REFERENCE ARCHITECTURE

Financial Services
Data Management and Analytics

Simplify and Accelerate Data Science at Scale with Bodo* and Intel

A unified analytics environment based on Python* and Bodo Engine* technology can accelerate analytics workflows, simplify infrastructure, and improve efficiency significantly.

What You'll Find in This Solution Reference Architecture

This solution provides a starting point for developing a data science infrastructure.

If you are responsible for:

- **Investment decisions and business strategy...**

  You'll learn how a unified data platform can help solve the pressing analytics challenges facing enterprises today.

- **Figuring out how to implement predictive analytics...**

  You'll learn about the architecture components and how they work together to create a cohesive business solution.

Executive Summary

The financial services industry has been at the forefront of using new technology to solve business challenges and exploring new ways of generating revenue. Many firms are embracing artificial intelligence, especially now with the advent of new generation of high performance compute power, network connectivity, new modeling techniques and readily available cloud infrastructure that is cheaper and easily accessible.

The new power of AI is changing the business landscape. Institutions are looking to artificial intelligence to create business value, increase efficiency, gain competitive advantage and establish a position in the market. Leading organizations are deepening their investment in AI and moving beyond isolated AI use cases toward an enterprise-wide program. However, they are discovering the challenges inherent in the scaling effort such as creating a strategic vision, assessing current capabilities, building AI supporting business processes and capable infrastructure, addressing data pipeline and model development, operational management, governance, ethics and integrating AI understanding into the business. While successful AI projects rely on many factors, in this paper we will focus on scalable AI infrastructure based on a framework that allows AI analytic models to be developed and deployed at scale using Python* programming language only.

This is done through compiling Python into binary code (machine executable) and using high performance computing (HPC) techniques so that the code is deployed onto parallel processing cores. In other words, data scientists develop their AI models in Python and will be able to deploy to production seamlessly and automatically without any need to re-write or re-engineer the code. This has many advantages per below:

1 - Lower development costs, fast and error free deployment and easy revisioning

2 - Enables scalable production with orders of magnitude performance improvement

3 - Allows near real-time analytics for time sensitive applications

4 - Enables AI deployment both in the cloud and at the edge

Introduction

Accelerated digital transformation is bringing exponentially larger amounts of data every year, but extracting business value from this data is challenging. Although the introduction of advanced analytics and AI techniques has enabled new data science methods, there are significant barriers for enterprise adoption which includes but is not limited to data quality, unproductive data science, quality of insights, complexity and redundancy; and performance and scalability as depicted in the Figure 1.

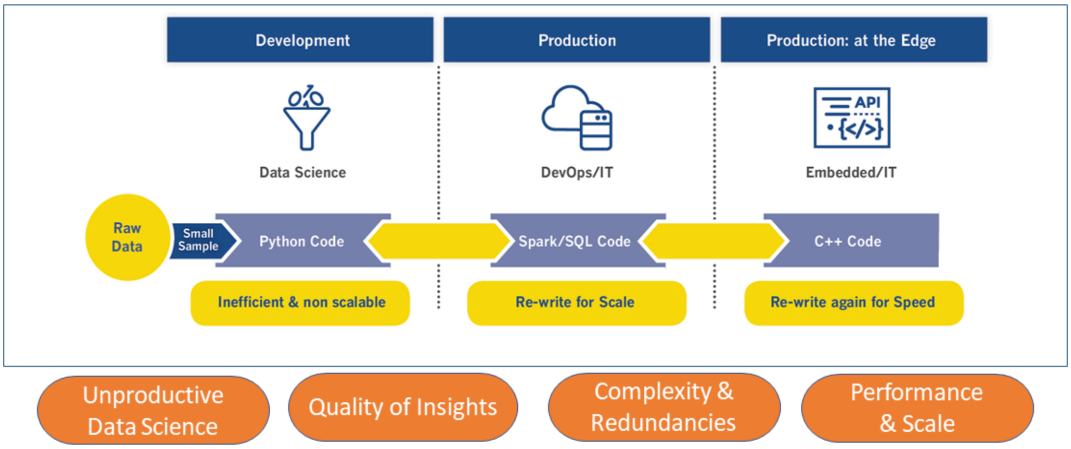


Figure 1: Enterprise AI adoption challenges according to Gartner

Analytics Workflow and Infrastructure Challenges

While advanced analytics and artificial intelligence (AI) is gaining wider acceptance in enterprises; complex workflow and infrastructure requirements inhibit effective analytics deployments that deliver high business value. A fundamental problem is segregation of environments between development (data science) and production (IT) teams. Data scientists usually write Python code on local workstations or small development clusters for productivity reasons, while IT teams rewrite the code in Java*, Scala* or C++*, using technologies such as Apache Spark* or MPI*, to achieve performance, scalability, and reliability on production clusters. This workflow is cumbersome, costly, slow and error-prone. Furthermore, requirements of each individual AI project might force IT to setup a new infrastructure, which exacerbates the problem. A unified environment, that can meet both development and production requirements and enable a continuous AI workflow, eliminates many of these barriers for enterprise analytics.

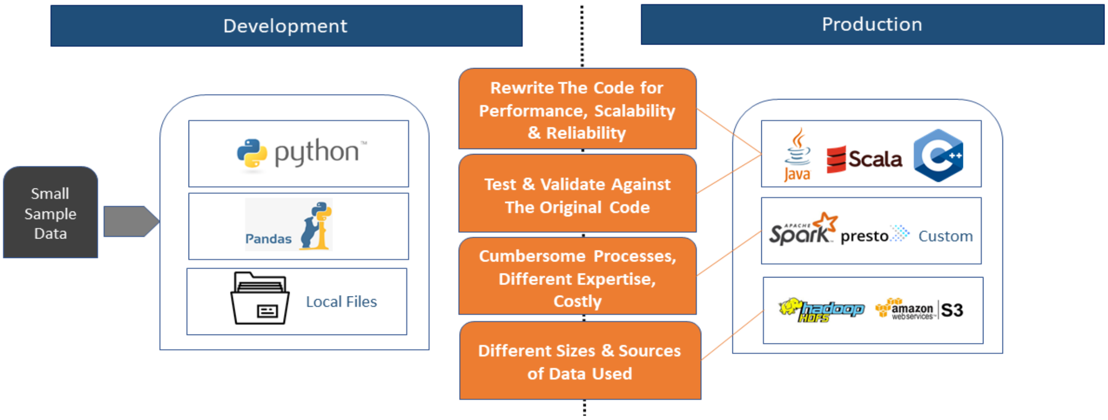

Development
Production
Small Sample Data
python
Pandas
Local Files
Rewrite The Code for Performance, Scalability & Reliability
Test & Validate Against The Original Code
Cumbersome Processes, Different Expertise, Costly
Different Sizes & Sources of Data Used
Java
Scala
Spark
presto
Custom
hadoop HDFS
amazon web services
S3

Figure 2: Existing Development and Production flow

Solution Architecture

The proposed architecture simplifies the infrastructure by unifying the production and development environments through the use of a single source code, software stack, and dataset. At the core of this solution is Bodo* analytics engine, which scales and optimizes Python analytics automatically, and can fully integrate with a variety of HPC architectures. This solution enables productive Data Science, more accurate insights, cost effective infrastructure through unparalleled efficiency, and is optimized to achieve real time insights from Cloud to Edge.

Unified Analytics Architecture overview

We propose a unified analytics reference architecture that uses the same Python source code and the same software stack for all development and production needs. Bodo and Intel provide the key technologies for performance, scalability, and reliability to enable such solutions. Using unified solutions, data science teams gain access to full data sets and can develop more accurate algorithms. Furthermore, IT teams can avoid code rewrites and redundant setups, and can focus on serving more high-value AI projects seamlessly. The portability and agility of this architecture enables utilization of existing HPC and big data infrastructures for analytics and AI, while also being able to take advantage of new technologies. Moreover, this architecture enables multi-cloud strategies since it is fully portable and not tied to any particular setting.

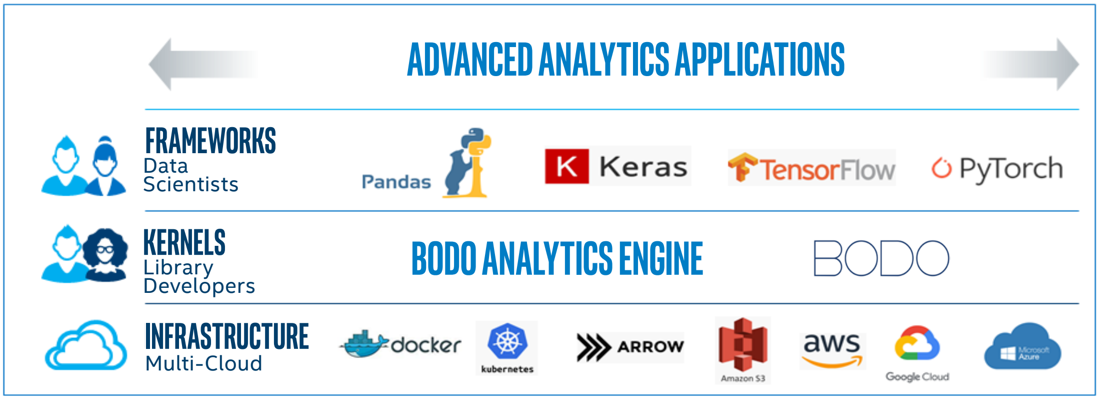


Figure 3: Bodo platform architecture

*Other names and brands may be claimed as the property of others. Developer personas shown above represent the primary user base for each row, but are not mutually-exclusive. All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.

Solution Architecture components

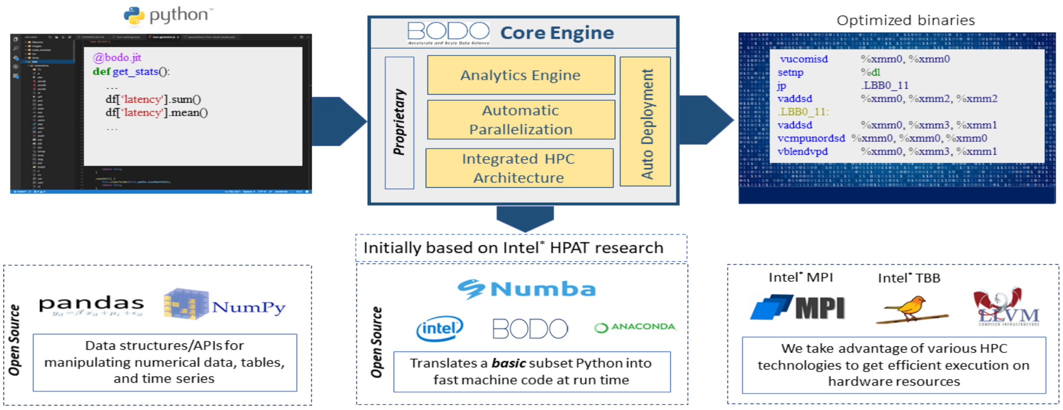
Modern technologies across the software stack are leveraged in this solution to provide the necessary development and deployment agility for effective data-centric organizations. Figure 4 demonstrates an example architecture using Bodo and Intel technologies, which includes several components:

- **Python Language:** a single Python source code is utilized for both development and production, which could be in Jupyter* notebooks, and is ideally managed in a single code repository (with version and access controls). Intel Distribution for Python, as well as Intel-optimized distributions such as Anaconda*, provide the required efficiency, reliability and security to enable robust IT environments. Bodo fully integrates with these distributions as well.
- **Data Frameworks:** various data frameworks and libraries, such as Pandas*, TensorFlow*, and Intel® Daal4py, are simply installed as Python packages. Bodo integrates with these packages seamlessly.
- **Engine and Connectors:** Bodo Core Engine scales and accelerates the Python code to achieve HPC performance and scalability. In addition, Bodo Connectors, based on Apache Arrow, provide high performance integrations with data sources and other components. Bodo Core Engine takes advantage of several high-performance Intel libraries, such as Intel® MPI, to improve efficiency.

- **Middleware:** container software such as Docker* and container orchestration frameworks such as Kubernetes* enable fast and robust deployment and infrastructure management. Various software products based on these technologies such as Red Hat OpenShift* are optimized for Intel® Architecture. Bodo is fully compatible with these environments.
- **Data lake/warehouse:** we recommend centralized data lake or data warehouse solutions that empower all teams to take advantage of data resources. Access control and security features of modern data storage technologies should be used to ensure integrity and safety of enterprise data. Bodo Connectors enable high-performance integration with various data storage solutions.

Bodo Core Engine

The Bodo Core Engine provides Python simplicity and HPC scalability simultaneously. The compiler technology parallelizes and optimizes Python code automatically, and produces HPC binaries for transparent execution. The engine uses open-source Numba* Just-In-Time (JIT) compiler technology (based on LLVM*), and integrates with HPC technologies such as MPI. Hence, the user can take full advantage of Python flexibility and capabilities, while exploiting advanced high-performance features in Intel products such as SIMD instructions, multicore servers, and distributed execution.



```
@bodo.jit
def get_stats():
    ...
    df['latency'].sum()
    df['latency'].mean()
    ...
```

```
vucomisd     %xmm0, %xmm0
setnp        %dl
jp           .LBB0_11
vaddsd       %xmm0, %xmm2, %xmm2
.LBB0_11:
vaddsd       %xmm0, %xmm3, %xmm1
vcmpunordsd  %xmm0, %xmm0, %xmm0
vblendvpd    %xmm0, %xmm3, %xmm1
```

Data structures/APIs for manipulating numerical data, tables, and time series

Translates a ***basic*** subset Python into fast machine code at run time

We take advantage of various HPC technologies to get efficient execution on hardware resources

Figure 4: Bodo core engine flow

Intel Hardware and Software Enhancements

Various Intel hardware and software technologies can be used in this architecture to improve performance.

Frameworks and Libraries: Intel continuously optimizes popular Python packages such as Tensorflow for Intel® Architecture. Furthermore, Bodo Core Engine takes advantage of several high-performance Intel libraries, such as Intel® MPI, to improve efficiency.

Hardware Platform:

- **Processors:** we recommend the Intel® Xeon® Scalable Processor family, especially Intel® Xeon® Platinum processors for high scalability, efficiency, agility, and robustness. Bodo optimizes for Intel processors at several levels:

  1) Code optimizations such as memory access improvements avoid processor stalls and enable high core performance

  2) Automatic vectorization enables fine-grained parallelism of Intel® Advanced Vector Extensions 512 (Intel® AVX-512)

  3) Effective parallelism enables multi-core efficiency, which can take advantage of high core count Intel® Xeon® processors

  4) Integration with Intel® MPI provides scalability and efficiency for multi-node execution.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer. Intel, the Intel logo, and Intel Xeon are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others

- **Networking:** Intel® Omni-Path Architecture (Intel® OPA) provides fast HPC network fabric, while Intel® Ethernet® products provide fast conventional networking.
- **Storage:** Intel® Optane™ DC Persistent Memory and Intel® Optane™ SSDs improve access to large datasets.

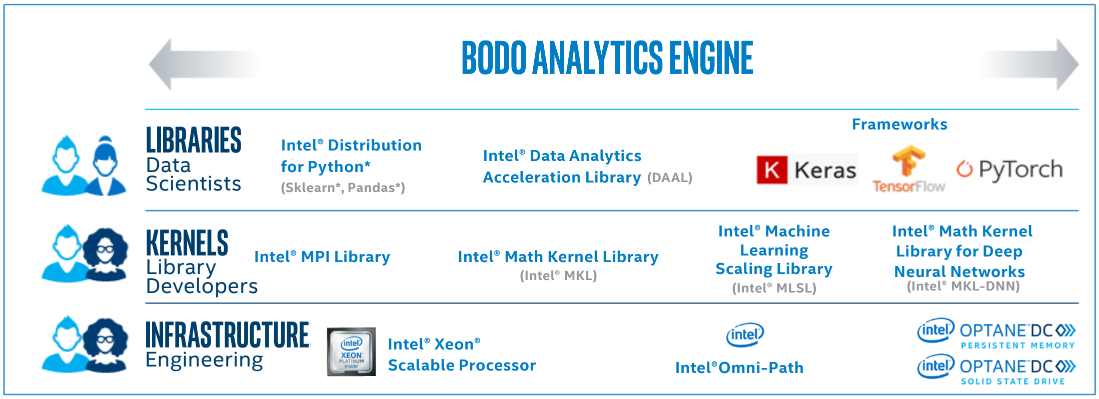

BODO ANALYTICS ENGINE
LIBRARIES
Data Scientists
Intel® Distribution for Python*
(Sklearn*, Pandas*)
Intel® Data Analytics Acceleration Library (DAAL)
Frameworks
Keras
TensorFlow
PyTorch
KERNELS
Library Developers
Intel® MPI Library
Intel® Math Kernel Library
(Intel® MKL)
Intel® Machine Learning Scaling Library
(Intel® MLSL)
Intel® Math Kernel Library for Deep Neural Networks
(Intel® MKL-DNN)
INFRASTRUCTURE
Engineering
intel XEON PLATINUM inside
Intel® Xeon® Scalable Processor
intel
Intel®Omni-Path
intel OPTANE DC PERSISTENT MEMORY
intel OPTANE DC SOLID STATE DRIVE

Figure 5: Intel optimized technology and BODO platform architecture

*Other names and brands may be claimed as the property of others.Developer personas shown above represent the primary user base for each row, but are not mutually-exclusive. All products, computer systems, dates, and figures are preliminary based on current expectations, and are subject to change without notice.

Use Cases and Benchmarks

Early experience with Bodo across many FSI customer POC engagements has shown its usefulness for processing large scale time series data and credit card data. Bodo has demonstrated ease of implementation, orders of magnitude performance gains, and Cloud to Edge efficiency on a half dozen uses cases from Fortune 500 enterprises. For example, Bodo let the quant team at a large bank scale their credit card risk analysis Python code automatically, avoiding code rewrite by IT teams which was a major obstacle, as they are resource constrained. Bodo's automatic scaling enabled the quant team to run their code on the full data set in real time instead of sampling 1% of data. In another engagement, Bodo enabled a major exchange to process transaction data much faster and meet their compute performance needs, operational time requirements, and also simplified their data infrastructure. The solution was also proved on several non-FSI use cases such as real time network security at the edge using AI, and analysis of network logs and load prediction with one of the largest US based telco-carriers. The technology has also been validated and benchmarked by the Securities Technology Analysis Center* (STAC*), which is a respected authority who maintains and publishes benchmarks of interest to the Financial Services Industry. Furthermore, Bodo demonstrated full portability to take advantage of various existing HPC and big data clusters, while enabling forward-looking multi-cloud strategies.

Deployment Recommendations

The platform can be deployed on-premises, in a public cloud or in hybrid cloud infrastructure, depending on specific business needs. Careful selection of infrastructure technologies can ensure scalability, high performance, security, and low cost. We highlight several key components of effective solutions below.

Installation and Incremental Upgrades

Bodo Core Engine is a standard Python package that can be installed in most environments seamlessly since Python is most often supported already, and Bodo is fully portable. We recommend enterprise-grade Python distributions such as Anaconda* Enterprise*, which provide robust package management and enhanced security by patching and upgrading packages in a timely manner.

We recommend installing Bodo in the existing environment and upgrading applications and infrastructure incrementally over time. Bodo scales existing Python applications and also accelerates legacy PySpark* workloads. However, applications and infrastructure can be enhanced to take full advantage of Bodo capabilities.

Scalable Storage Systems

Choosing the right storage systems for high performance analytics is important in this architecture, since Bodo eliminates compute bottlenecks and storage performance can become the limiting factor. For example, legacy relational database management systems (RDBMS) solutions can become a bottleneck for analytics on large datasets.

Public cloud platforms provide several storage options for analytics. For example, Amazon S3* is a scalable object storage system designed for large historical data. In addition, faster block storage such as Amazon Elastic Block Store* (EBS) is recommended for real time analytics. There are similar solutions for on-premise platforms such as OpenStack Swift object storage system*. Legacy Hadoop Distributed File System* (HDFS) storage can also be utilized if necessary, but performance tuning is recommended. Furthermore, latest storage hardware technologies such as Intel® Optane™ DC SSD can improve performance significantly.

Compute Platform

Bodo provides performance boost on all clusters, but performance features of latest hardware can enable much faster analytics. For example, Bodo can take advantage of high core count and AVX512 features of second-generation Intel® Xeon® Scalable Processors effectively. Furthermore, the Intel® Omni-Path network interconnect solution can be utilized effectively due to efficient distributed computing features of Bodo. Public cloud platforms also provide high performance instances, as well as high-bandwidth networks. We recommend taking advantage of enhanced networking and close instance placement features of cloud platforms for best performance.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer. Intel, the Intel logo, and Intel Xeon are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others

Conclusion

We have demonstrated that through a combination of Intel technology and Bodo, enterprises IT would be able to simplify their infrastructure by unifying the production and development environments through the use of a single source code and software stack. Bodo's analytics engine technology scales Python AI to HPC and is optimized to achieve real time insights from Cloud to Edge. The proposed architecture enables productive Data Science, more accurate insights, and cost-effective infrastructure through unparalleled efficiency. Additionally, this architecture is fully portable and enables running different workloads on different cloud and on-premises environments. Furthermore, this architecture provides the flexibility required to best support and adapt to emerging complexity of future workloads, and the requirements of the analytics economy.

Solutions Proven By Your Peers

The Bodo Platform, powered by Intel® technology, provides enterprises with advanced analytics and AI capabilities. This and other solutions are based on real-world experience gathered from customers who have successfully tested, piloted, and/or deployed these solutions in specific business use cases. Solution architects and technology experts for this solution reference architecture include:

- **Behzad Nasre**, Chief Executive Officer, Bodo Inc.
  behzad@bodo-inc.com
- **Ehsan Totoni**, Chief Technology Officer, Bodo Inc.
  ehsan@bodo-inc.com
- **Parviz Peiravi**, Chief Technology Officer,
  Financial Services Industry Solutions,
  Intel Corporation

Intel Solution Architects are technology experts who work with the world's largest and most successful companies to design business solutions that solve pressing business challenges. Bodo, a market leader in data science core technology, has worked closely with Intel Solution Architects to develop the solution described in this document.

Find the solution that's right for your organization.
Contact your Intel representative or visit intel.com/FSI.

Learn More

You may also find the following resource useful:

- **Intel:** intel.com/fsi

Solution Provided By: