Node Acquisition Bid Evaluation Formulas and Economic Model

Economic Model Subcommittee
Node Procurement Task Force

October 6, 2005

Abstract

We propose a structure for creating bid evaluation formulas associated with "farm node" acquisitions, and discuss the underlying economic model justifying the formulas. The proposed formulas include several components which were not quantitatively taken into account in earlier acquisitions, the most important of which is the cost of power consumption.

This document, when approved by the full task force, will constitue a recommendation to the Computing Division by the Node Aquisition and Procurement Task Force.

Contents

Chapter 1

Executive Summary

We are recommending a structure for creating the evaluation formula to be used in determining the winning bid for a given node acquisition. No two acquisitions are identical; the structure describes hwo to turn input about the current situation and acquisition into a qunatitative formula for choosing the winning bid.

The resulting formulas would differ from the typical evaluation formulas used in past acquisitions in the following ways:

1. The "figure of merit" is a ratio of Effective CPU Power to Effective Cost, rather than a score of CPU power alone.
2. Differences in electic power consumption (measured as Amps consumed) appear in the formula, heavily influencing the Effective Cost denominator.
3. Several other factors which are of some importance to our support of the acquired nodes are considered quantitatively.
4. By assigning quantitative values to considerations which might in the past have been treated as acceptance criteria, the evaluation formula makes for more flexibility in what vendors can bid (thus potentially opening the process to more and perhaps occasionally better bids).

1.1 The structure of each evaluation formula

For each bid, determine how many nodes would be acquired assuming the maximum number possible without direct cost exceeding the acquisition monetary limit. Also verify that all the accept/reject criteria are met. Then a figure of merit is determined for each bid, consisting of the ratio of an Effective Value to an Effective Cost for that bid. The highest figure of merits among nod-rejected bids is the winner.

1. Effective Cost, measured in dollars, starts out as the cost per node bid, times number of nodes.
2. Effective Value, measured in FU (Fermi Units) starts out as the CPU power per node (as measured by the benchmark suite) times the number of nodes.
3. Measure the average amps used by a node while running the benchmark suite. Add to the Effective Cost the power consumption cost of $367 per amp times the number of nodes.

4. Measure the maximum amps used by a node while running the various segments of a "maximum amps suite." Add to the Effective Cost the power avaibility cost of $101 per amp times the number of nodes.
5. Add to the Effective Cost a networking cost of $395 per node.
6. If the bid does not include vendor assembly of the nodes into racks on site at FNAL, add to the Effective Cost assembly costs of $200 per rack.
7. The Effective Value is adjusted by .5% per week of anticipated delay or more rapid in-service date caused by the proposed scheme of how the nodes will be assembled.
8. The Effective Value is adjusted by adding 1% if the discs bid are the more reliable SCSI discs.
9. *There may appear one additional input having to do with type of rack; we have not yet come to a resolution in this area.*

The numbers appearing above are not firmly determined and may vary slightly from one acquisition to another. In order to able to accurately produce the numbers used in these formulas, we should:

1. Always strive to make the evaluation benchmark suite as representative as possible of the anticipated program mix.
2. Create a suite suitable for measuring maximum amperage (which dictates the amount of power/cooling capacity required).
3. Attain the capability of reasonably accurate assessment of electric power a node consumes while performing the benchmark suite and the maximum aperage suite.
4. Settle in advance on accurate values to use for the costs of power and cooling, network connectivity, reliability characteristics, and various other features.

1.2 Unresolved issues

The following issues were discussed in task force meetings without any firm resolution. For the purposes of this document, we have assumed the resolutions indicated.

1. The issue of system assembly. The leaning was toward saying that final assembly of installation needs to be done on site but we can compare bids with the vendor doing the assembly to those wehre we must cost doing (or contracting) the ssembly ourselves. It is not feasible to allow for delivery of fully populated 40-node racks.

1.3 How accurate is "accurate enough"?

If we feel the precision of our bid formula leads to evaluation uncertainty which is small compared to the average distance between the best bids, then further improvements are not worthwhile. Until that point, each 1% improvement in evaluation accuracy translates to about a .15% expected advantage in true cost, and we should decide whether to expend further efforts in refinement with that level of potential gains in mind.

One consequence of these numbers, in concert with the fact that each week's delay of an acquisition is equivalent to a half percent decline in the value received, implies that the evaluatin procedure must be sufficiently simple. It would be misguided to complicate the evaluation procedure to the extent that it might delay bid acceptance and arrival of the nodes.

Chapter 2

Relevant Evaluation Considerations

The consequence of omitting an important consideration from the bid evaluation is that the extent to which the winning bid does well regarding that omitted consideration would be random with respect to all the bids. That mean we *might* bave done better by choosing some other bid.

The evaluation considerations deemed relevant are determined by consensus among the task force. A balance was struck, choosing to minimize the complexity of the bid evaluation formulas by incorporating only those considerations which are thought to have significant potential impact.

2.1 Quantified considerations

In rough order of importance, the following considerations will appear in the bid evaluation formula:

Bid cost : The cost for all the nodes if the bid is excercised to the monetary limit of this acquisition. This is the starting point for Effective Cost and is likely to nearly identical for all bids.

CPU power : The benchmarked CPU power, in FU (Fermi Units) for a node, times the number of nodes acquired if the bid is excercised to the monetary limit of this acquisition. This is the starting point for Effective Power. One of the most important factors in having an accurate and useful bid process is a choice of benchmarks that well reflects the mix of uses for the nodes acquired (as opposed to just SpecInts).

Current consumption The average current consumed by a node during execution of the benchmarks, times the number of nodes acquired if the bid is excercised to the monetary limit of this acquisition, times an effective cost per amp. This is added to the Effective Cost of the bid. The effective cost for current consumption is the sum of two components, each usiong a different measure of amperage:

1. The power consumption cost: Kilo-Dollars per KWH, times voltage (to translate amps to watts), times number of hours in the node lifetime, times a cost-of-cooling factor. This is to be multiplied by the amps consumed.

2. The power and cooling availability cost: Building and equipment costs amortized over their lifetimes, expressed in dollars per amp. This is to be multiplied by the "maximum amps" consumed.

Inserting reasonable numbers (§3.1.1), we arrive at a figure of $367 per amp plus $101 per maximum amp. Given that the typical node draws about 2.5 amps, this $1,200 correction is a substantial fraction of the base cost of the node, so there is considerable potential for gain in this refinement to the bid process.

Networking cost Cost per box of connecting to the network, times number of nodes acquired if the bid is excercised to the monetary limit of this acquisition. Assuming that a single network node would have satisfactory bandwidthfor even a dual-CPU dual-core node, this could give systems needing fewer network connections an advantage. The cost per node (attribution here to Phil DeMar) is $325 amortized switch cost, plus $50 for lateral cabling infrastructure, plus $20 for ed connections and jumper cables – a total of $395 per node.

Assembly Costs The bid proposal should allow for vendors doing the assembly and for vendors who do not wish to include assembly of the nodes into populated and set-up racks.

- If the assembly of nodes is to be done on site by the vendor, and the contract specifies as the "delivery date" the "nodes are here and assembled" date, then the assembly adds zero to the effective cost.
- If we will have to contract for assembly of the nodes into populated racks, we should add a fixed amount per node based on the cost to us of doing this contracting. The cost number to use for this is some figure (roughly $200) per rack (since the cost of conecting up the individual nodes is included in networking costs). *We are not solid on the $200 number; need to get info.*
- In addition, we should assign (pre-determined according to a chart) anticipated time to "nodes are here and assembled" for different schemes of assembly. (For example, we should have some estimate of how long it will take if we need to contract it ourselves.) We should add or subtract (for fair comparison) .5% per week of difference in anticipated "nodes are here and assembled" dates. This should be added or subtracted to the Effective Value.
- Because it would take considerable up-front effort to prepare to cope with a potential winning bid consisting of delivering fully populated racks to the site (and having riggers help move them into place in the buildings), we recommend *not* allowing this as a bid optoin.

Rack *I'm not sure what to put here but it seemed clear this made it into the consensus for being in the bid. There was talk about a preference for AMCO racks.*

SCSI Disks We have a preference for the additional reliability of SCSI disks, which will translate into more up-hours for the average node. In order to reflect this preference, yet not unduly constrain bids which prefer less costly discs, the Effective Power number per node should receive a 1% bonus for SCSI discs (as long as the disc capacity and speed meet the minimum acceptance criteria). We arrive at the 1% figure by saying that half of all non-burn-in downtime comes from disc problems, which will go away when SCSI is used. Figuring one down week per year for each node, we obtain the 1% figure.

2.2 Temporary: Considerations I'm not sure are to be included

- Reliability adjustment (other than SCSI disks)
- Effective Cost addition for an unfamiliar node administration scheme.

2.3 Recommendations concerning qualitative acceptance criteria

Disk space and memory should be phrased as minimum requirements, made with an eye toward levels which are fairly standard and won't require vendeors to bid (higher-priced) unusual configurations.

Floppy disk and CD are, by consensus, requirements.

The ability to run the operating system we want to stndardize on should be a requirement insisted upon, not a "bonus."

2.4 Considerations felt to be of low impact

The following considerations were discussed, but in the end felt to be of low impact (or just not proper), and thus are to be omitted in the interest of keeping the bid and formulas as simple:

- Effective power bonus for additional disk space or memory beyond the required levels.
- Effective power bonus for inclusion of cyclades, which might be felt to be a safer control/power alternative than IMPI.
- Single vendor point of support. Although there was some sentiment for avoiding potential finger-pointing circuits, by insisting that one vendor take responsibility for the system soup-to-nuts, this did not in the end appear to have many true advocates.

Chapter 3

Tailoring a Formula For an Acquisition

In the previous chapter, we specifieded the form of the bid evaluation formula. Several numbers appear which might not yet be well determined, or which might change between one bid and the next. Here, we will discuss how to determine those numbers.

3.1 Recommendations for determining input quantities

3.1.1 Effective cost per amp determination

The effective cost per amp consists of two parts. The larger part is the power consumption cost, which is very well justified and is precise to the degree that we know the future cost per KWH. The second part is the power availability cost, which represents a the cost of providing the bilding and power/cooling equipment. We know These costs very accurately, but when we use a power availability cost, we are assuming a model approximating the non-fungible asset of power capability by a fungible asset costing the same amount per unit power.

Thus, neither component is known perfectly, but each is known well enough to improve the fidelity of the bid evaluation process, which after all is the goal.

What should we measure for cost per amp?

For the power consumption comption cost, the appropriate amperage to use is the average current consumed while running a representative mix of programs; the benchmark suite is the ideal mix for this measurement. We shall call this simply the ‘amps” (per node for a given acquisition).

However, for the power availability cost, we need to supply power adequate for the maximum non-startup power usage of all the nodes. (We assume that we can stagger startups so as to avoid a huge spike due to simultaneous startups of many nodes.) For that purpose, we should measure the maximum power usage. This may be more difficult, but one thing we can do is to take the maximum of the current consumption in each of several small components of the benchmark, as being representative of the maximum current the node will actually draw. We shall call this the “maximum amps.”

Power consumption cost

For the power consumption cost, we first need to estimate the cost of electricity over the lifetime of the nodes. It is best neither to factor in anticipated inflation-tracking increases in power price, nor to discount future payments to account for inflation; those two effects cancel out. If, however, we know of some anticipated price increases based on other effects (such as coming deregulation) we should account for those. Our current number for the cost of electricity is 5.2 cents per KWH this year, but 7 cents each of the next two years, for an average cost of $.064 per KWH.

The next uncertain cost is what factor to use for the cost of cooling. Cooling cost clearly proportinal to power consumed, thus this is a "cost-of-cooling factor." For purely electrically air-conditioned buildings, this can perhaps be measured by comparing the total power going into the building (including air conditioning) to the power going to the nodes. This would have to be integrated over a full year, and might need to be adjusted downward to account for the cost of cooling the building even if the nodes were not present, and the savings due to needing less explicit heating in the Winter. If a building makes use of pond-cooling, that complicates matters further. Our current guess for this factor is 1.7 but this can easily be off by up to 30%.

The last component in power consumption cost is what to use for the node lifetime. The consensus appears to be that 3.5 years is about right.

Using these numbers, the power consumption cost works out to $367 per amp.

Power availability cost

The cost of making available building space with conditioned power and cooling for an amp worth of node power consumption can be determined pretty accurately: The last 4 buildings we have equipped for the purpose of housing nodes cost an average of $3169/KWH. If we amortize the 3.5 year node lifetime over a 12 year building/equipment lifetime, this works out to $101 per amp.

We cannot combine this with the power consumption cost since in the case of power availability

3.2 Recommended studies

For the first round of acquisitions, we don't have solid grounding (yet) on numbers which may not change much in subsequent acquisitions. Since thses will be used more than once, further careful refinement of the following quantities will be valuable:

1. Cost-of-cooling factor (which goes into power consumption cost): Either base the number on decent histories/measurements of power consumption and cooling costs, or do some more in-principle calculation. In either case, consider whether to include incidental heating savings as discussed earlier.

2. Power availability cost (which goes into power consumption cost): Recheck the frigure given for average cost per KWH in buildings, and decide whether the appropriate equipment/building lifetime to use for amortization is 12 years or something else.

3. The networking experts should supply a figure to use for per-node cost of network connection to a node.

4. We should be able to quantify the anticipated downtime due to EISA disks rather than SCSI so as to have a justifiable bonus for SCSI disks.
5. We need to decide what to do about local assembly and how to cost it.
6. We need to decide what to do about racks – just insist on AMCO, pre-qualify other brands, come up with mandated requirements (welds, not screws, steel not aluminum) – and whether and how much to adjust for better or worse choices on things we don't mandate.

None of these studies should involve much time or effort, so these numbers should be delivered as part of the product of this task force. At that point, the nature of this section ought to change from "*Recommended studies*" to "*Constants appearing in the formulas*".

3.2.1 Techniques we need to learn

In order to apply the bid evaluation formulas correctly for each acquisition, there is at least one evalaution step which we need to learn to do more accurately than has been done in the past. Also, there is one critical area (benchmarks) where it would be good to put in place procedures to make sure we continue to make measurements with sufficient accuracy.

The overall target is to try to make the evaluation formula meaningful at the 3% accuracy level. Beyond that, the payoff in terms of possibly choosing the better bid is small, compared to the FTE costs of trying to further refine the measurements. This is discussed in section 4.3.

- An accurate way to measure electric power consumed by a node while running the benchmark suite. Since the overal power cost will work out to be more than a third of the overal effective cost, we should shoot for reproducibility at a level of 5%. (That is, the measurements for two nodes with equal actual power consumptions should agree within 5%.) An overall scale uncertainty of up to 10% is acceptable, since we can't estimate future power costs more accurately than that.
- A segmentation of the benchmark suite, or a set of other defined programs, such that we can say that the maximimum current draw (which dictates the power availability requirement) is the maximum current measured over any one of N seqments of the suite.
- An ongoing process to make sure we don't forget to keep the benchmark suite(s) in step with the anticipated usage patterns.

Chapter 4

The Underlying Economic Model and Assumptions

4.1 The assumptions

discussion here

4.2 The economic model

brief discussion here

4.3 How precise a formula shoot we try for?

It is obvious that any evaluation formula is only an approximation to the true value, to our users, of the nodes acquired, and to the true effective cost of the bid. At one extreme, one could imagine that the formula had no useful resolution power at all; all bids would be equally likely to win, and we might as well just flip a coin. At the other extreme, the evaluation is absolutely accurate and we would always choose the best bid.

But an evaluation which is very slightly less than perfect will yield almost the same expected true value, since most of the time, no two bids will be close enough for the difference form perfection in evaluation to change the choice, and even when it does, the loss in value is tiny. Given that it takes real work to make a super-accurate evaluation formula, there must be some "sweet spot," some notion of how good is "good enough."

To get an estimate of the cost of imprecision, let us simplify be saying that we are going, at any rate, to choose one of the best *two* bids, and that the bids can be considered as Gaussian variates with uncertainty σ. (σ can be estimated by $\sqrt{\pi}$ times the average distance between the best two bids, based on past acquisitions. We should actually say that σ is some percentage of the total bid value.) Then relative to the null evaluation formula of choosing one of the two bids at random, the best we could do with an absolutely precise evaluation formula would be to gain (at average) $\sqrt{1/\pi\sigma}$.

This value is obtainted by integrating, over values of the difference in bids x (which is a Gaussian variable with uncertainty $\sqrt{2}\sigma$), the difference between the value obtained and the mean (which is $x/2$). For perfect accuracy, the integral will be broken into two parts:

When $x = V_1 - V_2$ is below zero, option 2 would be selected, so when $x < 0$ the integrand contains $-x$ instead of x.

The effect of an uncertainty $\sigma_e \neq 0$ in the evaluation formula will be to change the breakpoint in the integrals from zero to y, where y is a second Gaussian variable representing the evaluation error. Thus there is some component of $+x$ in the left-side integral, or of $+x$ in the left-side integral, depending on whether y is negative or positive. The expectation value becomes:

$$\frac{1}{2\sqrt{2}\pi\sigma\mu} \int_{y=-\infty}^{\infty} e^{\frac{y^2}{2\mu^2}} \left[\int_{x=-\infty}^{y} (-x) e^{\frac{x^2}{\sigma^2}} \, dx + \int_{x=y}^{\infty} (x) e^{\frac{x^2}{\sigma^2}} \, dx \right] dy$$

which can be integrated exactly:

$$\frac{\sigma^2}{2\sqrt{\pi}\sqrt{\sigma^2 + 2\mu^2}}$$

We can obtain the percentage differences between the two best bids in the last several acquisitions; for now let's assume that they differed by an average of 5% on million-dollar acquisitions. Then per million-dollar acquisition:

- The cost of extreme evaluation inaccuracy (assuming either of the two best bids would be chosen randomly) would be an average loss of $25,000.
- 10% evaluation inaccuracy would be $13,275 better than random.
- 7% evaluation inaccuracy would be $16,675 better than random.
- 5% evaluation inaccuracy would be $19,550 better than random.
- 3% evaluation inaccuracy would be $22,550 better than random.
- 2% evaluation inaccuracy would be $23,815 better than random.
- 1% evaluation inaccuracy would be $24,685 better than random.
- These savings scale with the 5% average bid difference.

The improvement per unit accuracy gain peaks strongly around 5%, the average difference between the two best bids. This is expected, since if the evaluation uncertainty is very small compared to that value, then you will probably pick the best bid anyway, and increased accuracy probably gains nothing.

Figuring that the evaluation process will be going on for at acquisitions totalling $10,000,000 over the lifetime of this process – and this is probably a high-end figure – then an improvement of 1% in accuracy in evaluation will yield a gain of $15,000 over the years the formula is used. So it is worth shooting for such gains only if they can be achieved with less than about .1 FTE of work.

And, remembering that all the sources of evaluation inaccuracy add in quadrature, it is very unlikely to be worth persuing any but the largest uncertainties.

An example: Since one type of system can easily be 30% more wattage efficient than another, and power consumption itself accounts for about a third of the overall true cost, omitting power consumption altogether would introduce a 10% error in evaluation, and this is well worth correcting. but once we have this measured with, say, 5% precision, further correction to absolute accuracy is only worth a gain of $7,500 or less, and is not worth the effort it would take.

4.3.1 The penalty for introducing delay

It is observed that the cost of delay introduced by the bid evaluation process would be about half a percent per week. Two schools of thought independantly reach this same number: One can say that the end of the node lifetime is dictated by the timing of some future acquisition, and thus if we get the nodes a week later, we will be getting about 1/200 less use out of them. Or one can say that there is a Moore's law decrease in the value of a given amount of computing power; assuming a doubling time of 3 years, this means that a week's delay is about like a half percent less value.

Given that a 1% improvement in accuracy of evaluation only leads to (at most) a .15% increase in expected value of the acquisition, this tells us that the evaluation criteria can't be too complex. Complicating the evaluation procedure to the extent that it might delay–even slightly–the bid acceptance and arrival of the nodes, is a calamity which must be avoided.

4.4 Fungible and non-fungible assets

The definition of a fungible asset is a cost associated with acquiring nodes in some way, which can be translated directly to a dollar cost. This dollar cost can then be added to the price of the nodes, contributing to some effective cost of the acquisition.

A non-fungible asset is one which is (or will be) limited in a manner that we can't easily buy our way out of. The relevant example is conditioned power and cooling capacity. (These are very closely related and are also coupled with space for racks.) We know with excellent precision how much conditioned power availability will be present for each of the next several years. If by acquiring "power-hungry" nodes today we place ourselves in the position of running out of conditioned power capacity in a future year, we could deal with this in one of three ways:

1. Plan to buy fewer nodes (or less power-hungry but less CPU powerful nodes) in the future than our budgets would otherwise support. We would take this into account as a decrease in the "value" of this year's acquisition.

2. Plan to retire existing nodes earlier to make for enough power headroom. This, too, represents a decrease in the "value" of the acquisition.

3. Plan to build additional power and cooling capacity, which might in the extreme be amortized only over the lifetime of this power-hungry acquistion. This would appear in the effective cost, as a very non-linear break in the cost of power and cooling.

Each of those options involves considerable complexity (and some shaky input) in the recipe for evaluation formulas. We would like, instead, to approximate the non-fungible asset by a nearly equeivalentt fungible asset (see the next section).

There is one extreme in the notion of non-fungible assets, which can in no sense be treated by a fungibility approximation. If the limitation associated with an asset is immediate (for example, if the DOE had dictated that we are not permitted to purchase more than 300 CPU's in a given year) then this could not look like a quantitative component of the bid formula. However, we are not discussing such extreme cases in the economic model or the bid evaluation formula; instead they appear in the bid process as bid rejection criteria.

4.5 Avoiding the non-fungible asset model

For this recipe for creating bid evaluation formulas, even if power and cooling capacities are projected to be tight in the near future, we still use the "fungible asset" model for power/cooling/space. That is, we assign some unit cost to the asset and ignore the fact that the cost becomes highly non-linear if we exceed the availability limit.

In particular, we add to the Effective Cost component of the figure of merit an 'amps available" cost A:

$$A = IE\frac{L_{\text{nodes}}}{L_{\text{bldg}}} \tag{4.1}$$

where I is the current drawn by the proposed nodes, E is the cost per amp of the building and equipment, and the fraction amortizes the building and equipment over the lifetimes of the nodes that will be served.

We make this choice because:

- In order to incorporate the effects of non-fungible assets any more accurately than this, we must deal with input concerning:
  - Predicted "Moore's Law" evolution of cost, CPU power, and current consumed for several potential future technology choices.
  - A justifiable model of uncertainties in the amount of computing power "needed" and of the nature of "diminishing returns" physics oportunities opened when additional CPU power is made available.
  - Quantative expression of the notion that physics enabled by computing acquired today is worth more than the physics enabled by the same amount of computing aquired some time later.
  - Reliable projections of future node acquisition profiles.

  Since there are no solid anwers to these issues, attempting to account for the non-fungible nature of future power and cooling capacity would do little to increase evaluation accuracy, and would subject the evalueation formula to considerable second-guessing about how to justify the numbers.

- Models which more seriously consider restricting this year's spending of our "power budget" are inconsistent with the policy that we spend the fixed amount of funds in the current fiscal year, rather than considering saving money for larger future acquisitions. This implies that we could never create a defensible recipe for creating evaluation formulas based on non-fungible assets, without also considering changes in the way acquistion funding is determined. Such considerations are quite beyond the scope of this committees recommendations.

- The model of a non-fungible asset with some fixed availability profile is itself imperfect. For example, given two years' warning, it is possible to do something about a shortage of power and cooling anticipated for a future acquisition (though it make be more awkward and costly than we would like).

- By considering power and cooling costs (including the amortized costs of building power and cooling capacity) in the effective cost for a bid, we already are doing much more to take this issue into account than previous bid processes. Further refinement is not worth the considerable complexity it would introduce.

In the case of conditioned power and cooling, the cost assigned is the cost of providing that capacity in a dedicated building, amortized over the lifetime of the equipment and building by mulitplying by the fraction of node lifetime over equipment/building lifetime. This estimate is reasonably unbiased and gives easily justified numbers.

Although this number appears in the bid evaluation formula in precisely the same way as power consumtion costs (and in fact is lumped in to form a single number) it really is a distinct contribution to the effective cost of the bid.

This is some sense in which the amortized equipment cost is an under-estimate: The collection of nodes will ultimately utilize less power than the total provided, yet this estimate is based only on the portion of power that is consumed. But there is also a sense in which this is an over-estimate of the marginal cost: Given hard-to-change plans for power and cooling installlation, it is not always possible to save the fair share of money by economizing on power availability. These two biases are in oppoite directions, so we can take this to be a relatively unbaised estimate of the cost of using up power and cooling availability.