

[SAP Community](#) > [Products and Technology](#) > [Technology](#) > [Technology Blogs by Members](#)> [BIG Data: Integrating SAP BusinessObjects with Hadoop](#)

Technology Blogs by Members

Explore a vibrant mix of technical expertise, industry insights, and tech buzz in member blogs covering SAP products, technology, and events. Get in the mix!

Blog

*What are you looking for today?*

BIG Data: Integrating SAP BusinessObjects with Hadoop 12.1

**poovin2749**

Explorer



2013 Jun 03 7:08 AM

**13 Kudos**

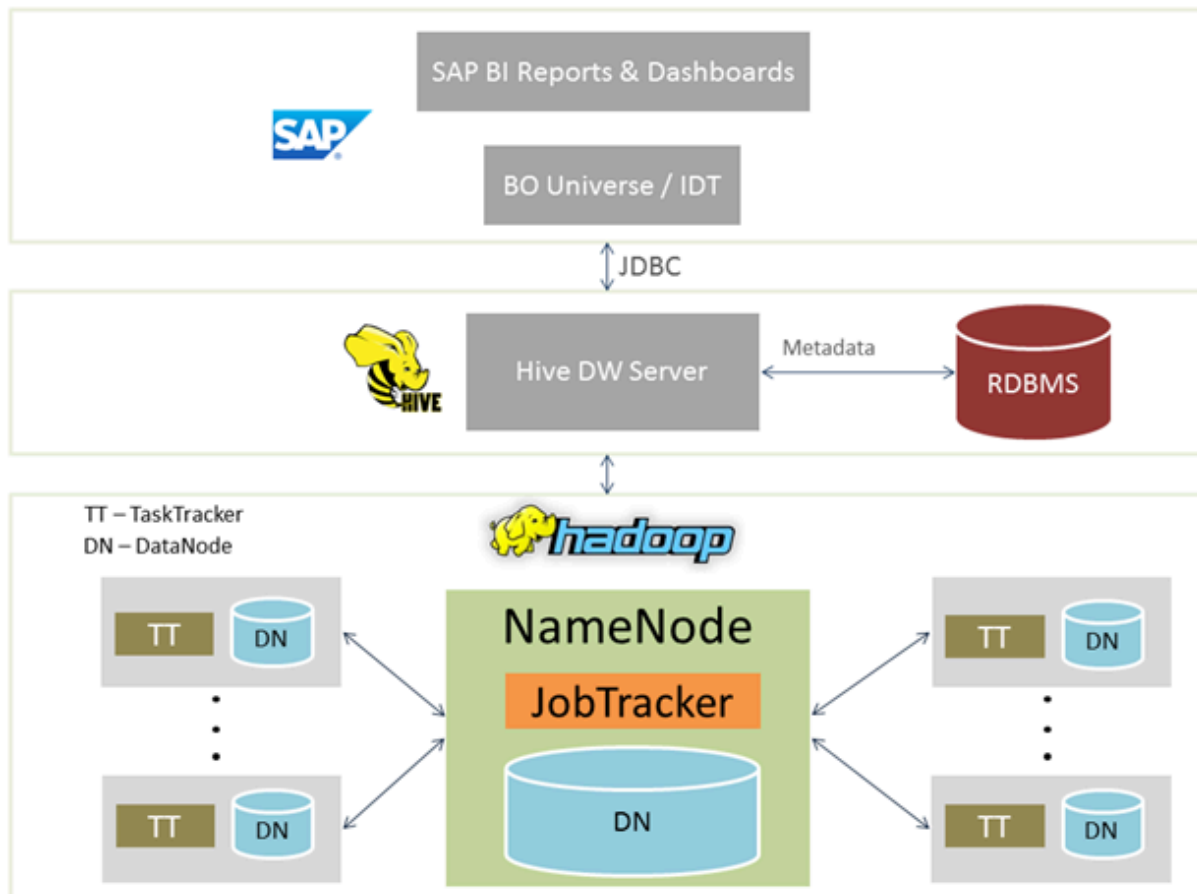
23,870

SAP Managed Tags: SAP BusinessObjects - Semantic Layer

Enterprises today are increasingly keen on harnessing the power of **BIG Data** for decision-making. While a few have made strong inroads into BIG Data and reaped benefits, realizing its true value has been elusive for most. For one, in order to benefit from BIG Data, it is essential to understand the business, and have a specific problem identified that needs solving.

Once the above is identified, one needs to look at the sources of information – structured and unstructured – that could come to aid. A mechanism to capture, analyze and interpret data has to be designed at this stage. An understanding of technology, tools and the various innovations in BIG Data is key to this step. Finally, the data may have to be integrated with structured information already available in the enterprise to make specific decisions.

We will focus on the last step mentioned above by integrating SAP BusinessObjects with **Hadoop**.



Keep in mind that despite having robust features, platforms such as Hadoop are dependent on batch processing, which make them unsuitable for real-time analytics and reporting. The powerful combination of SAP and Hadoop can bring together agility, scalability, flexibility and affordability needed to fully tap the potential of BIG data for effective decision making.

We used the following approach to achieve SAP BusinessObjects integration with Hadoop. [Click here to download step by step instructions.](#)

Step 1. Install and configure multi-node Hadoop Cluster

Hadoop cluster installation was done sequentially to ensure easy configuration and maintenance. The initial installation was done in a single machine with all services running in a single node (master node). It was then extended to two other nodes (slave nodes). Once the cluster became operational, all the different file types were loaded into the system. The Hadoop cluster replicated it across all available data nodes. The cluster was then tested with a conventional word count program. This verifies communication between various Hadoop nodes and validates their ability to perform distributed operations.

Step 2. Install and configure Hive Data Warehouse

In contrast to other data warehousing tools that extract, transform and store/load data into target systems, Hive extracts and transforms data at run time without having to load or store it in a target system. Hive converts and processes user tasks (defined using HiveQL statements) using the Map-Reduce distributed framework. Map-Reduce deploys the job tracker to split a single job to multiple sub tasks and returns the results back to the master node.

We then performed text analysis by running word count on a text file that contained a speech delivered by Richard Stallman on Free Software was performed. Hive schema was defined on the file which was then queried with HiveQL statements to obtain the respective word counts.

Step 3. Integrating SAP BusinessObjects with Hadoop

In order to run reports on Hadoop, SAP BusinessObjects was set up on top of Hadoop Hive-JDBC connectors. Information Design Tool (IDT) was used to configure the setup. Once you complete this, you would be able to use an array of BO front end tools that can report on Hadoop – including SAP Dashboards, SAP Web Intelligence, SAP Crystal Reports, SAP BO Explorer, Visual Intelligence and more.

Source : <http://www.visualbis.com/blogs/sap-businessobjects/2013/05/31/big-data-integrating-sap-businessobjec...>

Tags:

Big Data

big data geek challenge

bobj

hadoop

hadoop universe

hadoophive

hive

SAP

universe

17 Comments

**harshil_joshi**

Contributor



2014 Sep 15 8:46 AM



0 Kudos

Hi Poovin,

May I know hardware requirements for Apache Hadoop and Apache Hive ?

Both are open-source right ? Both should be install on different server ? Appreciate your reply. Thanks.



poovin2749

Explorer



2014 Sep 16 4:59 AM



0 Kudos

Hi Harshil,

There are no written standards for Hadoop/Hive Hardware requirements. Initially Hadoop/Hive was light and it used to run even in commodity hardware (ex: 2.0 GHZ Processor, 2 GB RAM and 20 GB HDD). However with increased functionalities, they demand more now. Still you can run Hadoop/Hive in your Desktop if it is of the recent configuration.

Yes, Both Hadoop and Hive are opensource. You can also get it from vendors like Cloudera and Hortonworks. They have both community edition(free) and enterprise edition. This installation will demand more resources due to increased functionality.

Hadoop and Hive can sit together in a same server. You can also scale out the installations across multiple nodes based on your need.

Thanks,

Poovin.

**harshil_joshi**

Contributor



2014 Sep 24 5:57 AM



0 Kudos

thanks poovin.

**former_member182541**

Active Contributor



2014 Dec 17 10:27 AM



0 Kudos

Hi Poovin,

I am in a Project with similar situation, the above provided link is not providing me the document or page would have broken. Can you please send me the document to my id.

Also how do we connect universe with impala hive tables, do we use ODBC or JDBC ?

Many Thanks,

sumanT

**poovin2749**

Explorer



2014 Dec 17 3:33 PM



0 Kudos

Hi Suman,

I have shared the document with you through E-Mail. I will fix the SCN link as well.

Reg Impala - We can use both ODBC and JDBC to connect with hive/impala. However i would prefer to go with JDBC drivers provided by respective Hadoop vendors. BO 4.1 SP05 has improved support for both hive and impala.

I will share the information very soon with step-by-step guides.

**Former Member**

2015 Feb 03 5:43 AM



0 Kudos

Hi Poovin,

I am pursuing my masters. I want to start my career in SAP BI or Hadoop. But I don't know which one to opt for. Recently, I have seen a job with title as SAP BI architect with hadoop hive. Can you please tell me about both technologies so that I can decide on which one to take up.

Thanks,

Shravani

**poovin2749**

Explorer



2015 Feb 03 10:28 AM



0 Kudos

Hi Shravani,

SAP BI is a large domain with different set of tools and features. SAP provides solutions (or) platforms for Business needs and that includes Business Warehousing, DataBase, ERP, Reporting tools and lot many.

Hadoop is an Open-Source product promoted by Apache. Hadoop is also a huge ecosystem with many allied tools like hive, impala, hbase etc.. However most of these tools cater towards Data Warehousing and Analytics.

A job role with Both SAP BI and Hadoop will be great and challenging.

Thanks,

Poovin

**Former Member**

2015 Feb 03 2:18 PM



0 Kudos

Hi Poovin,

How about the Performance of BO Reports, Is there any document on Performance.

**poovin2749**

Explorer



2015 Feb 03 3:04 PM



0 Kudos

Hi Venu,

Do you mean the performance of reports on top of Big Data (or) in general?..

**former_member193452**

Active Participant



2015 Feb 03 7:13 PM



0 Kudos

the latest Information and how to Connect the BI environment to Hadoop Hive may be found here on the SCN Semantic Team Page.

<http://wiki.scn.sap.com/wiki/x/FA-OFw>

Jacqueline

**Former Member**

2015 Feb 03 11:05 PM



0 Kudos

thanks poovin...

**Former Member**

2015 Apr 16 11:02 AM



0 Kudos

Hi Poovin,

Great Article keep it up

**Former Member**

2016 Nov 18 6:49 AM



1 Kudo

Hello Poovin,

We are currently on SAP Business Objects 4.1 SP8 and trying to establish a JDBC connection with Cloudera Impala and Hive through Simba drivers but we could not successful because cloudera enabled with Kerberos authentication so we are getting below error.

[Simba][ImpalaJDBCDriver](500164) Error initialized or created transport for authentication:

CONN_KERBEROS_AUTHENTICATION_ERROR_GET_TICKETCACHE.

Do you have any idea on this error? or any documents related to this?

Cloudera version 5.7

Imapala version 2.5

Hive version 1.1

BOBJ on Red hat linux 6.x

Thanks,

Gana



Former Member



2017 Feb 17 5:19 PM



0 Kudos

Impala query is much faster than Hive.

Do you see any performance issue in WebI report when it is trying to get data from Hadoop using Hive?

Regards,

vinesh



former_member227650

Explorer



2017 Mar 02 9:58 PM



0 Kudos

Hi Gana,

I'm seeing the same error. Did you ever get past this error?

**Former Member**

2017 Sep 13 9:59 PM



0 Kudos

Hi Brian,

We have resolved this issue by enabling the LDAP authentication on Impala. Please let me know, if you need more details.

Thanks,

Gana

**Former Member**

2018 Feb 15 2:06 PM



0 Kudos

Hi Poovin,

Thanks for the document its bit old now but still useful.

Thanks,

Prateek

You must be a registered user to add a comment. If you've already registered, sign in. Otherwise, register and sign in.

Comment

Labels In This Area

"Aging List of Receivables" 1 "as_written_by_Marian_Zeis" 1

"automatische backups" 1 "Data Source Migration" 1 "Integration Challenges" 1

"regelmäßige sicherung" 1 "SAP BW 7.4" 1 "SAP BW" 5

"SAP VARIANT CONFIGURAITION 2 "SAP_BUILD_APPS" 1 "SAPDatasphere" 2

"TypeScript" "Development" "FeedBack" 1 *SAP" 1

-147 Get CurrentUserInfo failed 1 2YM 1 3-TIER Extensibility 3 30 examples 1

505 Technology Updates 53 1 @expertsap 1 @hanasizing 1

@RetroDate_HireDateCorrection 1 @sapilm @archiving @sapiq 1

@SAPSupport 1 @SCPI 2

A Comprehensive Guide to Using OLE Objects in SAP ABAP 1 aATP 1 ABAP 41

ABAP 7.4 2 ABAP API 1 ABAP BAPI BAPI_FIXEDASSET_CREATE1 1

ABAP BTP 1 ABAP CDS VIEW 2 ABAP CDS Views 12

ABAP CDS Views - BW Extraction 3 ABAP CDS Views - CDC (Change Data Capture) 3

ABAP Class 3 ABAP Cloud 8 ABAP Cloud Developer Trial 1

ABAP DDIC CDS view 1 ABAP development 13 ABAP Editor 1

ABAP Environment & RAP 2 ABAP Extensibility 4 ABAP for EWM 1

ABAP in Eclipse 3 ABAP Interface 1 ABAP New Syntax 1 ABAP ODATA 2

ABAP on HANA 1 ABAP OOABAP 1 ABAP PLATFORM 1 ABAP Platform Trial 2

[ABAP Programming 8](#)[ABAP Push Channels 1](#)[ABAP Query 1](#)[ABAP RAP 3](#)[ABAP RAP custom action 2](#)[ABAP RAP\(RESTful Application Programming\) 5](#)[ABAP RESTful API 1](#)[ABAP RESTful Application Programming Model 2](#)[ABAP String functions 1](#)[abap technical 1](#)[ABAP test cockpit 1](#)[abap to xml 1](#)[abapGit 1](#)[absl 2](#)[Access data from datasphere to ADF Azure Data Factory 5](#)[access data from SAP Datasphere directly from Snowflake 1](#)[Access data from SAP datasphere to QlikSense 2](#)[Accessibility 1](#)[Accessibility in SAPUI5 1](#)[Accrual 1](#)[Acquire SAC Knowledge 3](#)[acquired 1](#)[action 1](#)[actions 1](#)[Activity 1](#)[Adaptation Project 1](#)[adapter 2](#)[adapter modules 2](#)[ADDING LEAN SERVICES 2](#)[Addon 2](#)[Adobe Document Services 1](#)[Adobe forms 1](#)[ADS 1](#)[ADS Config 1](#)[ADS with ABAP 1](#)[ADS with Java 1](#)[ADT 4](#)[Advance Shipping and Receiving 1](#)[Advanced ABAP 1](#)[Advanced Event Mesh 4](#)[Advanced formula 1](#)[Advanced Formulas 2](#)[Advanced Metric 1](#)[Advanced SAP Techniques 1](#)[Advanced Scripting in SAC 3](#)[Advanced Workflow 2](#)[AEM 1](#)[AEM Event Portal 1](#)

Related Content

SAP Business Technology Platform: “Change Agent” or “Scale Agent” Part 4 – Benefits

in Technology Blogs by SAP Wednesday

SAP Datasphere News in September

in Technology Blogs by SAP 2024 Sep 30

E-Learning: SAP BusinessObjects BI - Hybrid Integration

in Technology Blogs by SAP 2024 Sep 24

Comparison of SAP Analysis for Microsoft Office (SAC edition) vs. SAC add-in for Microsoft Excel

in Technology Blogs by SAP 2024 Aug 22

Unlock the full potential of SAP Analytics Cloud and SAP Datasphere

in Technology Blogs by SAP 2024 Jun 03

Popular Blog Posts



SAP PI for Beginners



former_member200339

Participant

👁 726458 💬 155 👍 388



ABAP 7.40 Quick Reference



jeffrey_towell2

Participant

👁 1213459 💬 75 👍 344



Difference between SAP S/4HANA :Public Vs Private edition : RISE with SAP



rajarajeswari_kaliyaperum

Active Participant

👁 185790 💬 47 👍 309

Top Kudoed Authors



IngoH

👍 11



santoshdwivedi

👍 8



dallas_marks

👍 7



Siva_Prakash_S

👍 7



Kamlesh_Rampal

👍 5



LucasMagriniRigo

👍 5

**jeffrey_towell2**

4

**Pradipta**

4

**PriyankaChak**

3

**LucaFalavigna**

3

[View all](#)[Privacy](#)[Terms of Use](#)[Copyright](#)[Legal Disclosure](#)[Trademark](#)[Support](#)[Cookie Preferences](#)[Follow](#)