

Free Starter lands developers on a managed vector database for RAG and search; Builder and Standard serverless usage convert PLG builders; Enterprise adds SLA, private networking, and BYOC — while Assistant, Inference, and Nexus extend Pinecone into an AI knowledge platform sold with AWS, GCP, and Microsoft marketplaces.

KEY PARTNERS

- AWS, GCP, Microsoft cloud marketplaces (pricing page).
- Orchestration / RAG frameworks (LangChain ecosystem and tutorials).
- Model providers feeding Inference + customer embedding pipelines.
- Data/ingestion and warehouse partners (Nexus / KnowQL alliances).
- SI and Marketplace app partners packaging domain knowledge apps.
- Named customer co-marketing (Gong, Vanguard, ZoomInfo, etc.).

KEY ACTIVITIES

- Operate serverless On-Demand indexes (object-storage backed).
- **Improve retrieval** — dense, sparse, full-text, hybrid, rerank.
- Ship DRN, backups, namespaces, and multi-tenancy controls.
- Build Inference, Assistant, Nexus, and Marketplace apps.
- Self-serve billing + enterprise security/compliance sales.
- **Developer education** — docs, examples, Discord community.

KEY RESOURCES

- Serverless vector DB architecture (storage/compute split).
- **Retrieval IP** — dense/sparse/full-text + reranking stack.
- Hosted Inference model catalog on Pinecone infra.
- Developer brand, docs, SDKs, Discord community.
- Enterprise security posture (SOC 2, HIPAA options, BYOC).
- **Investor capital** — a16z-led Series B (press, Apr 2023).

VALUE PROPOSITIONS

- **Managed vector DB** — no cluster capacity planning.
- **Hybrid retrieval** — dense, sparse, full-text, rerank.
- **Serverless cost model** — pay for storage + operations used.
- Free Starter land; Builder \$20 bridge to production.
- **Enterprise** — 99.95% SLA, SSO/RBAC, BYOC, private networking.
- Assistant + Nexus for managed RAG / agent knowledge.

CUSTOMER RELATIONSHIPS

- Self-serve signup and console for Starter/Builder.
- Community Discord on Starter; Free support on Std.
- **Support ladder** — Developer \$29 / Pro \$250 / Premium.
- Enterprise CSM, security reviews, and SLA management.
- Docs, examples, and learn content for developers.
- **Expansion** — Database → Inference/Assistant/Nexus inside accounts.

CHANNELS

- **Primary** — pinecone.io + console onboarding.
- Developer docs, SDKs, examples, Discord.
- Framework integrations and RAG tutorials.
- AWS / GCP / Microsoft marketplace listings.
- Pinecone Marketplace knowledge apps.
- Direct enterprise sales for BYOC / large ACV.

CUSTOMER SEGMENTS

- Indie developers / prototypes (Starter free).
- Solo & small teams (Builder \$20 flat).
- Production AI startups (Standard usage).
- Platform / ML teams needing DRN & SSO.
- Enterprises needing SLA, CMEK, BYOC.
- Agent / RAG product builders (Assistant, Nexus).

COST STRUCTURE

- **Cloud COGS** — object storage, query compute, egress.
- Free Starter allowances as demand-gen cost.
- **R&D** — retrieval algorithms, Inference, Assistant, Nexus.
- **Sales & marketing** — PLG content + enterprise field.
- Support / success (Free → Premium ladder).
- Compliance & BYOC operating overhead.

REVENUE STREAMS

- **Starter free** — demand gen (FACT — pricing).
- **Builder** — \$20/mo flat (FACT — pricing).
- **Standard** — \$50/mo min + \$0.33/GB + RU/WU PAYG (FACT).
- **Enterprise** — \$500/mo min + higher RU/WU; SLA/BYOC (FACT).
- **Inference embed/rerank token meters** (FACT — pricing).
- **Assistant token + storage meters** (FACT — pricing).
- **Support** — Developer \$29 / Pro \$250 / Premium (FACT).