



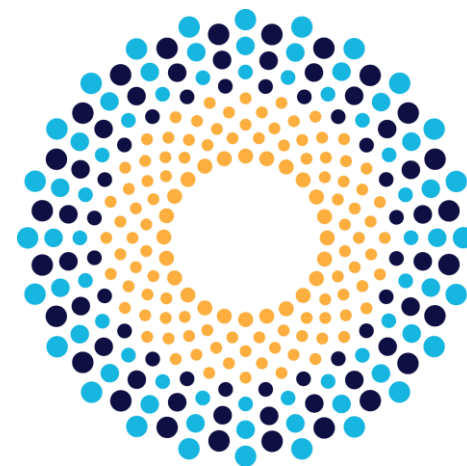
Science and  
Technology  
Facilities Council

Scientific Computing

# Accounting errors and where to find them

Inspired by IRIS accounting and beyond...

Prepared by D. Chekrygina  
with support from A. Coveney

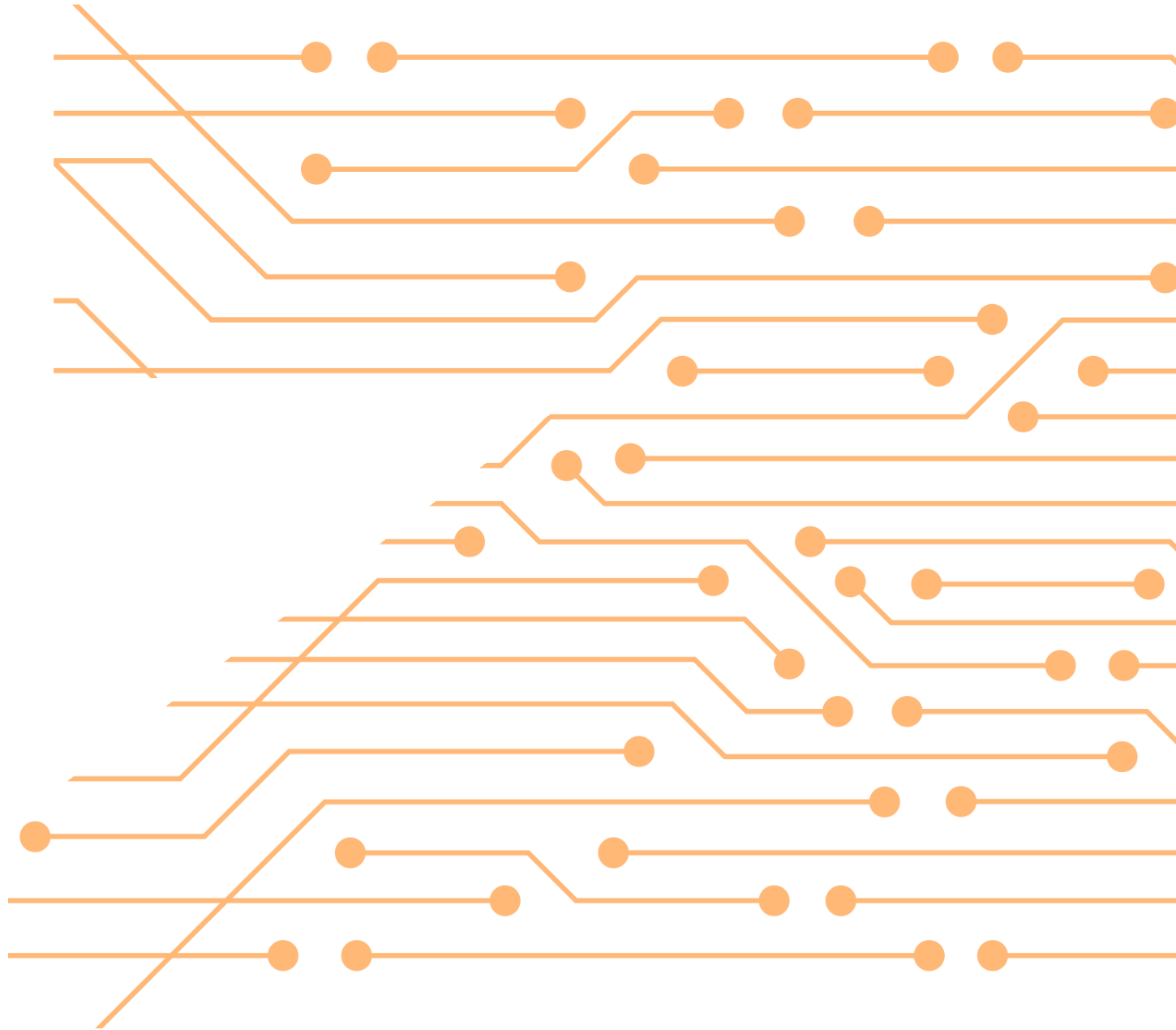


**iris**



Science and  
Technology  
Facilities Council

# Introduction

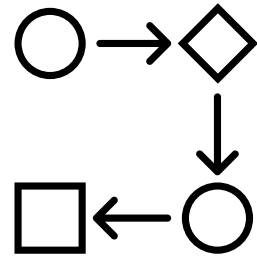


Science and  
Technology  
Facilities Council

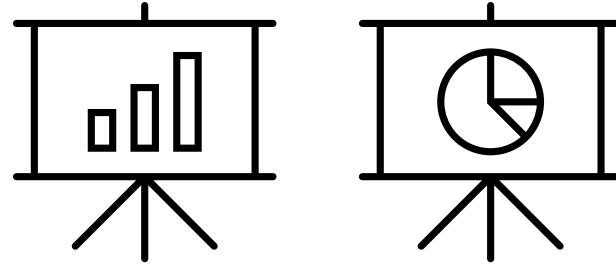
Scientific Computing

# Who we are and what we thought we have

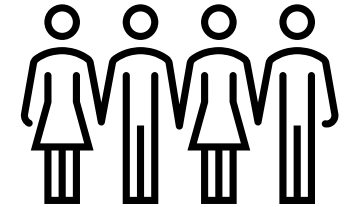
We have set processes for procurement, planning, allocation and resource scrutiny, etc



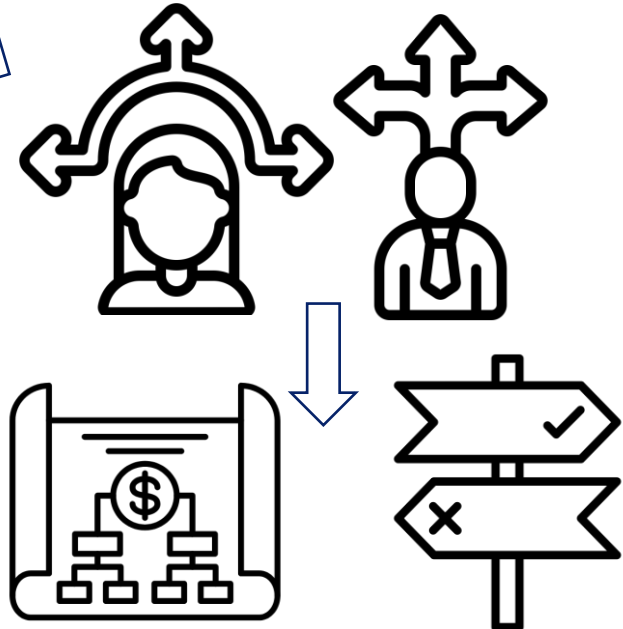
We have an accounting dashboard



We have an experienced community with shared understanding

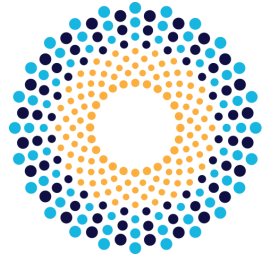


and decision-makers



And many more other necessary for DRI things, of course!

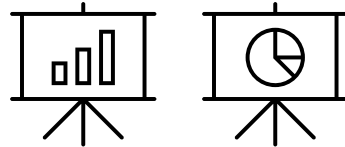
# Accounting report Q1 2023: how it was supposed to be



1<sup>st</sup> July



IRIS  
accounting  
dashboard



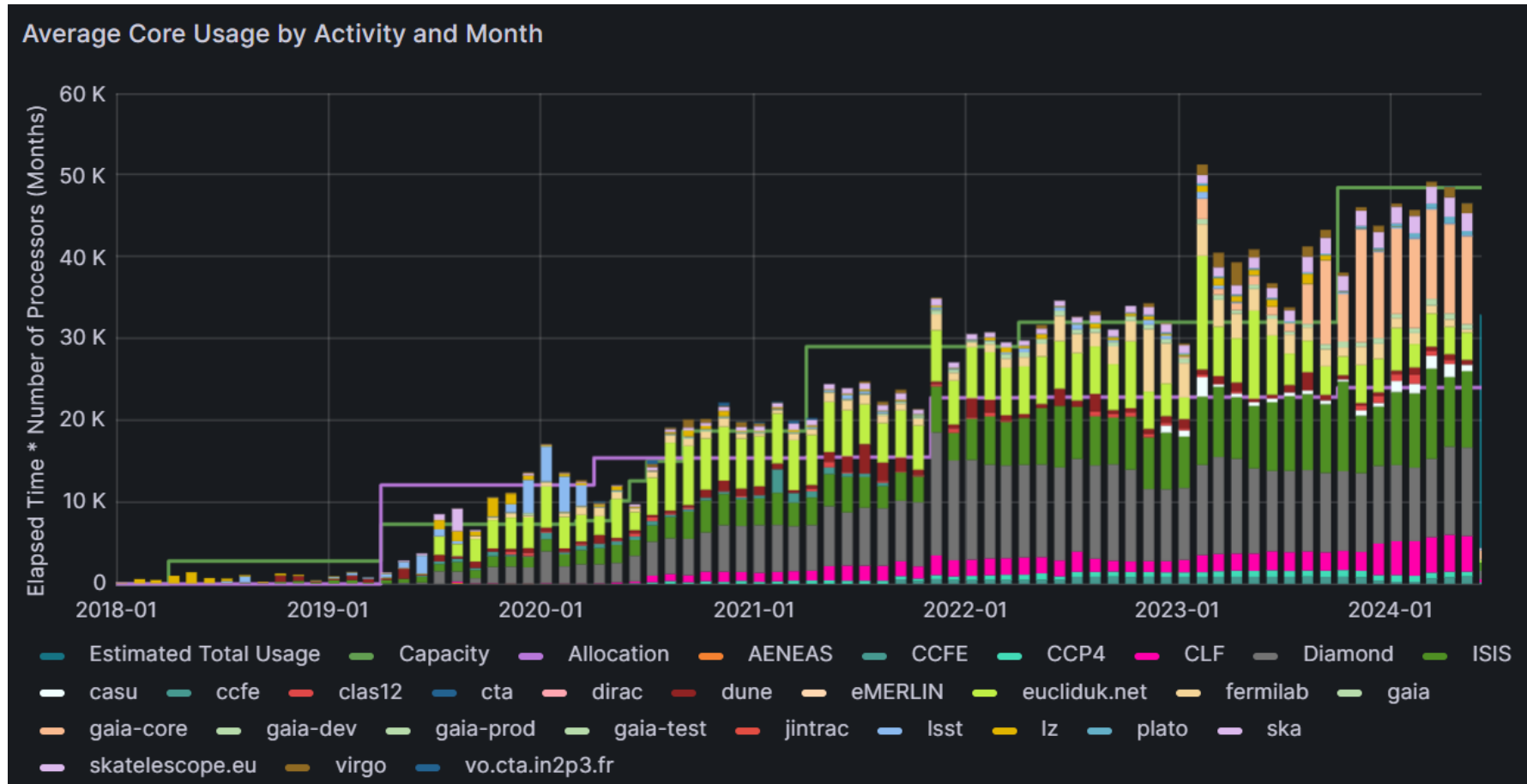
Hardware  
accounting  
report



15<sup>th</sup> July

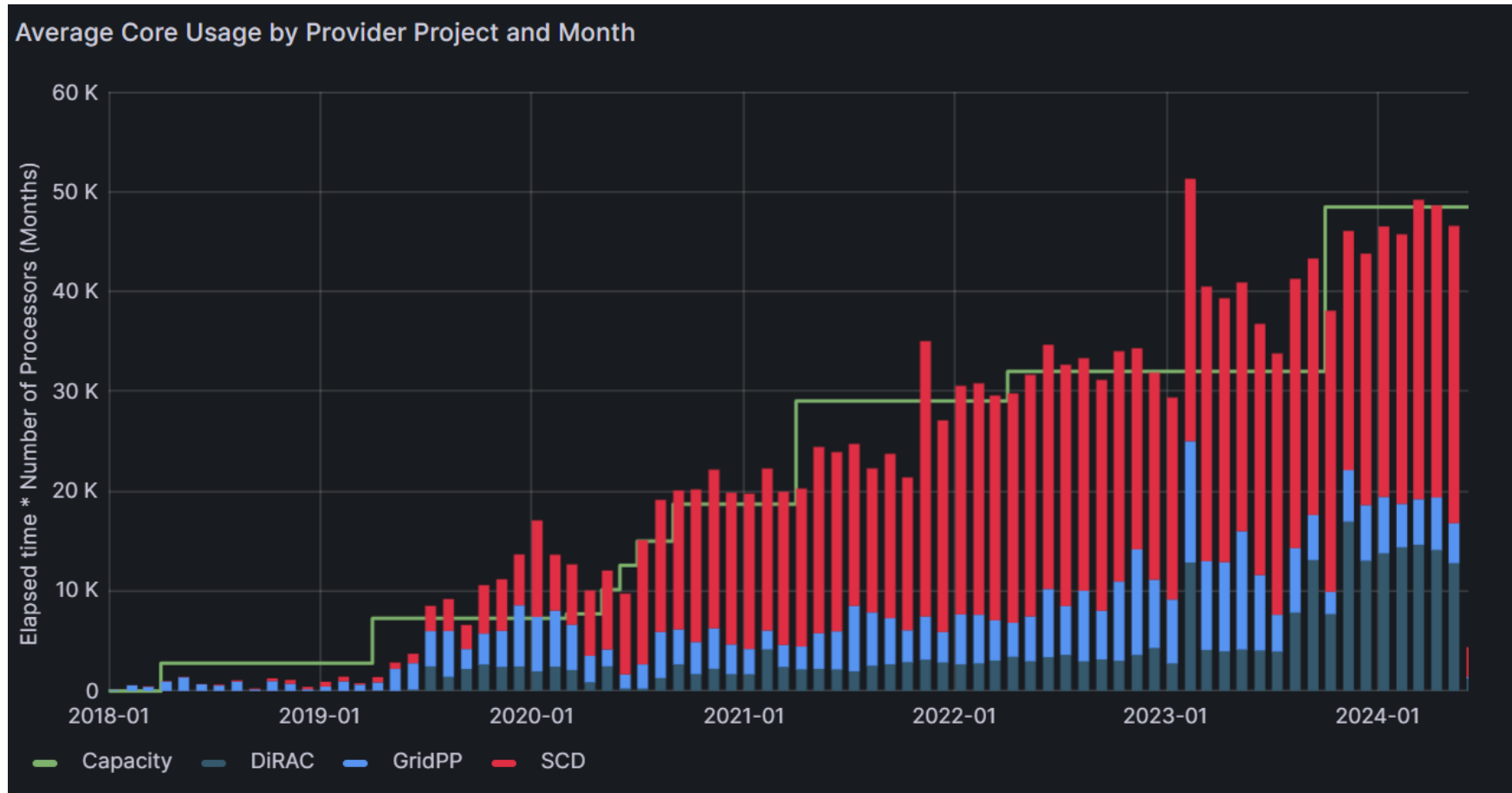


# IRIS accounting dashboard: users view



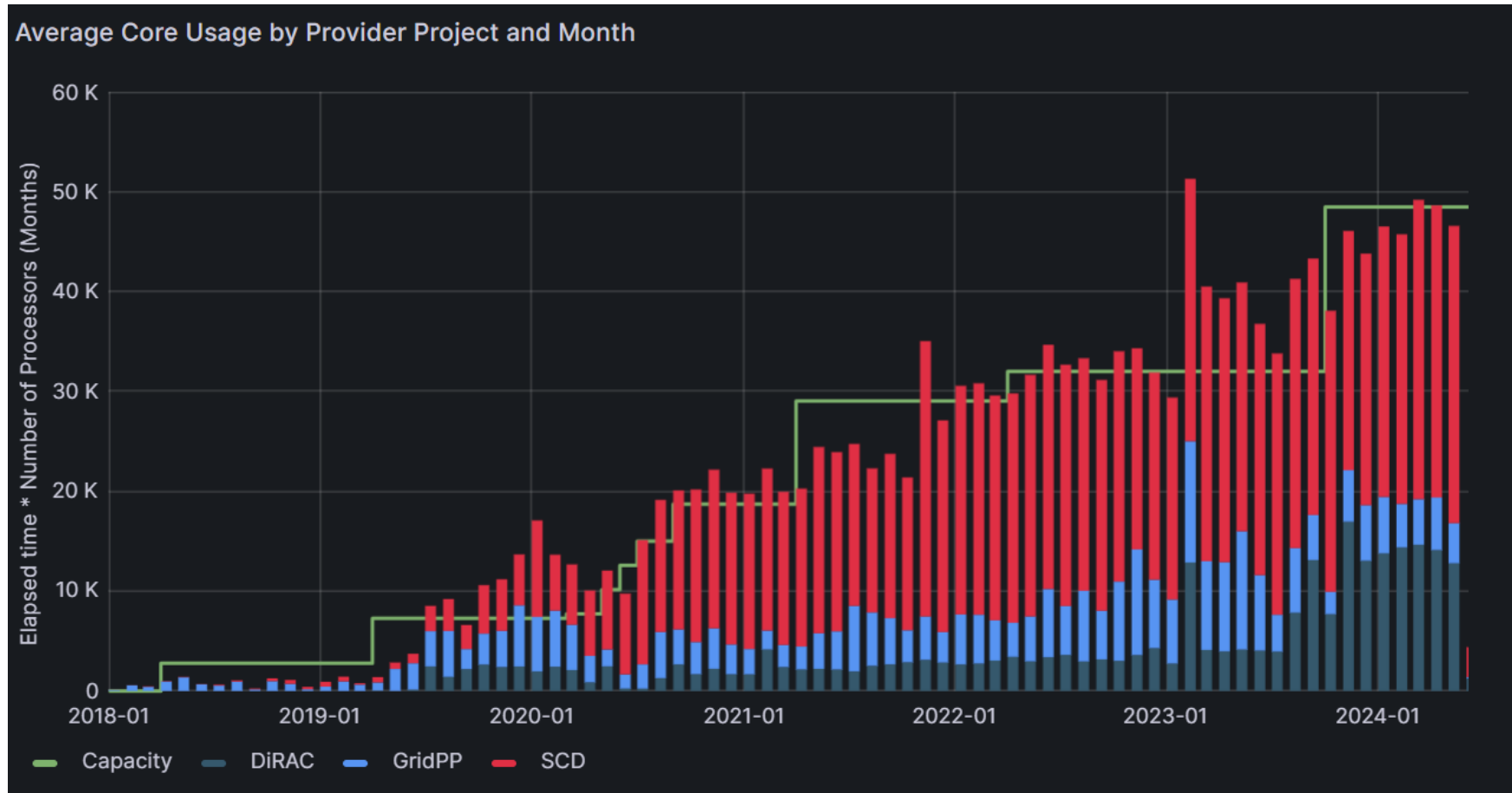
Use values are close to IRIS capacity.

# IRIS accounting dashboard: providers view



← Also looks reasonable for three main groups of IRIS providers.

# What it actually consists of



Physical cores  
(different cores,  
many generations,  
RAM varies from  
2GB/core to 20)



Virtual cores  
(hyperthreaded  
2:1)



Logical cores  
(hyperthreaded  
and not, so  
1.5ish:1)



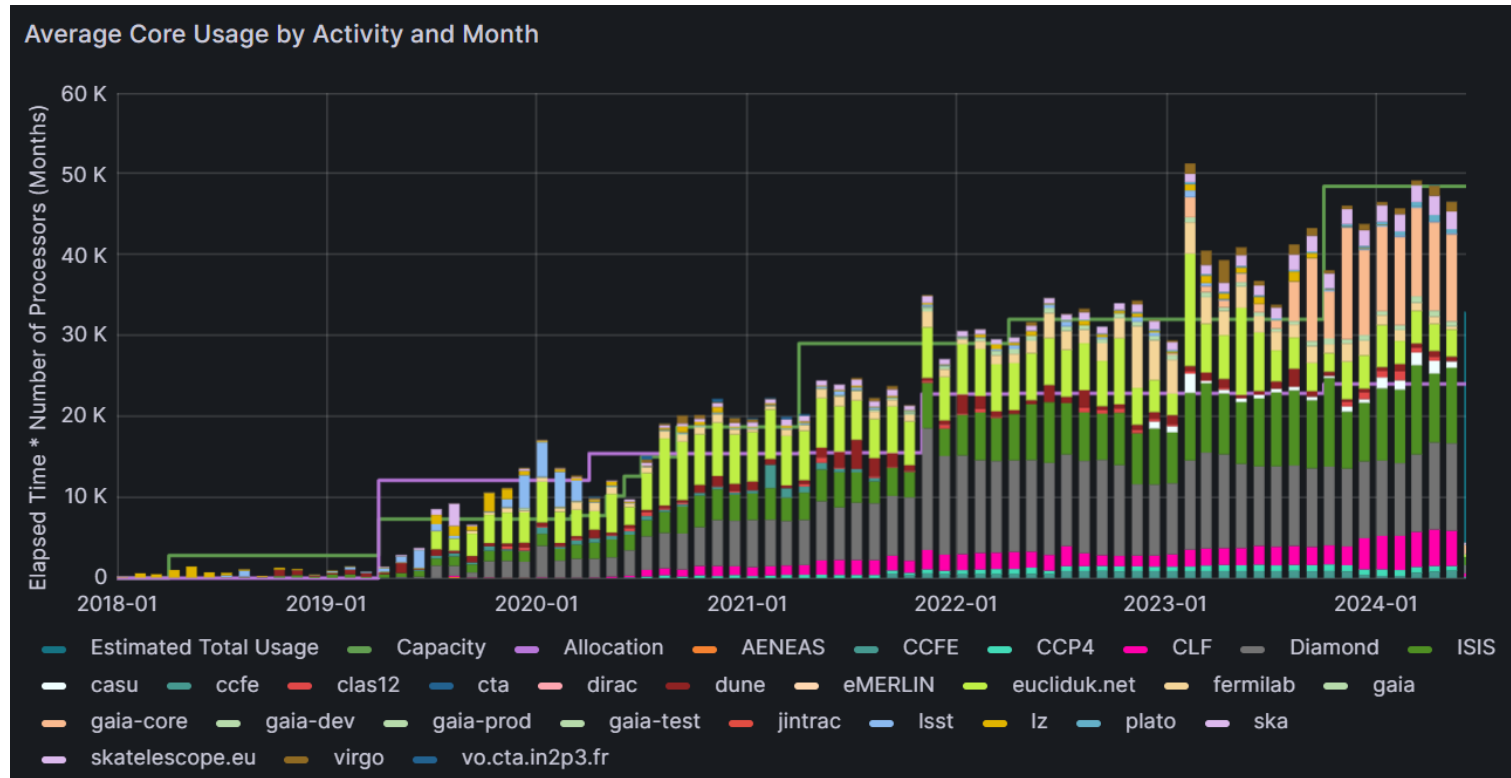
Virtual and  
physical (some  
2:1, other 1:1)



# What are the consequences

## Consequences

- Users think these are physical cores
- Users ask IRIS for physical cores
- IRIS might buy in some cases **TWICE** as much hardware as needed

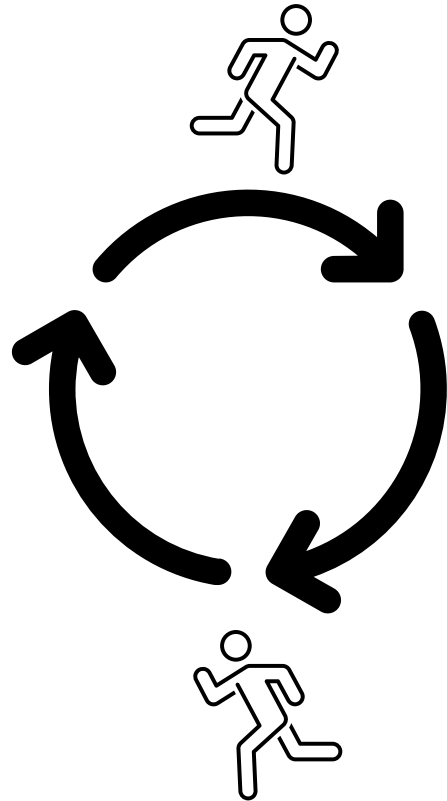
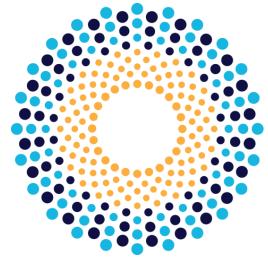


## Which means

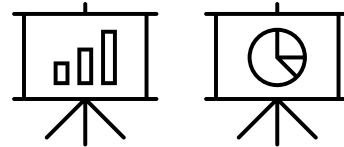
- Wasted money
- Wasted electricity
- Wasted resources



# Accounting report Q1 2023: how it actually went



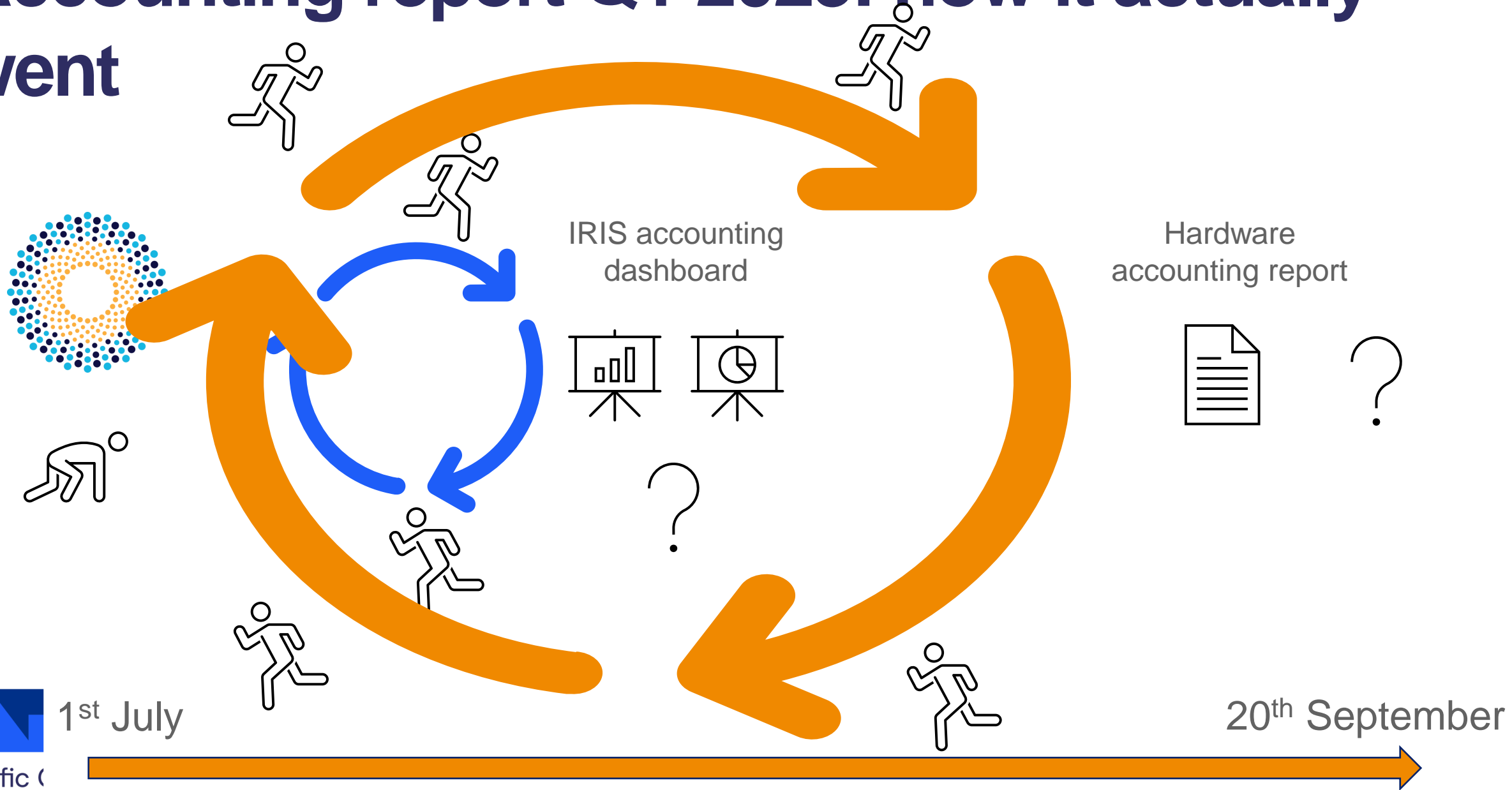
IRIS accounting dashboard



Hardware accounting report



# Accounting report Q1 2023: how it actually went



1st July

Scientific (

20th September



Science and  
Technology  
Facilities Council

# Now to some theory!



Science and  
Technology  
Facilities Council

Scientific Computing

# The value of information to a user is determined by its reliability.\*

Information must possess certain attributes—relevance, accuracy, completeness, summarization, and timeliness.\*

# In addition, information risks are increased due to

## the remoteness of the information

- When the source of the information is removed from the decision maker, the information stands a greater chance of being misstated.

## the volume and complexity of the underlying data

- When volume and complexity of the information increase-can result in the information overload. Increases the chance that misstated information may exist undetected.

## the motive of the preparer

- Those who prepare information may have goals different from those of the decision maker. The information may be slanted in favour of a particular viewpoint or incentive.

# What we discovered

## Accounting in finances has similar diseases to Accounting in IT.

### **Data entry errors.**

Typos, misplaced decimals and etc.

### **Errors of commission.**

Or an incorrect action error, can be when part of the record is incorrect, miscalculations, improperly reversed values, duplicates

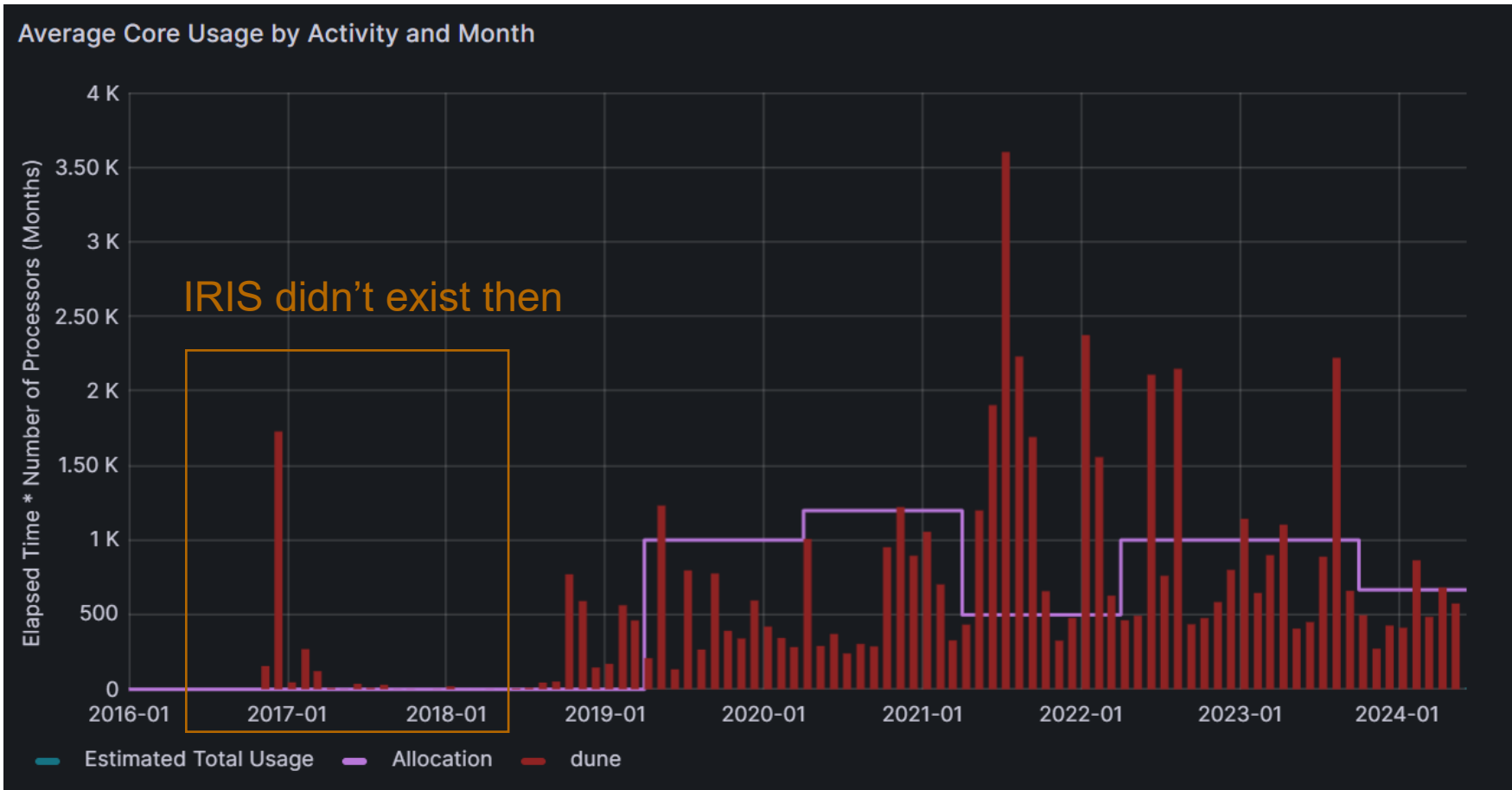
### **Errors of principle.**

When logic of calculation is incorrect, can be either because the wrong guidance was followed or followed incorrectly.

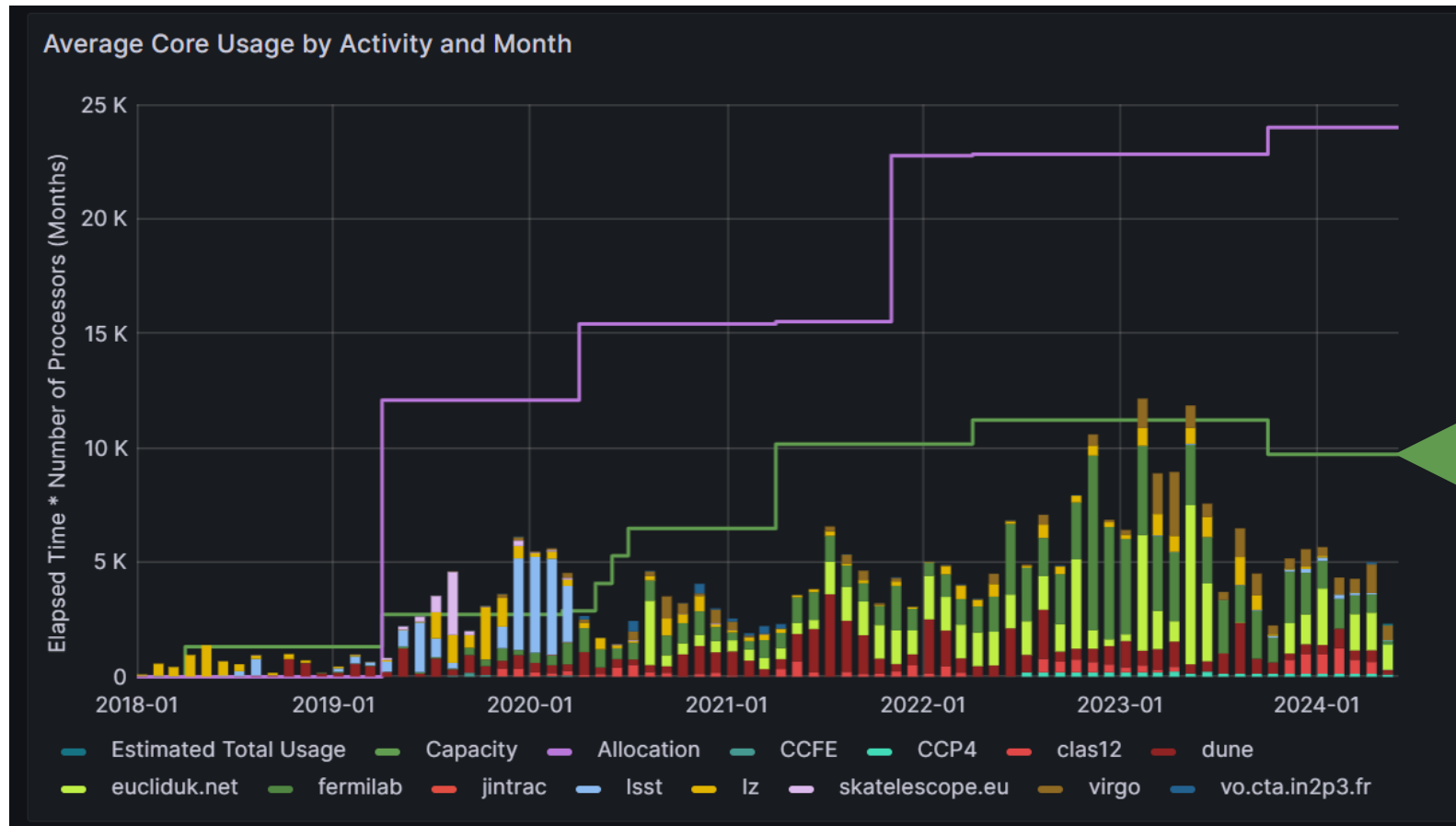
### **Errors of omission.**

When one of the transactions overlooked and not recorded

# Data entry error



# Error of commission/omission(both apply)

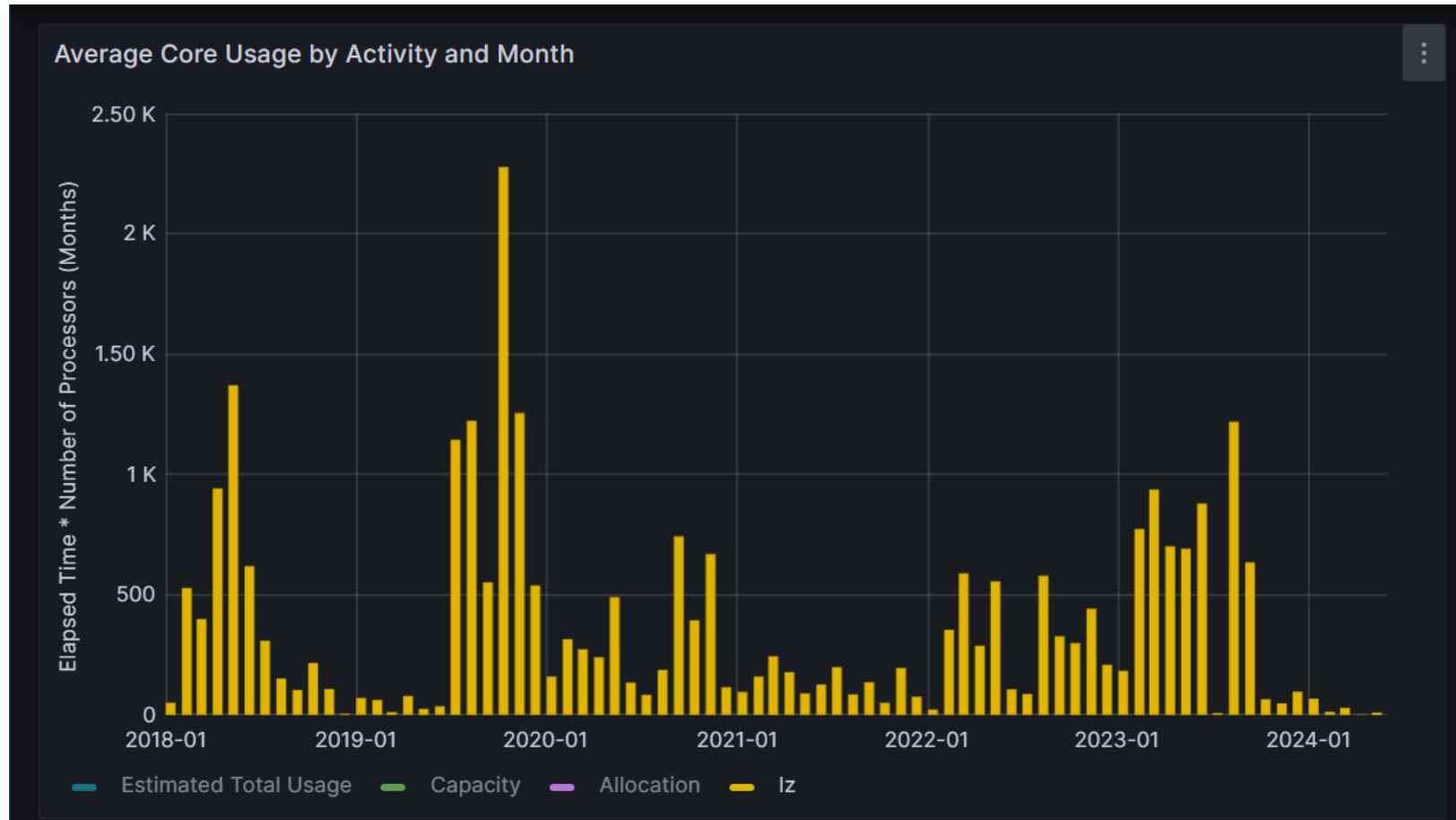


Allocation does not correspond to the GridPP( in this fig) users only, but for all. While capacity is correct (as much as it could be).

Commission: no functionality to do the calculation  
Omission: no relevant field yet introduced



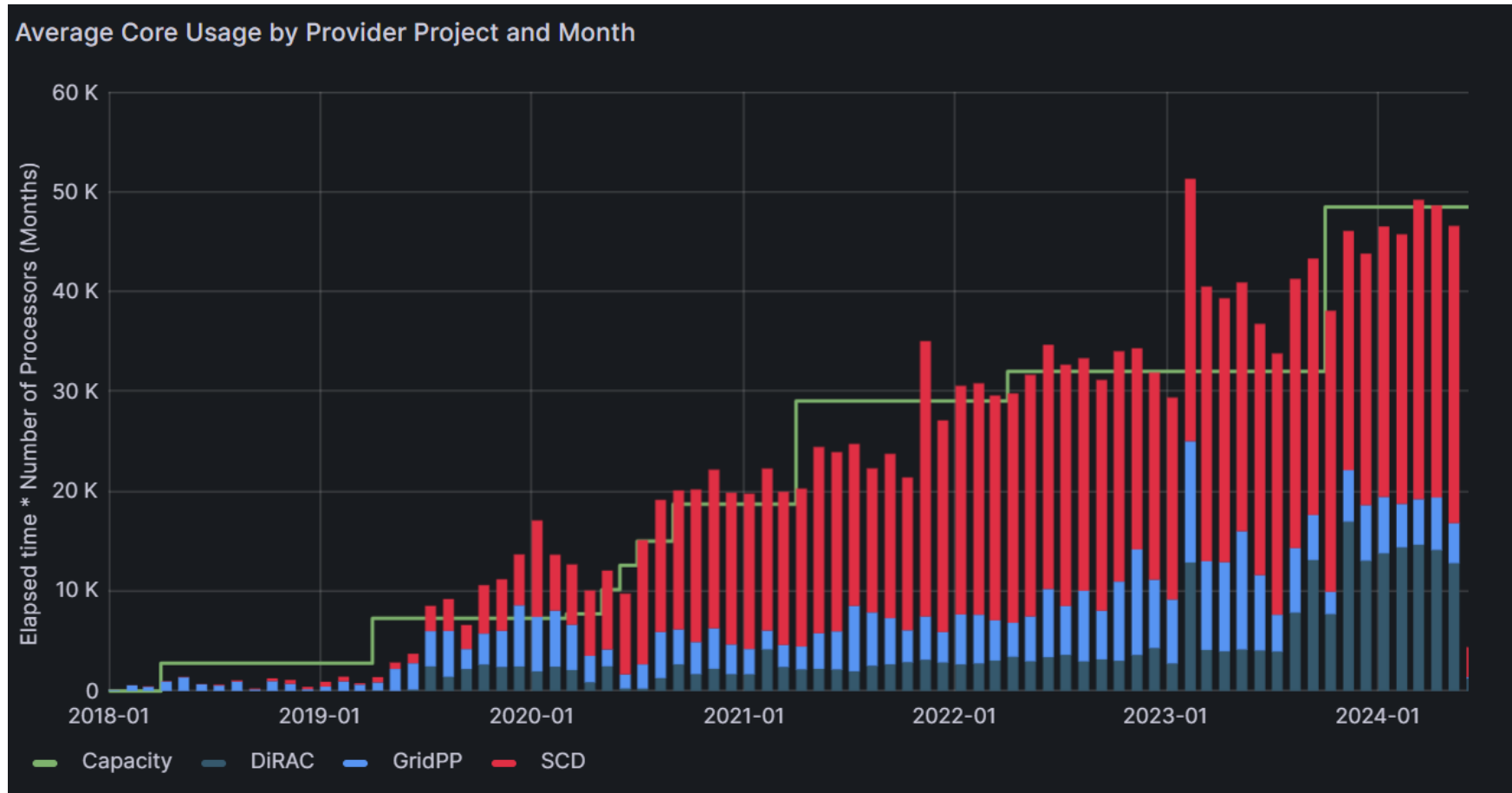
# Error of omission



Data seems reasonable, but this is only part of the full signal.

Not possible to know that this is not the full data without talking to the provider/user(in this case).

# And this case: Error of principle!



Physical cores  
(different cores,  
many generations,  
RAM varies from  
2GB/core to 20)



Virtual cores  
(hyperthreaded  
2:1)



Logical cores  
(hyperthreaded  
and not , so  
1.5ish:1)



Virtual and  
physical (some  
2:1, other 1:1)





Science and  
Technology  
Facilities Council

**These were just some common cases, but these are not inclusive**



# Some more historic examples

## Confused terminology

- Tebibytes vs terabytes
- Raw disk vs Fillable vs Usable
- Occupied size vs Used size in tape storage

## Data mismatch of different monitoring system for the same hardware: what is considered real?

- S3 data records for broken writing process
- Data deletion residues

## ... and the unexpected

- Historic data change

# Possible prevention methods

## General guidance

- Watch for external dependencies (e.g. alerts on public endpoints)
- Make reliable monitoring system out of many unreliable components: create dashboards with redundancies!
- Watch for anomalies!

Important, but would not prevent many of our errors

Great idea, but useless without regular auditing

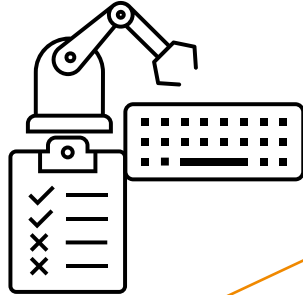
Another wonderful idea, but wouldn't have prevented errors above

# We already follow recommendations

Error detection and compensating error identification



Automated data entry and validation



Integration with accounting standards



Data centralisation



Real-time reporting and analytics

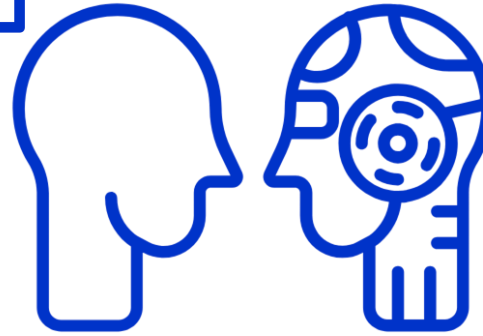
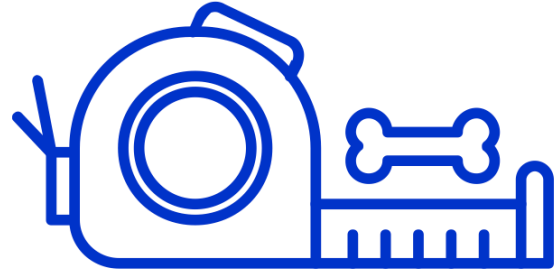


Enhanced audit trails

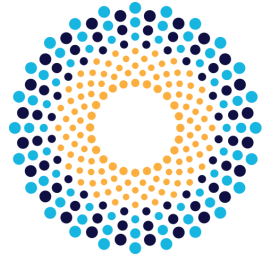


# For reflection

- Units
- Communication
- Audit
- Documentation



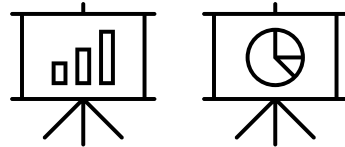
# And this utopian idea might become a reality!



Start



IRIS  
accounting  
dashboard



Hardware  
accounting  
report



In two week time







Science and  
Technology  
Facilities Council

# Thank you

*Special thanks to*

Adrian Coveney  
Rob Appleyard  
Andrew Sansum  
Rose Cooper

Hope to leave you  
with the sense of  
doom and desire  
to audit your  
monitoring ASAP



Science and Technology Facilities Council

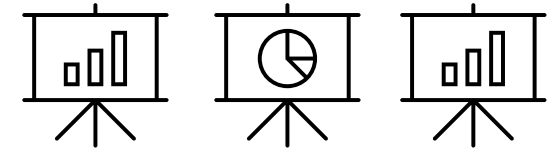
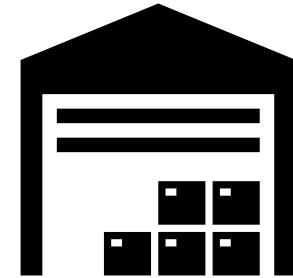
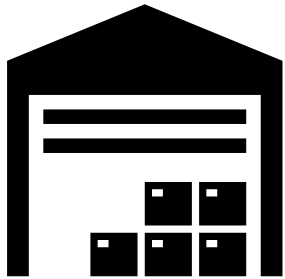


@STFC\_matters

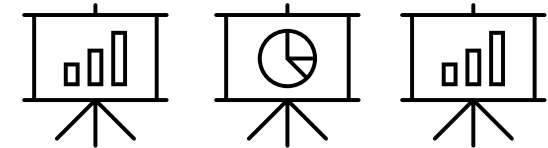
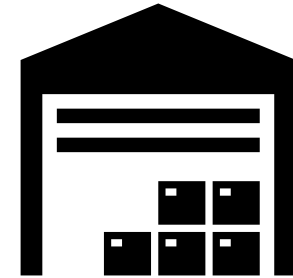


Science and Technology Facilities Council

# Accounting or no accounting?



# Accounting or no accounting?

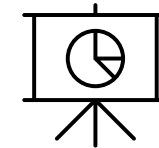
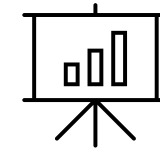


Hint: the correct response –  
Accounting!  
As much as possible!

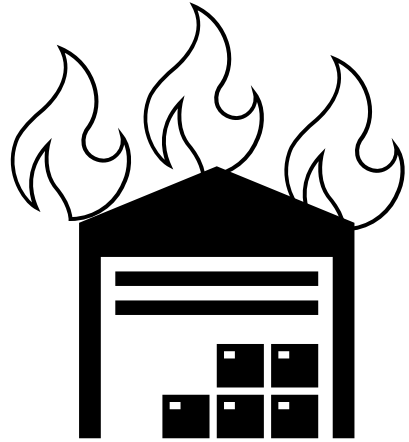
# We are following advised practices, but



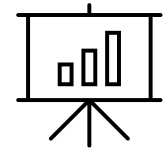
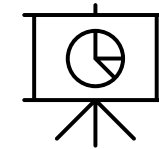
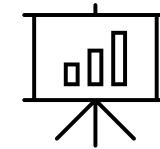
## accounting is hard!



# We follow advised practices, but



## accounting is hard!



# What we can do

- If you don't audit your data regularly, it might lose reliability
- Follow general guidance
- External audit
- Make it easy for users to report errors
- Have a proper easily accessible documentation
- Plan communication between parties