CHAPTER 16 / 3 MINUTE BRIEF

A Portfolio of Customers

Read the full chapter → / Read this brief online →

The main idea

One of the choices a capacity team has to make is how much work to ask of its customers. A major project that consumes a substantial part of a constrained pool may need recurring conversations with both the engineering team and the people who understand the business, and asking those questions is part of supporting it. Asking the same of every application on a shared platform would spend everyone's time, theirs and ours, on conversations that change nothing.

For a thousand small consumers, I may get much further with a platform owner who understands adoption, can change common defaults, and knows how to reach the members when something unusual happens. Those customers don't need to become capacity planners; the team providing the service may be best placed to understand what they will ask of it.

An experiment needs something else again: close attention for a while, as the organization learns what the system does and decides what it might become. I wouldn't describe that as less mature than a stable forecasting relationship; it suits a different kind of uncertainty. Specialist attention is the scarce resource in all three cases, and this chapter's process spends it deliberately, deciding who gets an individual forecast, who shares a group forecast, and what moves a consumer between them.

Explore this section in the full chapter →

3. Assign a planning treatment

Treatment	What the team maintains	Relationship needed
Individually managed projects	Workload model, business drivers, scenarios, commitments, and forecast revisions	Named technical and business owners
Cohorts of smaller projects	Membership, aggregate baseline, variation, shared drivers, and project-level exceptions	A cohort owner who can reach member teams
Temporary exceptions	A time-bounded investigation of a new, changing, or consequential workload	The person who can resolve the specific uncertainty

Explore this section in the full chapter →

4. Gather intelligence at the appropriate level

- **Record:** event `E-ROLLOUT`, revision 1, known June 25, three days before the forecast it informs.
- **Change and source:** product proposes wider customer access; the product lead supplied the plan.
- **Boundary:** additional demand on the web worker pool, measured as requests per second at the existing request mix.
- **Alternatives:** a full rollout adds an assumed 4,000; a staged rollout adds 2,000. Timing and scope remain negotiable.
- **Overlap:** zero of this increment is assumed to be present in the historical baseline used for this decision.
- **Dependencies:** the rollout proceeds on the stated schedule, the measured request mix remains representative, and the transferred capacity is qualified before expansion.
- **Review and expiry:** review by June 30; no forecast made after July 7 may use this revision.

Explore this section in the full chapter →

6. Promote, investigate, and release

A project or cohort needs closer attention when its growth, forecast error, contribution, or consequence crosses the threshold we've agreed to use. A small workload can qualify before its spend is large, by growing quickly, supporting a critical transaction, or needing a resource with a long lead time, and a suspected shared driver can qualify a cohort before any member is large. Sometimes, in the monthly meetings, a very small project that nobody thinks is a big deal starts giving signals that it is bigger than even its own team thinks. It may not be asking for capacity today, but it is going to need it. In each case I'd want a person responsible for resolving the uncertainty and a deadline tied to the decision, rather than an indefinite designation that the project is being watched.

Explore this section in the full chapter →

There also needs to be a way out of the special treatment. A system can stabilize enough to join an existing cohort, or an investigation can resolve the uncertainty that justified individual attention. We should agree what would bring it back, then release the extra effort. I'd set the release condition below the promotion trigger and require it to hold for two consecutive reviews, so a consumer near the line doesn't change treatment every month, and I'd return its history to the cohort baseline and its open events to the cohort owner. Otherwise every temporary exception becomes permanent, and the central team's limited attention goes to maintaining arrangements nobody has reconsidered.

Explore this section in the full chapter →