Sysplex Modeling:

Modeling Distributed Transaction Response Times As Impacted by In-Storage Buffer Accesses

Tim R. Norton, Sr. Staff Member

Information Technology
networkMCI Services
2424 Garden of the Gods Road
Colorado Springs, CO 80919

(719) 535-1163
Tim_R_Norton@mcimail.com

Rocky Mountain CMG
March 21, 1996

Agenda

- Background
- The Buffer Access Problem
- Modeling the Problem
- The Model
- Results
- Conclusion

Copyright © 1996, MCI Telecommunications Corporation

Background

- **Processor**
  - Contains One or More CPU's (Engines, IP, CP, x-way, etc.)
  - Shared Memory
  - IBM CEC - Central Electronic Complex
- **Growth**
  - Make Each CPU Faster
  - Add More CPU's
  - Connect Multiple Together
    - IBM Sysplex
    - IBM PTS - Parallel Transaction Server

Copyright © 1996, MCI Telecommunications Corporation

Background

- **Coupling Facility**
  - **Inter-connect Large Number of Processors**
  - **Part of IBM S/390 Parallel Sysplex**
  - **High Speed Fiber-Optic Links**
  - **Memory Holds Different Structures**
    - **Cache**
    - **Lock**
    - **List**

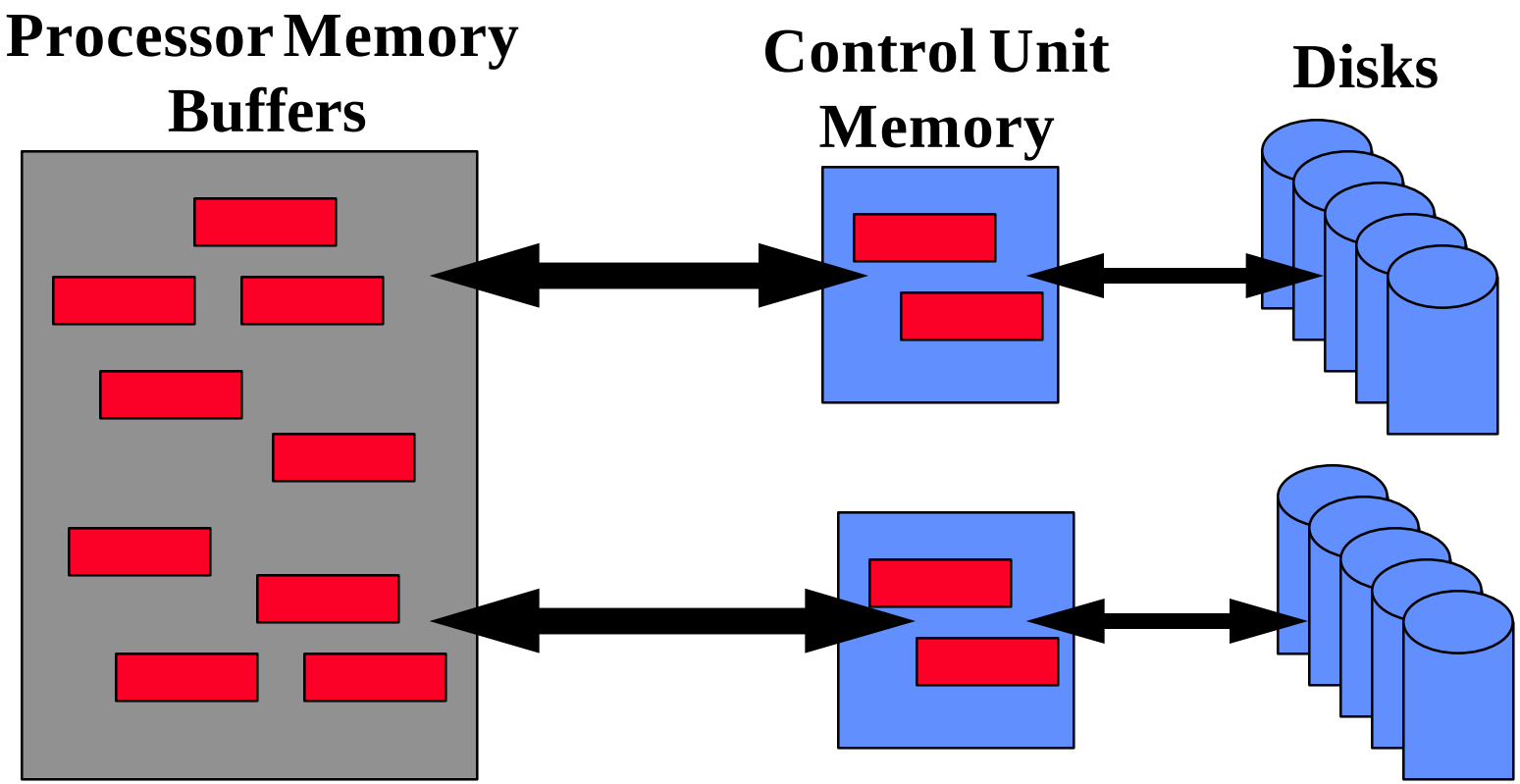
Copyright © 1996, MCI Telecommunications Corporation

Background

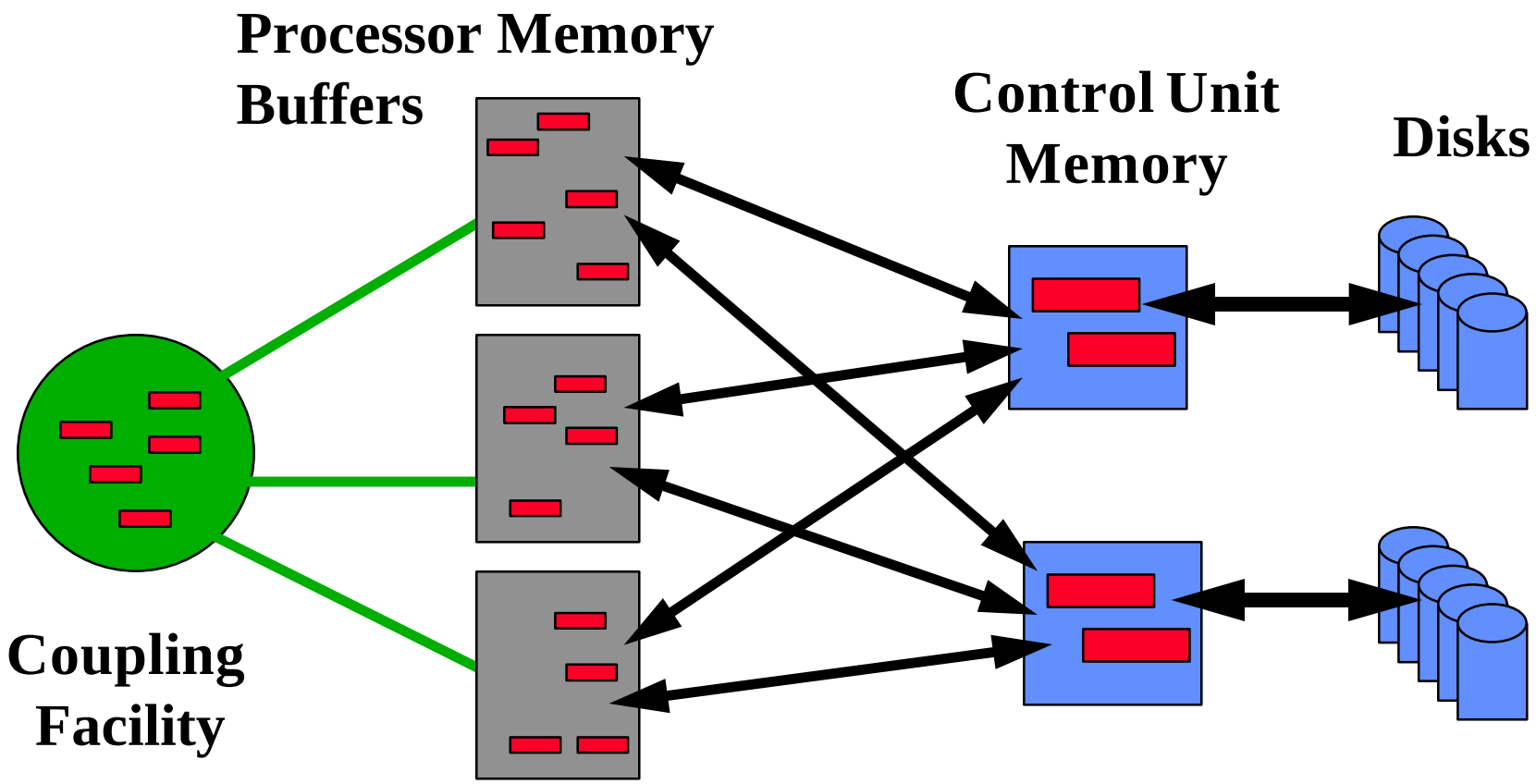
- **Data Base Systems**
  - **Optimized for the Current Environment**
  - **Data In Memory to Reduce I/O Times**
  - **Data Shared by Multiple Transactions**
  - **Data Shared by all CPU's in the Processor**
  - **Share Data Across Processors**
    - **Use Coupling Facility for:**
      - ✓ **Locks**
      - ✓ **Cache**

Copyright © 1996, MCI Telecommunications Corporation

Single Processor



Copyright © 1996, MCI Telecommunications Corporation



Copyright © 1996, MCI Telecommunications Corporation

The Buffer Access Problem

- **Overview**
  - **What Causes It?**
    - **Moving Applications**
    - **One Processor to Several Processors**
    - **Response Time Critical Transactions**
    - **Highly Optimized Workloads**
  - **Potential Environments**
    - **IBM PTS**
    - **Client/Server**
    - **“Getting Off the Mainframe”**

Copyright © 1996, MCI Telecommunications Corporation

The Buffer Access Problem

- **Elements of the Problem**
  - **CPU Size**
    - **Impact of Smaller CPU's**
    - **Single I/O Can Equal 10,000's Instructions**
  - **Buffer Hit Ratio**
    - **Hits vs. Misses**
    - **Locks vs. Cache**
  - **Transaction Routing - What to Optimize**
    - **Locality of Reference**
    - **Processor Utilization**

Copyright © 1996, MCI Telecommunications Corporation

The Buffer Access Problem

- **Details**
  - **What Can Happen**
    - **Buffer Hit**
    - **Buffer Miss / Cache Hit**
    - **DASD Access**
  - **What Causes Buffer Misses For Already Read Records**
    - **Time - LRU'ed Out**
    - **Update - Current Copy Invalid**
    - **Routing - Not in Shared Memory**
  - **Drivers**
    - **Locality of Reference**
    - **Number of Processors**
    - **Read/Write Ratio**
    - **Memory Size**

Copyright © 1996, MCI Telecommunications Corporation

The Buffer Access Problem

- **Not Considered**
  - **Lock delays**
  - **Enqueue delays**
  - **Queuing Caused by Smaller CPU**
  - **Security**
  - **Logging**

Copyright © 1996, MCI Telecommunications Corporation

Modeling the Problem

- **Overview**
  - **Large Changes in the Environment**
  - **Many Drivers of Change**
  - **Potential Impact is Large**
  - **But Impact Could be Minor**
  - **Application Changes Required?**
  - **Which Configuration Works Best?**

Copyright © 1996, MCI Telecommunications Corporation

Modeling the Problem

- **Example**
  - **Hypothetical System**
  - **Transaction Response Time Example**
    - **CPU Time: .005 seconds**
    - **Data Base Accesses - 100**
      - ✓ **Buffer hit: .0001 seconds**
      - ✓ **Buffer miss: .0300 seconds**
    - **Total Response Time**
      - ✓ **From 0.015 seconds (.005 + 100*.0001)**
      - ✓ **To 3.005 seconds (.005 + 100*.03)**
  - **Instructions Executed in the Time of an I/O**
    - **60 MIPS CPU**
    - **1.8M Instructions (60,000,000 * .03 = 1,800,000)**

Copyright © 1996, MCI Telecommunications Corporation

What Model to Use?

- **Analytic Model**
  - Deals With Averages
  - Assumes Homogenous Transactions
  - Generalized Cache Management Algorithms
  - Transaction Order Not Preserved
- **Simulation Model**
  - Describes Individual Behavior
  - Impact of Outliers
  - LRU Cache Management Algorithms
  - Transaction Order Preserved
    - Routing
    - Buffer Invalidation

Copyright © 1996, MCI Telecommunications Corporation

Designing the Model

- **Variable Number of Processors**
  - Use One Processor to Calibrate Model
  - Use More to Investigate Changes
- **Control Over Buffer Locality of Reference**
  - How to Identify the Buffer Accesses per Transaction?
  - How to Describe Locality of Reference?
- **Processor Memory Size**
  - One Large vs. Many Small
- **Transaction Profiles**
  - What Makes Transactions Different?

Copyright © 1996, MCI Telecommunications Corporation

Using the Model

- **What Should It Do?**
  - Response Time Predictions
  - Buffer Hit Ratio Predictions
  - Processor Memory Sizing
- **What are the Benefits?**
  - Additional Processors Analysis
  - Data Base Re-design Analysis
  - Additional Memory Analysis
- **What are the Problems?**
  - Data Collection: Which Buffers Does a Transaction Use?
  - Routing Techniques: Queue Length, Tran Content, Other?

Copyright © 1996, MCI Telecommunications Corporation

Modeling Tool

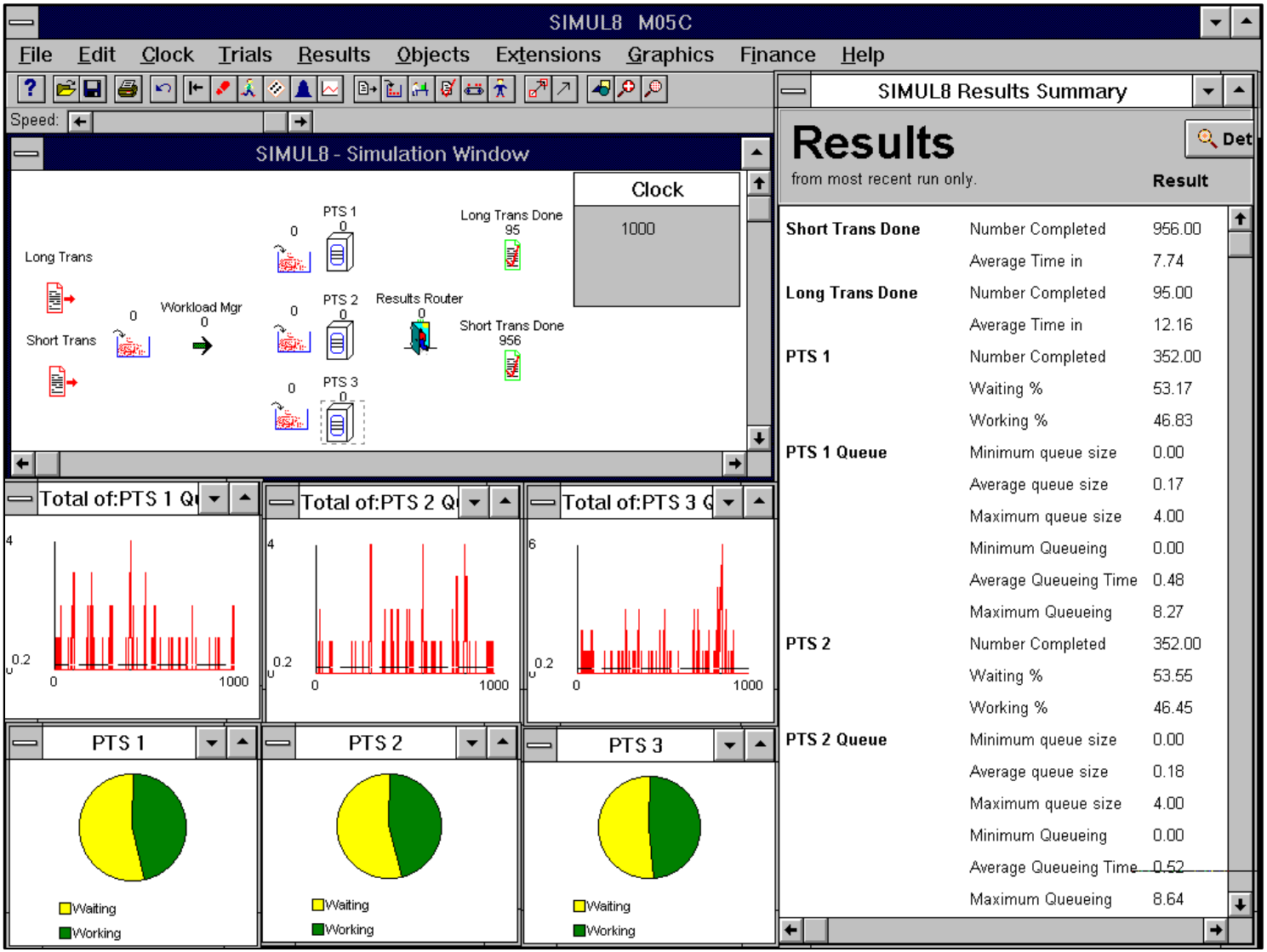
- **Simulation Tool**
  - **Simul8 from Visual Thinking International Ltd**
    - ✓ http://www.colloquium.co.uk/www/vti/
  - **General Purpose**
  - **GUI Interface**
  - **Links to Excel and Visual Basic**
  - **Animation Display**
  - **Inexpensive**

Copyright © 1996, MCI Telecommunications Corporation

Simul8 Model

Sysplex Modeling:

Modeling Distributed Transaction Response Times As Impacted by In-Storage Buffer Accesses



Copyright © 1996, MCI Telecommunications Corporation

Current State of the Model

- **Current Features**
  - Hit Ratio Controlled by Transaction Type
  - Arrival Rate and Distribution by Transaction Type
  - Service Time by Transaction Type
  - Some Routing Choices
- **Investigate Extremes**
- **Identify Areas For:**
  - Additional Research
    - Application Understanding
    - Sysplex Understanding
  - Model Development
    - What to Implement Next

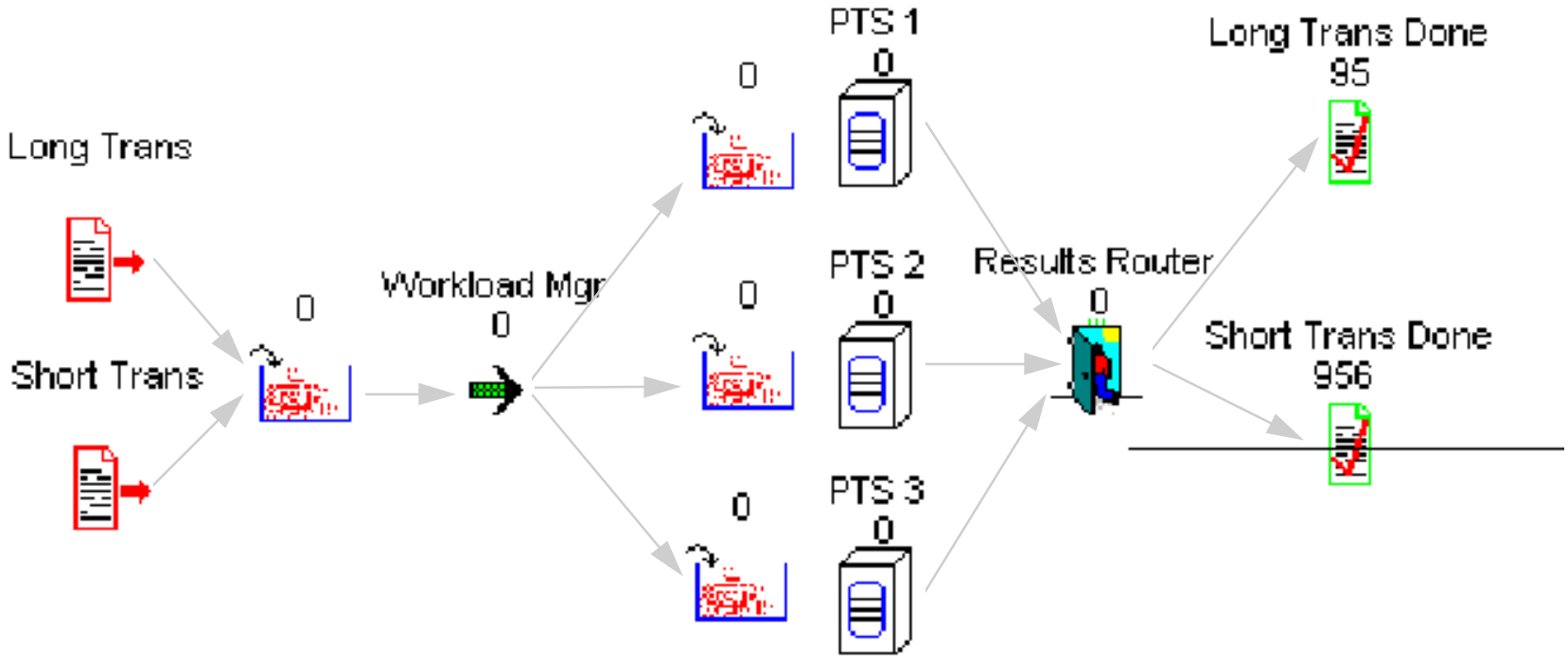
Copyright © 1996, MCI Telecommunications Corporation

Model Assumptions

- **Buffer Hit Ratio**
  - Shared Memory System (Big CEC) - 98.5%
  - Distributed Memory System (PTS) - 90.0%
- **Processor Speed**
  - PTS = Big CEC / 3
- **Transaction Service Time**
  - Long = x * Short
  - x = 5, 10 or 15 (depending on the run)
- **No Parallelism**
  - Single Transaction Active at a Time

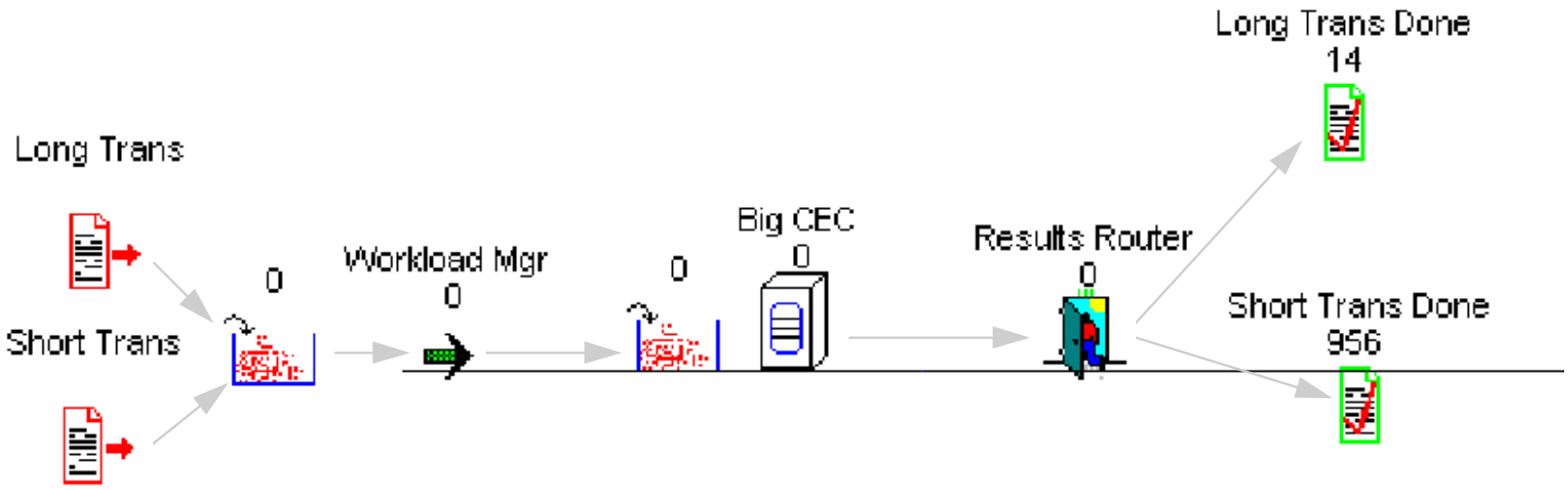
Copyright © 1996, MCI Telecommunications Corporation

Three PTS Sysplex



Copyright © 1996, MCI Telecommunications Corporation

Single Big CEC



Copyright © 1996, MCI Telecommunications Corporation

Model Runs

- **Run Groups**
  - PTS Sysplex with Shortest Queue Routing (SQ)
  - PTS Sysplex with Circulate Routing (C)
  - Single Big CEC (B)
- **Three Long Transaction Service Times**
  - 15 Time Units
  - 10 Time Units
  - 05 Time Units
- **Nine Total Runs**
  - Each Long Transaction Service Time for Each Group
  - Run Name: M<15/10/05><SQ/C/B>

Copyright © 1996, MCI Telecommunications Corporation

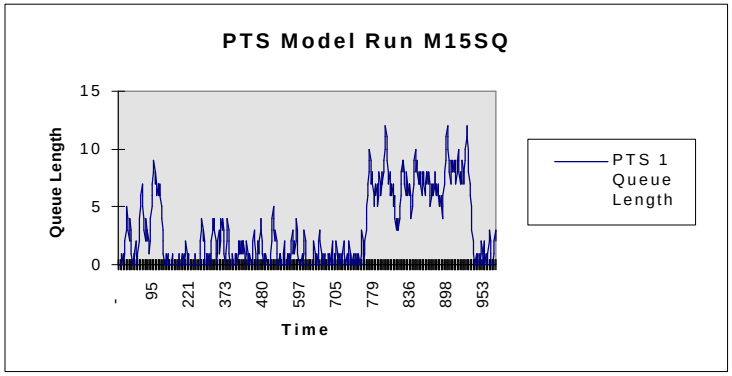
Model Parameters

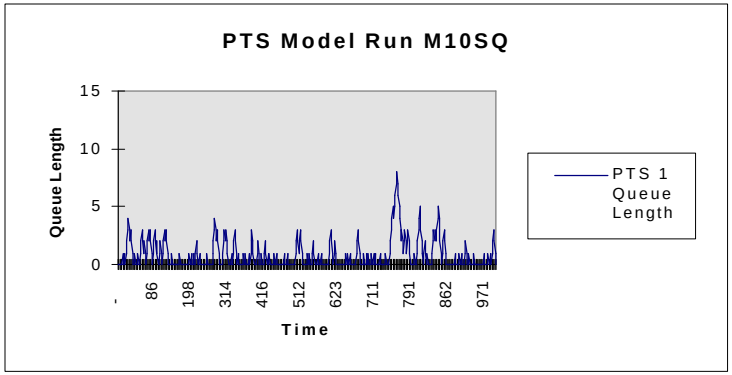
	Run:	M15SQ	M10SQ	M05SQ	M15C	M10C	M05C	M15B	M10B	M05B
Model Parameters										
Run Time		1000	1000	1000	1000	1000	1000	1000	1000	1000
Long Trans										
	Interarrival Time	10	10	10	10	10	10	75	75	75
	Std	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	Distribution	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp
Short Trans										
	Interarrival Time	1	1	1	1	1	1	1	1	1
	Std	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	Distribution	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp	Neg Exp
Workload Mgr										
	Timing	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Std	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Distribution	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal
	Routing	Shortest Q	Shortest Q	Shortest Q	Circulate	Circulate	Circulate	n/a	n/a	n/a
Service Times										
	Hit Timing	1	1	1	1	1	1	0.3	0.3	0.3
	Std	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Distribution	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal
	Miss Timing	15	10	5	15	10	5	14.3	9.3	4.3
	Std	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5
	Distribution	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal	Normal

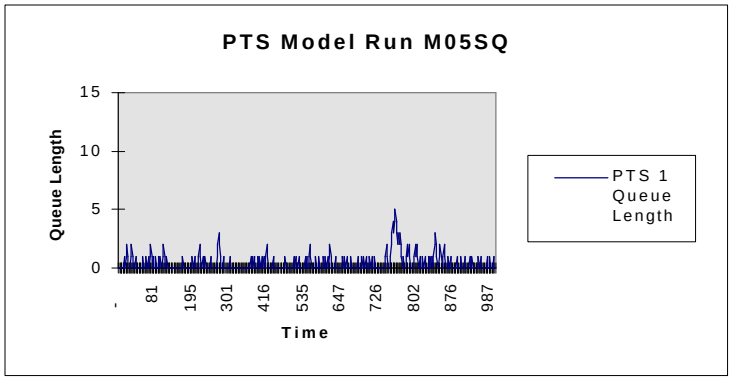
Copyright © 1996, MCI Telecommunications Corporation

Shortest Queue Model Run

- **Average Response Times**
  - **M15SQ - 9.9**
  - **M10SQ - 3.9**
  - **M05SQ - 2.0**
- **Longest Queue**
  - **M15SQ - 13**
  - **M10SQ - 8**
  - **M05SQ - 6**
- **Better than Circulate**
- **Long Trans Cause Backups**





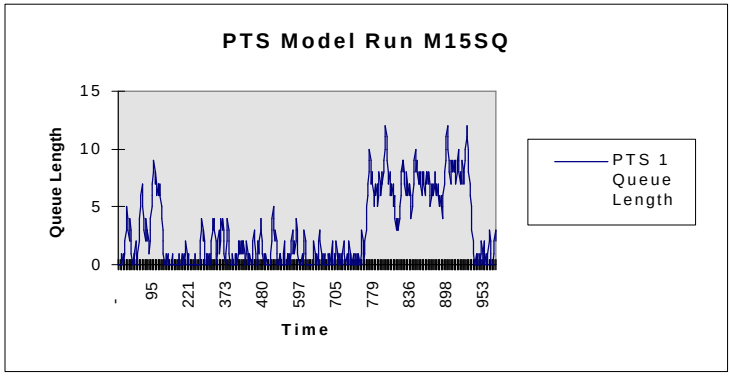


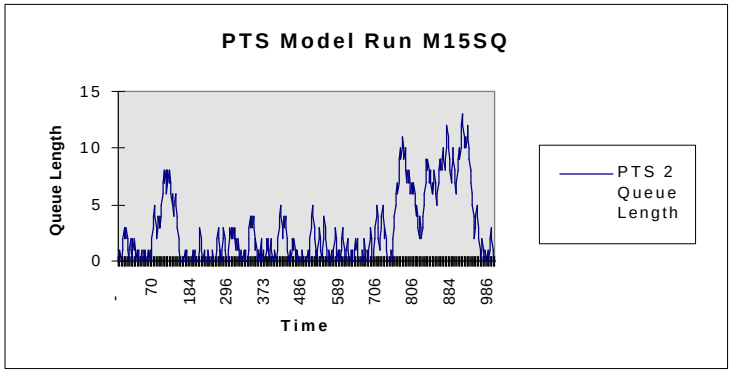
Copyright © 1996, MCI Telecommunications Corporation

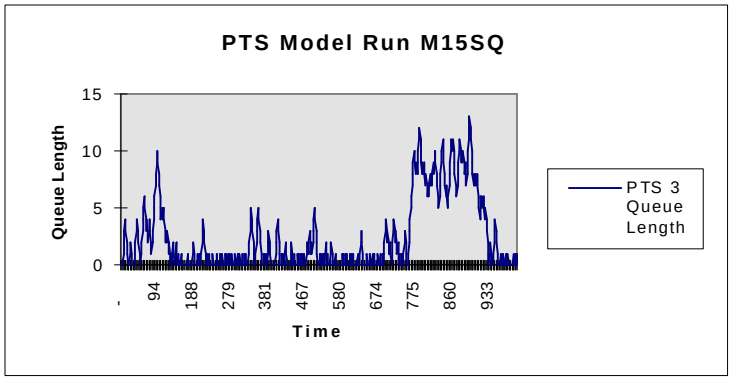
Shortest Queue 15 Model Run

- **Detailed Results:**

M15SQ	PTS 1	PTS 2	PTS 3
Minimum Queue Size	0.0	0.0	0.0
Average Queue Size	2.6	2.9	2.6
Maximum Queue Size	12.0	13.0	13.0
Minimum Queuing	0.0	0.0	0.0
Average Queue Time	6.9	9.3	6.8
Maximum Queuing	30.1	47.6	46.0
Number Completed	365.0	312.0	374.0
Waiting %	22.0	16.9	23.6
Working %	78.0	83.1	76.4
Aaverage Service Time	2.1	2.7	2.0





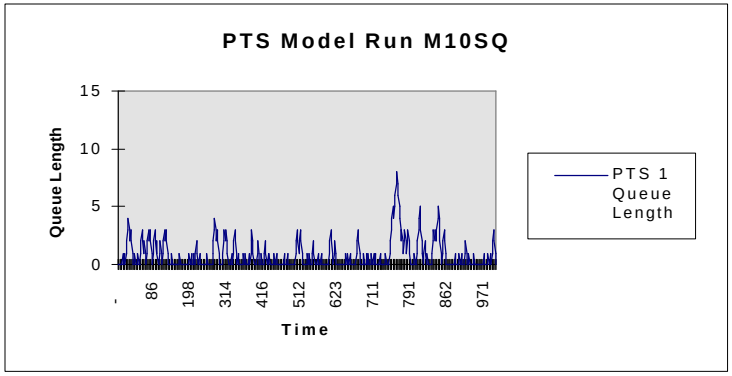


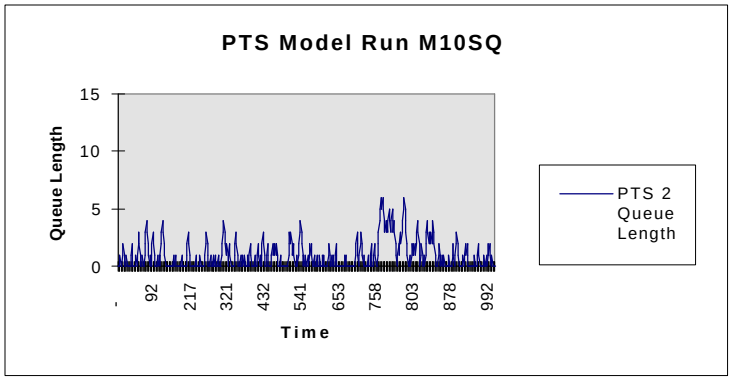
Copyright © 1996, MCI Telecommunications Corporation

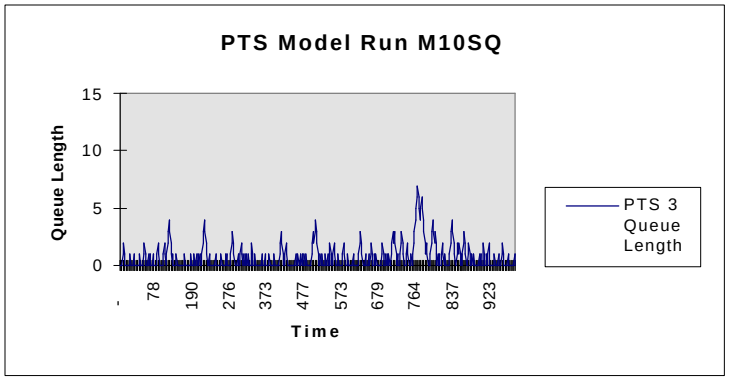
Shortest Queue 10 Model Run

- **Detailed Results:**

M10SQ	PTS 1	PTS 2	PTS 3
Minimum Queue Size	0.0	0.0	0.0
Average Queue Size	0.6	0.9	0.6
Maximum Queue Size	8.0	6.0	7.0
Minimum Queuing	0.0	0.0	0.0
Average Queue Time	1.9	2.5	1.8
Maximum Queuing	19.0	19.8	19.3
Number Completed	349.0	349.0	356.0
Waiting %	29.9	29.9	39.3
Working %	70.1	70.1	60.7
Aaverage Service Time	2.0	2.0	1.7





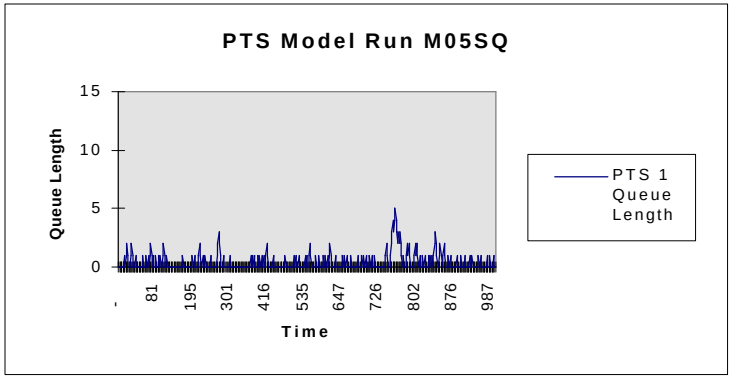


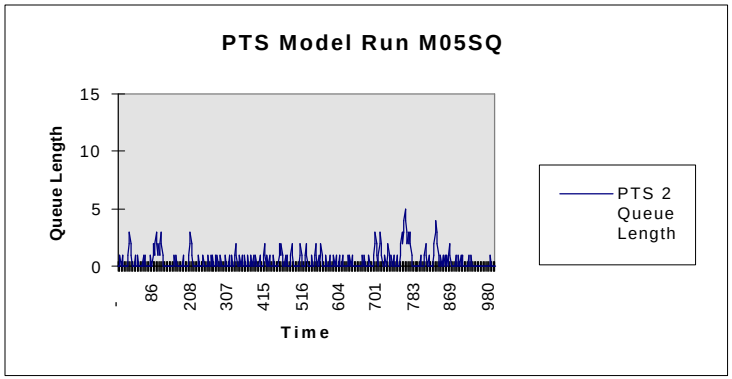
Copyright © 1996, MCI Telecommunications Corporation

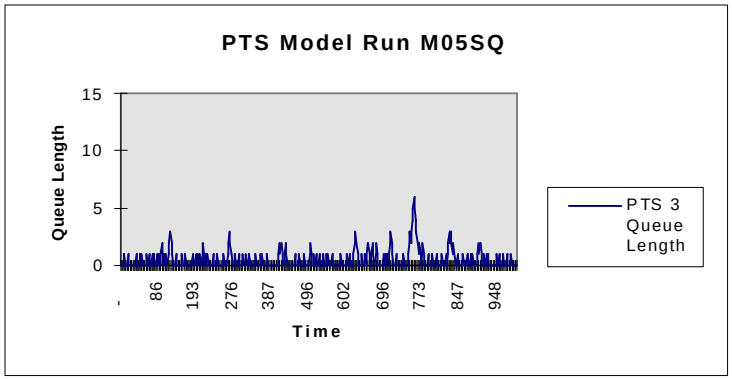
Shortest Queue 05 Model Run

- **Detailed Results:**

M05SQ	PTS 1	PTS 2	PTS 3
Minimum Queue Size	0.0	0.0	0.0
Average Queue Size	0.2	0.3	0.2
Maximum Queue Size	5.0	5.0	6.0
Minimum Queuing	0.0	0.0	0.0
Average Queue Time	0.4	0.8	0.7
Maximum Queuing	5.6	14.8	9.8
Number Completed	356.0	356.0	342.0
Waiting %	55.2	48.5	53.1
Working %	44.8	51.5	46.9
Aaverage Service Time	1.3	1.4	1.4



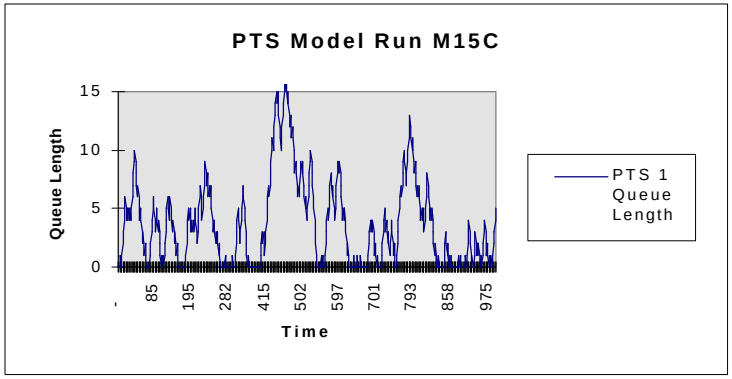


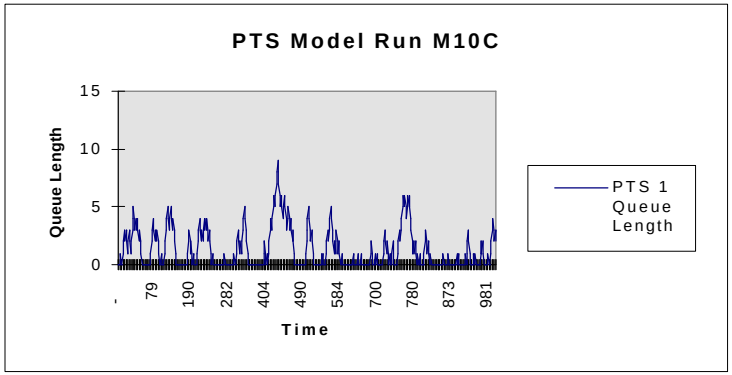


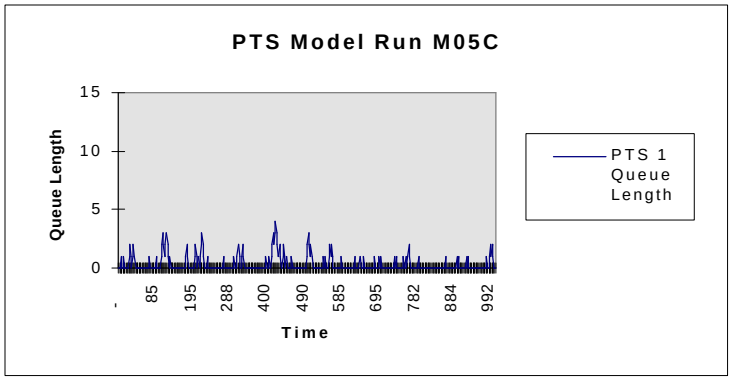
Copyright © 1996, MCI Telecommunications Corporation

Circulate Model Run

- **Average Response Times**
  - **M15C - 15.4**
  - **M10C - 5.3**
  - **M05C - 1.9**
- **Longest Queue**
  - **M15C - 31**
  - **M10C - 14**
  - **M05C - 6**
- **Worst Response Times**
- **Most Variation**
- **Some Smarts Better Than None**





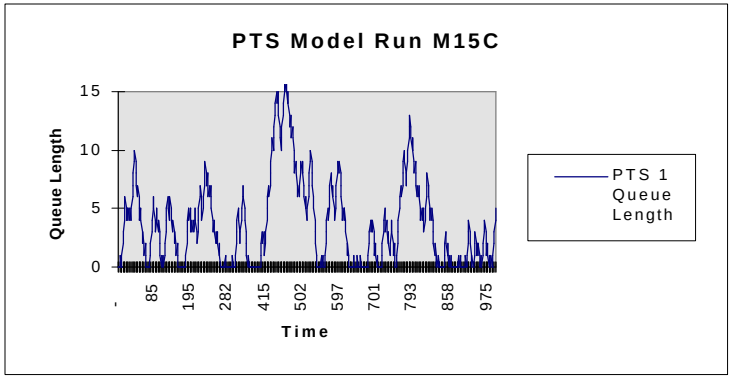


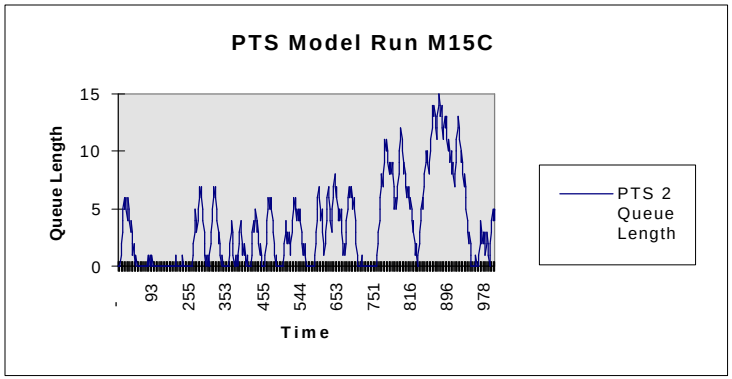
Copyright © 1996, MCI Telecommunications Corporation

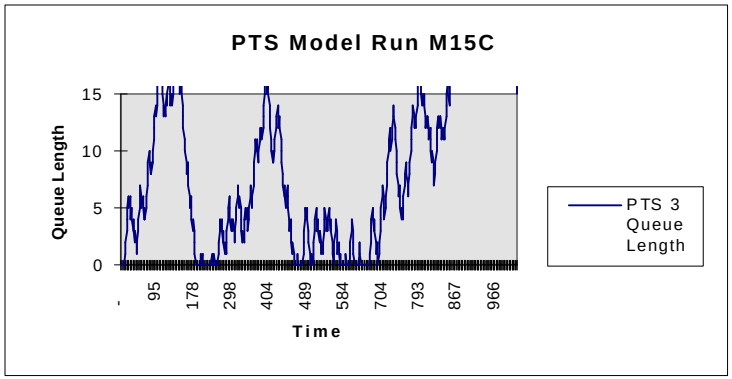
Circulate 15 Model Run

- **Detailed Results:**

M15C	PTS 1	PTS 2	PTS 3
Minimum Queue Size	0.0	0.0	0.0
Average Queue Size	3.0	2.5	8.3
Maximum Queue Size	16.0	15.0	31.0
Minimum Queuing	0.0	0.0	0.0
Average Queue Time	8.5	7.2	23.8
Maximum Queuing	43.7	41.3	84.2
Number Completed	346.0	347.0	335.0
Waiting %	21.9	29.9	13.7
Working %	78.1	70.1	86.3
Aaverage Service Time	2.3	2.0	2.6





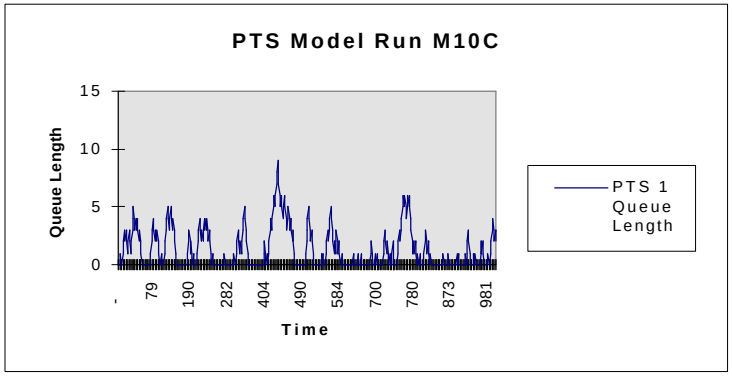


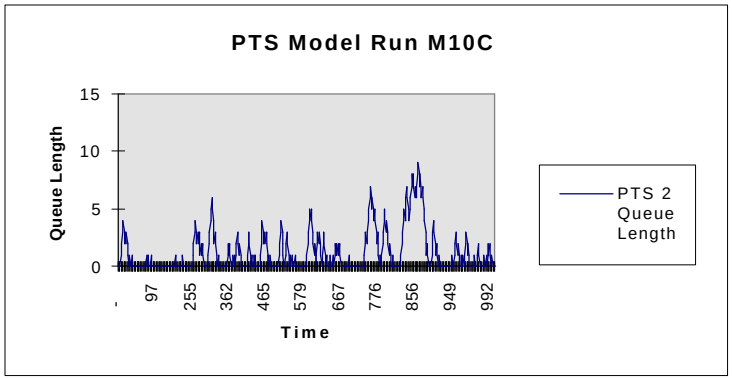
Copyright © 1996, MCI Telecommunications Corporation

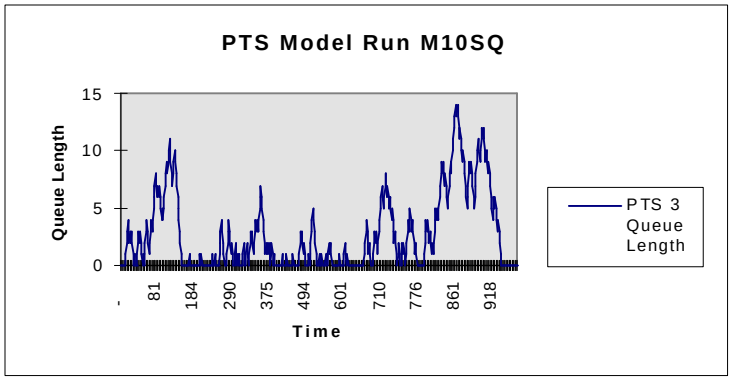
Circulate 10 Model Run

- **Detailed Results:**

M10C	PTS 1	PTS 2	PTS 3
Minimum Queue Size	0.0	0.0	0.0
Average Queue Size	0.9	0.9	1.9
Maximum Queue Size	9.0	9.0	14.0
Minimum Queuing	0.0	0.0	0.0
Average Queue Time	2.6	2.4	5.5
Maximum Queuing	23.0	23.6	34.5
Number Completed	349.0	352.0	351.0
Waiting %	37.4	41.6	31.6
Working %	62.6	58.5	68.4
Aaverage Service Time	1.8	1.7	1.9





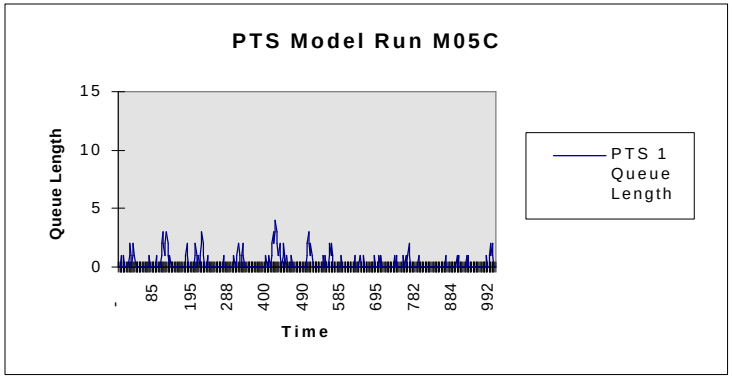


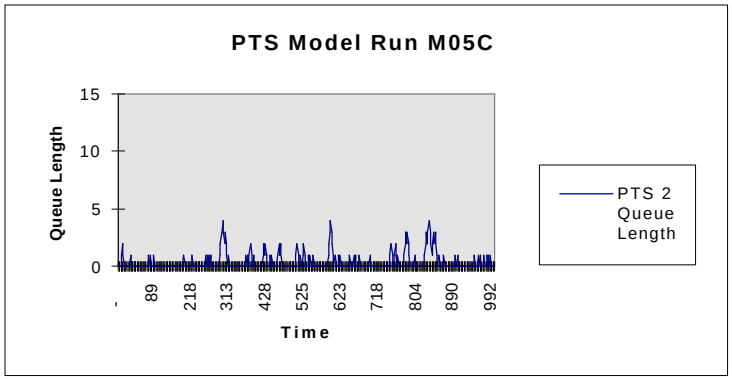
Copyright © 1996, MCI Telecommunications Corporation

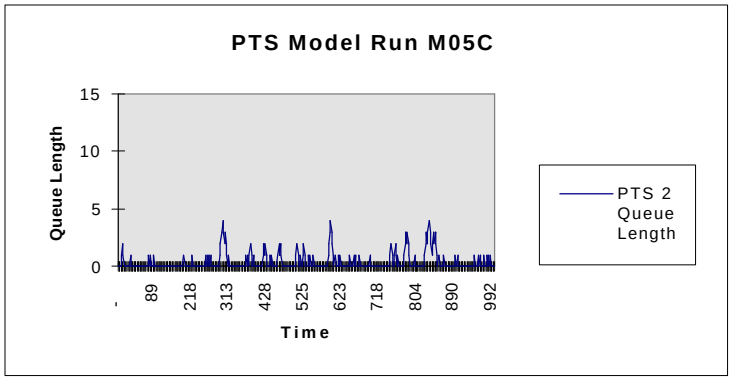
Circulate 05 Model Run

- **Detailed Results:**

M05C	PTS 1	PTS 2	PTS 3
Minimum Queue Size	0.0	0.0	0.0
Average Queue Size	0.2	0.2	0.2
Maximum Queue Size	4.0	4.0	6.0
Minimum Queuing	0.0	0.0	0.0
Average Queue Time	0.5	0.5	0.7
Maximum Queuing	8.3	8.6	13.1
Number Completed	352.0	352.0	351.0
Waiting %	53.2	53.6	51.1
Working %	46.8	46.5	48.9
Aaverage Service Time	1.3	1.3	1.4



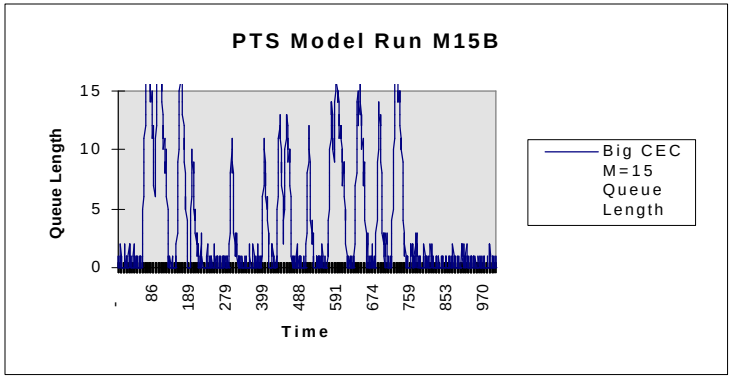


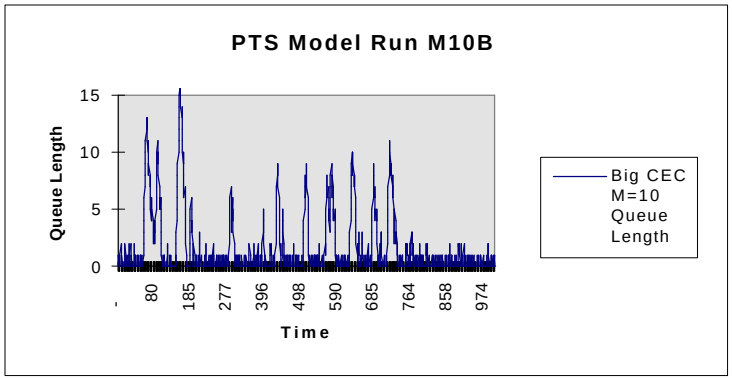


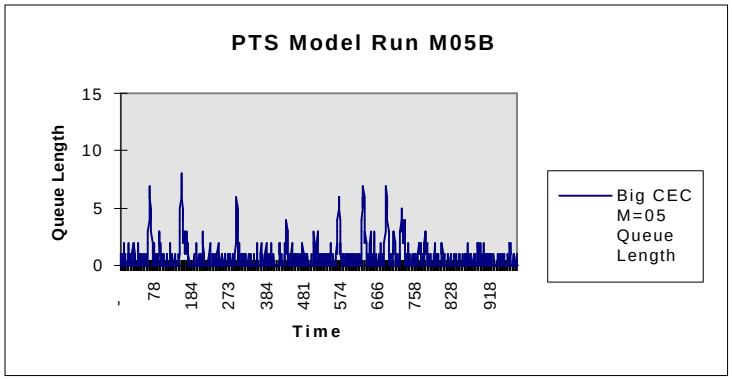
Copyright © 1996, MCI Telecommunications Corporation

Single CEC Model Run

- **Average Response Times**
  - **M15B - 2.9**
  - **M10B - 1.4**
  - **M05B - 0.6**
- **Longest Queue**
  - **M15B - 22**
  - **M10B - 16**
  - **M05B - 8**
- **Lots of Variation Not Reflected in Averages**
- **More Hits (Short Trans) Makes the Difference**





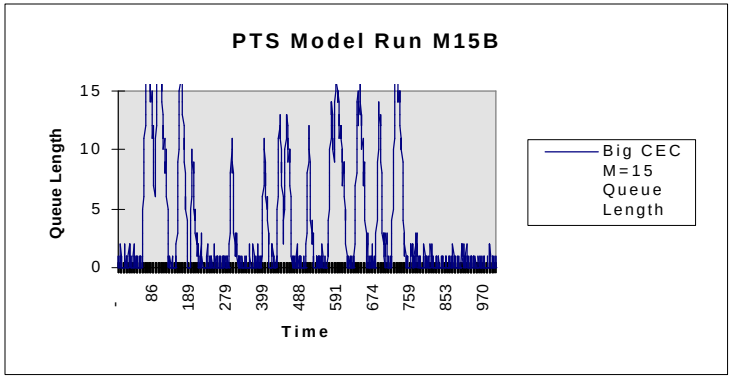


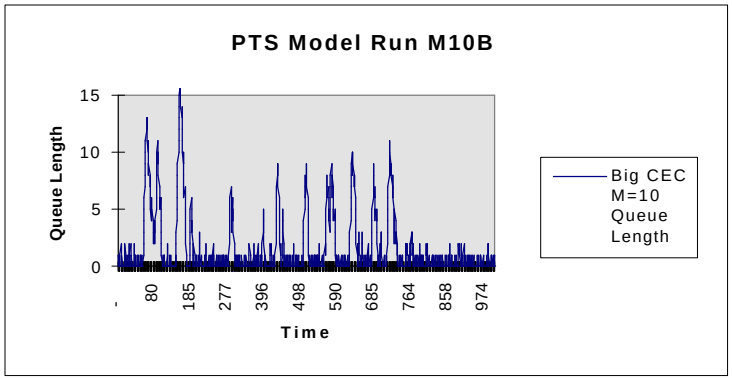
Copyright © 1996, MCI Telecommunications Corporation

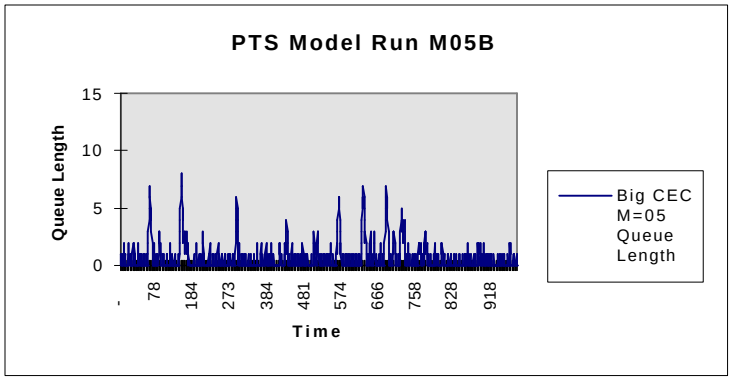
Single CEC Model Run

- **Detailed Results:**

Single CEC	M15B	M10B	M05B
Minimum Queue Size	0.0	0.0	0.0
Average Queue Size	2.3	1.0	0.3
Maximum Queue Size	22.0	16.0	8.0
Minimum Queuing	0.0	0.0	0.0
Average Queue Time	2.4	1.0	0.3
Maximum Queuing	21.0	12.8	7.8
Number Completed	974.0	974.0	974.0
Waiting %	52.0	59.0	66.0
Working %	48.0	41.0	34.0
Aaverage Service Time	0.5	0.4	0.3

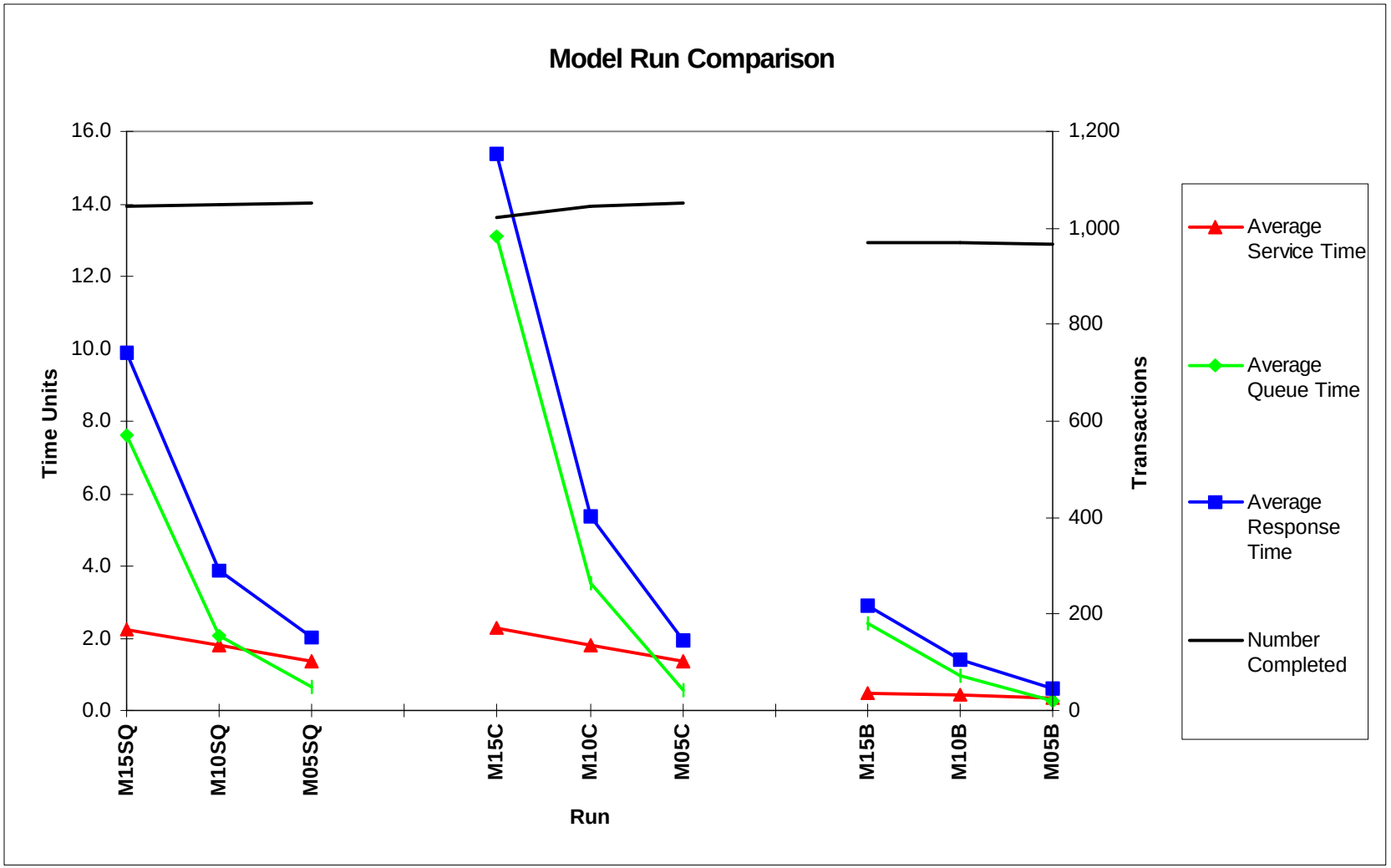






Copyright © 1996, MCI Telecommunications Corporation

Run Comparison



Copyright © 1996, MCI Telecommunications Corporation

Conclusions

- Simple Model Can Bracket the Problem
- Relative Differences vs Absolute Values
- Transaction Type More Important
- CPU Differences Less Important
  - But Difference Still Workload Dependent
- Routing Technique Will Make a Difference
  - Shortest Queue Better than Circulate
  - Will Content Routing Unbalance Queues?
    - Queues Continue to Grow
    - But Fewer Long Transactions

Copyright © 1996, MCI Telecommunications Corporation

Conclusions

- **More to Understand About the Workload**
  - Data Access Patterns
  - Hit Ratios
  - Locality of Reference
  - Arrival Patterns and Distributions
- **Key Questions:**
  - Is the Inter-arrival Time for Long Transactions Large Enough to Allow the Ones in the Queues to be Processed Before a New One Arrives?
    - Average vs True Distribution
  - Is the Number of Long Transactions to be Processed Concurrently Greater Than the Total Number of Servers?
    - If Yes - Timing Will Create Backups

Copyright © 1996, MCI Telecommunications Corporation

Conclusions

- Many Problems Moving Applications to Multiple Distributed Processors
- Industry Direction Increasing Problems
- Few Existing Tools
- Modeling Provides Information for Design and Implementation Decisions
  - How Much Difference Does a Choice Make?
- Assumptions Important
- Models Must Match Reality

Copyright © 1996, MCI Telecommunications Corporation

Next Steps For the Model

- **Increase Parallelism**
  - Add Servers to PTS's and CEC's to Represent the Maximum Number of CPU's or Tasks
- **Data Collection Techniques**
  - Application Instrumentation
  - Monitors?
  - Other?
- **Transaction Content Routing**
  - Data Collection an Issue
- **Buffer Access Emulation**
  - Data Collection an Issue
  - Buffer Management Technique (LRU or Other?)
  - Memory Sizing in the Servers (PTS's and CEC's)

Copyright © 1996, MCI Telecommunications Corporation

Questions

?

Copyright © 1996, MCI Telecommunications Corporation