

PART 3

USAGE-BASED / CONSUMPTION SAAS

When Revenue Follows Consumption

Committed vs. consumed ARR, variable consideration under ASC 606, consumption forecasting and ramp curves, infrastructure cost variability, gross margin management, overage economics, NRR in consumption models, take-or-pay accounting, digital services taxes, and the complete usage-based metrics framework.

SECTION 1

THE CONSUMPTION MODEL — A DIFFERENT KIND OF SAAS

Usage-Based SaaS: When Revenue Follows Consumption

Usage-based pricing — also called consumption-based pricing or pay-as-you-go — is the fastest-growing pricing model in enterprise software. Companies like Snowflake, Twilio, AWS, Stripe, Datadog, and MongoDB have demonstrated that aligning price with value delivered, measured in units of consumption, can produce extraordinary growth, high customer satisfaction, and superior net revenue retention. But the financial architecture of a usage-based business is fundamentally different from the seat-based SaaS model described in Part 2, and the differences matter enormously for how the CFO builds models, recognizes revenue, manages costs, and communicates with investors.

In seat-based SaaS, the customer commits to a fixed number of seats and pays a predictable monthly or annual fee regardless of how much they use the software. Revenue is highly predictable. In usage-based SaaS, the customer pays for what they actually consume — API calls, data processed, messages sent, compute hours used, rows queried. Revenue is inherently variable. It rises when customers use the product more and falls when they use it less. This variability is both the greatest commercial advantage of the model — customers feel safe expanding because they only pay for what they use — and the greatest financial challenge — the CFO cannot predict next month's revenue with the same confidence that a seat-based SaaS CFO can.

This part of the series covers the complete financial architecture of usage-based SaaS: pricing structure design, the distinction between committed and consumed ARR, revenue recognition under ASC 606 for variable consideration, the infrastructure cost variability problem, how to model and forecast consumption revenue, gross margin volatility, and the specialized metrics framework that usage-based CFOs must own. Every concept is grounded in practice, with formulas, benchmarks, and worked examples drawn from the leading consumption SaaS companies.

1.1 The Spectrum of Usage-Based Pricing Models

Usage-based pricing is not a single model — it is a spectrum. At one end is pure pay-as-you-go: no commitment, no minimum, customer pays exactly for what they consume in the billing period. At the other end is a committed spend model with a consumption mechanism: the customer commits to a minimum

annual spend (providing revenue predictability) and then draws down against that commitment based on actual usage. Between these poles sit hybrid models that combine elements of both.

Model Variant	Structure	Revenue Predictability	Customer Appeal	Example
Pure Pay-As-You-Go	No commitment; pay per unit consumed	Very Low	Lowest barrier to entry	Twilio SMS, AWS Lambda
Committed Spend / Drawdown	Annual \$ commitment; draw down by usage	High	Volume discounts; budget certainty	Snowflake, Google Cloud
Seat + Usage Hybrid	Base seat fee + overage for excess usage	Medium-High	Predictable base; flexible upside	Datadog, Zendesk
Tiered Consumption	Price per unit drops at volume thresholds	Medium	Incentivizes growth; volume economics	Stripe, SendGrid
Outcome-Based	Fee tied to business outcome achieved	Low	Risk-sharing; high alignment	Early-stage niche models

CFO INSIGHT

The choice of pricing model is one of the highest-stakes financial architecture decisions a SaaS CFO will make, and it is rarely revisited without significant disruption. Pure pay-as-you-go maximizes customer adoption but makes revenue forecasting extremely difficult and can produce revenue that falls sharply when customers reduce usage during downturns. Committed spend models provide revenue floor visibility but require more sophisticated customer success and renewal management. The CFO should model the revenue predictability, gross margin, and NRR implications of each variant before the company standardizes on a pricing architecture.

SECTION 2

COMMITTED ARR VS. CONSUMED ARR

The Two ARRs: Committed vs. Consumed

The most important conceptual framework for the usage-based SaaS CFO is the distinction between **Committed ARR** and **Consumed ARR** — and the gap between them. This gap is invisible in a seat-based SaaS model because what the customer commits to is exactly what they pay. In a usage-based model, the

commitment and the consumption can diverge significantly, and understanding that divergence is essential for accurate forecasting, investor communication, and customer success strategy.

2.1 Committed ARR

Committed ARR (sometimes called contracted ARR or bookings ARR) represents the annualized value of what customers have *agreed to spend* — the minimum committed drawdown in their contract, plus any baseline platform fees. It is the revenue floor that the company can count on regardless of whether customers hit their consumption targets. It is analogous to the seat-based SaaS ARR metric: a contractual guarantee.

However, committed ARR in a pure pay-as-you-go model is zero — there is no commitment, only consumption. In that case, the company has no ARR in the traditional sense. Some companies in this position define ARR as trailing 12-month revenue annualized, or as trailing 3-month revenue multiplied by 4, or as the prior month's revenue multiplied by 12. None of these is as clean or predictive as true contracted ARR. The CFO must be transparent about the definition used and its limitations.

COMMITTED ARR DEFINITION

Committed ARR = Annualized Value of Minimum Committed Contract Spend
+ Annualized Platform / Base Fees
 (excludes any overage or variable consumption above commitment)

For pure pay-as-you-go (no commitments):

Proxy ARR = Trailing 3-Month Revenue x 4 (disclose the definition clearly)
 or **Last Month Revenue x 12 (more volatile; use with caution)**

2.2 Consumed ARR

Consumed ARR (sometimes called run-rate revenue or consumption ARR) represents the annualized value of what customers are *actually consuming* at the current run rate. It is calculated by taking the most recent period's consumption revenue and annualizing it. When customers are using more than their committed amount, consumed ARR exceeds committed ARR — a positive signal. When customers are using less than their committed amount, consumed ARR is below committed ARR — a warning signal about future renewals.

CONSUMED ARR AND THE COMMITMENT GAP

Consumed ARR = Trailing 1-Month Consumption Revenue x 12

Commitment Gap = Committed ARR - Consumed ARR

If Consumed ARR > Committed ARR: Customer is over-consuming (positive)

-> Expansion opportunity at renewal; upsell higher commitment tier

If Consumed ARR < Committed ARR: Customer is under-consuming (risk signal)

-> Customer may not renew at full commitment; churn risk elevated

Commitment Utilization Rate = Consumed ARR / Committed ARR

Target: >85% utilization signals healthy renewal and expansion prospects

CFO INSIGHT

The Commitment Utilization Rate is the early warning metric that seat-based SaaS CFOs do not have and usage-based CFOs cannot live without. A customer with 60% utilization at month 9 of a 12-month committed contract is almost certain to either reduce their commitment at renewal or churn. The customer success team needs this signal at month 6, not month 11. Build a monthly commitment utilization report by customer and by cohort and use it as the primary input to renewal risk forecasting.

2.3 The Overage Economics

When customers consume more than their committed amount, they incur **overage** charges — additional usage billed at the per-unit rate (which may be the same as the committed rate or a higher rate for above-commitment consumption). Overage revenue is the most attractive revenue in a usage-based business: it requires no incremental sales cost, it signals deep customer engagement, and it often carries higher per-unit pricing than committed consumption.

However, overage revenue is also the most volatile. A customer who consumed 120% of commitment in Q4 due to a one-time data migration project will not reproduce that overage in Q1. If the CFO models overage as a recurring component of revenue without understanding the one-time drivers, the forecast will be wrong. Build a detailed overage analysis that categorizes overage by driver (structural growth vs. one-time event) and model each category separately.

OVERAGE REVENUE ANALYSIS

Total Revenue = Committed Revenue Recognized + Overage Revenue

Overage Rate = Overage Revenue / Total Revenue

Structural Overage: driven by customer business growth (model as recurring)

One-time Overage: driven by specific events (model as non-recurring)

Overage Gross Margin is often higher than committed margin because:

- No incremental sales cost (already in the account)
- Infrastructure cost per unit often lower at higher consumption volumes

SECTION 3

REVENUE RECOGNITION FOR USAGE-BASED MODELS

Revenue Recognition: Variable Consideration and ASC 606

Revenue recognition in a usage-based SaaS business is more complex than in a seat-based SaaS business, and the complexity stems from one source: **variable consideration**. Under ASC 606, when the amount of revenue the company will earn from a contract depends on something variable — the customer's consumption — the company must estimate that variable consideration and include it in the transaction price, subject to the constraint that recognized revenue cannot be subsequently reversed in a significant amount (the constraint on variable consideration).

3.1 The Variable Consideration Framework

ASC 606 provides two methods for estimating variable consideration: the **expected value method** (probability-weighted average of all possible outcomes) and the **most likely amount method** (the single most likely outcome). For usage-based SaaS, the expected value method is generally more appropriate because there is a range of possible consumption outcomes for each customer, each with its own probability.

Critically, the estimated variable consideration can only be included in recognized revenue if it is *highly probable* that a significant revenue reversal will not occur when the uncertainty resolves. In practice, this

constraint means that companies often recognize usage-based revenue only as the usage occurs — the most conservative and most common approach — rather than estimating and front-loading the full contract value. For a 12-month committed contract with variable consumption, this means recognizing the committed minimum ratably plus recognizing overage as it is consumed.

USAGE-BASED REVENUE RECOGNITION

Monthly Recognized Revenue = (Committed Amount / 12) + Overage Consumed in Month

For pure pay-as-you-go (no commitment):

Monthly Revenue = Units Consumed x Price per Unit

Recognition: at the point of consumption (when the service is delivered)

For committed + overage hybrid:

Committed portion: recognized ratably (like seat-based SaaS)

Overage portion: recognized as consumed (point-in-time each billing cycle)

ACCOUNTING ALERT

A common error in usage-based SaaS revenue recognition is treating the entire committed contract value as fixed and recognizing it ratably, then booking overage as a separate line item. This is correct for the committed portion, but the overage treatment depends on whether the contract structure creates a variable consideration arrangement or a separate performance obligation for usage above the minimum. Have your external auditors review the contract structure before establishing the recognition policy. Getting this wrong can require a restatement.

3.2 The Deferred Revenue Difference

In a seat-based SaaS model, the deferred revenue balance is relatively predictable — customers pay in advance and the company recognizes ratably. In a usage-based model, the deferred revenue dynamics are more complex. For committed contracts billed annually in advance, the company collects the committed amount upfront and creates a deferred revenue liability. But as usage is consumed, the recognized revenue may be faster or slower than the ratable amortization of the commitment, depending on the consumption pattern.

If a customer commits to \$120,000 per year and the billing is annual in advance, the company collects \$120,000 and defers it. If the customer's consumption is front-loaded — they use \$80,000 of the commitment in the first half of the year — the company recognizes \$10,000 per month in committed revenue (ratable) plus additional overage revenue in the high-consumption months. If consumption is back-loaded,

recognized revenue may trail the cash collection significantly in early months. The CFO must model the consumption curve, not just the commitment, to produce accurate revenue forecasts.

DEFERRED REVENUE WITH VARIABLE CONSUMPTION

Beginning Deferred Rev + Annual Billing - Revenue Recognized = Ending Deferred Rev

Example (front-loaded consumption customer):

Month 1: \$10K committed recognized + \$8K overage = \$18K revenue recognized

Month 12: \$10K committed recognized + \$0 overage = \$10K revenue recognized

Total = \$120K committed + \$48K overage = \$168K total revenue on \$120K commitment

SECTION 4

FORECASTING CONSUMPTION REVENUE

The Consumption Forecast: The CFO's Greatest Challenge

Forecasting revenue in a usage-based business is the hardest financial planning problem in enterprise software. Unlike seat-based SaaS, where next quarter's revenue is largely determined by the current deferred revenue balance and the renewal schedule, consumption revenue depends on how much each customer chooses to use the product — a behavioral variable that is influenced by the customer's own business activity, competitive alternatives, product quality, and external macro conditions. The CFO who builds a rigorous consumption forecast model has a significant advantage in capital allocation, investor credibility, and operational planning.

4.1 The Three-Layer Forecast Model

The most effective consumption revenue forecast is built in three layers: the **committed layer** (revenue locked in by contract, highly predictable), the **historical consumption layer** (revenue from existing customers based on trailing consumption patterns, medium predictability), and the **new consumption layer** (revenue from new customers and from expansion in existing accounts, lower predictability). Each layer requires different modeling inputs and carries different confidence intervals.

THREE-LAYER CONSUMPTION FORECAST

Layer 1 – Committed Revenue:

= Sum of (Monthly Committed Amount) across all active contracts

Confidence: High (90%+); only at risk from early termination clauses

Layer 2 – Historical Run-Rate Consumption:

= Trailing 3-month avg consumption x Consumption Growth Rate

Consumption Growth Rate = derived from same-customer cohort analysis

Confidence: Medium (70%–85%); depends on customer business stability

Layer 3 – New Customer + Expansion Consumption:

= New Logo Ramp Curve + Expected Expansion in Existing Accounts

Confidence: Lower (50%–70%); pipeline-dependent

4.2 The Ramp Curve: New Customer Consumption Modeling

New customers in a usage-based model do not immediately consume at their full potential. They ramp up over a period of weeks to months as they onboard, integrate the product into their workflows, and train their teams. The **ramp curve** — the trajectory from zero to steady-state consumption — is a critical input to the consumption forecast. It varies by customer size, product complexity, and industry.

Customer Segment	Month 1	Month 2	Month 3	Month 6	Steady State
SMB (self-serve)	30%–40%	55%–65%	70%–80%	85%–95%	Month 4–5
Mid-Market (guided)	15%–25%	35%–50%	55%–70%	80%–90%	Month 5–7
Enterprise (managed)	5%–15%	20%–35%	35%–55%	65%–80%	Month 8–12

The ramp curve has direct implications for how the company calculates and presents ARR. If a new enterprise customer signs a \$500,000 committed contract in month 1 and is only consuming at 10% of their commitment, the ARR waterfall correctly shows \$500,000 in new committed ARR — but the revenue recognized in month 1 is only \$41,667 (1/12 of the annual commitment). The gap between ARR bookings and near-term revenue recognition is a structural feature of usage-based enterprise SaaS that the CFO must explain clearly to investors.

4.3 Seasonal Consumption Patterns

Consumption-based revenue is significantly more seasonal than seat-based SaaS revenue. Many usage-based products are tied to their customers' own transaction volumes — an API call per customer purchase, a row of data per user interaction, a message per customer service ticket. When the customer's business slows down (holiday periods, budget cycles, off-peak seasons), consumption drops and so does revenue. The CFO must model these seasonal patterns explicitly and present seasonality-adjusted forecasts to avoid misleading QoQ comparisons.

CFO INSIGHT

Build a consumption seasonality index for your top 20 customers by comparing each month's consumption to the full-year average. If December consistently comes in at 75% of the average month and March at 115%, these indices should be applied to the forecast model and communicated to the board as a structural feature of the revenue stream — not as a variance to be explained after the fact. Investors in usage-based companies expect seasonality; what they cannot tolerate is being surprised by it.

SECTION 5

GROSS MARGIN VOLATILITY AND INFRASTRUCTURE COST

Infrastructure Costs: The Variable Cost Problem

The most structurally distinctive feature of usage-based SaaS — from a cost perspective — is that infrastructure costs vary with revenue. In seat-based SaaS, infrastructure costs are largely fixed regardless of how much customers use the software, because the licensing model does not create strong usage incentives. In usage-based SaaS, customers are directly incentivized to use more, and when they do, the company's compute, storage, network, and data processing costs rise correspondingly. This creates a **variable cost structure** that makes gross margin management far more complex.

5.1 The Infrastructure Cost Stack

In a usage-based business, the infrastructure cost stack has both fixed and variable components. The fixed component includes baseline compute (servers that must be running regardless of usage to ensure availability), network connectivity, monitoring and observability tooling, and engineering infrastructure. The variable component — the part that scales with customer consumption — includes compute for actual workload processing, storage for data generated by consumption, egress costs (data transferred out of the cloud), and third-party API costs that are themselves usage-based.

GROSS MARGIN IN USAGE-BASED SAAS

Revenue = Committed + Overage consumption in period
 Variable Infrastructure = Units Consumed x Infrastructure Cost per Unit
 Fixed Infrastructure = Base compute, monitoring, engineering tools
 COGS = Variable Infra + Fixed Infra + CS + Support
 Gross Profit = Revenue - COGS

Key ratio: Infrastructure Cost per Unit of Consumption
 Must decline over time as architecture improves and volume scales

Usage-Based Company	Gross Margin Range	Key COGS Driver	GM Trend
API / Communications (Twilio-type)	45%–58%	Carrier / telecom pass-through costs	Relatively stable; carrier negotiation drives improvement
Data / Analytics (Snowflake-type)	60%–72%	Cloud compute for query processing	Improving as architecture optimization scales
AI / ML Inference (OpenAI-type)	20%–50%	GPU compute — very high per inference	Improving rapidly as model efficiency scales
Payments / FinTech (Stripe-type)	40%–55%	Interchange and payment network fees	Stable; interchange negotiation at scale
Storage / CDN (Cloudflare-type)	65%–78%	Network infrastructure (largely fixed)	Strong leverage at scale; fixed cost amortization

BENCHMARK

Usage-based SaaS companies structurally carry lower gross margins than seat-based SaaS companies because of the variable infrastructure cost component. Investors understand this and apply appropriate valuation adjustments — but they expect to see gross margin improvement over time as the company optimizes its infrastructure architecture and negotiates better rates with cloud providers. A usage-based SaaS company with gross margins below 50% that cannot articulate a credible path to 60%+ within 24 months will face investor skepticism at any valuation.

5.2 Managing the Infrastructure Cost per Unit

The single most important operational metric for a usage-based SaaS CFO — from a cost perspective — is the **infrastructure cost per unit of consumption**. This metric must decline over time as the company

scales, because the revenue per unit of consumption is fixed (or declining due to volume discounts), and the only way to expand gross margins is to reduce the cost per unit. There are four levers available.

Optimization Lever	Description	Typical Impact	Timeline
Architecture optimization	Redesign processing pipelines for efficiency; reduce redundant compute steps	10%–30% cost reduction	6–18 months engineering effort
Reserved capacity / CUDs	Commit to cloud provider capacity in exchange for volume discounts	20%–40% vs. on-demand rates	Immediate; 1–3 year commitment
In-house vs. cloud balance	Move stable, predictable workloads to owned or colocation hardware	15%–30% for qualifying workloads	12–24 months capital and ops investment
Supplier renegotiation	Renegotiate carrier, API, and third-party data costs at volume milestones	5%–20% on third-party components	Ongoing; milestone-triggered

SECTION 6

UNIT ECONOMICS IN USAGE-BASED SAAS

Unit Economics: Redefining CAC, LTV, and NRR for Consumption Models

The unit economics framework developed in Part 2 for seat-based SaaS requires significant adaptation for usage-based SaaS. The fundamental challenge is that customer value — and therefore LTV — is not fixed at the time of sale. It grows as the customer consumes more, contracts as they consume less, and is shaped by the ramp curve and the expansion behavior of each customer cohort. This dynamic makes usage-based LTV a range, not a point estimate, and requires a probabilistic approach to unit economics modeling.

6.1 LTV in a Consumption Model

LTV in usage-based SaaS has two components: the **committed LTV** (the present value of minimum committed revenue over the expected customer relationship) and the **consumption LTV** (the present value of expected overage and expansion revenue above the commitment). Together these form the total LTV, but they have very different confidence levels and different implications for how aggressively the company should invest in acquisition.

USAGE-BASED LTV FORMULA

Committed LTV = Annual Committed ARR x GM% / Annual Churn Rate
 Consumption LTV = Annual Avg Overage Rev x GM% / Annual Churn Rate
 Total LTV = Committed LTV + Consumption LTV

Example:

\$50K committed ARR, \$15K avg overage, 72% GM, 10% churn

Committed LTV = (\$50K x 0.72) / 0.10 = \$360,000

Consumption LTV = (\$15K x 0.72) / 0.10 = \$108,000

Total LTV = \$468,000

6.2 NRR in Usage-Based SaaS: The Expansion Flywheel

Net Revenue Retention in usage-based SaaS is structurally higher than in seat-based SaaS — and this is one of the model's greatest financial advantages. Because customers can expand their consumption without a separate sales motion, and because the product is architected to deliver more value as usage grows, well-designed usage-based products generate NRR of 120% to 160% at maturity. Snowflake's NRR exceeded 170% at peak; Twilio sustained NRR above 130% for multiple years.

However, NRR in usage-based SaaS is also more volatile than in seat-based SaaS. During economic downturns, customers can reduce consumption immediately — no waiting for contract expiry — causing NRR to compress sharply. Snowflake's NRR declined from over 170% in 2022 to approximately 128% in 2024 as enterprise customers optimized their cloud spending. The CFO must model NRR under stress scenarios and ensure that the company's growth assumptions are not dangerously dependent on continued NRR expansion.

USAGE-BASED NRR CALCULATION

$$\text{NRR} = (\text{Beginning ARR} + \text{Expansion Consumption} + \text{Overage Revenue} - \text{Consumption Decline} - \text{Churn}) / \text{Beginning ARR}$$
Best-in-class usage-based NRR benchmarks:

>150%: Exceptional (Snowflake at peak, peak-era Twilio)

120%-150%: Excellent (top-quartile usage-based SaaS)

100%-120%: Healthy (solid usage-based model)

<100%: Concerning for a usage-based model (structural issue signal)

SECTION 7

TAX ISSUES FOR USAGE-BASED SAAS

Tax Architecture: Usage-Based Complexity

Usage-based SaaS companies face all of the tax challenges described in Part 2 for seat-based SaaS — sales tax nexus, R&D; credits, Section 174 amortization — plus several additional layers of complexity that arise specifically from the consumption model: the tax treatment of committed vs. variable revenue, the international tax implications of cross-border data flows, and the sourcing rules for usage-based digital services across state and international jurisdictions.

7.1 Sales Tax Sourcing for Consumption-Based Services

In seat-based SaaS, sales tax sourcing is relatively straightforward: the subscription is typically sourced to the customer's billing address or place of use. In usage-based SaaS, sourcing is more complex because the consumption may occur in multiple locations. A customer who uses an API processing service may route requests from data centers in multiple states or countries. The question of where the service is "used" — and therefore which state has taxing jurisdiction — is genuinely unsettled in many jurisdictions.

Most states that tax SaaS services apply a "place of use" or "customer location" sourcing rule — the service is taxed where the customer's primary place of business is located, or where the benefit of the service is received. For a usage-based company with enterprise customers whose usage spans multiple locations, determining the correct sourcing allocation can require detailed analysis of the customer's usage data and may necessitate multi-state apportionment of the tax base.

TAX ALERT

Digital services taxes (DSTs) are expanding globally and increasingly target consumption-based digital services specifically. The EU's approach, Canada's DST, the UK's digital services tax, and India's equalization levy all apply to companies providing digital services to customers in those jurisdictions — often with no minimum physical presence requirement. A usage-based SaaS company with \$5M or more in annual revenue from any of these jurisdictions must evaluate its DST exposure immediately. The CFO should conduct a global DST exposure analysis annually as revenue scales internationally.

7.2 Transfer Pricing for Cross-Border Data Processing

Usage-based SaaS companies that process data across borders — routing workloads through cloud infrastructure in multiple countries — face complex transfer pricing questions. If the US entity performs the computational work that generates revenue billed to a UK customer, but the UK entity holds the customer relationship, what portion of the revenue should be allocated to each entity? OECD transfer pricing rules require that the allocation reflect the functions performed, risks assumed, and assets used by each entity — the DEMPE (Development, Enhancement, Maintenance, Protection, Exploitation) framework for intangibles.

For usage-based SaaS companies with significant cross-border revenue, building a defensible transfer pricing model requires understanding not just where customers are located, but where the value-creating activities occur: where the engineering team builds the product, where the data is stored and processed, and where the customer success and sales functions operate. Engage specialized transfer pricing counsel before the international revenue base exceeds \$10 million annually.

SECTION 8

ACCOUNTING ISSUES FOR USAGE-BASED SAAS

Accounting Architecture: Usage-Based Complexity

Beyond revenue recognition, usage-based SaaS companies face several additional accounting challenges that do not arise in the seat-based model: the treatment of credits and free consumption granted to customers, the accounting for minimum commitment shortfalls (when customers do not hit their commitments), and the capitalization of the platform infrastructure that is both a product development asset and a revenue-generating cost center.

8.1 Credits and Free Consumption

Usage-based SaaS companies frequently grant credits to customers — free consumption units given as sales incentives, as remedies for service outages, or as part of promotional programs. The accounting treatment of these credits is nuanced. Credits given as part of the original contract negotiation (discount on the first \$X of consumption) must be reflected in the transaction price at inception, reducing total recognized revenue. Credits given after contract inception as a remedy for poor service performance (service level credits) may be treated as variable consideration adjustments or as a reduction in revenue in the period the credit is issued, depending on the circumstances.

CREDIT ACCOUNTING TREATMENT

Pre-contract credits (incentive): Reduce transaction price at contract inception
 = Gross contract value - Credit Value = Adjusted transaction price

Post-contract credits (SLA remedy): Reduce revenue in period issued

If material: treat as variable consideration and estimate upfront

If immaterial: recognize as revenue reduction when credit is issued

Free tier credits (product-led growth): Expense as S&M (customer acquisition cost)

NOT a revenue reduction (no revenue has been recognized on free usage)

8.2 Minimum Commitment Shortfalls

When a customer commits to a minimum annual spend of \$120,000 but only consumes \$80,000 during the contract year, the customer has a shortfall of \$40,000. If the contract requires the customer to pay the minimum regardless of consumption (a true take-or-pay commitment), the company must recognize the full \$120,000 as revenue — \$80,000 from actual consumption plus \$40,000 from the minimum commitment shortfall. The timing and recognition of the shortfall revenue requires careful analysis of when the performance obligation is satisfied and when the customer's obligation to pay becomes unconditional.

If the contract allows the customer to roll over unused commitment to the next period rather than forfeiting it, the accounting is different — the unearned portion creates a liability (or a reduction in deferred revenue) that the company carries forward to the next period. This treatment is common in cloud provider committed spend programs and must be carefully tracked at the customer level.

ACCOUNTING ALERT

The accounting for take-or-pay commitments in usage-based SaaS is an active area of auditor scrutiny. If the company has material take-or-pay contracts where customers consistently under-consume their commitments, the auditors will question whether the committed amount truly represents the transaction price, or whether the variable consideration constraint requires revenue to be limited to actual consumption. Document your position and the supporting ASC 606 analysis before the audit begins.

8.3 Infrastructure Capitalization vs. Expense

In usage-based SaaS, the cloud infrastructure serves a dual purpose: it is both the platform that generates revenue and a capital asset being enhanced through engineering work. When the engineering team builds optimizations to the infrastructure — new data processing architectures, caching layers, compression

algorithms — that reduce the cost per unit of consumption, are those costs capitalized under ASC 350-40 or expensed as operations?

The answer depends on whether the work represents the development of new functionality (which may be capitalizable) or the optimization of existing functionality (which is generally expensed as maintenance). The distinction matters because capitalizing infrastructure optimization costs can significantly improve the apparent profitability of the business in periods of heavy infrastructure investment. Build a clear policy — in consultation with your auditors — that distinguishes between capitalizable infrastructure development and expense-able infrastructure operations.

SECTION 9

BUILDING THE USAGE-BASED FINANCIAL MODEL

The Consumption Revenue Financial Model

The usage-based SaaS financial model is structurally more complex than the seat-based SaaS model because it requires modeling customer behavior (consumption patterns) in addition to customer count and contract size. The best-practice approach builds the model at the cohort level — modeling each customer cohort's consumption ramp, steady-state consumption, expansion trajectory, and churn probability separately — and then rolls up to the total company P&L;

9.1 The Consumption Forecast Architecture

Model Layer	Inputs Required	Output	Confidence Level
Committed Revenue	Active contracts, monthly committed amounts, renewal schedule	Monthly committed revenue by customer	High (>90%)
Current Cohort Consumption	Trailing 3-month consumption per customer, growth trend	Expected consumption for existing customers	Medium-High (75%–85%)
New Logo Ramp	Pipeline by segment, close rates, ACV, ramp curve by segment	New logo contribution by month	Medium (60%–75%)
Expansion / Upsell	Usage trends, product roadmap, CSM expansion pipeline	Expected expansion ARR and consumption	Medium (55%–70%)

Model Layer	Inputs Required	Output	Confidence Level
Churn / Contraction	Customer health scores, renewal schedule, utilization rates	Expected revenue at risk	Medium (60%–75%)
Overage / Seasonal	Historical overage patterns, seasonality index by customer	Expected overage revenue by month	Low-Medium (50%–65%)

9.2 The Burn Multiple in Usage-Based SaaS

The Burn Multiple — net cash burned divided by net new ARR added — is an important efficiency metric for any SaaS company, but it requires careful interpretation in a usage-based context. Because consumed ARR is not committed, the "ARR" denominator of the burn multiple is softer than in a seat-based SaaS model. A usage-based company should calculate burn multiple using both committed ARR and consumed ARR, and present both to the board with an explanation of the difference.

BURN MULTIPLE FOR USAGE-BASED SAAS

`Burn Multiple (Committed) = Net Cash Burned / Net New Committed ARR`

`Burn Multiple (Consumed) = Net Cash Burned / Net New Consumed ARR`

`Use Committed for: investor communication, efficiency benchmarking`

`Use Consumed for: internal planning, near-term cash flow modeling`

`Benchmark: <1.0x = excellent; 1.0x-1.5x = good; 1.5x-2.0x = manageable; >2.0x = inefficient; investigate S&M ROI immediately`

SECTION 10

COMPLETE METRICS FRAMEWORK — USAGE-BASED SAAS

The Usage-Based SaaS Metrics Framework

Usage-based SaaS requires the full seat-based SaaS metrics framework from Part 2, plus a set of consumption-specific metrics that capture the unique dynamics of the model. The following are the additional and modified metrics that the usage-based SaaS CFO must own, organized by category.

10.1 Consumption and Commitment Metrics

Metric	Formula	Benchmark / Target
Committed ARR	Annualized value of minimum contracted spend	Track trend; must grow with new logo adds
Consumed ARR	Trailing 1-month consumption revenue x 12	Should exceed committed ARR at healthy accounts
Commitment Utilization Rate	Consumed ARR / Committed ARR	>85% signals healthy renewal and expansion
Overage Revenue %	Overage Revenue / Total Revenue	15%–35% signals strong expansion motion
Consumption Growth Rate	(Current Consumption - Prior) / Prior	Same-customer growth; exclude new logos
Ramp to Steady-State (days)	Days from contract sign to 90% of expected consumption	Track by segment; declining = better onboarding
ARR at Risk	Committed ARR of customers with utilization <70%	Should be <15% of total committed ARR

10.2 Gross Margin and Infrastructure Metrics

Metric	Formula	Benchmark / Target
Gross Margin	Gross Profit / Revenue	50%–75% depending on model type
Infrastructure Cost per Unit	Total Variable Infrastructure Cost / Units Consumed	Must decline QoQ; track trend obsessively
Variable COGS % of Revenue	Variable Infrastructure / Revenue	Must decline as volume scales
Fixed COGS % of Revenue	Fixed Infrastructure + CS / Revenue	Declining with scale — operating leverage signal
Gross Margin per Customer	Customer Gross Profit / Active Customers	Track by cohort and segment
Overage Gross Margin	Overage Gross Profit / Overage Revenue	Should be higher than committed GM%

10.3 Forecasting Accuracy Metrics

Metric	Formula	Target
Revenue Forecast Accuracy	$1 - \text{Actual Rev} - \text{Forecasted Rev} / \text{Forecasted Rev}$	>90% for 30-day forecast; >80% for 90-day
Consumption Forecast Error	Avg absolute error vs. actual by customer	Track improvement over time; declining = better model
Billings Forecast Accuracy	$1 - \text{Actual Billings} - \text{Forecasted} / \text{Forecasted}$	>85% for 30-day forward
Pipeline to Revenue Conversion	Closed ARR / Beginning Pipeline	Varies; establish baseline and track trend
Seasonal Adjustment Accuracy	Forecast accuracy after applying seasonality index	Should materially improve unadjusted accuracy

10.4 Efficiency Metrics (Usage-Based Specific)

Metric	Formula	Benchmark
Burn Multiple (Committed)	Net Cash Burned / Net New Committed ARR	<1.5x healthy; <1.0x excellent
Burn Multiple (Consumed)	Net Cash Burned / Net New Consumed ARR	<2.0x acceptable; <1.5x healthy
NRR (Usage-Based)	$(\text{Beg ARR} + \text{Expansion} + \text{Overage} - \text{Decline} - \text{Churn}) / \text{Beg ARR}$	>120% strong; >150% exceptional
CAC Payback (Committed)	$\text{CAC} / (\text{Monthly Committed Rev} \times \text{GM}\%)$	<18 months target; <12 months excellent
Rule of 40 (Consumed)	Consumed Revenue Growth % + EBITDA Margin %	>40 healthy; >60 exceptional
Infrastructure Efficiency Ratio	Revenue / Variable Infrastructure Cost	Must improve (increase) QoQ as scale grows

SECTION 11

USAGE-BASED SAAS CFO CHECKLIST

The Usage-Based SaaS CFO Checklist

The following checklist covers the minimum set of capabilities the CFO of a usage-based SaaS company must maintain, in addition to the seat-based SaaS checklist from Part 2. These items are specific to the consumption model and address its unique financial, accounting, and operational challenges.

Revenue and Consumption Management

- ARR definition documented clearly — distinguishing between committed ARR and consumed ARR; definition disclosed consistently to all external parties.
- Commitment utilization rate calculated monthly for every customer with a committed contract; customers below 70% utilization flagged to customer success immediately.
- Three-layer consumption forecast model maintained (committed, historical run-rate, new logo ramp); updated monthly with actuals and rolling 90-day forward view.
- Overage revenue categorized monthly as structural vs. one-time; structural overage included in forward forecast; one-time overage excluded.
- Consumption seasonality index maintained by customer segment; applied to all forecasting and board reporting.
- Ramp curve actuals vs. assumptions reviewed quarterly; model updated when ramp curve behavior diverges from historical assumptions.

Accounting and Revenue Recognition

- ASC 606 variable consideration policy documented; auditors have reviewed and approved the approach for each contract type (committed, pure PAYG, hybrid).
- Take-or-pay commitment accounting policy documented; shortfall revenue recognition timing reviewed with auditors each quarter.
- Credit accounting policy documented distinguishing between pre-contract incentive credits, post-contract SLA credits, and free-tier product-led growth credits.
- Infrastructure capitalization vs. expense policy documented under ASC 350-40; engineering work on infrastructure optimization correctly classified.
- Deferred revenue waterfall maintained separately for committed portion and variable portion; reconciled to cash collected monthly.

Infrastructure Cost Management

- Infrastructure cost per unit of consumption calculated monthly and trended over rolling 8 quarters; declining trend required; flat or rising trend triggers investigation.

- Cloud provider committed use discount (CUD) and reserved instance contracts reviewed semi-annually; utilization rate of reservations monitored monthly.
- Variable vs. fixed infrastructure cost split tracked monthly; variable component modeled against revenue scenarios for gross margin stress testing.
- Infrastructure architecture review completed at least annually with engineering leadership; cost reduction roadmap with specific targets maintained.

Tax and Compliance

- Sales tax sourcing analysis completed for all major jurisdictions; place-of-use vs. billing address determination documented for each state.
- Global digital services tax (DST) exposure analysis completed for all jurisdictions with >\$5M in annual revenue; compliance calendar maintained.
- Cross-border data processing transfer pricing documentation current; DEMPE analysis completed for all jurisdictions where value-creating functions are performed.
- Section 174 R&D; amortization impact modeled annually; cash tax impact included in annual operating plan.

Closing Perspective: The Consumption CFO

Usage-based SaaS is, in many ways, the most honest of all SaaS business models. It aligns the company's revenue with the value it actually delivers — when customers derive more value, they consume more and pay more. There is no hiding behind annual prepayments from customers who are not using the product. The consumption data tells the truth every month, and the CFO's job is to read that truth clearly, forecast it accurately, and use it to allocate capital wisely.

The greatest risk in a usage-based business is not high churn in the traditional sense — it is consumption compression. A customer who stays but reduces consumption from \$100,000 to \$60,000 per year is not a churned customer by any logo metric, but they have created a \$40,000 annual revenue hole. The CFO who builds their metrics framework around customer count and logo retention will miss this signal entirely. The CFO who tracks commitment utilization, consumption growth rates, and ARR at risk by customer will see it at month 6 — in time to intervene.

The infrastructure cost management challenge — reducing the cost per unit as consumption scales — is equally distinctive. It requires the CFO to develop a working relationship with the engineering and

infrastructure teams that goes well beyond the typical finance-engineering dynamic. The CFO who can speak credibly about reserved instance strategy, architecture optimization ROI, and egress cost reduction is the CFO who can manage gross margin actively rather than reporting it passively.

Part 4 examines the Platform and API business — a model that shares structural features with usage-based SaaS but introduces developer economics, API rate limit monetization, collaborative arrangement accounting, and transfer pricing for global API traffic as distinctive additional layers of complexity.

End of Part 3: Usage-Based / Consumption SaaS | Financial Architecture of Different Business Models

eFuturesCFO | The Systems CFO Platform | efuturescfo.com