



31 August – 1 September 2026

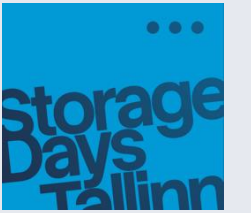
TALLINN, ESTONIA

The Disk Is Not the Truth: Forensics in the Capacity Era

Finding ground truth in modern storage systems

Pavel Chikuk

CTO – Trumo Finance



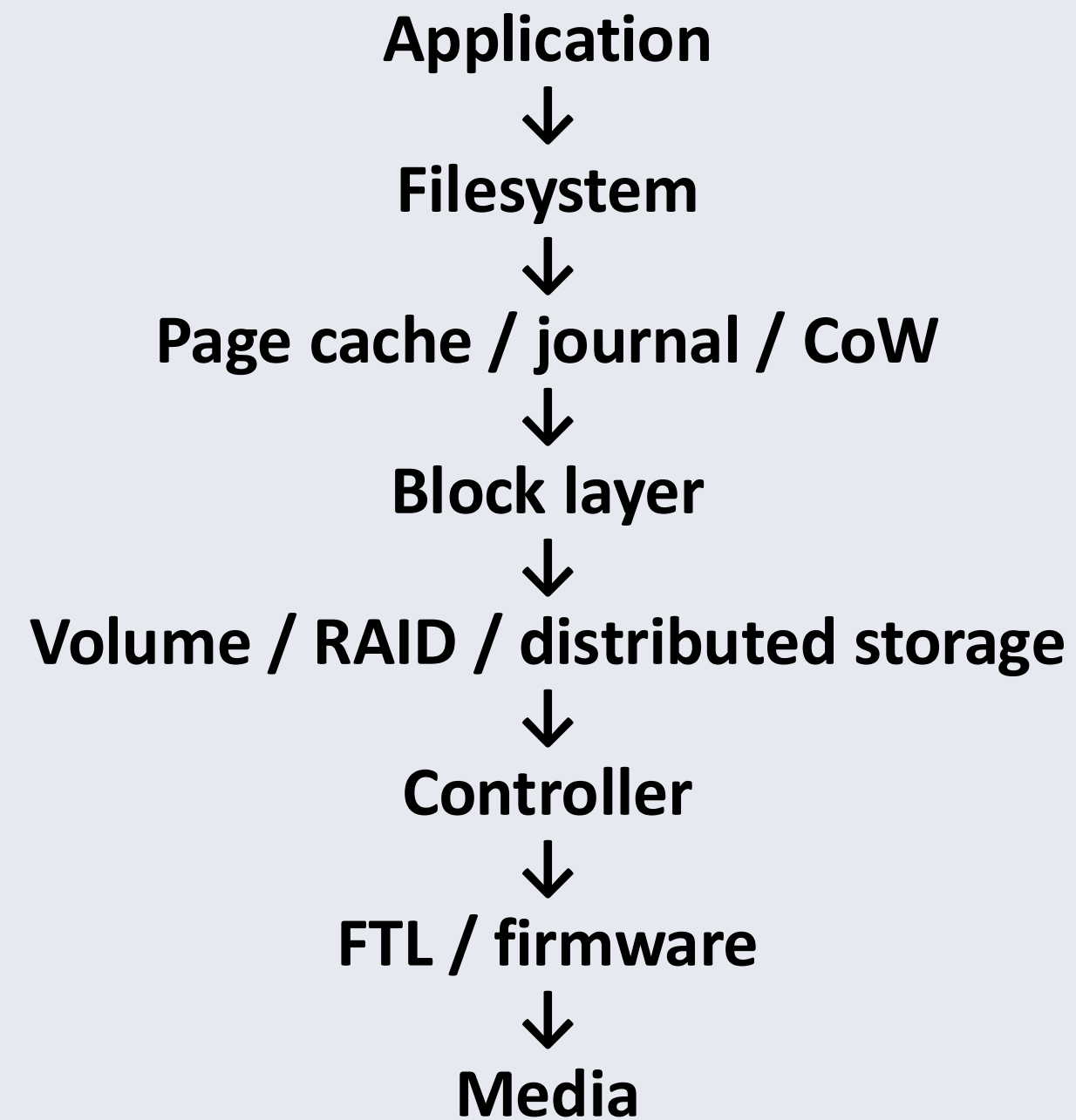
Where is the file?

An application writes a file.

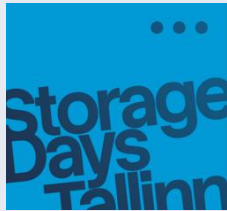
Where is it?

`write()` → ??? → persistence

The Stack



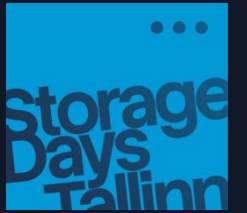
Which layer owns the truth?



Multiple Truths

TRUTH	EXAMPLE
Application truth	“write succeeded”
Filesystem truth	“block X belongs to file A”
Device truth	“LBA X contains ...”
Physical truth	“some pages currently encode ...”

All can be internally consistent. And still disagree.



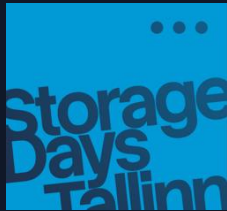
An Example Case

Application checkpoint is updated with a secure pattern:

```
write(temp)
fsync(temp)
rename(temp, target)
fsync(directory)
```

Logs read: **Configuration saved successfully**

Power loss occurs.



An Example Case

Clue #1:

Filesystem: clean
Journal replay: successful

Clue #2:

checkpoint.dat
size: 536870912
mtime: 02:14:36

Clue #3

In filesystem journal namespace update happened
New inode is bound to checkpoint.dat.

App start reports: **Checksum failure!**

offset	generation	
00000000	1842	
00001000	1842	
00002000	1842	
...		
0317F000	1842	
03180000	1841	<---
03181000	1841	
03182000	1841	
...		
03200000	1842	
03201000	1842	
...		
08A40000	1841	<---
08A41000	1841	
...		

Which layer is lying here?

An Example Case

Expectation

- fsync() succeeded
- rename completed
- therefore new file survived

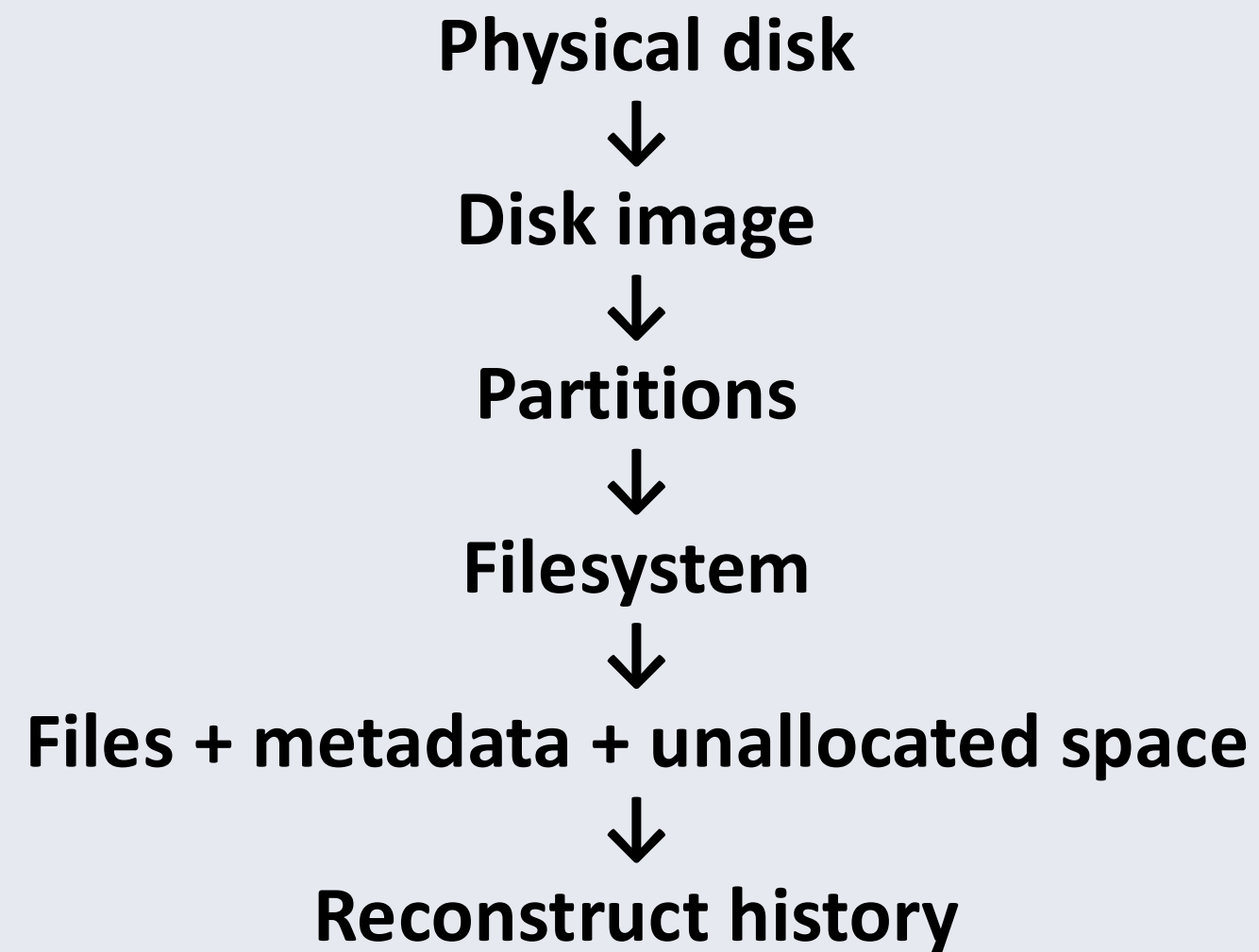
Observation

- directory state is new
- metadata is new
- some data is old

Explanation

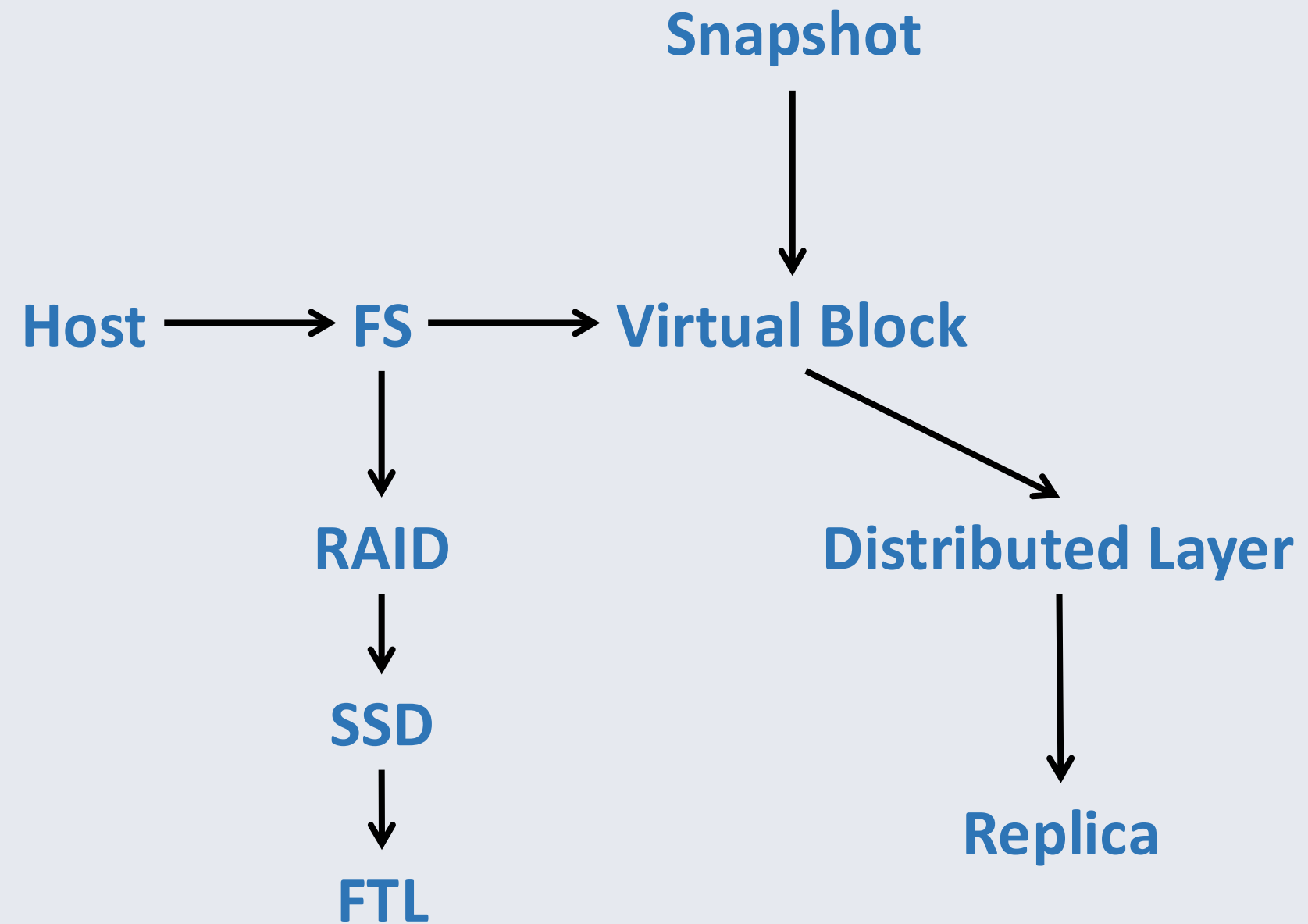
- the filesystem's durability guarantee guarantee
- depended on a lower layer honoring honoring persistence semantics

The Classical Forensic model



**Acquire everything.
Interpret later.**

After the File System



Where do you acquire “everything”?

You have the drives. Do you have the data?

I give you every physical drive from the system.

Do you have the data?

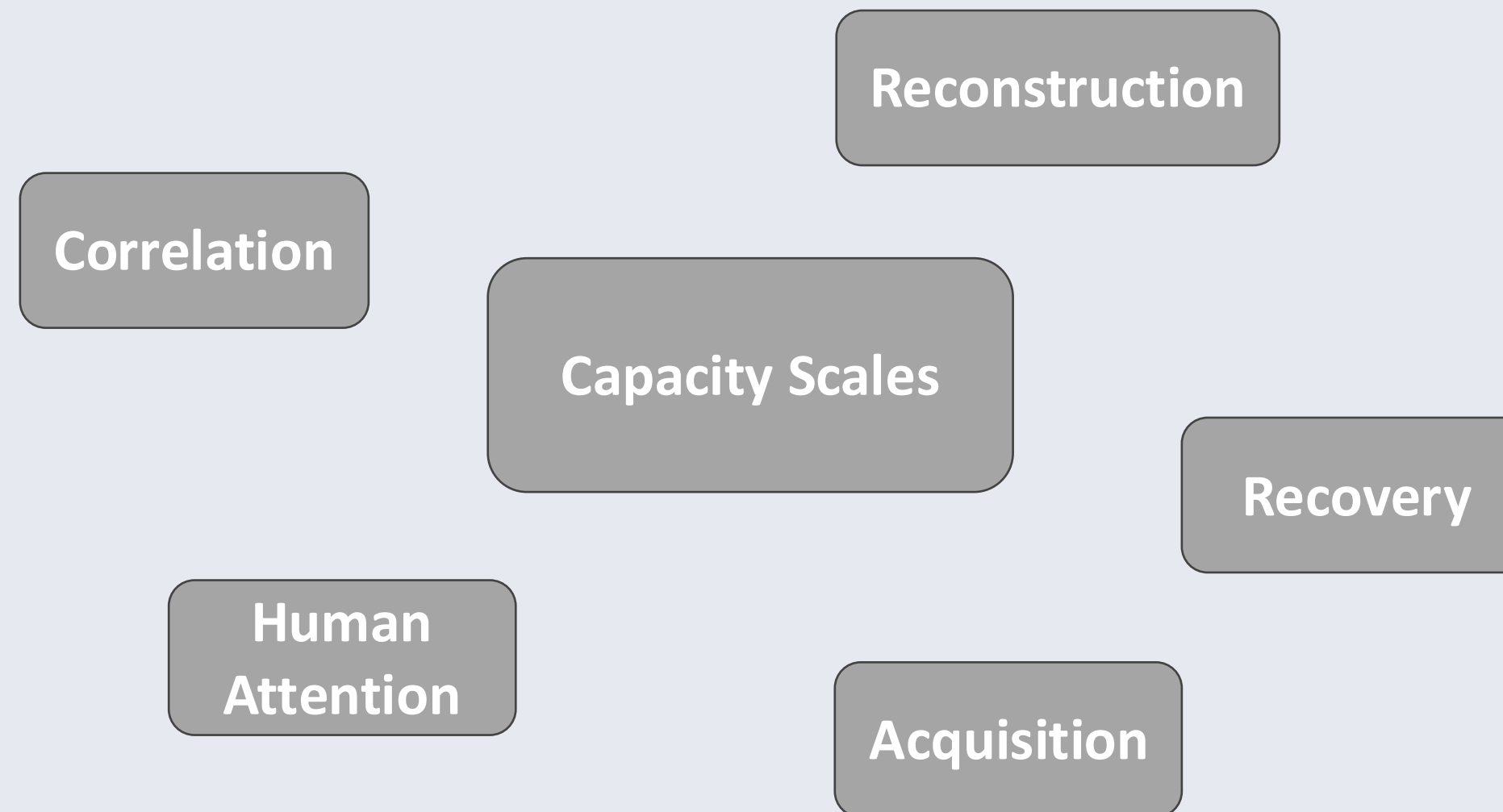
Maybe.

Then capacity happens

The old forensic primitive:
“Read everything.”

Does this still scale?

What doesn't scale?



Not at the same rate!

A Missing Storage Property?

- Capacity
- Latency
- Throughput
- Availability
- Durability
- Cost
- **Inspectability?**

Three things I learned from forensics

01 Persistence $\neq$ evidence

02 Recovery $\neq$ understanding

03 Scale changes the investigation

In the Capacity Era, are we building systems that can store more data - or systems we can still understand when something goes wrong?



Thank you

Questions and discussion

Speaker Name name@organisation.com

