

Case Study: Chargeback at Live-Event Scale

A subscription and payments platform sits behind the apps of some of the world's biggest sports and entertainment brands. It runs shared, multi-region infrastructure across AWS and GCP, serving more than 20 major customers whose traffic rises and falls around live events.

\$1.3M/month

CLOUD SPEND

20+

CUSTOMERS

AWS + GCP + AI

ENVIRONMENT

30 days

TIME TO VALUE

Challenge

Live events send customer apps into spikes lasting hours, with dozens of customers hitting shared databases in overlapping waves. Shared environments such as Aurora RDS clusters cannot be tagged per customer, so standard cloud tools could not split the bill. The platform's contracts allowed cost-plus chargeback, but there was no automated, audit-defensible method to support it, and every cost that could not be allocated came straight out of its own P&L.

Solution

- Turned traffic into a signal using load balancer logs to identify each customer by IP and weighting every call by response time as a live proxy for utilization.
- Split shared costs evenly, proportionally or by custom drivers such as that proxy — fully automated through the API framework.
- Ingested memory and disk utilization (which cloud bills don't show) and filtered them through a semantic layer of 1,850 metrics to right-size resources within SLA.

Benefits

- **Dynamic chargeback:** shared costs resolve automatically to the right customer, smoothing out monthly.
- **Audit-ready allocations:** every allocation traces back to actual calls, with 13 months of history to check against.
- **Hidden waste surfaced:** underused resources show up immediately for right-sizing or added load.
- **Profitability by any dimension:** costs segment per customer, vertical, region, application or event, across both clouds.

Cloud & AI Cost Management

Cloud and AI spend is now one of the fastest-growing lines on the technology cost base, and one of the least understood. Costs sit in shared environments that cannot be tagged, arrive as invoices that say what you paid but not why, and rarely resolve to the team, project or customer that drove them. Agents make it harder still, running continuously and decoupling spend from headcount.

CostPerform Cloud & AI ingests cost and usage data from every major cloud and AI provider, allocates it with audit-grade logic to the teams, clients and products behind it, and turns that visibility into showback, chargeback and unit economics. Trusted by over 200 organizations, including British Telecom, ING bank, and the USPTO, CostPerform delivers unparalleled financial clarity and cost optimization.

costperform.com/demo



w. costperform.com
t. +31 (0) 85 023 2114
e. info@costperform.com

CostPerform Cloud & AI for Enterprise

- **True Unit Economics:** Break spend down to the token, model, agent or workload, so you can measure the return on AI and cloud investment.
- **Allocate and Charge Back Instantly:** Split shared GPU pools, model inference and agent activity by team, business unit or client, so shared spend lands where it was actually used.
- **Proactive Governance:** Set budget thresholds by team, project or model, and catch the deviations you didn't predict with anomaly detection.
- **Connect Every Platform:** Native connections to 15+ cloud and infrastructure platforms, including AWS, Azure, Google Cloud, Oracle, Kubernetes, so nothing sits outside the model.
- **Capture Every AI Signal:** Ingest usage from all standard enterprise AI tools, including Anthropic, OpenAI, Gemini, Bedrock, and Vertex, so no spend goes unattributed.
- **Prove Every Number:** Trace every allocation back to transparent, versioned logic that holds up with customers and auditors.
- **See the Whole Cost Picture:** Carry the same model from Cloud and AI through to bill of IT, shared services and enterprise-wide profitability.

costperform.com/demo

CostPerform USA

+1 410 385 5325

infousa@costperform.com

CostPerform The Netherlands

+31 (0) 85 023 2114

info@costperform.com

