Big Data Storage Problems and Solutions

Elephant in the Room Problem.

Dr. Abzetdin ADAMOV

Director, Center for Data Analytics Research
School of IT & Engineering
ADA University
aadamov@ada.edu.az

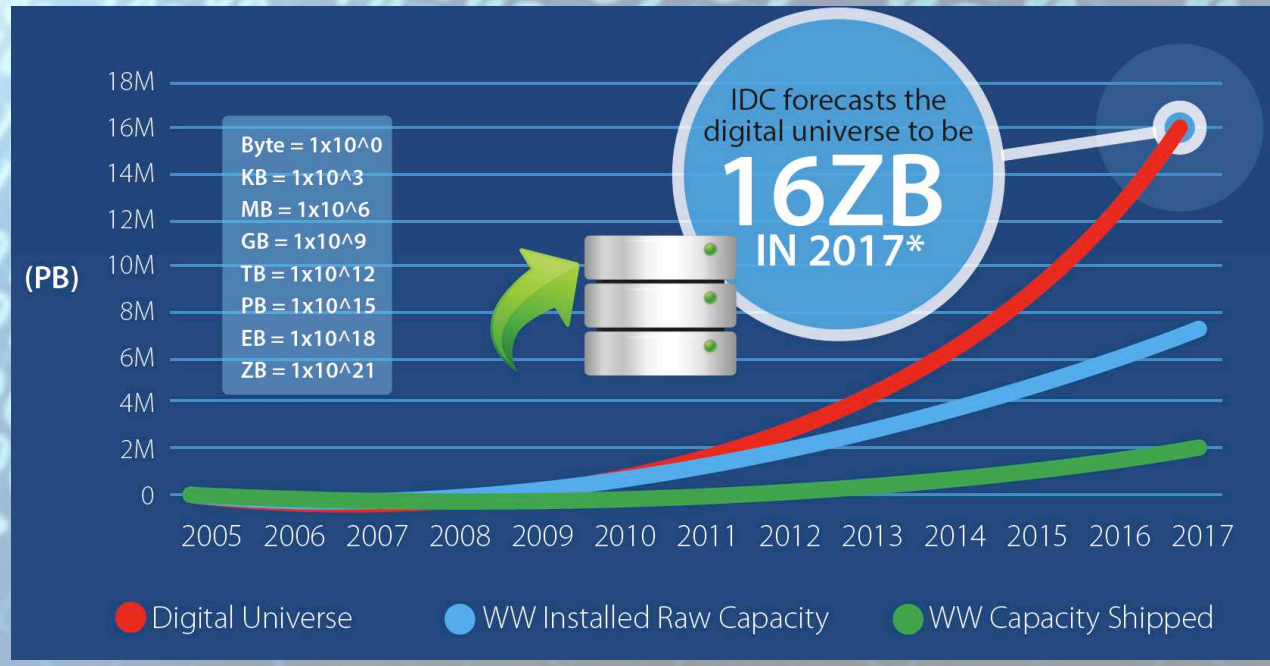
19 May 2018, BDDB2018, ADA University


BIG DATA is a PROBLEM
WHERE is SOLUTION?

Digital Universe Volume

Every day now we create as much information as we did from the dawn of civilization up until 2003

- 2003 – 5 exabytes from beginning of civilization
- 2005 – 130 exabytes
- 2008 – 480.000 petabytes (PB)
- 2009 – 800.000 PB
- 2010 – 1200 000 PB or 1.2 zettabyte (ZB)
- 2011 – 1.8 ZB
- 2012 – 2.7 ZB
- 2014 ~ 6.2 ZB
- 2015 ~ 10 ZB
- 2017 ~ 16 ZB
- 2018 > 20 ZB
- Expected to reach 44 ZB by 2020



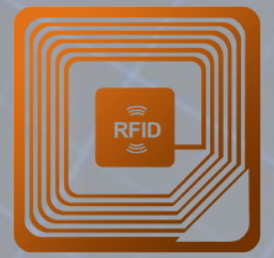
Where Data Comes From

Data is produced by:

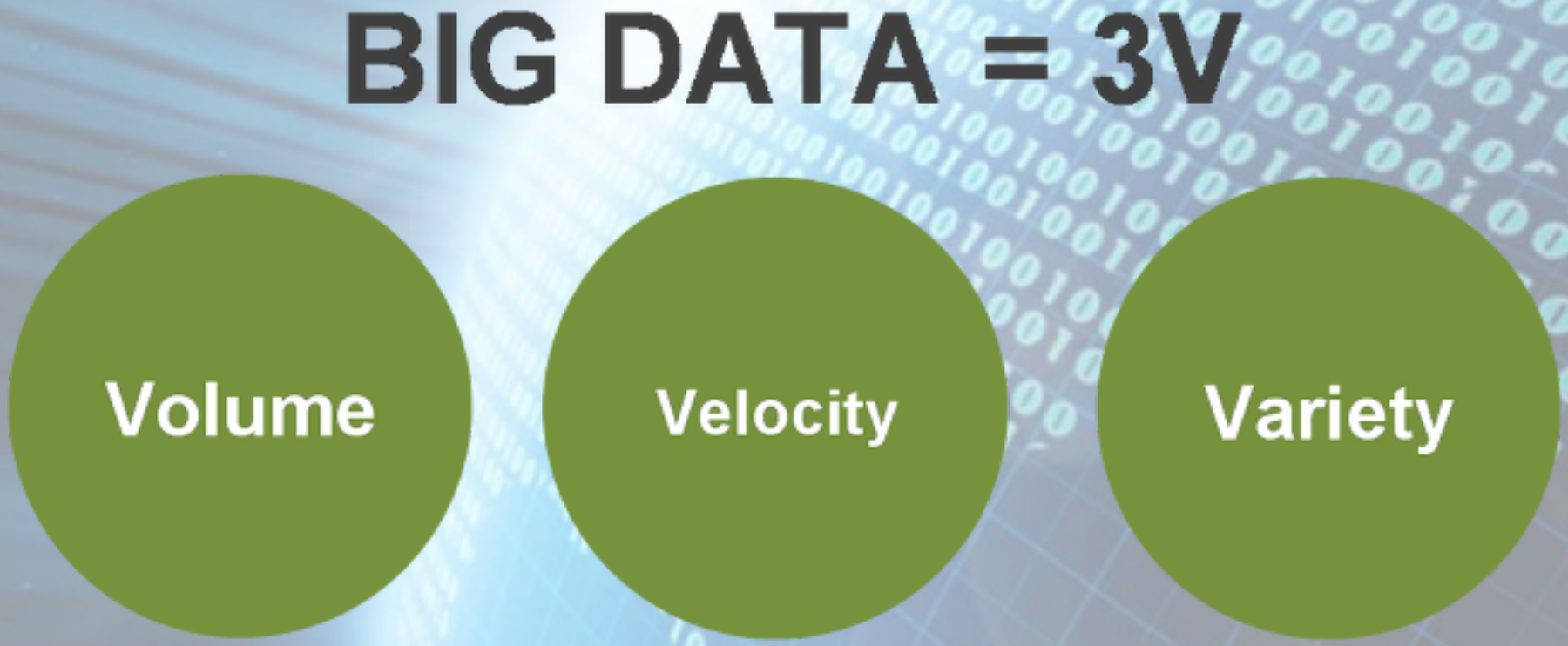
- People
  - Social Media, Public Web, Smartphones, …
- Organizations (Employer)
  - OLTP, OLAP, BI, …
- Machines
  - IoT, Satellites, Vehicles, Science, …

Modern Data Sources

- Internet of Anything (IoAT)
  - Oil Rigs, Wind Turbines, Cars
  - Weather Stations, Smart Grids
  - RFID Tags, Beacons, Wearables
- User Generated Content (Web & Mobile)
  - Twitter, Facebook, Snapchat, YouTube
  - Clickstream, Ads, User Engagement
  - Payments: Paypal, Venmo

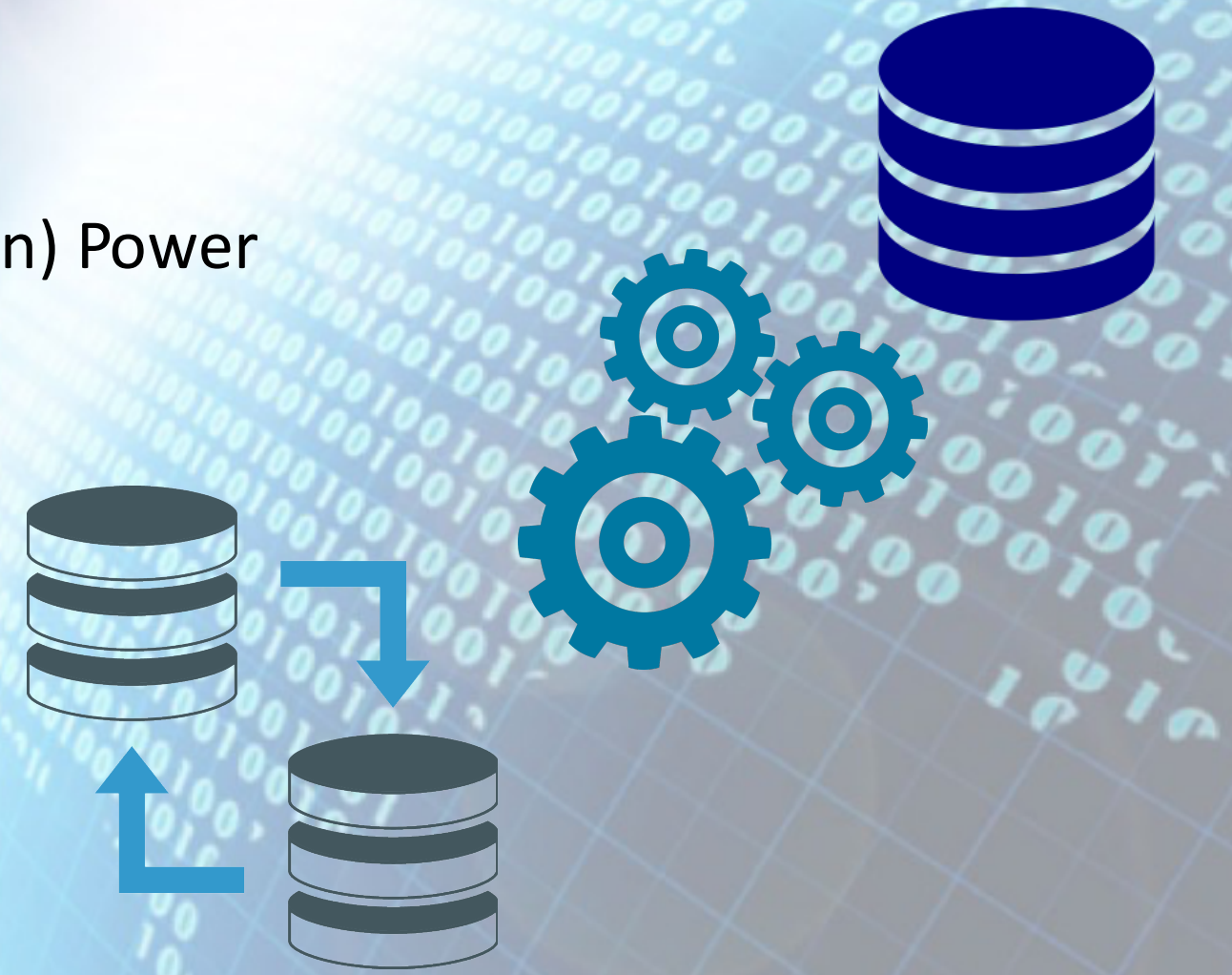


BIG DATA Problem



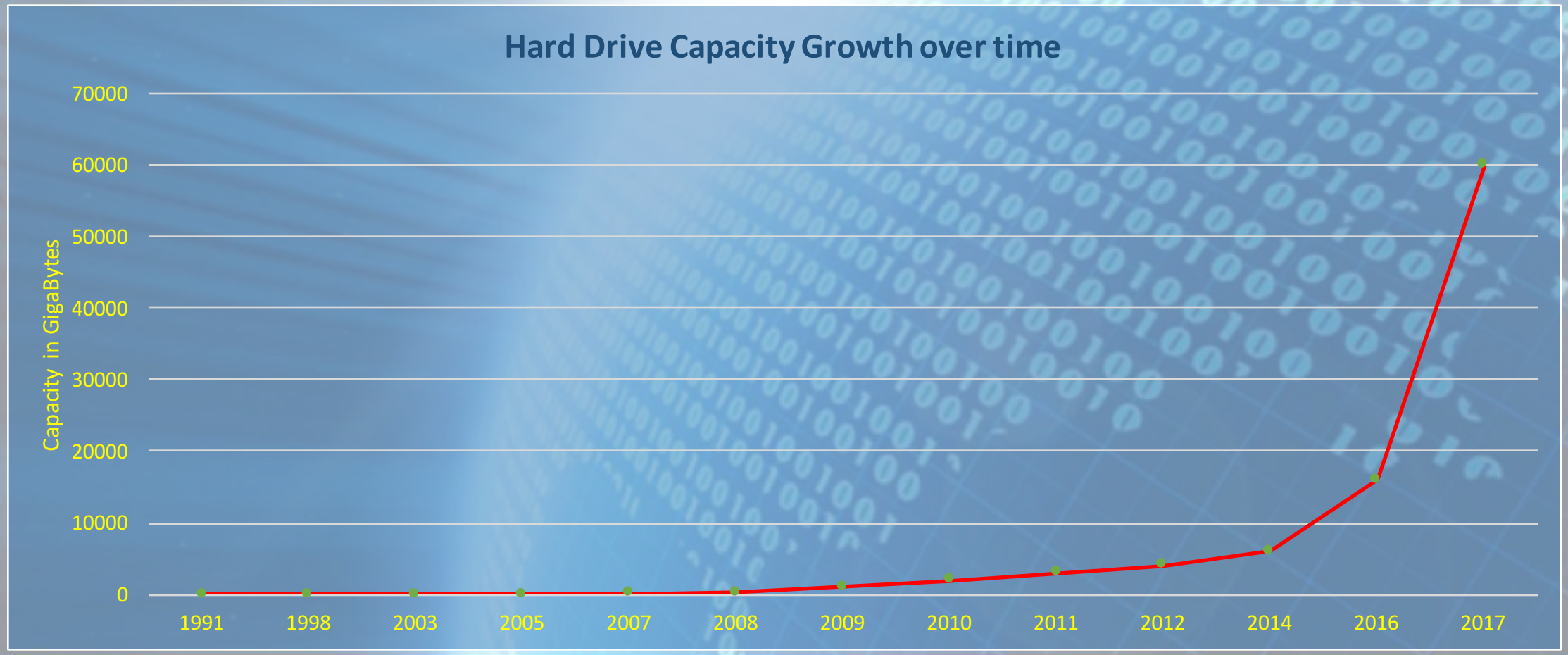
BIG DATA Management Requirements

1. Storage Capacity
2. Processing (Computation) Power
3. Data Transfer Speed

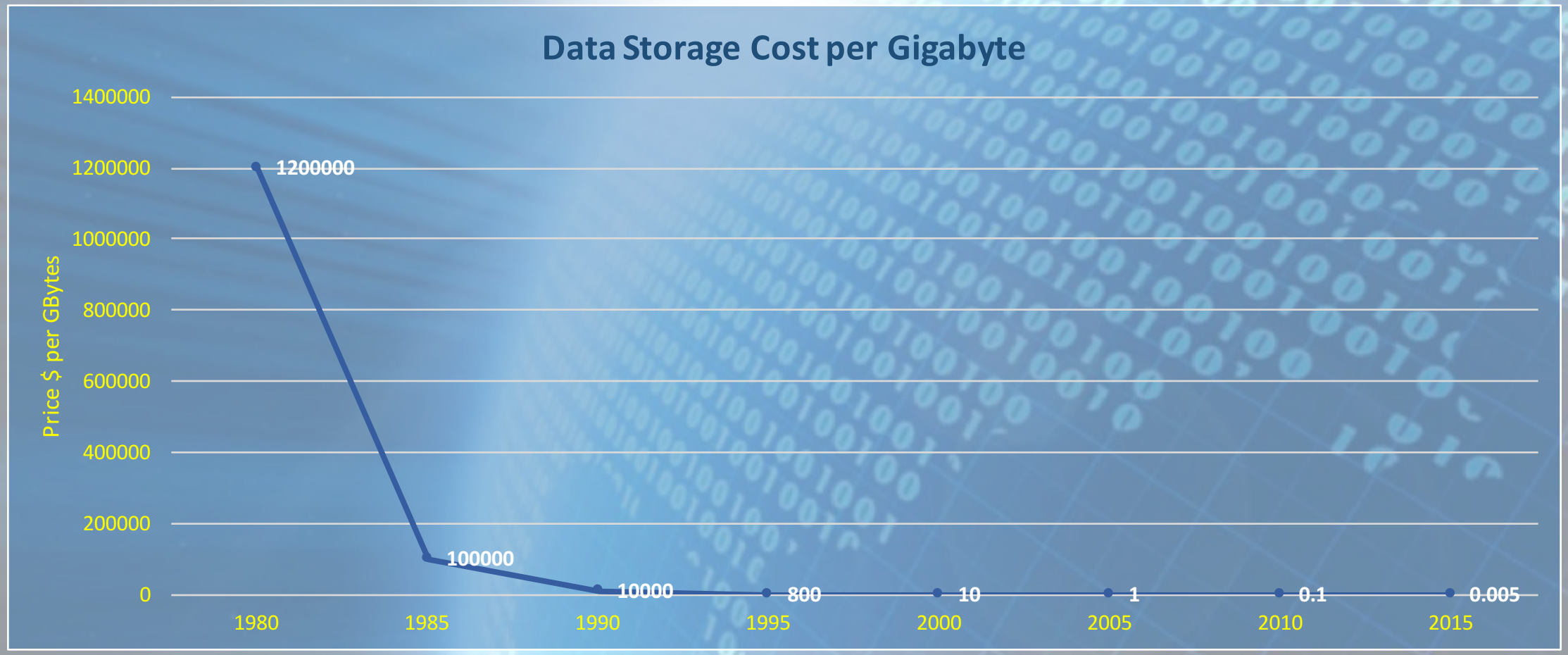



WHY DATA TRANSFER?

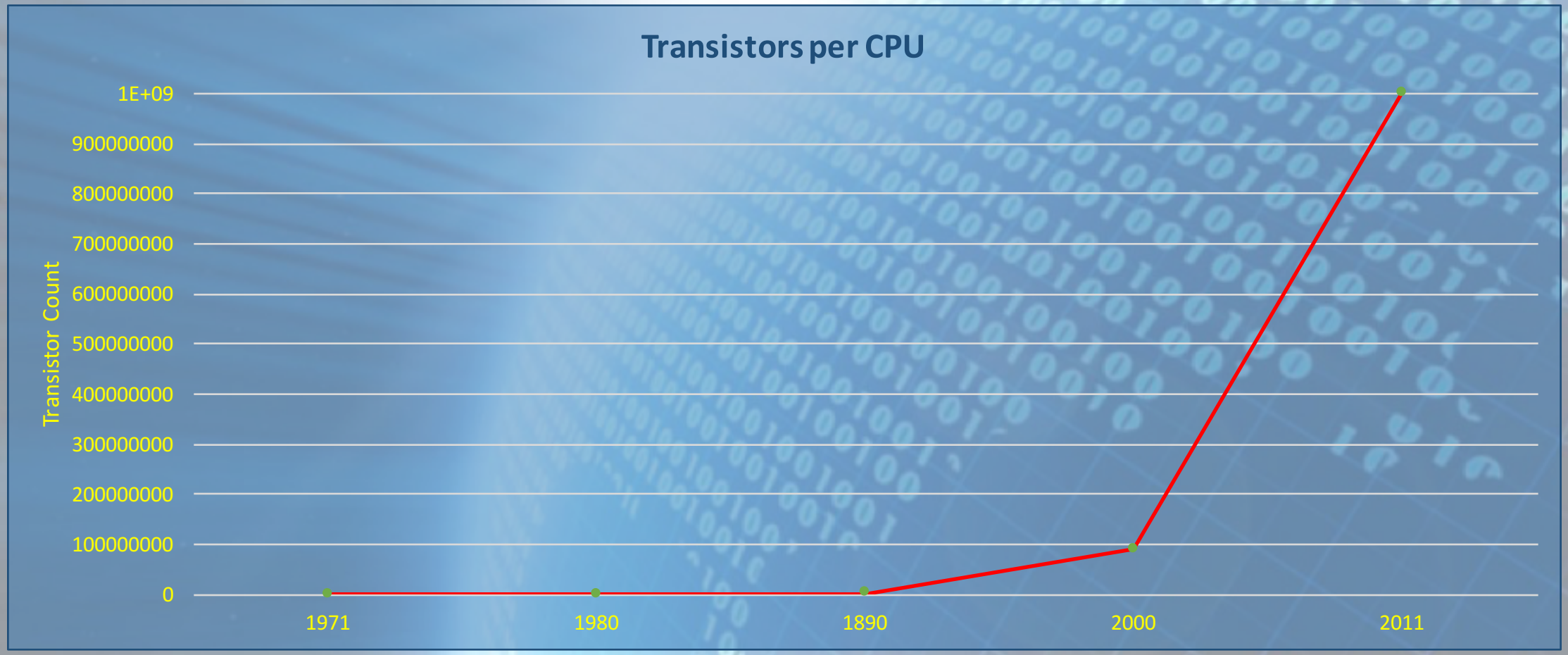
Addressing Data – Hard Disk Capacity



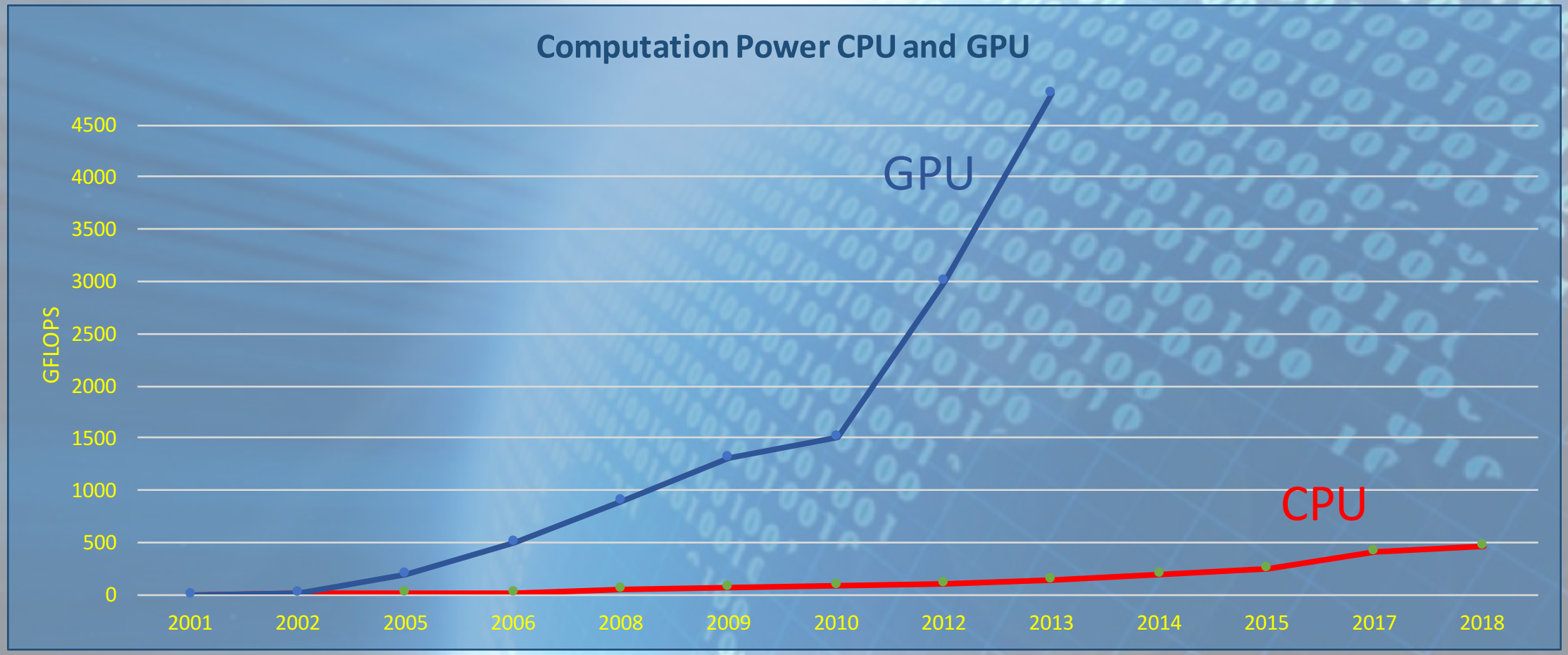
Addressing Data – Storage Cost



Transistors per CPU

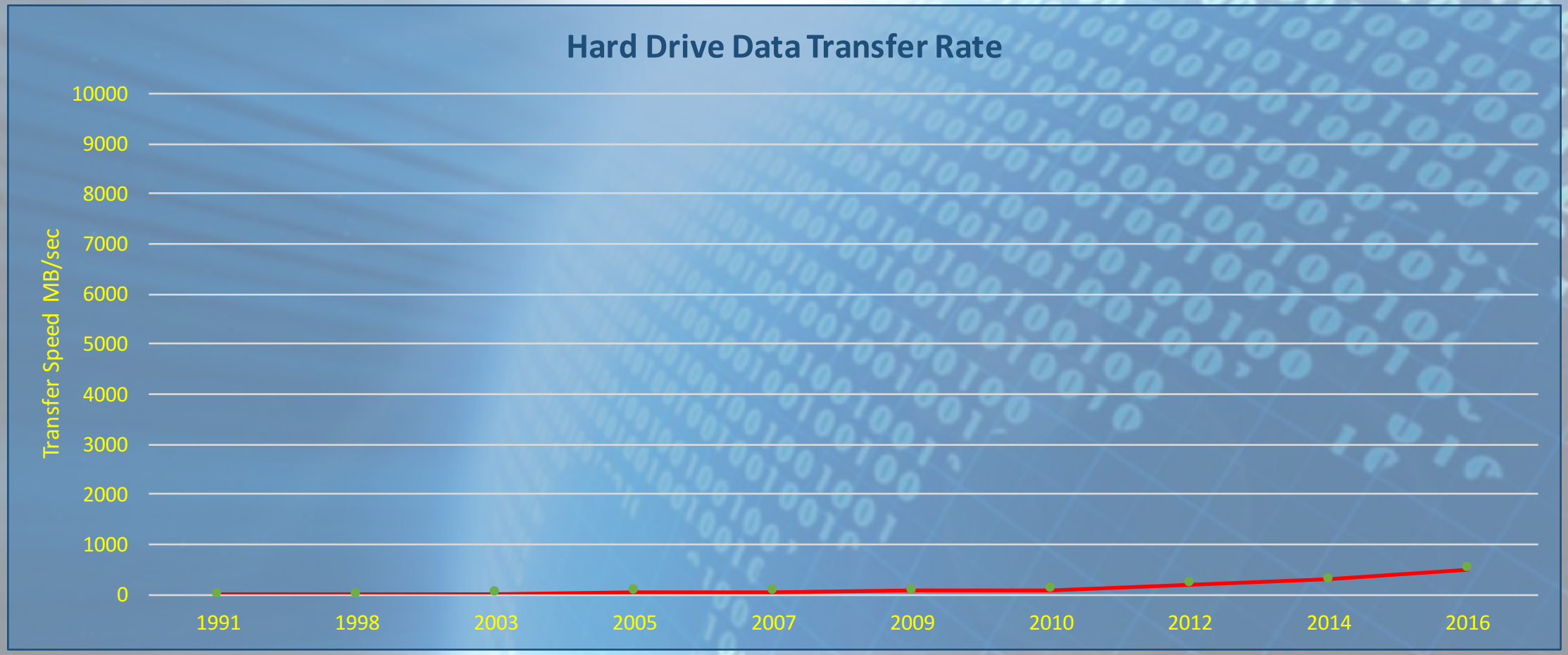


Computation Power CPU and GPU



Data Growth vs. Processing Power

Addressing Data – Transfer Rate



BIG DATA and TRADITIONAL SYSTEMS

The “Big Data” Problem Statement

Problem

- A single machine cannot process or even store all the data!

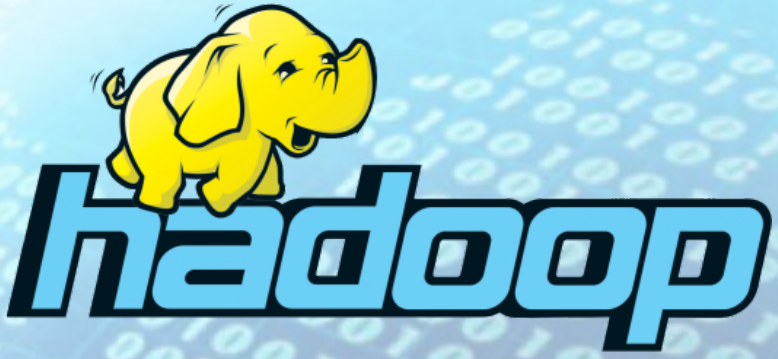
Solution

- Distribute data over many computers (large clusters)

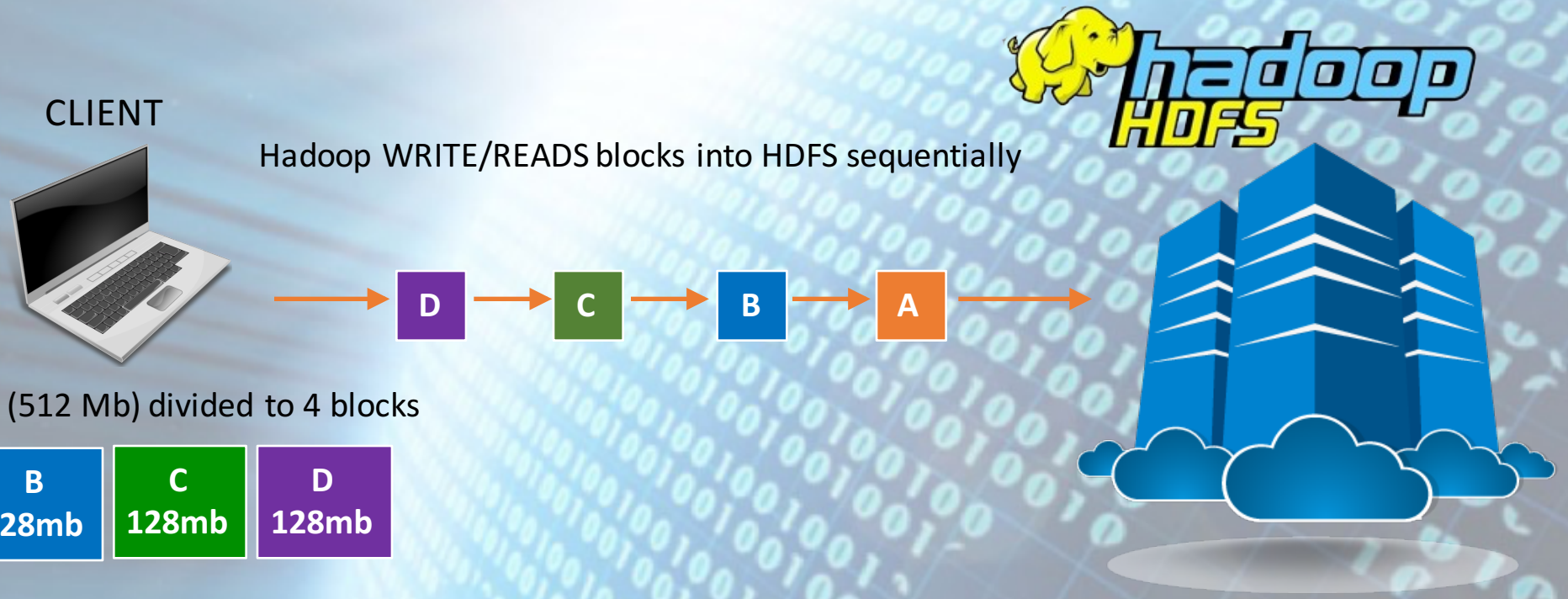
Difficulty

- How to split work across machines?
- Moving data over network is expensive
- How to deal with failures?
- How to deal with slow machines?

Hadoop Ecosystem Components



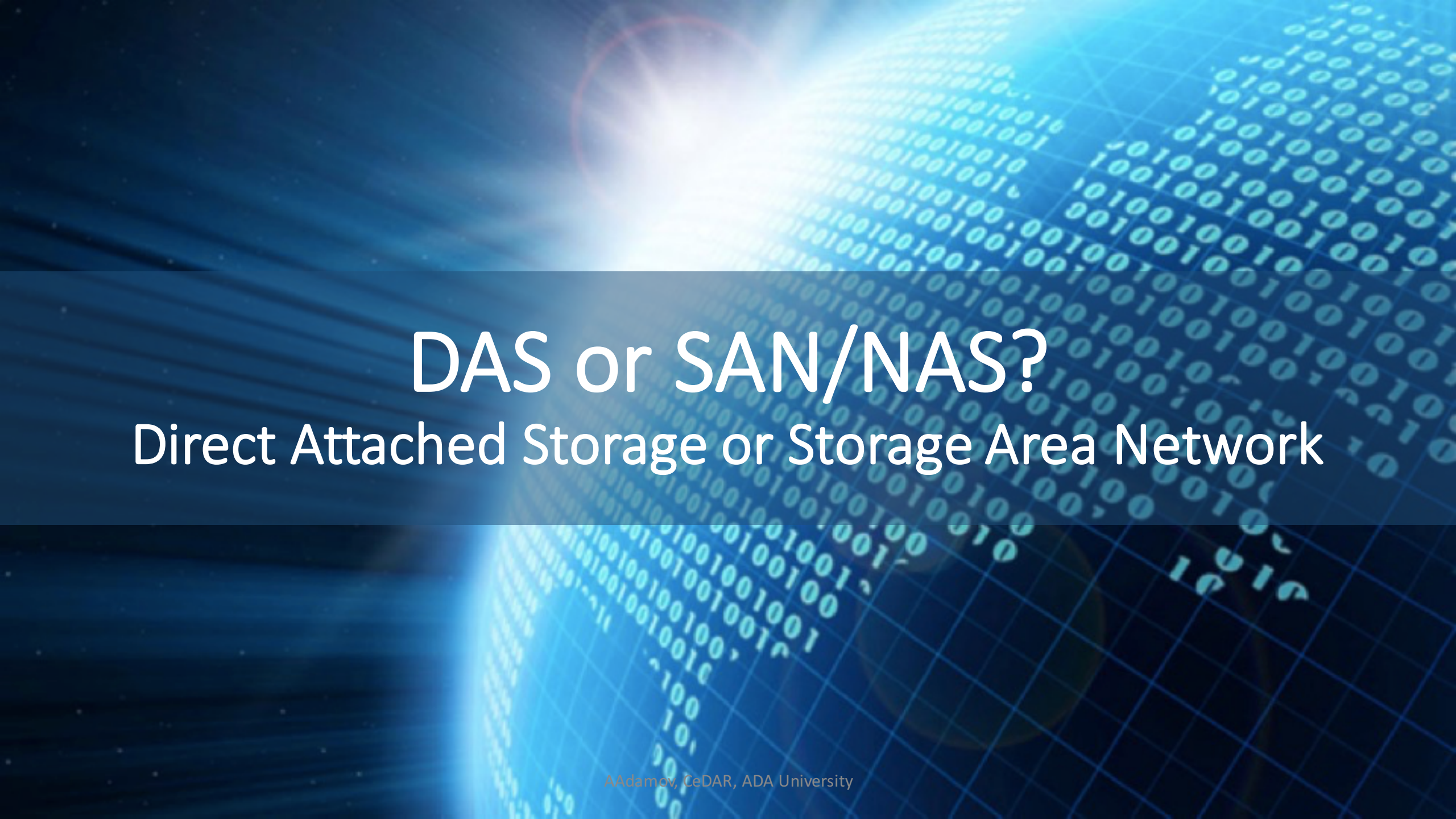
Hadoop Input/Output Model



Hadoop Reads/Writes blocks sequentially, not in parallel. Its why Hadoop does not affect IO performance significantly.

SOLUTION is Data Striping technique…

BIG DOES NOT MEAN SLOW

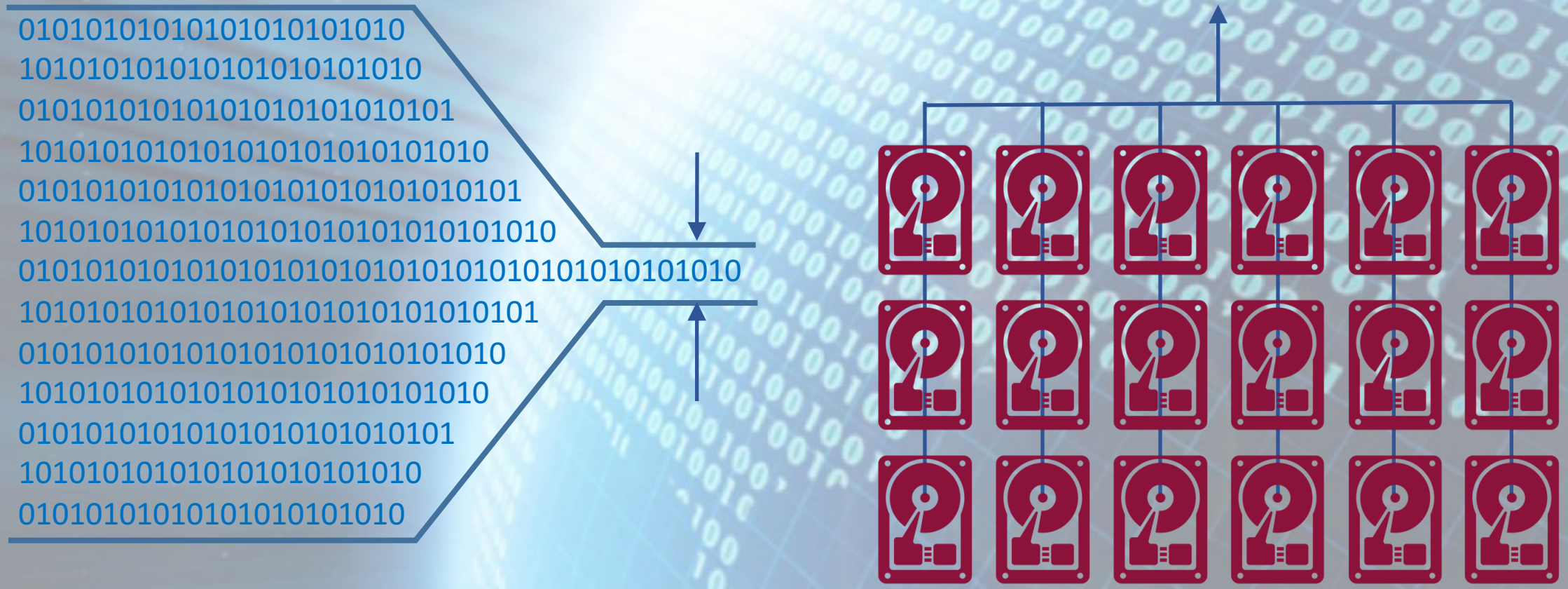


DAS or SAN/NAS?

Direct Attached Storage or Storage Area Network

Problem with Keeping Data Live

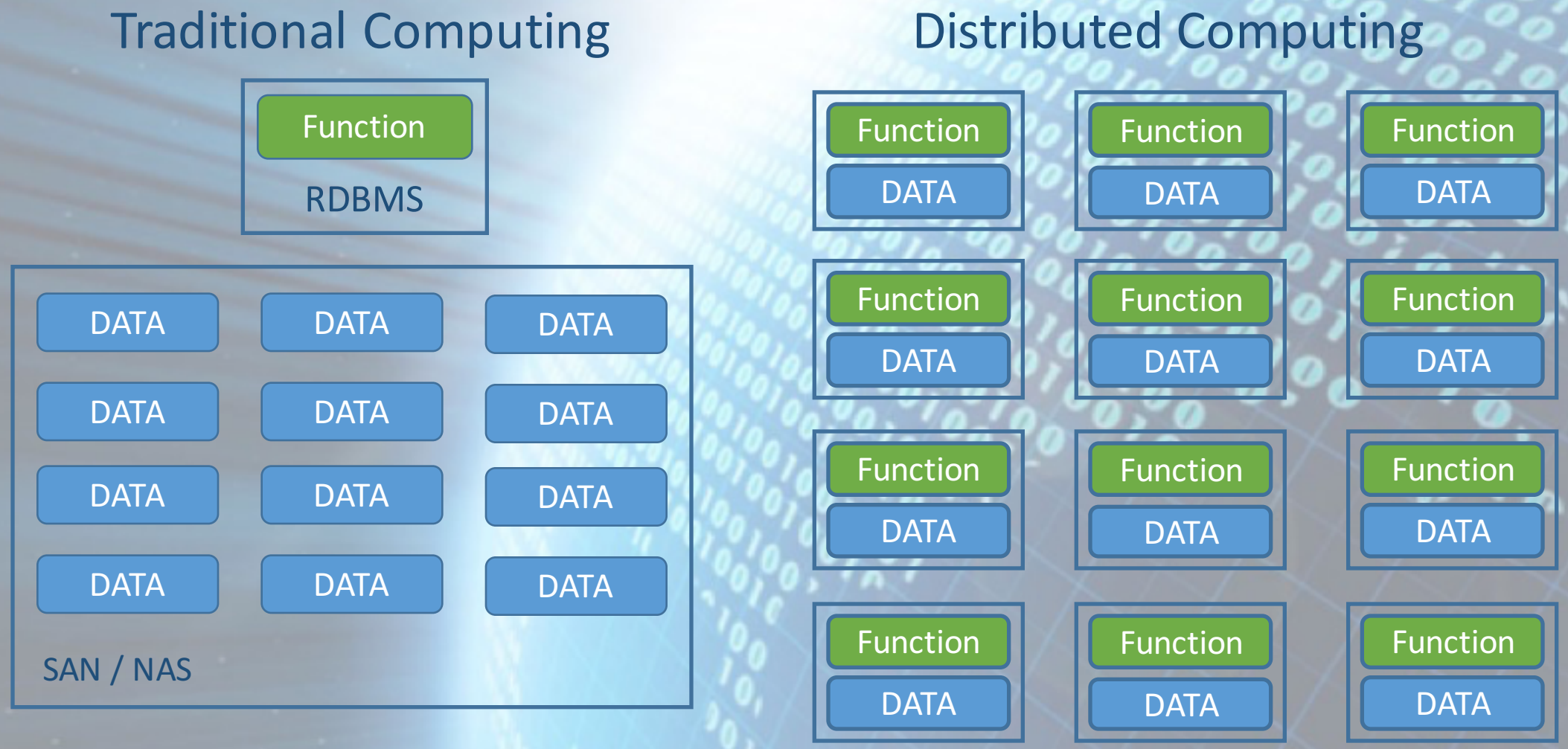
Network throughput is Bottleneck



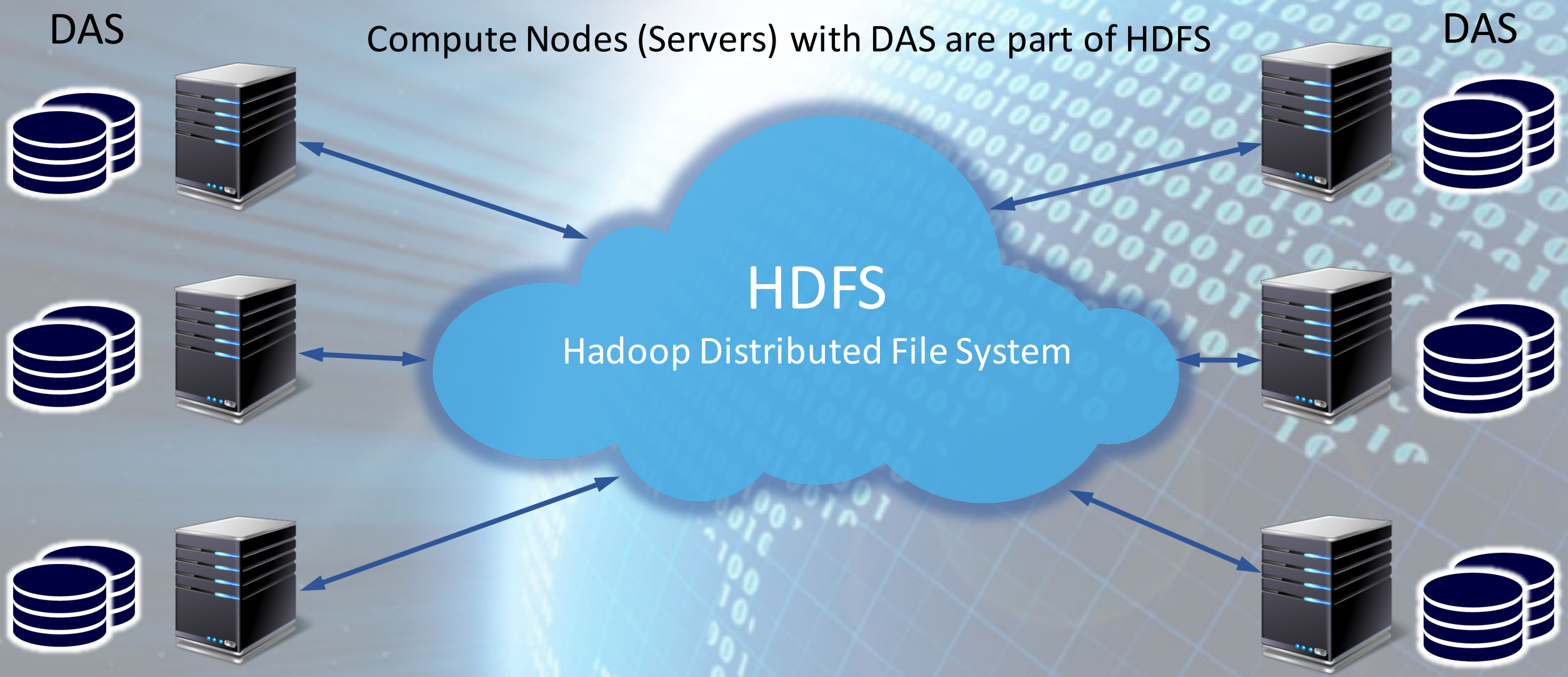
Cons with SAN/NAS

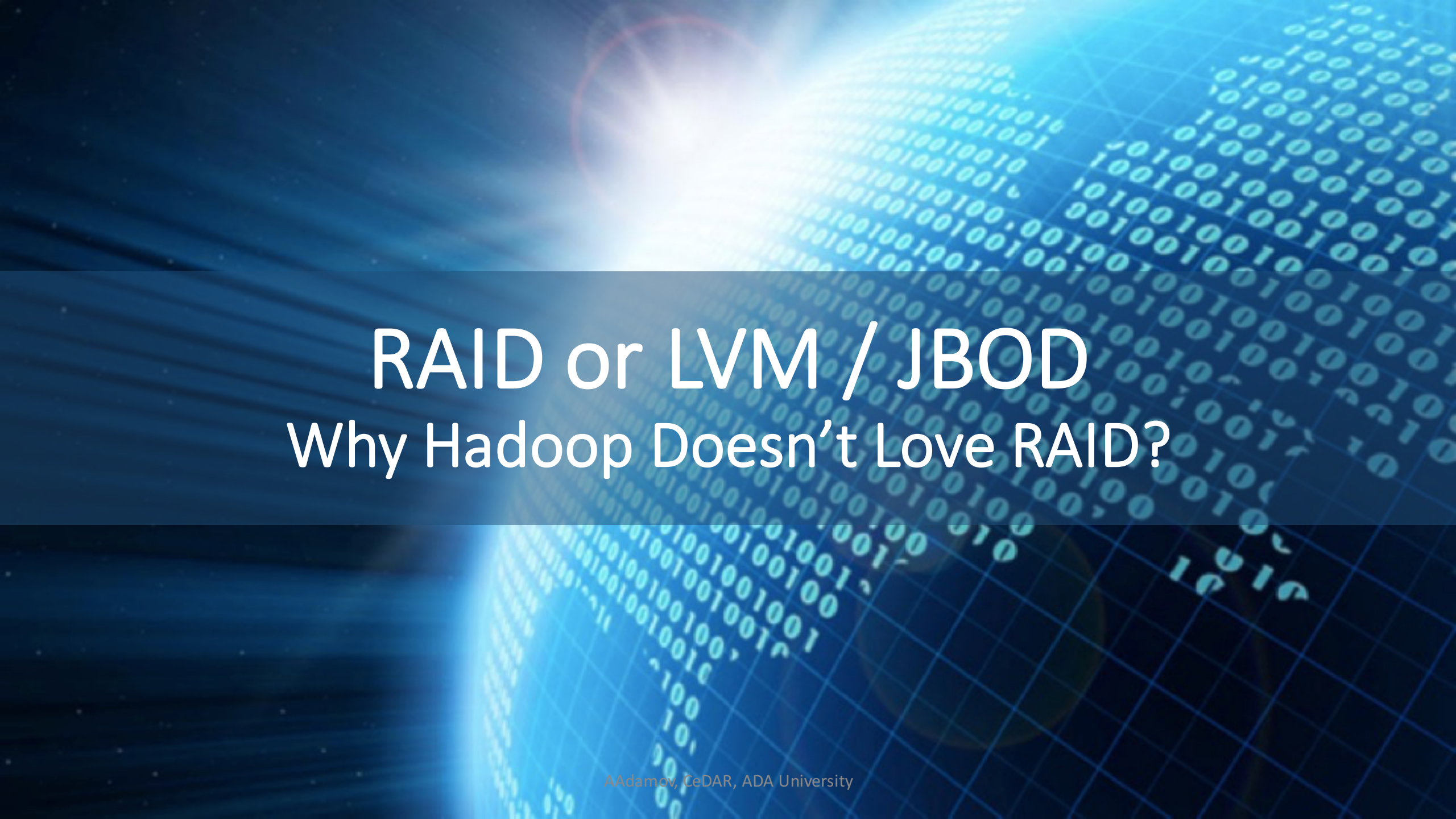
- Taking server's **components apart** decreases performance and Data transfer rates, at the same time increases workloads with high input/output operations;
- To run any query **all the data has to be moved** to the processing unit before any filtering can occur;
- Generally, requires same type and same size SAS **expensive drives**;
- Even "NAS needs less hardware", MapReduce jobs will **need to store intermediate data in local storage**.

Distributed vs Traditional Computing



HDFS Prefers DAS

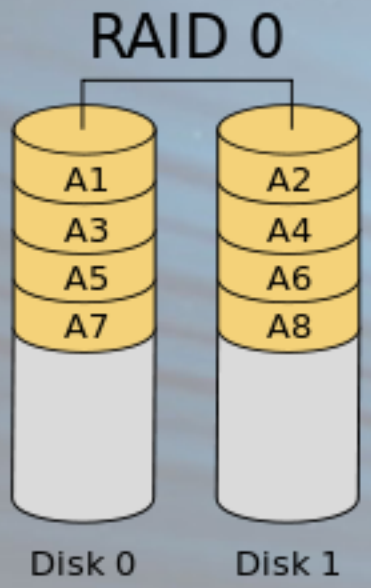




RAID or LVM / JBOD

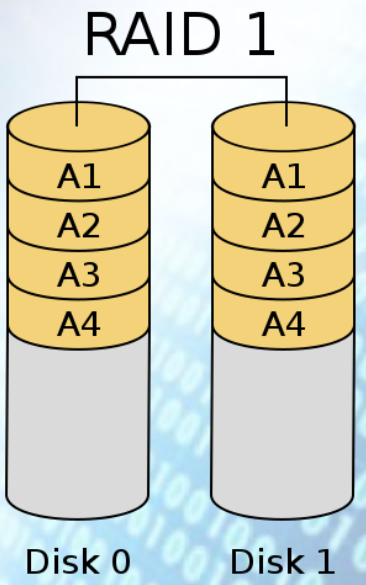
Why Hadoop Doesn’t Love RAID?

Truth about RAID



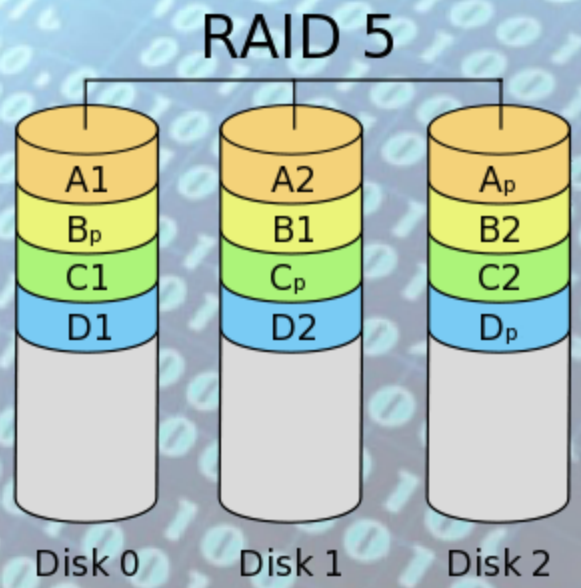
RAID 0, the data is striped (split/divided) across multiple disk drives.

Redundancy – NO
Performance – YES
Space Efficiency - YES



RAID 1 provides redundancy by storing exact copy of data in one disk on another disk

Redundancy – YES
Performance – NO
Space Efficiency - NO

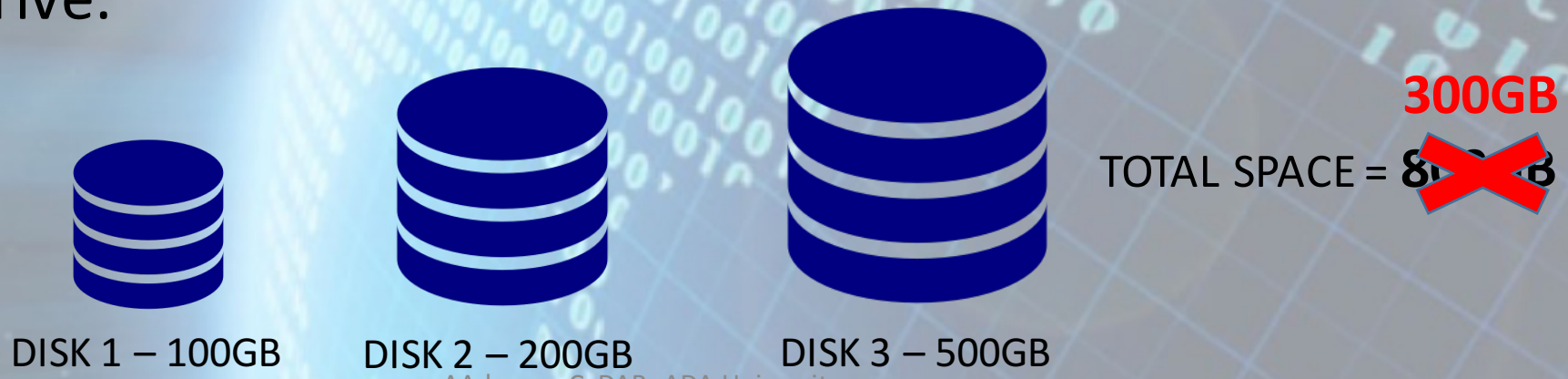


RAID 5, data and parity is used and are striped across the disks.

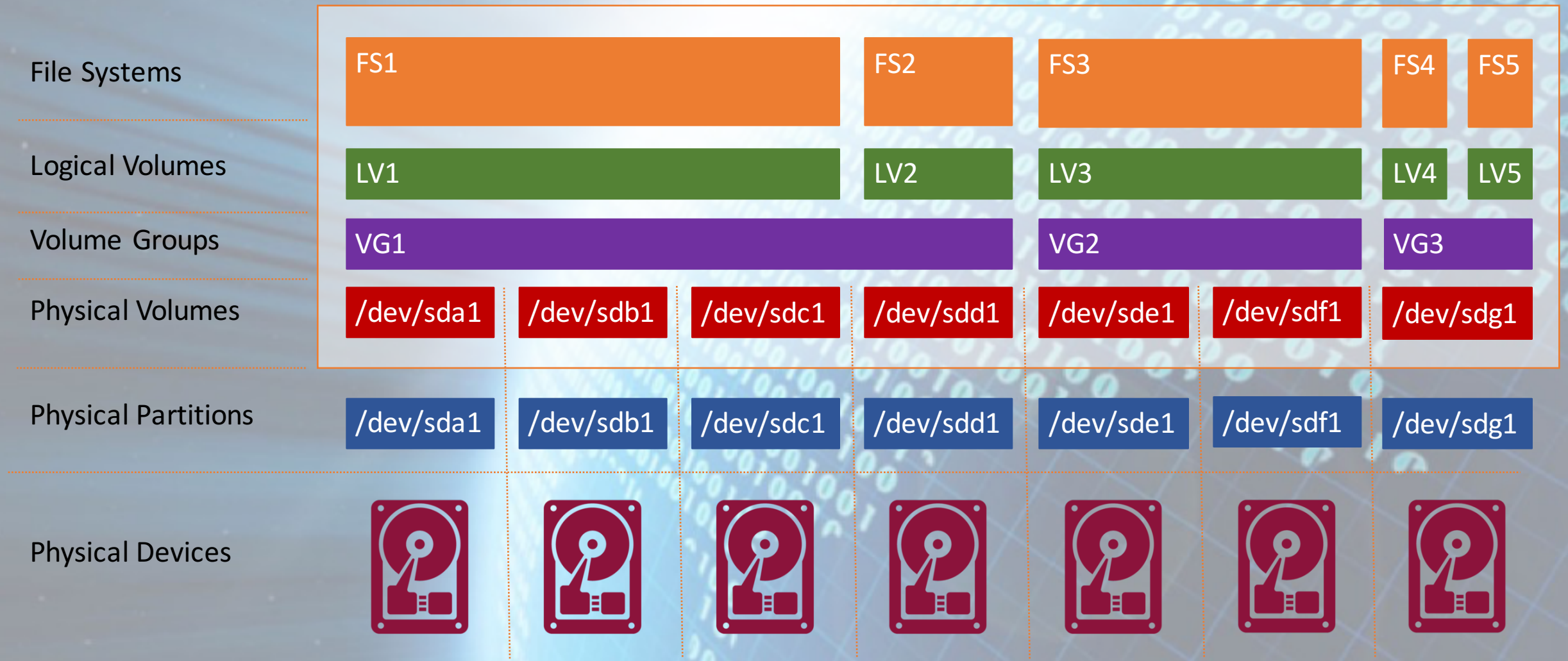
Redundancy – YES
Performance – NO NO
Space Efficiency – NO YES

Cons of RAID

- **Fault-tolerance** – There is no fault-tolerance in RAID 0
- **RAID technique is not reliable** – disks tend to fail. More discs in the array there is more chance for fail;
- **Lagging performance** – RAID delivers data at the rate of the slowest disk in the array. Disk speeds can vary up to 20%
- **Space Efficiency** – Total capacity is calculated based on the size of smallest drive.



LVM – Logical Volume Manager



LVM vs. RAID

- RAID requires all physical drives should have **same size**;
- RAID is a physical grouping of disks to present them to an OS **as one logical device**;
- RAID is used for **Redundancy or Performance** or a combination of the two;
- RAID is **NOT any kind of Data backup** solution
- RAID-0 **is not scalable** - what means RAID-0 group cannot grow without affected the Data

- LVM is a logical layer **to create, manipulate and expand a logical presentation** of a disk device(s) to an OS.
- LVM allows to **use entire space of each drive** if they are different sizes.
- LVM **enables growth** extending Logical Volumes through adding new drives **without affecting the Data**;
- LVM is a disk management approach that allows to create, extend, reduce, delete or resize the volume groups or logical volumes;
- LVM can be used to manage a large pool of disks what is called JBOD (Just-a-bunch-of-Disk)

Why Hadoop is against RAID?

- **The redundancy that RAID provides is not needed**, since HDFS handles it by replication between nodes;
- **JBOD (Just a Bunch Of Disks)** configuration recommended for HDFS, mostly is faster then RAID 0;
- If a disk fails in JBOD, **HDFS can continue to operate without it**, but in RAID if a disk fails the whole array becomes unavailable.

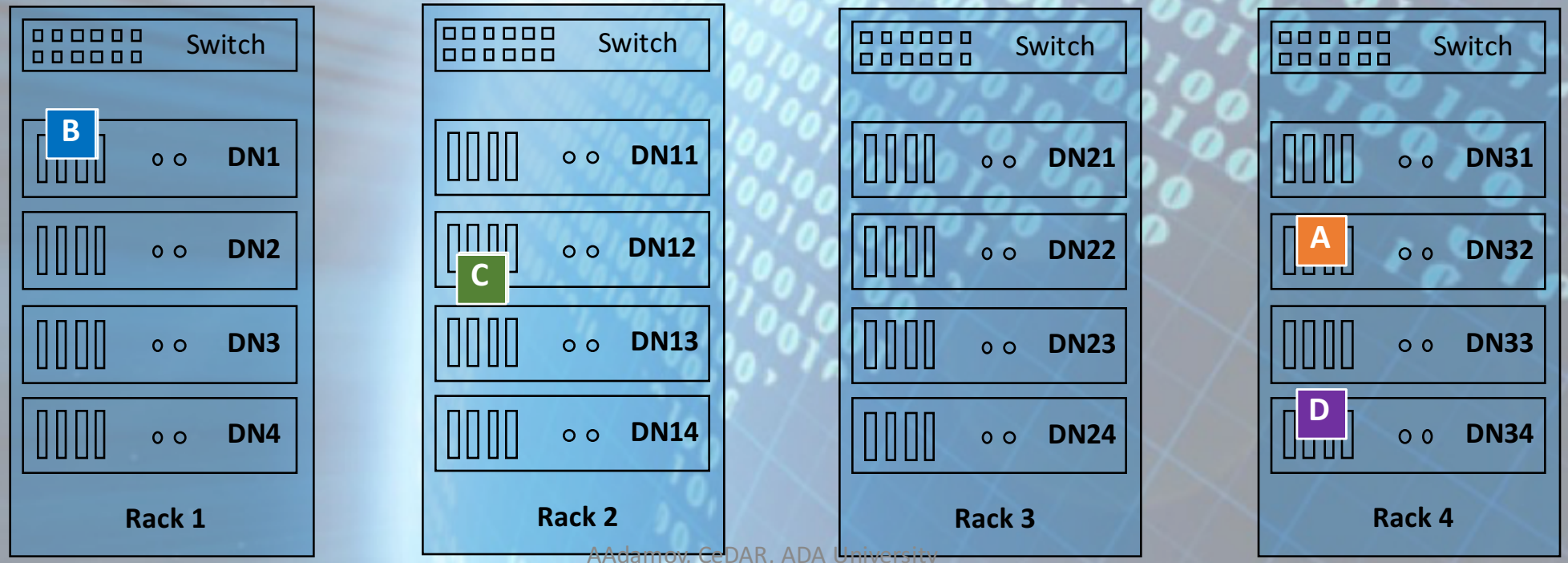
Distributed Architecture of HDFS

CLIENT

Where to write file ADA.txt (blocks A, B, C, D) in HDFS?

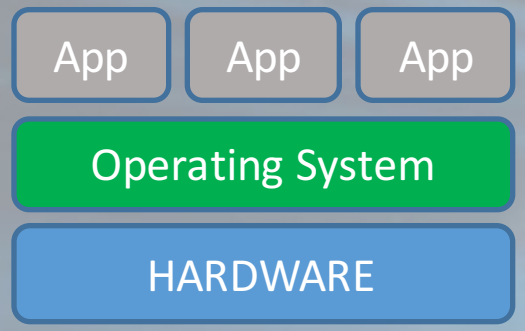
NAMENODE

A – DN32, 11, 14
B – DN01, 22, 23
C – DN12, 02, 04
D – DN34, 12, 14

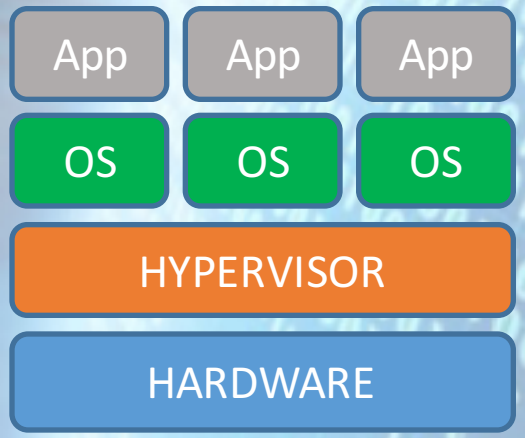


Big Data and Virtialization

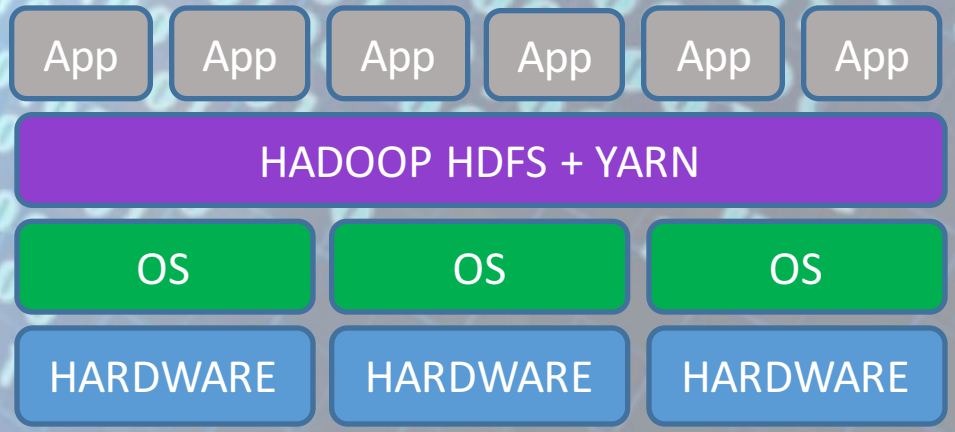
Traditional Architecture



Virtualized Architecture



Distributed Architecture



Computing Facilities at CeDAR

Computing Cluster - the primary component of the CeDAR. This is powerful, scalable and fault-tolerant computing cluster based on distributed architecture, which will operate totally on open-source software.

Characteristics of computing cluster:

- Processing CPU Cores: 102
- Combined Memory: 1,568 TB
- Total Storage: 136 TB

CEDAR.CLUSTER.ADA

Service	Host	Status	24-Hour	Response
All	Any	All		
Ambari	dnode01.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [12.14%, 6.5 GB], Capacity Total: [53.7 GB], path=/us...
Ambari	dnode02.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [8.12%, 4.4 GB], Capacity Total: [53.7 GB], path=/usr/...
Ambari	dnode03.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [7.98%, 4.3 GB], Capacity Total: [53.7 GB], path=/usr/...
Ambari	dnode04.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [7.79%, 4.2 GB], Capacity Total: [53.7 GB], path=/usr/...
Ambari	dnode05.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [7.65%, 4.1 GB], Capacity Total: [53.7 GB], path=/usr/...
Ambari	dnode06.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [7.28%, 3.9 GB], Capacity Total: [53.7 GB], path=/usr/...
Ambari	dnode07.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [6.89%, 3.7 GB], Capacity Total: [53.7 GB], path=/usr/...
Ambari	dnode08.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [7.47%, 4.0 GB], Capacity Total: [53.7 GB], path=/usr/...
Ambari	dnode09.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [7.82%, 4.2 GB], Capacity Total: [53.7 GB], path=/usr/...
Ambari	dnode10.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [8.64%, 4.6 GB], Capacity Total: [53.7 GB], path=/usr/...
Ambari	dnode11.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [10.50%, 5.6 GB], Capacity Total: [53.7 GB], path=/us...
Ambari	dnode12.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [11.62%, 6.2 GB], Capacity Total: [53.7 GB], path=/us...
Ambari	dnode13.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [11.43%, 6.1 GB], Capacity Total: [53.7 GB], path=/us...
Ambari	dnode14.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [11.61%, 6.2 GB], Capacity Total: [53.7 GB], path=/us...
Ambari	dnode15.cedar.cluster.ada	OK for 2 days	0	Capacity Used: [12.46%, 6.7 GB], Capacity Total: [53.7 GB], path=/us...

DATA is the NEW OIL!

But do you have the capacity to refine it?

Q & A

Dr. Abzetdin Adamov,
Email me at: aadamov@ada.edu.az
Follow me at: @
Link to me at: www.linkedin.com/in/adamov
Visit my blog at: aadamov.wordpress.com