregation helps

---

Do the Design Problem:
Find the link capacity between the stat mux and the server such that the delay is 20 ms.

\[ E[T] = \frac{1}{\mu - \lambda} = 0.02 = \frac{1}{50} \]

\[ \mu - \lambda = 50 \]
\[ \mu = 550 \]

\[ C = (1000 \text{bits/packet})(550 \text{packets/sec}) = 550 \text{kb/sec} \]
Network Performance Evaluation: Design Example

Assume each customer and printer is connected using Ethernet, i.e. at 1 Gb/s.

How fast does the link between Youngberg and the computer center have to be to guarantee all the customers can use the 1 Gb/s.

- C = 55 Gb/s
- Too expensive
Network Performance Evaluation: Design Example

- Customer performance requirements:
  - Delay < 100 ms (use M/M/1 analysis to find C)
  - Loss < 10% (use M/M/1/N analysis to find N)
- Assume customer traffic:
  - Average packet length = 9000 bytes/packet
  - 55 sources
  - Packets are generated at a rate of 2 per second/source

\[ \lambda = 55 \times 2 = 110 \text{ packets/sec} \]
\[ \mu = 120 \]
\[ C = \mu \times L = 120 \times 9000 \text{ Bytes/packet} \times 8 \text{ bits/Byte} = 8.64 \text{ Mb/s} \]

\[ \rho = \frac{\text{Rate_in}}{C} = \frac{7.92 \text{ Mb/s}}{8.64 \text{ Mb/s}} = 0.916 \]
\[ \rho = 0.916 \text{ and Blocking Prob = 0.1} \Rightarrow K = 7 \]

Final design is:
- C = 8.64 Mb/s
- Average system size > 7 packets