

インターシステムズのベンチマークレポート:

InterSystems EMPI



リリース:
2025年11月

 **InterSystems**[®]
Creative data technology

InterSystems EMPI では、一括取り込みの持続的なスループットが平均で **5400 TPS(トランザクション毎秒)**を超えています。

エグゼクティブサマリー

InterSystems EMPI™は、次世代型のエンタープライズ・マスター・パーソン・インデックスであり、統合の容易な、ID 管理の自動化ソリューションです。スタンドアロンのソリューションとして利用することも、**InterSystems HealthShare Unified Care Record®**とシームレスに連携して、マルチソースデータを 1 つの正確な縦断的健康記録に統合することもできます。

インターシステムズは最近、最新バージョンの InterSystems EMPI のパフォーマンスと拡張性のベンチマークテストを実行しました。テストプロファイルは、**一括取り込みと人物検索**という 2 つの個別のワークロードで構成しました。テストでは、一括のデータ取り込み、大容量データ、高速の持続的 TPS(トランザクション毎秒)などを含む複数の主要領域に重点を置きました。このドキュメントでは、このテストから観察された特性について説明し、InterSystems EMPI の一般的な設定やサイズ設定のガイドラインについても解説します。

このベンチマークでシミュレーションするワークロードは一般的な製品導入にほぼ一致するように設計されており、最初に、Synthea(FHIR 向けの最も一般的な合成データ生成ツール)を使用して生成された個人属性データセット 3 億 800 万件の大規模な一括読み込みを行いました。データ取り込みの完了後、さまざまな人物検索ワークロードを混合して同時に実行し、スケールとパフォーマンスに加え、決定論的検索方法と確率論的検索方法の違いを実証しました。ワークロードシミュレーションの詳細については、このレポートの「ワークロードの説明と方法論」のセクションをご覧ください。

InterSystems EMPI の 2025.2 バージョンでは、一括取り込みの持続的なスループットが平均で **5400 TPS(トランザクション毎秒)**を超えています。このハイペースの取り込みにより、13.5 時間で 3 億 800 万件以上の人物データを読み込むことができました。人物検索の往復レイテンシの平均は、決定論的検索で 100 ミリ秒、確率論的検索でも 400 ミリ秒の低さです。**人物検索の応答時間の総平均は 263 ミリ秒**であり、全国レベルの大規模な人口ボリュームでも検索の応答時間は一貫して 1 秒以内であることが示されています。

ワークロードの説明と方法論

負荷生成ツールには、Apache JMeter が使用されました。JMeter は Java ベースのオープンソースツールで、静的/動的リソースの負荷とパフォーマンスのテストに使用されています。マルチスレッドのフレームワークで重い負荷のシミュレーションを行い、さまざまなプロトコルやシナリオでテストを実行し、スループットとレイテンシのレポートを生成することができます。

ベンチマークアーキテクチャの概略

ベンチマークテストは 3 つの JMeter ドライバサーバーで構成し、これらのサーバーと InterSystems EMPI インスタンス上の標準の RESTful API エンドポイントの間で通信を行いました。RESTful API エンドポイントは Apache HTTPD 2.4 Web サーバーを通じて提示され、InterSystems EMPI と通信するために InterSystems Web Gateway と統合されました。図 1 は、取り込みワークロードと検索ワークロードの両方の概略図です。

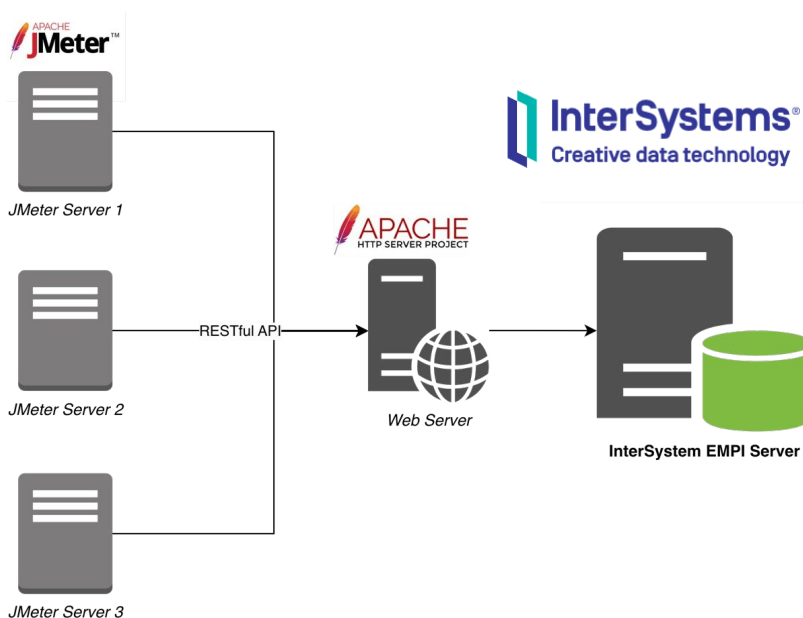


図 1: 取り込みワークロードと検索ワークロードの概略図

一括取り込み

Synthea を使用して、InterSystems EMPI に読み込まれる、3 億件を超える合成の人物データが生成されました。複数のレコードが生成され、テストに使用する 3 つの JMeter ドライバサーバーのそれぞれに配布されました。各 JMeter ドライバワークフローでは、InterSystems EMPI インスタンス上の **AddPerson(...)** の RESTful API エンドポイントを使用しました。取り込みプロセスは、データベースにレコードを挿入するだけのものではありません。取り込みプロセスの過程で、ID 照合エンジンは、リンク付けまたは一致の可能性に対するリンク付け、レビュー検証のためのレコードのマーキング、参照照合のための候補の特定などのオペレーションを実行します。

人物検索

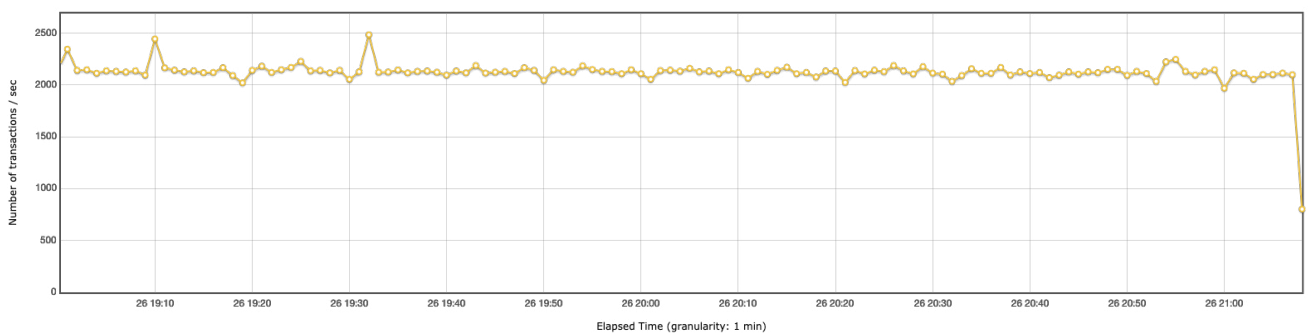
InterSystems EMPI は、決定論的アルゴリズムと確率論的アルゴリズムを活用してレコードを照合します。これらのアルゴリズムは、それぞれワークフローと複雑さが異なります。**PersonSearch(...)** のベンチマークワークロードでは、2 つの異なるタイプの検索アルゴリズム間で 50/50 の分割比を使用して、より現実的な混合ワークロードを示しました。

ベンチマークの結果

一括取り込みワークロード

前述したとおり、InterSystems EMPI に対する人物データの高速追加/更新をシミュレーションするために、3 つの JMeter ドライバサーバーが使用されました。各ワークロードドライバには処理対象の 1 億件以上の人物レコードが存在し、合計で 3 億 800 万件の人物レコードが InterSystems EMPI に読み込まれました。持続的なスループットの総平均は、最大 6400 TPS(ドライバあたり最大 2200)、平均レイテンシは 60 ミリ秒以下で成功率は 100%でした。下の図 2 のグラフは、いずれかのワークロードドライバによる持続的 TPS と予測可能なレイテンシの例を示しています。

TPS



レイテンシ

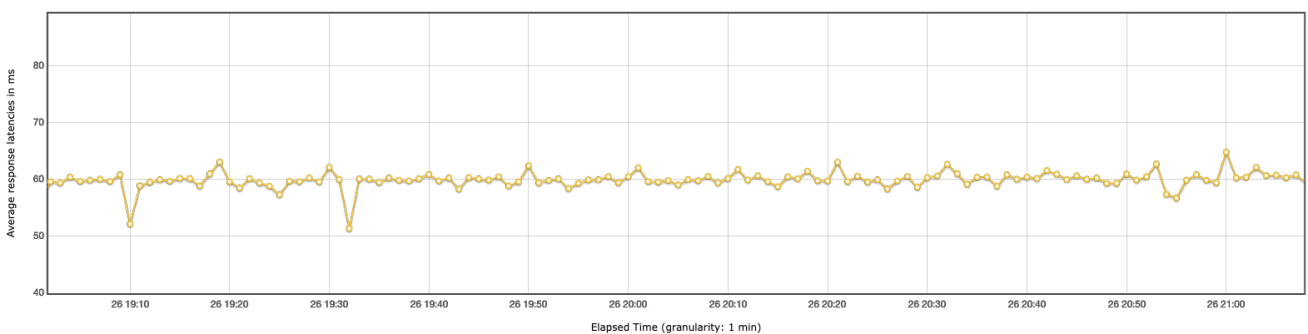


図 2: ワークロードドライバあたりの経時的な TPS とレイテンシ

取り込み中、サーバーの CPU 使用率は効率的で一貫していたため、取り込みの TPS 速度の線形スケールリングが可能になりました。図 3 は経時的な CPU 使用率を示しており、一括取り込みワークロードにはフラットな高い使用率が予想されることを示しています。

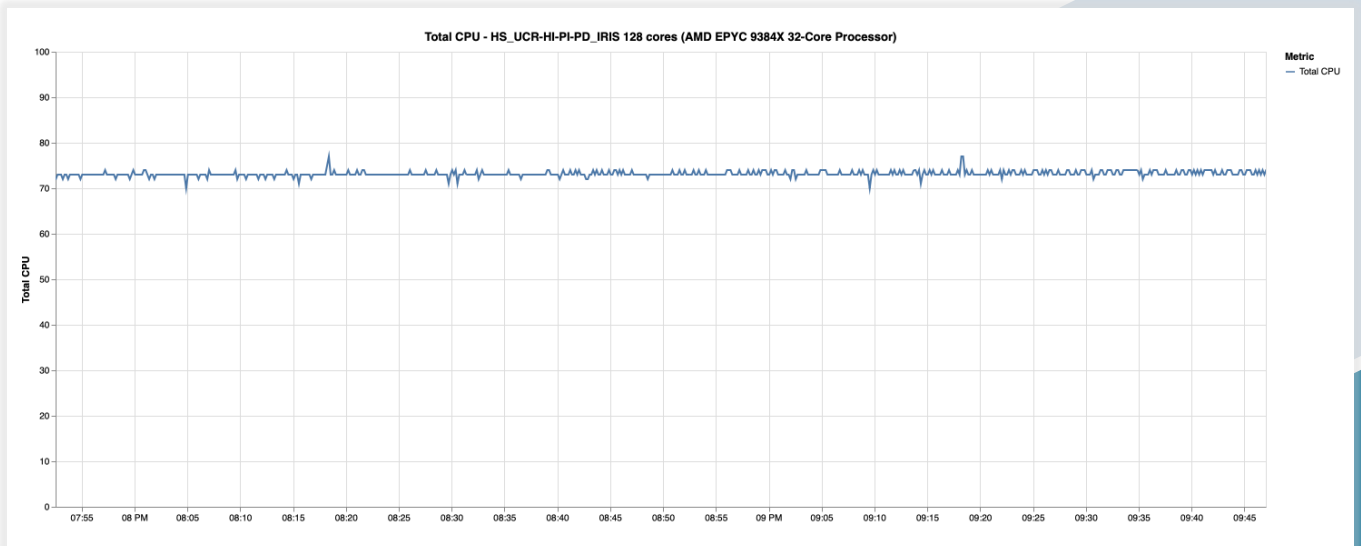


図 3: InterSystems EMPI サーバーの CPU 使用率 - 取り込み

人物検索ワークロード

高速の人物検索トランザクションのシミュレーションには、同じ 3 つのワークロードドライバが使用されました。決定論的アルゴリズムと確率論的アルゴリズムを 50/50 の分割比で使用して、より現実的な使用例を提供しています。この比率は使用例によって顧客ごとに異なるため、分割比 50/50 はこのベンチマークに対して公正な提示であると考えられます。

この人物検索テストでは、長期のランタイムを使用して経時的なパフォーマンスが示されました。これにより、アクティビティの最初の短い急上昇は単にデータベースキャッシュからデータを取り出しているだけではなく、実際にはデータベースの物理ディスクの読み出しを高速で行いながら、極めて低いレイテンシを維持していることがわかります。



図 4: 往復レイテンシ - 人物検索

上の図 4 のグラフに示すように、平均の往復レイテンシはおおよそ 4 時間のテスト実行期間にわたって 263 ミリ秒以下になっています。

この後の図 5 のグラフは物理ディスクの読み出し速度を示すもので、数時間にわたるテストランにおいても 1 秒あたりのディスク読み出しが平均で 80,000 件程度に維持されています。重要なのは、このメトリクスが経時的に示されており、高い TPS 速度と低い往復レイテンシが計測されたのは、データベースキャッシュ内に事前にキャッシュされたデータのみを対象とする短い期間ではないということです。

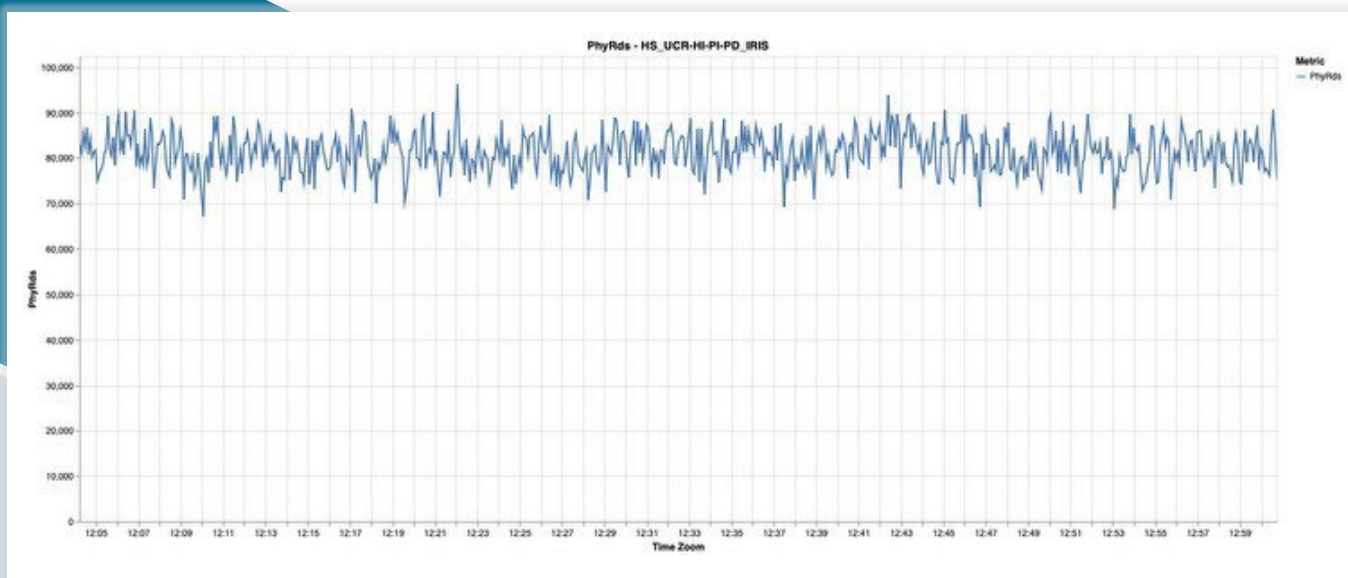


図5: 物理ディスクの読み出し - 人物検索

現実的な物理ディスクの読み出し速度に加えて、1秒あたりの実際のデータベースアクセス数(Glorefs)についても、1秒あたり800万回を超えるデータアクセス速度を維持しています。

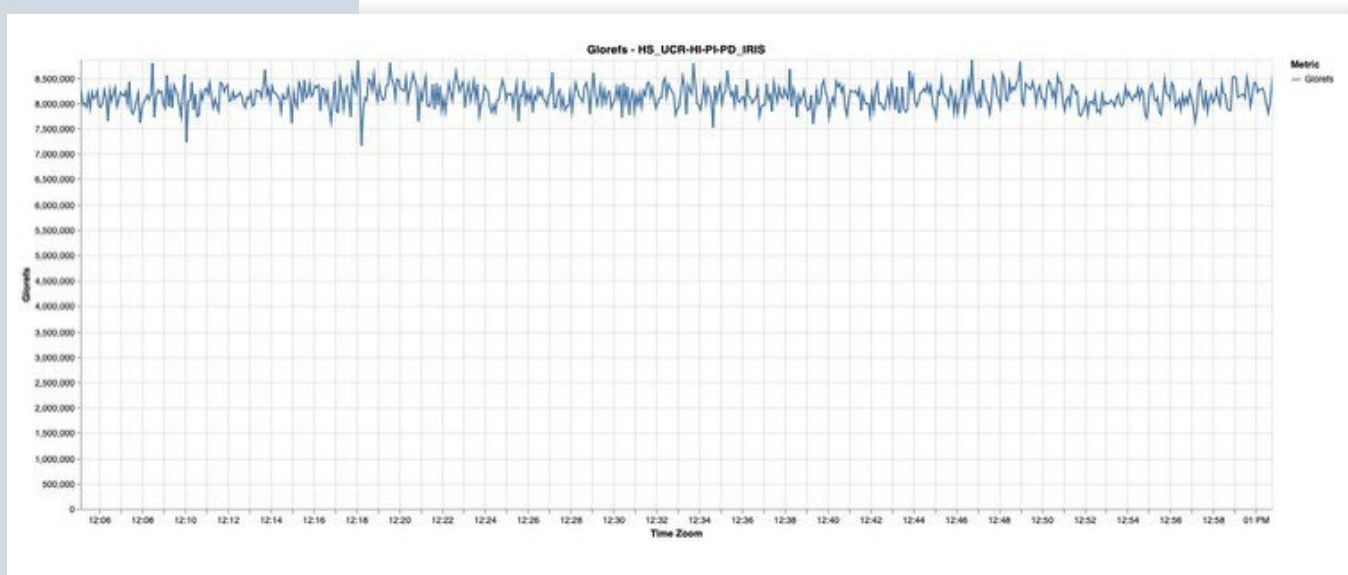


図6: 1秒あたりのデータベースアクセス数 - 人物検索

使用したハードウェア

サーバー構成

InterSystems EMPI のテストでは、2 基の AMD EPYC 9384X 32 コアプロセッサを使用しました。これは 3.1Ghz のベースクロック速度、1.5TB の RAM、10GbE のネットワークインターフェースで、合計 64 コア/128 スレッドを提供するものです。テストには Ubuntu 24.04.1 LTS オペレーティングシステムを使用しました。

ディスク構成

サーバーは 7.6TB のローカル接続された NVMe SSD を 6 つ活用しており、インターシステムズのベストプラクティスに沿って、InterSystems IRIS Data Platform データベースとトランザクションジャーナルボリュームが別々のディスクデバイスに分離されています。3 億 800 万件の人物データの取り込みワークロードが完了した後のデータベースのサイズは 622GB 程度でした。

サマリー

次の表は、ワークロードのスループットとレイテンシをまとめたものです。

ワークロード	ドライバスレッド数	データボリューム	JMeter ドライバの TPS (トランザクション毎秒)	平均レイテンシ (ミリ秒)
取り込み	384	人物データ 3 億 800 万件の一括読み込み	6400 TPS	60 ミリ秒
検索	192	毎分 44,000 回の検索	732 TPS	263 ミリ秒

まとめ

このベンチマークレポートは、InterSystems EMPI がエンタープライズレベルの ID 解決に対して卓越した拡張性とパフォーマンスをもたらすことを示しています。持続的な一括取り込み速度が 6,400 TPS を超え、3 億 800 万件のレコードを 14 時間以下で読み込む機能を備えていることは、このソリューションが国家規模のデプロイに対応可能であることの証です。人物検索ワークロードにおいて、長期で高ボリュームな条件下にあっても平均 263 ミリ秒のレイテンシを維持していることは、大規模データセット間で迅速かつ正確な ID マッチングを実現するプラットフォームの能力を証明しています。

これらの結果は、InterSystems EMPI がスピード、信頼性、効率性を主眼として設計されたことを裏付けています。これにより医療機関は、患者データを統合し、臨床試験や集団健康管理などのイニシアチブを大規模に促進できるようになります。

InterSystems EMPI の詳細は[こちらをクリック](#)してください。

