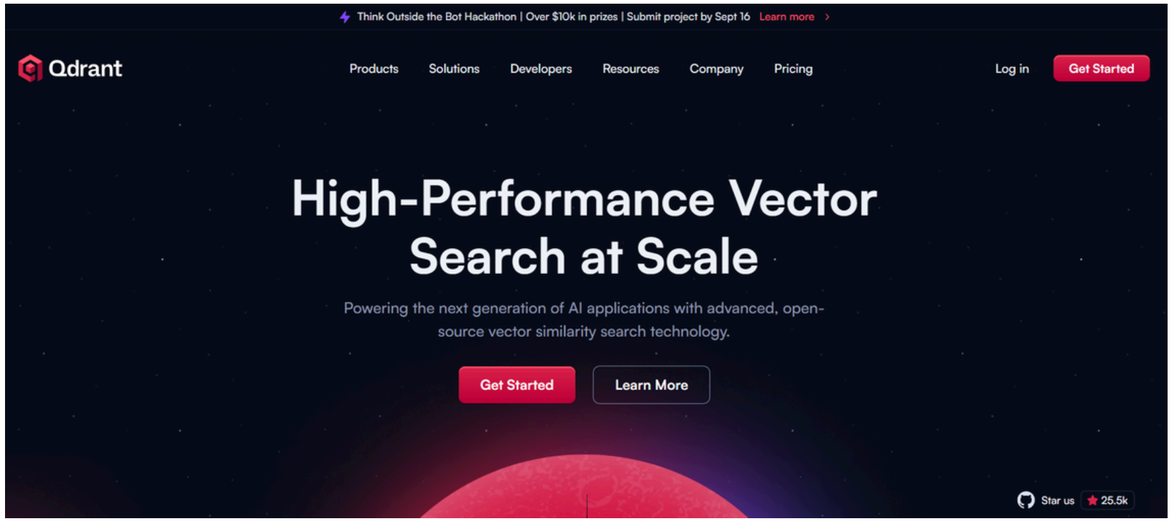
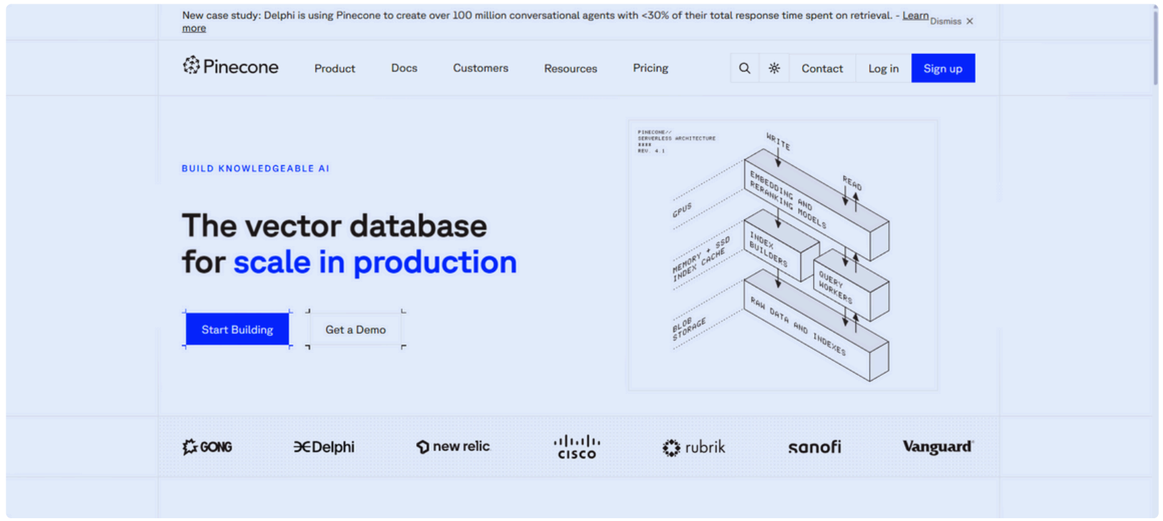
Qdrant vs Pinecone

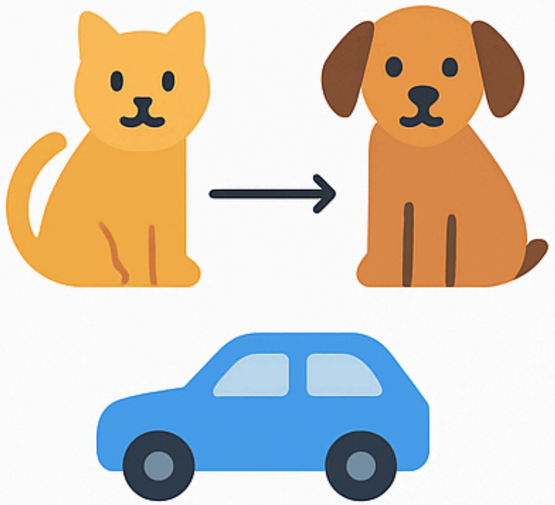
Choosing the right Vector Database





Why Vector Databases?

- AI search & chatbots need meaning-based search
- Vectors = numbers that capture meaning
- Example: “cat” is closer to “dog” than “car”



Meet the Players

Qdrant

- Open-source
- flexible
- runs anywhere

Pinecone

- Cloud-first
- serverless,
- zero-ops

Ease of Use & Hosting

Qdrant

1. Docker
2. Kubernetes
3. self-host or
4. Cloud

Pinecone

1. SaaS,
2. pay-as-you-go,
3. no setup

Data Model (Qdrant)

- **Collections** : holds collection of points
- **Points (vector data + metadata payload + ID)** : Point can be thought of individual data entities and it holds vector or multiple named vectors. Additionally points can have an optional metadata.
  - **Named Vectors** allow us to hold vectors for a point in a different type.
  - **Example:** A single point about a product can have an embedding for the product description, another embedding vector for product image and so on.

Data Model (Pinecone)

- **Index**: A container which can hold data from one or more pods (hardware unit)
- **namespace:** For supporting multi tenancy. can be though of as a collection for a given purpose
- **Record** => ID + VECTOR + PAYLOAD
- **Note**: No support for named vectors within the same record.

Filtering & Search

- **Qdrant**:
  - full-text
  - geo
  - nested filters
- **Pinecone**
  - metadata filters
  - lexical indexes
- **Both**
  - hybrid search = semantic + keyword

Scaling and Performance

- **Qdrant**:
  - Manual Scaling
  - Sharding and Replicas
  - We need to do this manually unless we are on Qdrant Cloud Paid
- **Pinecone**
  - Auto Scaling
  - Server and Replicas
  - No need to worry about anything, everything is taken care of by pinecone

Pricing

- **Qdrant**:
  - Open Source Code available for anyone to use
  - Can go for paid qdrant cloud
  - if self managing, cost depends on infra and servers
- **Pinecone**
  - Usage based
    - Read Usage (RU)
    - Write Usage (WU)
  - Storage based
    - GB storage data in index

Strengths

- **Qdrant**:
  - control,
  - filters,
  - multi-vectors,
  - geo
  - Can go for paid qdrant cloud
- **Pinecone**
    - no-ops,
    - auto-scale,
    - usage billing

Use Case Mapping

- Enterprise / Gov → Qdrant
- Startups / MVP → Pinecone
- Hybrid Search → Both

Final Recommendation

Choose based on your ops model, not hype.