

# Ray vs. Anyscale



Ray is the world's most adopted AI compute engine. Anyscale makes it production and AI-coding agent ready.

Agentic-First Experience		
AI coding tool - e.g. Claude Code, Cursor - integration accessible via CLI	—	✓
Anyscale Agent Skills to develop, debug and guide infra setup of Ray workloads	—	✓
Agents deploy clusters & debug failures through logs, metrics & code analysis	—	✓

Multi-Cloud Orchestration		
Priority-aware scheduler for job queues	—	✓
Multi-cloud GPU capacity search and routing	—	✓
Deploy Ray clusters across AWS, GCP, Azure, Nebius, CoreWeave and other K8s environment from one control plane	—	✓

Security & Governance		
SSO / SAML / SCIM	—	✓
RBAC for orgs, projects, Workspaces, clusters, Jobs, Services	—	✓
Audit logs across cluster + workload actions	—	✓

Interactive Dev Environments		
Multi-node cluster backed IDE (VSCode, Jupyter) with Workspaces	—	✓
Elastically scalable Ray clusters for development integrated with AI coding tools	—	✓
Automated dependency management across nodes with uv	—	✓

Workload Observability		
Persistent logs after cluster shutdown	—	✓
Workload-aware dashboards for Ray Data (DAG/Tree)	—	✓
GPU/memory profiling for Ray Train	—	✓
Model lineage with integration with MLFlow + W&B	—	✓

Budgets & Billing		
Per-team/project budgets & quotas	—	✓
Usage & spend breakdown for all cluster types (Workspace, Job, Service)	—	✓
Cost guardrails & alerts	—	✓

Production-ready Ray Clusters		
Managed Jobs with queues, retries, alerting	—	✓
Managed Services for online inference (zero-downtime rollouts)	—	✓
Promote Workspace code > Jobs > Services without rewrites	—	✓

Cluster Monitoring		
Unified view of cluster health, utilization, uptime	—	✓
Machine history dashboard (node failures, replacements)	—	✓
Export to Grafana, Prometheus, Datadog, CloudWatch	—	✓

Committer-led Support		
Mission-critical SLAs	—	✓
Professional Services	—	✓
Dedicated Technical Account Manager	—	✓



Anyscale

Explore Anyscale today for free with \$100 credits

## runway | Case Study

Anyscale empowers **40+ engineers** to build data curation pipelines feeding models that generate video **4x faster**.

“Anyscale enables us to push the boundaries of what’s possible in generative AI by giving us the flexibility to scale workloads seamlessly.”

Anastasis Germanidis | Co-Founder & CTO, Runway