

# NY STATE DEPARTMENT OF HEALTH BRINGS MEDICAID ANALYTICS INTO THE FUTURE

*Double query performance for Medicaid Data Warehouse*

## CASE STUDY



Department  
of Health



Mellanox  
Technologies



The New York State Department of Health provides more than 6.6 million residents enrolled in Medicaid with access to essential medical care. Its central Medicaid Data Warehouse (MDW) was designed to facilitate the state's mission to enhance healthcare quality for its residents. Over 3,000 users simultaneously query its data stored in an Oracle RAC system using Oracle Business Intelligence - Enterprise Edition (OBIEE) to find the proverbial needle in the haystack of an insight that may help improve care delivery. With data from 10 billion Medicaid claims and encounters, MDW is massive and its growth and costs were becoming unwieldy.

To facilitate future growth, improve database query performance and lower overall TCO, the New York State Department of Health (NYSDOH) chose to modernize its MDW by deploying CMA's Mosaic SD NVMe, an NVMe-based scale-out storage architecture featuring Excelero's NVMesh® software and Mellanox Infini-band-based switches. The new DWH infrastructure enables NYSDOH to increase the percentage of ad hoc queries completed in under 1 minute to 98%, up from 50%. The new architecture provides freedom to scale out with cost-effective commodity hardware purchases from any vendor. This helps future-proof NYSDOH's storage architecture for the next sea change in technology, whatever it may be.

## Mosaic – NVMesh Benefits

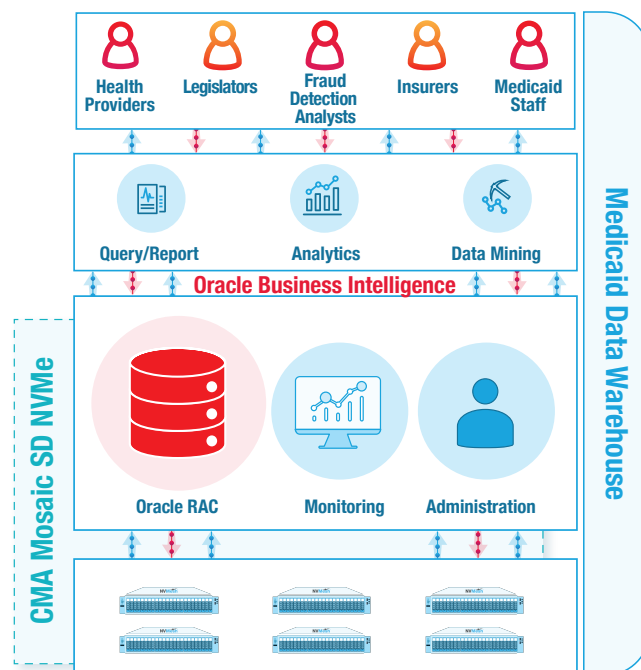
- **Spend less money on hardware:**  
Use standard Hardware Components, maximize NVMe efficiency
- **Lower the datacenter footprint:**  
Less rack space, less power and cooling costs
- **Reduce the software cost:**  
Cost-efficient SDS and fewer Oracle licenses
- **Simple management:**  
NVMesh comes with an intuitive interface, runs on standard Linux



## NEXT-GENERATION DATA WAREHOUSE FACILITATES BETTER HEALTHCARE

Debuted seven years ago, Medicaid Data Warehouse is the centerpiece of New York State's Medicaid redesign initiative to enhance healthcare quality and reduce costs across its \$70 billion Medicaid spending – the second largest in the US. The MDW is a revolutionary architecture, designed to enable faster and more insightful analytics by providing instant access to over a decade's worth of medical claims data, reports, and interactive dashboards that facilitate easier analytics. Over 3,000 simultaneous users, including healthcare providers, payers, legislators, fraud detection analysts, Federal Centers for Medicare and Medicaid Systems (CMS) staff, and insurers query its data to find the proverbial needle in the haystack of an insight that may help improve care delivery. MDW data is stored in an Oracle RAC system; Oracle Business Intelligence- Enterprise Edition (OBIEE) provides the analytics functionality.

The Medicaid Data Warehouse currently holds double the volume the first-generation MDW architecture was designed to handle. The main claim transactions fact table alone now requires 10 billion rows and 15 TB of storage. On top of that, ad hoc queries often include data from two other fact tables on claim diagnosis codes and claim procedures (services performed), which hold another 20 billion rows of data. As a result, ad hoc query response times were becoming an issue: just 50% of ad hoc queries could be completed and returned to end users in under 1 minute on the EMC VMAX-based system, a key metric of end user performance, with about 30% of ad hoc queries requiring over 10 minutes. To improve database query performance, facilitate further growth and lower the overall MDW TCO - including storage infrastructure, compute and Oracle licensing - the New York State Department of Health (NYS-DOH) chose to modernize its MDW by deploying Mosaic SD NVMe, a CMA-designed, NVMe-based, scale-out storage architecture featuring Excelero's NVMesh® software and Mellanox Infiniband-based switches.



## FUTURE-PROOFING THE NYS MEDICAID DATA WAREHOUSE

The CMA Mosaic solution deployed by NYSDOH features an Oracle RAC cluster with 8 Dell R740xd storage nodes running Excelero's NVMesh software, connected through Mellanox Infiniband-based switches. The new DWH infrastructure enables NYSDOH to increase the percentage of ad hoc queries completed in under 1 minute to 98%. 99% are completed within 3 minutes. The new architecture provides freedom to scale out with cost-effective commodity hardware purchases from any vendor. This helps future-proof NYSDOH's storage architecture for the next sea change in technology, whatever it may be.

# NY STATE DEPARTMENT OF HEALTH BRINGS MEDICAID ANALYTICS INTO THE FUTURE

*Double query performance for Medicaid Data Warehouse*

The new MDW architecture enables the pooling of NVMe Flash resources enterprise-wide, to support greater storage capacity utilization and cost-efficiency with health care analytics data that MDW retains from the 6 million residents enrolled in Medicaid. Because its storage architecture now is based on commodity-based servers and storage software, the NYSDOH future-proofed the MDW system so it can grow incrementally and affordably, with ready extensibility to support any next generation of storage technology without fork-lift upgrades. The new storage architecture proves organizations can achieve Oracle Exadata-class performance using a commonly deployed Oracle Enterprise Edition Real Application Clusters (RAC) software license on cost effective, readily sourced commodity hardware.

## PERFORMANCE TEST RESULTS

The CMA team assisting the NY State Department of Health scans about 1 petabyte of claims per day, or about 4 trillion claim rows. Looking at the results of a 1 hour sample timeframe:

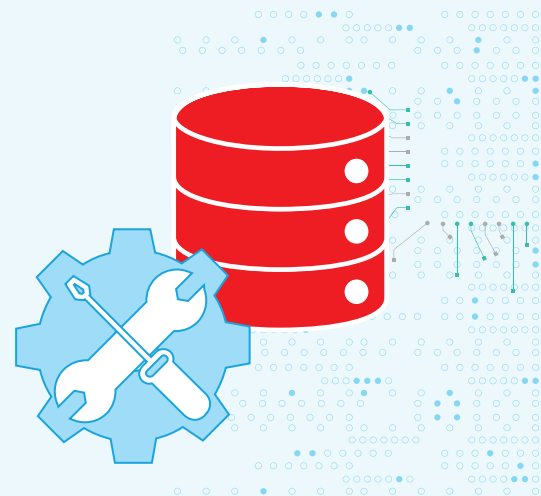
The average latency per read for an Oracle 8k block was 57 $\mu$ s. To put that in perspective, state-of-the-art NVMe devices can deliver 20 $\mu$ s reads for a 4k block. Multiplied by two that gives 40 $\mu$ s. This means Excelero's NVMesh adds only 17 $\mu$ s to access remote, pooled NVMe devices over the RDMA network all the way to the Oracle application.

## CMA MOSAIC FOR ORACLE RAC

CMA's Mosaic SD NVMe, is an end-to-end enterprise business analytics platform that integrates best-of-breed hardware and software technologies with CMA expertise to create a high-performance, fully-functional Oracle Real Application Cluster (RAC), a clustered version of Oracle Database based on a comprehensive high-availability stack. This packaged "data warehouse in a box" solution reduces integration risk, accelerates project delivery, and is scalable to meet customers' current and future reporting and analysis needs.

### THE SOLUTION IS BUILT WITH VERY SPECIFIC DESIGN GOALS TO MAXIMIZE CUSTOMERS' INFRASTRUCTURE ROI:

- **Enhance the database user experience**  
by enabling faster and predictable response times
- **Increase multi-tenancy capabilities**  
by allowing multiple analysts to submit simultaneous queries
- **Maximize database infrastructure flexibility**  
to support changing and increased workloads
- **Lower the Oracle RAC TCO**  
by reducing software licensing costs, optimizing hardware usage, and minimizing the datacenter footprint





Mosaic also enhances the database user-experience by providing predictable, high performance response times. Predictability is often even more important than performance: when an application delivers fast response times on some queries but slow response for others, this is considered as inconsistent performance. This hints at inherent architectural design flaws, which leverage and expect certain query plans. With Excelero's NVMesh, CMA designed a solution that delivers extreme IOPs per node, in a simple, lightweight, and very predictable platform. Likewise, abundant IO capability allows for increased multi-tenancy with larger numbers of simultaneous queries.

## NVMesh FOR THE NEXT-GENERATION DATA WAREHOUSE

NVMesh is a Software-Defined Block Storage solution that features Elastic NVMe, a distributed block layer that allows applications to utilize pooled NVMe storage devices across a network at local speeds and latencies. NVMe storage resources are pooled with the ability to create arbitrary, dynamic block volumes that can be utilized by any host running the NVMesh block client. In short, database applications can enjoy the latency, throughput, and IOPS of a local NVMe device while at the same time getting the benefits of centralized, redundant storage.

NVMesh is a 100% software solution, allowing customers to choose the hardware that best fits their requirements. NVMesh uses zero CPU on the target side, which generates cost savings in many different ways: customers can use standard servers with standard CPUs – of which they need less – and they need fewer Oracle licenses to achieve better performance results. NVMesh has a controller-less architecture, with intelligence on the client side. This takes away the controller as the traditional performance bottleneck and enables near 100% efficiency in scaling, without the need for forklift upgrades.

