

DOI: 10.11991/yykj.201802006

<http://kns.cnki.net/kcms/detail/23.1191.U.20180328.1333.002.html>

曲英涛

哈尔滨工程大学 信息化处, 黑龙江 哈尔滨 150001

DAMA

JYT1002_

JYT1003_

JYT1006_

Hadoop

Spark

Elasticsearch

G434

A

1009-671X(2018)02-0086-05

Construction and research of big data platform in Colleges and Universities

QU Yingtao

Information Office, Harbin Engineering University, Harbin 150001, China

Abstract: Data resource construction is an important part of information construction in Colleges and universities. The construction of big data management and service platform will provide effective support for data analysis, construction, data, management, service and technology. Based on the DAMA data management service system methodology, with the contents of standard documents including JYT1002 Basic Information on Education Management, JYT1003 Information on Administrative Management of Education and JYT1006_Information on Higher Education Management issued by the Ministry of Education as a reference, in combination with the characteristics of university data, this paper deeply integrated Hadoop framework system, Spark technology, Elasticsearch data engine technology and other big data technologies on basis of the related technologies of traditional university data sharing platform, proposed a train of thoughts for building a one-stop big data service platform integrating data storage, unified management, unified processing and unified service, providing a reference for the construction of big data platform in universities.

Keywords: data management; data center; data standard system; data service; big data technology; construction of data platform; data governance of university; data security of university

HDFS

1 高校数据现状分析

[1]

收稿日期: 2018-02-12. 网络出版日期: 2018-03-28.

作者简介: 1980-

通信作者: E-mail quyingtao@hrbeu.edu.cn.

[2]

[4]

2

3

2 高校大数据平台建设思路

3 高校大数据平台总体架构设计

3

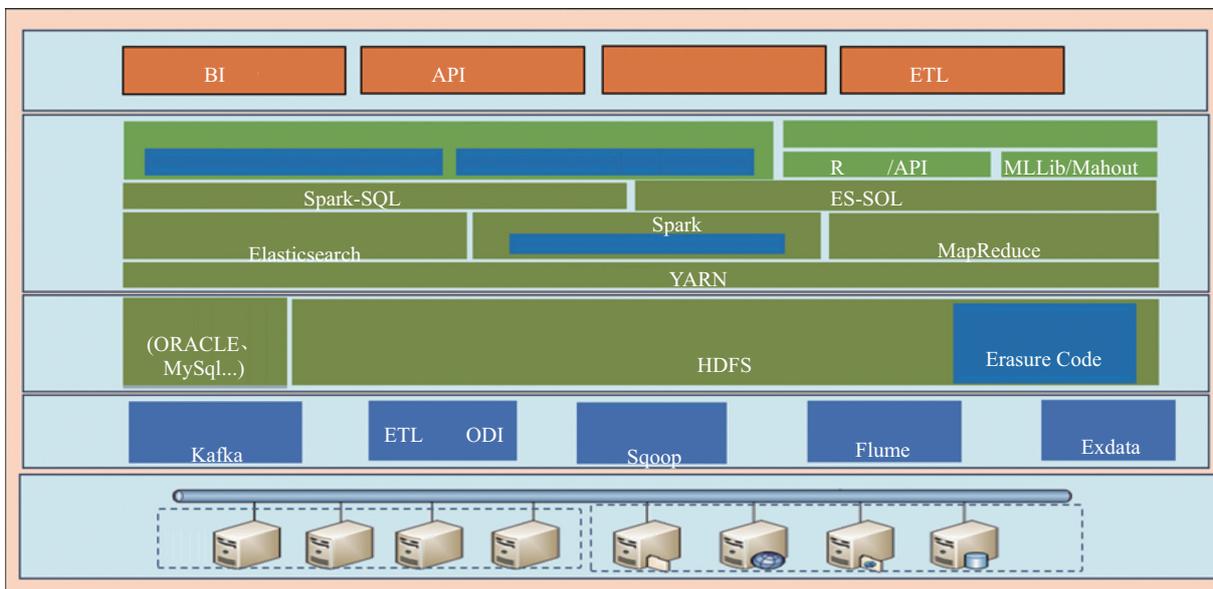
3.1

[3]

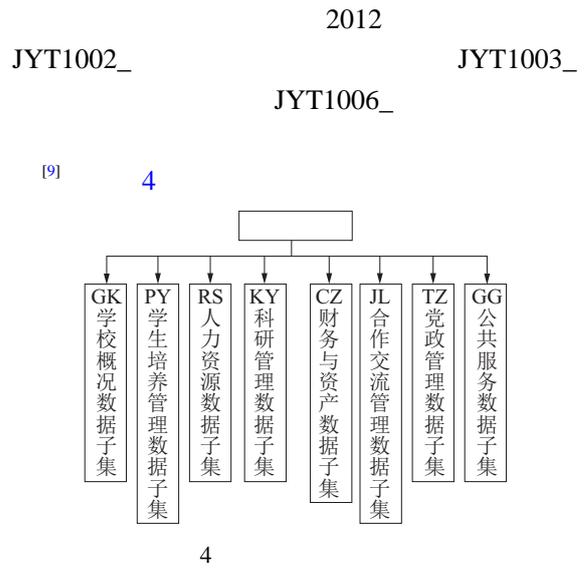
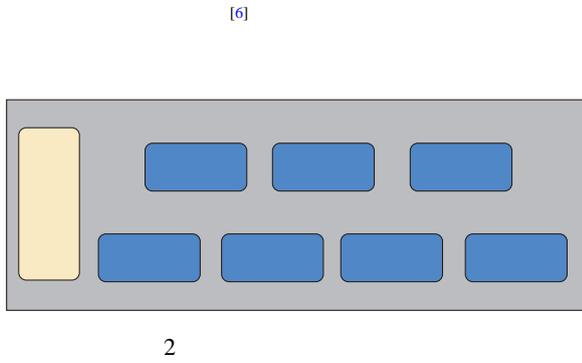
1

[5]

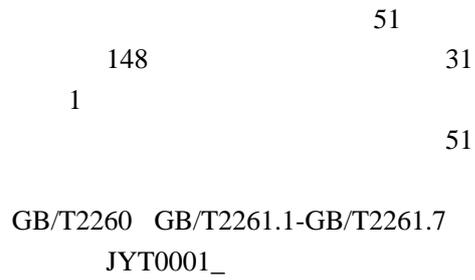
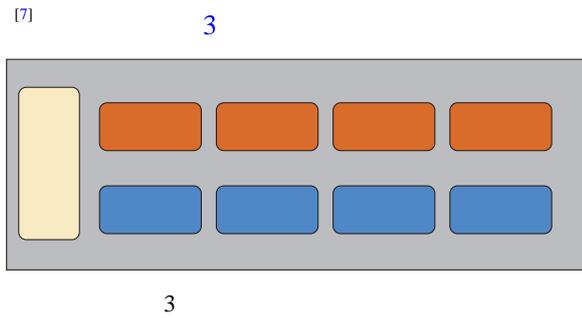
1



3.2



3.3



4 高校大数据平台建设内容

4.1

JYT0001_ 3

4.2

EXCEL

WORD

12

[8]

4.2.1

5

HDFS

6

4.2.2

4.4.2

4.4.3

[10]

4.3

5 平台建设原则

ETL

[11]

1

2

ETL

3
7 d×24 h

API

[12] SQL

4

BI

[15]

[13]

6 结论

4.4

4.4.1

1

1

2

[14]

2

IP

3

IP

3

4)

4

Uk o wnkpm 5MG [OCUVGT]L_0" ."4228."45*5+<"76:0

[7] ." ." ." 0" Uk o wnkpm TGNCR7]L_0" ."4229."4:*8+< 33463380

[8] 0" Uk o wnkpm/u]L_0" ."4237."59*4+<"556560

[3] ." ." ." 0"]L_0" ."4237."54*4+<"47364770

[4] ." 0" 5MG [OCUVGT]L_0" ."4232."4*4+<"426450

[5] ." ." ." 0" 5MG [OCUVGT]L_0" ."4237."5:*7+<"3242632560

[6] ." ." ." 0"Uk o wnkpm Ykpfqyu]L_0" 4234<"42264630

[32] 0"]O_0" <" ."422;< 36763740

." ." ." 0"]L_0" ."423:."67*4+<"26;70

\ J C Q " J w e p j w e p . " Y C P I " \ k . " U W P " R g k y g k . " g v " c n 0 " F g u k i p " c p f " x c n k f c v k q p " q h " p w e n g e t " r q y g t " r n c p v " e q p v t q n " u { u v g o " r n c v h q t o " d c u g f " q p R Y T " u k o w n c v q t] L _ 0 " C r r n k g f " u e k g p e g " c p f " v g e j p q n q i { . " 4 2 3 : . " 6 7 * 4 + < ; 2 6 ; 7 0

* ; ; + 423:*3+<"7;6850

[9] ." ." ." 0"]L_0" ."4234."4:*33+<"32;63340

[:] ." 0"]L_0" ."422:*54+<"34463450

[;] ." ." 0"]L_0" 4239*6+<"3363;0

[32] 0" ö - ö /]L_0" ."4239*43+<"6866:0

[3] 0"]L_0" 423:*+<"3563;0

[33] 0" QFK]L_0" ."4236*4+<"47;647;0

[4] " 0"]L_0" ."4239*64+<"7633."820

[34] ." 0"]L_0" 4239*6+<"46: :0

[5] ." ." ." 0"]L_0" ."4236*4+<"5:65;0

[35] 0"]L_0" ."4239*43+< 3456346."34:0

[6] 0"]L_0" ."4238*3+<"326350

[36] ." ." ." 0"]L_0" <" ."4239*7+<"36340

[7] ." ." ." 0"]L_0" ."4238*3+<"3680

[37] ." ." ." 0"]L_0" ."4235*4+< 5465:0

[8] 0"]L_0" 0"]L_0" ."423:."67*4+<"86:.;:;70

SW" [k p i v c q 0 " E q p u v t w e k q p " c p f " t g u g e t e j " q h " d k i " f c v c " r n c v h q t o " k p " E q m g i g u " c p f " W p k x g t u k v k g u] L _ 0 " C r r n k g f " u e k g p e g " c p f " v g e j p q n q i { . " 4 2 3 : . 6 7 * 4 + < ; 8 6 : . ; ; 7 0