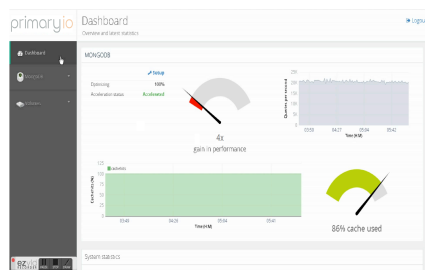# primaryio

PrimaryIO APA™

for Amazon Web Services (AWS)

Highlights

- Reduce AWS rental costs by 45%
- Improve Tier 1 database application performance by 15%
- Create no adverse effect on database semantics or transaction reliability
- Respond to changing business requirements with no business disruption
- Integrate seamlessly into existing AWS server and storage infrastructure
- Simple installation and management for application administrators and DBAs
- Support bare-metal, hypervisor or cloud deployment modes

"IT teams are increasingly leveraging AWS IaaS in order to meet users needs for faster delivery of IT services. The PrimaryIO APA software offers an innovative and cost-effective way for admins to accelerate performance of databases applications such as MongoDB that are deployed on AWS by leveraging lower cost EC2 instance stores to intelligently optimize IO and accelerate application performance in a very cost-effective manner."

Mark Peters, Senior Analyst
Enterprise Strategy Group

At PrimaryIO, we are performance optimization specialists. Our sole focus is accelerating the performance of database applications whether they are deployed on-premise or in the cloud in the most cost effective and least disruptive way imaginable in order to maximize business results.

Balancing Performance and Cost Challenges

Public cloud IaaS providers such as Amazon Web Services (AWS) offer enterprise grade infrastructure to run new age, distributed scale out databases such as MongoDB. While AWS can be cost effective for small deployments, rental costs can become quite expensive for production scale deployments especially for a cost conscious organization. In the case of MongoDB, most of the I/O is random. Therefore, if your working set is much larger than memory, with random access patterns you may need many thousands of IOPS from your storage layer to satisfy demand. AWS offers two broad choices for customers: Amazon Elastic Block Store (EBS) and Amazon EC2 instance store.

An Amazon AMI (Amazon Machine Image) for MongoDB can be provisioned to deliver up to 4,000 IOPS per EBS volume, but there are often very high reoccurring monthly costs associated with this option that can rise into the tens or hundreds of thousands of dollars for larger deployments. An instance store, on the other hand, provides temporary block-level storage, which depicts very good random IO property like flash and comes free with most of the latest Amazon instance types. The instance store can be efficiently leveraged for IO optimization to provide a better user experience, especially for data that is replicated across a fleet of instances, such as a MongoDB shared cluster.

A New Approach - PrimaryIO APA™

Traditional EBS on AWS has many drawbacks when it comes to accelerating I/O for an application like MongoDB. Since the AWS EBS infrastructure is not application aware, it writes more data than is necessary to optimize MongoDB performance resulting in inefficient flash utilization. Application Performance Acceleration (APA) technology from PrimaryIO solves these limitations by intelligently and dynamically writing only the most relevant business analytics data to the free instance store within the EC2 host servers.

The PrimaryIO APA technology is an application tier plug-in that integrates tightly with MongoDB. It identifies component I/O blocks such as collections, databases, indexes, etc., in the primary I/O stream that are important to tuning the performance of a MongoDB data store. The APA software intelligently caches only the latest incremental information within the instance store that will result in higher workload performance.

Industry's First Application Aware Solution
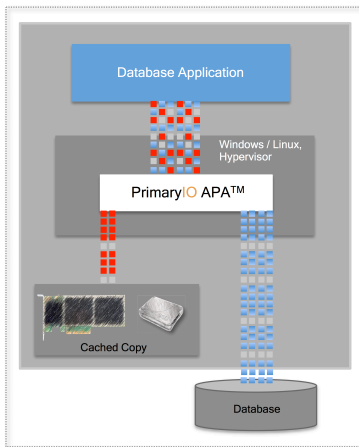
As illustrated in Figure 1, PrimaryIO APA is a software-only, hardware agnostic application acceleration solution for database applications. Regardless of whether the database is deployed on a Windows or Linux bare-metal server or hypervisor the PrimaryIO APA software is designed to improve performance to the highest level.

The PrimaryIO APA solution is simple to install and the GUI is intuitive and tailored for both system or application administrators. The PrimaryIO Application Performance Acceleration (APA) technology intelligently intercepts and caches a copy of the critical components of the incremental data set based on their ability to accelerate performance, such as frequently accessed collections or indexes that speed up queries. Application admins in particular benefit from easier management since they do not have to become system experts. Rather, they can optimize performance based on familiar database application components without the risk of adversely impacting existing server and storage infrastructure. The PrimaryIO solution requires no changes to the underlying system infrastructure and supports multiple cache modes including write-through to ensure transaction reliability.

Maximize Performance and Minimize TCO

Figure 2 below depicts a lab experiment comparing the performance and TCO of a MongoDB based AMI versus a MongoDB AWS deployment optimized with PrimaryIO APA. In this use case, MongoDB is used to run analytics on a web server log repository to get answers for various business intelligence questions. The data load phase includes importing the data into the MongoDB database while the run phase includes running analytics on the job on the latest inserted data set.

Figure 2
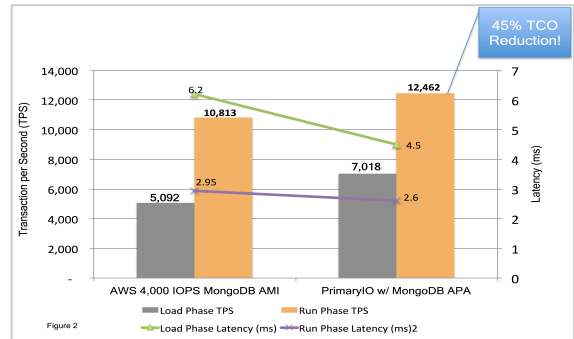
The results show that the PrimaryIO APA technology for MongoDB improved TPS of the load phase by 38% and latency by 27%, as well as the run phase TPS by 15% and latency by 12%.

What is even more compelling is that the PrimaryIO solution was able to reduce the monthly rental cost by per node by 45% through more intelligent and efficient utilization of the EC2 flash instance stores vs. relying entirely on the much more costly Amazon AMI for MongoDB provisioned to deliver up to 4,000 IOPS per EBS volume.

Product Specifications and System Requirements	
Components	PrimaryIO APA Core, Web GUI and APA plug-in for MongoDB, MySQL, Oracle and MS SQL Server
Cache Type	Fully Associative Cache
Cache Mode	Write-through, Write-back, Write-around, Read-around
Cache Size	Maximum 1 TB
Management	GUI, CLI
Operating System	64 bit x86 Linux, Microsoft Windows 2008/2012
Environment	Bare metal, VMWare, KVM, Xen, AWS/EC2
RAM	40MB minimum for optimal performance
Processor	Two dual-cores Processor 2.2 GHz or more
Supported SSDs	SSD supporting PCIe, SAS or SATA interface

About PrimaryIO

Big data, web and database applications are stressing the limits of traditional IT architectures because storage performance has not kept pace with compute capabilities. Eliminating this bottleneck has become a top priority for IT organizations globally. PrimaryIO provides software solutions to dramatically increase application performance by accelerating data I/O throughput. Our first product, PrimaryIO APA, is a caching software solution that non-intrusively increases application performance while avoiding costly server and storage upgrades. For further details on PrimaryIO APA or to download the product, please contact: info@PrimaryIO.com or visit www.PrimaryIO.com.