

## Cassandraが限界に達したことを示す5つの兆候

最初のNoSQLプロジェクトにCassandraを使用することは良くあることですが、稼働後に負荷やデータ量が増えると状況が一変します。サービス品質保証契約(SLA)違反や所有コストの増大は一般的に発生し、予算超過、不安定なオペレーション、顧客サービスの問題が明らかになります。そのような時こそ、Aerospikeがお役に立ちます。

NoSQLキーバリュー・ストアであるAerospikeは、大規模な読み取り/書き込み負荷に対して超高速で予測可能なパフォーマンスを提供し、ほぼ全ての場合、総所有コスト(TCO)を大幅に削減します。Aerospikeの処理速度、高可用性、運用効率、自己管理機能により、多くのCassandraユーザーがAerospikeに移行しました。ある企業では、450ノードのCassandraクラスタを60ノードのAerospikeクラスタに置き換え、3年間で690万ドルのTCO削減を見込んでいます。

では、Cassandraの限界に達した可能性を示す5つの兆候とはどのようなもののでしょうか？

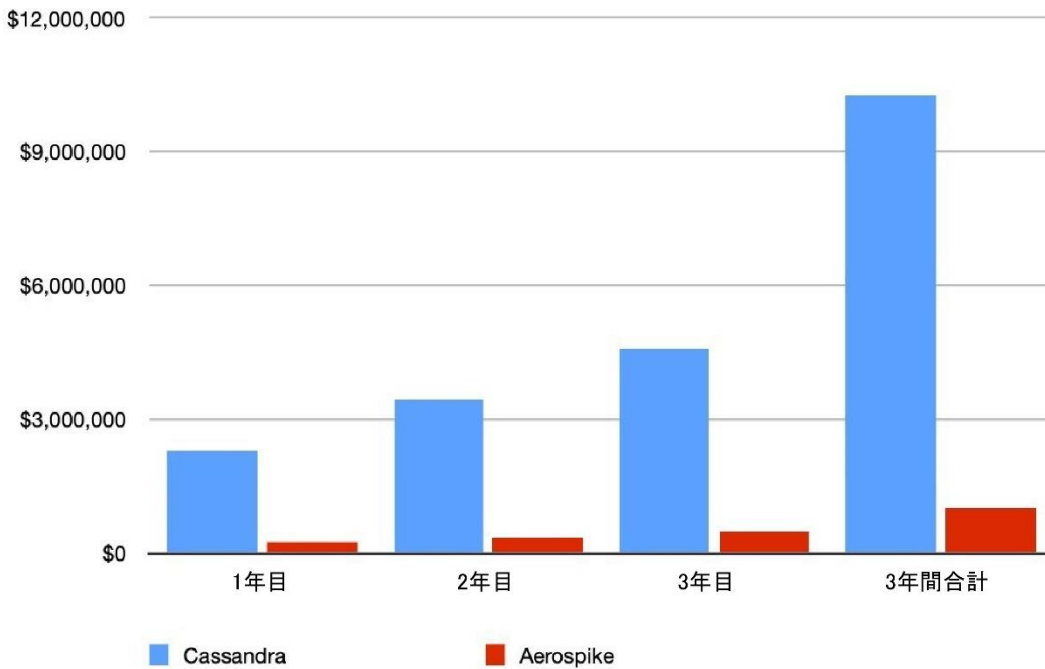
兆候1	ノードが制御不能なほど増え、TCO増大を懸念している
兆候2	負荷ピーク時にSLAを満たせない
兆候3	カスケード障害にうんざりしている
兆候4	運用チームの人員が増え続け、コストが増大し続けている
兆候5	Cassandraの専門知識の習得と維持に苦労している

### 兆候1: ノードが制御不能なほど増え、TCO増大を懸念している

これは公然の事実ですが、Cassandraは水平方向の拡張性には優れていますが、リソースの使用はあまり効率的ではありません。これが、多くの本番環境においてユーザーが大規模なクラスタを必要としている理由です。ノード単位での課金のため、ベンダーにとっては好都合な話ですが、クラスタが大規模になればなるほど、稼働環境の監視、調整、管理の労力が増えるだけでなく、より頻繁にコンポーネントやハードウェアの障害が発生します。

Aerospikeが特許を取得しているHybrid Memory Architecture™ (HMA)は極めて効率的に動作するため、TCOをCassandraの十分の1程度まで削減できるケースが頻繁にあります。両データベースの利用費用を、データ量36TB、2年目に50%、3年目に33%の増加率を想定し、計3年間のコストを算出しました。このシナリオでは、読み取り/書き込みが混在する負荷を想定し、Micron SSDを搭載したDell R730dxサーバ費用、電力、冷却、設置スペース、人件費などのコストを検証しました。その結果、AerospikeのコストはCassandraに比べて90%低く、3年間で約927万ドルが節約されることが明らかになりました。

コスト比較: Cassandra vs. Aerospike

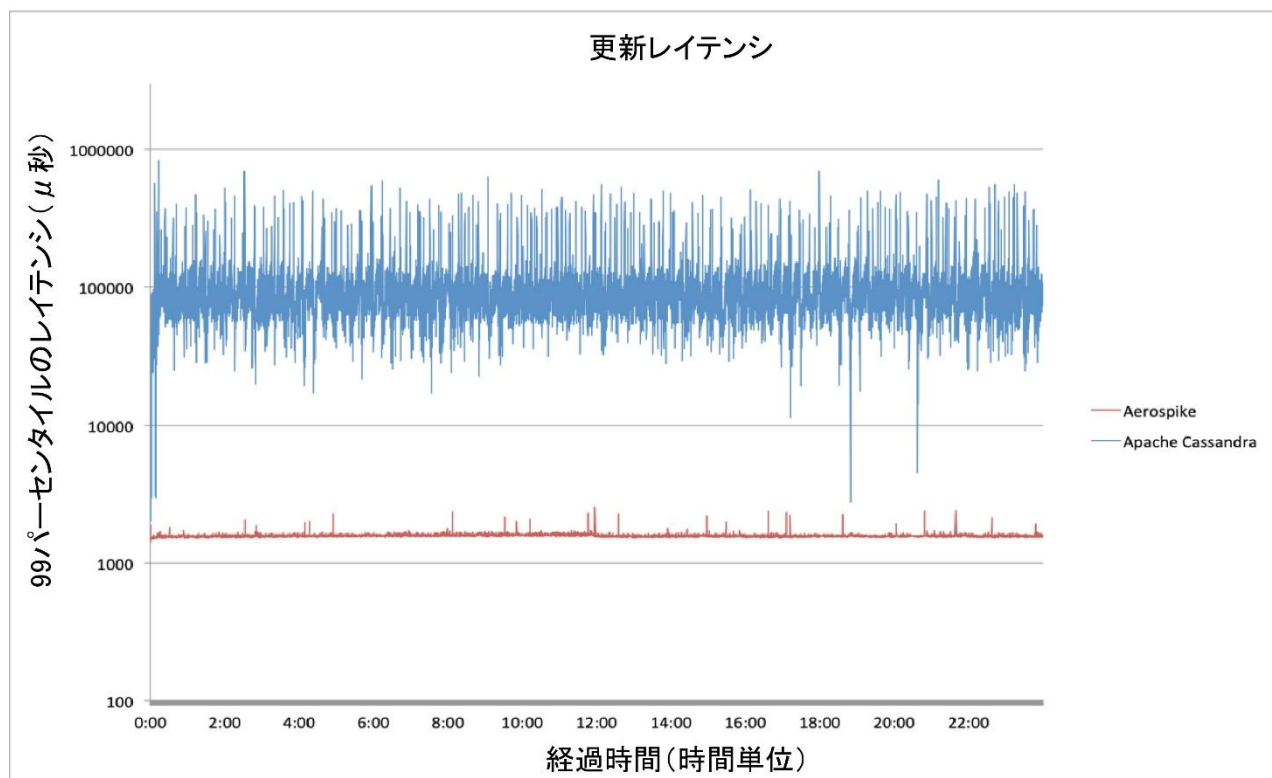
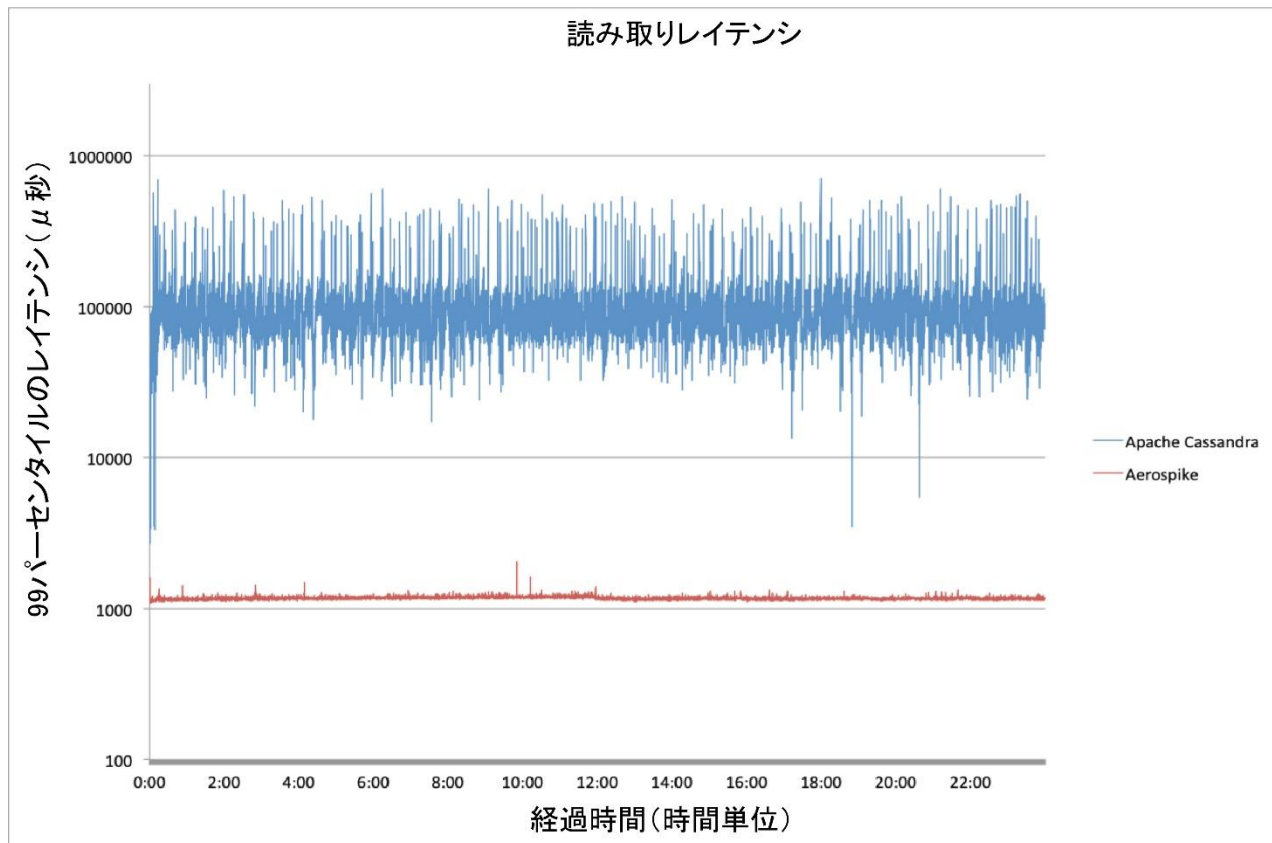


1年目は、200億件のユニーク・オブジェクト、レプリケーション係数は2、総データサイズは36TBになります。2年目にはデータが50%増加し、3年目にはさらに33%増加する想定です。Aerospikeによる節約額の合計は、927万ドルに達しています。

## 兆候2: 負荷ピーク時にSLAを満たせない

読み取り/書き込みが混在する負荷状況でCassandraを実行する場合、高速で一貫性のあるパフォーマンスを達成することが困難なケースが頻繁に発生します。ログ構造のマージツリー・ストレージ設計に基づきJavaで書かれたCassandraでは、パフォーマンスを最大化するために、Java仮想マシンの調整、ガベージ・コレクタの管理、圧縮の最適化、データモデルの慎重な構築をユーザー側で実施する必要があります。ミスがあれば、システムが不安定になり予期しないレイテンシが発生する可能性があります。

さらに、全てのCassandraノードは、対象となるデータを管理していなくても、読み取り/書き込み要求を受け入れることができるため、余分なネットワーク・トラフィックの発生を回避できません。これは、レイテンシの変動を引き起こします。実際、我々のベンチマークでは、Cassandraの読み取り/書き込みレイテンシが大きくばらつくことが明らかになりました。一方、Aerospikeのレイテンシは常に安定した予測可能な状態を持続し、応答時間もはるかに高速でした。



### 兆候3:カスケード障害にうんざりしている

Cassandraは、スムーズに動作させることが難しいケースがあまりに多く発生します。Cassandraユーザーが直面する問題として、連鎖するパフォーマンスの問題、ノードの障害、データの喪失などが挙げられ、これには次のような原因が考えられます。

- ・ 障害復旧中の暗示的な引継ぎによるメモリ圧迫。
- ・ 圧縮により発生する、キャッシュ、I/O、CPUの負荷超過。
- ・ メモリを使用により発生する頻繁なガベージ・コレクション、その結果引き起こされるCPUの負荷と高いテイル・レイテンシ。
- ・ メモリ負荷の原因となる多数のテーブル。
- ・ パーティションのサイズが適切でないことによる、Javaのヒープとガベージ・コレクションの圧迫。
- ・ 多くのブログ、プレゼンテーション、記事、ビデオで上記のような話題が取り上げられているのはこのためです。次から次へと障害が発生するのを防ごうとしたり、障害を修正しても次の障害が発生するだけだったりする状況に飽き飽きしているのであれば、他の選択肢を評価してみてもいいでしょうか。以前にCassandraを使用していたユーザーが既に気付いているように、Aerospikeはこのような問題を回避する設計となっています。

### 兆候4:運用チームの人員が増え続け、コストが増大し続けている

アクセス負荷とデータ量の増加に伴い、Cassandraクラスタは急速に拡張することが予想されます。極めて大規模なクラスタの管理は容易ではなく、コストも高くなります。そのような環境を最適設定するために、何が必要になるか考えてみましょう。あるアーキテクトは次のように述べています。「…Cassandraでは、使い始める前に多くの構成と調整作業が必要です。ではAerospikeはどうでしょうか？ 数点の変更だけで、即座に利用開始できますよ」<sup>1</sup>

また、極めて大規模なクラスタにおいて問題を診断し修正する場合、人手による膨大な作業が必要になります。あるCTOは次のように述べています。「Aerospikeの導入前は、Cassandraの管理に費やす時間ばかりが増え続け、新製品の開発に費やす時間は減る一方でした。Aerospikeでは…新しい機能を追加することに集中できています」<sup>2</sup>

### 兆候5:Cassandraの専門知識の習得と維持に苦勞している

技術的なリソースの獲得と維持は常に困難であり、Cassandraに精通した人材を雇うことも例外ではありません。最近の調査によると、たとえ適切な人材が見つかったとしても、他のDBMSのスキルを持つ人材に比べ、Cassandraの場合は15~20%高い報酬を支払う必要があることが明らかになっています。

また、負荷に合わせて異なるバージョンのCassandraを調整する必要があるかもしれません。さらに、代理店からCassandraを購入している場合はベンダーからパッチが提供されるまで待つか、新しいApacheリリースに移行するか（この場合、サポート契約は破棄しなければなりません）、現在のコードベースを修正するか、いずれかを選択する必要に迫られることとなります。これらの解決方法には、各々課題があります。

社内に専門知識を蓄えた人材を育成したり、独立したコンサルタントを雇ったりする時間や予算がないことも考えられます。Aerospikeでは、サーバの設置面積が小さく自動管理と自動修復の機能を備えたクラスタにより、スタッフを確保する必要性が大きく減少し、容易に需要を満たすことが可能になります。

<sup>1</sup> <https://youtu.be/ljWnG8dWTms?t=203>

<sup>2</sup> <https://youtu.be/tGPqvl3BwzY?t=126>

## Aerospikeで将来に備える

現在Cassandraを使用していて、要望の達成に苦労していたり、これまでに説明した5つの兆候のいずれかを経験していたりする場合は、Aerospikeに何ができるのかご確認いただくことをお勧めします。[info@aerospike.com](mailto:info@aerospike.com)までお問合せいただくか、「Cassandraが限界に達したことを示す5つの兆候」に関する詳細な[ホワイトペーパー](#)をご参照ください。

### Aerospikeについて

Aerospikeを活用する企業は、解決不可能にみえるデータによるボトルネックを解消し、インフラを極めてシンプルにすることにより、従来型のNoSQLデータベースの数分の1のコストで競争力を強化できます。Aerospikeが特許を取得しているHybrid Memory Architecture™により、最新のハードウェアが備えるポテンシャルを最大限に引き出し、比類なき競合優位性をもたらされます。その結果、エッジからコアやクラウドに存在する膨大なデータから、これまで想像もできなかった価値を創造します。Aerospikeは、不正行為への即座の対抗措置、ショッピングカートサイズの飛躍的な拡大、デジタルペイメントネットワークのグローバルなデプロイ、数百万人規模の顧客ひとり一人に向けた瞬時のパーソナライゼーションなど、幅広くお客様を支援します。

Aerospikeは、Airtel、Baidu、Banca d' Italia、Nielsen、PayPal、Snap、Verizon Media、Wayfairをはじめとする世界有数の企業に採用されています。米国カリフォルニア州マウンテンビューに本社を構え、ロンドン、バンガロール(インド)、オル・イェフダ(イスラエル)に拠点を展開しています。