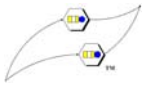
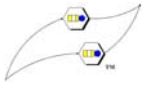


Modeling Virtualized Environments in Simalytic® Models by Computing Missing Service Demand Parameters

CMG2009 Paper 9103, December 11, 2009

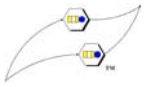
Dr. Tim R. Norton
Simalytic Solutions, LLC

719-635-5825
email: tim.norton@simalytic.com
http://www.simalytic.com



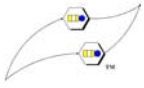
OR

Using Virtual Data to Model Virtual Systems



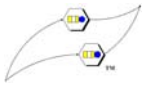
Agenda

- **Introduction**
  - Capacity Planning and Virtualization
- **The Problem**
  - Missing Measurements in Virtualization Guests
- **A Proposed Solution**
  - Computing Missing Service Demand Parameters
- **Conclusion**



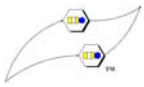
Introduction

- **Capacity Planning**
  - Capacity is Measured by Business Performance Objectives
    - Making decisions about resource requirements
  - What do we have to buy and when do we have to buy it to make sure that the business applications perform at the level required to insure the business succeeds?



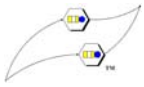
Introduction

- **Capacity Planning**
  - Two key aspects:
    - Demand for available resources
      - What do we have to buy
    - Effective completion of business work
      - When to buy it?
  - Requires some predictive technique
    - Usually some form of model
      - Simple: trend or other statistical techniques
      - Advanced: simulation or queuing network

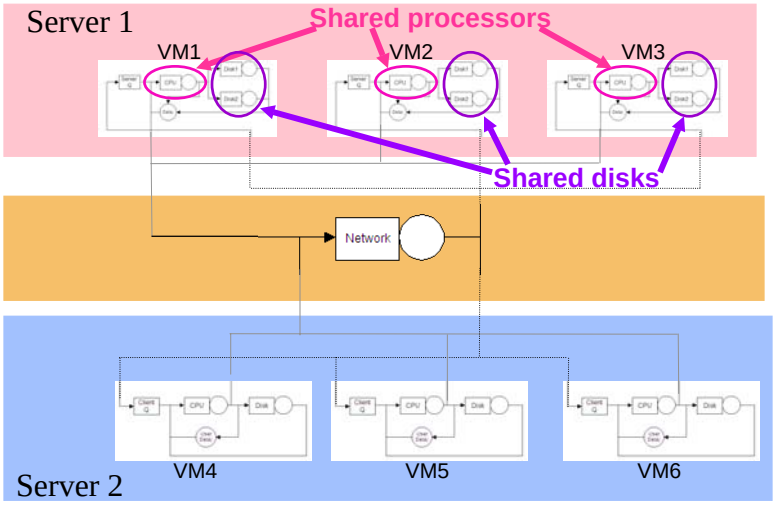


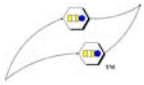
Introduction

- **Virtualization**
  - Increasing parallelization within the host system
    - Increase productive business-related work
    - Increase the usage of resources
  - Virtualized environment control program
    - Hypervisor
      - Usually implies a hardware implementation
    - VMM (Virtual Machine Monitor)
      - VMM often implies a software implementation
        - Will use VMM for all virtualization control programs



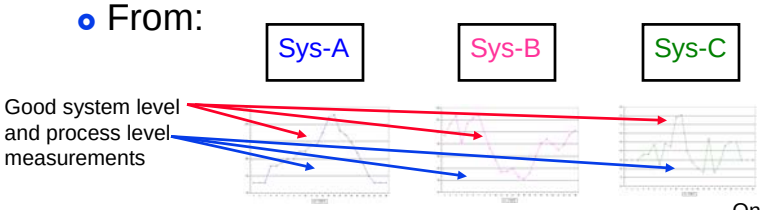
Virtualization Model



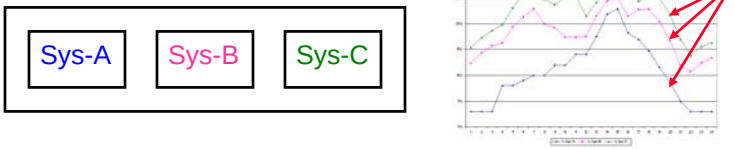


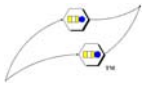
Introduction

- **Virtualization**
  - From:



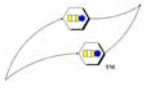
  - To:





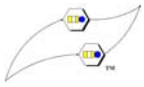
The Problem

- **Measurements in Virtualization Environments**
  - Guest Operating Systems
    - Most not Virtualization aware
      - Incorrect accounting for time when guest VM not active
        - Rate based measurements incorrect
        - Count based measurements valid
  - Hypervisor
    - Accurately measures Guest VM active time
      - Cannot measure processes within the Guest O/S
  - Affects measurements needed for models
    - VM System Level Service Times good
    - Process Level Service Times incorrect



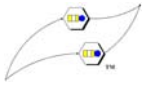
Capacity Models

- **What Virtualization does to Models**
  - Reduces accuracy – Measurement issues:
    - System clock
    - Accounting for dispatch time of other VMs
    - Interrupts – delays and re-driven to VM
    - VM vs. process priority
    - Delays – where are they accounted for
    - Virtualization overhead
      - VMM (Virtual Machine Monitor) and VM context switch
    - Interference from other VMs
  - Complicates Workload Characterization
    - Shared resources



Where's the Data

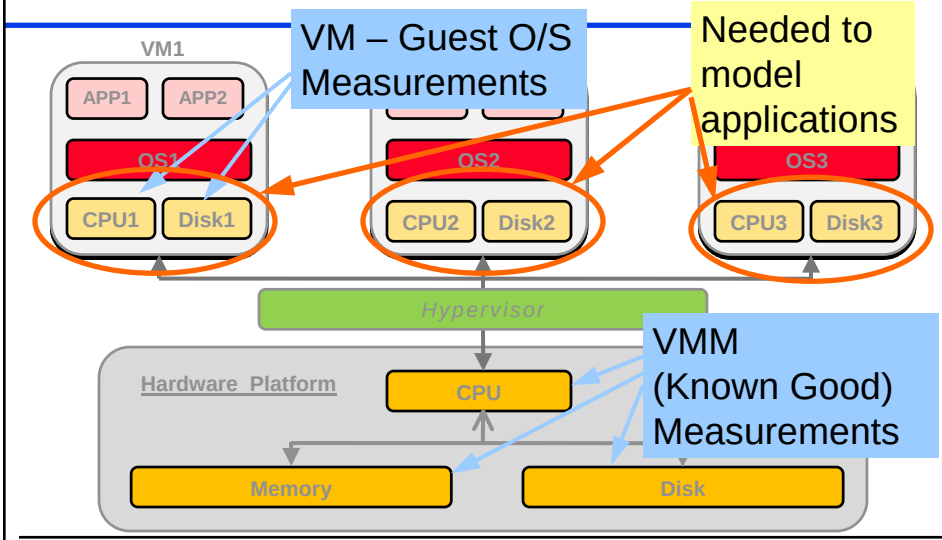
- **If It's Not There…**
  - It's missing!
- **If It's Wrong…**
  - It's missing!
- **If You Can't Trust It…**
  - It's missing!



A Proposed Solution

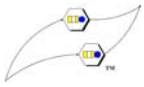
- **Use known good measurements**
  - VM utilization
    - from the VMM (Virtual Machine Monitor)
  - Transaction arrivals
    - from application measurements
  - Transaction response times
    - from application measurements
- **Use a Simalytic Model**
  - Builds relationships
  - Iterative design
  - Leverages existing tools

Virtualization Metrics



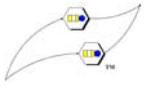
Model Requirements

- **Application Transaction Arrivals**
  - Count of transactions over many intervals
- **Application Response Times**
  - Always needed for validation of model results
    - Now needed for Solver calculations
- **Service Demand**
  - Menascé technique to compute class service demand from total service demand
    - VMM measurements used to compute VM level measurements that guest O/S cannot provide



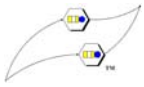
Implementation Approach

- **Collect Measurement Data**
  - Application – counts and response times
  - VMM resource usage
- **Compute Service Demand by Application**
  - Validate against measured response times
- **Build Model of Overall Virtualization Environment**
  - Using a Simalytic Model to express the relationships between applications



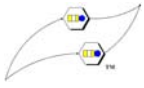
Implementation

- **Measurements**
  - Collect measurements
    - From the applications
      - Actual response times
      - Actual transaction counts
    - From the VMM
      - Total utilization for each resource (CPU, disks, etc.)
  - Multiple intervals
    - Measurements are needed for many intervals
      - Variety is important!
        – but avoid problem areas like low utilization effect
        - Low to high counts for each application
        - Low to high resource utilization



Implementation

- **Compute Service Demand**
  - Use collected measurements to construct multiple Open Multiclass QN formulae
    - Each uses the transaction response time and count for each class (application) along with the total service demand for each resource (CPU, disks, etc.) for one measurement interval.
  - Solve the non-linear constraint problem
    - *Computing Missing Service Demand Parameters for Performance Models.* Danny Menascé, CMG 2008.
    - Can be done with Microsoft Excel Solver

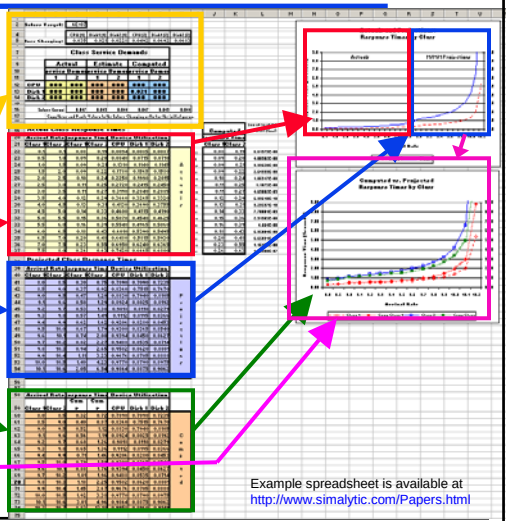


Implementation

- **Model Beyond What was Measured**
  - Measurement data provides historical view
    - Many intervals available
      - But at lower than expected future traffic volume
  - Model results provide future view
    - Answer the classic questions:
      - When does response time become unacceptable?
      - What resource saturates first?
  - The following example takes this approach
    - Measure system and applications at lower volumes
    - Predict behavior at higher volumes
      - Different than example in the paper

Excel Solver Spreadsheet

- **Solves Multiple Equations**
  - Initial Guess and Solver Results
  - Known Good Measurements
  - "Actual" measurements
  - Computed measurements
  - Validation



Example spreadsheet is available at http://www.simalytic.com/Papers.html

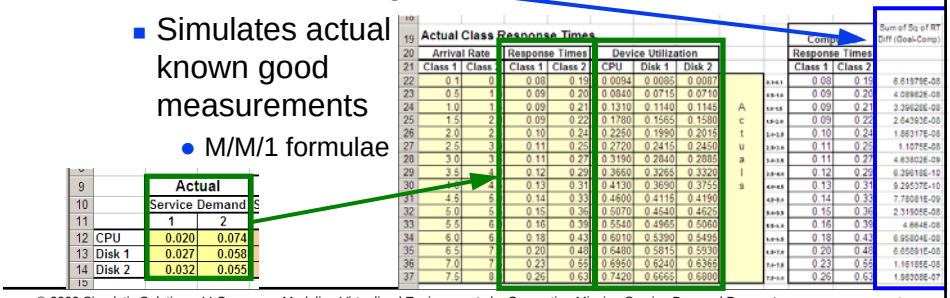
Using Excel Solver

- **Solver Results – Example Spreadsheets**
  - Start with initial guess
    - Same for both classes – calculated from VMM resource measurements divided sum of trans in both classes.
  - Computes values for both classes
    - non-linear constraints
  - Actual Service Demand Data
    - Not real measurements
      - Validate approach
      - Generate data for testing

	A	B	C	D	E	F	G	H
2	Solver Target:		6.3E-07					
4			CPU (1)	Disk1 (1)	Disk2 (1)	CPU (2)	Disk1 (2)	Disk2 (2)
5	Solver Changing:		0.035	0.021	0.0228	0.04918	0.06924	0.06832
7	Class Service Demands							
9		Actual		Estimate		Computed		
10		Service Demand		Service Demand		Service Demand		
11		1	2	1	2	1	2	
12	CPU	0.020	0.074	0.047	0.047	0.035	0.049	
13	Disk 1	0.027	0.058	0.0	0.043	0.021	0.069	
14	Disk 2	0.032	0.055	0.04	0.044	0.023	0.068	
16	Solver Guess:	0.047	0.043	0.044	0.047	0.043	0.044	
17	Copy these and Paste Values to the Solver Changing cells for the initial guess.							

Using Excel Solver

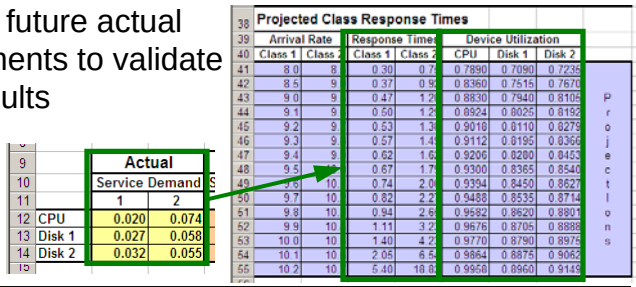
- **"Actual" Service Demand Values**
  - Used at low transaction volume to create known good response times and device utilizations
    - Used for Solver goals
    - Simulates actual known good measurements
      - M/M/1 formulae



	Actual Service Demand	
	1	2
CPU	0.020	0.074
Disk 1	0.027	0.058
Disk 2	0.032	0.055

Using Excel Solver

- **"Actual" Service Demand Values**
  - Used at high transaction volume to create "actual" (projected) response times and device utilizations
    - Predicts results at higher arrival rates
    - Simulates future actual measurements to validate Solver results



	Actual Service Demand	
	1	2
CPU	0.020	0.074
Disk 1	0.027	0.058
Disk 2	0.032	0.055

Using Excel Solver

- **Computed Service Demand Values**
  - Used to create "computed" response times and device utilizations
    - Solver results
    - Validated against "projected" values

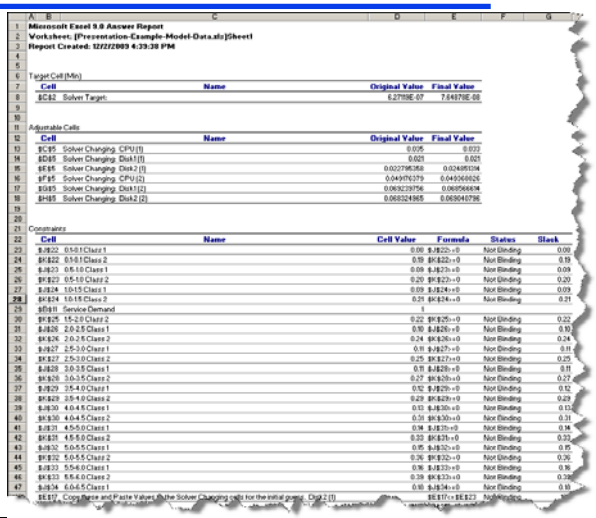
	Actual		Estimate		Computed	
	Service Demand		Service Demand		Service Demand	
	1	2	1	2	1	2
CPU	0.020	0.074	0.047	0.047	0.035	0.049
Disk 1	0.027	0.058	0.043	0.043	0.021	0.069
Disk 2	0.032	0.055	0.044	0.044	0.023	0.068

	Arrival Rate		Response Times		Device Utilization			
			Comp	Comp				
	Class 1	Class 2	Class 1	Class 2	CPU	Disk 1	Disk 2	
60	8.0	8.[illegible]	0.32	0.72	0.7890	0.7090	0.7235	
61	8.5	9.[illegible]	0.40	0.87	0.8360	0.7515	0.7670	
62	9.0	9.[illegible]	0.52	1.12	0.8830	0.7940	0.8105	
63	9.1	9.[illegible]	0.56	1.19	0.8924	0.8025	0.8192	
64	9.2	9.[illegible]	0.60	1.26	0.9018	0.8110	0.8279	
65	9.3	9.[illegible]	0.65	1.36	0.9112	0.8195	0.8366	
66	9.4	9.[illegible]	0.71	1.46	0.9206	0.8280	0.8453	
67	9.5	9.[illegible]	0.79	1.59	0.9300	0.8365	0.8540	
68	9.6	10.[illegible]	0.88	1.76	0.9394	0.8450	0.8627	
69	9.7	10.[illegible]	1.01	1.96	0.9488	0.8535	0.8714	
70	9.8	10.[illegible]	1.18	2.25	0.9582	0.8620	0.8801	
71	9.9	10.[illegible]	1.45	2.67	0.9676	0.8705	0.8888	
72	10.0	10.[illegible]	1.92	3.38	0.9770	0.8790	0.8975	
73	10.1	10.[illegible]	3.01	4.96	0.9864	0.8875	0.9062	
74	10.2	10.[illegible]	8.83	13.18	0.9958	0.8960	0.9149	

Computed

Using Excel Solver

- **Solver Answer Report**
  - Shows details behind solution


Microsoft Excel 9.0 Answer Report
Worksheet: [Presentation-Example-Model-Data.xls]Sheet1
Report Created: 12/27/2009 4:39:38 PM

Target Cell (Min)

Cell	Name	Original Value	Final Value
C2	Solver Target:	6.2709E-07	7.1497E-08

Adjustable Cells

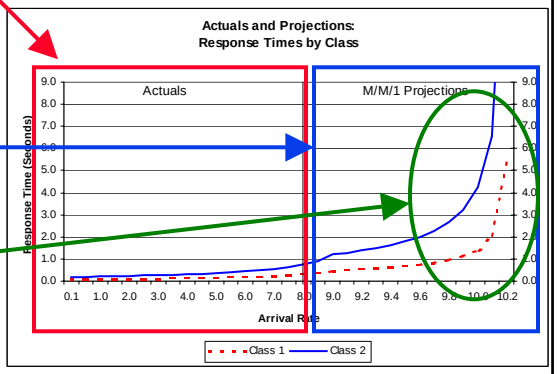
Cell	Name	Original Value	Final Value
C5	Solver Changing: CPU (1)	0.035	0.022
D5	Solver Changing: Disk1 (1)	0.021	0.021
E5	Solver Changing: Disk2 (1)	0.022795358	0.024051014
F5	Solver Changing: CPU (2)	0.049376379	0.049366826
G5	Solver Changing: Disk1 (2)	0.069239756	0.069566674
H5	Solver Changing: Disk2 (2)	0.068324965	0.068040796

Constraints

Cell	Name	Cell Value	Formula	Status	Slack
J22	0.5-0.1 Class 1	0.00	J22>=0	Not Binding	0.00
K22	0.5-0.1 Class 2	0.19	K22>=0	Not Binding	0.19
J23	0.6-1.0 Class 1	0.09	J23>=0	Not Binding	0.09
K23	0.6-1.0 Class 2	0.20	K23>=0	Not Binding	0.20
J24	1.0-1.5 Class 1	0.09	J24>=0	Not Binding	0.09
K24	1.0-1.5 Class 2	0.21	K24>=0	Not Binding	0.21
D31	Service Demand	1			
K25	1.5-2.0 Class 2	0.22	K25>=0	Not Binding	0.22
J26	2.0-2.5 Class 1	0.10	J26>=0	Not Binding	0.10
K26	2.0-2.5 Class 2	0.24	K26>=0	Not Binding	0.24
J27	2.5-3.0 Class 1	0.11	J27>=0	Not Binding	0.11
K27	2.5-3.0 Class 2	0.25	K27>=0	Not Binding	0.25
J28	3.0-3.5 Class 1	0.11	J28>=0	Not Binding	0.11
K28	3.0-3.5 Class 2	0.27	K28>=0	Not Binding	0.27
J29	3.5-4.0 Class 1	0.12	J29>=0	Not Binding	0.12
K29	3.5-4.0 Class 2	0.29	K29>=0	Not Binding	0.29
J30	4.0-4.5 Class 1	0.13	J30>=0	Not Binding	0.13
K30	4.0-4.5 Class 2	0.31	K30>=0	Not Binding	0.31
J31	4.5-5.0 Class 1	0.14	J31>=0	Not Binding	0.14
K31	4.5-5.0 Class 2	0.33	K31>=0	Not Binding	0.33
J32	5.0-5.5 Class 1	0.15	J32>=0	Not Binding	0.15
K32	5.0-5.5 Class 2	0.36	K32>=0	Not Binding	0.36
J33	5.5-6.0 Class 1	0.16	J33>=0	Not Binding	0.16
K33	5.5-6.0 Class 2	0.39	K33>=0	Not Binding	0.39
J34	6.0-6.5 Class 1	0.18	J34>=0	Not Binding	0.18
E17	Copy these and Paste Values (to the Solver Changing cells for the initial guess): Disk 2 (1)	E17<=E23		Not Binding	

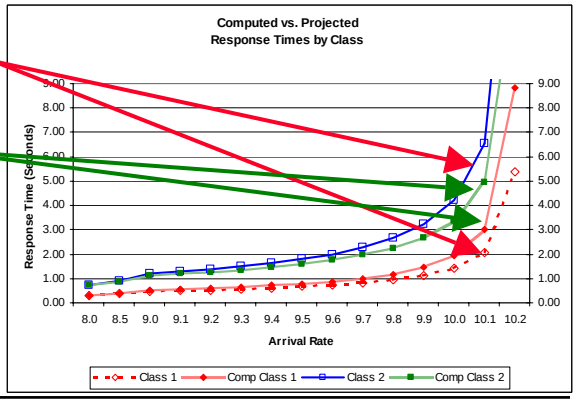
Application Performance

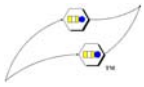
- **Two groups of measurements**
  - "Actuals" are data used to create the models
  - "Projections" are data used as post-model measurements
  - Each class (application) shows response time increase as utilizations go up



Application Prediction

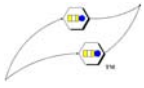
- **How Well do Computed Results Match**
  - Actual Measurements
  - Results using Computed Service Demands
  - Some difference but same trend





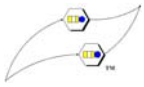
Usage

- **How Can Computed Service Demands be Used?**
  - Stand-alone models
    - Same as measured service demands
      - Adjustments may be needed for VMM overhead and other interference – similar to other model calibrations
  - Simalytic Models
    - Enhanced Simalytic Function for multi-tier models
      - Dynamic calculations to
        - simulate complex usage patterns
        - account for effects of spikes in other workloads



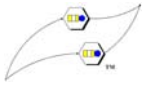
Conclusion

- **Future Work**
  - Explore criteria around measurement collection
    - How many intervals needed for minimal effectiveness
    - What improves accuracy
  - Incorporate solver into modeling tool
    - Possibly as enhanced Simalytic Function



Conclusion

- **Valid Approach**
  - Works with synthetic test data
    - Some differences
      - between projected and computed results
    - But trends usable for planning
  - Refinement needed
    - account for VMM overhead
  - Model virtual systems with virtual data



Missing Data Doesn't Stop a Real Modeler!

Presentation and spreadsheet will be available at: http://www.simalytic.com/Papers.html