AirMettle™

Real-Time Smart Data Lake for Accelerated In-Place Analytics of Scientific Data

Donpaul C. Stephens
donpaul@airmettle.com
Founder, AirMettle, Inc.
+1-646-872-2124

HDF5 User Group Meeting

August 2023

AirMettle's Mission

AirMettle has developed a real-time smart data lake solution that simplifies big data analytics and accelerates processing by an order of magnitude, or more.

It is implemented in the data lake storage layer and performs basic analytics tasks that:

- Reduce network traffic - Extract only what is needed before returning
- Improve data freshness - All data can be rapidly queried
- Enable real-time operation - Sub-second queries on GByte objects from storage

AirMettle's Mission

AirMettle has developed a real-time smart data lake solution that simplifies big data analytics and accelerates processing by an order of magnitude, or more.

It is implemented in the data lake storage layer and performs basic analytics tasks that:

- Reduce network traffic - Extract only what is needed before returning
- Improve data freshness - All data can be rapidly queried
- Enable real-time operation - Sub-second queries on GByte objects from storage

STTR

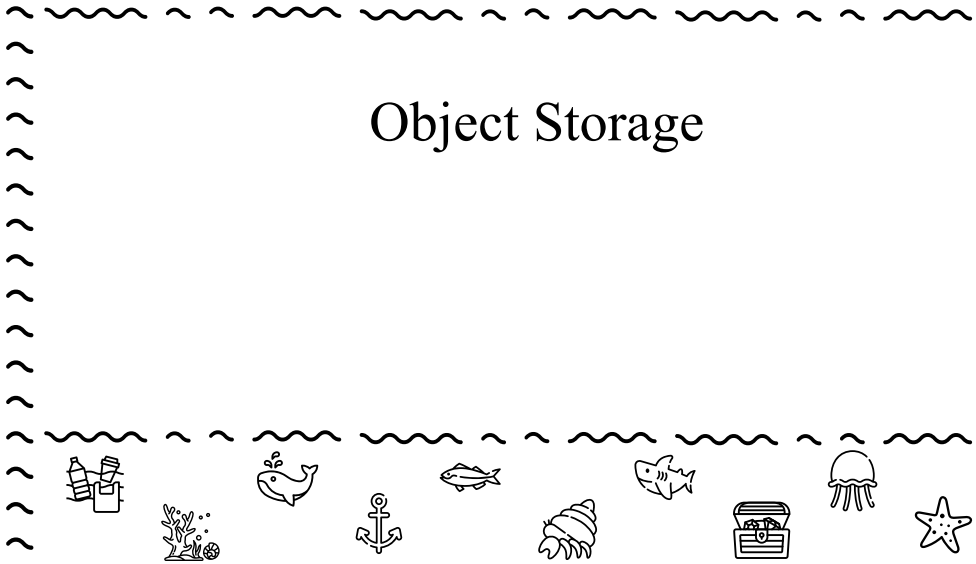
Los Alamos NATIONAL LABORATORY

SBIR

https://www.nsf.gov/awardsearch/showAward?AWD_ID=2135007&HistoricalAwards=false

https://techpartnerships.noaa.gov/sbir/awards/

Traditional Data Lake



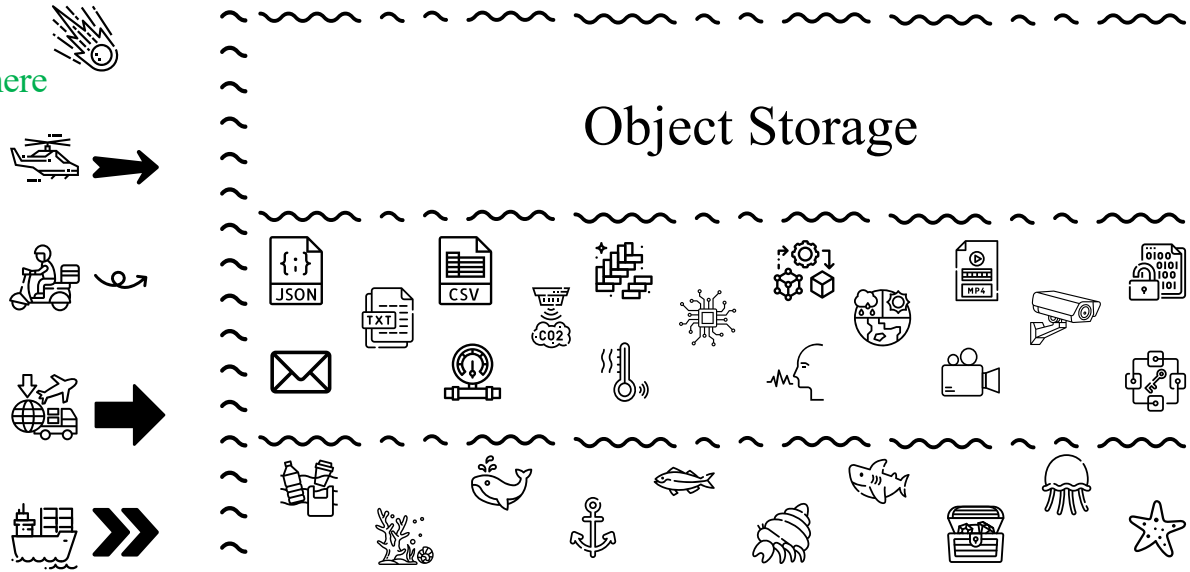
Objects are internally partitioned
For storage in parallel

Data Lake

Traditional Data Lake:

Data generally arrives semi-structured

Comes from Everywhere



Objects are internally partitioned
For storage in parallel

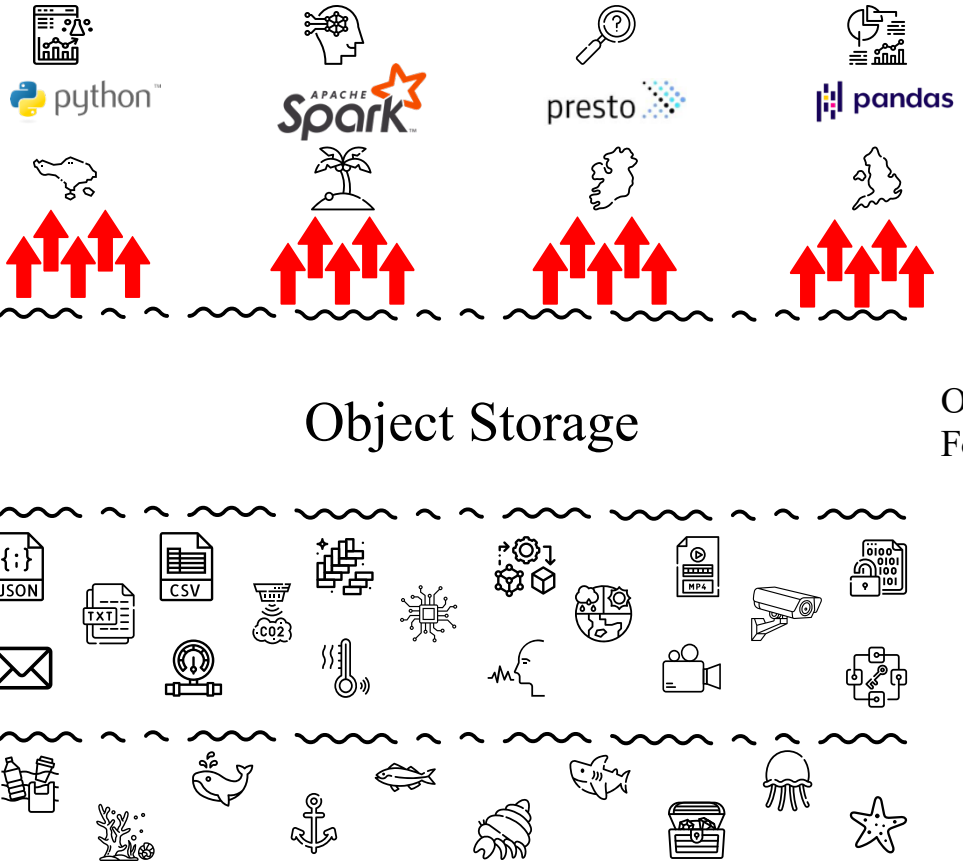
Primarily
Semi-structured data

Data Lake

Traditional Data Lake:

Data must be moved to gain value from it

Analyzed
In Islands



Applications retrieve full objects*
To their own (small) clusters
for processing

Comes
from
Everywhere

Objects are internally partitioned
For storage in parallel

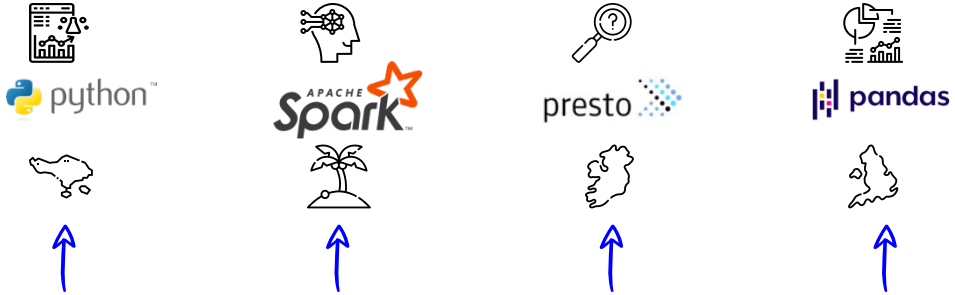
Primarily
Semi-structured data

Data Lake

Smart Data Lake: Get only what is needed 100x Faster

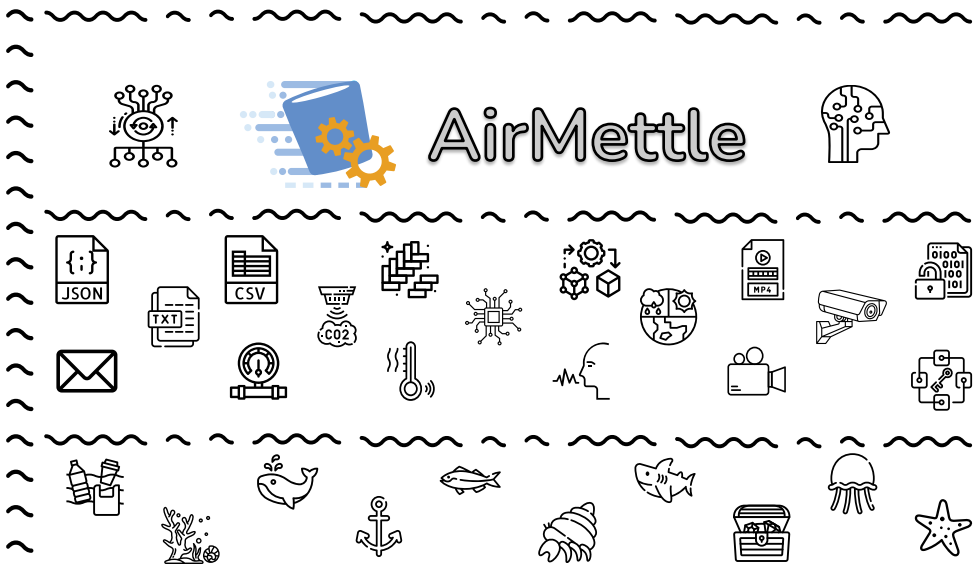
Delegate simple processing to storage thereby **eliminating the need** to move 90%+ of the data

Analyzed **Faster**
In Islands



Retrieve what they need in an **Immediately** usable form

Comes from Everywhere



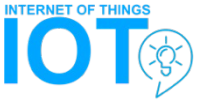
Objects are internally partitioned
For storage **& processing** in parallel

Primarily
Semi-structured data

Data Lake

Accelerated analytics of classic tabular data

Security Information & Event Management



- Scan historical data to diagnose current events
  - Determine how many records might be relevant before retrieving any

Natural Language Processing

- Search for key-words
  - Gather statistics of usage
  - Extract text if required for further analysis

Accelerated analytics of classic tabular data (S3 Select API)

Security Information & Event Management

SIEM

INTERNET OF THINGS IOT

- Scan historical data to diagnose current events
  - Determine how many records might be relevant before retrieving any

Natural Language Processing

- Search for key-words
  - Gather statistics of usage
  - Extract text if required for further analysis

Validated with WISCONSIN & MIT

Star Schema Benchmark
Utilized 223 Select queries to Object Storage

100 X faster

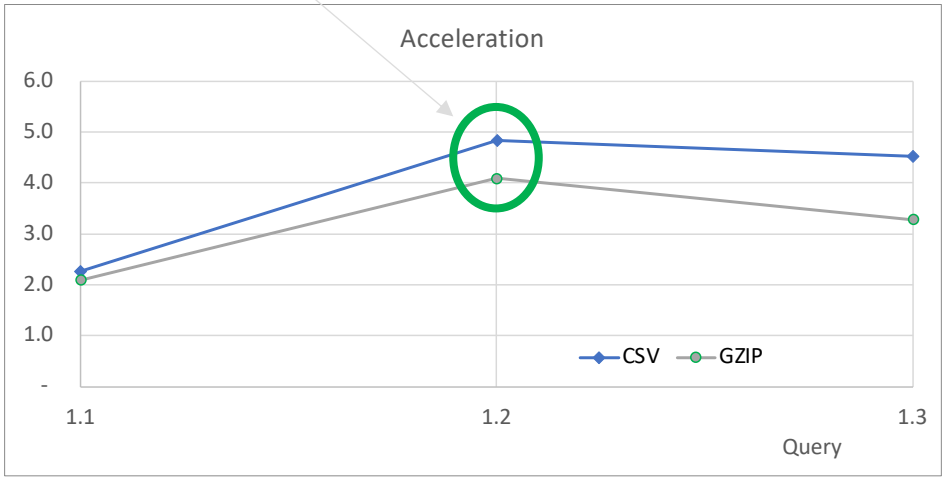
Under a minute vs. 1 hour 45min

Unprecedented speed of analysis: Directly from storage

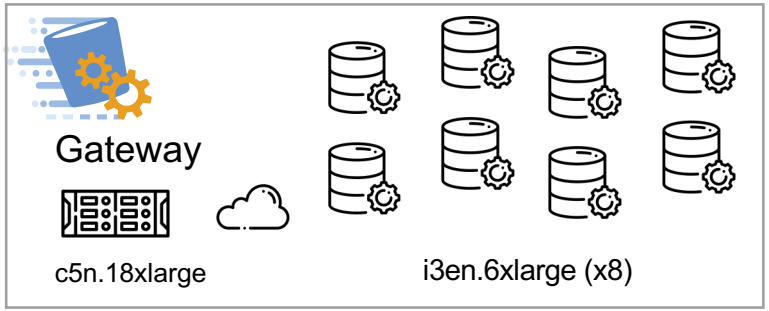
No data warehouse required

AirMettle Accelerates

5x Acceleration on Complete Queries, today… just by using a different storage



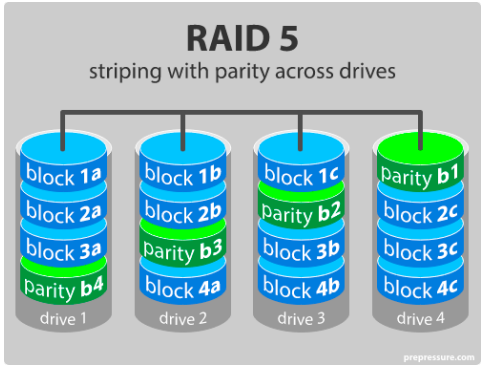
Star Schema Benchmark, Scale Factor 1 with 1 object per table

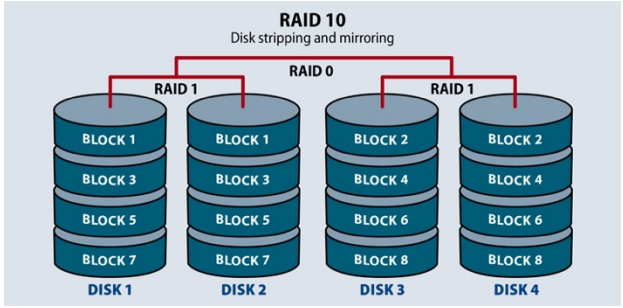


S3 Select API enables comparison vs. major cloud's object storage

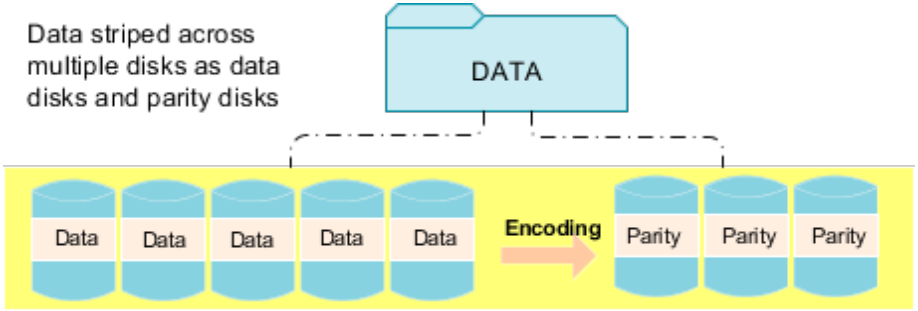
Resiliency 101: How do storage solutions protect data?

RAID:





Erasure Coding:

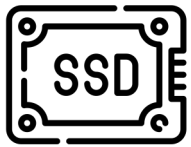


Data protection algorithms designed for HDD

What that means for data reliably placed in storage: First 4 devices shown…

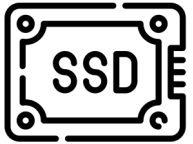
Simple Table:

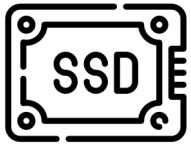
1	155190	7706	1	17	21168.23	0.04	0.02	N	O	3/13/96	2/12/96	3/22/96	DELIVER_IN_PERSON	TRUCK	egular_courts_above_the
1	67310	7311	2	36	45983.16	0.09	0.06	N	O	4/12/96	2/28/96	4/20/96	TAKE_BACK_RETURN	MAIL	ly_final_dependencies:_slyly_bold_
1	63700	3701	3	8	13309.6	0.1	0.02	N	O	1/29/96	3/5/96	1/31/96	TAKE_BACK_RETURN	REG_AIR	riously._regular|_express_dep
1	2132	4633	4	28	28955.64	0.09	0.06	N	O	4/21/96	3/30/96	5/16/96	NONE	AIR	lites._fluffily_even_de
1	24027	1534	5	24	22824.48	0.1	0.04	N	O	3/30/96	3/14/96	4/1/96	NONE	FOB	_pending_foxes._slyly_re
1	15635	638	6	32	49620.16	0.07	0.02	N	O	1/30/96	2/7/96	2/3/96	DELIVER_IN_PERSON	MAIL	arefully_slyly_ex
2	106170	1191	1	38	44694.46	0	0.05	N	O	1/28/97	1/14/97	2/2/97	TAKE_BACK_RETURN	RAIL	ven_requests._deposits_breach_a



#1

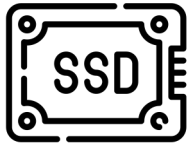
#2





#3

#4



Bytes of data divided evenly across SSDs!

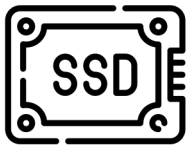
Data protection and streaming performance!

Supports data protection algorithms designed for HDD!

What that means for data reliably placed in storage: First 4 devices shown…

Simple Table:

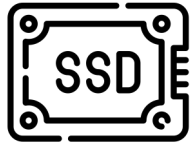
1	155190	7706	1	17	21168.23	0.04	0.02	N	O	3/13/96	2/12/96	3/22/96	DELIVER_IN_PERSON	TRUCK	egular_courts_above_the
1	67310	7311	2	36	45983.16	0.09	0.06	N	O	4/12/96	2/28/96	4/20/96	TAKE_BACK_RETURN	MAIL	ly_final_dependencies:_slyly_bold_
1	63700	3701	3	8	13309.6	0.1	0.02	N	O	1/29/96	3/5/96	1/31/96	TAKE_BACK_RETURN	REG_AIR	riously._regular|_express_dep
1	2132	4633	4	28	28955.64	0.09	0.06	N	O	4/21/96	3/30/96	5/16/96	NONE	AIR	lites._fluffily_even_de
1	24027	1534	5	24	22824.48	0.1	0.04	N	O	3/30/96	3/14/96	4/1/96	NONE	FOB	_pending_foxes._slyly_re
1	15635	638	6	32	49620.16	0.07	0.02	N	O	1/30/96	2/7/96	2/3/96	DELIVER_IN_PERSON	MAIL	arefully_slyly_ex
2	106170	1191	1	38	44694.46	0	0.05	N	O	1/28/97	1/14/97	2/2/97	TAKE_BACK_RETURN	RAIL	ven_requests._deposits_breach_a



#1

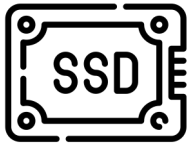
1,155190,7706,1,17,21168.23,0.04,0.02,N,O,1996-03-13
,1996-02-12,1996-03-22,DELIVER_IN_PERSON,TRUCK,egula
r_courts_above_the, 1,67310,7311,2,36,45983.16,0.09,
0.06,N,O,1996-04-12,1996-02-28,1996-04-20,TAKE_BACK_

#2



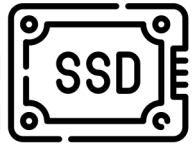
RETURN,MAIL,ly_final_dependencies:_slyly_bold_, 1,63
700,3701,3,8,13309.60,0.10,0.02,N,O,1996-01-29,1996-
03-05,1996-01-31,TAKE_BACK_RETURN,REG_AIR,riously._r
egular|_express_dep, 1,2132,4633,4,28,28955.64,0.09,

#3



0.06,N,O,1996-04-21,1996-03-30,1996-05-16,NONE,AIR,l
ites._fluffily_even_de, 1,24027,1534,5,24,22824.48,0
.10,0.04,N,O,1996-03-30,1996-03-14,1996-04-01,NONE,F
OB,_pending_foxes._slyly_re, 1,15635,638,6,32,49620.

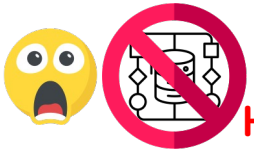
#4



16,0.07,0.02,N,O,1996-01-30,1996-02-07,1996-02-03,DE
LIVER_IN_PERSON,MAIL,arefully_slyly_ex, 2,106170,119
1,1,38,44694.46,0.00,0.05,N,O,1997-01-28,1997-01-14,
1997-02-02,TAKE_BACK_RETURN,RAIL,ven_requests._depo

Bytes of data divided evenly across SSDs!

Data protection and streaming performance!



HDD-centric RAID/Erasure Coding prevent in-storage analytics

AirMettle: Data partitioning for processing AND protecting data

AirMettle

- Data is unchanged for client
- Each internal component can be processed in parallel

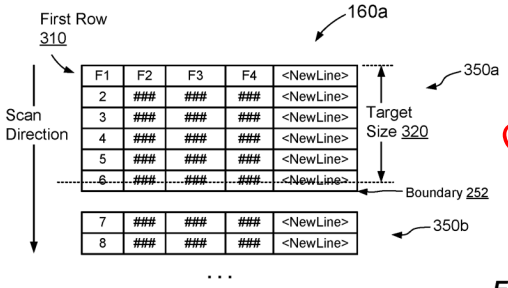


FIG. 3A

Object's own metadata

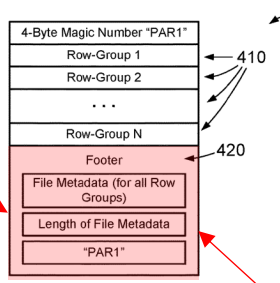


FIG. 4A

internal metadata

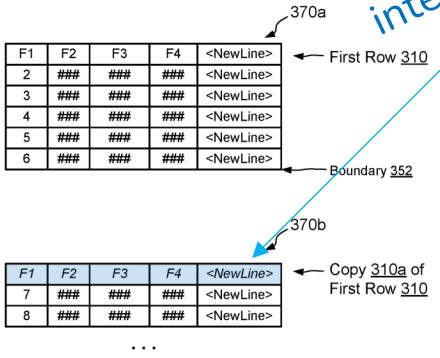


FIG. 3B

Not to scale!
Meta-data typically <0.1% of data

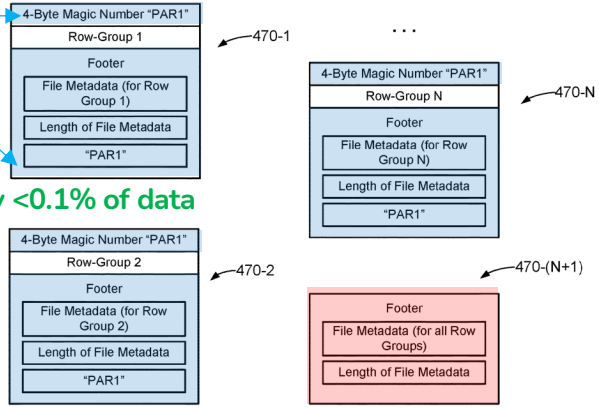


FIG. 4B

AirMettle Patented IP

AirMettle internal metadata enables parallel in-storage analytics

AirMettle: Data partitioning for processing AND protecting data

- Data is unchanged for client
- Each internal component can be processed in parallel

AirMettle Patented IP

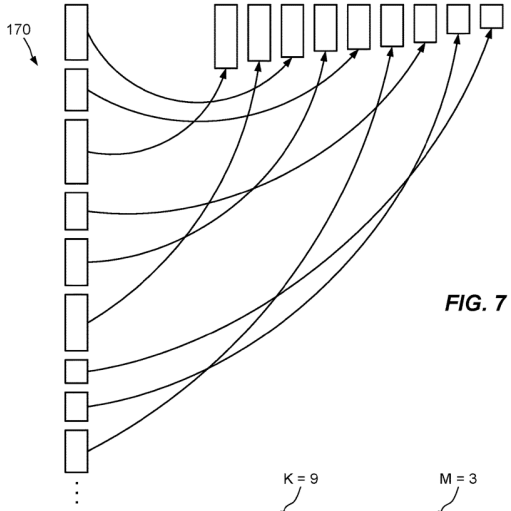


FIG. 7

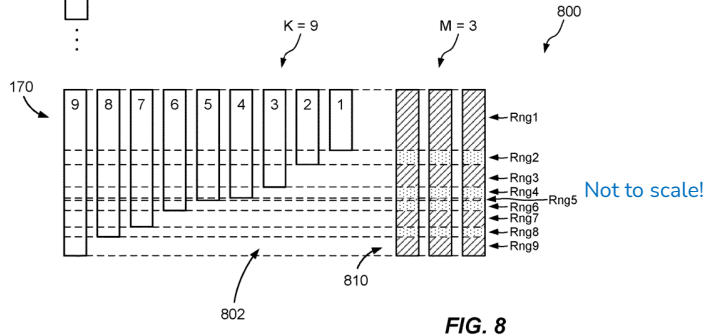


FIG. 8

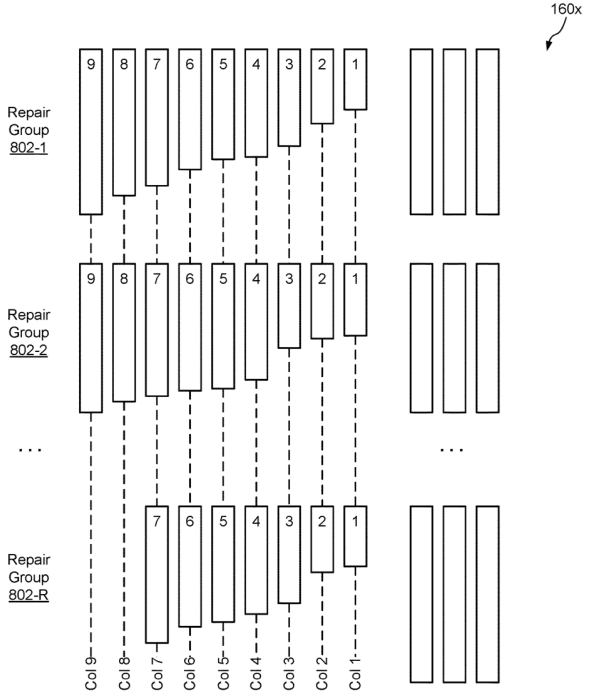
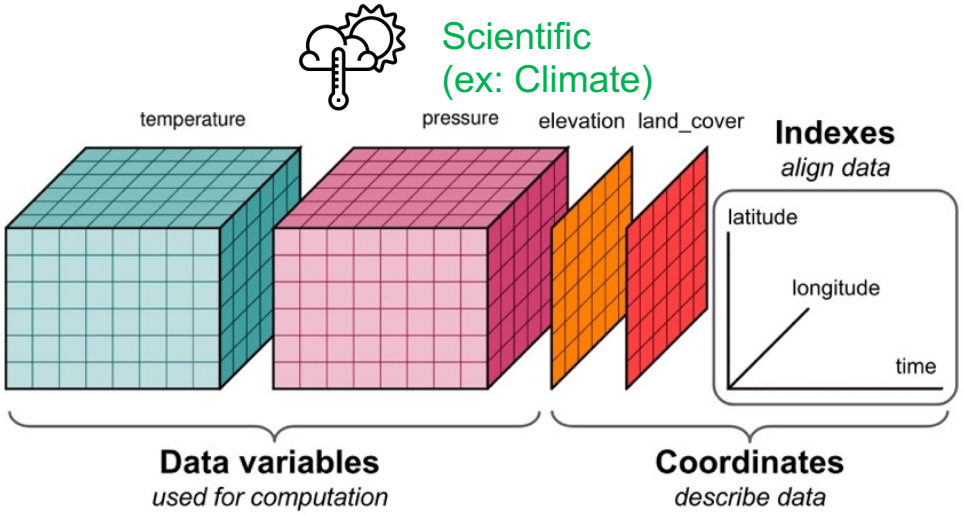


FIG. 9

AirMettle protection algorithms are designed for our non-uniform data segments

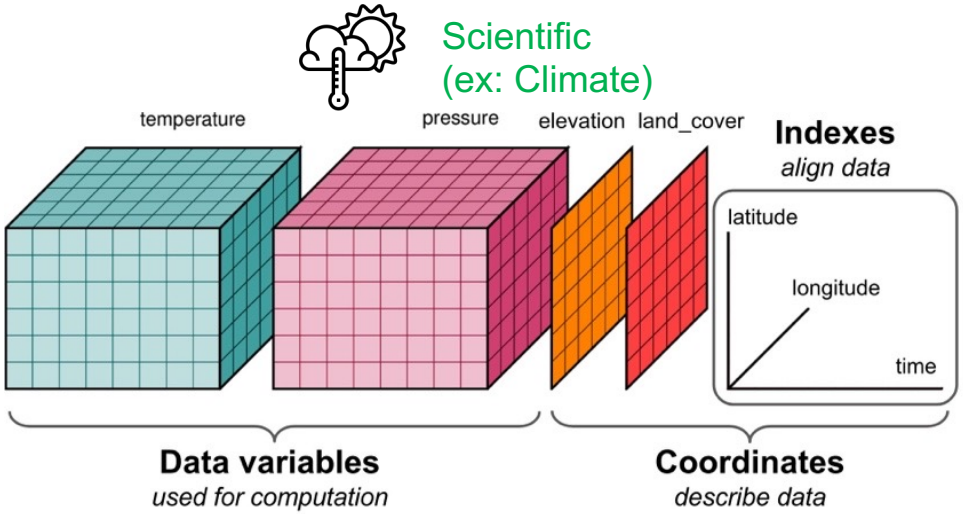
AirMettle Accelerates Multi-dimensional data

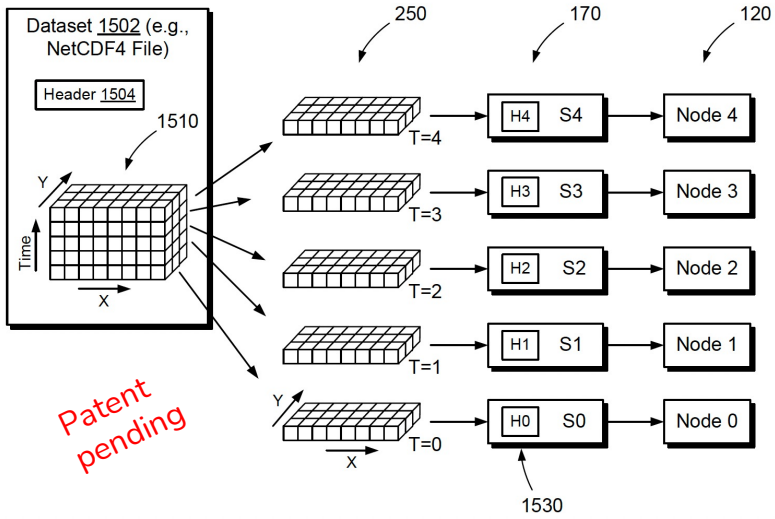


- Complex data format!
- Even the coordinates are multi-dimensional

- Scientific analysis also requires:
  - Sampling of different subsets for each query
  - Gathering statistical properties

AirMettle Accelerates Multi-dimensional data



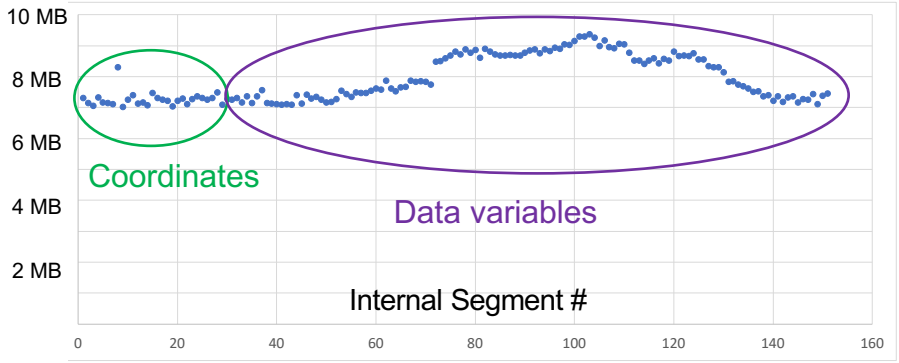


- Complex data format!
- Even the coordinates are multi-dimensional

- Scientific analysis also requires:
  - Sampling of different subsets for each query
  - Gathering statistical properties

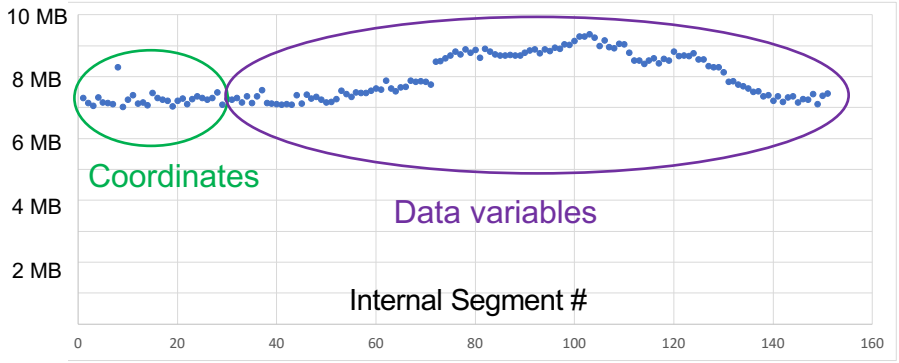
- Data is stored in partitions, based on semi-structural boundaries & size of segment
  - Enables distributed in-place parallel processing
  - Validated in NOAA Phase I
  - Queries are more complex than events w/ SQL
    - 2-3 Stage queries typical, but each stage can be executed in parallel

Initial Support: Selection & Simple Aggregations via REST APIs



- Selection returned as "smaller" NetCDF4
- Partitioned tensor data can be efficiently scanned in parallel
- Multi-stage processing:
  - Check coordinates
    - Bounding box & potential mask
    - Identifies components holding relevant data
  - Scan / Filter desired data
  - Integrate result for return

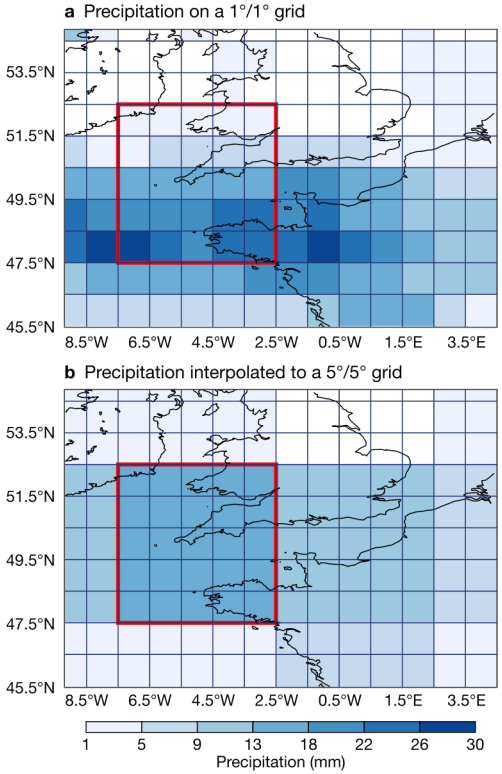
Initial Support: Selection & Simple Aggregations via REST APIs



- Selection returned as "smaller" NetCDF4
- Partitioned tensor data can be efficiently scanned in parallel
- Multi-stage processing:
  - Check coordinates
    - Bounding box & potential mask
    - Identifies components holding relevant data
  - Scan / Filter desired data
  - Integrate result for return

- Challenges included:
  - User-transparent partitioning
    - Internal data placement was more complex than originally anticipated
    - Typical internal overhead <0.3%
      - Dwarfed by data protection overheads
        - Erasure coding typically 20% to 35%
  - HDF5 does not have a re-entrant library
    - Developed support for concurrent analytics
    - Enables massive parallelism required for a shared storage service
  - Query engine required extensions for multi-stage execution & tensors

NOAA Phase II: Regridding



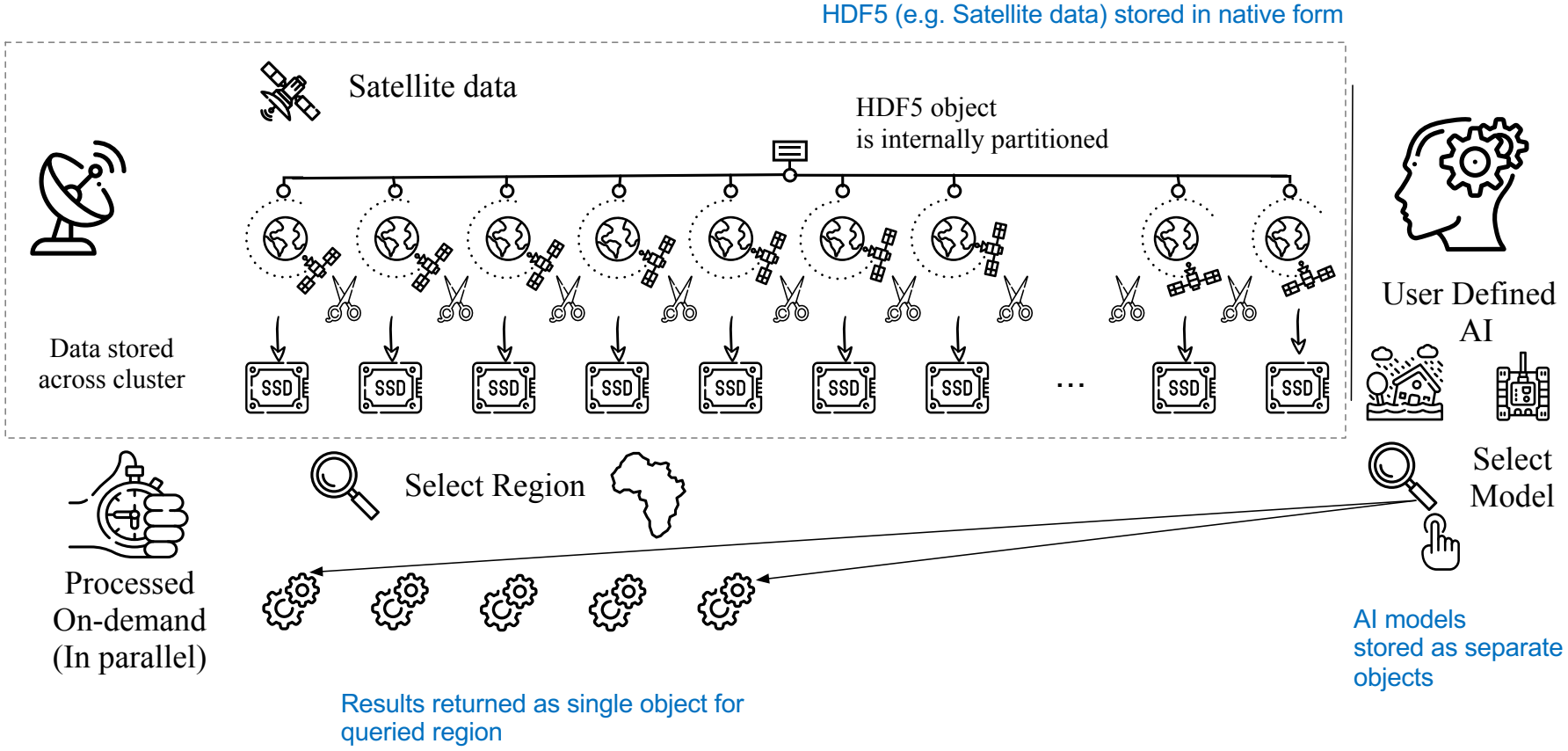
Precipitation on 1º/1º and 5 º/5 º grids

- Very common operation to convert from higher-resolution to lower-resolution.
  - When combining data from different sources, the final result can only be the lowest common resolution among the different sources.
- Project formally commenced Aug 1st
  - Developing parallel re-gridding on internal partitions

- Coming Summer ’24
  - Multi-dimensional in-place analytics

- Plan to explore alternative aggregation functions (Min/Max/Std. Dev) to enable faster analysis
  - Potential storm fronts, etc.

Roadmap: Hierarchical support & User Defined Functions! AI in-place



Acknowledgement

This work was supported with government support under
- 2135007 awarded by the National Science Foundation
- NA22OAR0210591 awarded by the National Oceanic and Atmospheric Administration
- NA23OAR0210342 awarded by the National Oceanic and Atmospheric Administration

AirMettle™

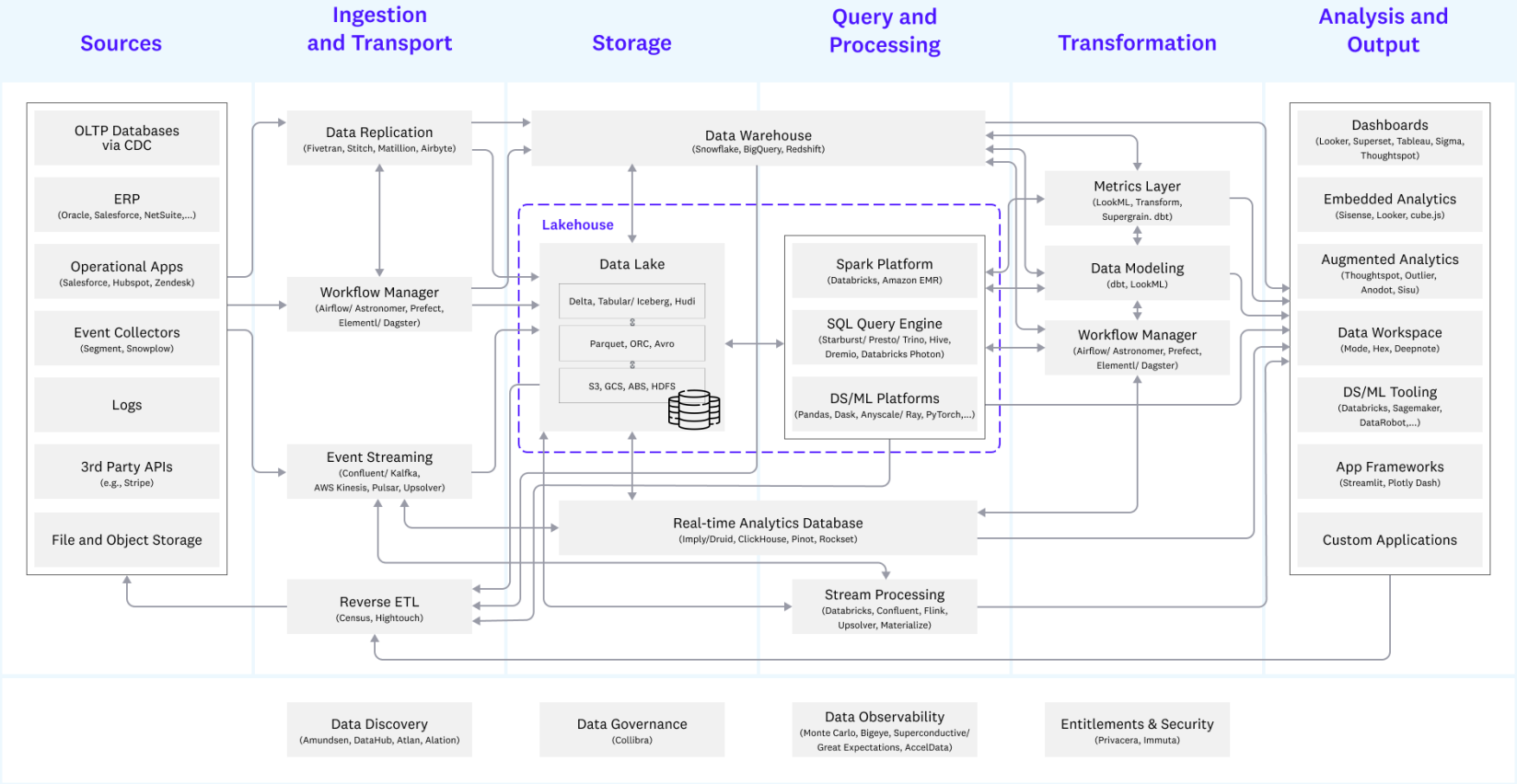
Backups follow...

Donpaul C. Stephens
donpaul@airmettle.com
Founder, AirMettle, Inc.
+1-646-872-2124

Troy Trenchard
troy@airmettle.com
Chief Product Officer
+1-408-823-1486

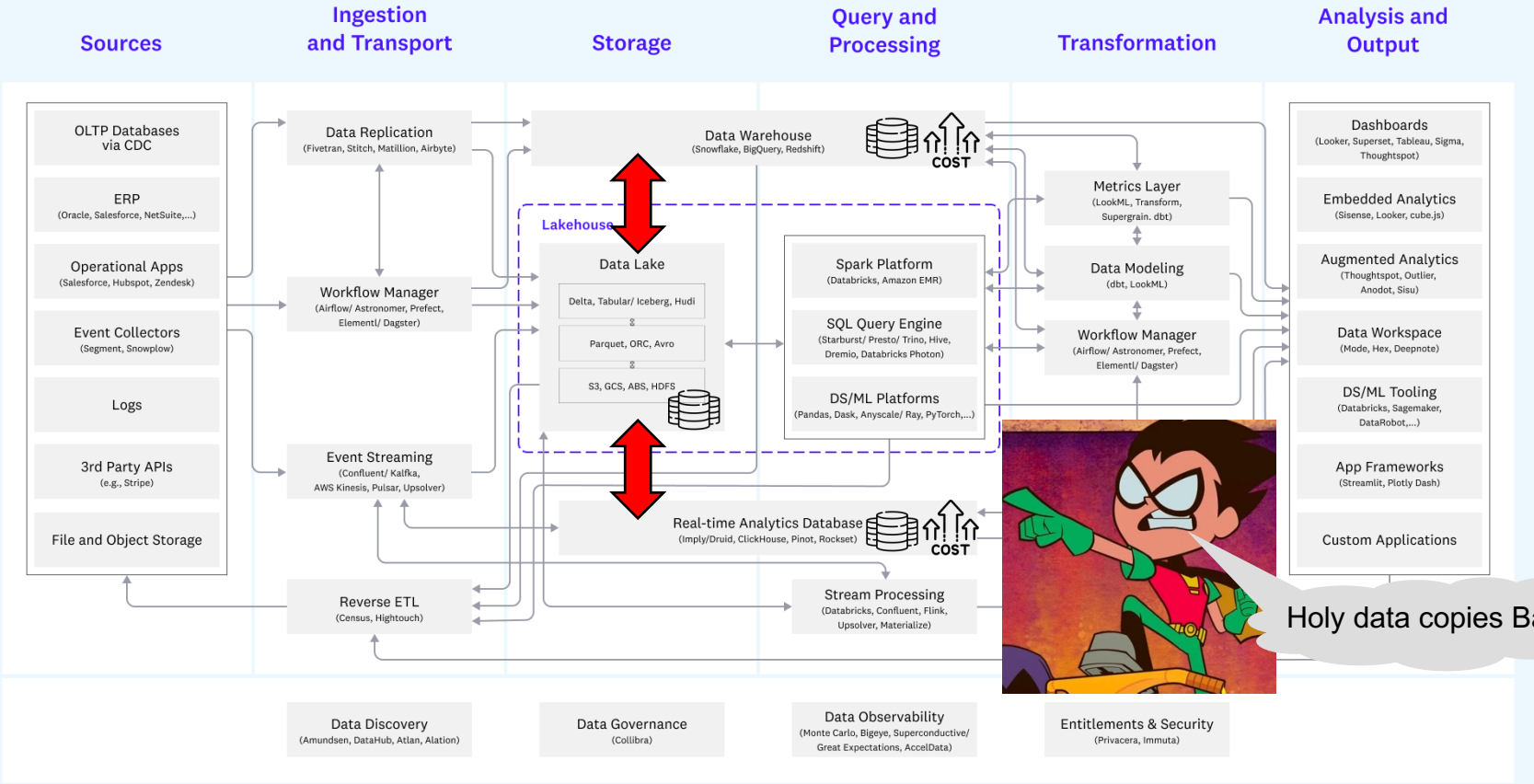
August 2023

Unified Data Infrastructure (2.0)



Source: https://future.a16z.com/emerging-architectures-modern-data-infrastructure/

Unified Data Infrastructure (2.0)

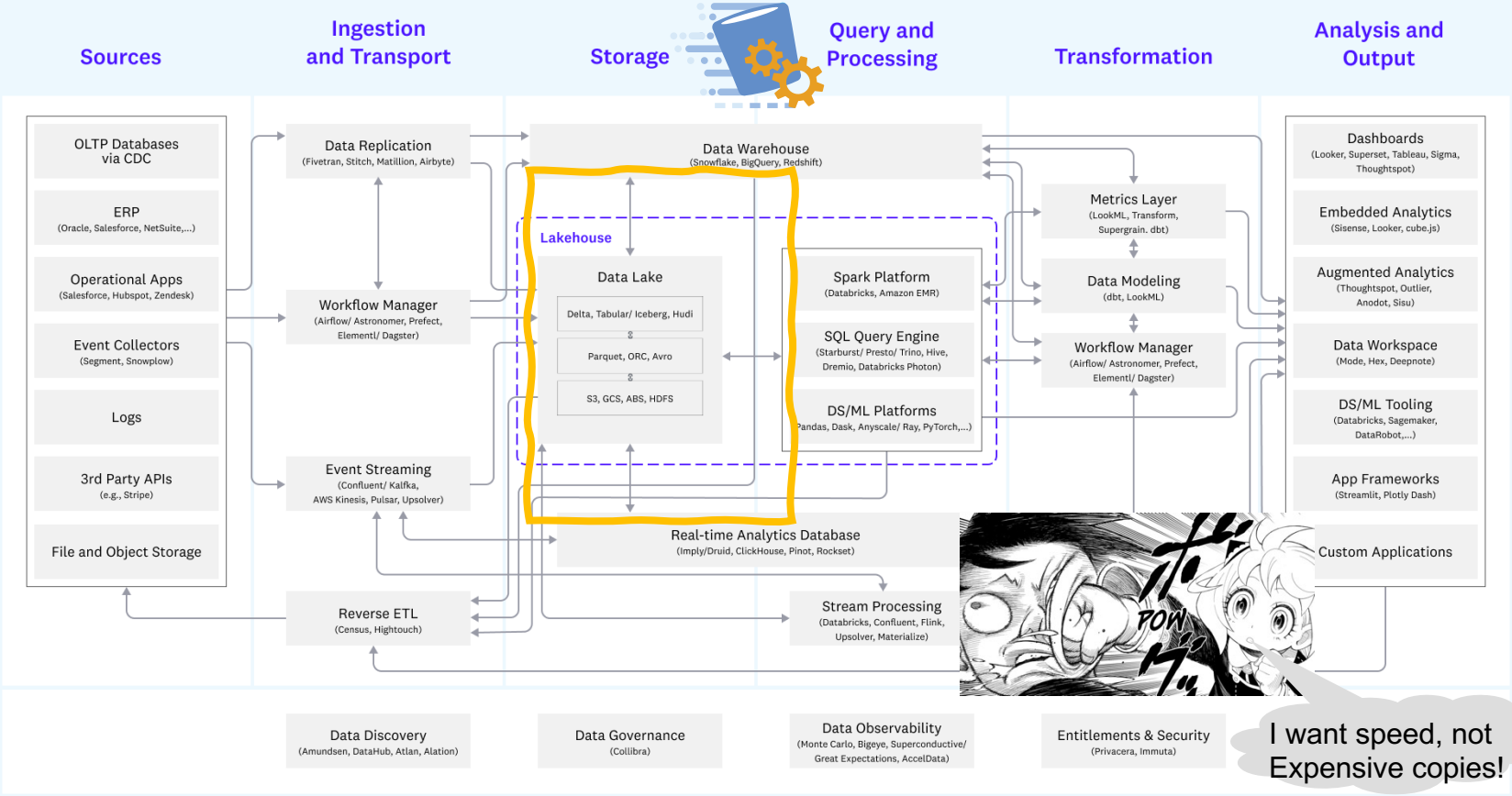


Source: https://future.a16z.com/emerging-architectures-modern-data-infrastructure/

Contemporary

Unified Data Infrastructure (3.0)

- Eliminate Data Warehouse for simple queries
- Real-time analytics from semi-structured data



Source: https://future.a16z.com/emerging-architectures-modern-data-infrastructure/