Optimizing Performance in Networked Systems

Julien Granger (UW-Madison)
Ananth Krishnamurthy (RPI)
Stephen M. Robinson (UW-Madison)

Work Sponsored by ARO and AFOSR

U.S. Army Conference on Applied Statistics
Napa, CA, 30 October 2003

Outline

- Problem of improving performance in a network with uncertainty
- Difficulty with repeated simulation
- Using approximation methods to avoid repeated simulation
- Example: Analytical Airfield Model (AAM)
  - Small numerical example that nonetheless gives significant managerial insight
- Conclusion

The Problem

- Given a network with uncertainty, decide how to invest resources to improve its performance
  - These often arise in logistics, but occur also in many other areas
- Performance measure: often throughput
- Investment: change system parameters to
  - Decrease times for critical operations such as transshipment, refueling, maintenance
  - Increase capacity of intermediate stations that perform these operations

Measuring Performance

- We usually have to measure performance with simulation models
  - Significant uncertainties in these systems
  - Deterministic models are easy to use but they ignore the uncertainties and may give seriously wrong answers
- Simulation models give "snapshots" of performance of a given system, but are slow
  - More so, if system is large and complex
  - And we need to simulate *every time we change parameter values,* to check improvement

Problem: Slowness

- These methods simulate the system at each step
  – On complex systems they can be very slow
- How can we get around this?
  – One way: speed up the simulations
    » Network of workstations (Condor Project)
  – Another: maybe avoid simulating each time
    » Use approximation methods instead

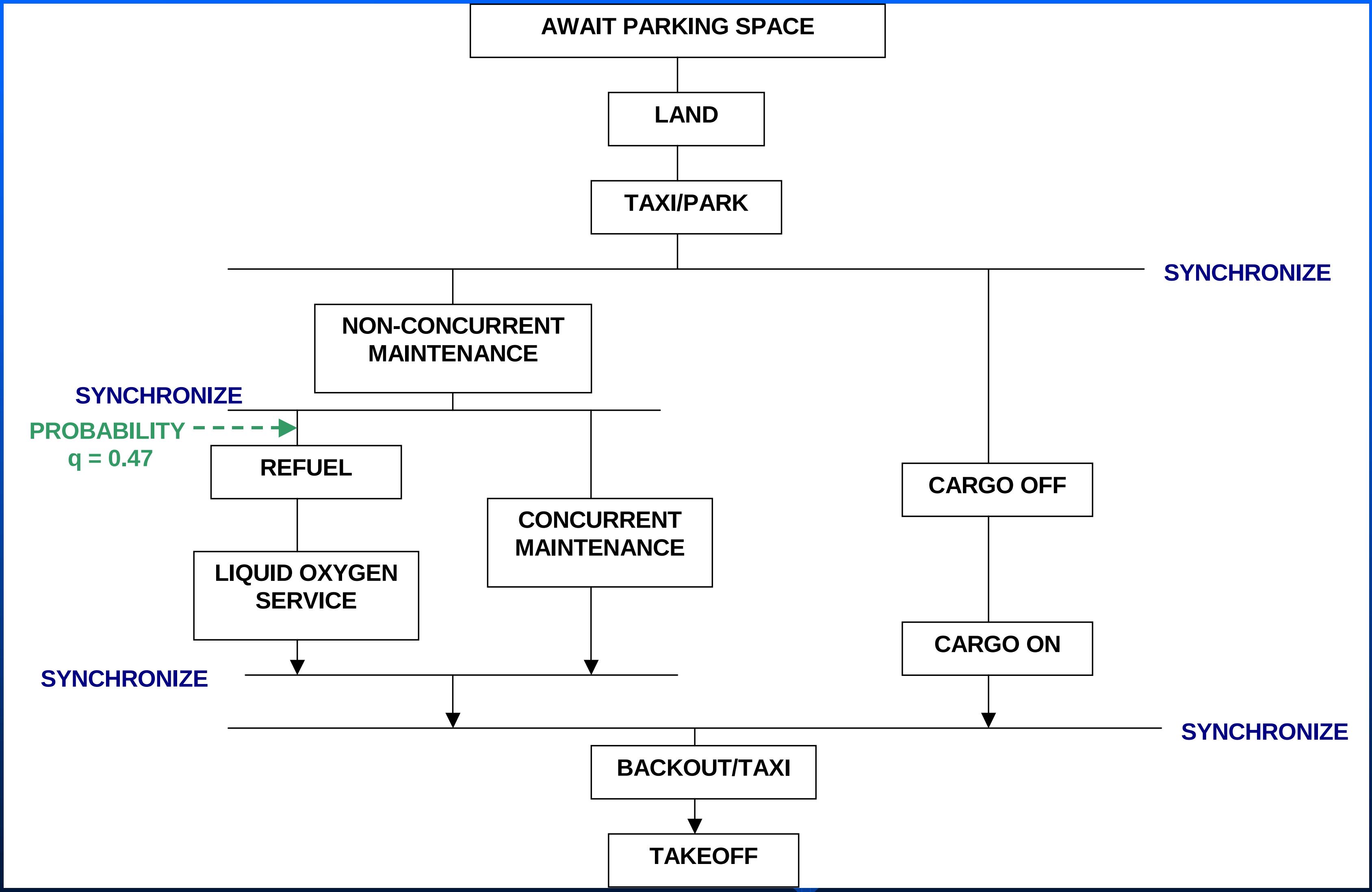
Network Approximations

- Idea: don't simulate the system exactly
  - Approximation requires *only solving a system of nonlinear equations*
    - » Much faster than simulating
  - On various problems we've tried, ratio of simulation time to approximation time has been in range 16:1 to 80:1 or more
- So: use approximations to improve the system, then simulate once to confirm the results

Example

- Analytical airfield model (AAM)
  - Ref: Dietz, D. C., *J. Aircraft* (1999) *et seq*.
- Relatively complex network of immediate military interest
  - We'll illustrate how to use optimization to answer managerial questions
  - What's new here is better control methods through use of new 2-moment approximations
  - Tools apply just as well to many other networks

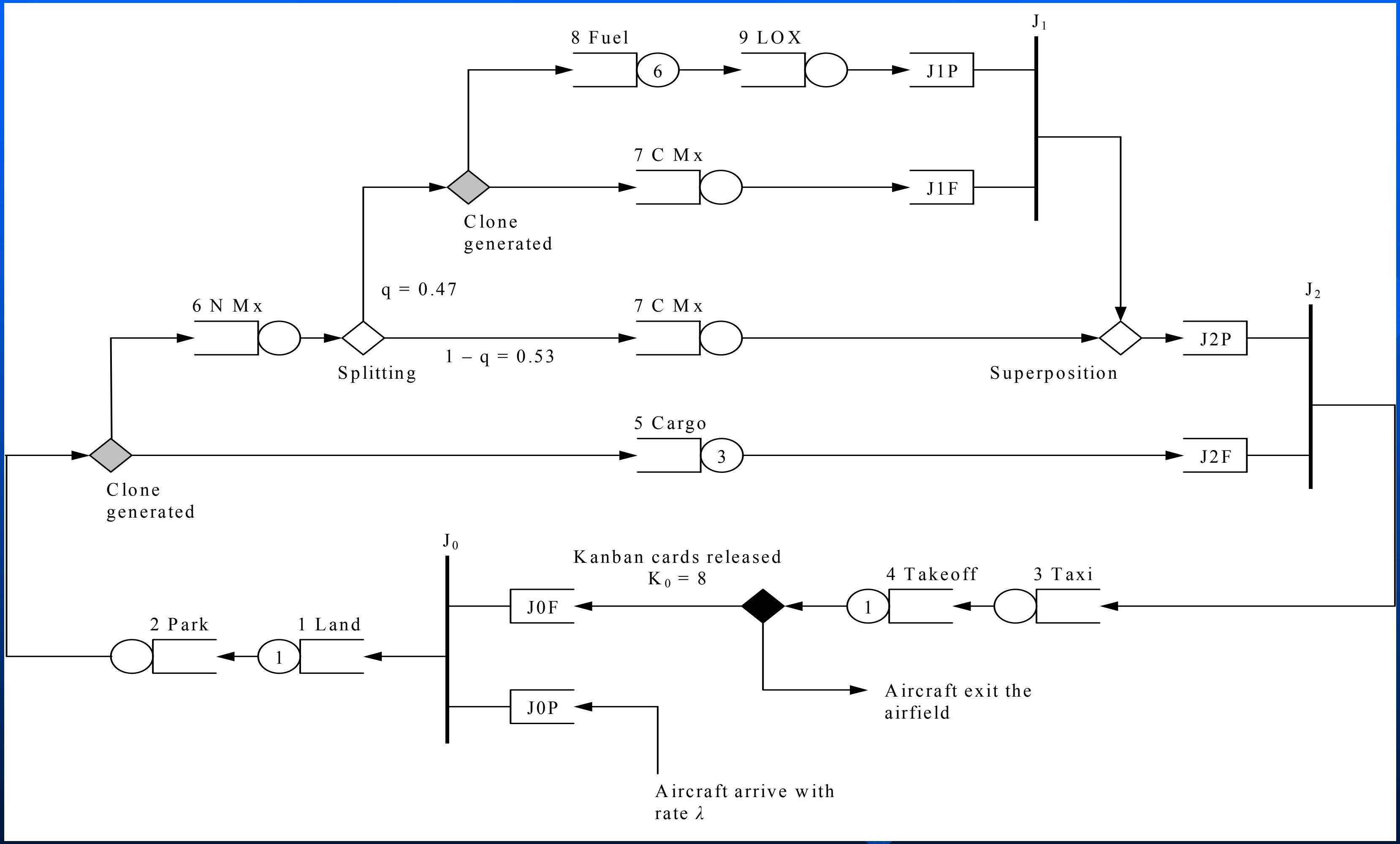
Task Precedence Graph for the AAM Problem



Modeling Essential Features

- Concurrent use of resources can be modeled as a fork/join mechanism:
  - Aircraft generate temporary clones with identical attributes
  - Overall waiting time is the maximum over all clone paths
- Aircraft waiting authorization for landing obey a join mechanism governed by Kanban cards
  - Number of Kanban cards equals maximum number of aircraft allowed on ground simultaneously

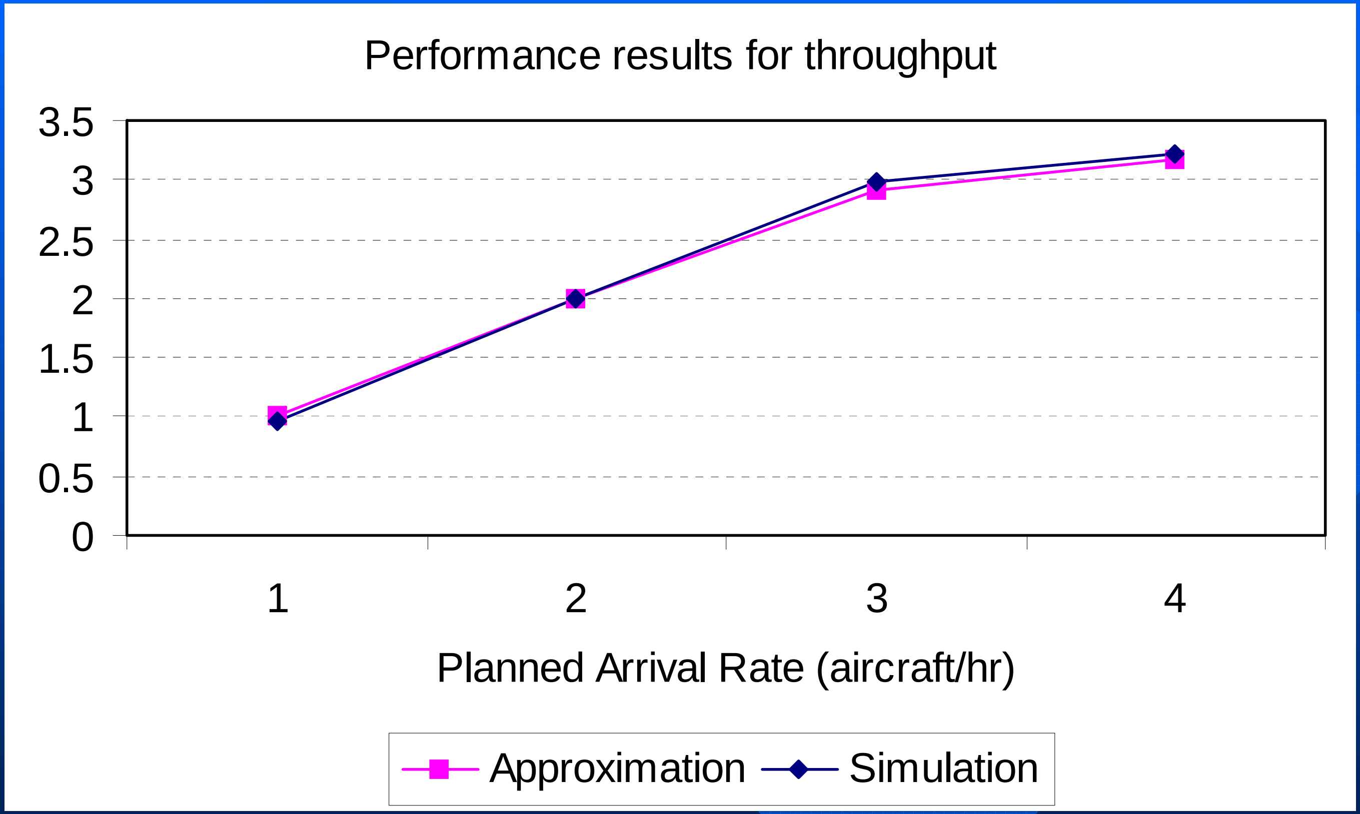
Layout of Queuing Model



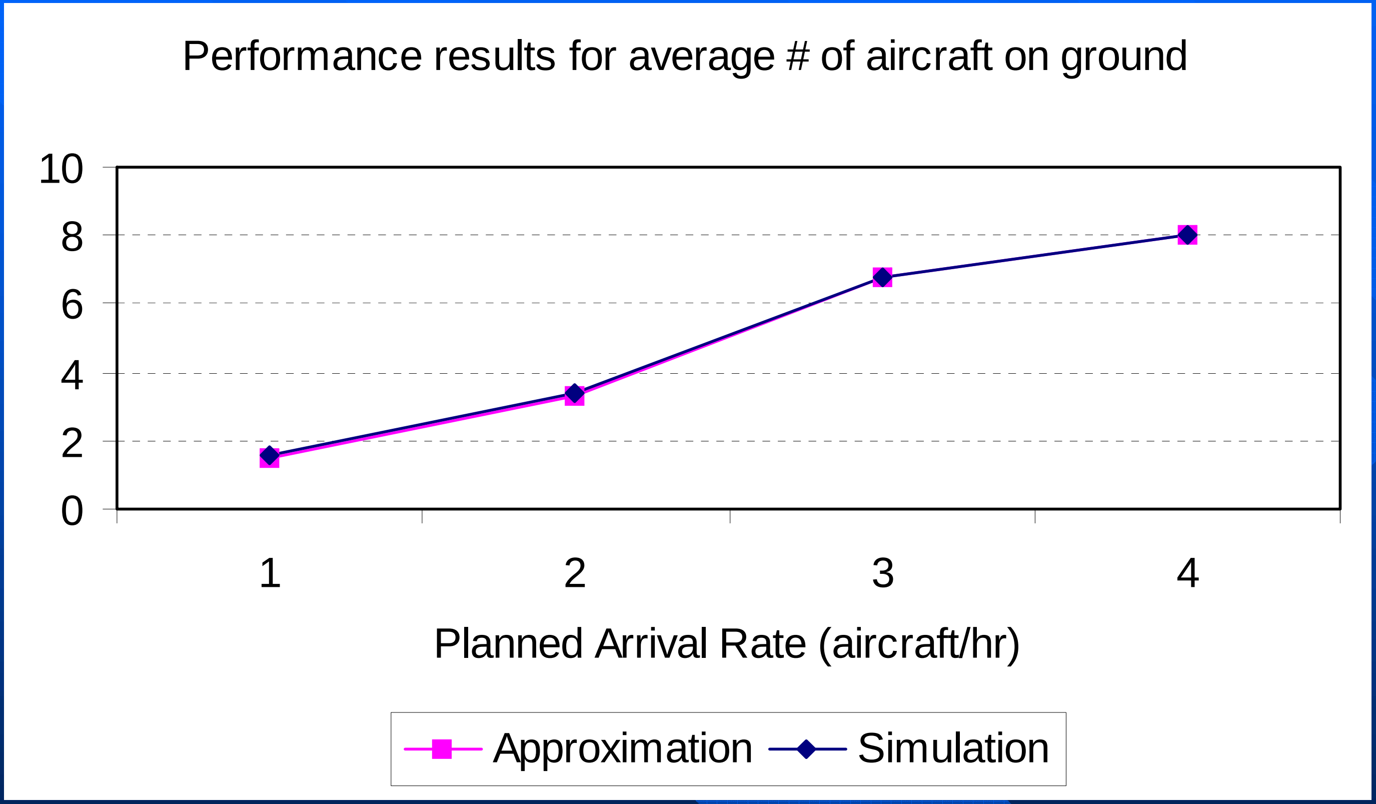
Approximating the AAM

- Capacity planners need fast answers to changes in the system
  - Performance measures are usually estimated with simulation, which is accurate but costly
- Instead, we do most of the analysis calculations with 2-moment approximations
  - Very fast: solution in < 5 seconds on a PC
  - Relatively accurate (relative error < 10 %)
- Kanban modeling better captures real constraint for landing authorization than current modeling

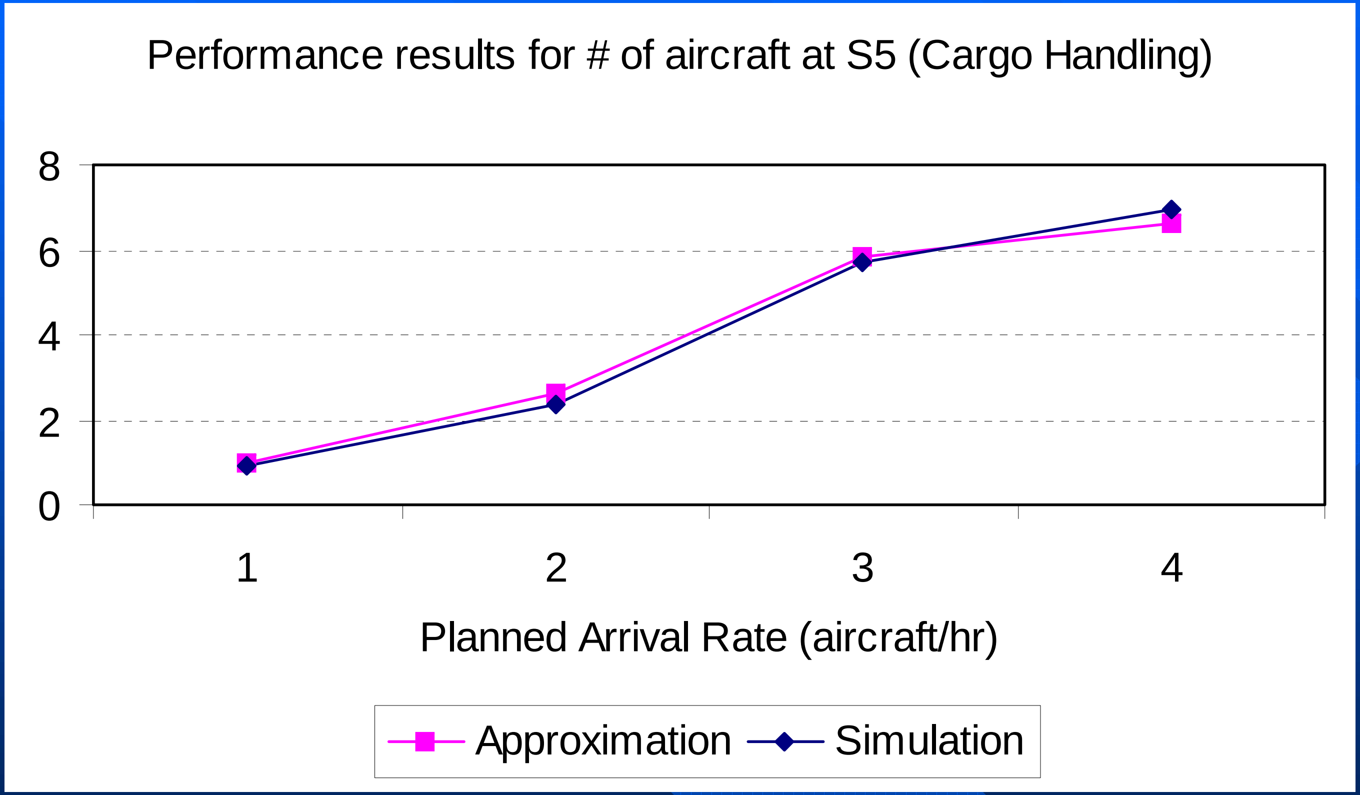
Example: Throughput



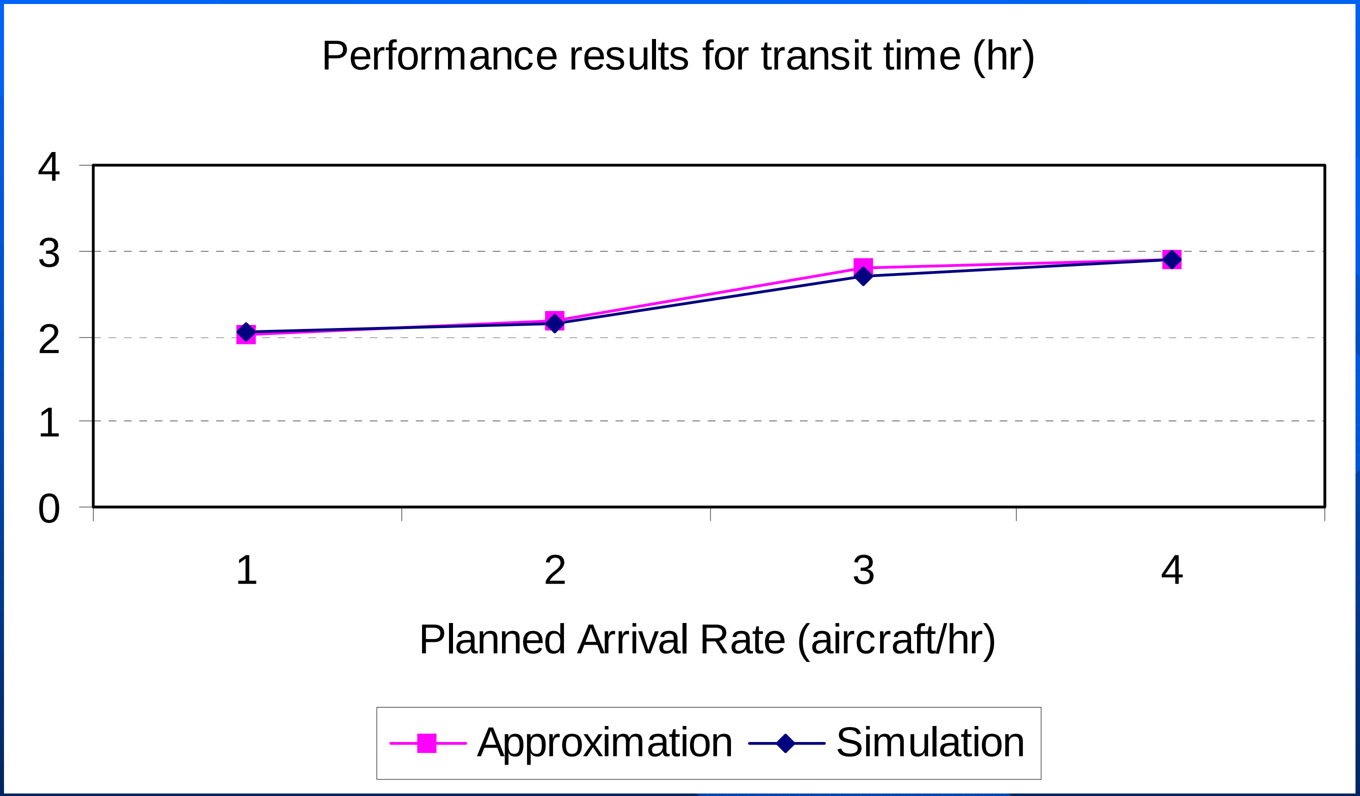
Example: Average on Ground



Cargo Handling is a Bottleneck



Transit Time Up By > 40 %



Network Improvement

- Initial situation:
  - λ = 3.0 aircraft/hr
  - K_0 = 6 (# allowed on ground)
  - Transit time (Approximation) = 2.482 hr.
- Constraint: transit time should be decreased by 20 %.
- Question: What is the minimum # of capacity units to invest at S5 (Cargo Handling) to have transit time decreased by 20 %?

Network Improvement

- Answer: 2, so that number of servers at Cargo Handling increases from 3 to 5.

Transit time (hr.)

Number of servers at Cargo Handling	Approximation	Simulation
3	2.48	1.84
5	1.89	1.49

Network Improvement

- Additional insight: Number of aircraft at Cargo Handling decreases by 31 %.

Number of aircraft at Cargo Handling

Number of servers at Cargo Handling	Approximation	Simulation
3	4.67	4.25
5	3.22	2.92

-31%

Conclusion

- Fast numerical procedure that gives *quick,* reasonably accurate results
  - Can validate results by using just one simulation run instead of many
- New technical results allow modeling of more complex systems
  - Better modeling of join stations
- Results
  - We can ask and answer managerial questions quickly and easily
  - Improve operations by investing optimally

(Backup slides follow)

Main Assumptions for 2-Moment Approximations

- Traffic processes are approximated as renewal processes
- Traffic process rate and SCV (squared coefficient of variation) are enough to completely characterize the performance measures of the network
- Example: for a GI/G/1 queue

$$W_q = \left(\frac{c_a^2 + c_s^2}{2}\right)\left(\frac{\lambda\tau^2}{1-\lambda\tau}\right)$$

W_q : delay in queue

c_a^2 : SCV of arrival process

c_s^2 : SCV of service process

λ : arrival rate

τ : mean service time

New Result for a Join Station

- Developed in (Krishnamurthy, 2002)
- Rate and SCV of departure process of a Join station as a function of the 6-tuple:

$$\left(\lambda_1, c_1^2, K_1, \lambda_2, c_2^2, K_2\right)$$

- Example:

If $\rho \equiv \dfrac{\lambda_1}{\lambda_2} \neq 1:$

$$\lambda_D = \lambda_1 \left[\frac{1-\rho^{K_1+K_2}}{1-\rho^{K_1+K_2+1}}\right]\left[1-0.5\left(\frac{c_1^2+c_2^2}{2}\right)\left(\frac{1-\rho}{1-\rho^{2(K_1+K_2)+1}}\right)\rho^{2(K_1+K_2)}\right]$$

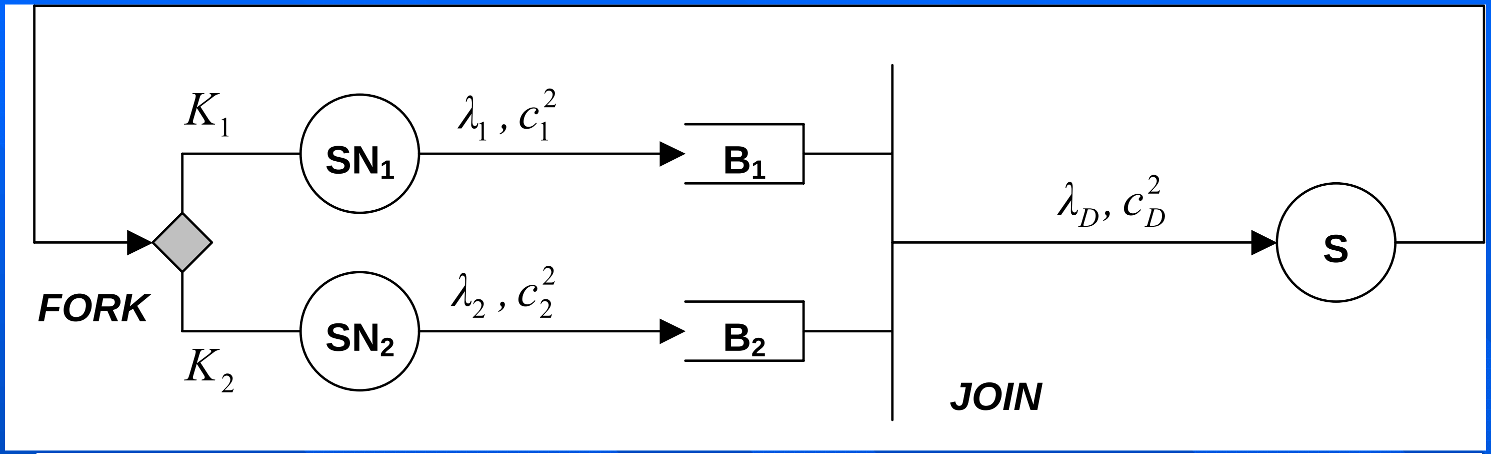
References

- New results for a Join station:
  - A. Krishnamurthy, Analytical performance models for material control strategies for manufacturing systems, Ph.D. Dissertation, University of Wisconsin-Madison, 2002
- Known results for other types of queues and overall performance evaluation of the network:
  - R. Suri, J.L. Sanders and M. Kamath, Performance evaluation of production networks, in: S.C. Graves et al. (Eds.), *Logistics of Production and Inventory* (Elsevier, Amsterdam, 1993)

Reference for the AAM Problem

- Dietz, D. C., Mean value analysis of military airfield operations at an individual airfield, *Journal of Aircraft* **36,** No. 5, 1999.
- Start with:
  - Total fleet size equals 20.
  - Maximum number of aircraft allowed on ground $K_0 = 8$.
- Vary planned arrival rate from 1.0 to 4.0 aircraft/hr.
- Compute performance measures of interest:
  - Throughput.
  - Transit time on airfield.
  - Average number of aircraft on ground.
  - Queue length at bottlenecks.

Fork-join Dynamics



SN_1 : Sub-network where entity 1 travels independently

SN_2 : Sub-network where entity 2 travels independently

B_1 : Buffer where entity 1 waits for an arrival of entity 2

B_2 : Buffer where entity 2 waits for an arrival of entity 1

S : Sub-network where joined entity travels

λ_1, c_1^2, K_1 : Rate, SCV and shut-down level of arrival process of entity 1

λ_2, c_2^2, K_2 : Rate, SCV and shut-down level of arrival process of entity 2

λ_D, c_D^2 : Rate, SCV of departure process of joined entity