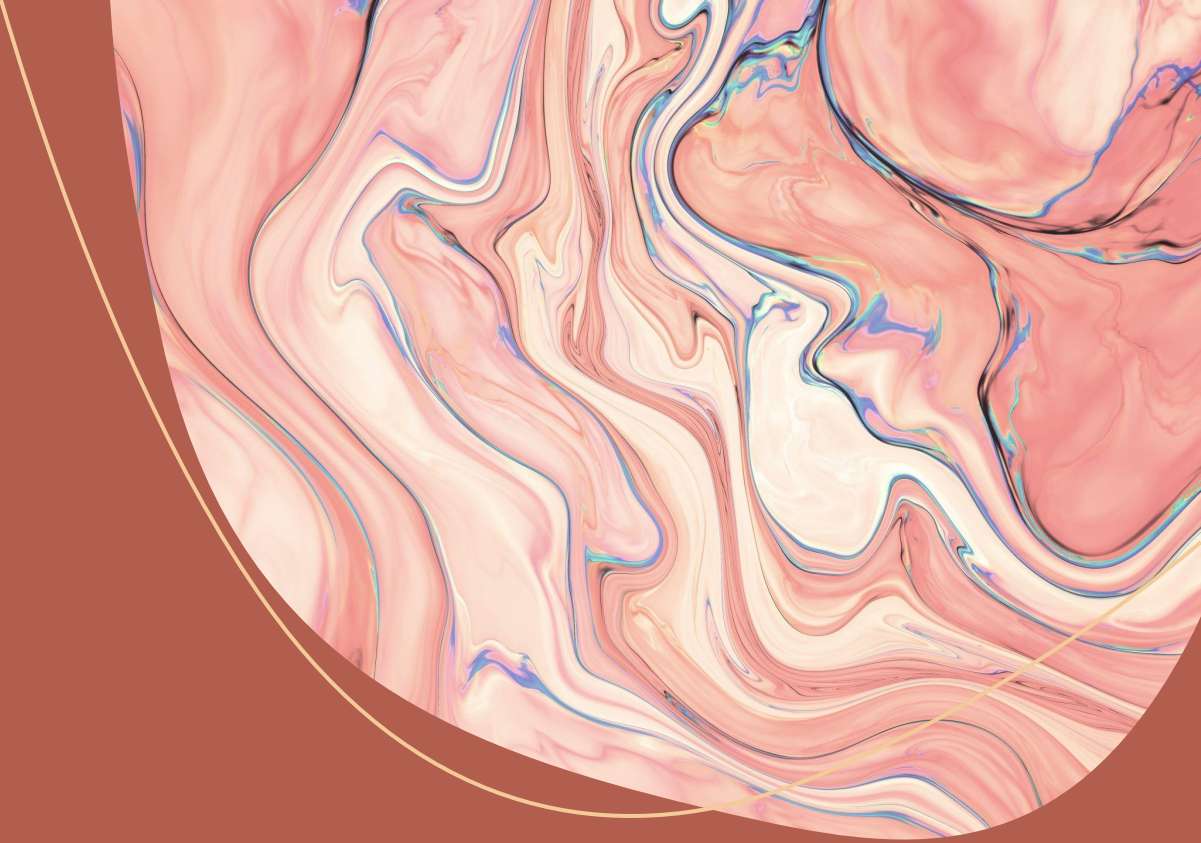


Emulated Distributed File System

Jinqiu Fan, Yating Yang



Overview

- The main objective is to build an EDFS and implement PMR on data stored in EDFS.
- Our project is about California Wildfires in 2018 and weather factors that might lead to the wildfires.

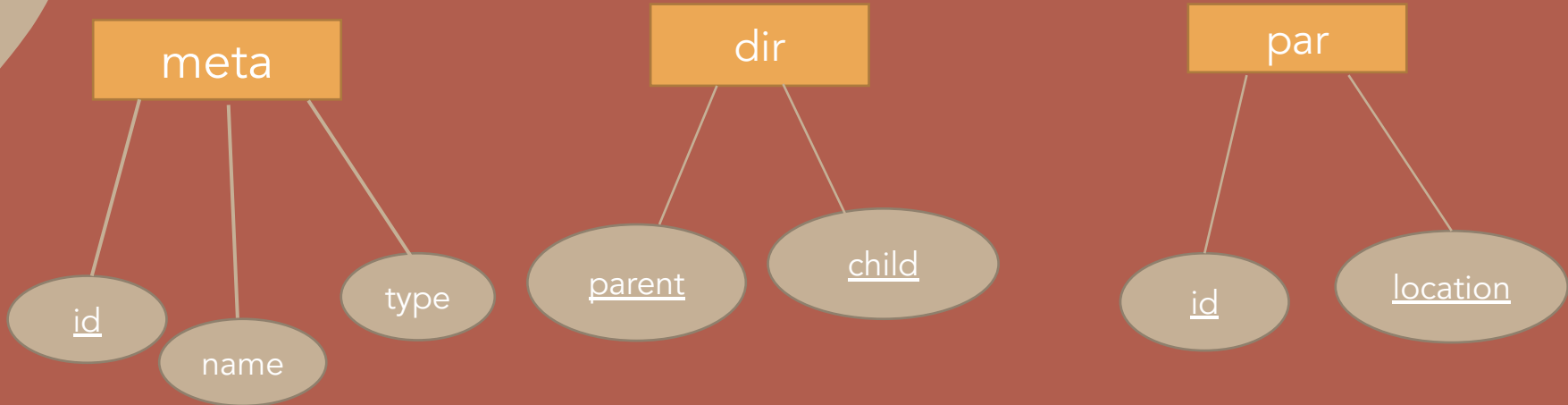
wildfire					
AcresBurned	ArchiveYear	Counties	Name	Started	Extinguished
0	2018	Santa Barbara	Montecito Flooding / Mudflows	2018-01-09	2019-01-03
0	2018	Ventura	Montecito Flooding / Mudflows	2018-01-09	2019-01-03
50	2018	Monterey	Haycamp Fire	2018-02-07	2019-01-03
33	2018	Tulare	Wilcox Fire	2018-02-18	2019-01-03

weather				
Region	Months	Precipitation(in)	AvgTemperature(F)	AvgHumidity(%)
Bishop	Jan	0.00	41.79	45.57
Bishop	Feb	0.00	40.26	32.43
Bishop	Mar	0.01	46.72	32.86
Bishop	Apr	0.01	58.05	25.39

- Task1: My-SQL based EDFS
- Task2: Search function: search_wildfire(lower, upper)
Search function: search_weather(choice, month)
Analytics function: analyze_wildfire(lower, upper, num)
- Task3: Web application based on Jupyter Notebook

Database Structure

Metadata:



Actual Data:

Partitioned file:

wildfire1

wildfire2

wildfire3

AcresBurned	ArchiveYear	Counties	Name	Started	Extinguished
0	2018	Santa Barbara	Montecito Flooding / Mudflows	1/9/18	1/3/19
0	2018	Ventura	Montecito Flooding / Mudflows	1/9/18	1/3/19
50	2018	Monterey	Haycamp Fire	2/7/18	1/3/19
33	2018	Tulare	Wilcox Fire	2/18/18	1/3/19
2070	2018	Inyo	Pleasant Fire	2/18/18	1/3/19
258	2018	Santa Barbara	Santa Cruz Fire	3/27/18	1/3/19
11	2018	San Luis Obispo	San Luis Obispo Fire	3/27/18	1/3/19

mkdir(directory)

eg. mkdir /user

meta

id	name	type
1	/	d
2	user	d

dir

parent	child
1	2

partition

id	location

ls(directory)

eg. ls /

meta

id	name	type
1	/	d
2	user	d

dir

parent	child
1	2

partition

id	location

put(file name, directory, k = # of partition)

put(wildfire.csv, /user, 2)

Meta Data

meta

id	name	type
1	/	d
2	user	d
3	wildfire.csv	f

Actual Data

wildfire1
wildfire2

wildfire1

AcresBurned	ArchiveYear	Counties	Name	Started	Extinguished
0	2018	Santa Barbara	Montecito Flooding / Mudflows	1/9/18	1/3/19
0	2018	Ventura	Montecito Flooding / Mudflows	1/9/18	1/3/19
50	2018	Monterey	Haycamp Fire	2/7/18	1/3/19
33	2018	Tulare	Wilcox Fire	2/18/18	1/3/19
2070	2018	Inyo	Pleasant Fire	2/18/18	1/3/19
258	2018	Santa Barbara	Santa Cruz Fire	3/27/18	1/3/19
14	2018	Chico	Rock Fire	3/27/18	1/3/19

parent	child
1	2
2	3

partition

id	location
3	wildfire1
3	wildfire2

cat(file path)

cat /user/wildfire.csv

Meta Data

meta

id	name	type
1	/	d
2	user	d
3	wildfire.csv	f

Actual Data

wildfire1
wildfire2

wildfire1

AcresBurned	ArchiveYear	Counties	Name	Started	Extinguished
0	2018	Santa Barbara	Montecito Flooding / Mudflows	1/9/18	1/3/19
0	2018	Ventura	Montecito Flooding / Mudflows	1/9/18	1/3/19
50	2018	Monterey	Haycamp Fire	2/7/18	1/3/19
33	2018	Tulare	Wilcox Fire	2/18/18	1/3/19
2070	2018	Inyo	Pleasant Fire	2/18/18	1/3/19
258	2018	Santa Barbara	Santa Cruz Fire	3/27/18	1/3/19
14	2018	Chico	Rock Fire	3/27/18	1/3/19

dir

parent	child
1	2
2	3

partition

id	location
3	wildfire1
3	wildfire2

getPartitionLocations(file path)

getPartitionLocations(/user/wildfire.csv) dir

Meta Data

meta

id	name	type
1	/	d
2	user	d
3	wildfire.csv	f

parent	child
1	2
2	3

partition

id	location
3	wildfire1
3	wildfire2

Actual Data

wildfire1
wildfire2

wildfire1

AcresBurned	ArchiveYear	Counties	Name	Started	Extinguished
0	2018	Santa Barbara	Montecito Flooding / Mudflows	1/9/18	1/3/19
0	2018	Ventura	Montecito Flooding / Mudflows	1/9/18	1/3/19
50	2018	Monterey	Haycamp Fire	2/7/18	1/3/19
33	2018	Tulare	Wilcox Fire	2/18/18	1/3/19
2070	2018	Inyo	Pleasant Fire	2/18/18	1/3/19
258	2018	Santa Barbara	Santa Cruz Fire	3/27/18	1/3/19
14	2018	Chico	Rock Fire	3/27/18	1/3/19

readPartition(file path, k)

readPartition(/use/wildfire.csv, 2)

Meta Data

meta

id	name	type
1	/	d
2	user	d
3	wildfire.csv	f

Actual Data

wildfire1
wildfire2

wildfire1

AcresBurned	ArchiveYear	Counties	Name	Started	Extinguished
0	2018	Santa Barbara	Montecito Flooding / Mudflows	1/9/18	1/3/19
0	2018	Ventura	Montecito Flooding / Mudflows	1/9/18	1/3/19
50	2018	Monterey	Haycamp Fire	2/7/18	1/3/19
33	2018	Tulare	Wilcox Fire	2/18/18	1/3/19
2070	2018	Inyo	Pleasant Fire	2/18/18	1/3/19
258	2018	Santa Barbara	Santa Cruz Fire	3/27/18	1/3/19
14	2018	Chico	Rock Fire	3/27/18	1/3/19

dir

parent	child
1	2
2	3

partition

id	location
3	wildfire1
3	wildfire2

rm(file path)

rm /user/wildfire.csv

Meta Data

meta

id	name	type
1	/	d
2	user	d
3	wildfire.csv	f

Actual Data

wildfire1
wildfire2

dir

parent	child
1	2
2	3

partition

id	location
3	wildfire1
3	wildfire2

wildfire1

AreaBurned	ArchiveYear	Counties	Name	Started	Extinguished
0	2018	Santa Barbara	Montecito Flooding / Mudflows	1/9/18	1/3/19
0	2018	Ventura	Montecito Flooding / Mudflows	1/9/18	1/3/19
50	2018	Monterey	Haycock Fire	2/7/18	1/3/19
33	2018	Tulare	Wilcox Fire	2/18/18	1/3/19
2070	2018	Inyo	Pleasant Fire	2/18/18	1/3/19
258	2018	Santa Barbara	Santa Cruz Fire	3/27/18	1/3/19
11	2018	Chico	Rock Fire	3/27/18	1/3/19