



The New Enterprise Blueprint



“By 2016, 30% of businesses will have begun directly or indirectly monetizing their information assets via bartering or selling them outright.”¹

Solix delivers the first unified platform for Enterprise Archiving, Enterprise Data Lake, and Analytics for all enterprise data.

Executive Summary

CIOs are in a bind. Demands for operational efficiencies and better business intelligence (BI) seem to be at odds, and CIOs are caught in the middle. They are expected to not only manage data growth and infrastructure costs effectively, but also provide modern tools to analyze and extract more value from their information assets. Now more than ever, CMOs, CFOs, and VPs of Sales are counting on the CIO to help them succeed.

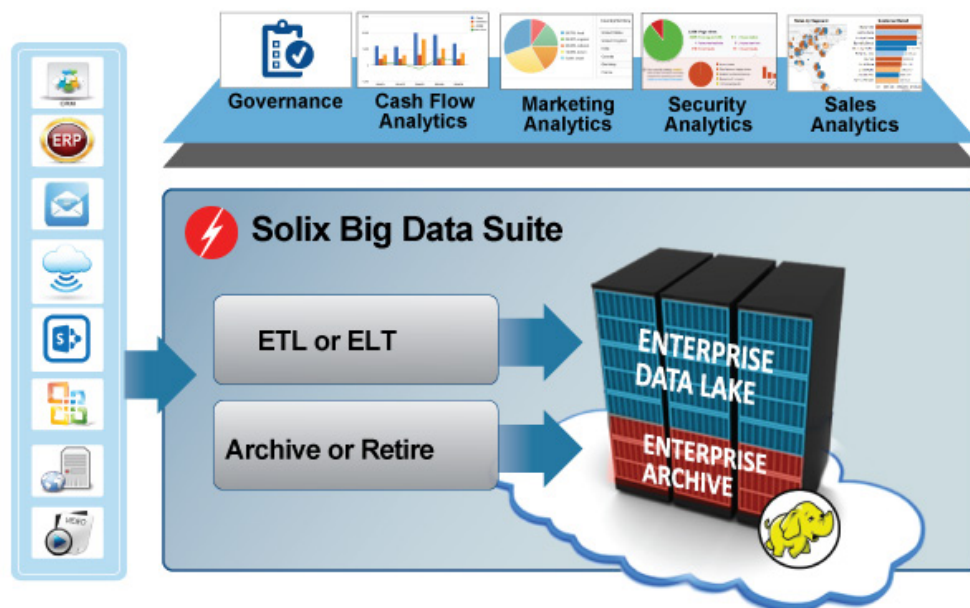
Gartner states that: “By 2016, 30% of businesses will have begun directly or indirectly monetizing their information assets via bartering or selling them outright.”¹ Other businesses simply want to mine enterprise information to improve effectiveness and bring more value to the business. Either way CIOs now hold the keys to organizational success and have gained responsibility for business growth. Caught in the bind are challenging requirements to control infrastructure costs, meet compliance objectives and also deliver better business intelligence. Process duplication, incremental hardware requirements and scarce resources

combine to threaten IT success. Few enterprise blueprints are capable to dramatically improve enterprise analytics and also reduce the cost of information assets.

Now more than ever, CMOs, CFOs, and VPs of Sales are counting on the CIO to help them succeed.

Solix Technologies has found a new way to deliver dramatically better business intelligence and still meet process, compliance, and infrastructure cost challenges. Apache Hadoop not only offers an exciting new framework for enterprise analytics, it also stores and processes large data sets at the lowest possible cost. According to a survey by Gartner, “big data investments in 2013 continued to rise, with 64% of organizations investing or planning to invest in big data technology compared with 58% in 2012.”²

FIGURE 1 Solix Big Data Suite



Source: Solix

^{1,2} “Big Data Benefits are Hampered by ‘Culture Clash,’” Gartner, September 12, 2013

Big data helps CIOs out of the bind, and we've built the Solix Big Data Suite to end competing interests between operational efficiency and BI. Built on Apache Hadoop, the Solix Big Data Suite is the first unified platform for Enterprise Archiving and Enterprise Data Lake to deliver analytics applications across the enterprise. Secure, compliant, low-cost, bulk storage of both structured and unstructured data is now available on a petabyte scale and capable to meet the most demanding requirements for next generation analytics.

Gartner states; "By 2016, 75% of structured data archiving applications will incorporate support for big data analytics."³

Enterprise Data Management solutions are changing rapidly and offering important new opportunities. IT costs may now be managed by the same enterprise blueprint that delivers improved BI. Analytics are the foundation upon which organizations are making business decisions to not only survive, but to thrive. Gartner believes BI has become a business asset in itself, and effective data management is at the heart of it all. The new enterprise blueprint manages all enterprise data at the lowest possible cost, improves operational efficiency and enables improved BI as well.

Source: Solix

³ "Magic Quadrant for Structured Data Archiving and Application Retirement," Gartner June 11, 2014

The Data Growth Crisis

CIOs are in a difficult position. The demand for operational efficiencies and the demand for Business Intelligence seem to be odds. Traditional thinking says the processes and infrastructure required to control costs, meet compliance objectives and keep costs in check do not mesh with the processes and structures needed for BI. Duplication of effort, hardware and operational costs could cripple IT.

CIOs must change the paradigm to ensure organizations can compete in today's economic climate as structured and unstructured data continue to explode. The time has come for a new enterprise blueprint. The tools are here. Apache Hadoop makes it possible. Solix Big Data Suite makes it a reality.

The Paradox of Moore's Law

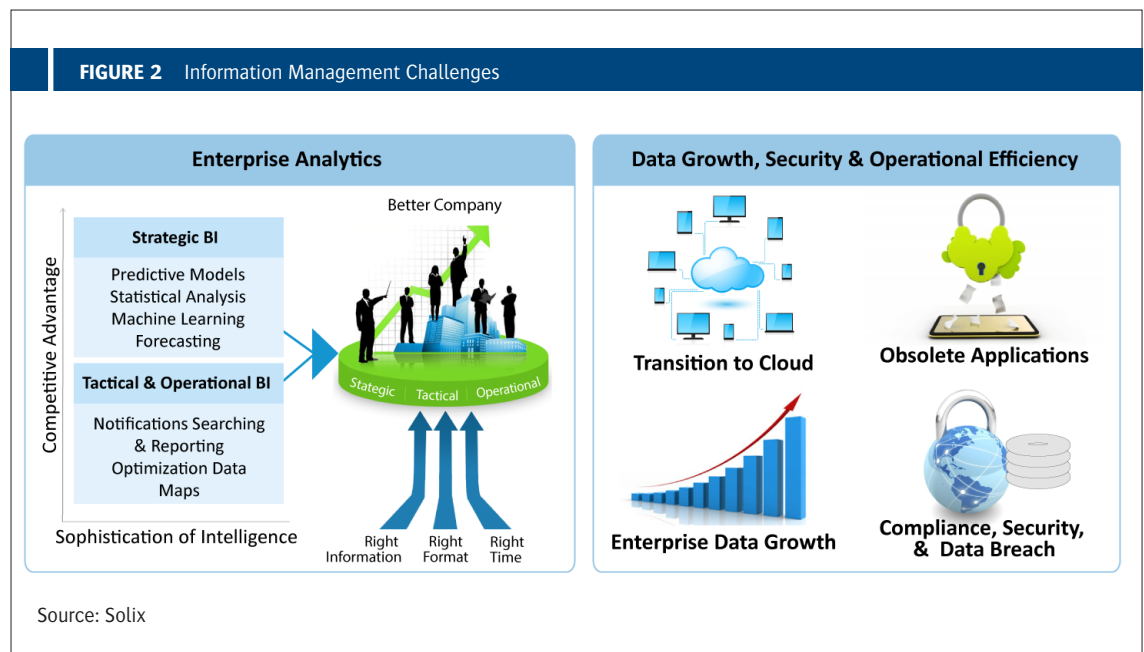
While data growth explodes exponentially, Moore's Law continues to astound as processor and integrated circuit performance doubles every few years, just as the visionary Intel founder predicted. The benefits of such dramatic technological advances cannot be overstated: Today, more processing power is packed into a smartphone than yesterday's mainframe, and the cost to store a terabyte of data in the cloud has fallen to as low as \$10 per month.

But rather than saving, organizations are spending more to gain improved business value by processing mission-critical enterprise data faster and faster. Until recently, few ERP users would ever have imagined paying the high cost to process enterprise data using full flash arrays, yet today, most have either already installed or are evaluating premium performance enterprise platforms with solid state disks (SSD) such as SAP Hana and Oracle Exadata.

Despite such spectacular gains in semiconductor price/performance, the overall cost of IT continues to rise, largely because we are now processing so much more data. Challenged by this dilemma, CIOs must continually find new ways to reduce the cost of data growth, so they may afford to fund more mission-critical applications that improve business results.

A New Enterprise Blueprint

Advances in semiconductor technology have indeed enabled "commodity" hardware to process and store extraordinary amounts of data at lower unit costs. Through virtualization, this low-cost infrastructure may now be utilized with extraordinary efficiency.



Apache Hadoop is a framework for distributed processing and storage of large data sets on virtualized commodity hardware. Hadoop has been designed to provide analytics over petabyte-scale data sets, and the infrastructure can be leveraged for long-term storage and analytics.

The Hadoop Distributed File System (HDFS) has rapidly emerged as the leading storage platform because it provides secure, stable storage for structured and unstructured enterprise data with enhanced access. Hadoop's MapReduce framework can process large data sets across distributed compute nodes in parallel, allowing "commodity" hardware to be used as the most efficient and cost-effective bulk data storage solution available.

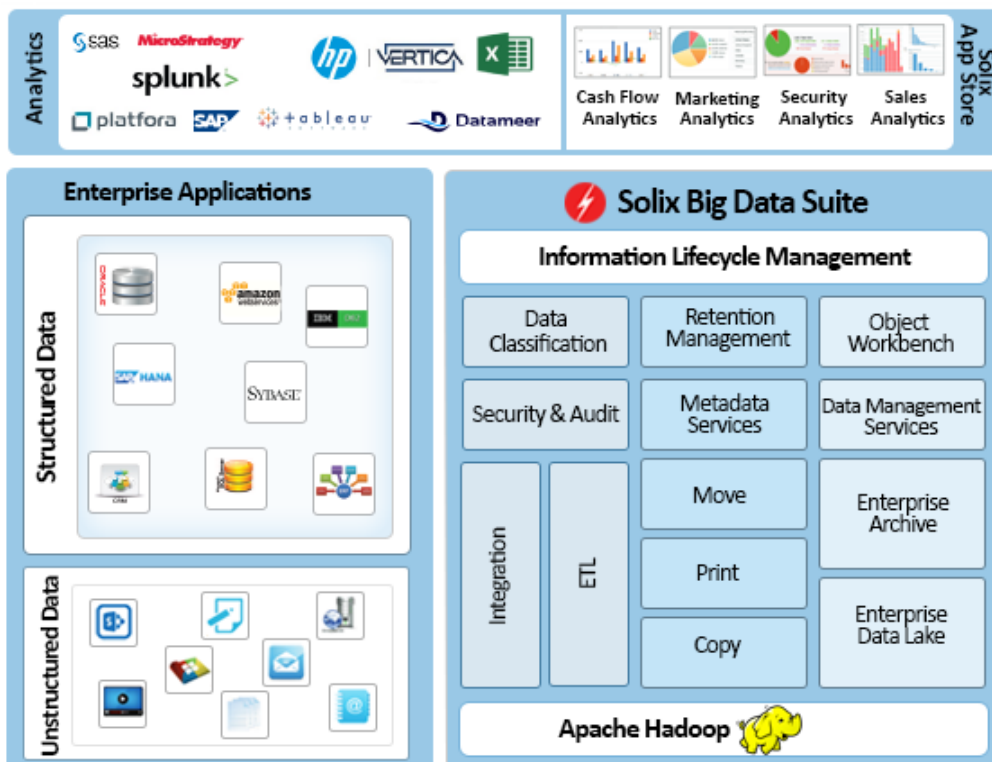
Experts agree that as much as 80 percent of production data in ERP, CRM, file servers and other mission-critical applications may not be in active use, and both structured and unstructured data becomes less active as they age. Large amounts of inactive data stored online for too long reduces the performance of production applications, increases costs and creates compliance challenges.

Apache Hadoop is a better alternative to storing inactive enterprise data online. By moving inactive data to nearline storage, application performance is improved and costs are reduced because data sets are smaller and workloads are more manageable. Universal data access is maintained by analytics applications, structured query and reporting, or just simple text search.

This new enterprise blueprint reduces infrastructure costs and enables organizations to gain improved value from their data. Enterprise data warehouse (EDW) and analytics applications leverage big data for better described views of critical information. As a low-cost data repository to store copies of enterprise data, big data is an ideal platform to stage critical enterprise data for later use by EDW and analytics applications.

Source: Solix

FIGURE 3 The New Enterprise Blue Print



Source: Solix

Solix Big Data Suite provides an Information Lifecycle Management (ILM) continuum that ensures an efficient and compliant extract transform and load (ETL) process.

Solix Big Data Suite — The New ILM Continuum

Capture. Organize. Analyze.

The Solix Big Data Suite delivers a new enterprise blueprint to help CIOs drive down infrastructure costs and make BI flourish, all while ensuring data is safe and secure. Solix Big Data Suite provides an Information Lifecycle Management (ILM) continuum that ensures an efficient and compliant extract transform and load (ETL) process. CIOs no longer need to choose between compliance, efficient processes and BI.

ILM Best-Practice

Gartner says “any organization thinking of simply applying existing information governance practices to big data will likely fail — not least because much data is ungoverned; or governed by others according to a different set of objectives.”⁴

Solix Big Data Suite provides the first true ILM continuum to address the complexity of data governance as data is copied and moved between production and Big Data. The ILM framework not only manages the ETL process, it also manages the data within HDFS and HBASE and provides an integrated retention management and legal-hold

capability for Apache Hadoop. Data from structured and unstructured sources is classified, migrated into HDFS/HBASE and validated with audit reports. These reports provide defensibility and chain of custody for compliance and data governance.

Solix’s ingestion engine captures application metadata and integrates with enterprise authentication systems, such as Active Directory, LDAP and others to establish role-based security. Solix provides a single user interface to create a seamless user experience for administrators and end-users. Through role-based-security, the Solix interface allows users to access different facets of the product. For example, a compliance officer could be given access to create retention policies, apply legal-holds or generate reports on cases. End-users, however, might only be allowed to perform keyword-searches on projects assigned to them.

This extensive ILM framework allows Solix Big Data Suite to create a unified repository to **capture** all enterprise data and optimally **organize** it for **analytics** tools offered through the Solix App Store.

FIGURE 4 Solix Big Data Suite - Advantages of Apache Hadoop

| | |
|----------------------|--|
| Scalability | Supports petabyte-scale enterprise data |
| Cost | Built on low-cost commodity storage |
| Analytics | 1000s of tools to unlock value of enterprise data |
| Modern | Supports petabyte-scale enterprise data |
| Comprehensive | Supports structured, unstructured, social data, etc. |

⁴“Big Data Governance From Truth to Trust,” published July 18, 2013

The highly scalable Solix Big Data Suite offers an extensible connector framework to ingest all forms of enterprise data from any source. Archiving, application retirement and flexible ETL capabilities improve the speed of deployment, reduce costs and optimize available infrastructure. Solix supports on-premise and cloud-based deployment on a variety of Hadoop distributions.

The Solix Big Data Suite harnesses the capabilities of Hadoop to create a comprehensive and efficient platform that produces unified and cost-effective ETL, ILM and BI infrastructures for all data, requiring smaller teams with fewer IT skills, while allowing quicker rollouts and faster results.

The Solix Big Data Suite includes:

- Solix Enterprise Archiving to improve enterprise application performance and reduce infrastructure costs. Enterprise application data running online is first moved, and then purged from its source location according to ILM policies to ensure governance, risk and compliance objectives are met.
- Solix Enterprise Data Lake reduces the complexity and processing burden to stage enterprise data warehouse (EDW) and analytics applications, and provides highly efficient, bulk storage of enterprise data for later use when it is needed. Solix Data Lake provides a copy of production data and stores it “as is” in bulk for later use.
- Solix App Store offers pre-integrated analytics tools for data within the Enterprise Archiving and Enterprise Data Lake.

Source: Solix

FIGURE 5 Information Lifecycle Management Framework



Source: Solix

Experts estimate that up to 40 percent of applications are candidates for retirement, migration or rationalization.

Solix Big Data Suite — Enterprise Archiving

Solix Enterprise Archiving offers a unified ILM framework to archive and retire all enterprise data — structured and unstructured — to optimize application performance, reduce cost and achieve compliance goals. Solix's ILM framework meets all the necessary aspects of governance, including legal hold, retention management, eDiscovery, and auditing. Solix takes governance to the next step by including validation of all archiving and retirement actions taken within the platform.

The Enterprise Archive has two components, Data Archiving and Application Retirement.

Data Archiving

Database archiving has emerged as a key component to an ILM best-practice framework to meet data growth challenges.

Data archiving best-practice requires that move and purge processes be coordinated and validated. All data deletions are done in a defensible manner with audit records and chain of custody. Solix Enterprise Archiving ensures proper data governance since enterprise data is ingested and stored based on retention management policies with support for custom business rules. Archive data is classified for security and compliance requirements, such as legal hold, and universal access is provided for business users through structured reports and full text search for business objects.

Application Retirement

Retiring legacy applications creates instant cost savings by freeing staff for more important assignments, eliminating licensing and maintenance fees, and enabling IT to shut down unnecessary hardware saving power, floor space and air conditioning load.

Retired data — both structured and unstructured — is compressed by as much as 90 percent and stored at the lowest possible cost in an immutable format which is easily accessed by users. Yet, active data can be migrated to the production database and accessed by new core applications. Hundreds of legacy applications, in a variety of configurations and on different platforms, can be decommissioned.

Print and Purge for Database Archiving and Application Retirement

Solix Print Report Archiving offers the capability to generate printed reports from native applications in readable formats such as PDF or Text for archiving or application retirement. The tool allows import of previously created reports from folders. Users can access these reports for compliance, eDiscovery and other regulatory requirements through Solix's search and reporting interfaces.

Solix Enterprise Archive Benefits

The features of database archiving and application retirement ensure that Solix Enterprise Archive meets ILM best practice for both structured and unstructured data. All move and purge processes can be automated based on organization policies and schedules. Only Solix Big Data Suite ensures all move and purge actions are validated to ensure the proper data has been processed.

User access to the data occurs through reporting and search tools utilizing metadata tags and keywords. Furthermore, archived data is classified for security and compliance requirements and access is role-based and predicated on clearance.

Solix Enterprise Archive also benefits organizations through:

- Improved application performance by reducing the amount of production data core applications must process.
- Smaller pools of active data means faster backups, minimized downtime during upgrades and faster return time for disaster recovery. Cloning time will also be reduced.
- Reduced operational costs and the ability to reallocate staff time to more mission-critical operations.
- Decommissioning legacy applications, which allows for the elimination of infrastructure, maintenance, and support costs.

Source: Solix

Solix Big Data Suite — Enterprise Data Lake

Solix Enterprise Data Lake is a repository to integrate all formats of data from structured and unstructured sources. The Solix Enterprise Data Lake offers data integration and ETL capabilities on a low-cost, highly scalable platform.

Traditional EDW platforms deliver highly specific data views based on corporate strategy. This canonical top-down enterprise approach is deductive and has merits, yet it does not allow for the inductive analytical style that drives many BI proponents. The fact is, the volume, velocity and variety of data in organizations is simply too much for the human mind to grasp. Yet, Gartner says these two approaches are locked in a “culture clash” that must be resolved.⁵

The Solix Enterprise Data Lake reduces the complexity and processing burden to stage EDW and analytics applications, and it provides highly efficient, bulk storage of enterprise data for later use when it is needed. Proponents of both deductive and inductive analytics can be satisfied.

With the Solix Enterprise Data Lake, the data is stored on “commodity” servers that are far less expensive than the Tier 1 infrastructure enterprise data traditionally utilizes. These commodity servers

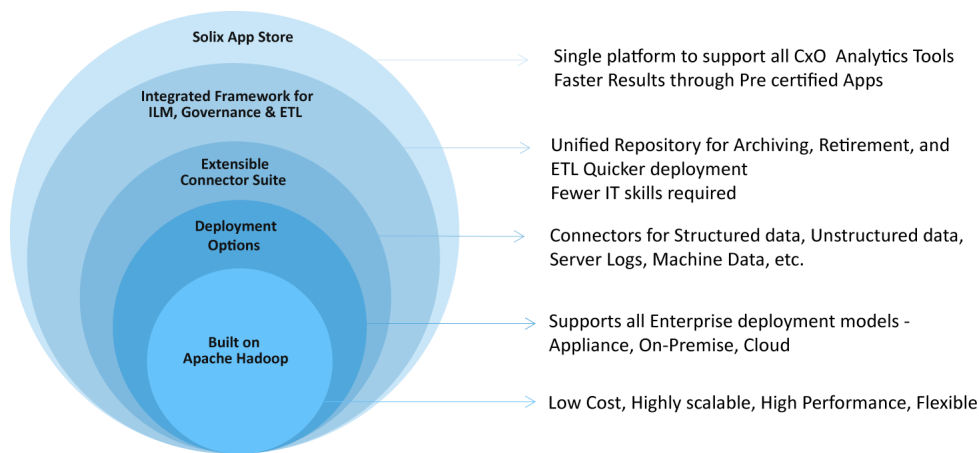
can save organizations more than 80 percent on infrastructure costs alone. Add in the reduced expense for licenses and utility costs to keep Tier 1 infrastructures running and the Data Lake pays for itself quickly.

The Solix Enterprise Data Lake provides a copy of production data and stores it “as is” in bulk to be better described and distilled later. This simple copy process eliminates the need for heavy ETL processing during ingestion. Once resident within the Hadoop file system (HDFS), enterprise data can be better described or transformed for use with business analytics applications based on user need, which allows organizations to get the most out of the data.

Utilizing the Solix Enterprise Data Lake, organizations can have access to larger pools of data to drive business decisions at far lower cost. Simple tools could leverage server and machine logs to predict system failures, alerts and the like. However, organizations need to mine data at a far more strategic level. Sophisticated BI tools can become key decision drivers and bring

Utilizing the Solix Enterprise Data Lake, organizations can have access to larger pools of data to drive business decisions at far lower cost.

FIGURE 6 Solix Big Data Suite - Key Benefits



Source: Solix

significant value to impact the revenue and growth of the organization. Solix Enterprise Data Lake also minimizes the “culture clash” with BI users where access to large quantities of data once considered unmanageable is now possible to drive deeper levels of analysis without sacrificing overall organizational needs.

Solix Enterprise Data Lake also offers the same extensive Solix ILM framework to ensure data retention and classification policies are deployed and data meets governance, risk and compliance objectives. Gartner analysts say “today’s

information governance practices focus on controlling various information quality attributes, typically for information governed behind the firewall. Organizations are either following the same approach with big data (which is proving impractical) or looking for new ways to govern that data to reduce risks.”^{5,6}

Solix Big Data Suite creates the new standard for governance in a Big Data world. ILM policies and business rules may be pre-configured to meet industry standard compliance objectives, such as COBIT, or custom designed to meet more specific requirements.

Source: Solix

^{5,6}“Big Data Governance From Truth to Trust,” published July 18, 2013

Solix App Store

The Solix App Store makes inductive BI user-friendly. The App Store offers out-of-the-box analytics through pre-integrated applications and offers the opportunity to utilize third-party apps. Organizations can also build custom and mobile apps through the App Store.

The App Store offers:

- **Cash-flow analytics**
- **Sales/pipeline analytics**
- **Marketing and social analytics**
- **Integration with Tableau, Splunk, etc.**
- **An executive dashboard**
- **Ad-hoc reporting and searching tools**

Connector Framework

Solix Big Data Suite has an extensive connector framework that allows the ingestion of data from all sources. The framework can take heterogenous data streams and allow the organization to manage the data according to business policies.

The connector framework encompasses:

- **Structured Data:** Core enterprise applications including Oracle E-Business Suite, PeopleSoft, Baan, JD Edwards, Siebel and custom applications
- **Unstructured Data:** MicroSoft Sharepoint, file systems, web servers, desktops, laptops etc.
- **Social feeds:** Including facebook, twitter, Instagram
- **Log files:** Security Logs, Network Server Logs, Mobile Logs, etc.
- **Print stream:** Native Report Archiving, Fax/Scan Data, etc.

The various streams of data are brought together through Solix Data Integration Studio, which provides a comprehensive set of ETL capabilities to ingest any enterprise data into the Solix Big Data Suite.

Source: Solix

Conclusion

Solix Big Data Suite brings the new enterprise blueprint within the grasp of every CIO.

With Solix Enterprise Archiving and Solix Enterprise Data Lake, CIOs find a path to start the enterprise Hadoop journey without having to choose between operational efficiency or BI. Solix Big Data Suite manages the entire big data lifecycle without sacrificing governance, compliance or performance.

Solix supports a variety of commercial Apache Hadoop distributions including Cloudera and Hortonworks. Deployment methodologies include a hardware appliance and public/private/hybrid cloud to fit into all enterprises.

The challenge to exploit Apache Hadoop technologies has been solved. Solix has the products and expertise to complement CIOs and their teams to leverage emerging Apache big data technologies and stay ahead of their competition.

CIOs are in a bind no longer.

Case Studies

Enterprise Archiving Customer:

Publically Traded Telecom Company

Business Objectives:

The customer has a fast growing Siebel application. The performance and cost of managing the environment is growing year-over-year. Customer would like to reduce the size of the production database through archiving, maintain access to the data, and apply retention policies to the data.

Solix Solution & Benefits:

Solix has archived more than 50 months of data, amounting to more than 19 Billion records. Due to the high transaction volume of the telecom customer, Solix archives at a rate of 300,000 records per minute. With Solix's archiving solution, the customer has recovered a significant portion of their expensive Tier-1 storage, along with improved performance. Cost of operations have been reduced by including infrastructure and support costs. Solix's Application Portfolio Manager (APM) provides a reporting and search capabilities for the data with full ILM capability for compliance and regulatory purposes.

Enterprise Archiving Customer:

Fortune 100 Technology and Manufacturing Company

Business Objectives:

The customer has more than 50 legacy enterprise applications, including JD Edwards, which are eligible for decommissioning. The primary objectives are to reduce IT cost, centralize retired data, and maintain regulatory compliance for data retention.

Solix Solution & Benefits:

Solix Enterprise Archiving platform was leveraged to migrate and consolidate data from all the enterprise applications into a single repository for centralized management. The archive provides retention and legal hold capability to address governance requirements. Additionally, for business users, an integrated user interface provides seamless role-based access through searches and reports. With Solix's solution, the customer is projected to save \$12 Million over 3 years.

Other notable use cases of the Solix Big Data Suite

- **Enterprise Data Lake** - One of Solix's manufacturing customers is deploying the Solix Big Data Suite to aggregate security log files from heterogeneous data sources to analyze them for unauthorized access to intellectual property and prevention of IP theft.
- **Enterprise Archiving** - A Fortune 500 distributor of building supplies is leveraging Solix Big Data Suite to archive SharePoint sites and consolidate documents into a unified repository for easier access and also retire an ERP application.

About Us

Solix Technologies, Inc., the leading provider of Enterprise Data Management (EDM) solutions, is transforming information management with the first enterprise archiving and data lake application suite for big data: The Solix Big Data Suite. Solix is helping organizations learn more from their data with enterprise analytics and achieve Information Lifecycle Management (ILM) goals. The Solix Enterprise Data Management Suite (Solix EDMS) and Solix Enterprise Standard Edition (SE) enable organizations to improve application performance, meet compliance objectives and reduce the cost of data management across the enterprise. Solix Technologies, Inc. is headquartered in Santa Clara, California and operates worldwide through an established network of value added resellers (VARs) and systems integrators.



Biography

Vikram Gaitonde heads the technology leadership and product strategy for Solix as Vice President of Products. Vikram is responsible for product management and engineering for Solix's next generation information governance solutions. Mr. Vikram brings 15 years of product management, engineering, and software delivery experience and has held several leadership positions within technology companies. Prior to Solix he was the Director of Enterprise Solutions at EMC, where he led the product solutions and product management teams for the information intelligence, data governance, eDiscovery, and content management products. Additionally, he has held numerous technical and product leadership positions at Silicon Valley startups and industry giants including Hotmail/Microsoft, Wipro, Kazeon (acquired by EMC), and Cast Iron Systems (acquired by IBM). Mr. Vikram holds an MBA from The Wharton School of the University of Pennsylvania and a B.S. in Computer Science from the University of Pune, India.