Failing with the Best Intentions

Enterprise application capacity planning is a difficult juggling act. On one side of the equation you have business demand, looking for innovative technology to help improve business performance and increase profitability. The IT organization stands on the other side of the equation, responsible for satisfying these demands. The capacity of this team is limited by the organization's facilities, the number of developers and their specific skills, and the infrastructure and tools they use. This leaves the business and technology executives in the unenviable position of trying to balance the demand for IT development with the current capacity levels.

Both the business and IT leadership need to understand the demand for new development and the organization's current and forecasted capacity to fulfill these demands. Organizations have implemented project and portfolio management (PPM) systems to track these two different but connected factors. These PPM systems are great for allocating known resources against approved projects. They can also aggregate the data into a portfolio to give a current picture of how the IT capacity satisfies business demands. When the current demand exceeds the capacity of the IT organization, the IT team must delay or eliminate projects which have a negative impact on the business. These types of surprises are all too common, which leads us to question why this is.

The first explanation is that the estimation methods used by most organizations often only address the total effort expended by each role, and does not provide a schedule for how those resources will be applied over time. It is the estimator's job to review the requirements and estimate the number of labor hours required for each task. The task estimates are rolled up into a total labor estimate, which is then allocated out to labor categories. However, it doesn't show how they ramp up and taper off the project, an important component of capacity planning. So using today's most common practices, one ends up with a project estimate that looks something like Figure 1.

Bottoms-Up Estimate

Figure 1: Typical Bottoms-Up Estimate without Project Schedule

Unfortunately, this method also does not include a schedule. The schedule is usually decided arbitrarily by the stakeholders, and often does not align with the estimated effort. Quantitative Software Management research shows that unrealistic schedule expectations are the number one reason why projects fail.

This brings us to the second issue with current approaches to estimation. Task-based estimation is not well matched for the information that is available during the early stages of a project when the estimates and business decisions are typically made. The initial estimates are generally performed when there are only a few pages of business requirements. Rather than task-based estimation, top-down parametric estimation tends to be the most effective in the early stages of a project lifecycle.

Getting it Right Early in the Lifecycle

Parametric estimation utilizes gross sizing and productivity assumptions to produce schedule and effort estimates. If the estimates do not match expectations, top-down estimation tools can quickly evaluate what can be accomplished with the available resources, or determine how much additional time and effort is required to achieve the business objectives.

The beauty of the top-down method is that it uses crude size measures like business requirements, agile epics, use cases, or function points as a direct input. Coincidentally, these are the design artifacts most likely to be available during the early stage of estimation. With this basic information, top-down methods can identify both

high risk and conservative expectations, and if appropriate, suggest more reasonable alternatives.

Negotiation Based in Reality

Let's take a look at a typical estimation situation early in the lifecycle. For example, suppose a business has identified some new capabilities it would like to implement. It has a five-page document that describes 19 business requirements or capabilities. They would like to have the capabilities available within three months and have a budget of $250,000. Figure 2 shows its desired outcome.

Software Estimate Risk

Figure 2: Risky desired outcome compared with the recommended, more realistic estimate.

Compared to historical data, this solution is considered risky, and an alternative more typical estimate of 5.7 months and $571,000 is recommended. Now a negotiation between the stakeholders and the IT department can take place to explore more viable options. The main purpose for estimating at this stage of the demand management process is to smoke out grossly unrealistic expectations and negotiate realistic options in order for all parties to be winners in this development.

Multiple Planes for Forecasting

What is needed from the beginning is an estimate of how the various skills will ramp up and taper off over the course of the project. With this information, PPM systems can match individuals and their respective skills to the project at the time they are required, and release them when their work is done. See Figure 3. It is important to recognize that one size does not fit all projects. For a large enterprise there might be several skill profiles or templates to support the different project types and methodologies.
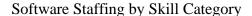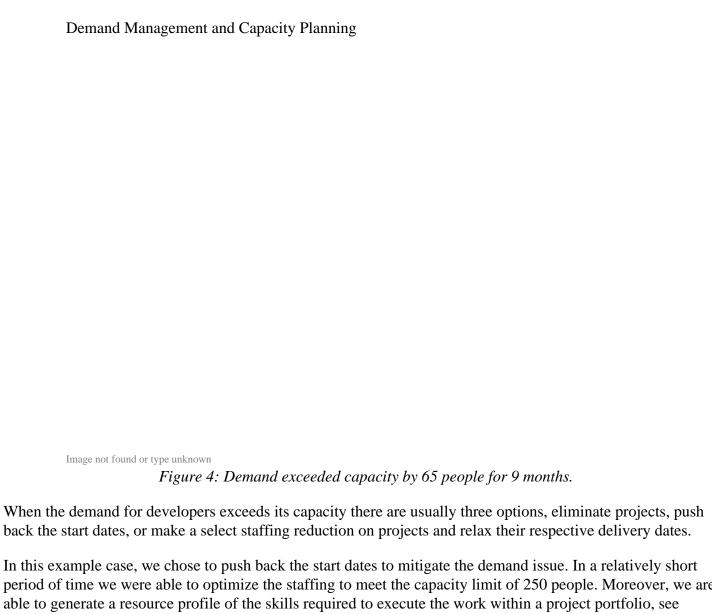
Image not found or type unknown

Figure 3: The monthly staffing of skill categories shows how resources roll on and off a project.

An Answer to "What If"

With a way to realistically estimate schedule, effort, cost, and skill requirements over the duration of a project, and feed that information into a PPM system, which enables insight into how the demand matches capacity at the enterprise level and the ability to evaluate practical options for matching demand and capacity.

For example, imagine an enterprise with three different development centers, one each in the U.S., India, and Ireland. The total capacity for these three centers is limited to 250 IT developers. 35 projects are currently approved for production.

Each of these projects are in a different stage of production, with some in the early planning stages, others beginning coding, and some approaching delivery. Each project was estimated individually and they can be rolled up to show the aggregate data at either the data center level or the enterprise to help determine the capacity and demand levels. When the data is aggregated at the enterprise level, we notice that for a nine month period, between August 2014 and May 2015, the demand for developers exceeds capacity by 65 full time equivalents, see Figure 4.

Image not found or type unknown

Figure 4: Demand exceeded capacity by 65 people for 9 months.

When the demand for developers exceeds its capacity there are usually three options, eliminate projects, push back the start dates, or make a select staffing reduction on projects and relax their respective delivery dates.

In this example case, we chose to push back the start dates to mitigate the demand issue. In a relatively short period of time we were able to optimize the staffing to meet the capacity limit of 250 people. Moreover, we are able to generate a resource profile of the skills required to execute the work within a project portfolio, see Figure 5.
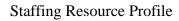
Staffing Resource Profile

Image not found or type unknown

Figure 5. The staffing resource profile shows the necessary skills that will be needed across the project portfolio to execute the work.

Setting up a PPM system is a great start to managing capacity and demand, but often it is not enough. To manage capacity and demand effectively, one needs the ability to realistically forecast demand early in the lifecycle; negotiate functionality; schedule and effort with business stakeholders when their expectations are unrealistic; have the ability to forecast effort by skill category, month, or week, and feed this information to PPM systems; and be in a position to perform quick what-if analyses on the portfolio when the demand exceeds capacity.

Utilizing the top-down approach of parametric estimation enables all four and helps organizations find true value in their planning intentions.