STORAGE DEVELOPER CONFERENCE

SDC 23

BY Developers FOR Developers

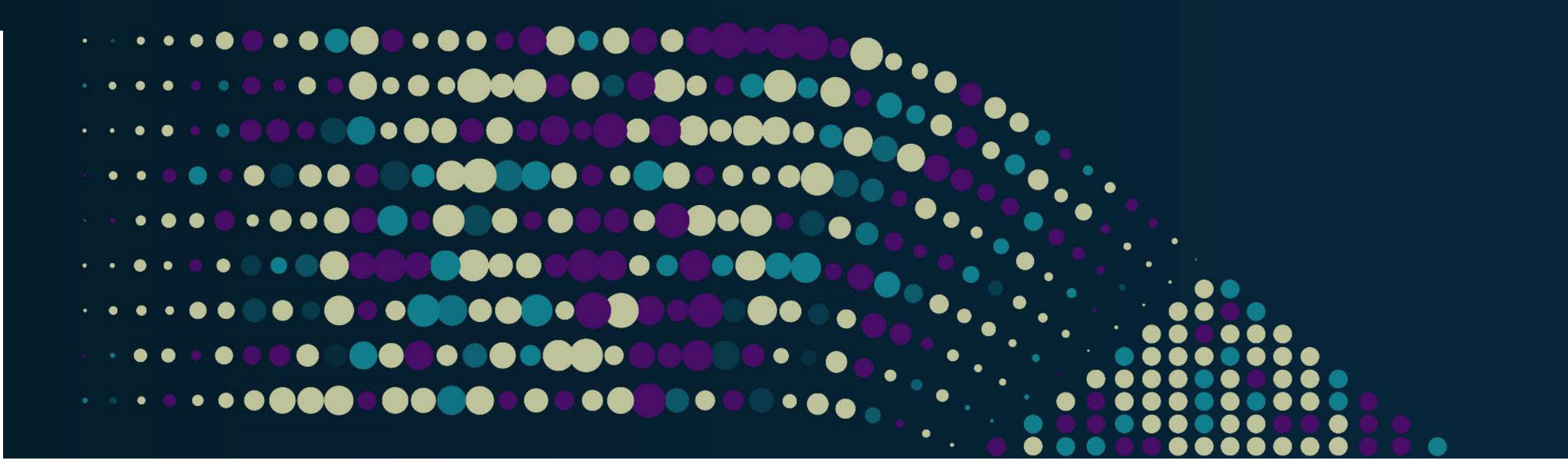
Computational Storage Service

A Real-Time Smart Data Lake

Presented by: Donpaul Stephens

Agenda

- What is Big Data?
- Computational Storage: challenges
- Computational Storage Service
- Reference Design

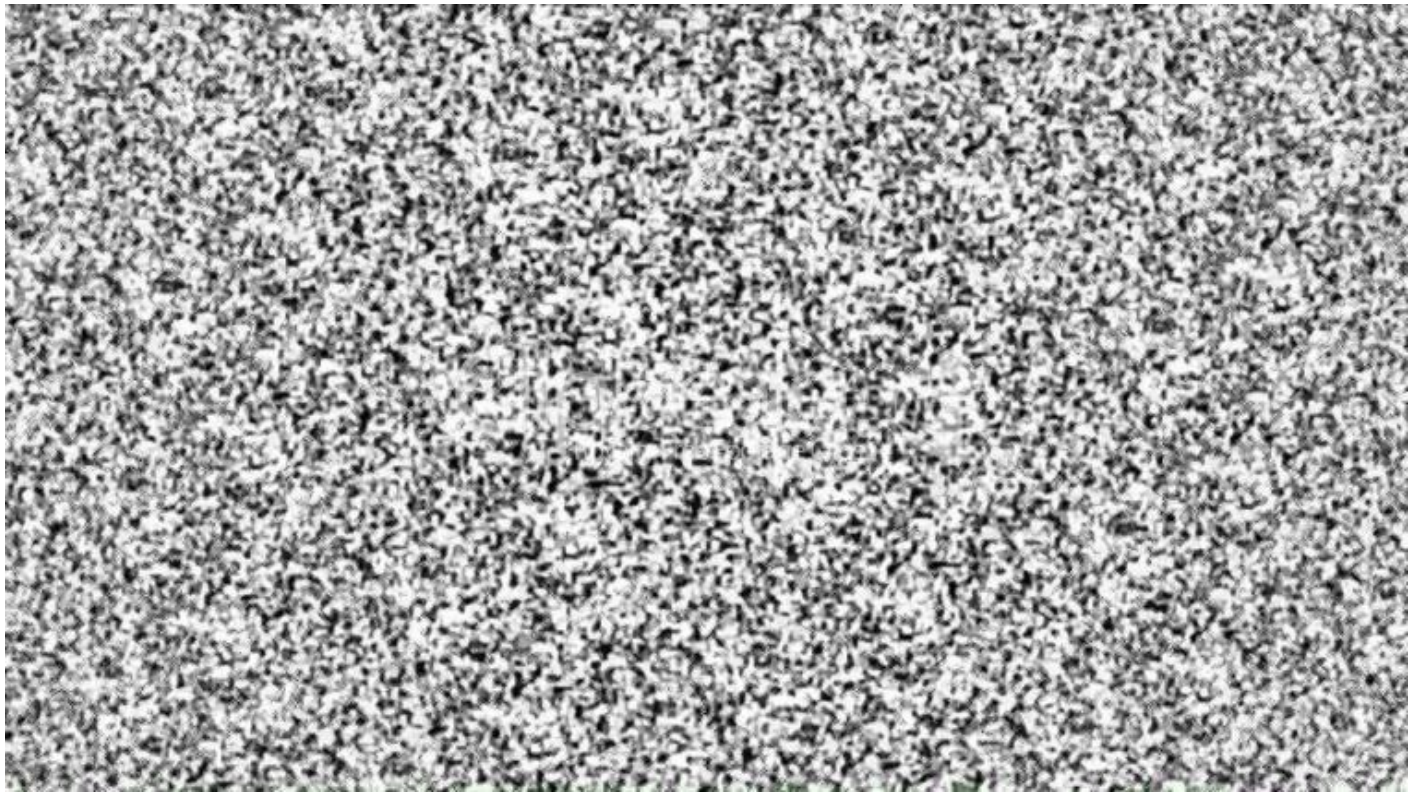


What is Big Data?

Digital Packratism?

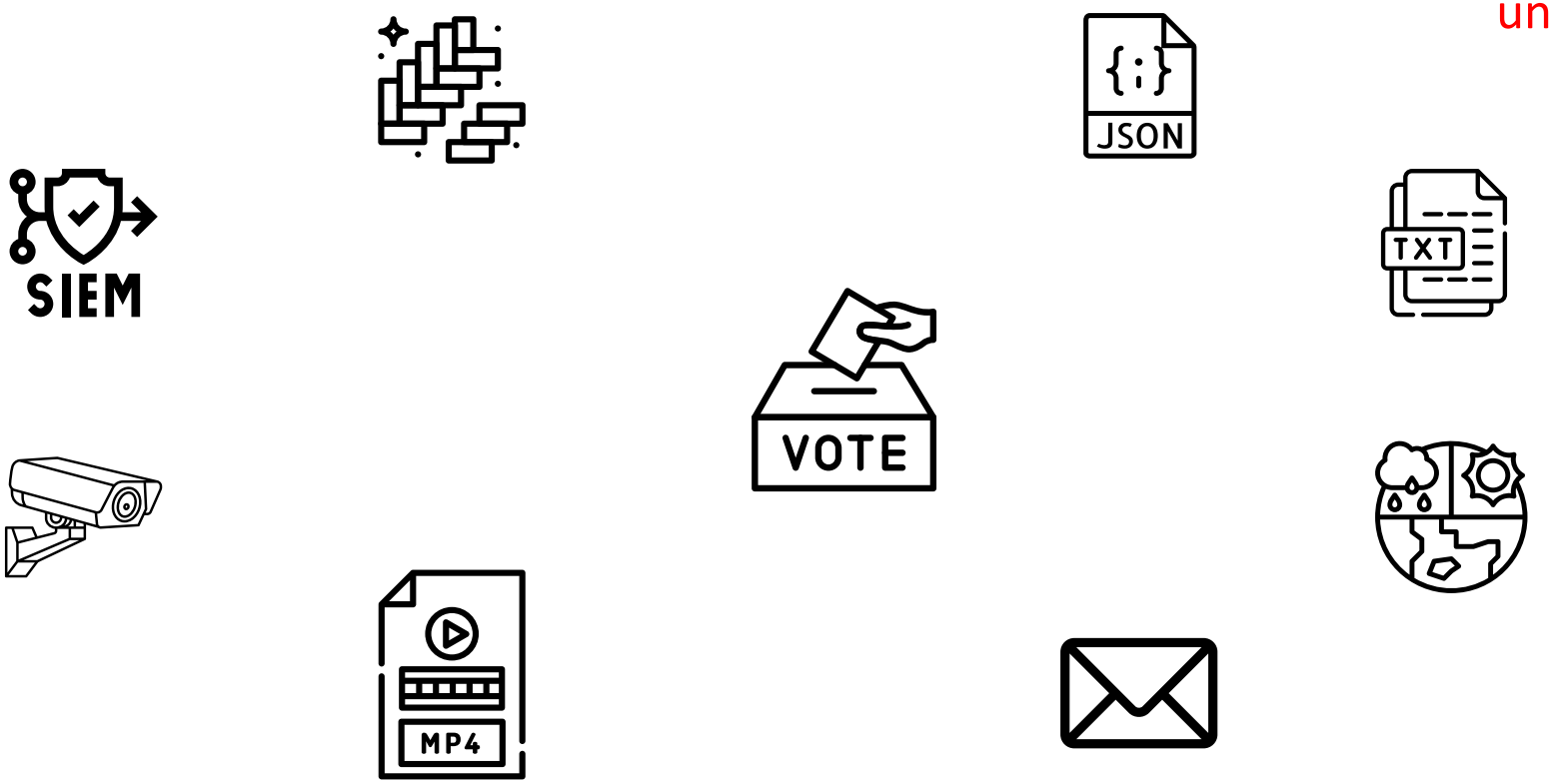
What is Big Data?

Unstructured?



Most data is Semi-Structured

Encrypted data is closest to uncompressible white noise



Stored in a formatted ‘file’

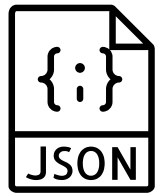
Object! Because historical records can be appended,

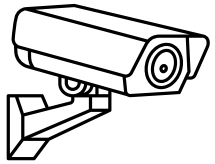
But you can’t rewrite the past, corrections must be trackable

How BIG is the data?

They don’t call it ‘Big Data’ for nuthin!

0.4 to 1GB+ per file:





Video: 1.5 GB to 4GB per hour

SDC 23

Extracting insights from Tabular Data (via SQL)

Security Information & Event Management

Collects sample measurements with certain flags and arguments and groups them by minute.
Returns the number of samples, average duration and standard deviation of duration for each group.

```
select to_string(event_ts, 'yyyy-mm-dd hh24:mi') as interval, count(*), avg(cast(event_dur as int)), stddev_samp(cast(event_dur as int))
from events
 where flgs like 'C__'
 and regexp_contains(args, 'JY.')
 and event_ts between to_timestamp('2000-01-01 00') and to_timestamp('2000-01-01 01')
 group by interval
```


Extracting insights from Tabular Data (via SQL)

Security Information & Event Management

Collects sample measurements with certain flags and arguments and groups them by minute.
Returns the number of samples, average duration and standard deviation of duration for each group.

```
select to_string(event_ts, 'yyyy-mm-dd hh24:mi') as interval, count(*), avg(cast(event_dur as int)), stddev_samp(cast(event_dur as int))
from events
 where flgs like 'C__'
 and regexp_contains(args, 'JY.')
 and event_ts between to_timestamp('2000-01-01 00') and to_timestamp('2000-01-01 01')
 group by interval
```

Star Schema Benchmark

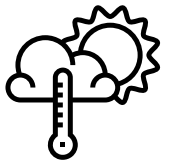
Comparison of revenue for some product classes, for suppliers in a certain region, grouped by product brand and year.

```
select sum(lo_revenue), d_year, p_brand1
from lineorder, date, part, supplier
 where lo_orderdate = d_datekey
 and lo_partkey = p_partkey
 and lo_suppkey = s_suppkey
 and p_category = 'MFGR#12'
 and s_region = 'AMERICA'
 group by d_year, p_brand1
 order by d_year, p_brand1
```

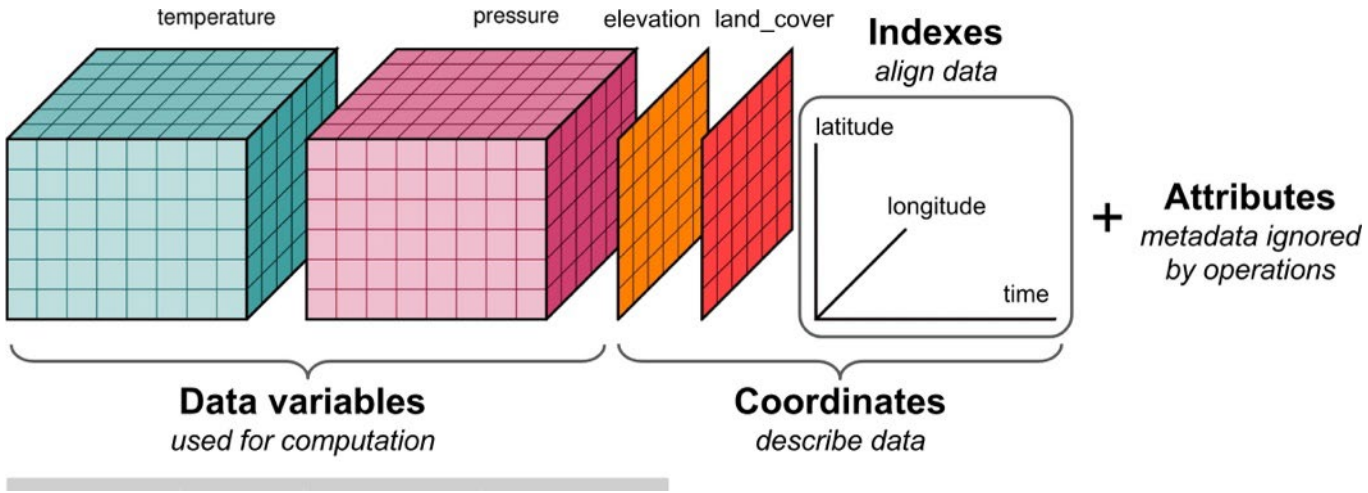
Award from: NSF

SDC 23

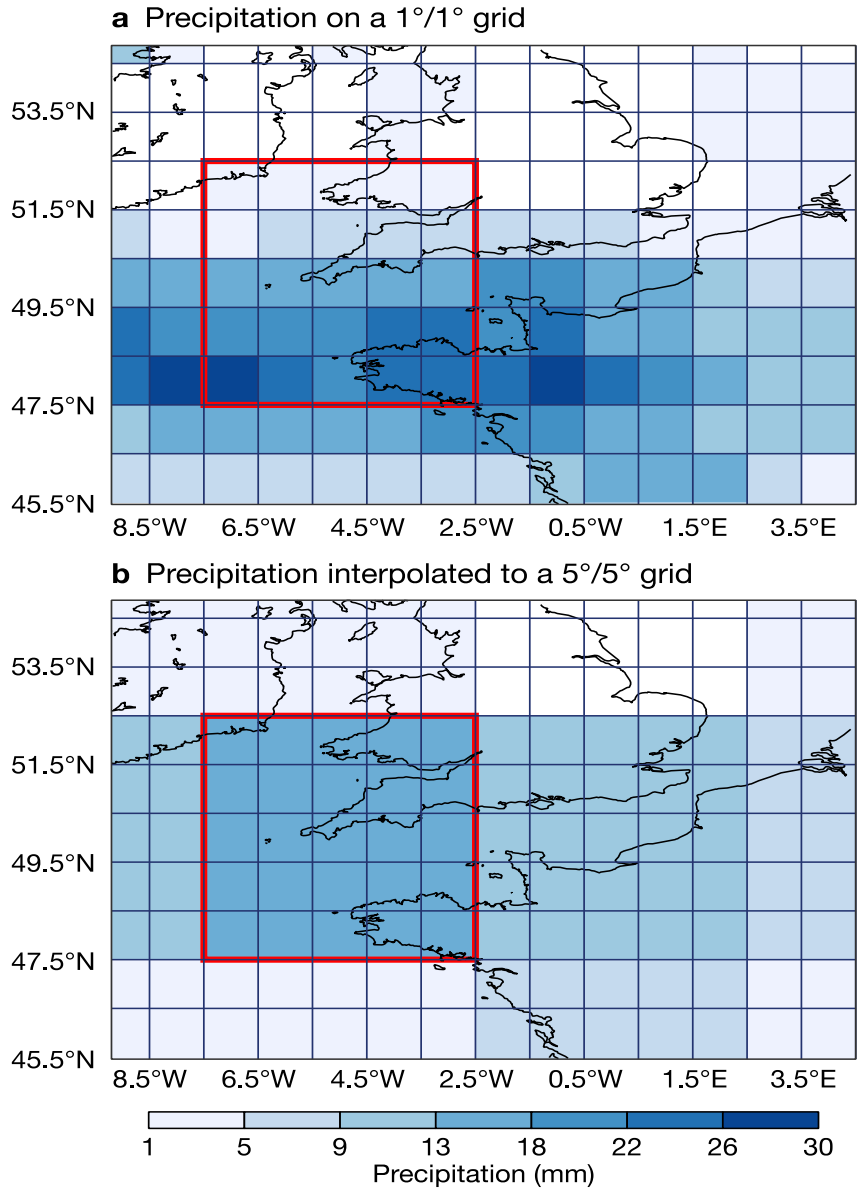
Re-scaling weather data?



Scientific
(ex: Climate)



Awards from:



Could we search video?

Private Sector

Fantasy Adventure show

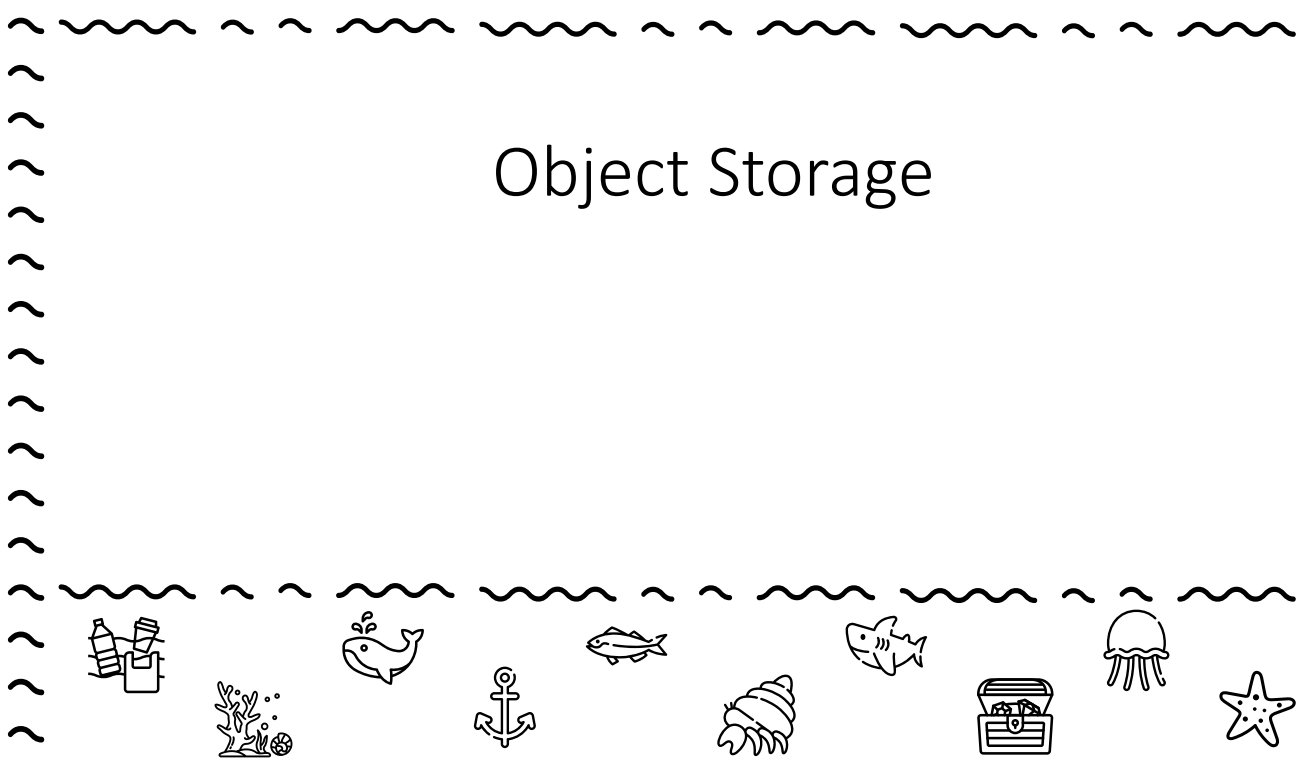
Make sure current-day items
(e.g. coffee cup)
Do not appear on screen

Public Sector

Find missing people / validate alibi

- Child not found at amusement park
- Parent/guardian has pictures…

Traditional Data Lake



Objects are internally partitioned
For storage in parallel

Data Lake

Traditional Data Lake

Comes from Everywhere

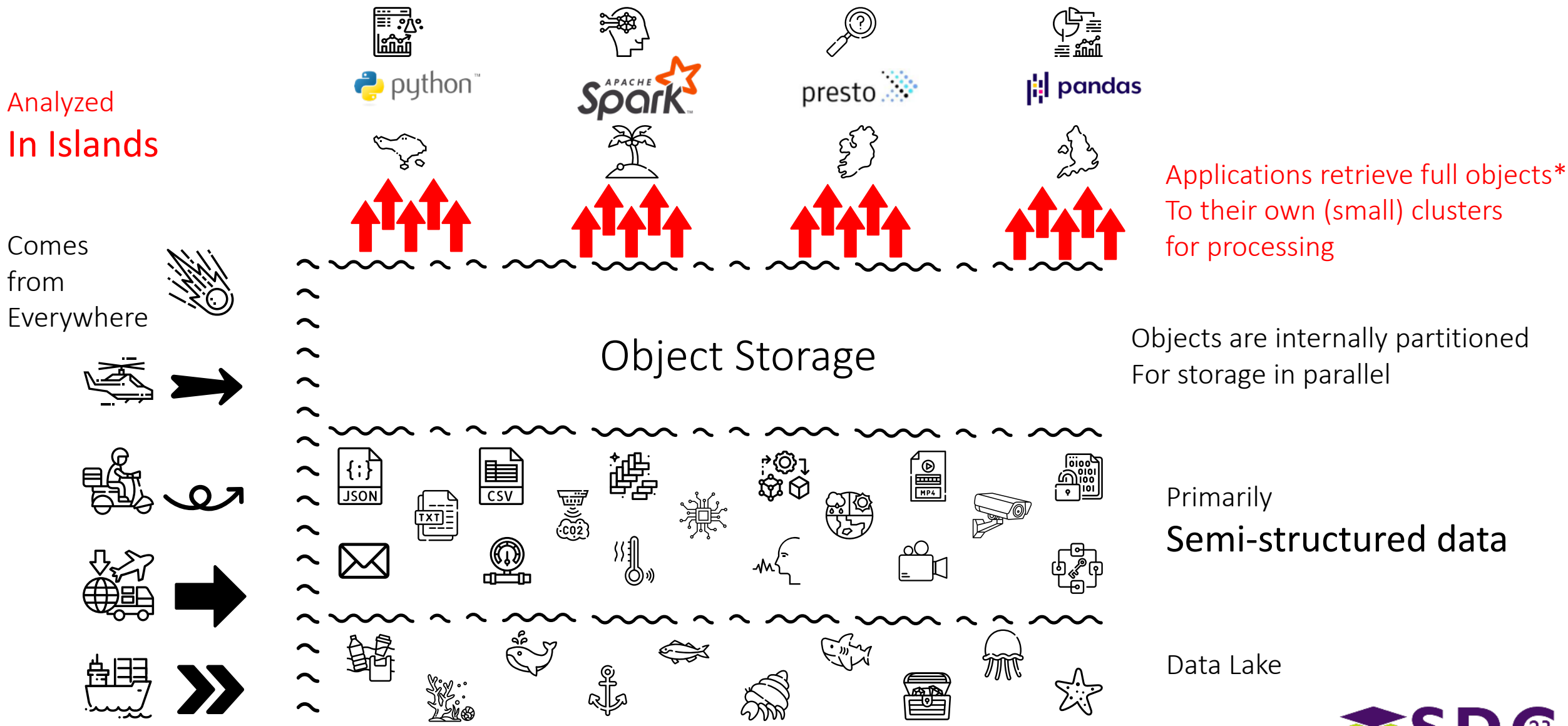
Objects are internally partitioned
For storage in parallel

Primarily

Semi-structured data

Data Lake

Traditional Data Lake



This is AWESOME for selling networking gear!!!

Time to Move the data!

https://www.youtube.com/watch?v=IYRurPB4WA0&list=PPSV

29sec

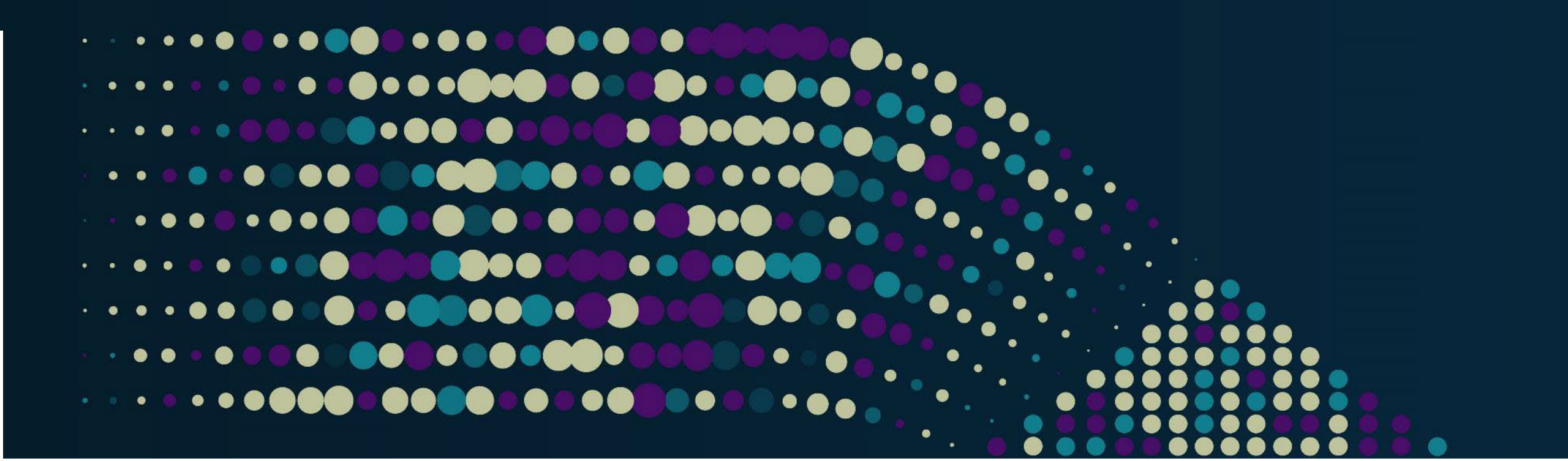
This is AWESOME for selling networking gear!!!

Time to make the donuts!

I ne donuts!

Time to Move the data!

I Moved the data!



Why not process where the data is?

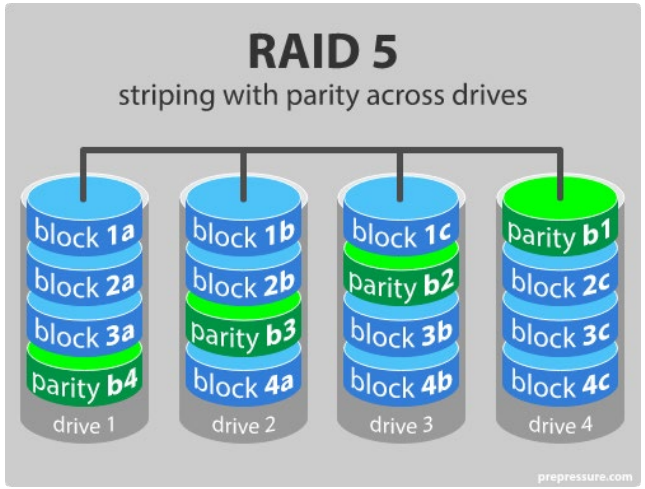
Computational Storage

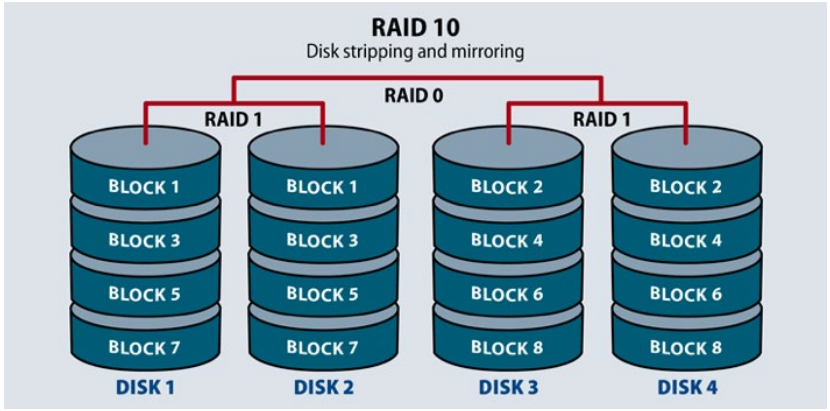
Active Disks… circa ‘98… wha’ happened?!?!?

SDC 23

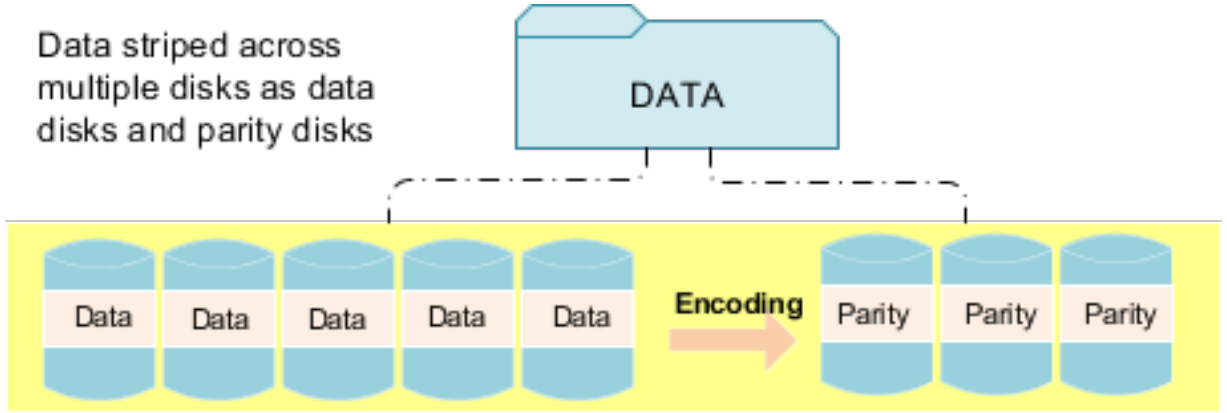
Resiliency 101: How do storage solutions protect data?

RAID:





Erasure Coding:

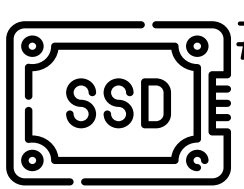


Data protection algorithms designed for HDD

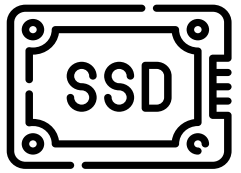
What that means for data reliably placed in storage:
First 4 devices shown…

Simple Table:

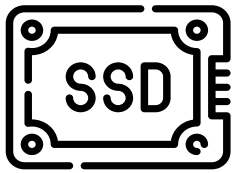
1	155190	7706	1	17	21168.23	0.04	0.02	N	O	3/13/96	2/12/96	3/22/96	DELIVER_IN_PERSON	TRUCK	egular_courts_above_the
1	67310	7311	2	36	45983.16	0.09	0.06	N	O	4/12/96	2/28/96	4/20/96	TAKE_BACK_RETURN	MAIL	ly_final_dependencies:_slyly_bold_
1	63700	3701	3	8	13309.6	0.1	0.02	N	O	1/29/96	3/5/96	1/31/96	TAKE_BACK_RETURN	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



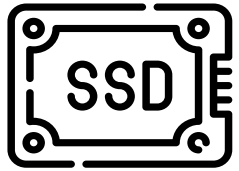
#1



#2



#3



#4

Bytes of data divided evenly across SSDs!

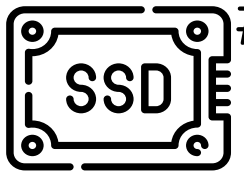
Data protection and streaming performance!

Supports data protection algorithms designed for HDD!

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

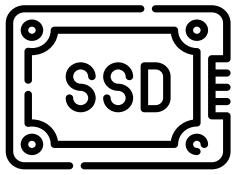
Simple Table:

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



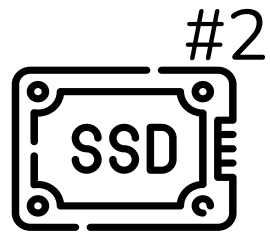
#1

```
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_
```



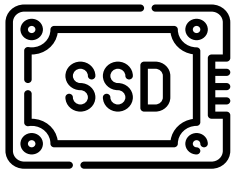
#3

```
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.
```



#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,
```

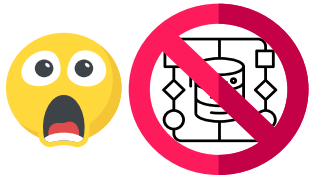


#4

```
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._depos
```

Bytes of data divided evenly across SSDs!

Data protection and streaming performance!

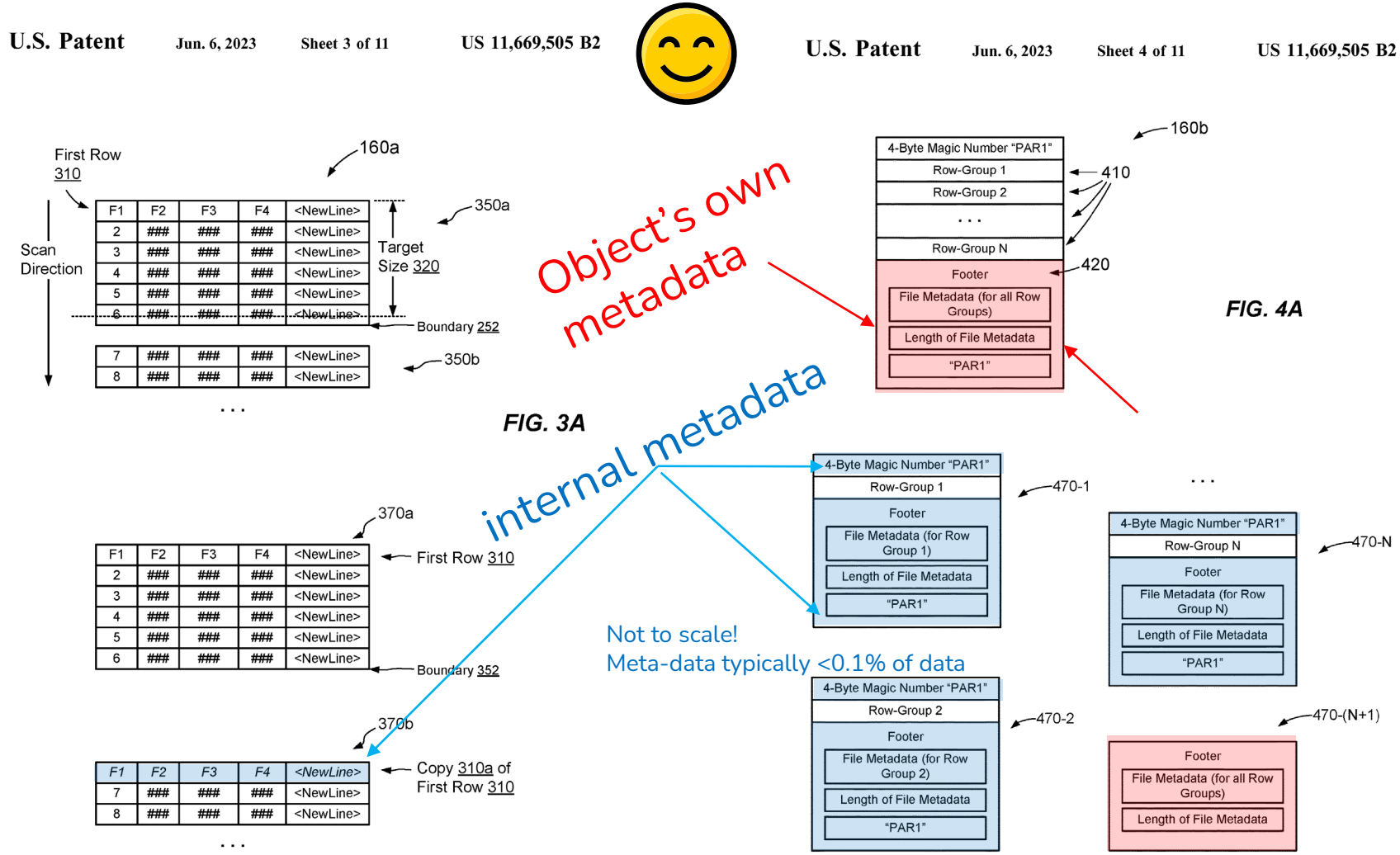


HDD-centric RAID/Erasure Coding prevent in-storage analytics

AirMettle: Data partitioning for processing AND protecting data

AirMettle

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



AirMettle Patented IP

AirMettle internal metadata enables parallel 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

AirMettle Patented IP

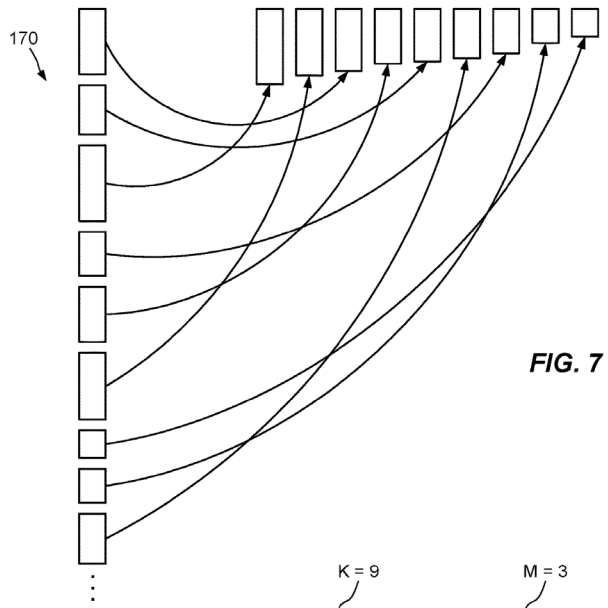
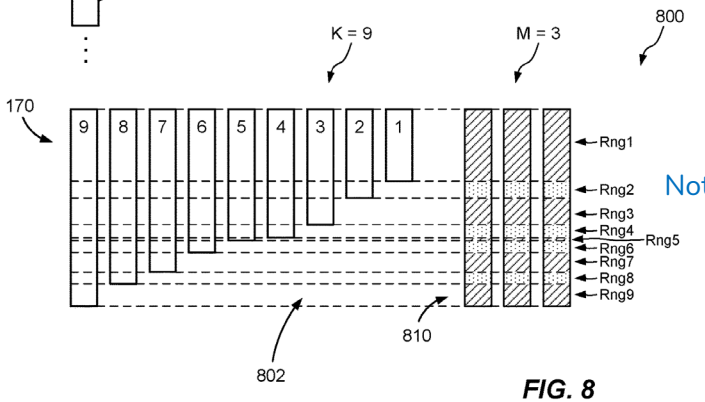


FIG. 7



Not to scale!

FIG. 8

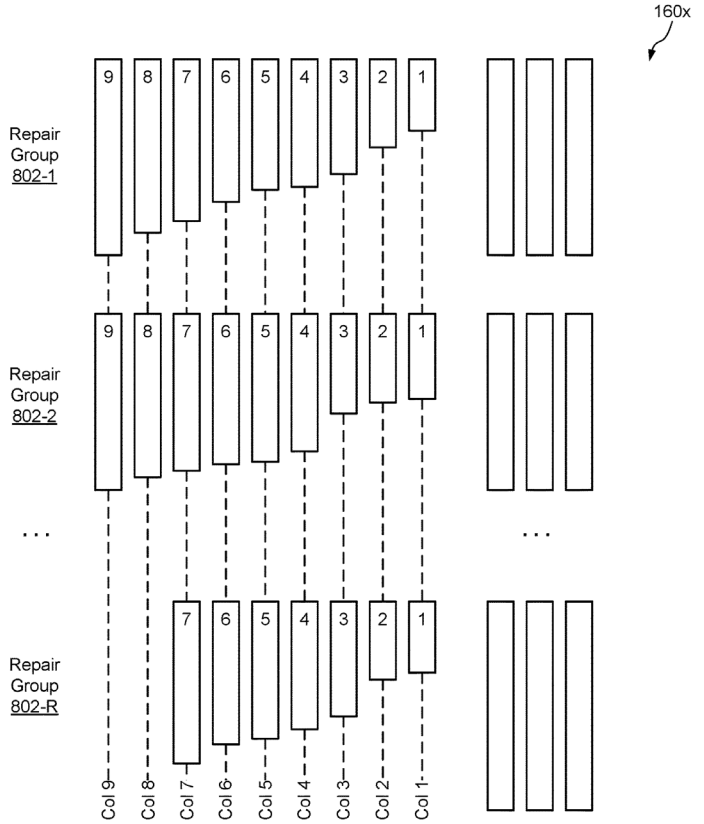
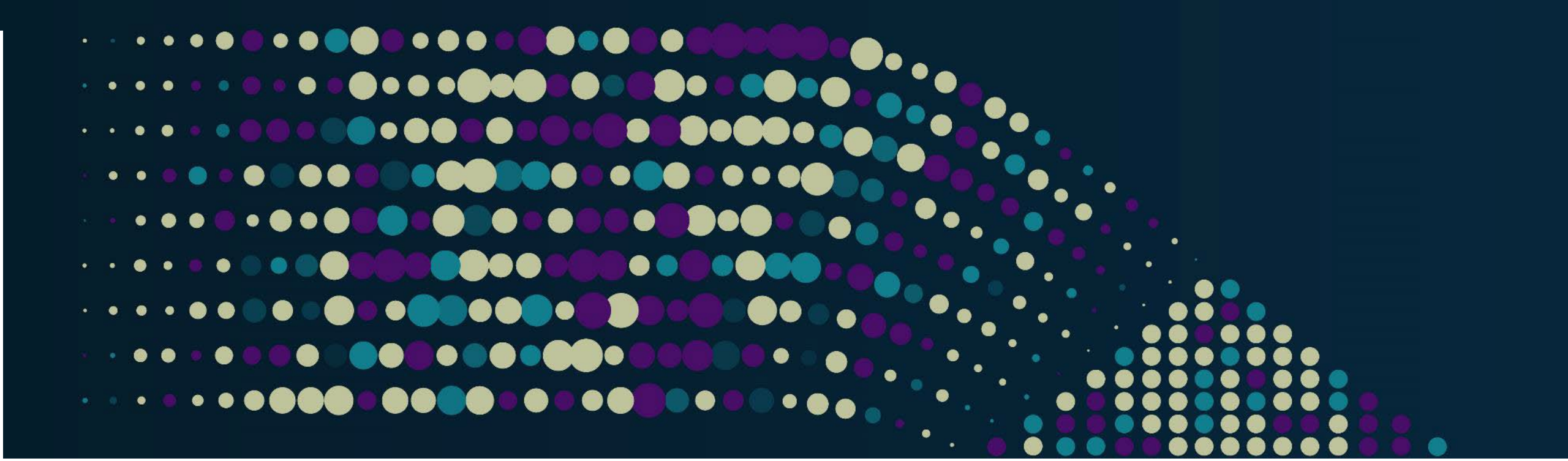


FIG. 9

Erasure coding for our non-uniform data segments

SDC 23

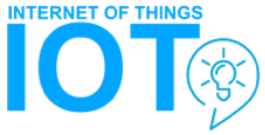


In practice

Initial Results

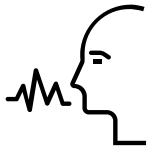
Accelerated analytics of classic tabular data

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 UNIVERSITY OF WISCONSIN-MADISON & MIT

Star Schema Benchmark

Utilized 223 Select queries to Object Storage:

100 X faster

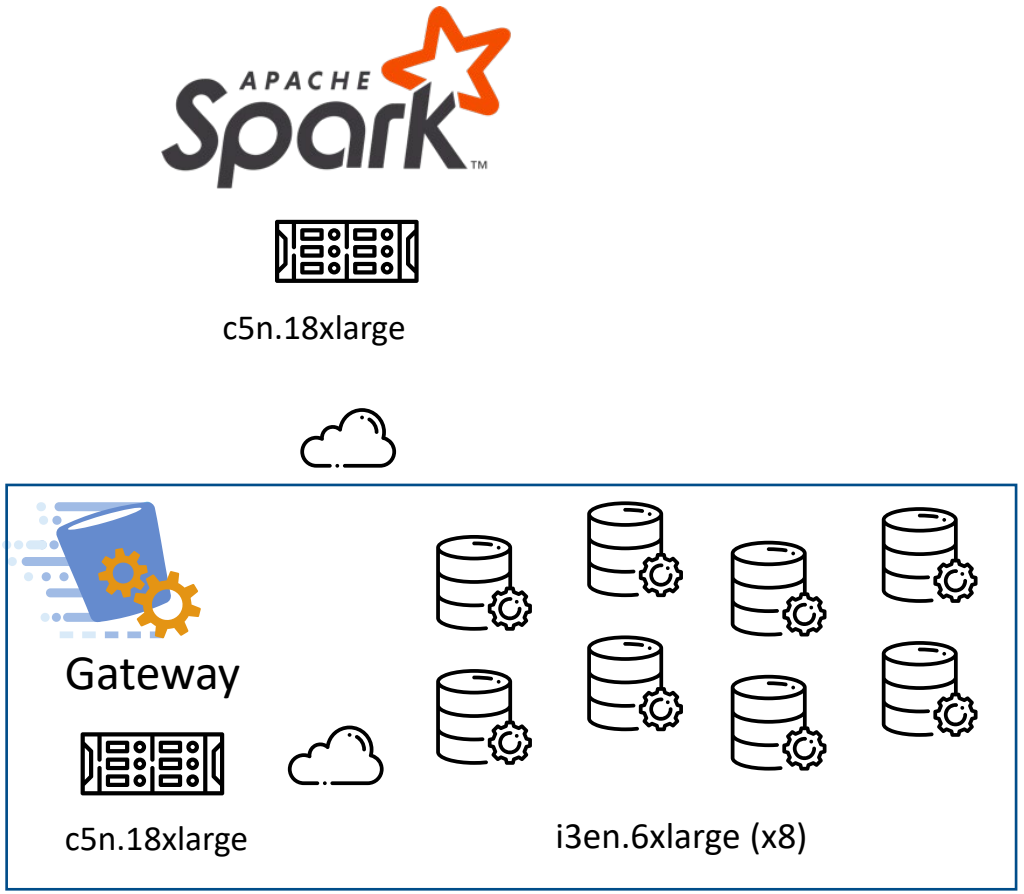
Under a minute vs. 1 hour 45min

Unprecedented speed of analysis: Directly from storage

No data warehouse required

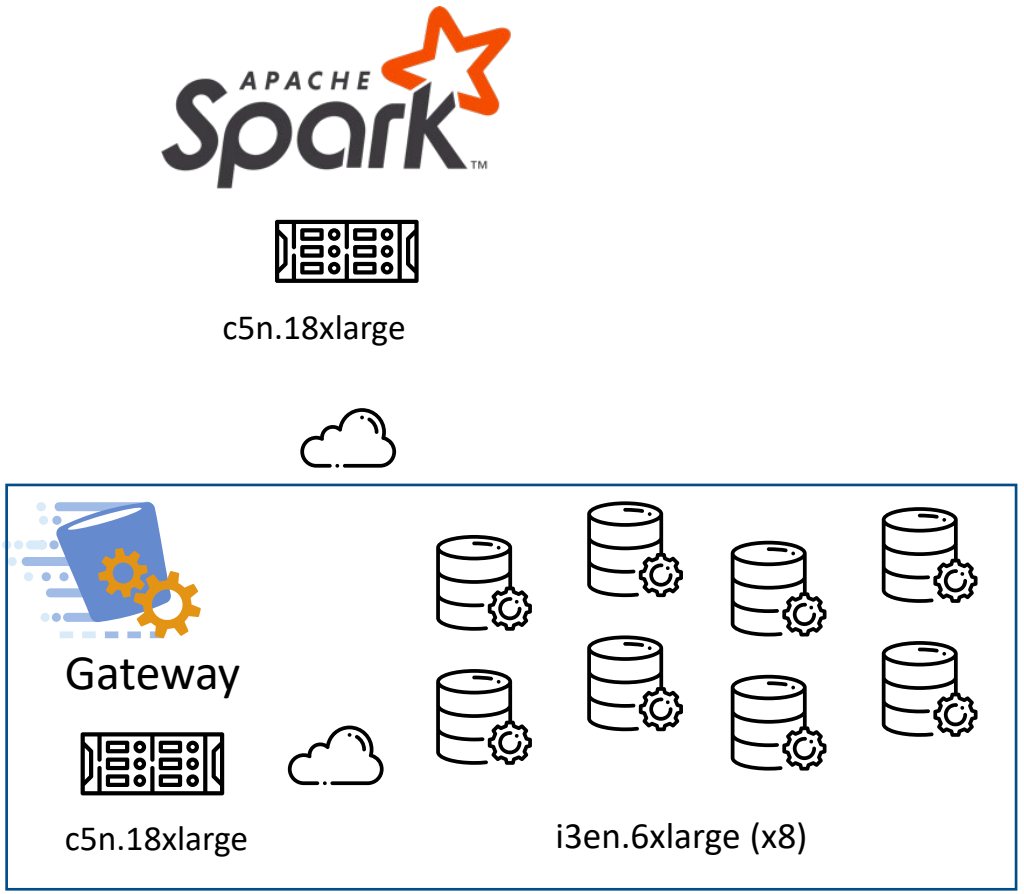
SDC 23

AirMettle Accelerates



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

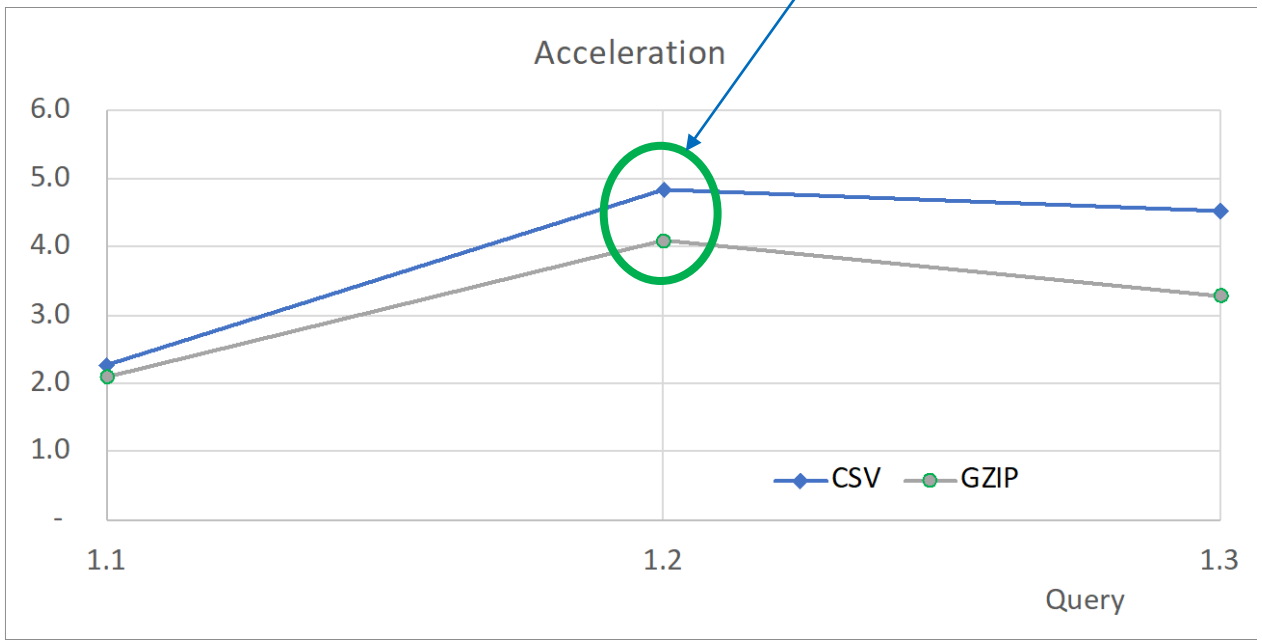
AirMettle Accelerates



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

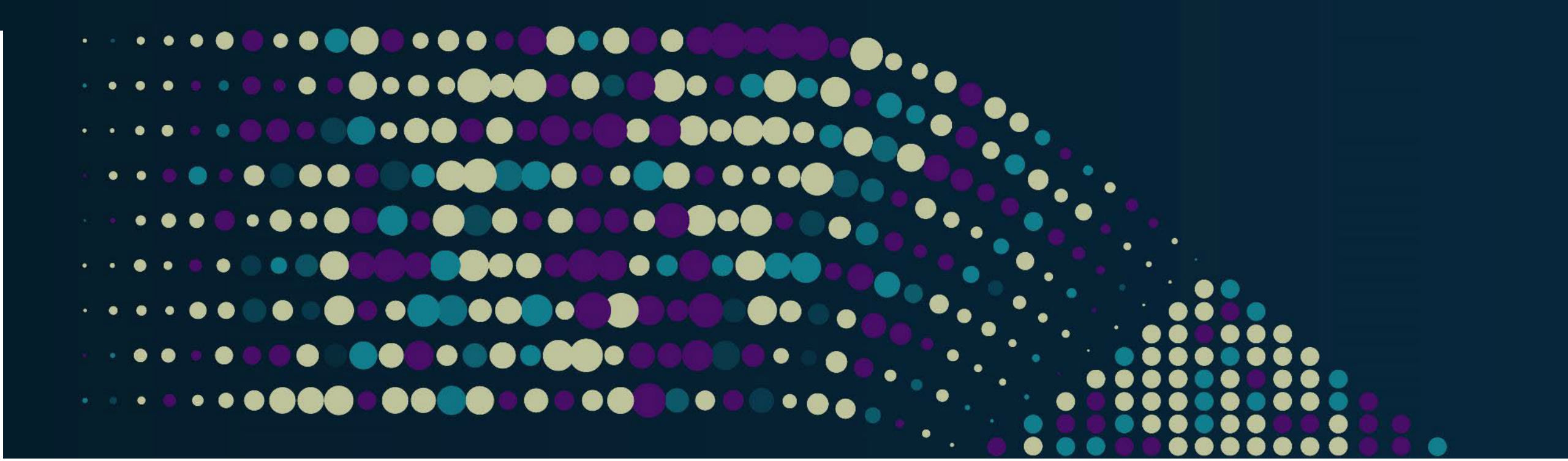
For more complex queries, acceleration depends on how much time was spent in portions we offload:

- Q1.1: 50% of time… 100x faster: 2x overall
- Q1.2: 80% of time… 100x faster: 5x overall



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

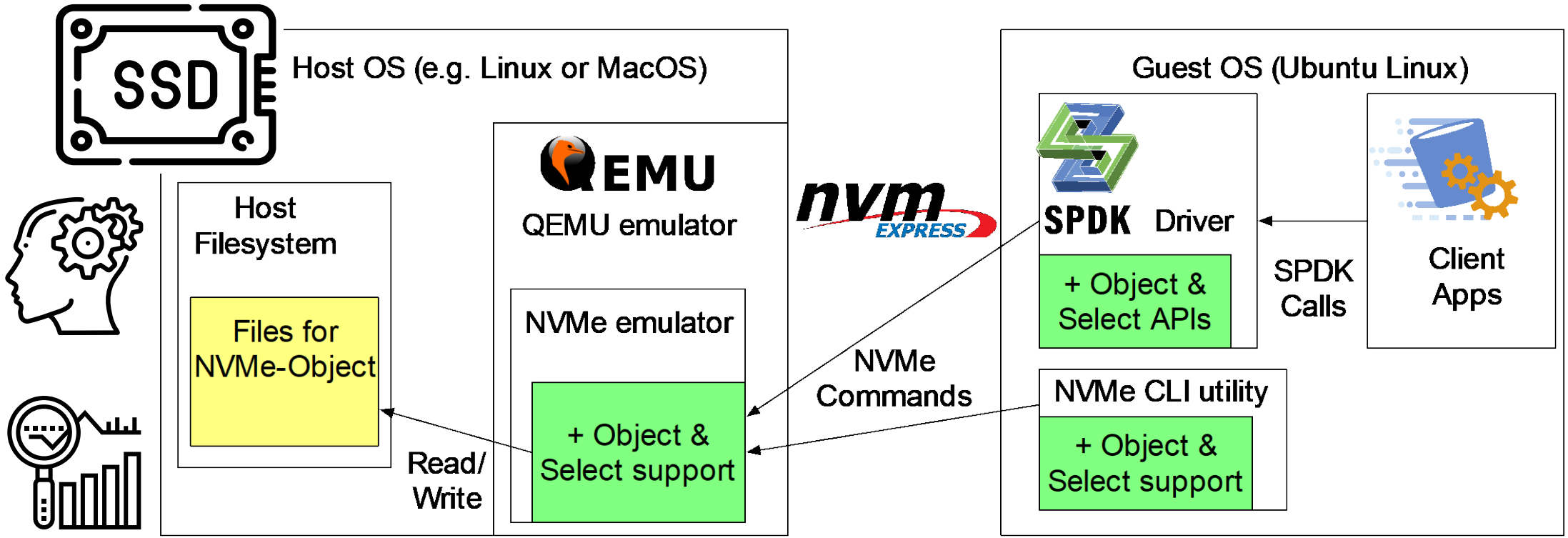
SDC 23



Computational Storage Devices

For *GASP* computation

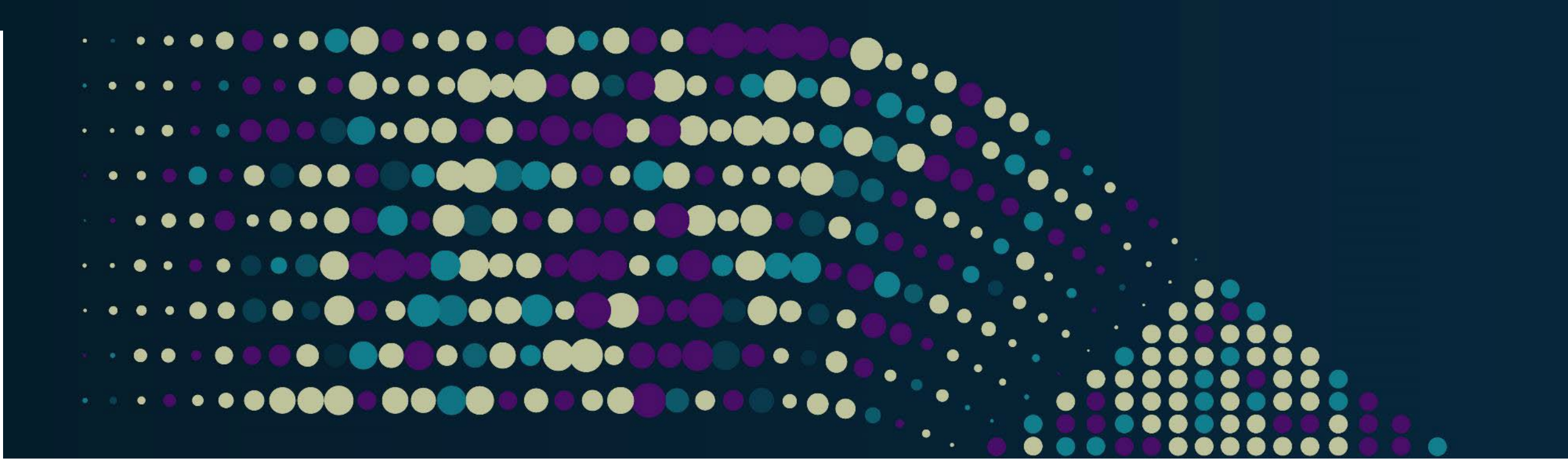
NVMe-Object SSD: yes, we can do it!



https://github.com/AirMettle/csd

https://github.com/AirMettle/qemu-csd

https://github.com/AirMettle/spdk-csd



Please take a moment to rate this session.

Your feedback is important to us.

