Solution Brief

VAST AI OS バージョン 5.5の最新機能

高速検索、スマート分析、統合パイプライン基盤 すべてを、ひとつのリリースで

次世代のエンタープライズAIに必要なのは、個々のコンポーVAST AI OS 5.5は、まさにこの発想を具現化するプラットフォームとして、ハイパースケールなベクトル検索、ネイティブ分析、クラスタ内のマネージド・パイプラインコンピュート、ゼロコピーのデータ処理をひとつのプラットフォームに統合しました。数兆ベクトル規模を扱うRAG、エンタープライズデータウェアハウスのクエリ高速化、独立したKubernetes基盤の運用負荷の削減などエンタープライズAIの課題に対して、バージョン5.5は少ないシステム運用負荷で、効率的なAIワークロードを推進するための基盤になります。

大規模なリアルタイムAIのためのプラットフォーム

Hyperscale Vector Index：オーバーヘッドのない大規模ベクトル検索

ベクトルデータベースには長らく厳しいトレードオフがありました。パフォーマンスのために規模を犠牲にするか、エンタープライズ規模を実現するために複雑な運用を受け入れるか、という選択です。たとえば、メモリ常駐型のインデックスはある規模まではうまく機能しますが、データ量が増えると限界を迎えます。シャーディングを使えば拡張の余地は生まれるものの、ワークロードが増えるにつれてリバランスの負荷、ガバナンスの複雑化、予測できないレイテンシを引き起こします。外部のベクトルストアを導入しても、統合・保護・保守すべき別のシステムが増えるだけです。

バージョン5.5では、VAST DataBaseに直接組み込まれた独自の階層クラスタリングアーキテクチャ「VAST Hyperscale Vector Index」を新たに導入しました。ベクトルは距離に基づくクラスタの多層階層に整理され、クエリ実行時には最も有望な候補群のみを探索し、量子化されたベクトルの対象部分だけをメモリに読み込みます。これにより、クエリのレイテンシはデータ量が増えても一定の水準にとどまり、インデックス全体は構造化データ・半構造化データとともにフラッシュ上に配置され、単一のガバナンスモデルの下で管理されます。

その結果、数兆ベクトル規模でもメモリ使用量を抑えながら、高い並列性と安定したベクトル検索性能を実現します。外部のベクトルストアやシャーディングも不要で、元データとそのベクトル表現との間でアクセス権限がずれる心配もありません。こうした特長は、動画解析、セマンティック検索、不正検知、大規模なRAGといった用途において、より高速かつ文脈的に精度の高い検索という形で効果を発揮します。

VAST Native Query Engine：データがある場所で分析する

従来のエンタープライズ分析では、実行エンジンとストレージプラットフォームが分離されているのが一般的でした。データは処理のために外部へ移動されるため、レイテンシの増加やパイプラインの複雑化、アーキテクチャの肥大化を招いていました。VASTはその逆のアプローチをとり、データがある場所で処理を行います。

VAST Native Query Engineは、これまでVASTプラットフォーム内でのベクトル検索を支えてきました。バージョン5.5では、より広範な分析ワークロードへと大幅に拡張され、50を超える新しい集計・統計関数をVAST内でネイティブに実行できるようになりました。

これにはSUM、MIN、MAX、COUNTといった基本操作、分散・標準偏差・歪度・尖度・共分散などの統計量、回帰・相関関数、分位数・パーセンタイルの推定、FILTER句による条件付き集計が含まれ、ブール演算やビット単位の集計もサポートされています。

これまで外部エンジンへデータを書き出す必要があった複雑な統計分析、機械学習の特徴量作成、探索的データ分析も、プラットフォーム上で直接実行できるようになりました。分析ロジックはデータのある場所で、同じガバナンスモデルの下で実行されるため、データ移動や追加のインフラは不要です。このネイティブなVAST Engineにより、ベクトル検索とSQLフィルタを組み合わせたハイブリッド分析も同じクエリパス内で実行できます。

大規模テーブルのパーティショニング：スケールしても、運用はシンプルなまま

大規模テーブルでクエリ性能を維持するためには、これまで多くのトレードオフを伴っていました。複雑なシャーディング構成やリバランス作業を管理するか、データが増えるにつれてクエリの実行が遅くなるのを受け入れるか、という選択です。バージョン5.5では、このトレードオフを解消します。

新たに追加されたSQLネイティブなパーティショニングにより、大規模な分析ワークロードに高いクエリ性能を実現します。ユーザーが複雑な運用を行う必要もありません。時間範囲、カテゴリ、あるいは任意の論理的な軸に沿って、高レベルなパーティショニングを宣言するだけで、VAST Engineがクエリ実行時にパーティションプルーニング、join最適化、効率的な削除処理を自動的に行います。シャードの管理やデータリバランス、ノードのオーナーシップの考慮も不要です。

VAST Native Compute：クラスタ内で完結するパイプラインコンピュート

これまで、データパイプラインとAIパイプラインを動かすには、保存・検索を担うデータプラットフォームと、計算負荷の高いパイプラインワークロードを担う別のKubernetes環境という、2つの環境を並行して運用する必要がありました。この境界を管理することでコスト、レイテンシ、運用上の摩擦が生まれ、パイプラインを動かし続けるためだけに、システム間でデータを移動させる必要もありました。

バージョン5.5では、DataEngineのサーバーレスPython関数に対応したVAST Native Computeを新たに導入、コンテナオーケストレーションをVAST AI Operating System内に実装しました。管理者は、VASTクラスタ内のCNode上にマネージドKubernetesクラスタをプロビジョニングし、コンテナ化されたAIアプリケーション、データ処理ワークロード、イベント駆動型サービス、サーバーレス関数向けに専用のコンピュートリソースの割り当てや、ベクトル化・チャンキング・埋め込み生成のための外部AIパイプラインもサポートします。プロビジョニング、スケーリング、アップグレード、ライフサイクル管理は、すべてプラットフォームが担います。

パイプライン関数はVASTクラスタ内で、処理対象のデータのすぐそばで実行され、VAST上のファイル・オブジェクト・データベース・イベントサービスに直接アクセスできます。その結果、データとパイプラインコンピュートが同じ運用基盤を共有する統合された実行モデルが実現します。「データはここ、計算はそこ」というような分断されたアーキテクチャに比べて、シンプルな運用、強固なガバナンス、低いレイテンシを実現できます。

効率的なデータ操作とパイプライン自動化

バージョン5.5にはさらに2つの機能が追加され、プラットフォーム全体で自動化を拡張し、不要なデータ移動をなくします。

NFS向けファイルトリガー・パイプライン： VAST DataEngineのイベント駆動型自動化を、S3だけでなくネイティブなNFSv4ワークフローにも拡張しました。ファイルの作成・削除・メタデータの変更が直接パイプラインをトリガーとして指定できるため、プロトコル変

換やオブジェクトストレージへの再構築が不要となりました。アプリケーションデータ、メディア処理、科学技術系ワークロード、あるいはレガシーなファイルベースの環境でNFSを利用している組織にとって、既存環境を活かしながら、リアルタイムな自動化を実現する、モダナイゼーションの取り組みが進められるようになります。

ブロック向けインスタント・データクローニング：ブロックレベルのコピー操作を、データ移動を伴う処理から軽量なメタデータ操作として実行できるようになりました。従来のようなソースデータを読み込んでターゲットに書き込むのではなく、ターゲットが同じ基盤となるフラッシュ上のデータを参照できるようになります。これにより、VMのフルクローン、Storage vMotion、SMBのサーバーサイドコピーといった大規模なコピー操作も、ストレージ帯域や書き込みバッファのリソースを消費せずに、ほぼ即時に完了します。また、ソースとターゲットは論理的に独立した状態を保つため、その後の書き込みや削除がクローンされたデータに影響することはありません。

VAST 5.5：パフォーマンスが変える会話

機能	結果	比較対象
ベクトル検索速度	10倍高速	OSSベクトル（AISAQ Index、10億ベクトル、128次元）
1,000件のベクトル検索あたりのコスト	91%削減	主要なOSSベクトルデータベース、500億ベクトル
選択的クエリ性能	60%以上高速	最適化されたIceberg
テーブル更新速度	200倍高速	最適化されたIceberg
クエリ性能（EDW）	20%高速	Icebergベースのテーブル、実際の製造業ワークロード

ベンチマーク結果はVAST社内テストによるものです。結果はワークロードや構成によって異なる場合があります。

VAST AI OS 5.5にアップグレードする理由

1兆ベクトル規模：
Hyperscale Vector Indexは、従来のベクトルデータベースでは難しい大規模ベクトル検索においても、高速かつメモリ使用量を抑えて実現。外部のベクトルストアやシャーディングも不要。

データを外に出さない分析：
50を超える新しいネイティブの集計・統計関数により、複雑な分析ワークロードをプラットフォーム内で処理。複雑な分析処理のための、外部分析エンジンへのデータ書き出しも不要。

大規模テーブルでの高速クエリ：
SQLネイティブなパーティショニングにより、大規模データセットのスキャン・join・更新を、運用負荷や手動でのデータ管理なしに高速化。

統合されたパイプラインコンピュート：
VAST Native Computeにより、DataEngine関数用に別途Kubernetesクラスタが不要。パイプラインを、データのすぐそばのVASTクラスタ内で実行可能。

NFSパイプライン自動化：
ファイルトリガー型のパイプラインにより、リアルタイムのイベント駆動自動化をNFSv4ワークフローにも拡張。

VASTについて

VAST Dataは、AI時代に最適化された「AIオペレーティングシステム（AI OS）」を提供する企業です。データ基盤とコンピュートサービス、そしてAIエージェントの実行環境を統合したスケーラブルな「VAST AI OS」により、組織はAIエージェントの展開と連携、リアルタイムデータを活用した推論、複雑なワークフローの自動化を実現できます。その中核を担うのが、性能・拡張性・シンプルさ・堅牢性のトレードオフを解消した、世界初の真の並列分散アーキテクチャ「DASE」です。従来の制約を打破するこのアーキテクチャを基盤に、VAST Dataは企業がAIの可能性を最大限に引き出せるよう支援しています。

VAST AI OSが自社の環境で何を実現できるか、確認してみませんか？すべてのデータから新たな知見を得る方法をお探しの方は、hello@vastdata.com までお気軽にお問い合わせください。

©2026 VAST Data, Inc. All rights reserved. All trademarks belong to their respective owners.
8/26 | 13393