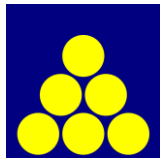


Capacity Management: Key Activities for Capacity Planning

Simplify process documentation and instruction by describing the structure for the key activities performed in capacity planning as a series of steps or activities - scoping, collecting, analysing, modelling, reporting, implementing, tracking and overseeing.



CapProcess

Processes that work!

Philip Bailey
Director and Principal Consultant

Email: pbailey@capprocess.com
Phone: +61 4 13 487 701 (or 0413 487 701)

Introduction

The capacity management knowledge domain manages IT capital resources to minimise IT capital spend. It typically lacks the staff to manage operational expenses, such as the purchasing of consumable items (such as tapes, paper and toner). Operational expenses (consumables) can be cost-effectively managed by using economic order quantity (EOQ) formulas through operational procedures, rather than using capacity planning approaches (expensive staff, specialised tools and catering for the significant impacts caused by projects).

Capital spend influences other costs, such as depreciation, software costs and maintenance. Capacity management includes activities to make best use of capital resources through tuning and translating the business requirements and plans into the capital items required to support the business needs. Several processes are required. These processes can include raise performance change, produce performance assessment, produce performance report, produce capacity assessment, produce capacity plan and produce capacity report. (See our whitepaper “Capacity Management: Process Types”).

The key capacity planning process should be called the produce capacity plan process. It integrates the capacity planning activities into a purpose built process. A simple structure can document the produce capacity plan process as a flow (like the development waterfall methodology) – the reality is that the boundaries between the activities are fuzzy and loopbacks and re-work are frequent (so the activities operate more like an agile methodology). The simple standardised process structure also works for the other capacity planning and performance management processes. This approach can be extended to all of the other ITIL knowledge domain processes.



Divide the produce capacity plan process into several steps or key activities such as:

- scope – understand the context and stakeholder requirements/needs, purpose and expected outcomes
- collect – gather relevant data
- analyse – understand the environment, assess problems and build a model
- model – make predictions and assess possible solutions
- report – inform stakeholders and propose actions via output products and artefacts
- implement – take relevant actions (trigger and drive separate processes)
- track – track and investigate variances
- oversee (governance) – report to management and stakeholders; ensure the expected outcomes are achieved.

Segregating the process into discrete steps is only for visualisation and ease of documentation.

1. Scope

The intent of the scope step is to define the context and the stakeholder expectations with the intention that these are checked/confirmed as being met or not in the later Track step. It defines or confirms the process purpose or expected outcomes. The manager provides a vision of the work required similar to a project kick-off meeting, except the review occurs at the start of the produce capacity plan process cycle each year. The process execution should concentrate on the cost areas or areas that currently restrict/reduce organisational performance or increase risk. These areas are identified through engagement with the stakeholders. It should also conceptualise possible solutions and the required data inputs.

The step should cover the following:

- any scope changes such as:
 - new categories of devices
 - new applications/services
- any capital budget constraints
- any new information (changes in corporate plans and strategies)
- the time-line for producing the capacity plan, such as whether the due date has changed – communicate the start date and capacity plan delivery dates to stakeholders
- the assignment of tasks – who is undertaking the various tasks/activities?

The step should also consider the key components to include in the report and presentation. It will include the required information and its sources, the format (source, business unit, geography etc.) and other rules such as corporate standards for reporting (such as colours and chart types).

2. Collect

The data collection step should be based on the requirements identified in the scope step. In practice the data collection process overlaps with the analysis step. In fact all the steps overlap to varying degrees – depending on the maturity in each step and the process. (See the chevron figure in the introduction, which shows overlap between each step and its following step – such as Scope–Collect and Collect–Analyse.)

The data collection step matures over time and is affected by role allocation. This maturity is dependent on organisational size, e.g. a separate performance team or risks or issues that need to be managed. The initial focus is usually on obtaining system metrics, such as server utilization, memory paging rates and disk space usage. It leads to collecting other capacity measures such as server ratings (megahertz [MHz] and/or Standard Performance Evaluation Corporation [SPEC]) and disk space capacity, as utilization figures on their own are not much use for long term forecasting.

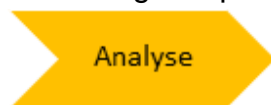
Other devices can be added later based on priorities such as cost, organisational performance impacts and risk.

As the capacity management processes mature, automated reporting is required to simplify, standardise, speed up the process and reduce costs.

The trends are only one piece of the forecast. The other key piece is the business plans: the additions, removals or changes to the environment. The capacity analyst must obtain various documents, such as the business strategy, business plans and where possible historical and planned business volumes, by engaging demand or business so that the forecast can reflect the variations.

3. Analyse

Producing a capacity plan is based on the collection of some workload data into the performance database (PDB) / capacity database (CDB) or somewhere else where it is available for processing.



The capacity analyst can use the results of any analysis exercise to build models of the environment based on the relationships identified. These relationships can be between resource, service and business measures or any combination.

A model represents some aspect of the real world. These days when the term “model” is used in capacity management, most people immediately think of queuing theory models, but the formula $Y = a + bx$ is also a model.

Model building

The key aspects to model building are:

- build – select representative data and create a model of some type:
 - trend
 - analytical
 - simulation
- calibrate/validate – test the accuracy using known data, such as behaviour at the extremities. Is it realistic / believable?

Test the model to verify its accuracy before it is used, for instance, by comparing known outcomes with the model predictions for the same event.

Workload Projections

The workload data is projected out into the future to the limit of the capacity plan to align it with the capital budget horizon. The initial forecasts usually use some form of regression (linear, quadratic, other). Linear trends assume that the frequency and increase in changes in the environment that occurred in the past will continue in the future, at least for the duration of the capacity plan.

4. Model

The initial workload growth predictions can use a straight line regression. If more information is available, such as business changes or changes to existing workloads or new workloads, make adjustments to the projection:



- either as the trend and adjustment of the resource variables
- or treat the trend and adjustment as input variables into the analytical and simulation models of the resources.

Incorporating significant planned changes into the model improves its accuracy and provides for better decision making.

Base the adjustments on the results of surveys for application, business or external changes that are significant enough to affect the projections. Adjustments are frequently due to new applications, but there may be other reasons, e.g. a change to a high-use program in an on-line environment, a significant marketing effort, or some change in government regulations. For marketing efforts or government regulations, business input is needed – this should not be abdicated to the capacity planner.

When assessing response time-based service level agreements (SLAs) through modelling (both analytical and simulation), capture some response time data from the various applications. Unfortunately, building a response time capture mechanism into application design is rare.

Model execution

The key aspects to modelling are:

- model – match the workloads that are being forecast to the modelled workloads, to predict resource requirements (a mismatch will entail a lot of work to decipher/split and rebuild)
- sensitivity analysis – how do small changes in the input variables affect the predictions? Where does the model break? If it is well beyond the current loads there is no need for concern
- what if – how long do the various options last? The most cost effective option? A different growth rate than the expected growth rate?

The time available may restrict the modelling. Of all the scenarios, you should test a few – how much time do you have? Restrict the number of what ifs/scenarios modelled to the most likely or highest-impact scenarios.

The modelling provides a figure, or set of figures, for the capital budget estimates.

5. Report

The capacity plan and reports are documents or slides and should include charts and tables as justification for the included commentary. The format ultimately depends on the organisation's standards (if they exist) and audience preferences, so discuss the report with the audience and take note of their preferences. The capacity plan and presentation should be tailored to focus on the needs of the audience. Focus on services for IT management and focus on business for an audience of business unit staff, rather than focusing on resources. The preferred structure should be business (decision) – services – resources – cost including discussion on options and risks as appropriate.

Reports should cover the following:

- what equipment is required in the planning horizon timeframe? (the difference between current state and planned state)
- when is it required?
- why is it required?
- how much will it cost?
- how long will it last?
- is it cost effective?
- a recommendation, based on the answers to these questions
- and as the report and process matures, a new section – accuracy of the previous forecasts (validation against the previous plan and lessons learnt).

6. (Implement)

Implementing the recommendations from the capacity plan requires obtaining the approval of the various key stakeholders that control the organisation's budget. Capacity plan recommendations rarely initiate a single implementation project. Prioritised recommendations are initiated over time as the management team determines the need. Each recommendation may be treated as a single project with an assigned project leader or manager coordinating the necessary activities of the vendor and the technical staff. Project initiation is a whole new process and is separate from the capital budget cycle.



The amount of confidence that the organisation has in the capacity plan (and the capacity planner) will affect whether the recommendations will be funded and later initiated.

Sharp and McDermott (2009) determined that process boundaries occurred where cycles differed, such as when the relationship was multiple executions to one execution (M:1 or 1:M). A one to one (1:1) relationship would indicate that there was probably a single process. For example, the produce capacity plan process and other business units (M) provide inputs into the capital budget process (1) preceding the acquisition process at different times throughout the year (M) and subsequent installation for use (M).

7. Track

Keep the resulting forecasts in some form of accessible library. Store the forecasts from the modelling steps in a dataset and merge them with the actual values on a monthly basis so that the key stakeholders can monitor the forecast accuracy. Note: The comparison of actual values and the previous forecast is called 'testing the capacity plan'. (ISO IEC 20000-2:2012 Sect. 4.5.5.2 (e) 'Test of plans, e.g. service continuity tests').



Normal practice should include predicting future events using the model and reviewing the result afterwards (and should lead to regular refinement/revision of the model). It should also form part of the regular monthly reporting process.

Monthly reporting

Review the report to identify significant changes:

- a unit of work behaving abnormally
- the processing of new work
- the peak has changed (different time, larger or smaller)
- produce a report (commentary) after any significant event.

These significant changes in the environment should be discussed with the performance team or systems administrators when identified.

Report significant events including:

- peak processing periods (especially any new records/highs)
- new applications including the effects of transaction volumes, resource usage (cpu, disk [space and IOPS]), response times (if measured)
- changes in technology (improvement and degradation) – resource usage (cpu, disk [space and IOPS]), response times (if measured)

These should be included in reports to IT management.

8. **Oversee**

Some form of oversight is required to ensure that the processes perform to expectation and outcomes are achieved, i.e. that the participants are meeting their commitments in terms of quality, timeliness, effort and cost. This is called governance. It creates a feedback loop between actual outcomes and planned outcomes as defined in the earlier Scope step.

There are at least two levels of governance – the capacity governance (reporting to the capacity process owner) and service management governance (reporting to the service manager/SLM manager). These should report through to the organisation's corporate governance process.

One aspect of governance is to conduct reviews of the process to identify potential improvements in the process. These improvements should include process performance, process maturity and process effectiveness. The review should determine whether the process is failing – and if so where is it failing? Is it due to poor or unclear procedural instructions, poor process definitions/descriptions, a lack of skills, an insufficient number of staff, no motivation, no co-ordination, no funds, no tools or inappropriate tools? Follow-on questions are: Who is responsible? Are the roles and responsibilities allocated effectively? Are teams working appropriately together? Who should be trained or re-trained? Are the capacity planning manager, capacity planning staff, produce capacity plan process, infrastructure staff, or others such as the business units not providing forecasts or providing incorrect forecasts? Has the message of the reports been communicated properly?

Most organisations focus on the middle process steps (Collect through to Report) they often ignore the Scope and Track steps. It is the Scope and Track step interactions that review whether the expected/required outcomes

are being achieved, which determines whether any changes need to be made to the process to improve its performance.



Operating governance properly requires a regular feed of key performance indicators (KPIs) to see how the various aspects of the process are working (as well as the other input processes). The ongoing monitoring of the process should lead to continual service improvement (CSI) activities to regularly realign the actual outcomes with the expected outcomes.

West (2004) takes a quality management approach and suggests that an ‘across the board’ approach to maturity is a waste of effort – only mature those elements that improve the support for the business. As such, any maturity review of capacity management should report on the maturity required in the various activities as well as the achieved maturity, such as data scope, collect, analyse, model, report, implement, track, oversight.

The service management governance or oversight process is the SLM process reporting on the performance of the process domains – as well as the overall performance.

Conclusion

Structuring the produce capacity plan process in a standardised way with a common set of key activities should improve both the process documentation and our understanding of the ITIL processes. Identify missing activities that potentially seriously affect the process by comparing the current process documentation with the steps in the process model template. Many process implementations do not define or describe the Scope and Track activities, but these are probably the two most important activities in any process.

Structure the other capacity-related processes (produce performance report, raise performance change, produce performance assessment, produce capacity report and produce capacity assessment) in a similar way with similar activities. The other ITIL processes can also be structured with these common activities. They will then be documented, managed and governed better because the deficiencies will be more obvious due to the flows and linkages.

Bibliography

[Hoyle, D. (2009)] *ISO 9000 Quality Systems Handbook: Using the Standards as a Framework for Business Improvement (Sixth Ed.)*, Butterworth-Heinemann, Oxford, UK

[ISO IEC 20000, (2012)] *ISO IEC 20000-2:2012 Information Technology – Service Management – Part 2 – Guidance on the application of service management systems (second edition)*, ISO copyright office, Geneva, Switzerland

[Protzen, J-P. and D.J. Harris (2010)] *The Universe of Design: Horst Rittel's Theories of Design and Planning*, Routledge, London

[Sharp, A. and P. McDermott (2009)] *Workflow Modeling: Tools for Process Improvement and Application Development (Second Edition)*, Artech House, Norwood, Massachusetts

[West, M. (2004)] *Real Process Improvement Using the CMMI®*, Auerbach Publications, Boca Raton, Florida

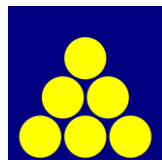
Philip Bailey



Philip Bailey is an expert in the field of IT capacity management. Phil has implemented, guided and improved the capacity management processes in several organisations (financial and outsourcing). In addition to 25 years of experience, he brings first-rate conceptual / process skills, a creative approach and proven leadership to the role as consultant.

Phil has established a capacity planning methodology at various organisations. He has also trained several capacity practitioners over the years. He is an ITIL V3 Expert and ISO 20000 Consultant certified. He has written articles and spoken extensively about the methodology he developed relating to the IT capacity management process and its underlying activities.

Phil has a habit for acquiring and reading non-fiction books.



CapProcess

Processes that work!

CapProcess is a company that provides expert advice in the capacity management process domain. Expertise that includes how to improve capacity management activities so that IT service improves and costs are effectively managed.

Our purpose is to transfer our years of experience to your staff are more self-sufficient and provide the support that your customers expect.

CapProcess is based in Sydney, Australia.

To find out more about how CapProcess expertise can help your organisation contact us on +61 4 13 487 701 or send an email to info@capprocess.com or visit our website at www.capprocess.com