CHAPTER 3 / 4 MINUTE BRIEF

The Players and the Vocabulary

Read the full chapter → / Read this brief online →

The main idea

As an engineer, I can learn the difference between a capital expense and an operating expense, and I can put together a comparison that explains why changing one into the other looks attractive. Then I talk to someone in Finance and discover how much of the question I have left out. The timing matters, the planning process matters, the person who can approve the commitment matters. "It depends" begins to sound less like an obstacle and more like the answer I should have expected.[1]

Explore this section in the full chapter →

The nouns that change a meeting

A **forecast** estimates future demand under stated assumptions. Give it an as-of date, because the information available at the time is part of what makes the estimate interpretable. Revising it after a customer changes plans may be exactly the right thing to do; the revision is a new forecast, not evidence that the old one was accurate. Check the unit, too: in a Finance meeting "the forecast" usually means expected spending for the year, and in a capacity meeting, demand in requests, cores, or tokens.

A **target** is something the organization intends to achieve. A sales target may be deliberately ambitious to encourage people to pursue an opportunity, while the most likely estimate of realized demand is lower. Capacity needs to examine the target as a scenario without silently converting ambition into a prediction.

A **budget** is an approved plan for spending: an amount, for a period, held by someone accountable for it, usually set months before the demand it funds. It doesn't approve a particular purchase. Who can approve that kind of spending, and whether the amount

[1] Tim O'Brien, "It's Not That Simple", Medium, June 21, 2026. The opening passage adapts the author's description of recurring conversations with Finance.

crosses a threshold that sends it through a different process, are separate questions, and often what Finance means by "it depends." Whether the money will support the service's obligations is an engineering and business question, and a discrepancy between the two needs a conversation rather than a revised label.

A **request** says what someone wants and starts an investigation or a negotiation. It doesn't establish that the resources exist, that funding has been approved, or that another customer will give up an allocation, which sounds obvious until a requested amount starts appearing in other people's plans as confirmed supply. Requests also lose their reasons on the way: the SRE book's account of Auxon, a Google planning tool, describes requests arriving as "X cores in cluster Y," their degrees of freedom "long lost."[2] In our launch, the degrees of freedom were most of the answer. A quota increase says no more about supply: it raises the limit on what we may request without putting any machines behind it.

A **reservation** needs to be understood in its particular context. It might be a provider product, an internal hold, or a position in a planning queue, and a provider's reservation may buy a discount, a hold on capacity, or both. I want to know what has been reserved, for whom, where, for which dates, and with what release rule before I use it to support a customer promise.

A **commitment** is a promise with defined parties and conditions. In Volume II, Chapter 20 we will separate a promise to spend from a promise to supply usable capacity, because each can exist without the other. An optimistic forecast in a presentation establishes neither.

One more noun is not a planning word but a product word, and meetings trip over it as often as the others: the **SKU**. The term is stock-keeping unit, borrowed from retail, where it names one distinct item a store sells — one product, in one configuration, that someone can order by a code. The cloud providers borrowed it for their catalogs, and in this book a SKU is one sellable configuration of a service: one size, one feature set, one price. Every line on the bill in Chapter 1 names a SKU.

[2] Beyer et al., *Site Reliability Engineering*, Chapter 18, "Software Engineering in SRE", quoted from the Auxon case study. Auxon was an internal Google tool; the connection to this chapter's launch is the author's.

The name is worth reading slowly, because providers encode the machine into it. In Microsoft's convention, `Standard_D4s_v5` decodes as a general-purpose family (D), four vCPUs, premium-storage capable (the s), fifth version (v5), and the encoded extras run to local disks, network bandwidth, memory ratios, processors, and, for the accelerated families, the accelerator model itself: the GPU size Chapter 10 works with, `Standard_ND96isr_H200_v5`, announces the H200 in its own name.[3] Other providers publish naming conventions of their own, and the skill of reading them transfers even where the grammar differs.

Explore this section in the full chapter →

Who can change the promise?

For each condition the plan relies on, I want to know who supplied it, who can check it, who can change it, and who else needs to hear when it changes. For our constructed launch, a first version might look like this:

Condition	Supplied by	Checked by	Can change it	Must also hear
First stage adds 2,000 requests/s	Product lead	Engineering and operations	Product and the business owner	Operations, Finance
Batch work can wait while eight workers move	Batch owner	Batch owner, against a catch-up plan	Batch owner	Product, operations
Moved workers ready after two days' testing	Operations	Operations, by testing	Operations	Product
Sixteen new workers in about six weeks	Supplier, via Procurement	Operations, at qualification	Supplier and operations	Product, batch owner
Money for the new workers	Budget owner	Finance	Budget owner, within a delegated limit	Product, Procurement
Service survives loss of one failure group	Business owner	SRE, by testing	Responsible business executive	Everyone relying on it

Explore this section in the full chapter →

[3] Microsoft Learn, "Azure virtual machine sizes naming conventions", checked October 6, 2026. Field order (family, sub-family, vCPUs, additive features, accelerator, version), the additive-feature letters, and the rule that only specialized SKUs launched from the third quarter of 2020 carry the accelerator in the name. The `Standard_D4s_v5` reading follows the page's conventions and examples.