Demand Pricing & Resource Allocation in Market-based Compute Grids: A Model and Initial Results

Vladimir Marbukh and Kevin Mills
National Institute of Standards and Technology
100 Bureau Drive, Stop 8920
Gaithersburg, MD 20899-8920
E-mail: marbukh@nist.gov

Abstract- Market-based compute grids encompass service providers offering limited resources to potential users with varying quality of service demands and willingness to pay. Providers face problems of pricing and allocating resources to maximize revenue. Previous research proposed and analyzed a method for allocating resources based on joint optimization of access control and scheduling strategies. This paper proposes a tractable analytical model for joint optimization of job pricing and scheduling strategies with the objective of maximizing provider revenue. The paper provides initial results for the case of a single provider serving price-sensitive users whose utilities decay linearly with increasing service delay. The paper also shows that providers must combine both pricing and admission control to achieve maximum revenue.

I INTRODUCTION

Emerging Grid technologies pose a challenging problem of efficient resource allocation in complex, decentralized systems with strategically behaving users and service providers. In market-based compute grids efficient resource allocation can be achieved through a combination of demand pricing and resource management. While demand pricing matches average (long-term) demand with available resources, resource management ensures that the system can accommodate instantaneous (short-term) demand fluctuations. This paper, a continuation of previous work [1], considers market-based compute grids where providers attempt to maximize their profit through quality-of-service-dependent pricing and resource allocation.

Researchers who investigate market-based compute grids typically model users willingness to pay (utility) as a reward for completing a job by a deadline and a decay rate, which defines the slope of a linearly decreasing function of the reward over time for late jobs [2]-[7]. As reward decays beyond zero, user utility becomes negative and a provider must pay a corresponding penalty. Many researchers devise heuristics for pricing, admission control and scheduling of jobs by service providers and then use simulation to evaluate performance of those heuristics when subjected to a mix of job classes. Each job class is defined by a deadline and associated reward, along with a decay rate for exceeding the deadline (and possibly a bound on the penalty for late jobs).

Providers may increase revenue by controlling demand through pricing rather than through admission control. Game theory provides a natural analytical framework for modeling a market of providers and users. However, the presence of multiple providers, who compete for users on price and quality of service, greatly complicates analysis of the corresponding game. This paper proposes a tractable analytical model for joint optimization of job pricing and scheduling aimed at maximizing revenue given a single provider. We solve this model under the assumption that potential provider penalties are unbounded, and we analyze key model parameters. In this particular case the scheduling optimization problem can be solved explicitly, yielding priority scheduling with priorities determined by job urgency.

We assume that a user submits a service (job) request if the corresponding net utility, which is the user utility minus the price of the service, is positive. If net utility is negative, the user does not submit this job request. A user obtains maximum utility if a job completes by a specified deadline, and a discounted utility for a late job. Late jobs may require the provider to pay a penalty; thus, revenue for a job could be negative. Given limited resources, a provider maximizes its revenue through pricing and allocating resources. We assume the provider knows the price-demand curve, which is stable on the scale of individual job arrivals and departures. This assumption allows us to use steady-state formulas for queuing delays.

Our analytical solution may be applied to determine maximum achievable provider revenue for a given mix of jobs, characterized by delay sensitivity, demand potential and price elasticity, assuming unbounded provider penalties. Our model can also be used to understand the operating limits of heuristics for pricing, admission control and scheduling, and to investigate the implication of varying job mixes. The major conclusion we derive from analysis of our model is the necessity to combine both pricing and admission control for adequate resource allocation. While pricing matches average demand with available resources on the “slow” timescale, admission control reacts to demand variations on the “fast” timescale.

The paper is organized as follows. Section II introduces the user model, which assumes that each user attempts to maximize his net delay-sensitive utility. Section III describes the provider model, which assumes that a provider attempts to maximize earned revenue, where maximization is performed over job pricing and scheduling. Section IV shows, for a particular case of linear user utilities, that the provider revenue maximization problem can be decomposed into finding optimal scheduling and optimal pricing, where the optimal scheduling, which turns out to be priority scheduling, can be determined explicitly. Section V provides some numerical results derived from the model. Section VI compares the effects of pricing and admission control on provider revenue. Finally, Section VIII summarizes results and outlines directions for future research.

II USER MODEL

We model users as job submitters, where each job includes an expressed willingness to pay that consists of two parts: base value and delay-dependent decay, which can be seen as diminished value for jobs completed late. Thus, we model jobs as being delay sensitive.

We assume that there are S classes of jobs, where all jobs of each class $s = 1,..,S$ have the same pattern of delay sensitivity. Delay sensitivity of a job of class s is characterized by the non-increasing utility function $u_s(\tau)$ of the queuing delay τ. Function $u_s(\tau)$ has a form shown in Figure 1 and is often used [2]-[7] in grid computing research.

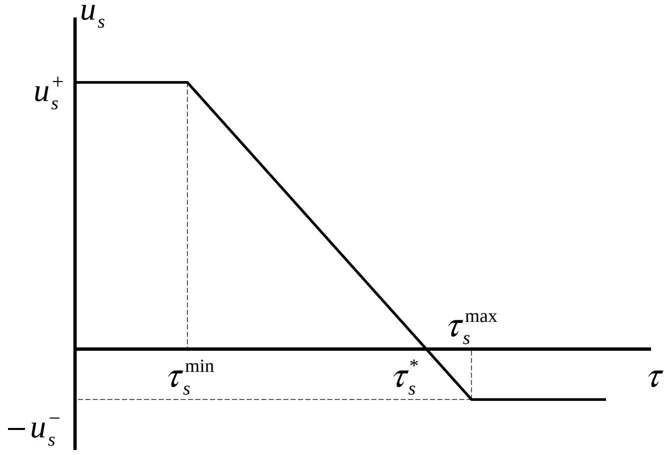

Figure 1. Generic utility function

Utility function $u_s(\tau)$ can be interpreted as the user willingness to pay.

We assume that utility functions $u_s(\tau)$ of all jobs of the same class $s = 1,..,S$ have the same cut-off and break-even parameters $0 \le \tau_s^{\min} \le \tau_s^* \le \tau_s^{\max}$ representing delay-sensitivity, but may have different parameters $u_s^-, u_s^+ \ge 0$ representing budget. Under our assumptions, utility of a user of class $s = 1,..,S$ as a function of the job queuing delay τ is

$$u_s(\tau) = u(\tau, \beta; \tau_s^{\min}, \tau_s^*, \tau_s^{\max}) \tag{1}$$

where function

$$u(\tau, \beta; \tau_s^{\min}, \tau_s^*, \tau_s^{\max}) = \begin{cases} u^+ & if \quad 0 < \tau \le \tau^{\min} \\ (\tau^* - \tau)\beta & if \quad \tau^{\min} < \tau \le \tau^{\max} \\ -u^- & if \quad \tau^{\max} < \tau \end{cases}$$

parameter $u^+ = (\tau^* - \tau^{\min})\beta$ represents the user budget, while parameter $u^- = (\tau^{\max} - \tau^*)\beta$ represents user dissatisfaction when the job is not completed. Piecewise linear utility function (1) represents willingness to pay as a base value (u^+) for completing a job on time (by $\tau^{\min}$) with a decreasing value for late jobs, up to some bound ($-u^-$).

We can simplify (1) to a linear utility function,

$$u_s(\tau) = u(\tau, \beta; \tau_s^*) \overset{def}{=} (\tau_s^* - \tau)\beta \tag{2}$$

as depicted in Fig. 2, where u_{0s}^+ represents base value for completing a job of class s without queuing delay and β represents the rate of decay in job value due to queuing delay. Linear utility function (2) ignores the offset ($\tau^{\min}$) and removes the penalty bound ($-u^-$). We will see that these restrictions simplify the analysis. If desired, the offset can be reintroduced later when using numerical methods.

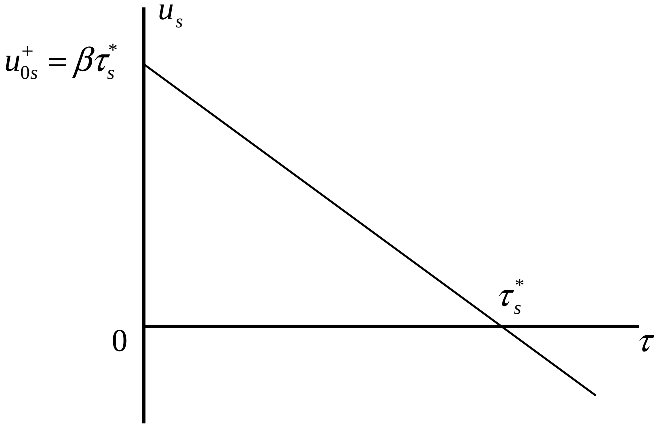

Figure 2. Linear utility function

Given queuing delay time τ, we assume that a provider charges a job of class $s = 1,..,S$ amount

$$p_s(\tau) = (\tau_s^* - \tau)q_s \tag{3}$$

where $p_{0s} = p_s(0) = \tau_s^* q_s$ is the base price for the job and q_s is a price reduction rate with increase in the queuing delay incurred by the job. Note that when $p_s(\tau) < 0$ the user is reimbursed for poor service. A user will submit job $j = 1,.., J_s$ of class $s = 1,..,S$ to a provider only when

$p_{0s} \le u_{0s}^{+}$. Thus, the service provider can control demand by varying base prices p_{0s}, $s = 1,..,S$.

III Provider Model: Pricing and Scheduling

A service provider can maximize revenue by controlling demand through pricing and by controlling relative job queuing delays through scheduling of jobs of various classes. The problem of finding optimal price and a related schedule for these jobs is nontrivial. In this section we show how solving two interrelated optimization problems (determining optimal price and a related, optimal schedule) allows a service provider to maximize revenue.

We assume the service provider uses a pricing scheme (3), which can be rewritten as follows:

$$p_s(\tau) = (1 - \tau / \tau_s^*)\, p_{0s} \qquad (4)$$

where base prices p_{0s} are intended to maximize revenue. We consider the following demand function [8]:

$$\lambda_s(p_{os}) = A_s\, p^{-\alpha_s} \qquad (5)$$

where A_s is demand potential and $\alpha_s > 1$ is price elasticity. For telecommunication data traffic [8], price elasticity has been measured to be $\alpha \in [1.3, 1.7]$. We also assume that jobs of class s arrive according to a Poisson process of rate λ_s. We assume an $M/G/1$ service model: all accepted requests are serviced by a single server of capacity C, with service time for requests of class s being a random variable with probability distribution $B_s(t)$ with moments $b_s^{(i)} = \int t^{(i)} dB_s(t)$.

The provider employs pricing and scheduling, as shown in Fig. 4. Since a completed job $j = 1,.., J_s$ of class $s = 1,.., S$ brings revenue (4), given queuing time $\tau = \tau_{js}$, the average provider revenue is

$$R = \sum_s p_{0s} \lambda_s(p_{0s})(1 - T_s / \tau_s^*)\,, \qquad (6)$$

where the average queuing delay for a job of class s is $T_s = E[\tau_s]$. Expression (6) represents the sum over all job classes of the base price for jobs in a class minus the decay in value determined by the expected queuing delay for jobs in the class, weighted by the proportional arrival rate for jobs in the class. The goal of the provider is to maximize the average total revenue

$$\max_{pricing} \max_{scheduling} R \qquad (7)$$

where the average total revenue is given by (6).

The case of a linear utility function (2) is comparatively simple because (a) average utility is equal to the utility of the average queuing delay, and (b) optimization problem (7) can be solved explicitly yielding the optimal scheduling. Indeed, in this case optimization problem (7) can be rewritten as follows:

$$\max_{(p_{0s})} \sum_s p_{0s} \lambda_s(p_{0s})[1 - T_s^*(p_0) / \tau_s^*], \qquad (8)$$

where the vector of average delays $T^*(p_0) = \left(T_1^*(p_0),.., T_S^*(p_0)\right)$ for a vector of base prices $p_0 = (p_{01},.., p_{0S})$ is determined by solution to the following optimization problem

$$\min_{(T_s)} \sum_s (p_{0s} / \tau_s^*)\, \lambda_s T_s \qquad (9)$$

Fig. 4 depicts decomposition of optimization problem (7) into problem (8) for optimal pricing and problem (9) for optimal scheduling.

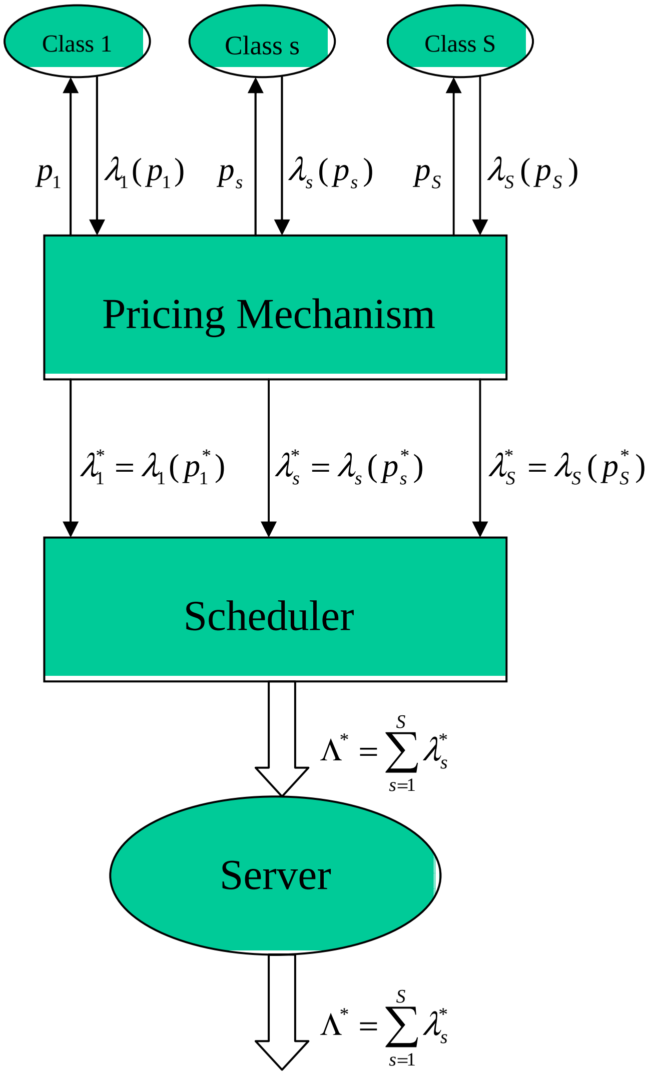
Class 1
Class s
Class S
p_1
$\lambda_1(p_1)$
p_s
$\lambda_s(p_s)$
p_S
$\lambda_S(p_S)$
Pricing Mechanism
$\lambda_1^* = \lambda_1(p_1^*)$
$\lambda_s^* = \lambda_s(p_s^*)$
$\lambda_S^* = \lambda_S(p_S^*)$
Scheduler
$\Lambda^* = \sum_{s=1}^{S} \lambda_s^*$
Server
$\Lambda^* = \sum_{s=1}^{S} \lambda_s^*$

Figure 4. Pricing and Scheduling

IV Decomposition: Pricing and Scheduling

Kleinrock [9] shows that, for the class of non-preemptive, work-conserving scheduling disciplines, the solution to optimization problem (9) is given by priority scheduling where priorities are assigned according to values:

$$f_s \overset{def}{=} p_{0s} \big/ (\tau_s^* b_s^{(1)}), \; s = 1,..,S \qquad (10)$$

Assuming that job classes are arranged in the following order:

$$f_1 \le f_2 \le .. \le f_S \, , \qquad (11)$$

the optimal scheduling discipline assigns priority to job class $i = 2,..,S$ over job class $j = 1,..,S$ if $i > j$.

The corresponding optimal average waiting times are

$$T_s^* = \frac{T_0}{(1-\sigma_s)(1-\sigma_{s+1})} \qquad (12)$$

where

$$T_0 = \frac{1}{2}\sum_s \lambda_s b_s^{(2)} \qquad (13)$$

server utilization by a job class $s = 1,..,S$ is

$$\rho_s = \lambda_s b_s^{(1)} \qquad (14)$$

and server utilization by job classes $i = s,..,S$ is

$$\sigma_s = \sum_{i=s}^{S} \rho_i \qquad (15)$$

For the optimal scheduling discipline, the average total revenue (6) becomes

$$R = \sum_s R_{0s}(p_{0s}) \left\{ 1 - \frac{T_0}{[1-\sigma_s(p_0)][1-\sigma_{s+1}(p_0)]} \frac{1}{\tau_s^*} \right\} (16)$$

where the base revenue rate (assuming that capacity is so high that the queuing delay is negligible) from class s is

$$R_{0s}(p_{0s}) = p_{0s} \lambda_s (p_{0s}) \qquad (17)$$

The optimal price vector $p^* = (p_{01}^*,.., p_{0S}^*)$ is determined by solution to the following optimization problem:

$$\max_{p_0} \sum_s R_{0s}(p_{0s}) \left\{ 1 - \frac{\sum_i \lambda_i(p_0) b_i^{(2)}}{2\tau_s^* [1-\sigma_s(p_0)][1-\sigma_{s+1}(p_0)]} \right\} (18)$$

In the case of a single job class, $S = 1$, provider revenue (16) becomes

$$R(p_0) \sim p_0 \rho(p_0) \left[1 - \frac{\theta}{\tau^*} \frac{\rho(p_0)}{1-\rho(p_0)} \right] \qquad (19)$$

and thus, optimization problem (18) takes the following form:

$$\max_{p_0} p_0 \rho(p_0) \left[1 - \frac{\theta}{\tau^*} \frac{\rho(p_0)}{1-\rho(p_0)} \right] \qquad (20)$$

subject to $p_0 > (Ab^{(1)})^{1/\alpha}$, where

$$\theta = \frac{b^{(2)}}{2b^{(1)}} \qquad (21)$$

and

$$\rho(p_o) = Ab^{(1)} p_0^{-\alpha} \qquad (22)$$

Optimization problem (20)-(22) has unique solution: $p_0^{opt} = p_0^{opt}(A, \alpha, \theta, \tau^*)$, which can be easily determined numerically.

Consider a particular case of jobs with low delay sensitivity: $\tau^* \to \infty$. Analysis shows that in this case the optimal price is:

$$p_0^{opt}(\tau^*) = \left(Ab^{(1)}\right)^{1/\alpha} + \left(Ab^{(1)}\right)^{1/(2\alpha)} \left(\frac{\theta}{\alpha\tau^*} \frac{1}{\alpha - 1} \right)^{1/2} \qquad (23)$$

the corresponding optimal utilization (24) is

$$\rho^{opt} \overset{def}{=} \rho(p_0^{opt}) = 1 - \left(\frac{\theta}{\tau^*} \frac{\alpha}{\alpha - 1} \right)^{1/2} \qquad (24)$$

Expressions (23)-(24) are instructive. If jobs are completely insensitive to delay, then the optimal price maximizing the provider revenue is determined by the condition that the provider is completely utilized: $\rho(p_o) = 1$. Combining this equation with (22) we obtain the first term in expression (23) for the optimal price: $p_0^{opt}(\tau^*)\Big|_{\tau^* = \infty} = \left(Ab^{(1)}\right)^{1/\alpha}$. Optimal price (23) increases with decrease in the job delay sensitivity $1/\tau^*$ as $(1/\tau^*)^{1/2}$, which results in server underutilization by a margin of $\sim (1/\tau^*)^{1/2}$ and job delays of the order of $\sim (\tau^*)^{1/2}$.

V NUMERICAL RESULTS: SINGLE CLASS

This section provides some initial numerical results. Due to limited space we consider only a case of a single job class. Figure 5 shows provider revenue as a function of the demand potential for fixed pricing without admission control and for several cases of delay sensitivity.

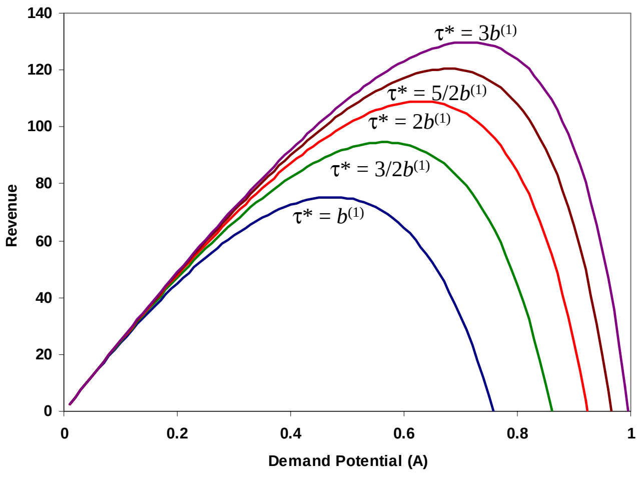

Figure 5. Revenue: Fixed Pricing without Admission Control

As admitted demand increases, provider revenue is a subject to two opposing trends: on the one hand revenue increases due to more jobs admitted, but on the other hand revenue decreases due to increase in the "delay penalty". The first trend is dominant for light demand while the second trend becomes dominant for heavy demand. Accordingly, as demand increases, the provider revenue first increases, then peaks, and after that decreases. The location of the peak depends on the delay sensitivity of users.

Figure 6 plots provider revenue as a function of the demand potential for fixed pricing and optimal admission control (described elsewhere [1]) as a function of the demand potential. The optimal admission control kicks in for sufficiently large demand, when revenue without admission control (see Figure 5) peaks and starts deteriorating due to the delay penalty. Optimal admission control rejects excessive demand to ensure that the revenue is kept at the maximum. The lower the users delay sensitivity the higher the maximum revenue available to the provider.

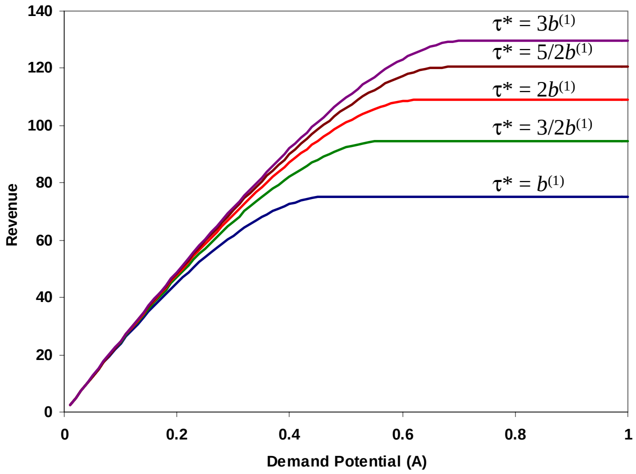

Figure 6. Revenue: Fixed Pricing with Admission Control

Figure 7 shows provider revenue for optimal pricing without admission control as a function of demand potential. Here, price elasticity is fixed at $\alpha = 1.5$. As demand increases, the provider is able to raise the price and extract more revenue. As demand potential increases, the provider can charge a higher price, while still generating sufficient customers to increase revenue.

Figure 8 compares provider revenue in all three cases: fixed pricing without admission control, fixed pricing with optimal admission control, and optimal pricing without admission control. Here, delay sensitivity is fixed at $\tau^* = 3b^{(1)}$ and price elasticity is fixed at $\alpha = 1.5$. As Figure 8 demonstrates, optimal admission control prevents deterioration of the provider revenue but does not take advantage of the possibility of increasing revenue by raising price. Thus, it may appear that optimal pricing eliminates the need for admission control. In the next section, however, we argue that in practical situations admission control may be necessary even when a provider is capable of price optimization.

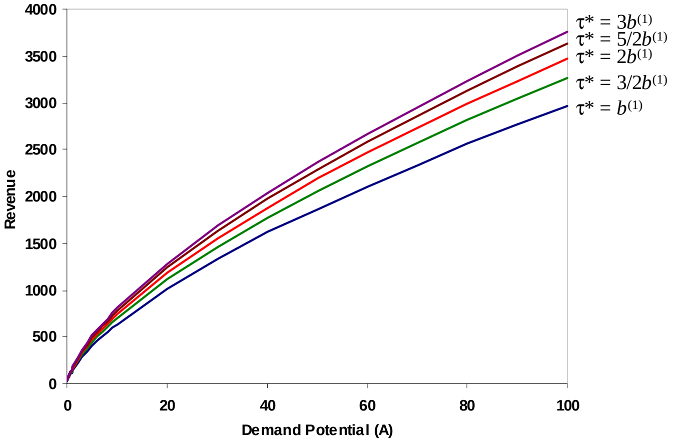

Figure 7. Revenue: Optimal Pricing without Admission Control

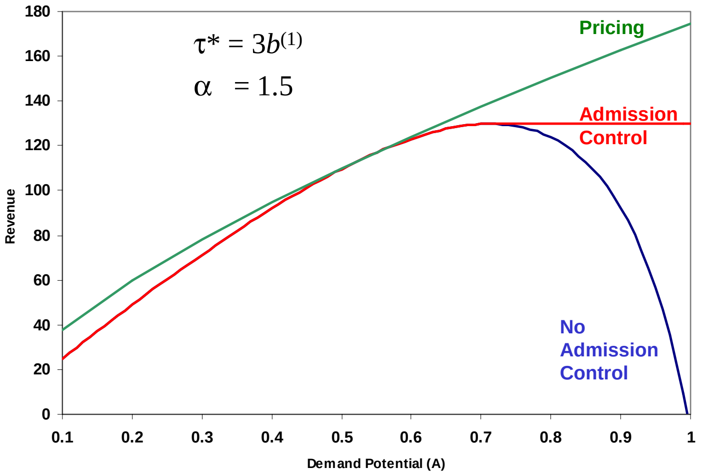

Figure 8. Revenue: Impact of Pricing and Admission Control

VI. Discussion: Pricing vs. Admission Control

Figure 9 shows provider revenue as a function of price for three different values of price elasticity and fixed demand potential. The general shape of the curves is the same. As price decreases, provider revenue is a subjected to two opposing trends: on the one hand, revenue increases due to increased demand (since demand elasticity $\alpha > 1$), and on the other hand, revenue decreases due to increases in the delay penalty, which arises as increased demand raises provider utilization. However, as Figure 9 demonstrates, the second trend is much sharper, and thus even slight underestimation of the provider price from the optimal value causes a sharp deterioration in provider revenue or even causes a provider to pay penalties due to sharp increases in the job delays.

This high sensitivity of provider revenue to pricing presents a serious problem since the optimal price depends on the price-demand curve, which may be a subject to statistical uncertainty and variability. Also, it may be difficult or even impossible, e.g., due to regulations, to vary pricing sufficiently fast to control delays. Admission control could alleviate this problem by reducing sensitivity of provider revenue to pricing non-optimality at the expense of some loss in revenue. Also, admission control operates on a sufficiently fast timescale to be able to control delays.

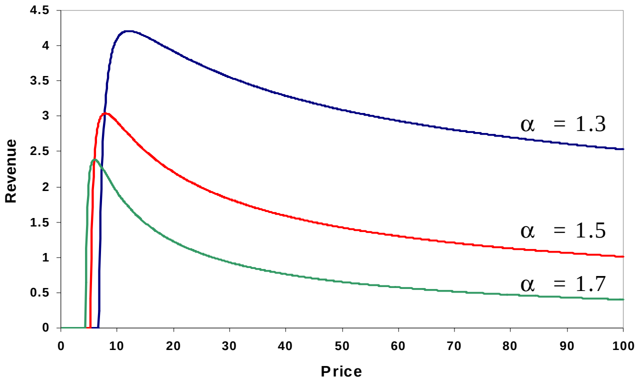

Figure 9. Revenue Sensitivity to Price

This suggests that a combination of pricing and admission control should be used. While pricing operates on a comparatively slow timescale, estimating the *average* price-demand curve and tracking the optimal price, admission control operates on a comparatively fast timescale, consistent with the tolerable delays. Admission control can absorb variations in demand that arise from the time lag in varying and disseminating prices. In practical situations, both pricing and admission control should operate jointly and adaptively depending on the average and instantaneous demand as well as queue sizes.

VII Conclusion and Future Research

Market-based compute grids encompass service providers offering limited resources to potential users with varying quality of service demands and utility (willingness to pay). Researchers typically model utility as a reward for completing a job by a deadline and a decay rate, which defines the slope of a linearly decreasing function of the reward over time if a job is late. As reward decays beyond zero, user utility becomes negative and a provider must pay a corresponding penalty. Under such conditions, providers face difficult job pricing, admission and scheduling decisions.

While previous research [2-7], typically, investigated various heuristics, this paper has proposed a tractable analytical model for joint optimization of job pricing and scheduling strategies aimed at maximizing provider revenue. We solved this model under the assumption that potential provider penalties are unbounded, and we analyzed key model parameters. We demonstrated how the model could be used to compute optimal pricing under a complex mix of jobs. Our model could be used to understand the operating limits of proposed heuristics for pricing, admission control and scheduling, and could also be used to investigate the implication of varying job mixes and workloads. Our results suggest that combination of pricing, operating on a slow timescale, and admission control, operating on a fast timescale, will be required for revenue maximization in practical applications.

Further work remains to investigate and develop such combined schemes, which should be closed-loop, measurement-based strategies. Though our optimization framework is applicable for an arbitrary number of service classes, solving the corresponding optimization problem for a case of more than one class may present a difficulty. Indeed, optimal priority scheduling (10)-(11) solving optimization problem (9) depends on the optimal pricing obtained by solving problem (8). Thus, attempt to decompose original optimization problem (7) into optimization problems (8) and (9) may result in an unstable, non-convergent process, which indicates that the optimal job scheduling lies outside the class of priority scheduling disciplines. We are currently modifying the decomposition procedure to include dynamic priority scheduling disciplines. We plan to develop an optimization framework, which will yield an optimal combination of pricing and admission control. We also plan to validate our optimization results using simulation and also against available data [10] on current web services.

VIII. Acknowlegments

The authors acknowledge support from the NIST “Complex Systems” program.

References

[1] V. Marbukh, K. Mills. “On Maximizing Provider Revenue in Market-based Compute Grids”, International Conference on Network Services (ICNS’07).

[2] B. N. Chun and D. E. Culler. User-centric performance analysis of market-based cluster batch schedulers. *Proceedings of the 2nd IEEE International Symposium on Cluster Coniputing and the Grid* May 2002; pp. 30-38.

[3] D.E. Irwin, L. E. Grit and J. S. Chase. Balancing Risk and Reward in a Market-Based Task Service. *Proceedings of the 13th IEEE International Symposium on High Performance Distributed Computing* 2004; pp. 160-169.

[4] F. I. Popovici and J. Wilkes. Profitable services in an uncertain world. *Proceedings of the ACM/IEEE Supercomputing Conference*. 2005; pp. 36-47.

[5] C.S Yeo and R. Buyya. Service level agreement based allocation of cluster resources: Handling penalty to enhance utility. *Proceedings of the 7th IEEE International Conference on Cluster Computing* 2005.

[6] D. Vengerov Adaptive Utility-Based Scheduling in Resource-Constrained Systems. *Proceedings of the 18th Australian Joint Conference on Artificial Intelligence*, December 2005; LNCS 3809: pp. 477-488.

[7] A. AuYoung, L. Grit, J. Wiener and J. Wilkes. Service contracts and aggregate utility functions. *Proceedings of the 15th IEEE International Symposium on High Performance Distributed Computing* 2006; pp. 119-131.

[8] D. Mitra, K. Ramikrishnan, and Q. Wang, “Combined economic modeling and traffic engineering: joint optimization of pricing and routing in multi-service networks,” *Proc. ITC*, Brazil 2001.

[9] L. Kleinrock, Queueing Systems, Volume II:Compute Applications, John Willey, 1976.

[10] A.K. Iyengar, M.S. Squillante, and L. Zhang. Analysis and characterization of large-scale Web Server Access patterns and performance. *Proceedings of the Conference on World Wide Web*, 1999.