

SBP Studio

Web GUI-based IDE for Hadoop Data Processing and Analysis Applications

HIGHLIGHTS

- Hadoop-based integrated solution that provides data discovery, data processing and machine learning
- No coding needed for application development with 30 pre-built components and GUI
- Components optimally developed with native MapReduce, which execute faster than the applications developed with Pig or Hive
- Prevent errors with impact analysis between applications
- Troubleshooting with drill-down job tracking
- Compatible with major Hadoop distributions: Apache Hadoop, CDH, HDP and MapR

Intuitive GUI-based Development

SBP Studio allows users to easily develop Hadoop applications by placing and connecting visual components on a canvas with drag-and-drop even if users don't know how to write code or scripts. With SBP Studio, a business analyst who does not have any software development skills can develop applications with 2 days of training.

Also, users can easily understand the applications developed by other developers because users can see the visualized structure of an application.

High Performance Components

SBP Studio provides about 30 pre-built components (I/O, control, action and machine learning, such as log analysis and data mining) for the development of data processing applications of various types. These components were developed with Java MapReduce, whose execution speed is faster on Hadoop platform than any other technology, and were optimized by MapReduce experts in terms of performance, thus running two to four times faster than the programs developed with Pig or Hive scripts.

Product Overview

It is difficult to develop and operate applications on the existing Hadoop platform.

- Difficult to find developers or learn advanced technologies required for application development in a Hadoop environment, (e.g. MapReduce, Java)
- Configuring a development environment is complicated when using various technologies, (e.g. Eclipse, SVN, job scheduler)
- Difficult to maintain applications because content developed by others is not readily understandable at a glance.
- Developed modules often do not meet performance standards.
- Difficult to measure the impact between jobs as the number of jobs increase.
- Difficult to identify the cause of errors and take appropriate measures in a timely manner

SBP Studio is a web-GUI based integrated solution designed to solve these problems and develop and operate applications on the Hadoop platform effectively. This solution has been tested and is stable enough to use in the enterprise environment.

Key Features

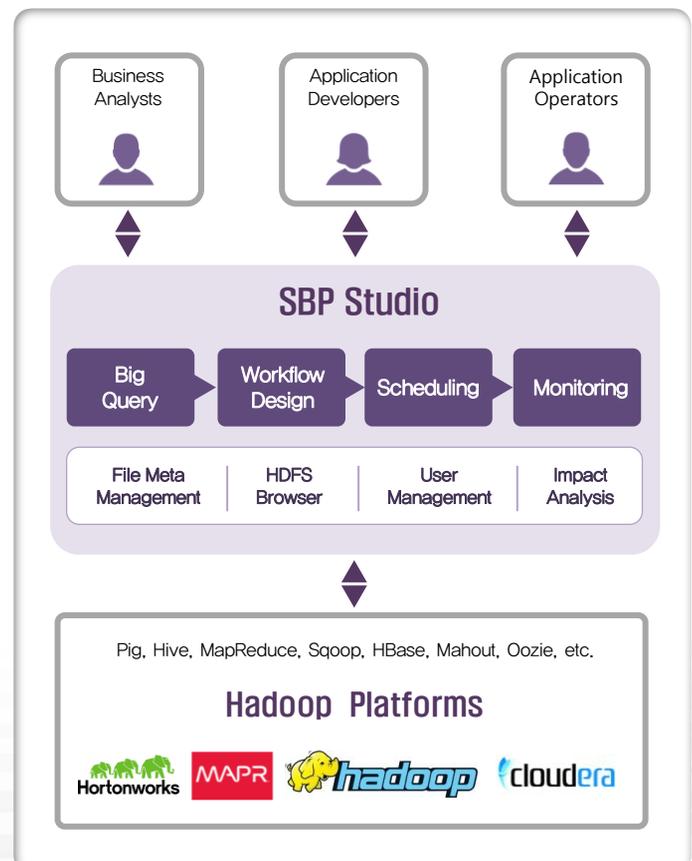
Integrated Solution for Hadoop

SBP Studio offers the functions required for the development and operation of Hadoop applications including user management, HDFS browsing, metadata management, data processing and machine learning in a single Web UI environment. Therefore, there is no need for separate solutions for module sharing and job scheduling.

Moreover, SBP Studio, once installed in the server, allows multiple people to develop applications on the same environment and it is easy to maintain the web applications and share the development modules among developers, making it easy to collaborate among developers.

Components for Advanced Analytics

Pre-built components include the components for advanced analytics. These components use Mahout, a Hadoop-based software for machine learning, and allow users to implement advanced analysis applications easily and quickly without having to program them using Mahout Java API.



Key Features

Extension with Custom Codes

SBP Studio provides functions that allow developers to enter source codes directly. Thanks to these functions, the existing codes deployed in the Hadoop platform or the ones developed in other systems can easily be incorporated into applications.

That is, developers can select the programming language that they think is most familiar and productive for development as diverse custom code components are provided, including Pig, Hive, MapReduce and Java.

Impact Analysis of Applications

To add or change a data processing job, it is important to understand its connection with the existing jobs. Otherwise, it may result in data errors or operational defects. However, it is practically impossible to manually analyze the impact between jobs with more jobs deployed in Hadoop. By using Impact Analysis in SBP Studio, users can easily and accurately understand the impact between jobs with a click on the GUI environment.

Efficient Drill-Down Job Tracking

Problems can be easily detected as SBP Studio allows users to monitor job results or progress on a single display. Also, for specific jobs that have caused an error or a defect during operation, users can identify the problem with the drill-down and advanced log search functions. During monitoring, if a particular job is detected to cause problems for other jobs or the entire system, users can pause or kill the job on the SBP Studio screen without accessing the Hadoop system console.

Compatible with Major Hadoop Distributions

SBP Studio features excellent compatibility as it uses standard open sources provided in most Hadoop distributions, including Oozie, MapReduce, Pig and Hive. Compatibility tests were completed for major distributions such as Apache Hadoop, Cloudera CDH, Hortonworks HDP and MapR Technology MapR, and compatibility certification is underway for each distribution. In terms of investment protection, SBP Studio offers great advantages for companies that use the Hadoop platform.

Additional Benefits

SBP Studio offers all functions required for the development and operation of Hadoop applications in a single web UI environment and is easy to use, improving the productivity of development and operation.

By using the powerful and comprehensive pre-built components, users can develop high-performance and high-quality applications.

The drag & drop development system based on easy-to-use and intuitive GUI makes it possible for a business analyst to do end-user computing.

Technical Specification

Supporting Hadoop Distributions

- Apache Hadoop
- Cloudera CDH
- Hortonworks HDP
- MapR Technology MapR
- LG CNS SBP

Supporting Operating Systems

- CentOS
- Redhat Linux
- Suse Linux

Supporting Web Browsers

- Internet Explorer 9, 10, 11
- Google Chrome

Pre-Built Components

- Control – Start, End, Fork, Join, Decision
- IO – HDFS Input, HDFS Output, SqoopImport, SqoopExport, HBaseExport, HBaseImport
- Action – Expression, Filter, Aggregate, Sort, SubWorkflow, Upsert, Grep, Relocate, Union, Join, Pivot, Pattern, Pig, Hive, MapReduce, Java, Ssh, Fs(HDFS)
- Mahout(machine learning) – CF KeyGen, CF Parser, ItemBased CF, ParallelALS, SVD CF, CL Preprocess, CL Postprocess, Clustering

About LG CNS

LG CNS provides top-class integrated solutions to customers from consulting and system implementation to system operation through smart convergence with all business areas based on IT services.

LG CNS runs the Big Data Business Unit a specialized organization in charge of big data business and technical support. At the Big Data Business Unit, specialized big data consultants, analysts and platform engineers work together. The LG CNS Big Data Business Unit offers big data services optimized for customers via its technical capabilities and solutions based on the company's traditional proven technical support system.

http://bigdata.lgcns.com/Products/SBP_Studio