

使用 SelectDB 构建 企业级向量基础设施

衣国垒

飞轮科技技术副总裁、Apache Doris PMC 成员



超越检索

使用 SelectDB 构建企业级向量基础设施，重塑 AI 时代的数据检索与分析能力

AI 正在创造新的数据工作负载

数据形态演变



结构化数据：传统基石

以数据库表、行列结构为核心，具备严格的格式规范，是企业最基础、最规范的数据存储形态。



非结构化数据：海量涌现

文本通过倒排索引或向量检索，图片被转化为密集向量矩阵，打破了传统格式限制，成为数据主体。



向量：连接异构数据的基础设施

作为检索文本、图片等非结构化及半结构化数据的关键技术，实现了多模态数据的统一理解与处理。

新兴工作负载



RAG

检索增强生成



语义搜索

上下文相似匹配



推荐系统

个性化项目建议



照片搜索

视觉内容理解



欺诈检测

模式异常检测



可观测性

系统健康监控



模型训练

向量输入微调



智能体记忆

长期上下文保留



聚类分析

过去几年努力解决的问题



ANN 算法支撑

通过近似最近邻算法，高效处理高维向量空间的相似性搜索，奠定检索基础。



Top-K 快速定位

在海量非结构化数据集中，精准筛选并快速返回最相似的前K个结果，提升效率。



极致低延迟响应

优化底层查询路径与索引结构，确保用户请求在毫秒级完成响应，保障体验。



高召回率平衡

在追求检索速度的同时，兼顾结果的准确性，实现速度与召回率的最佳平衡。

问题 #1: 大规模应用时的成本问题

内存优先的架构成本迅速增加

1亿规模

原始大小: ~400 GB
所需内存: 数百GB DRAM

10亿规模

原始大小: ~4 TB
所需内存: 多TB DRAM集群

100亿规模

原始大小: ~40 TB
DRAM 成本不切实际



嵌入持续增长

数据量随时间呈指数级增长



冷数据占主导

大部分数据不常访问



DRAM 成本爆炸

线性扩展成本高昂

处方: 瓶颈已从算法转向经济成本

问题 #2: 向量计算的负载多样

⚡ Online Retrieval

面向实时服务场景，涵盖RAG、图片搜索、在线推荐及实时风控等业务，对系统响应速度有着极致要求。

🔍 Ad-hoc Exploration

面向人工分析场景，涉及Embedding分布分析、训练数据导出及异常样本探索，模式灵活且依赖人工交互。

💾 Offline Processing

面向海量数据批处理，包括索引重建、全量数据质量扫描及向量聚类分析，核心追求高吞吐量与计算效率。

类别	核心诉求	时间要求	查询模式	数据规模	典型特点
Online Retrieval	低延迟响应，高可用性	毫秒 - 秒级	高频、小查询请求	中大规模	Serving-oriented
Ad-hoc Exploration	灵活分析，结果可解释	秒 - 分钟级	模式不固定，按需查询	大规模	Human-in-the-loop
Offline Processing	批量处理，高计算效率	分钟 - 小时级	大任务、吞吐量优先	超大规模	Throughput-oriented

问题 #3: 架构碎片化



企业数据平台

作为中心化的数据平台，汇聚结构化与非结构化数据，支撑BI报表、趋势分析等离线数据分析场景。



向量数据库

专门用于存储高维向量嵌入，核心能力是支持大规模、低延迟的近似最近邻（ANN）相似性搜索。

重复与浪费：成本高昂的效率陷阱

数据在多系统间流转时被重复存储和处理，不仅占用大量存储资源，还造成计算能力的冗余消耗，推高整体运营成本。

治理混乱：难以统一的管控难题

复杂的跨系统工作流导致数据标准不一、版本混乱，缺乏统一的元数据管理与质量监控，给数据治理带来极大挑战。

SelectDB愿景：统一的AI数据基础设施

将向量作为另一种列类型，而非另一个系统，构建极简、高效的AI数据底座

全类型数据原生支持



结构化数据

完美兼容传统关系型数据，提供极致的分析查询性能，保障业务数据的高效处理。



半结构化数据

原生解析JSON、Parquet等格式，无需额外ETL转换，直接对嵌套数据进行高效分析。



向量嵌入 (Vector Embedding)

将向量视为普通列类型，与业务数据同库存储，打破向量与标量数据的存储壁垒。



Apache Doris

高性能统一分析引擎，在单一平台实现存储、索引与处理一体化，彻底简化数据技术栈。

One Engine, All Data

四大核心能力赋能业务



混合检索

通过标准SQL，融合标量过滤与向量相似度检索，实现精准+语义的多维查询。



线性扩展

MPP 的分布式架构，灵活的支持从百万级，到百亿级别的向量



低成本

结合量化压缩和DiskANN，冷热分离等技术，大幅度降低成本。



多负载隔离

通过存算分离的架构，灵活的隔离，在线，adhoc，离线等多种类型的任务

完备的向量能力

```
CREATE TABLE sift_1M (  
  id int NOT NULL,  
  embedding array<float> NOT NULL COMMENT "",  
  INDEX ann_index (embedding) USING ANN PROPERTIES(  
    "index_type"="hnsw",  
    "metric_type"="l2_distance",  
    "dim"="128",  
    "quantizer"="pq",      -- Specify using PQ for quant  
    "pq_m"="2",            -- Required when using PQ, in  
                           sub-vectors  
    "pq_nbits"="2"        -- Required when using PQ, in  
  )  
) ENGINE=OLAP  
DUPLICATE KEY(id) COMMENT "OLAP"  
DISTRIBUTED BY HASH(id) BUCKETS 1  
PROPERTIES (  
  "replication_num" = "1"  
);
```

Top N Query

```
SELECT id,  
       L2_distance(  
         embedding,  
         [0,11,77,24,3,0,0,0,28,  
          0,6,92,8,14,73,125,29,  
          50,25,70,64,7,59,18,7,  
        ] AS distance  
      )  
FROM sift_1m  
ORDER BY distance  
LIMIT 10;
```

Range Query

```
SELECT count(*)  
FROM sift_1m  
WHERE l2_distance_approximate(  
  embedding,  
  [0,11,77,24,3,0,0,0,28,70,  
   0,6,92,8,14,73,125,29,7,0,  
   50,25,70,64,7,59,18,7,16,0,  
  ]  
> 300;
```

结合元数据、关键词和向量实现精准搜索

```
SELECT
  *
FROM
  auto_drive_data
WHERE
  MATCH(data_desc, '雨天')
  AND road_status = '路口' AND sensor_type = '前视摄像头' AND speed < 30
ORDER BY cosine_distance(embedding, [0.22, 0.56, 0.18, 0.71]) ASC
LIMIT 100;
```



结构化元数据过滤

通过结构化属性（如类别、日期、传感器类型、行驶速度等）进行前置过滤，快速缩小检索范围，提升查询效率。



关键词全文检索

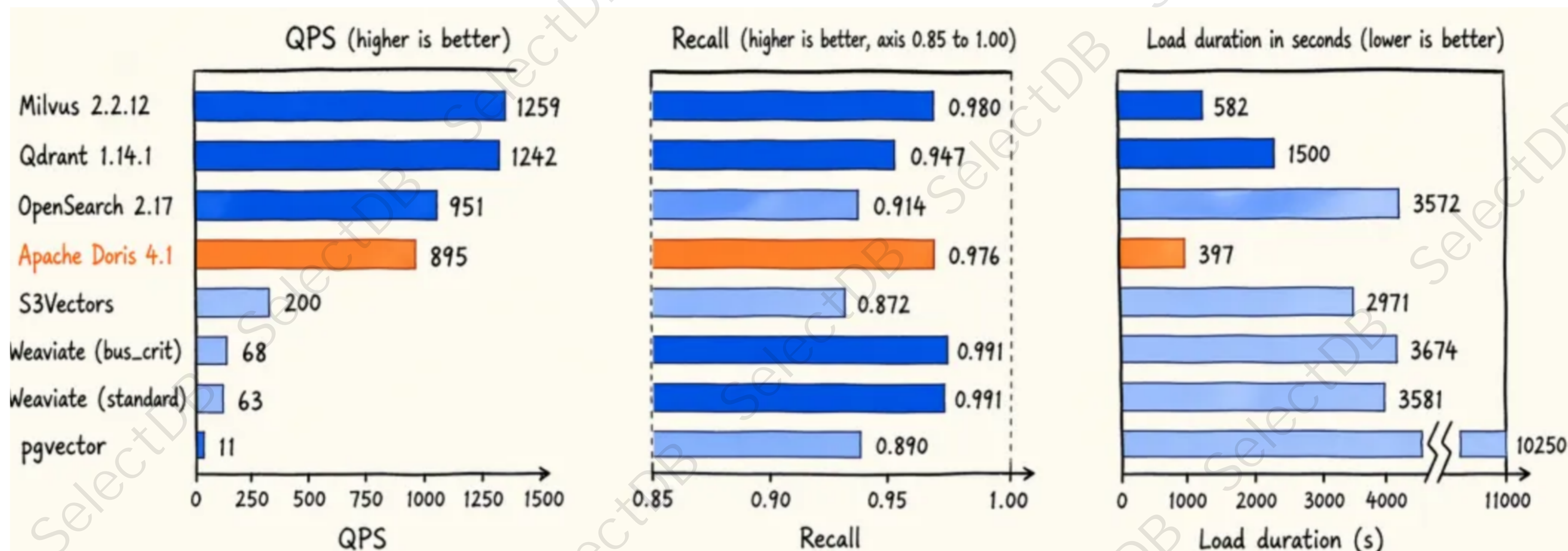
从非结构化的文本描述中，通过全文索引技术查找精确或模糊匹配项，精准定位包含特定信息的目标数据。



向量语义相似度排序

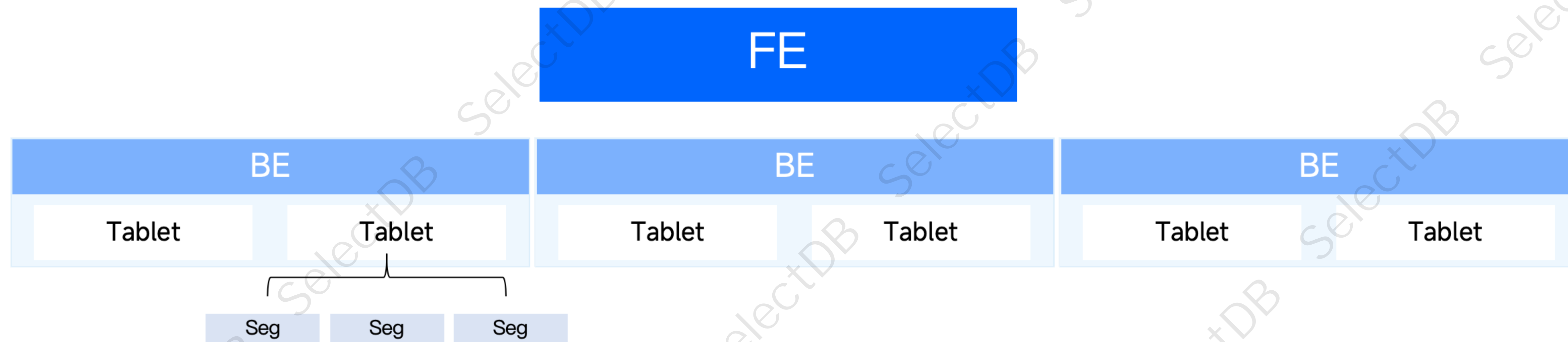
基于数据的语义向量，计算向量间的余弦距离，按照语义相似度进行排序，返回与查询意图最匹配的结果集。

可与专用向量数据库媲美的性能



Doris 4.1 有最快的导入性能，中等偏上的QPS 和 召回率支持

从百万扩展到**百亿**



集群级横向扩展： 按 Tablet 分布多机

数据按 Tablet 切分后分散在多台 BE 节点上，查询时并行下发计划。数据量越大，只需线性增加 BE 节点即可支撑更大规模的索引。

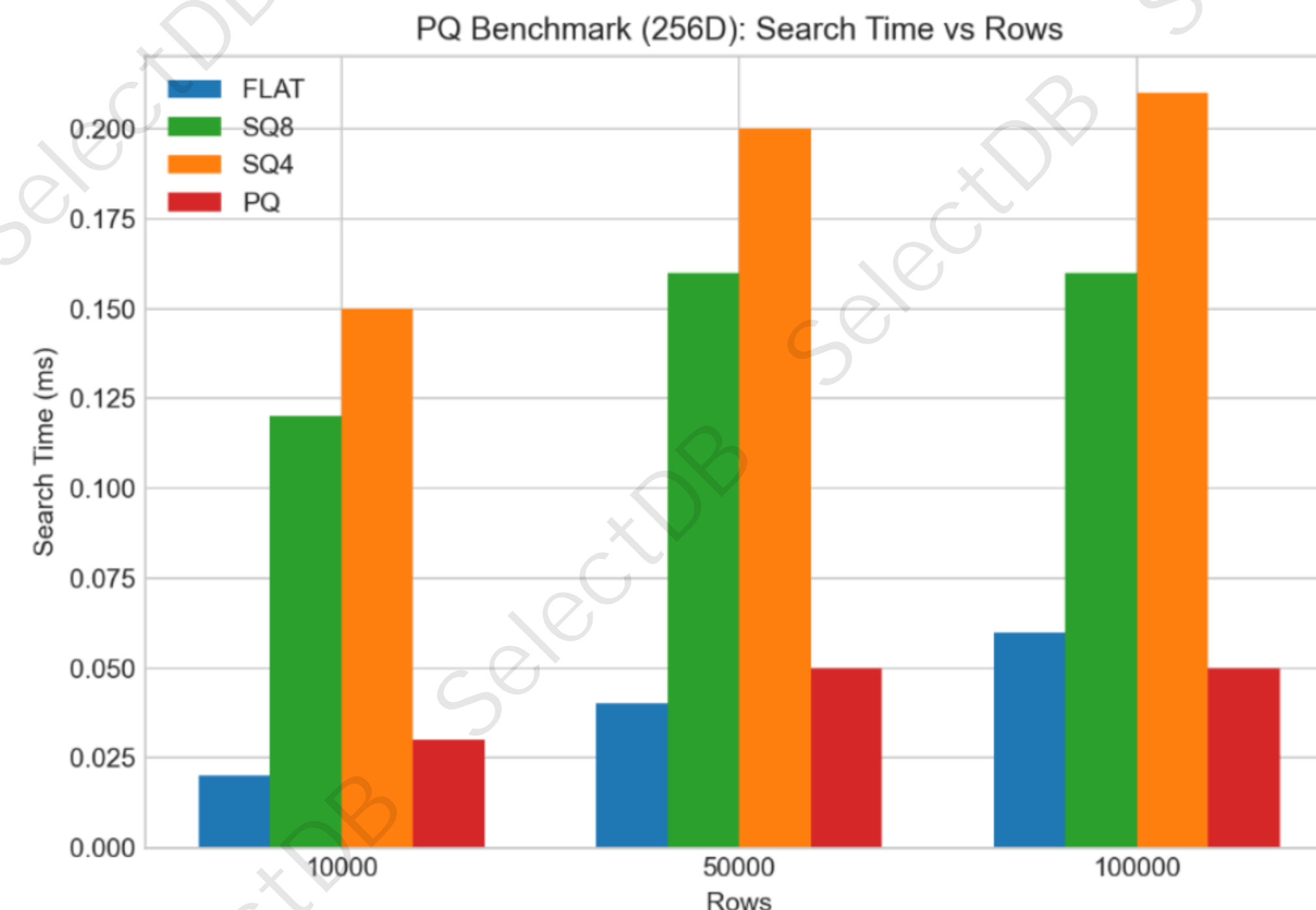
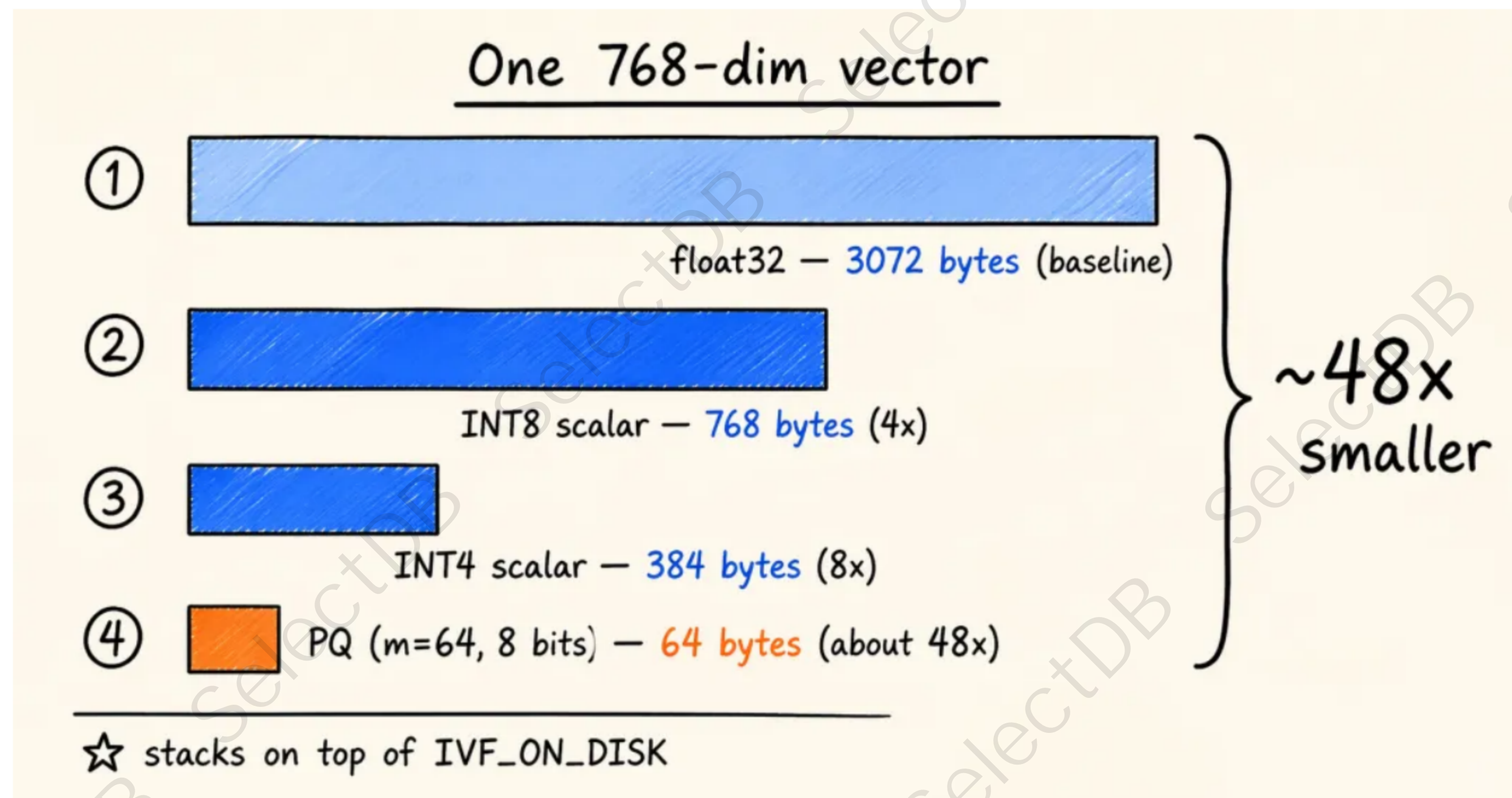
单机级纵向扩展： 按 Segment 多线程并行

单机数据进一步切分为 Segment，每个 Segment 独立建索引。查询时启动多线程并发扫描，充分利用多核 CPU 资源，最大化单机检索吞吐。

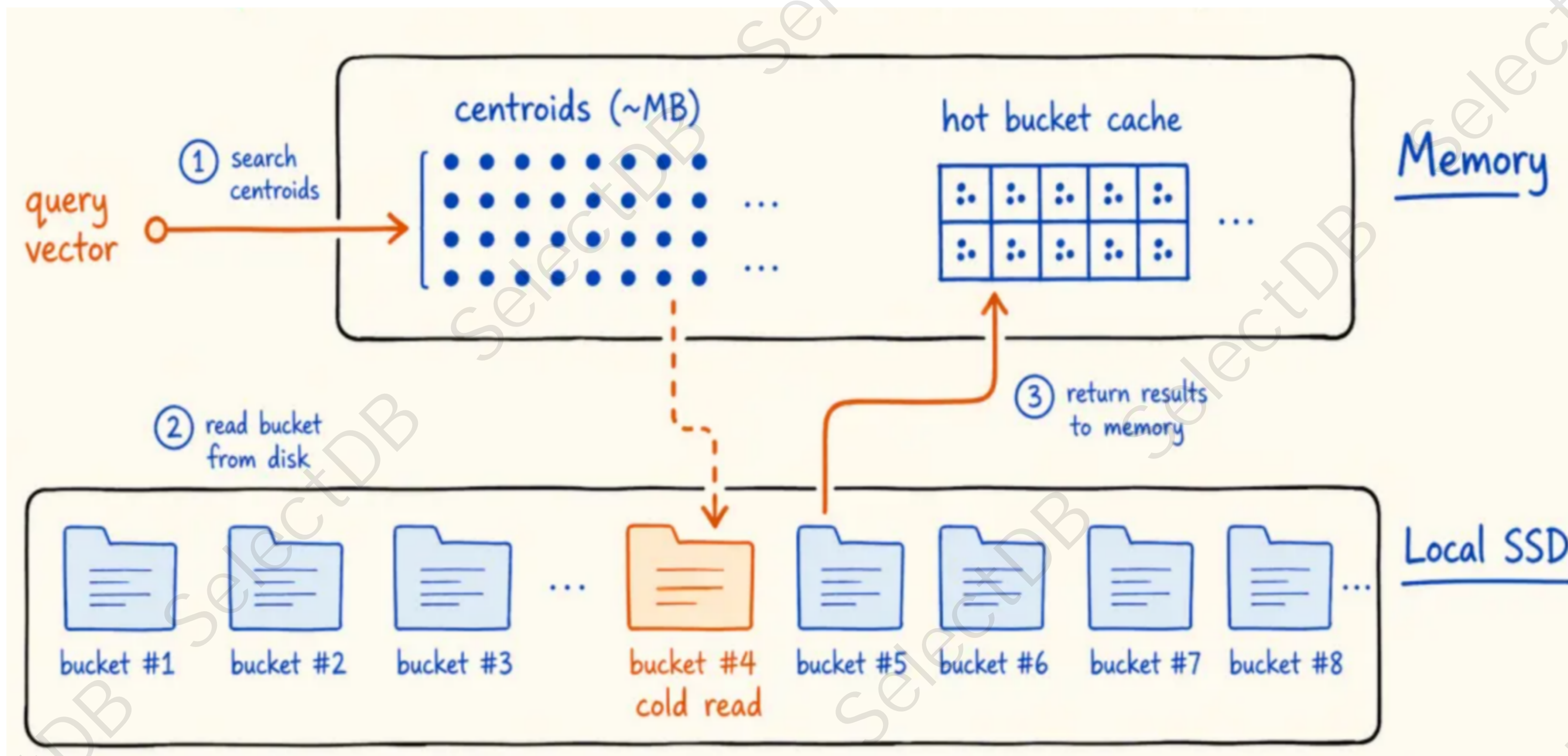
二维扩展叠加： 平滑支撑百亿规模

通过“机器数 × 线程数”的二维并行模型，既解决了容量问题，也解决了性能问题，实现从百万到百亿向量数据的无缝平滑扩展。

量化压缩，降低成本



IVF based DiskANN



内存中保存聚类中心和热数据缓存



磁盘中保存倒排列表和向量数据



结合存算分离技术，将本地磁盘中的冷数据进一步offload 到对象存储

多工作负载隔离

计算与存储分离带来弹性与隔离



Online Cluster



Ad-hoc Cluster



Offline Cluster



弹性伸缩

为每个工作负载独立调整资源，确保最佳性能与成本效率。



共享存储层

统一、持久化的存储层，聚合所有工作负载数据，确保单一事实来源，实现一致的分析和搜索。



工作负载隔离

不同任务互不干扰，避免资源抢占，保障稳定性。

通过解耦计算和存储，确保高资源利用率，同时提供灵活性以经济高效地处理各种工作负载。

超越检索

企业级AI基础设施需要的远不止ANN

统一存储 + 分析 + 向量搜索是AI数据基础设施的下一阶段。