



PPT

<http://dblab.xmu.edu.cn/node/422>

=

=



厦门大学计算机科学系教师 林子雨 编著

<http://www.cs.xmu.edu.cn/linziyu>

2013 9

1 / 19

第 13 章 Google Dreme l

厦门大学计算机科学系教师 林子雨 编著

<http://www.cs.xmu.edu.cn/linziyu>

<http://dblab.xmu.edu.cn/node/422>

2013 9

第 13 章 Google Dremel

Dremel

nested

CPU

Google

PB

2 3

PB

Dremel

MapReduce

Dremel

- Dremel
- Dremel
-
-
-

13.1 Dremel 概述

13.1.1 大规模数据分析

1

1TB

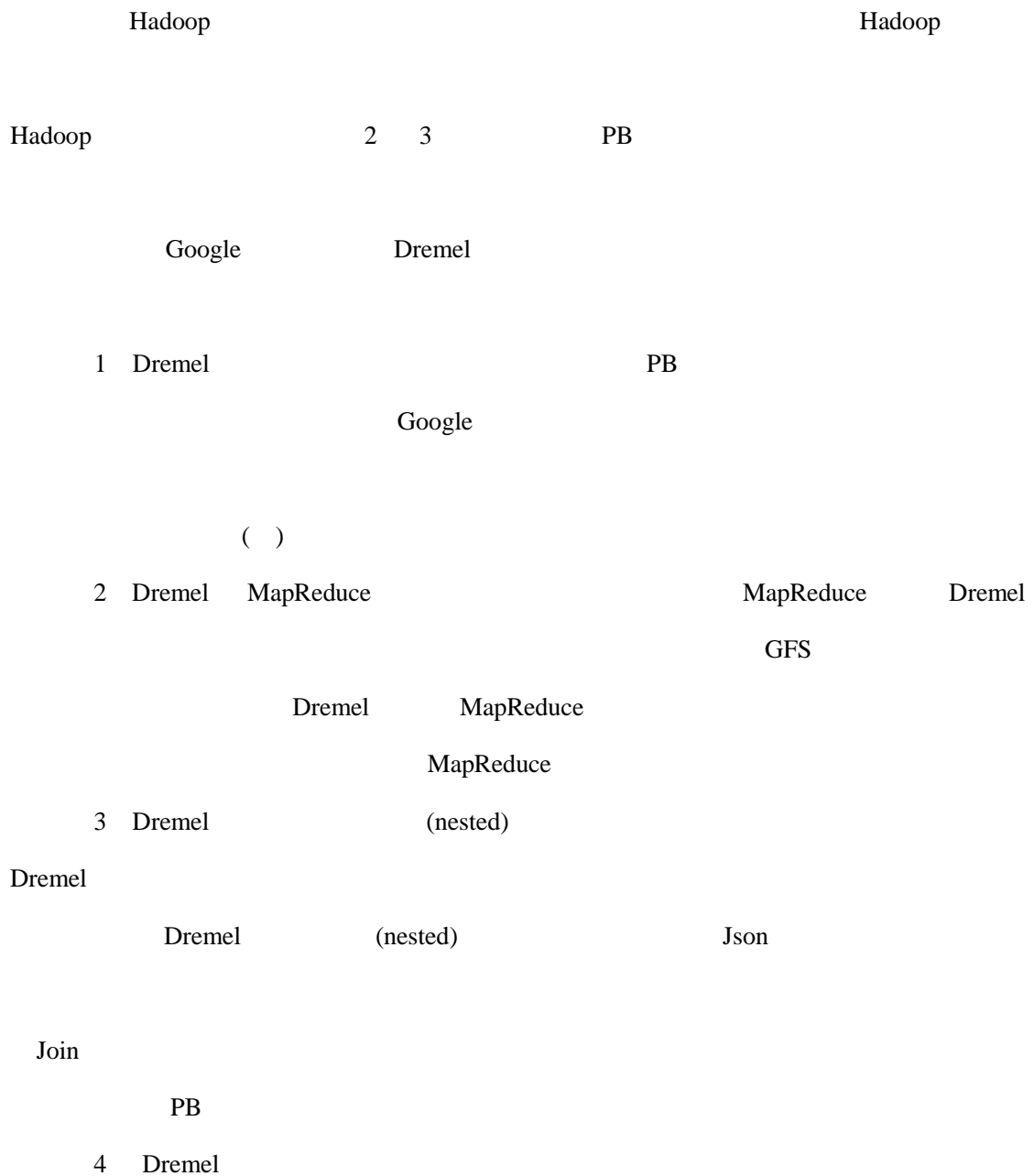
CPU

Google

PC

Google

13.1.2 Dremel 的特点



CPU

CPU

(nested) Dremel

5 Dremel Web DBMS Web

DBMS Dremel

SQL Hive Pig

Dremel 2006 Google

Dremel Google

Dremel

-
- Android
- Google
- Google Books OCR
-
- Google Maps
- Bigtable Tablet
- Google
- IO
- Google
- Google

13.1.3 Dremel 的应用场景

Google Alice

MapReduce

Dremel

```
DEFINE TABLE t AS /path/to/data/*
```

SELECT TOP(signal1, 100), COUNT(*) FROM t

signal1

FlumeJava

SQL

GFS Google File System

GFS

DBMS

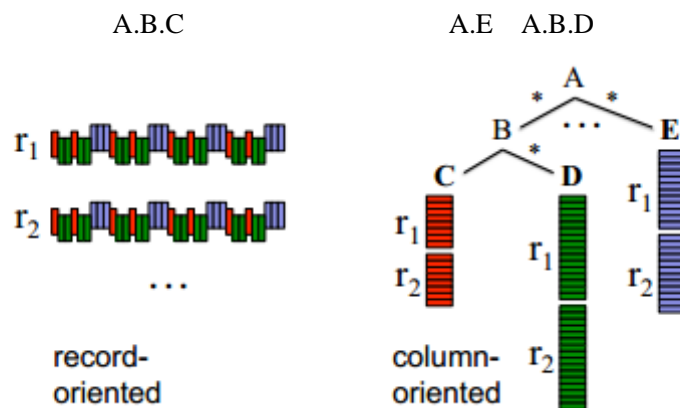
MapReduce

GFS

Google

13-1

A.B.C



13-1

13.2 Dremel 数据模型

Dremel

C++ Java

Java MapReduce

C++

MapReduce

13.3 嵌套列式存储

13-1

13.3.1 重复深度、定义深度

13-1

Dremel

13-2 Document

Google

Dremel

13-2 r1 r2 URL

["http://A", "http://B", "http://C"]

Name

DocId			Name.Url			Links.Forward			Links.Backward		
value	r	d	value	r	d	value	r	d	value	r	d
10	0	0	http://A	0	2	20	0	2	NULL	0	1
20	0	0	http://B	1	2	40	1	2	10	0	2
			NULL	1	1	60	1	2	30	1	2
			http://C	0	2	80	0	2			

Name.Language.Code			Name.Language.Country		
value	r	d	value	r	d
en-us	0	2	us	0	3
en	2	2	NULL	2	2
NULL	1	1	NULL	1	1
en-gb	1	2	gb	1	3
NULL	0	1	NULL	0	1

13-3

● **Repetition Level**

13-2	Code	r1	3	'en-us'	'en'
Name	'en-gb'	Name			
		Name.Language.Code		Name	
Language	Code	0	2	0	
		r1		'en-us'	
		0		'en'	Language
Language	Name.Language.Code	2		2	
'en-gb'	Name	Name	Language		
Language	Name	Name.Language.Code	1		
1		r1	r1	Code	'en-us'
'en-gb'		0	2	1	
	r1	Name	Code		
'en-gb'	Name	Name	NULL	'en'	
'en-gb'	13-3	Language	Code		
Language		Code		Code	

Language	r1	Name	Code
Language			Name.Language.Code

● **Definition Level**

NULL

Dremel r1

r2

NULL

NULL

Name.Language.Country	Name	Language	repeated
Country optional		Country	
Language			

r1 Backward

Links	1		Links.Backward
-------	---	--	----------------

NULL	1	Links	r2
------	---	-------	----

NULL	Name.Language.Country	1	Country
------	-----------------------	---	---------

Name	Name	1	
------	------	---	--

Name.Language.Country	1	r1	NULL
-----------------------	---	----	------

Name.Language.Country		2	Name.Language	1	Name
-----------------------	--	---	---------------	---	------

● **Encoding**

NULL

NULL

0 0

13-3	DocId	bit
	3	2 bit

13.3.2 将记录转换为列式存储

column-stripe 13-1

13-4

Google

schema

column-stripe

fieldWriter

schema

fieldWriter

fieldWriter

fieldWriter

13-4

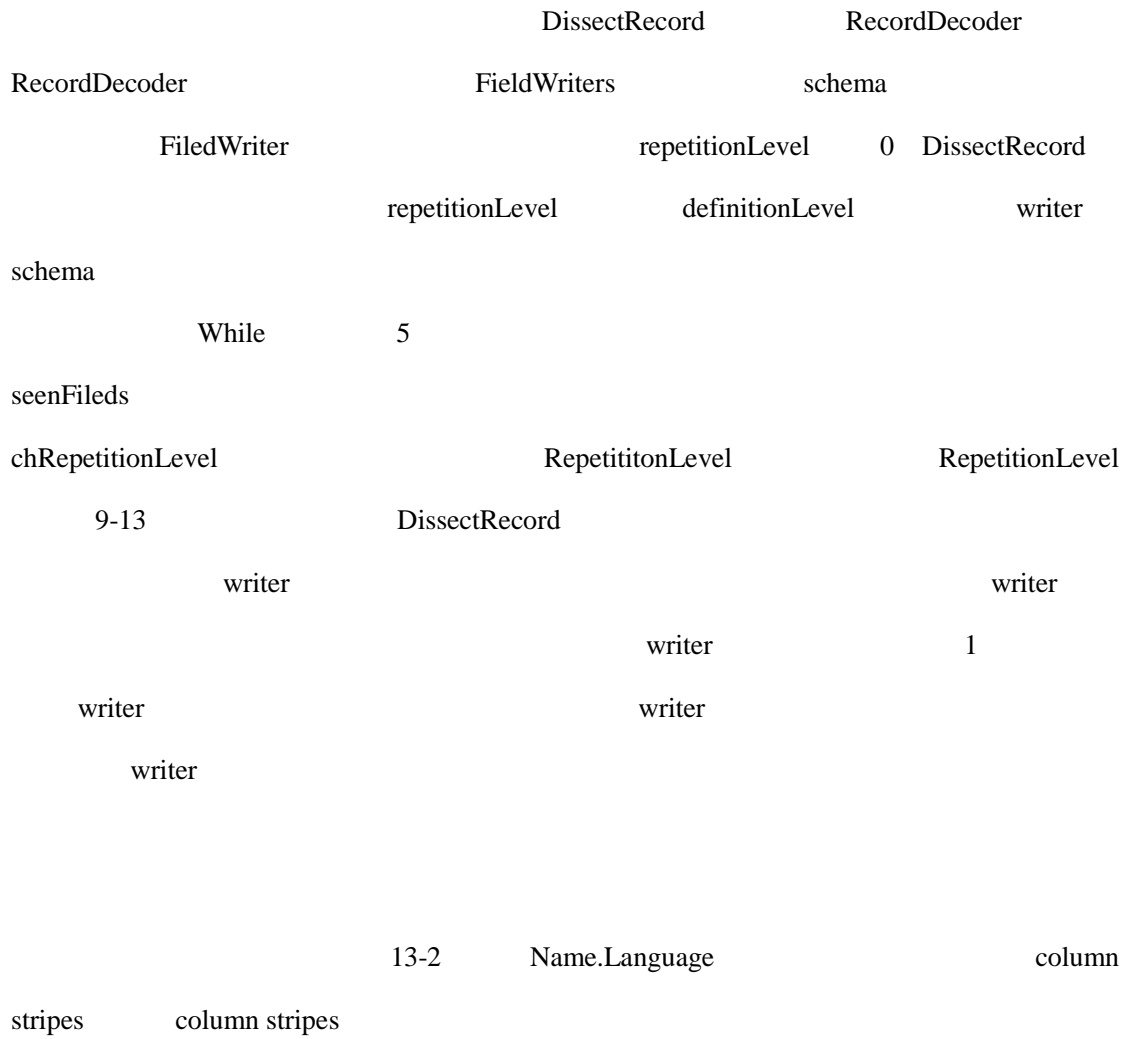
```

1 procedure DissectRecord(RecordDecoder decoder,
2     FieldWriter writer, int repetitionLevel):
3     Add current repetitionLevel and definition level to writer
4     seenFields = {} // empty set of integers
5     while decoder has more field values
6         FieldWriter chWriter =
7             child of writer for field read by decoder
8         int chRepetitionLevel = repetitionLevel
9         if set seenFields contains field ID of chWriter
10            chRepetitionLevel = tree depth of chWriter
11        else
12            Add field ID of chWriter to seenFields
13        end if
14        if chWriter corresponds to an atomic field
15            Write value of current field read by decoder
16            using chWriter at chRepetitionLevel
17        else
18            DissectRecord(new RecordDecoder for nested record
19                read by decoder, chWriter, chRepetitionLevel)
20        end if
21    end while
22 end procedure

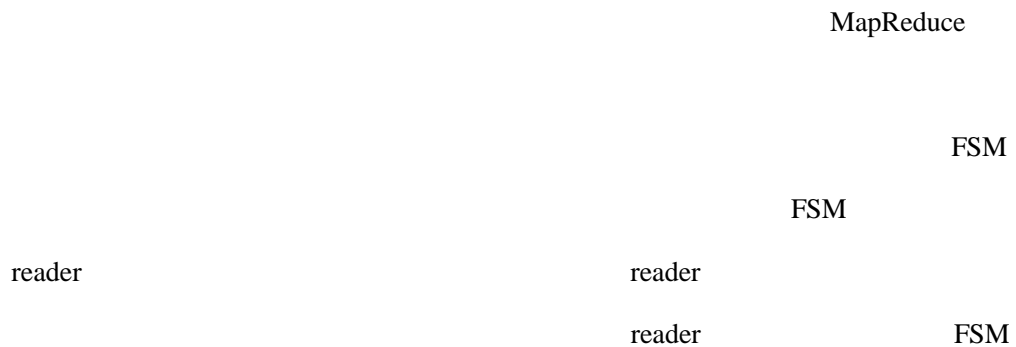
```

13-4

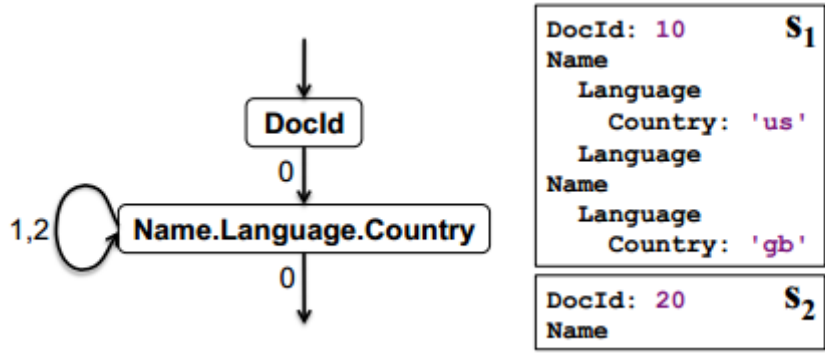
13-4



13.3.3 记录的装配



Name	Language	Name	Language
Name.Url	2	Name	Language
Language		Name.Language.Code	
Name	Language		
FSM		<i>l</i>	<i>f</i>
schema		<i>l</i>	
<i>n</i>	FSM	$(f;l) \rightarrow n$	$l=1 \quad f=Name.Language.Country$
		1	Name
$n=Name.Url$			



13-6

DocId	Name.Language.Country	FSM	13-6	FSM
	Country		s1 s2	Dremel
Country	Name	Language	XPath	
$/Name[2]/Language[1]/Country$				

13.4 查询语言

Dremel	SQL
	SQL
schema	schema
	13-7
	13-2 $t = \{r1,r2\}$

SQL

```

SELECT DocId AS Id,
       COUNT(Name.Language.Code) WITHIN Name AS Cnt,
       Name.Url + ',' + Name.Language.Code AS Str
FROM t
WHERE REGEXP(Name.Url, '^http') AND DocId < 20;

```

<pre> Id: 10 Name Cnt: 2 Language Str: 'http://A,en-us' Str: 'http://A,en' Name Cnt: 0 </pre>	<p>t₁</p>	<pre> message QueryResult { required int64 Id; repeated group Name { optional uint64 Cnt; repeated group Language { optional string Str; }}} </pre>
---	-----------------------------	---

13-7

WHERE

Name.Url

'http'

SELECT

schema

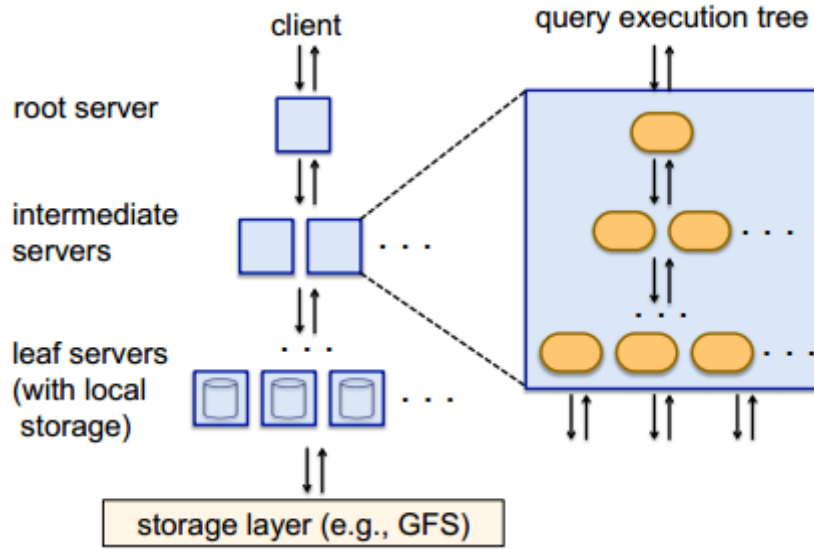
Name.Language.Code	3	Str	Count
Name			64
Name.Language.Code	Name		

top-k joins

13.5查询的执行

serving-

tree



13-8

-

Dremel

13-8

SELECT A, COUNT(B) FROM T GROUP BY A

tablet

column-stripe

table

T table

T

tablet

T

SELECT A, SUM(c) FROM (R_1^1 UNION ALL ... R_n^1) GROUP BY A

R_1^1 R_n^1

1

1

n

R_i^1 = SELECT A, COUNT(B) AS c FROM T_i^1 GROUP BY A

T_i^1

T

1

i

tablet

rewrite

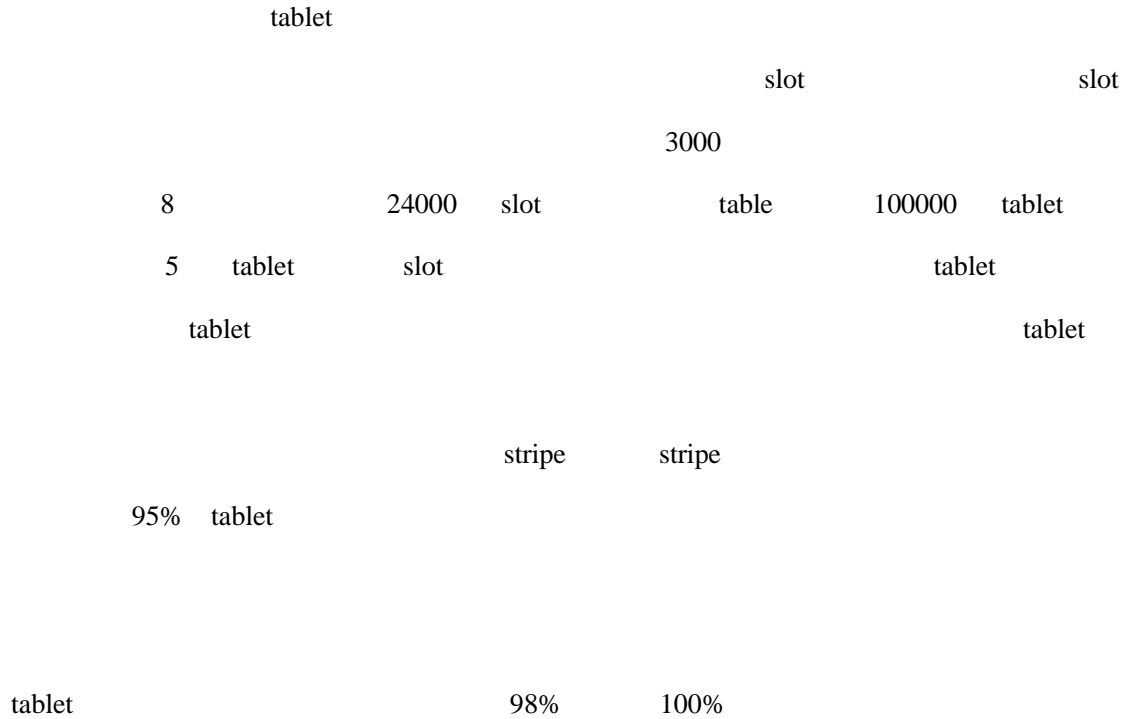
T

tablet

DBMS MR



Dremel



本章小结

Dremel Dremel

Dremel

MapReduce

参考文献

[1] Melnik, Sergey, et al. "Dremel: interactive analysis of web-scale datasets." Proceedings of the VLDB Endowment 3.1-2 (2010): 330-339.

[2] . Dremel: Interactive Analysis of WebScale Datasets .
<http://blog.csdn.net/macyang/article/details/8566105>

[3] . Google Dremel 3 1PB. <http://blog.jobbole.com/29561/#jtss-tsina>

