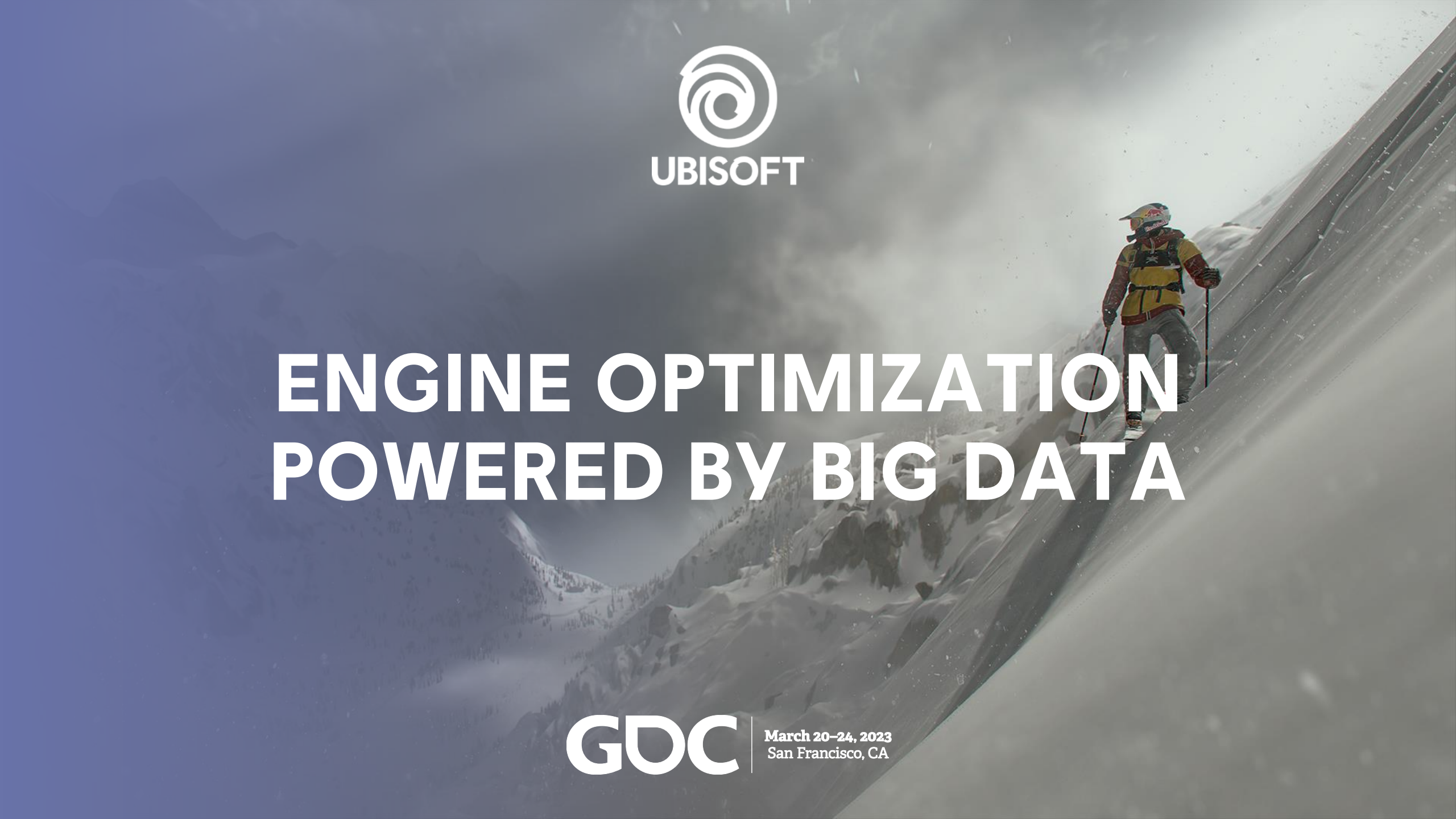

UBISOFT
ENGINE OPTIMIZATION
POWERED BY BIG DATA
GDC
March 20–24, 2023
San Francisco, CA

HELLO GDC!

VINCENT LECLERCQ

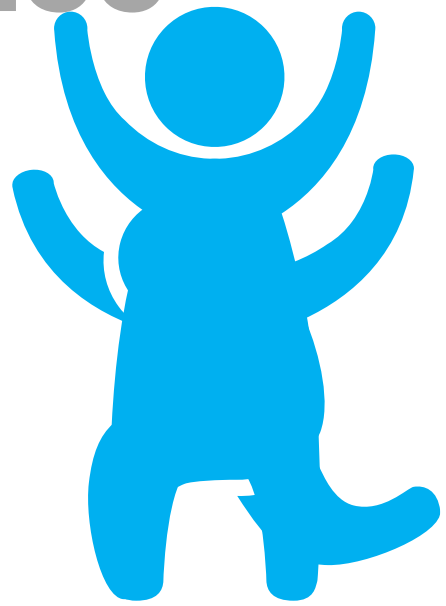
Lead Programmer
Ubisoft Paris

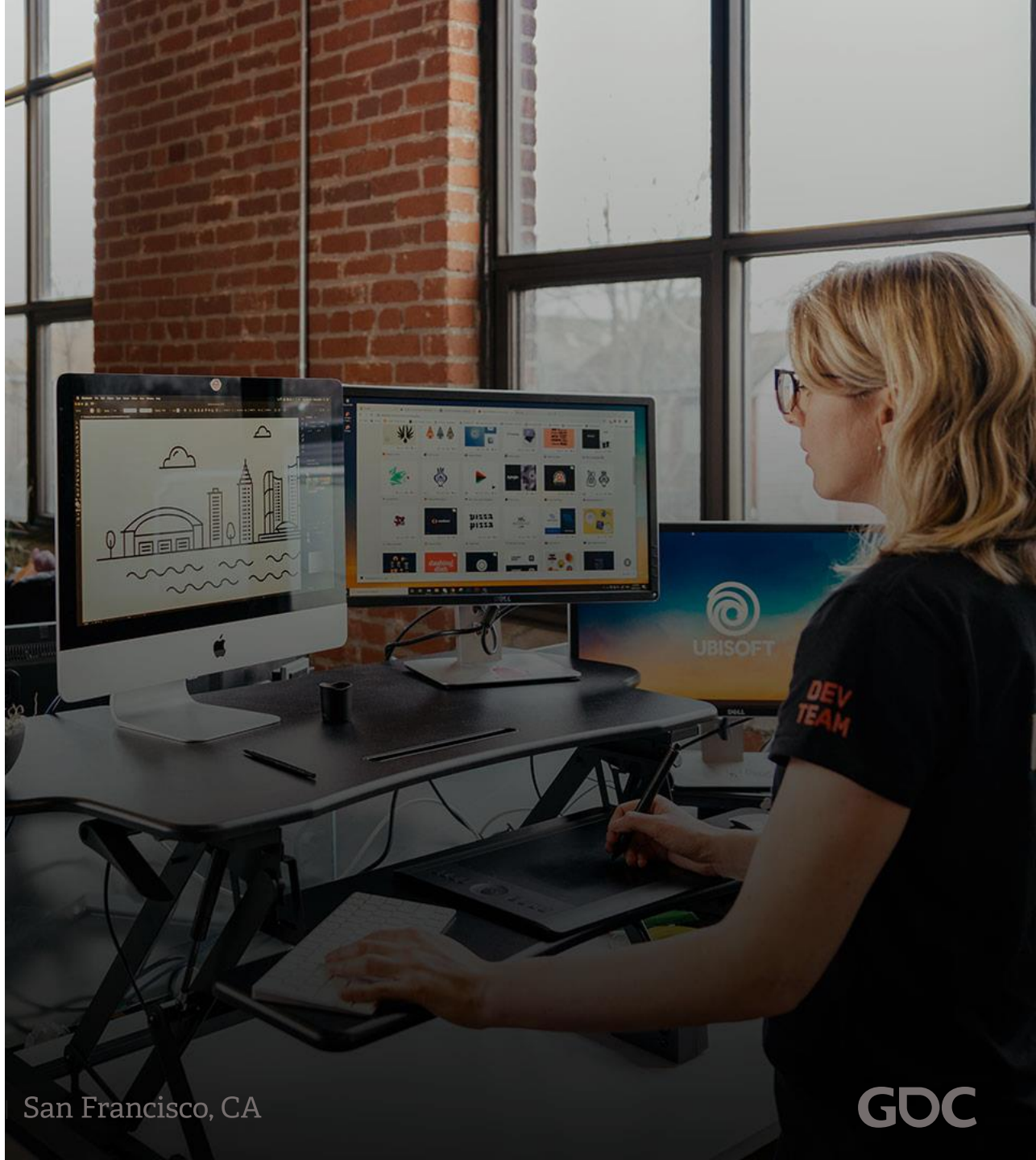
VLADIMIR KATCHADOURIAN

Senior Data Scientist
Ubisoft Paris

SHIPPING A GAME

DOWNGRADES
CUTS
PERFORMANCE
ISSUES
STRESS
BUGS



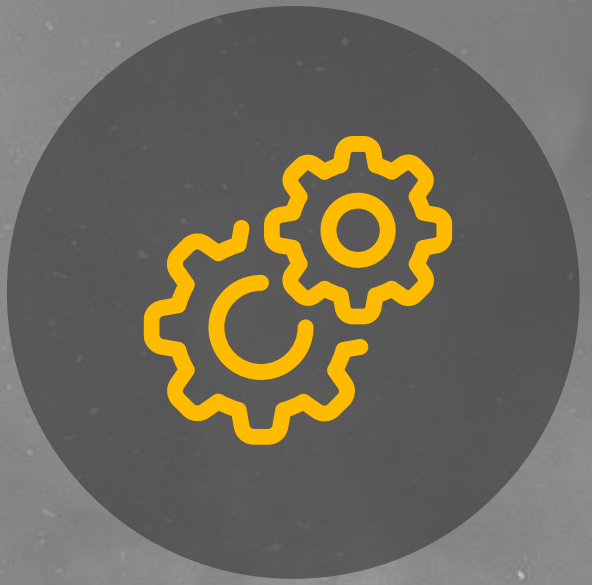


ANALYSIS ARE
(MOSTLY)
HANDMADE BY A FEW
(VERY BUSY) EXPERTS
IT SHOULD BE A SHARED RESPONSABILITY
BETWEEN THE TEAM MEMBERS
March 20–24, 2023 | San Francisco, CA
GDC

DATA ANALYTICS AS A SOLUTION ?

DATA ANALYSIS

MACHINE LEARNING MODEL



INVESTIGATION PROCESSES

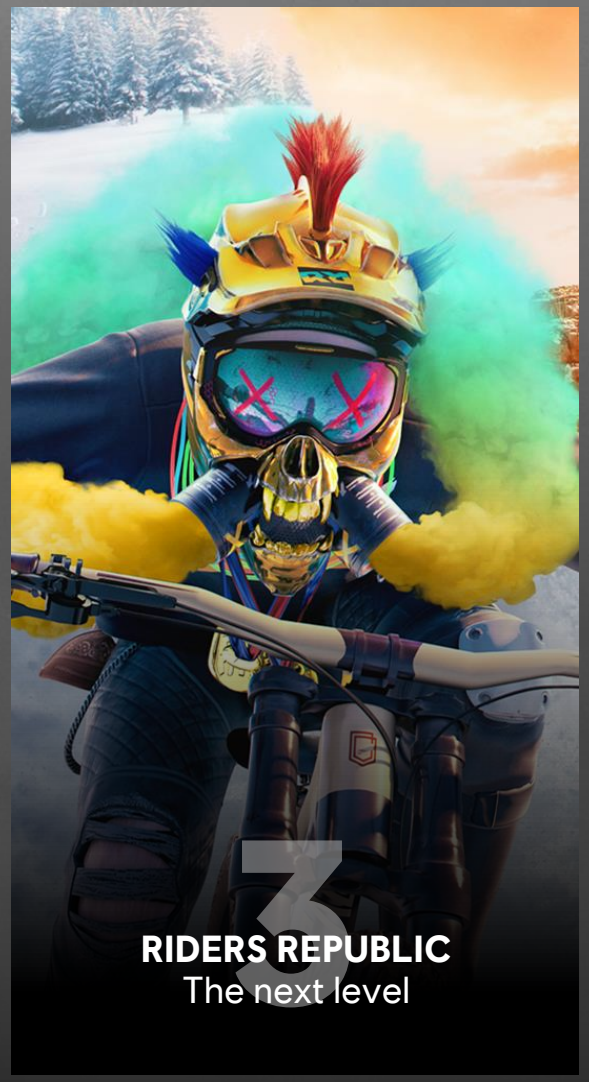
SUMMARY

1

THE DIVISION 2
The premises

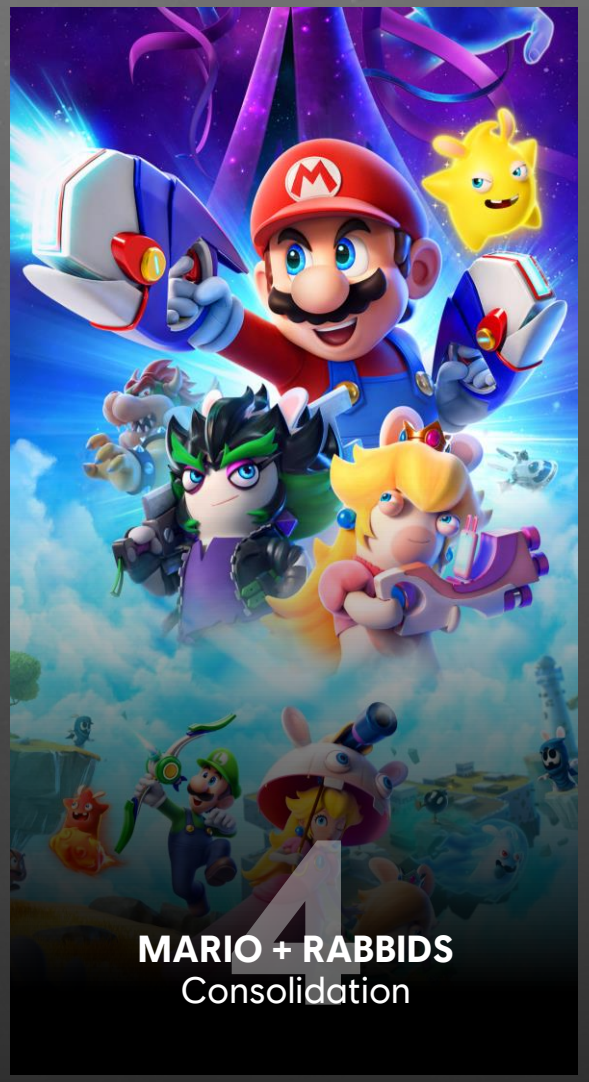
2

GHOST RECON BREAKPOINT
The first tool



3

RIDERS REPUBLIC
The next level



4

MARIO + RABBIDS
Consolidation

THE DIVISION 2

THE BEGINNING OF PERFORMANCE DATA ANALYSIS

GDC

THE DIVISION 2 : CONTEXT

Multiplayer Action RPG

Open World

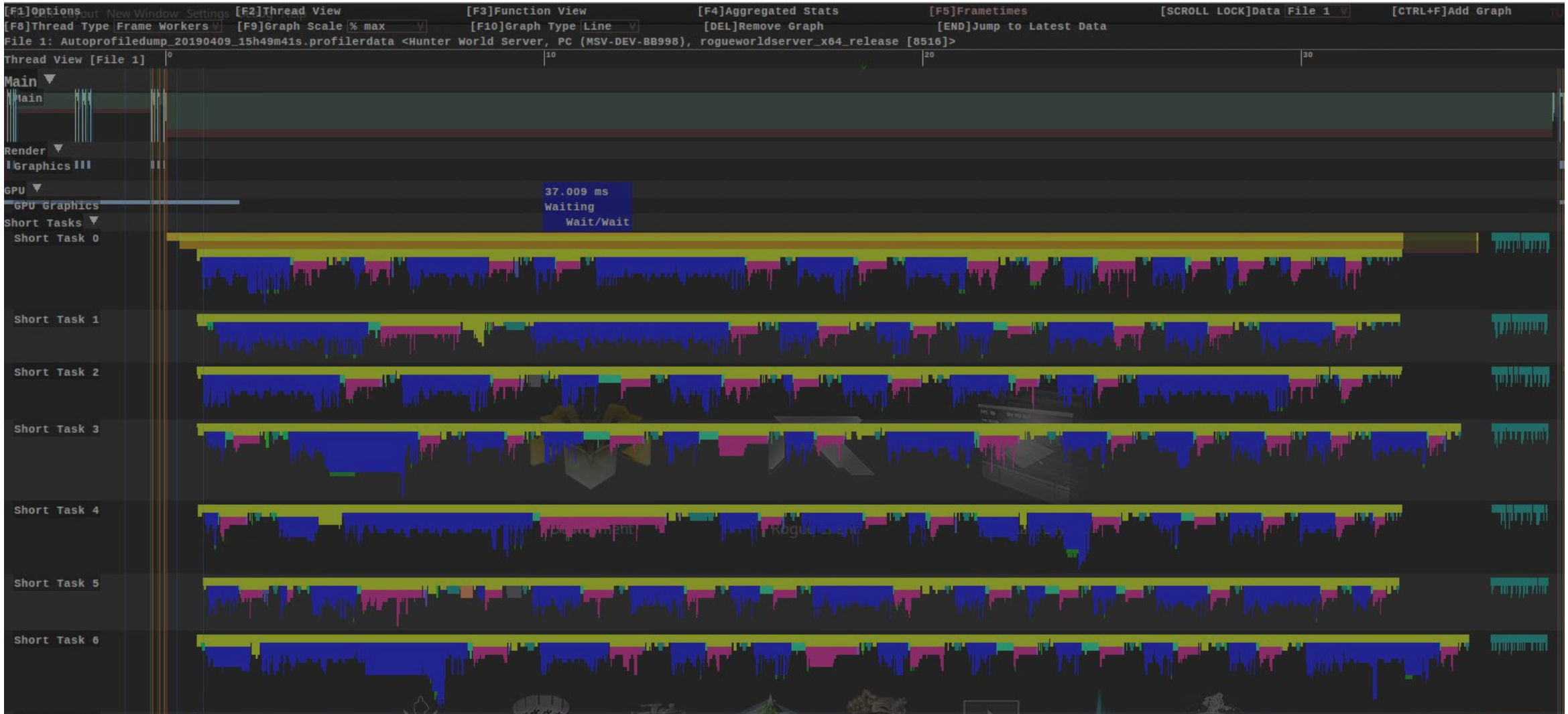
PVP, PVE, PVPVE

TOO HIGH-LEVEL INFORMATION

SERVER MONITORING

TOO LOW-LEVEL INFORMATION

PROFILER DATA



LIFE OF AN EXPERT

1 **FIND** a suspicious function

2 **GATHER** more information

3 **COMPARE** & dig the code

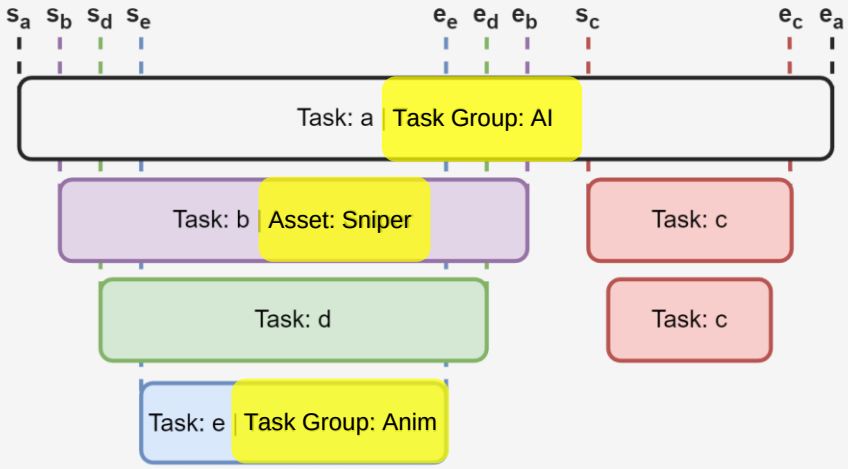
4 **ISSUE IDENTIFIED?**

YES

NO

IS THERE A
BETTER WAY?
A DATA ANALYTICS SOLUTION
BETWEEN HIGH & LOW LEVEL
March 20–24, 2023 | San Francisco, CA
GDC

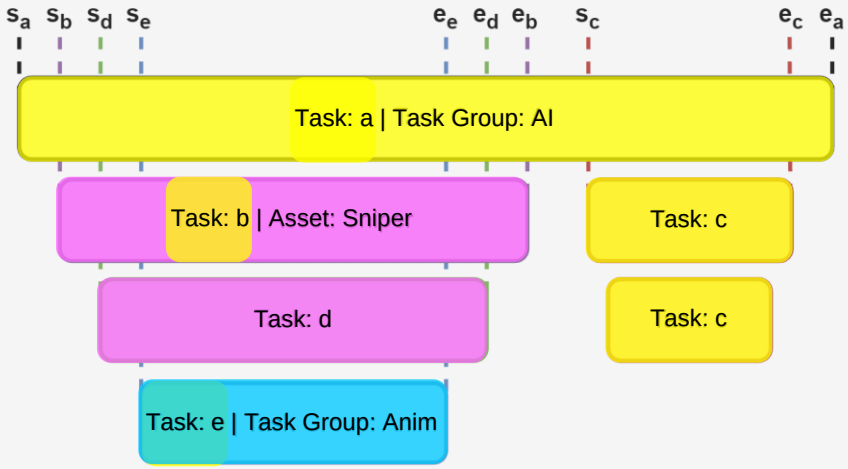
FROM PERFORMANCE PROFILING FILE…



…TO ANALYTICAL TABULAR DATA

Task ID	Task Name	Task Group	Asset Name	First task name	Path	Is_first_task group	Is_first_asset	n_recursions	Root Task
1	a	AI		a	a	True	False	0	a
2	b	AI	Sniper	b	b‖d	False	True	0	a
3	c	AI		a	a‖c	False	False	1	a
4	d	AI	Sniper	b	b	False	False	0	a
5	e	Anim	Sniper	e	e	True	False	0	a

FROM PERFORMANCE PROFILING FILE…

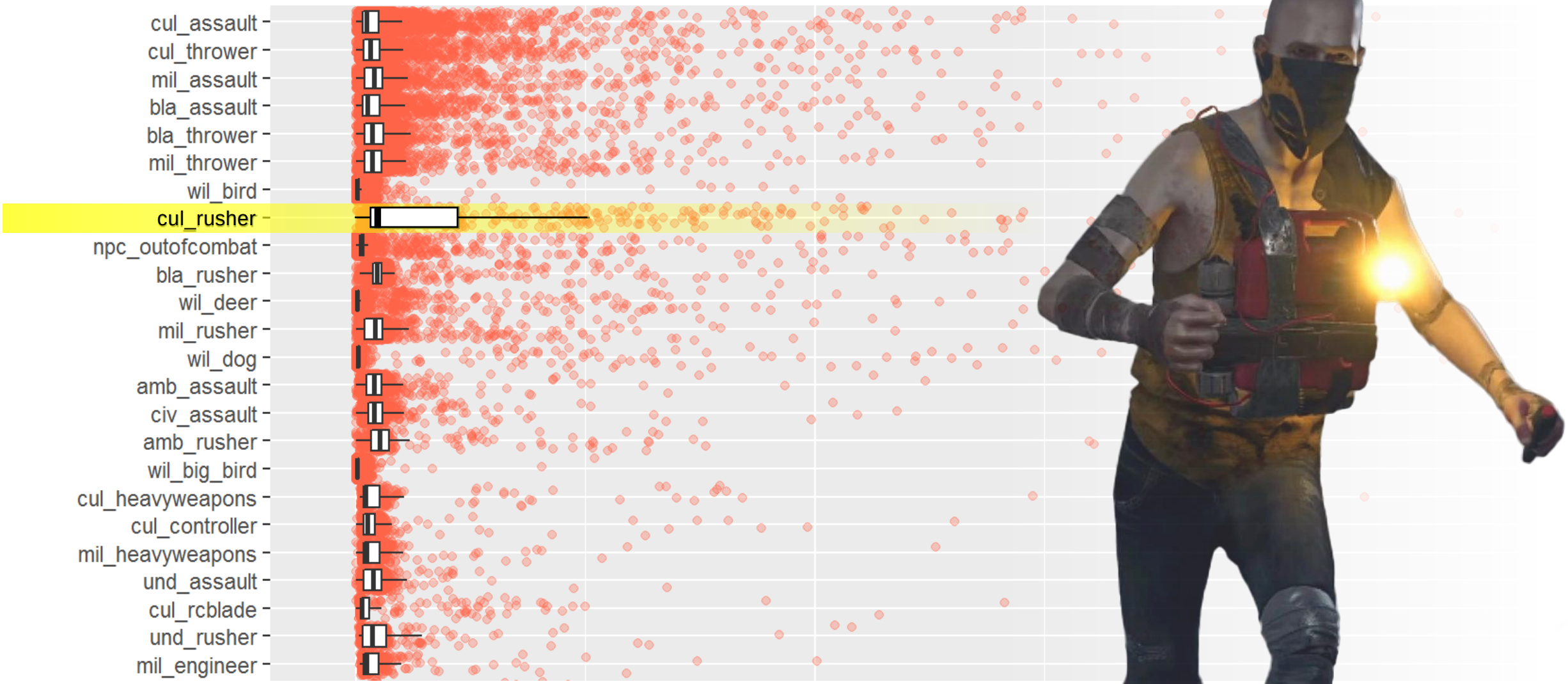


…TO ANALYTICAL TABULAR DATA

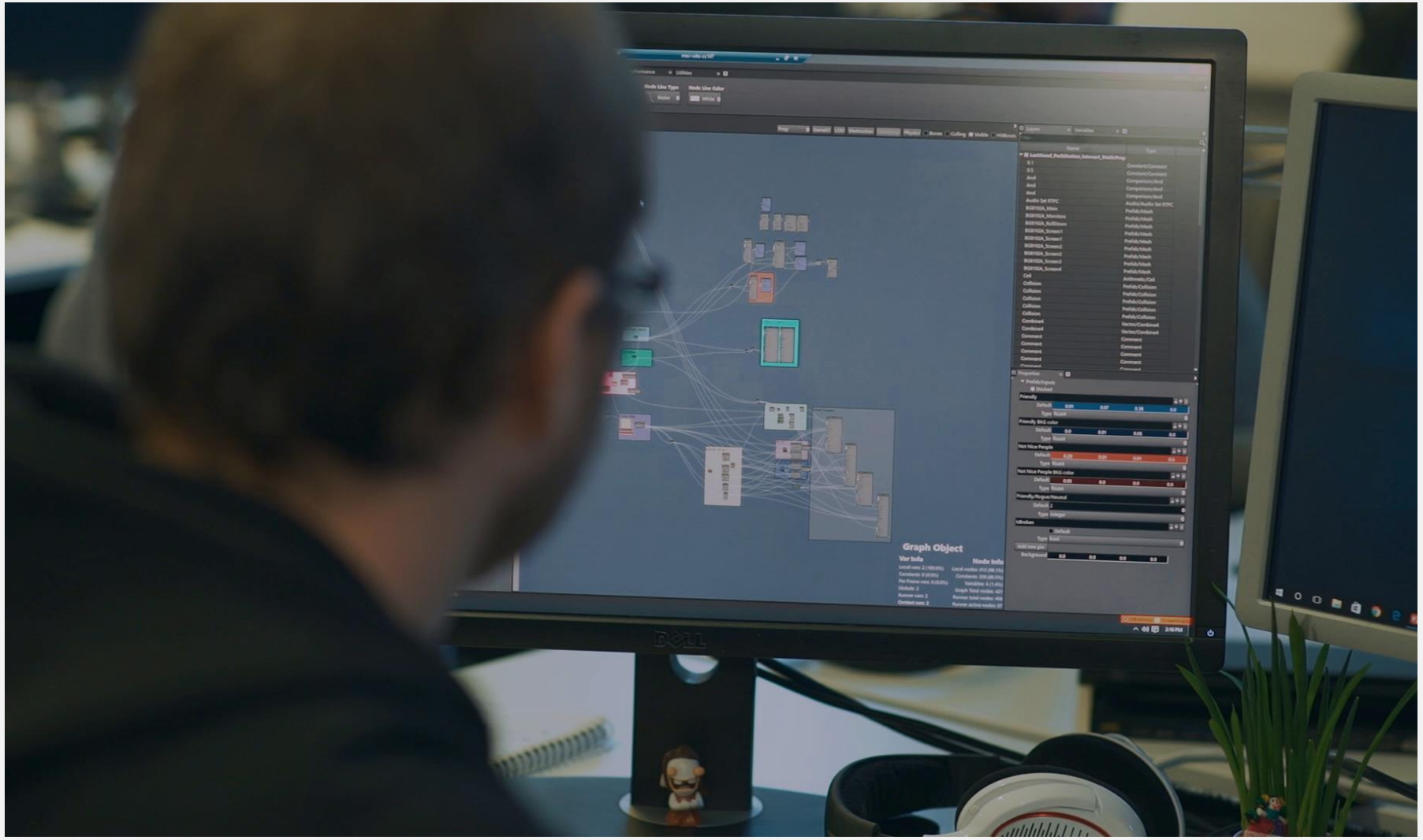
Task ID	Task Name	Task Group	Asset Name	First task name	Path	Is_first_task group	Is_first_asset	n_recursions	Root Task
1	a	AI		a	a	True	False	0	a
2	b	AI	Sniper	b	b‖d	False	True	0	a
3	c	AI		a	a‖c	False	False	1	a
4	d	AI	Sniper	b	b	False	False	0	a
5	e	Anim	Sniper	e	e	True	False	0	a

CULTIST RUSHER ASSET ISSUE DETECTION

Behavior Update Duration per Asset



THE DIVISION 2: CONCLUSION



We Still need
an expert but
it's promising

Objective:

900

PLAYERS
PER SERVER

GHOST RECON BREAKPOINT

INDUSTRIALIZATION & PERFORMANCE ANALYTICS PIPELINE

From prototype to new performance analytics tool

GDC

GHOST RECON : CONTEXT

NEXT STEPS

1

AUTOMATIZE REPORTING

2

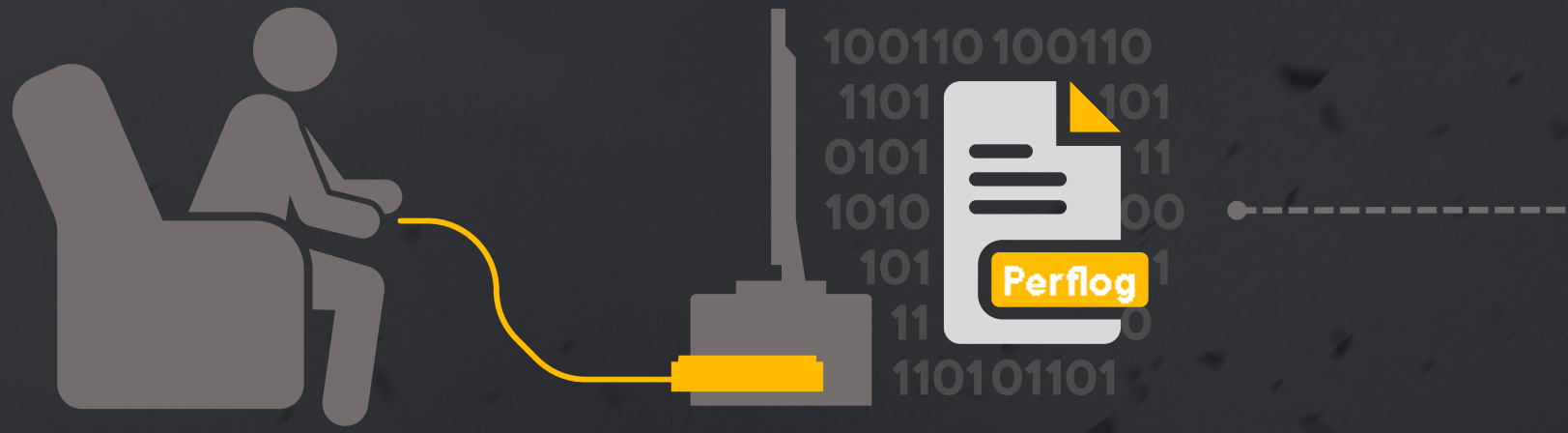
DEMOCRATIZE INVESTIGATION

LET'S BUILD A PIPELINE!

THE PIPELINE

GENERATE DATA

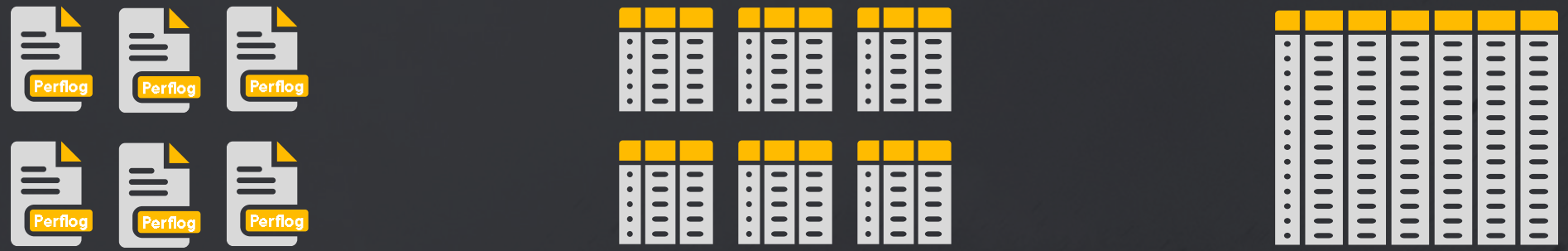
Daily Auto-tests
Daily Manual tests
Playtest sessions



Continuous profiling

Record 1 sec every 20 seconds
Add files when slow frames are detected

DATA PROCESSING



File level transform

- Deserialization
- Inheritance
- Frame-wide aggregation

Session level transform

- Session-wide aggregation

Analytics Database

Exploration,
Iteration, R&D

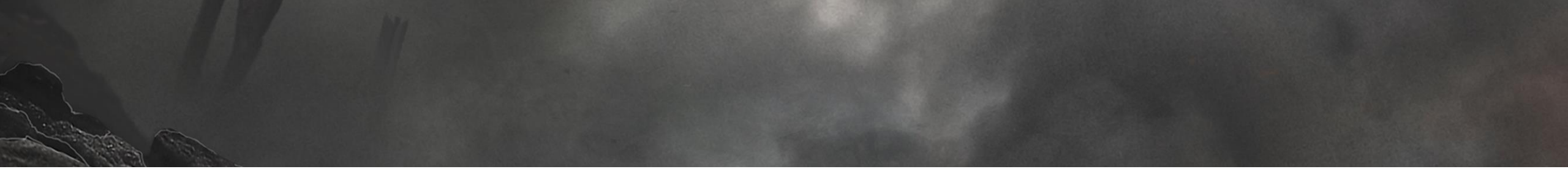
Output is saved in 2 formats

Applicative Database

Data services
& exposition

ANALYZE & VISUALIZE

- R&D, iteration, exploration
- Production agnostic applicative usage



AD-HOC ANALYSIS

TO ANSWER SPECIFIC QUESTIONS

REGRESSION INVESTIGATION

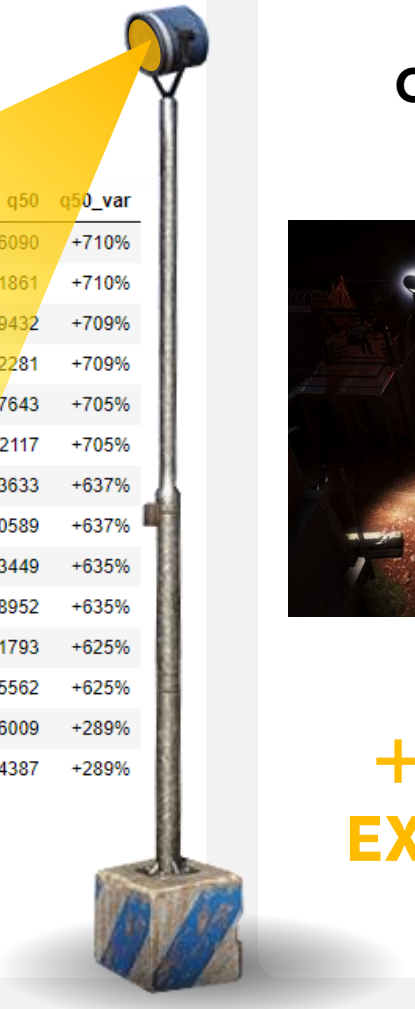
Ad-Hoc analysis example 1

REGRESSION SPOTTED AFTER VERSIONS MERGE

- Auto test detection
- Historical data available

INVESTIGATION ON DATA SYSTEMS

asset_name	q50	q50_var
TGT_SearchLight_BaseEntity	96090	+710%
TGT_SearchLight_BaseEntity	11861	+710%
TGT_SearchLight_BaseEntity	99432	+709%
TGT_SearchLight_BaseEntity	12281	+709%
TGT_SearchLight_BaseEntity_NoPole_Vertical	97643	+705%
TGT_SearchLight_BaseEntity_NoPole_Vertical	12117	+705%
TGT_SearchLight_BaseEntity	13633	+637%
TGT_SearchLight_BaseEntity	100589	+637%
TGT_SearchLight_BaseEntity_NoPole_Vertical	13449	+635%
TGT_SearchLight_BaseEntity_NoPole_Vertical	98952	+635%
TGT_SearchLight_BaseEntity_NoPole_Vertical	11793	+625%
TGT_SearchLight_BaseEntity_NoPole_Vertical	85562	+625%
TGT_SearchLight_BaseEntity	56009	+289%
TGT_SearchLight_BaseEntity	14387	+289%



OFFENDING ASSET IDENTIFIED

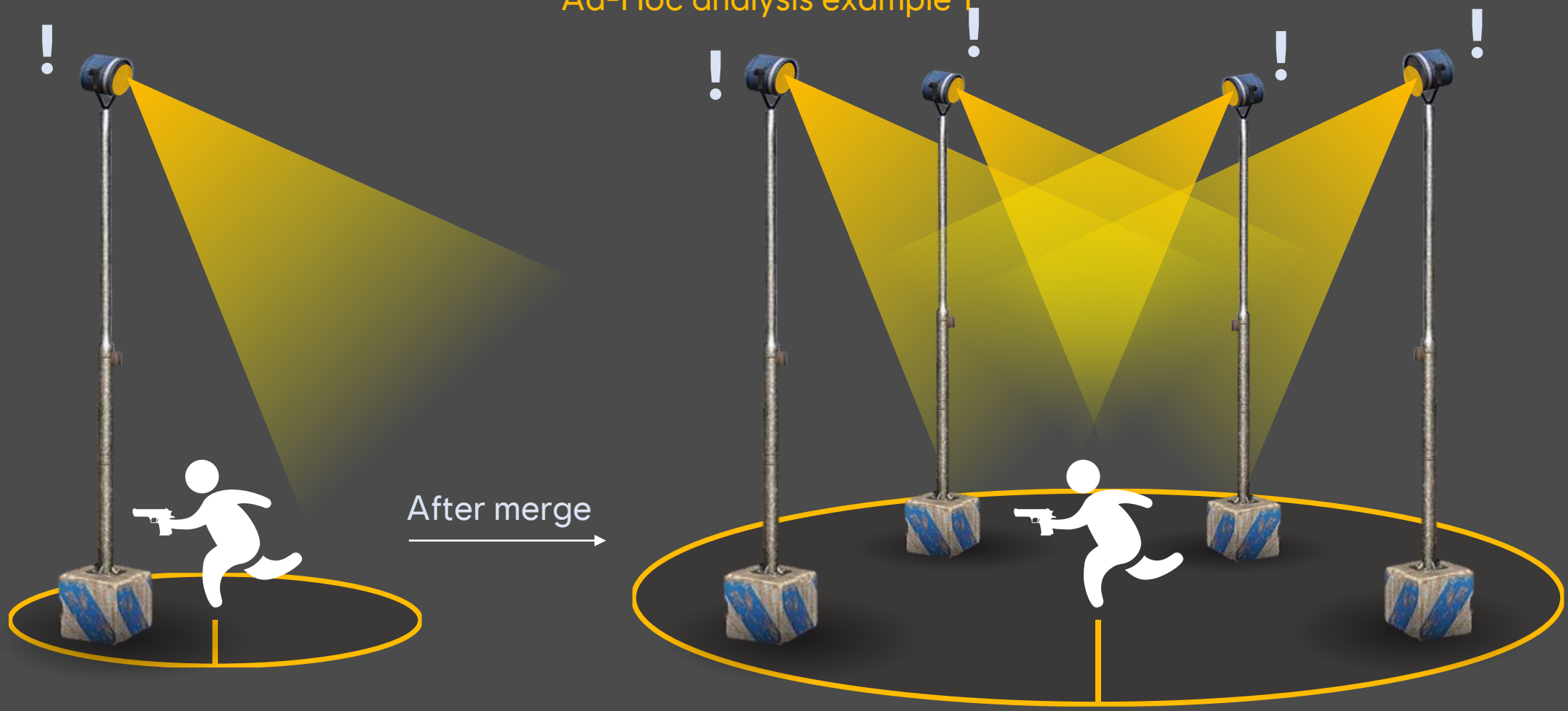
+700%

EXECUTION TIME

F(Search_Light)

REGRESSION INVESTIGATION

Ad-Hoc analysis example 1

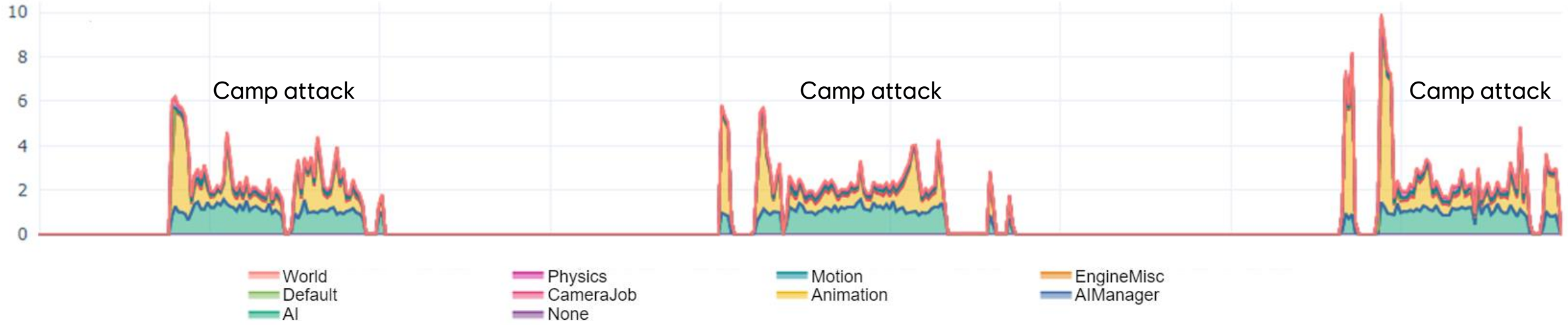


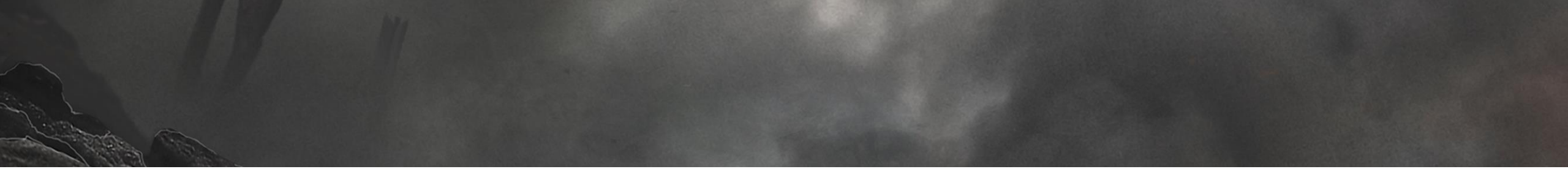
ANTICIPATE NEW FEATURE COST

Ad-Hoc analysis example 2

AI TEAMMATES

Estimated asset's CPU cost
& FPS impact

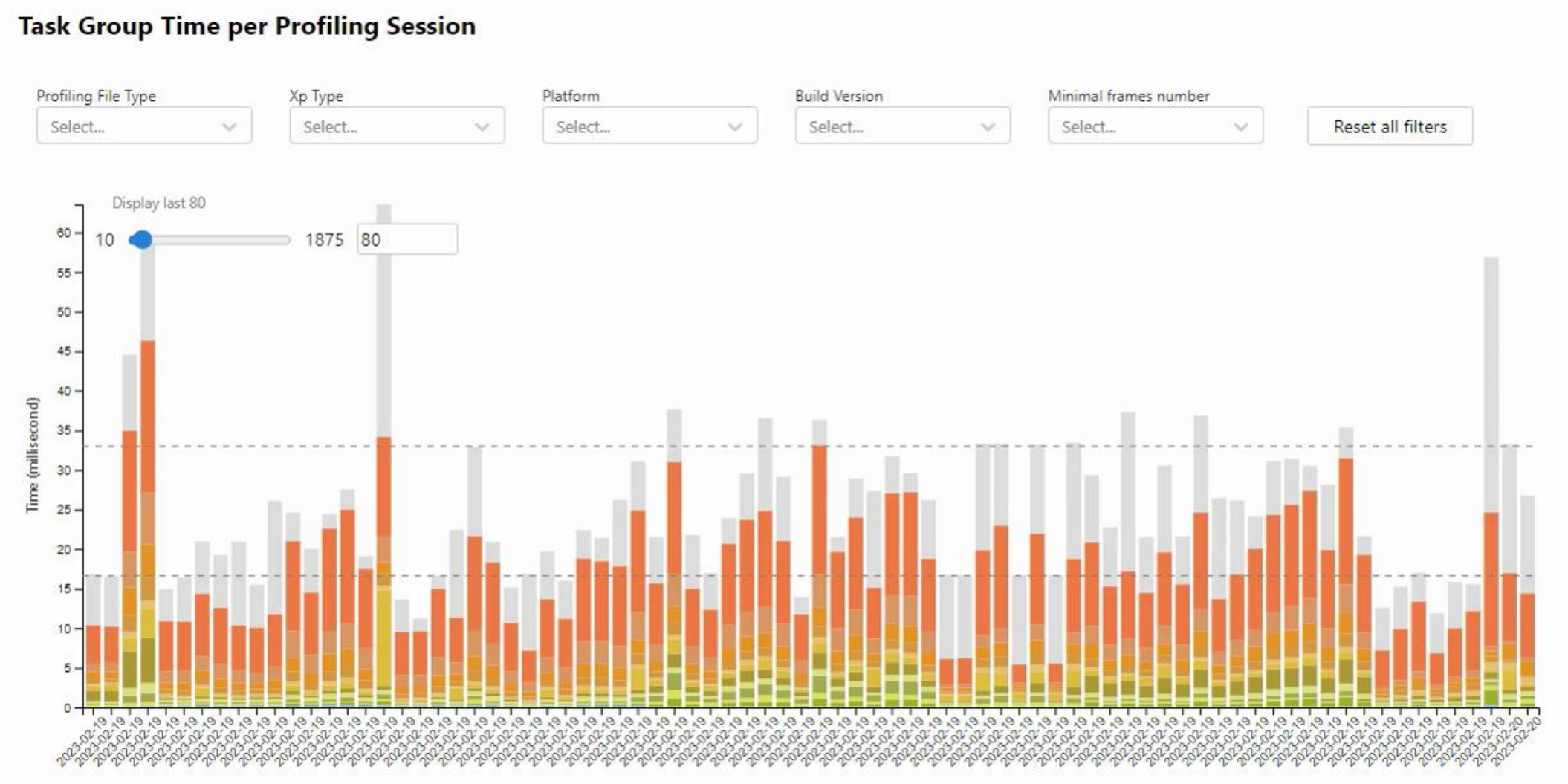




STANDARD DASHBOARDS

TO MUTUALIZE THE BENEFITS.

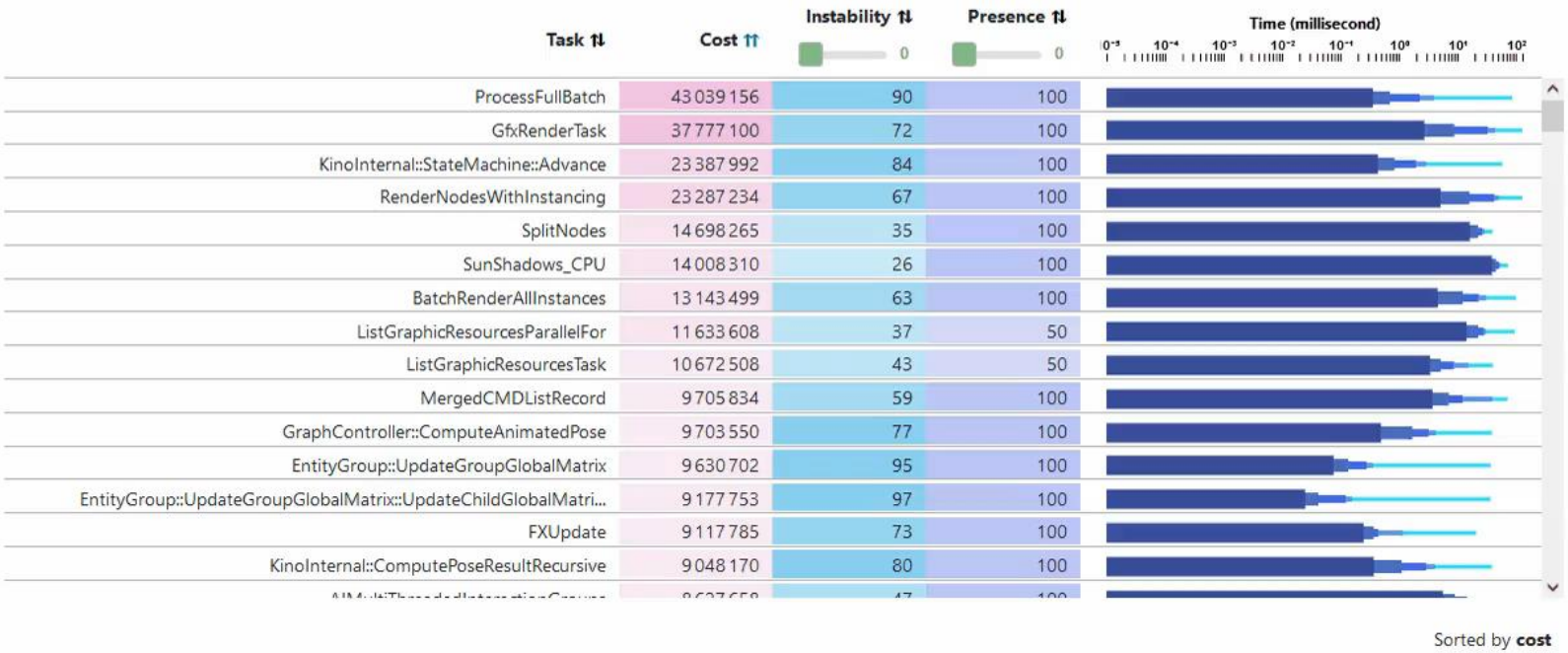
SELECT A SESSION



ANALYZE A SESSION

STANDARD DASHBOARD

ANALYSE SESSION: TASKS



Task	Cost	Instability	Presence	Time (millisecond)
ProcessFullBatch	43 039 156	90	100	
GfxRenderTask	37 777 100	72	100	
KinoInternal::StateMachine::Advance	23 387 992	84	100	
RenderNodesWithInstancing	23 287 234	67	100	
SplitNodes	14 698 265	35	100	
SunShadows_CPU	14 008 310	26	100	
BatchRenderAllInstances	13 143 499	63	100	
ListGraphicResourcesParallelFor	11 633 608	37	50	
ListGraphicResourcesTask	10 672 508	43	50	
MergedCMDListRecord	9 705 834	59	100	
GraphController::ComputeAnimatedPose	9 703 550	77	100	
EntityGroup::UpdateGroupGlobalMatrix	9 630 702	95	100	
EntityGroup::UpdateGroupGlobalMatrix::UpdateChildGlobalMatri...	9 177 753	97	100	
FXUpdate	9 117 785	73	100	
KinoInternal::ComputePoseResultRecursive	9 048 170	80	100	

Sorted by **cost**

STANDARD DASHBOARD

ANALYSE SESSION: TASKS

Task	Cost	Instability	Presence	Time (millisecond)
ProcessFullBatch	43 039 156	90	100	
GfxRenderTask	37 777 100	72	100	
KinoInternal::StateMachine::Advance	23 387 992	84	100	
RenderNodesWithInstancing	23 287 234	67	100	
SplitNodes	14 698 265	35	100	
SunShadows_CPU	14 008 310	26	100	
BatchRenderAllInstances	13 143 499	63	100	
ListGraphicResourcesParallelFor	11 633 608	37	50	
ListGraphicResourcesTask	10 672 508	43	50	
MergedCMDListRecord	9 705 834	59	100	
GraphController::ComputeAnimatedPose	9 703 550	77	100	
EntityGroup::UpdateGroupGlobalMatrix	9 630 702	95	100	
EntityGroup::UpdateGroupGlobalMatrix::UpdateChildGlobalMatri...	9 177 753	97	100	
FXUpdate	9 117 785	73	100	
KinoInternal::ComputePoseResultRecursive	9 048 170	80	100	

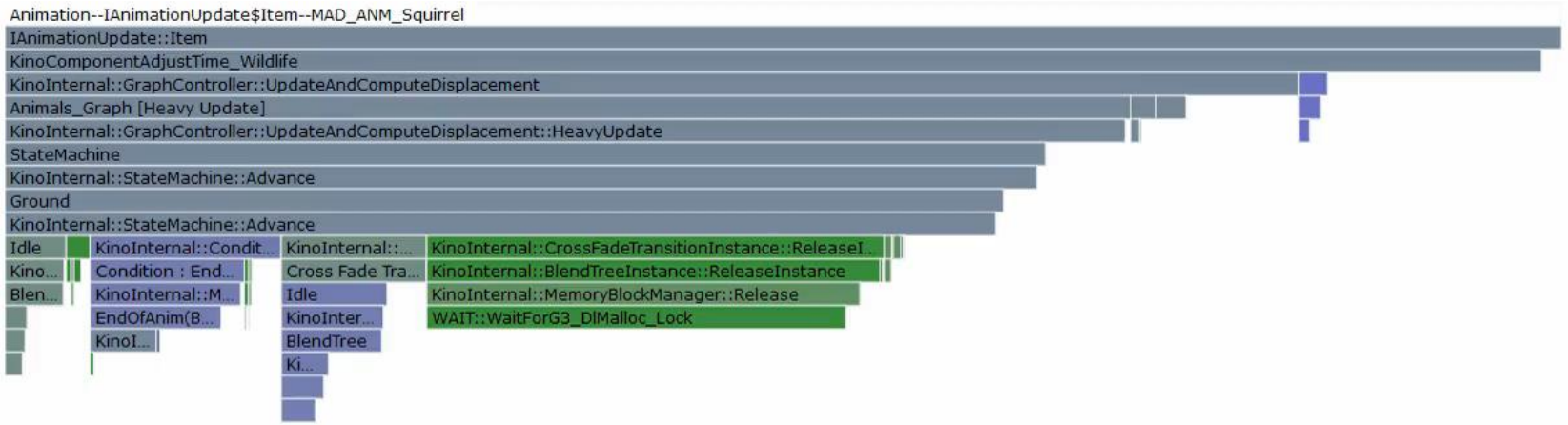
Sorted by **cost**

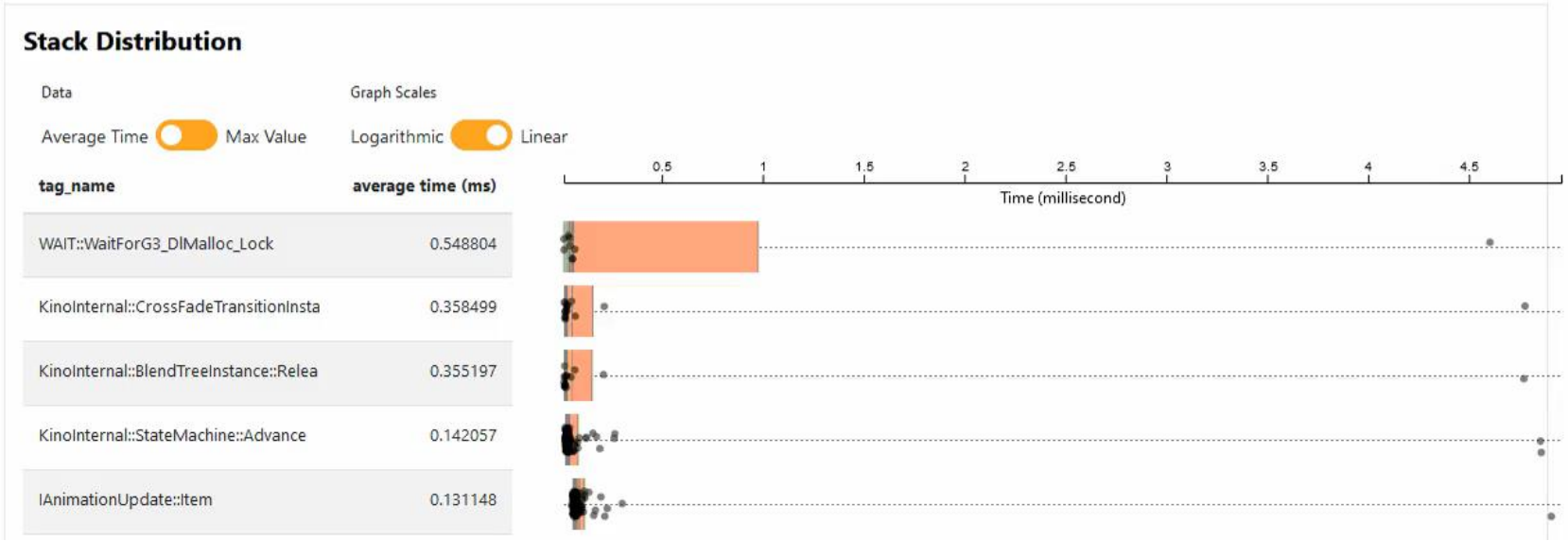
Task time per system for **KinoInternal::StateMachine::Advance**

Task Group	First Task	Asset Name	Flamegraph	Cost	Instability	Presence	Time (millisecond)
Animation	IAnimationUp...	MAD_CHR_Te...		22 971 844	81	100	
Animation	IAnimationUp...	MAD_CHR_G...		608 793	39	100	
Animation	IAnimationUp...	MAD_ANM_B...		107 276	30	91	
Animation	IAnimationUp...	MAD_ANM_...		80 608	23	5	
Animation	IAnimationUp...	MAD_ANM_S...		78 975	81	41	
Animation	IAnimationUp...	MAD_ANM_B...		56 693	24	39	
Animation	IAnimationUp...	MAD_ANM_B...		48 786	26	58	
Animation	IAnimationUp...	MAD_ANM_B...		48 212	26	46	
Animation	IAnimationUp...	MAD_ANM_...		47 508	24	100	
Animation	IAnimationUp...	MAD_ANM_B...		43 394	23	95	
Animation	IAnimationUp...	MAD_CHR_N...		41 837	15	50	

STANDARD DASHBOARD

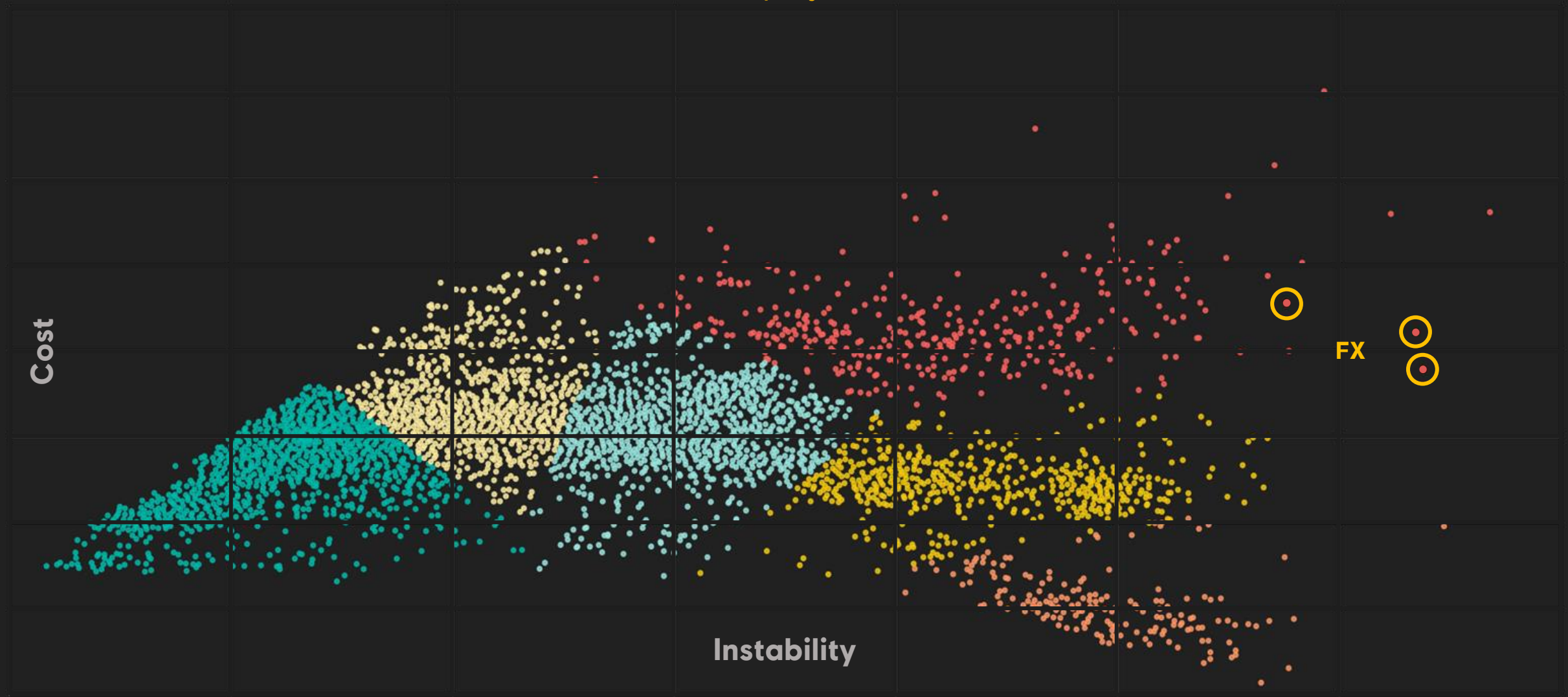
ANALYSE SESSION:
DATA SYSTEMS



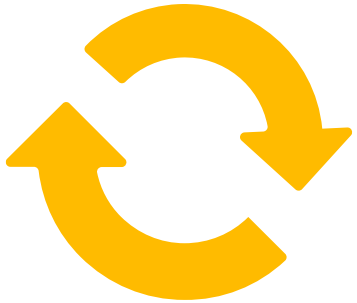


FX ISSUES

Assets projection



FX ISSUES SOLVED



100+

USELESS UPDATE REMOVED

25%

EXECUTION TIME SAVED

75%

BETTER STABILITY ACHIEVED

RIDERS REPUBLIC

MACHINE LEARNING USING CPU PROFILING DATA

How to save 50% of experts time using ML

GDC

RIDERS REPUBLIC : CONTEXT

Anvil engine /Scimitar

Open World

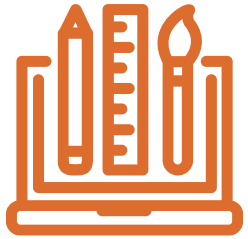
Easy transition

DASHBOARDS ARE GOOD FOR MONITORING:

Builds health

Engine behaviors

Tasks



Assets

BUT ARE LIMITED TO PRIORITIZE OPTIMISATION EFFORTS

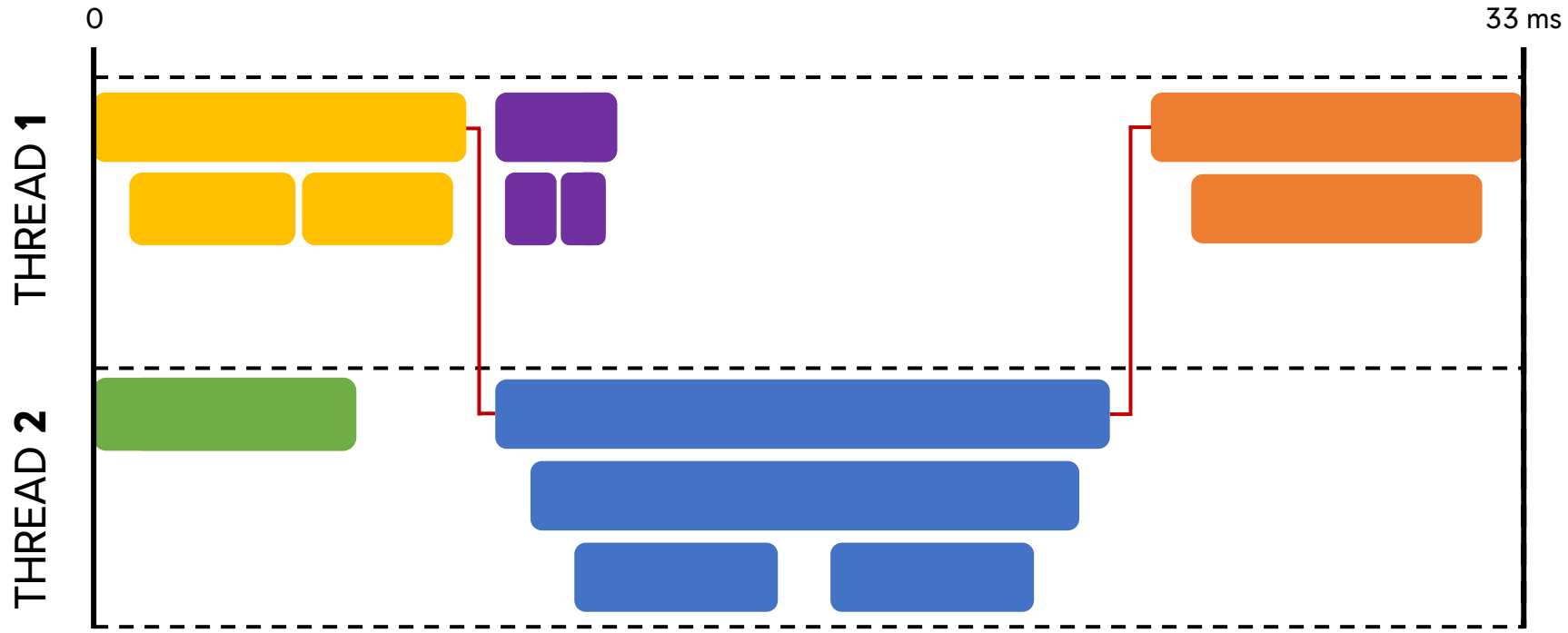
EXPERTS ARE BOTTLENECK

and overwhelmed

We need to ease
& automate part of their job

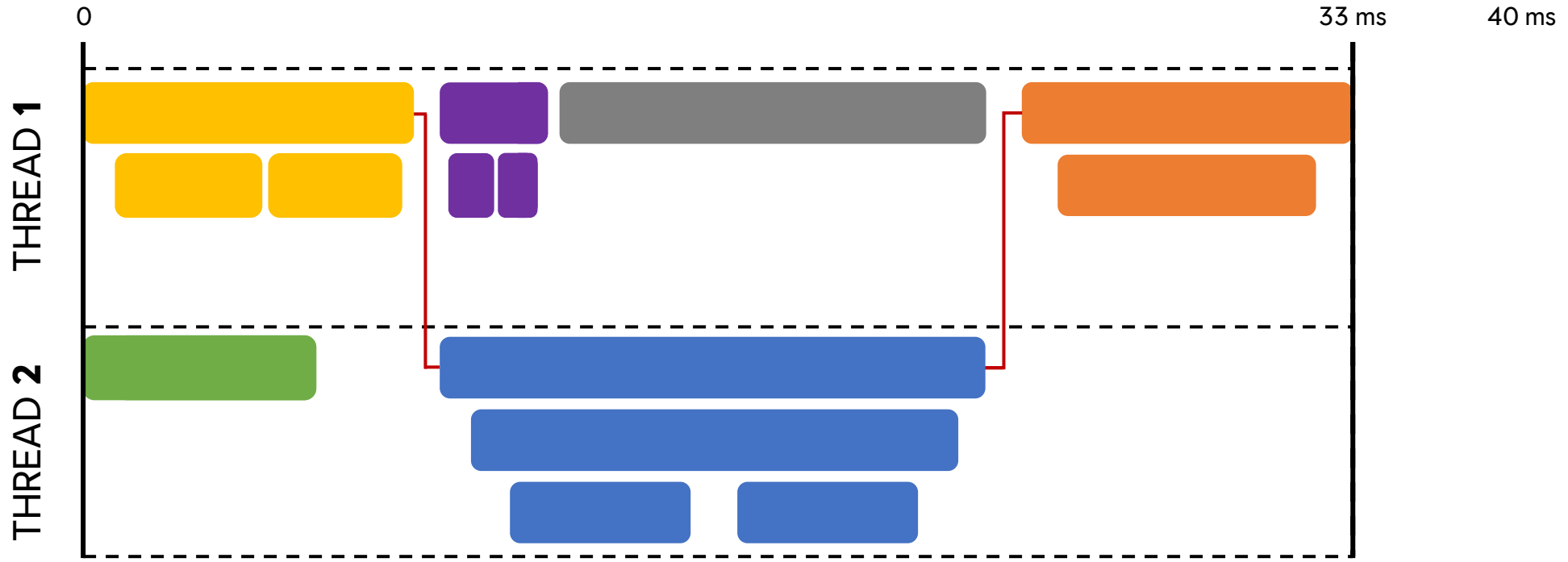
INSIDE A CPU FRAME

The problem



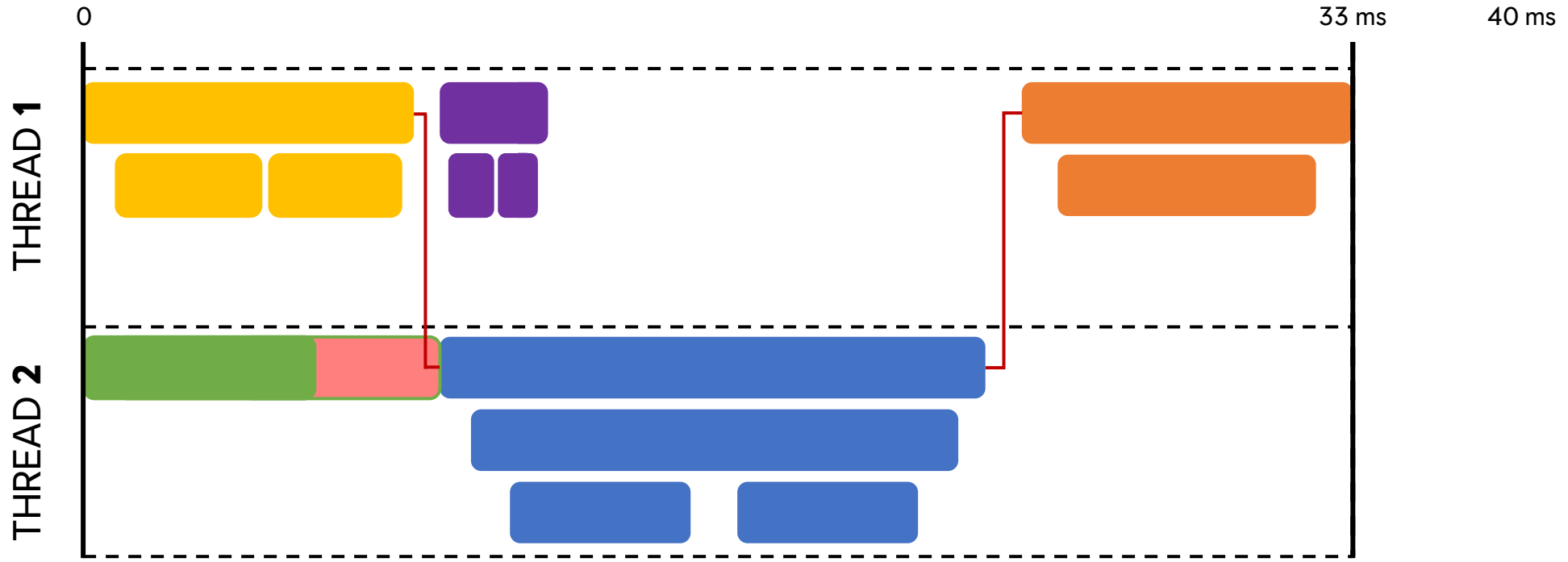
INSIDE A CPU FRAME

The problem



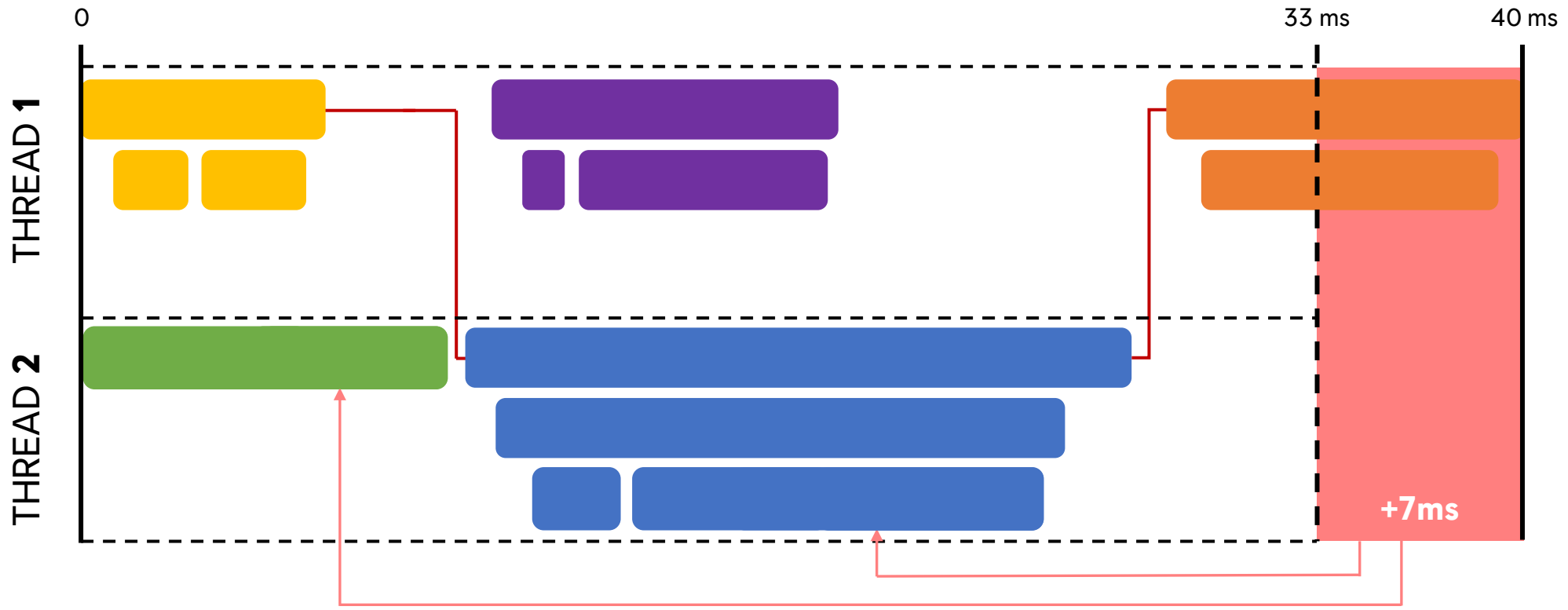
INSIDE A CPU FRAME

The problem

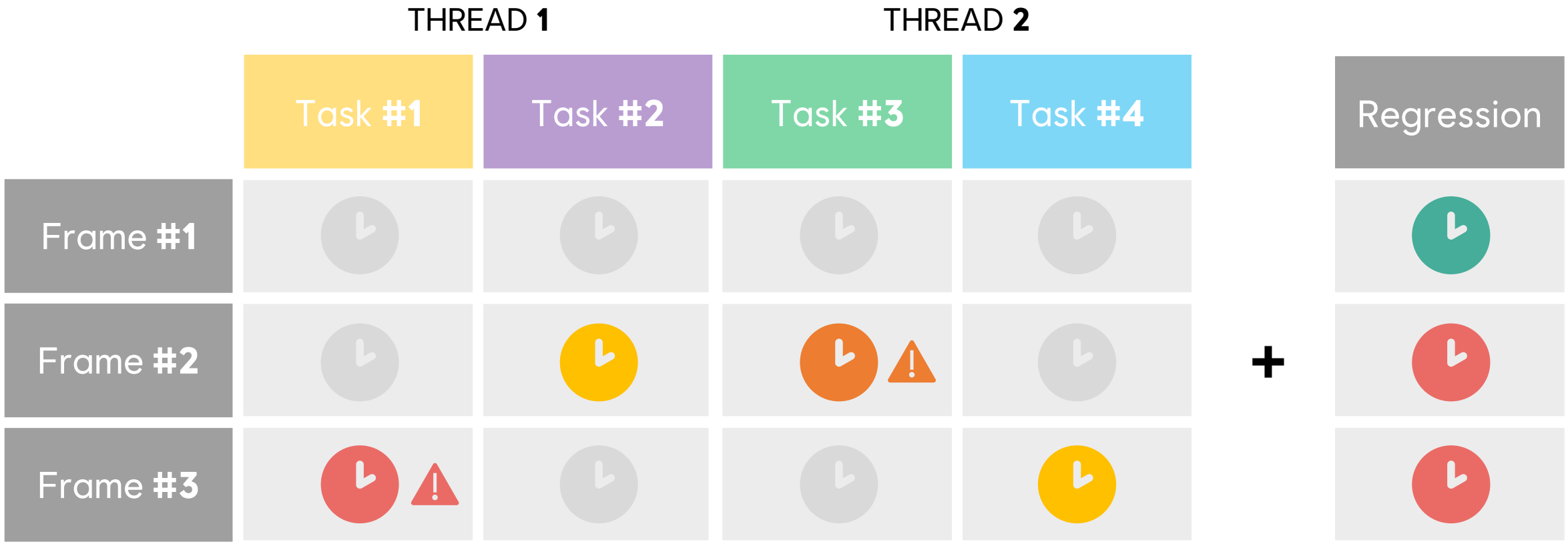


INSIDE A CPU FRAME

The objectives



FLATTEN VIEW OF THE PROBLEM



Estimate the **time impact** of guilty tasks for each frame

CONSTRAINTS

FRAME-WISE EXPLICATIVE

MEANINGFUL OUTPUTS

ROBUST & TRUSTWORTHY

CANDIDATE SOLUTIONS

Hand-made correlation metric

CANDIDATE SOLUTIONS

Hand-made correlation metric

It was a dead end

Causal profiling, Frame simulation

CANDIDATE SOLUTIONS

Hand-made correlation metric

It was a dead end

Causal profiling, Frame simulation

Lack resources & data

Machine learning XAI: eXplainable AI

METHODOLOGY

1

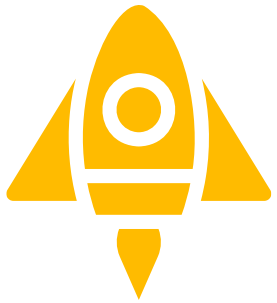
DEVELOP
different approaches

2

EVALUATE
which works better

3

DEPLOY
the best one

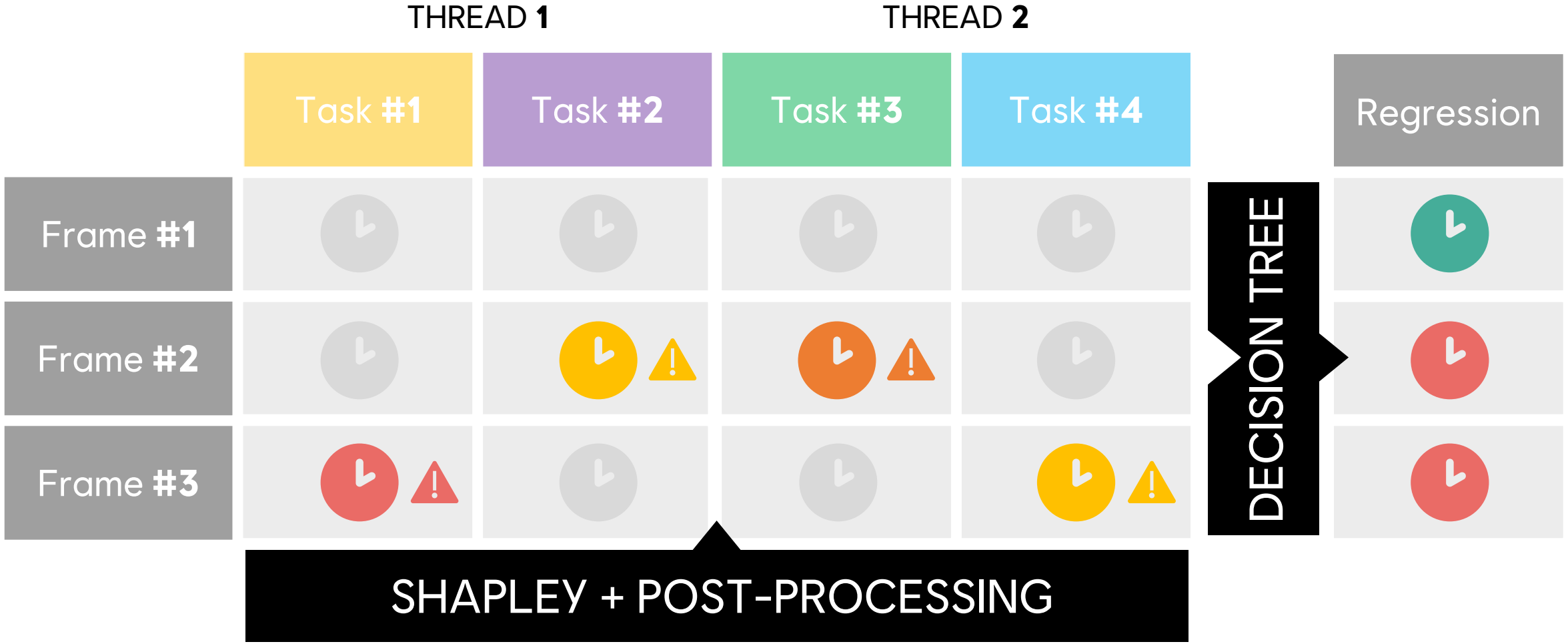


4

REPORTS
provided to team

DECISION TREES & SHAPLEY VALUES

Method 1



DECISION TREES & SHAPLEY VALUES

Method 1

+ CUMULATIVE

FRAME-WISE EXPLICATIVE

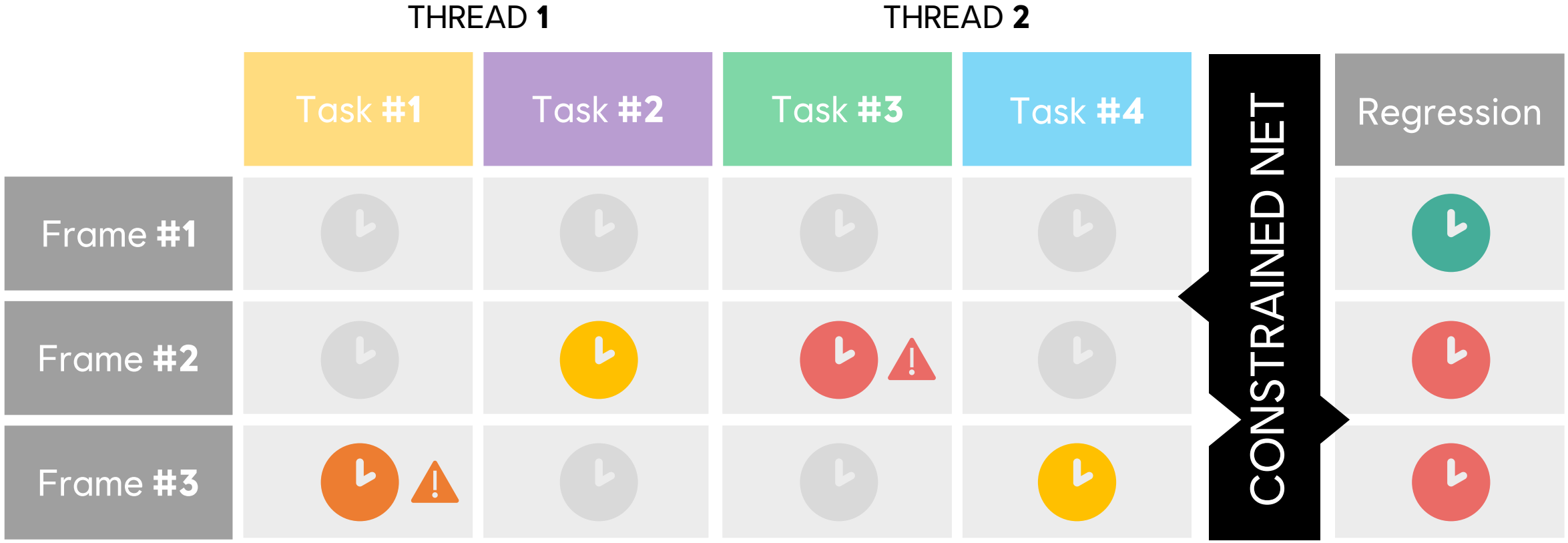
MEANINGFUL OUTPUTS

Heavy post-processes needed

ROBUST AND TRUSTWORTHY

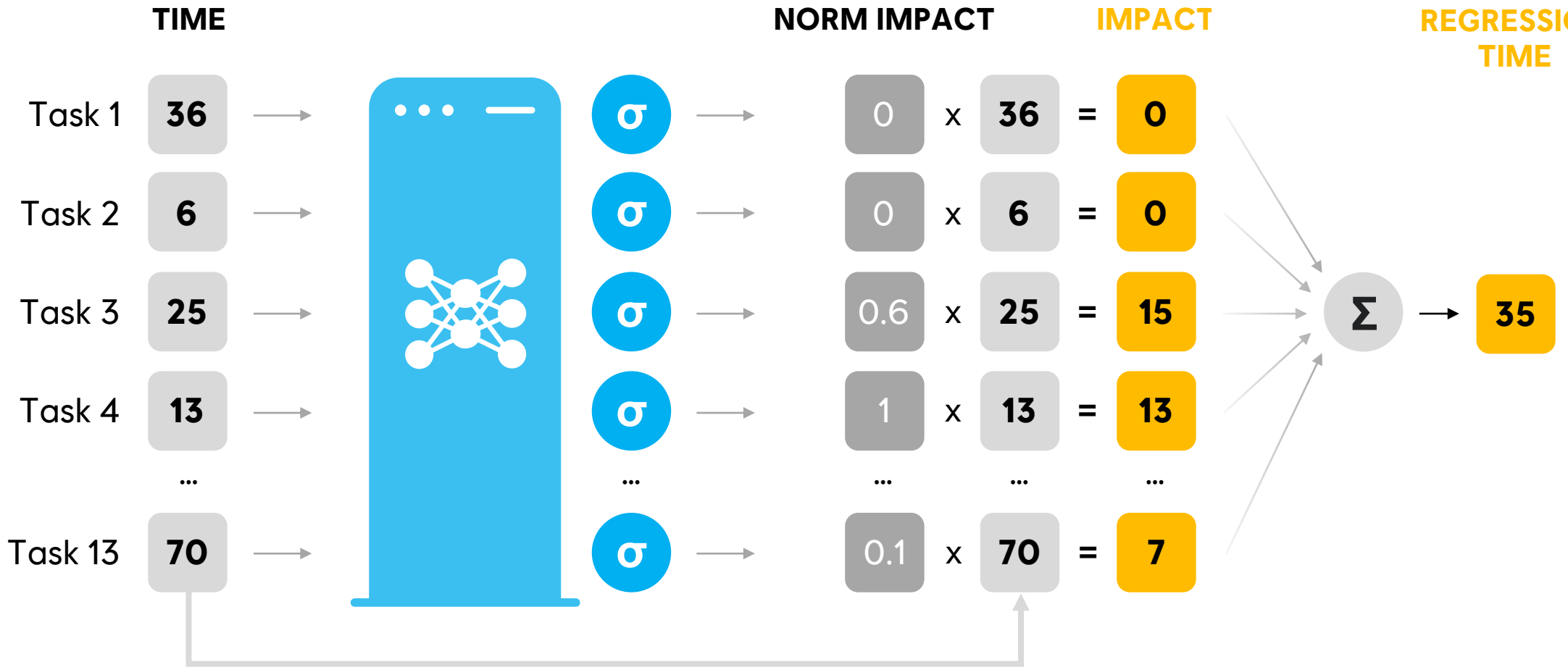
CONSTRAINED NEURAL NETWORK

Method 2



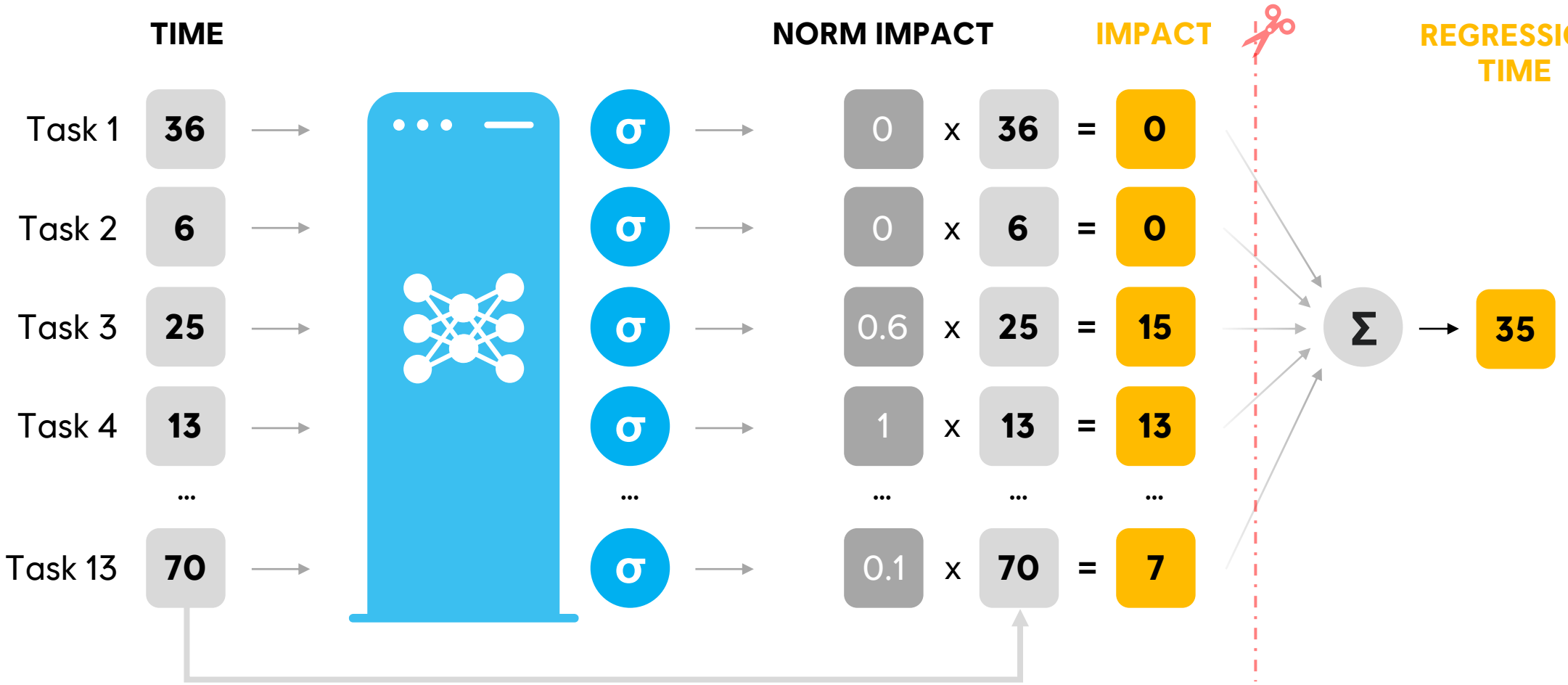
CONSTRAINED NEURAL NETWORK

Method 2



CONSTRAINED NEURAL NETWORK

Method 2



DECISION TREES & SHAPLEY VALUES

Method 1

+ CUMULATIVE

+ DATA FORMAT FLEXIBILITY

+ COST EFFECTIVE

FRAME-WISE EXPLICATIVE

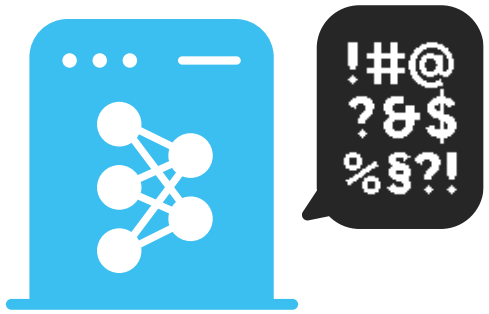
MEANINGFUL OUTPUTS

ROBUST AND TRUSTWORTHY

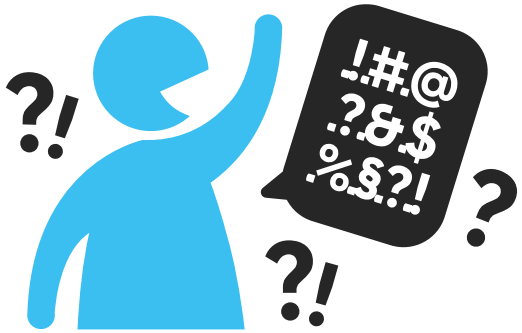
EVALUATION

- CHOOSE THE BEST METHOD
- MEASURE RELEVANCE
- BUILD TRUST

EVALUATION CHALLENGES



No standard methods for XAI evaluation



Good prediction ≠ good explanation

Single output evaluation = 1-5min of busy expert

Solution: Applicative survey to limit evaluation cost & minimize expert time

EVALUATION

Explainable models validation

Evaluation How-to

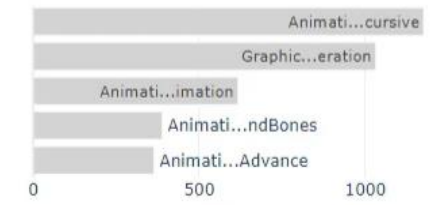
1. **Open GearStudio**
2. **Select** best offenders estimation
3. **Evaluate** relevance of each estimation
4. **Validate** or **inform** all true offenders you are sure about (click on bars in charts or "search and insert" just bellow)
5. **Comment** (optional)
6. **Confirm** or **skip** evaluation (when skipping, comment section will be stored and example won't be proposed again if filled-in)

Session description :

`Platform` : PS4
`Version` : 6597877
`Ubi session id` : a7b508d6-4664-4c05-86bb-d4ee35836672
`Xp type` : WalkHub32Pautotests
`Frame id` : 20161
`Frame length` : 37975µs
`Regression time` : 4642µs

Open in GearStudio

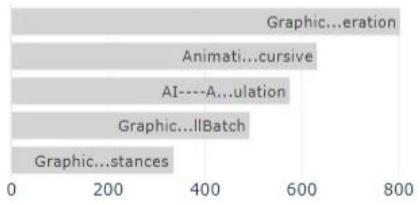
Top contributors to regression



Do you find those estimations relevant ?

No Yes

Top contributors to regression



Do you find those estimations relevant ?

No Yes

DEPLOYMENT

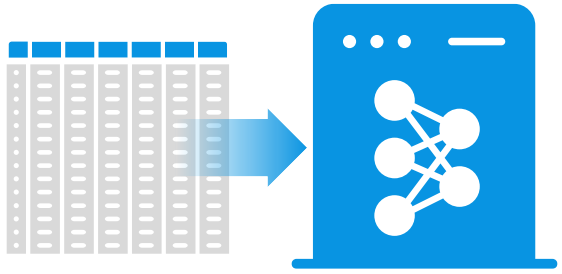
For each new game version & platform, daily

1

GATHER
sessions data

2

TRAIN
on multiple game versions



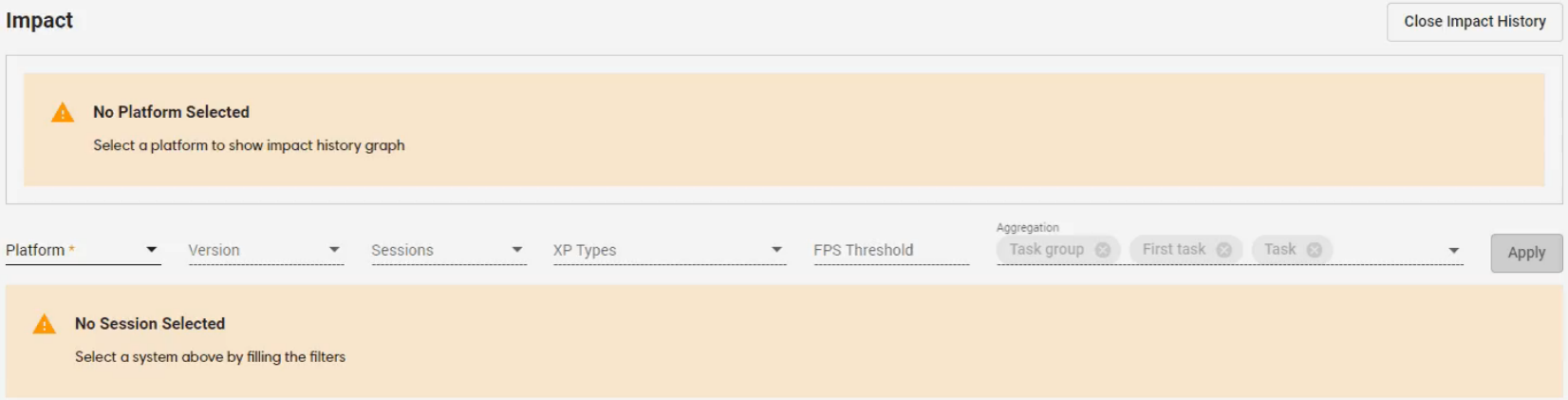
3

EXTRACT
& serve impact

Hyperparameters are updated **Bi-monthly**

REPORTS

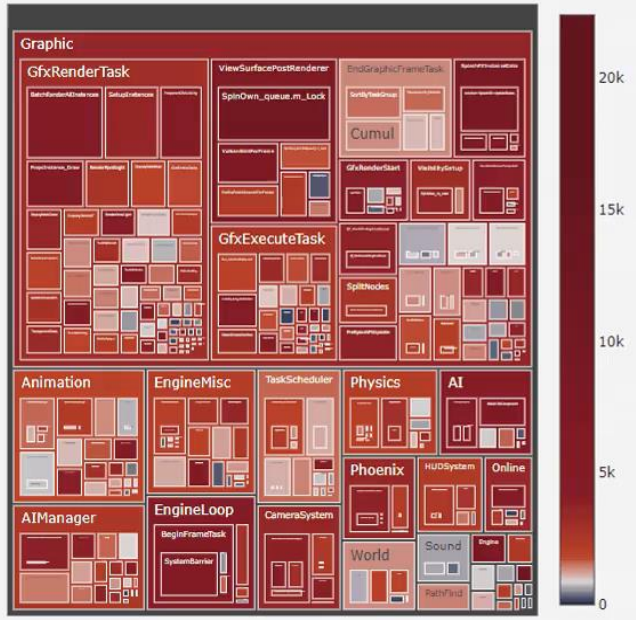
NAVIGATE GAME VERSIONS



REPORTS

EXPLORE GAME VERSIONS

Platform * WinCitrix | Version * 7365970 | Sessions * continouscapture,... | XP Types * InGameDead, PAFInsertion, Ti... | FPS Threshold * 56 | Aggregation * Task group, First task, Task | Apply

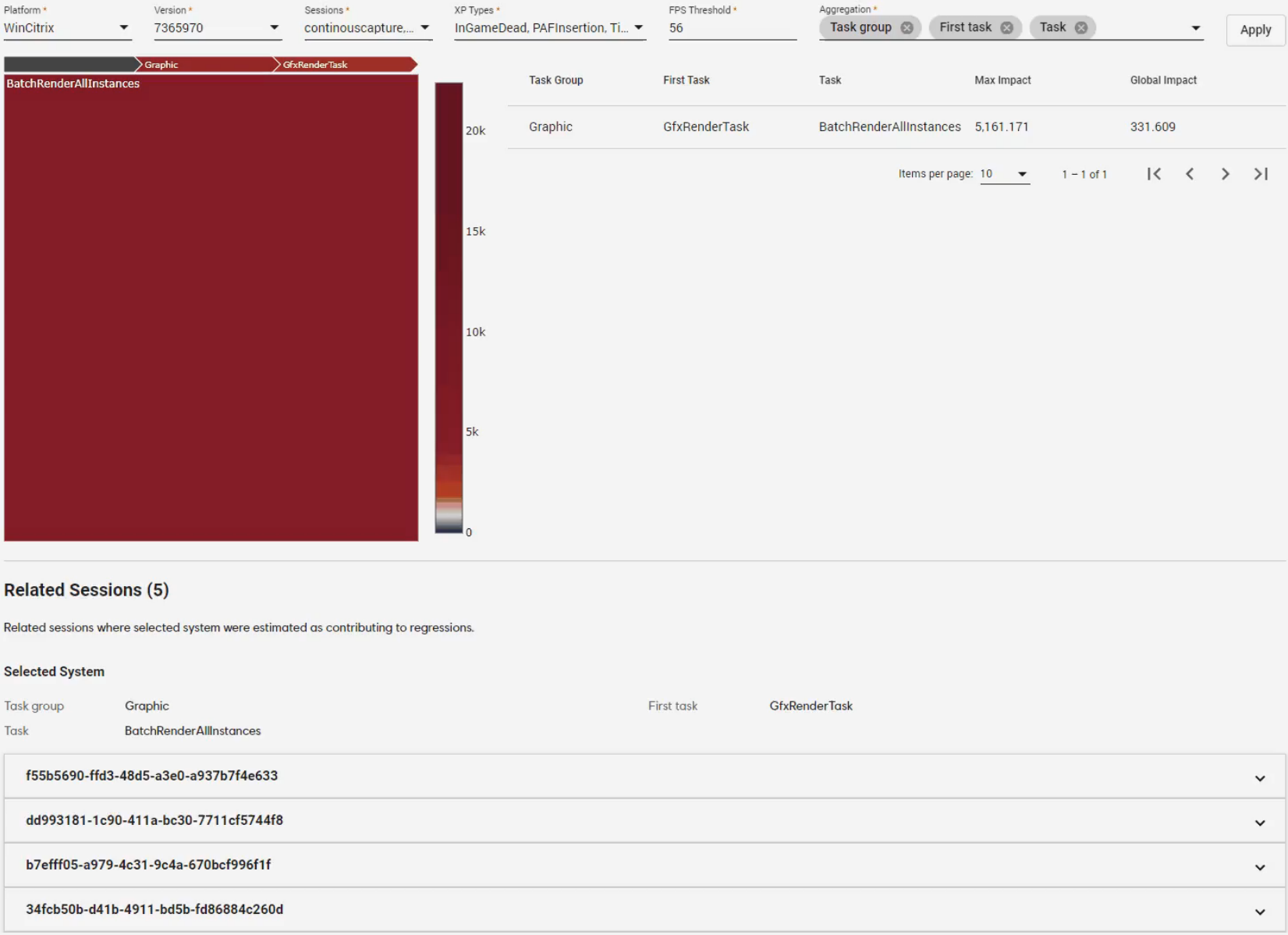


Task Group	First Task	Task	Max Impact	Global Impact
Graphic	ViewSurfacePostRenderer	SpinOwn_queue.m_Lock	11,303.997	430.499
EngineLoop	BeginFrameTask	SystemBarrier	6,046.418	389.3
Graphic	GfxRenderTask	BatchRenderAllInstances	5,161.171	331.609
Graphic	SplashFXInstanceData	scimitar::SplashFX::Updat...	9,337.21	250.747
Graphic	GfxRenderTask	SetupInstances	3,546.097	234.489
Graphic	GfxRenderTask	PrepareHZVisibility	4,496.753	185.194
Graphic	ViewSurfacePostRenderer	VulkanWaitForFrame	7,031.244	177.049
Graphic	GfxRenderTask	PropsInstance_Draw	5,510.992	165.355
Graphic	EndGraphicFrameTask	SortByTaskGroup	1,573.664	158.423
Graphic	QT_VisualsCullingTrivialA...	QT_MainCameraCullingTri...	3,765.921	153.224

Items per page: 10 | 1 – 10 of 808

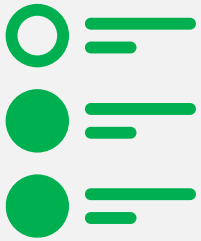
REPORTS

FIND EXAMPLES



A NEW WORK CYCLE

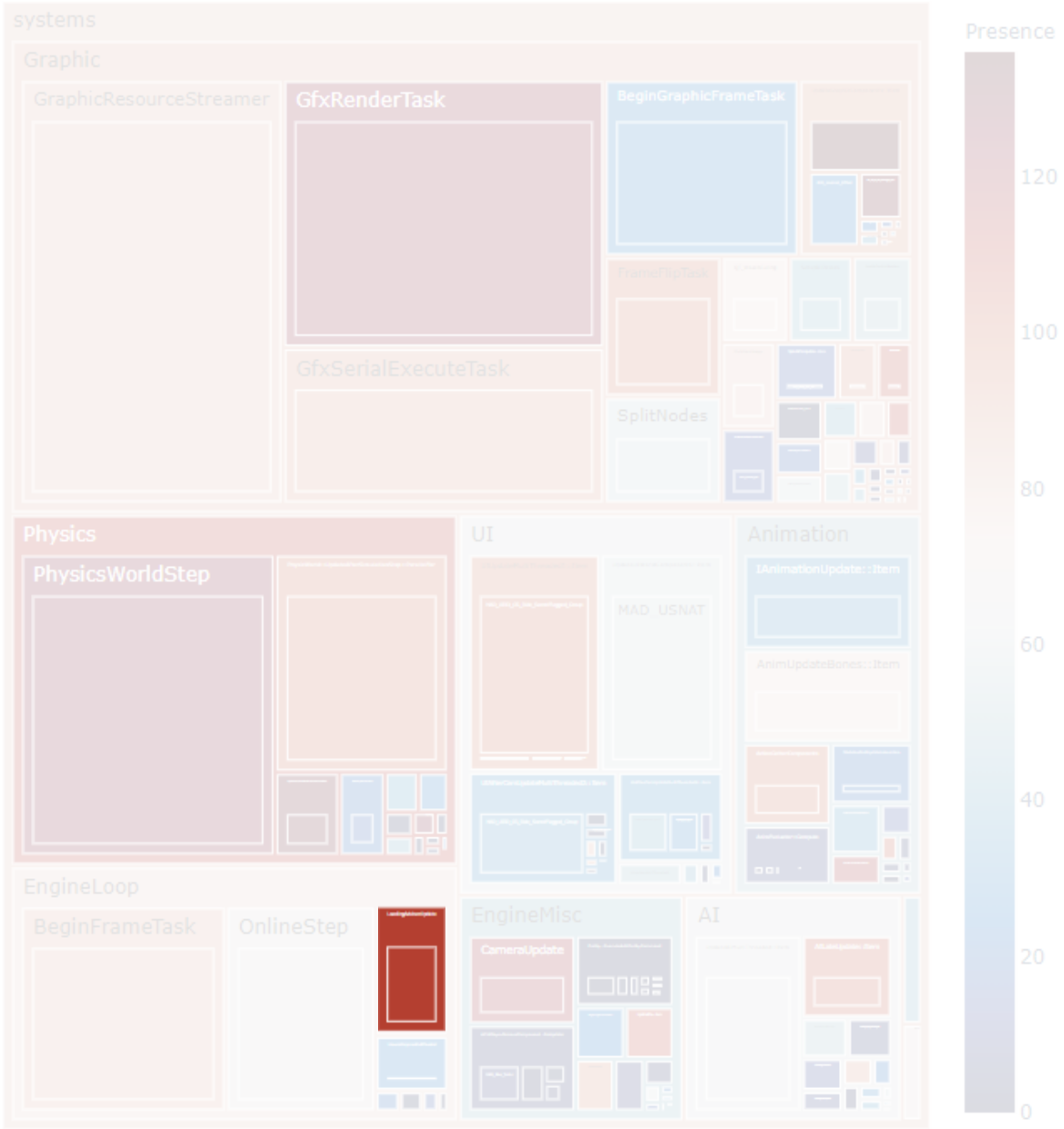
Look at the results



Check & Prioritize.

Distribute to the team

Repeat every week



« LOADING ADVISOR UPDATE » SYSTEM

Riders Republic, example 1

BEFORE

LoadingAdvisorUpdate :

- Impact **30% frames**
- Avg regression of **1ms**

AFTER

LoadingAdvisorUpdate :

- Impact **10% frames**
- Avg regression of **0,1ms**

FRAMEWORK : THE TOOLS

Dashboard

Ad-Hoc analysis

Machine learning model

MARIO RABBIDS : SPARKS OF HOPE

A NEW ENGINE, A NEW MINDSET

MARIO + RABBIDS : CONTEXT

PERFORMANCE TEAM

Key roles

PERFORMANCE CHAMPION

Ensure data quality, check data availability, train the team

PROJECT CHAMPION

Identify priorities, reports to core team about game's health

DATA ANALYST

Create custom dashboards & ad-hoc analysis

THE SUCCESSES

STRUCTURE

Thanks to recommended organization

EARLY DETECTION

Of issues

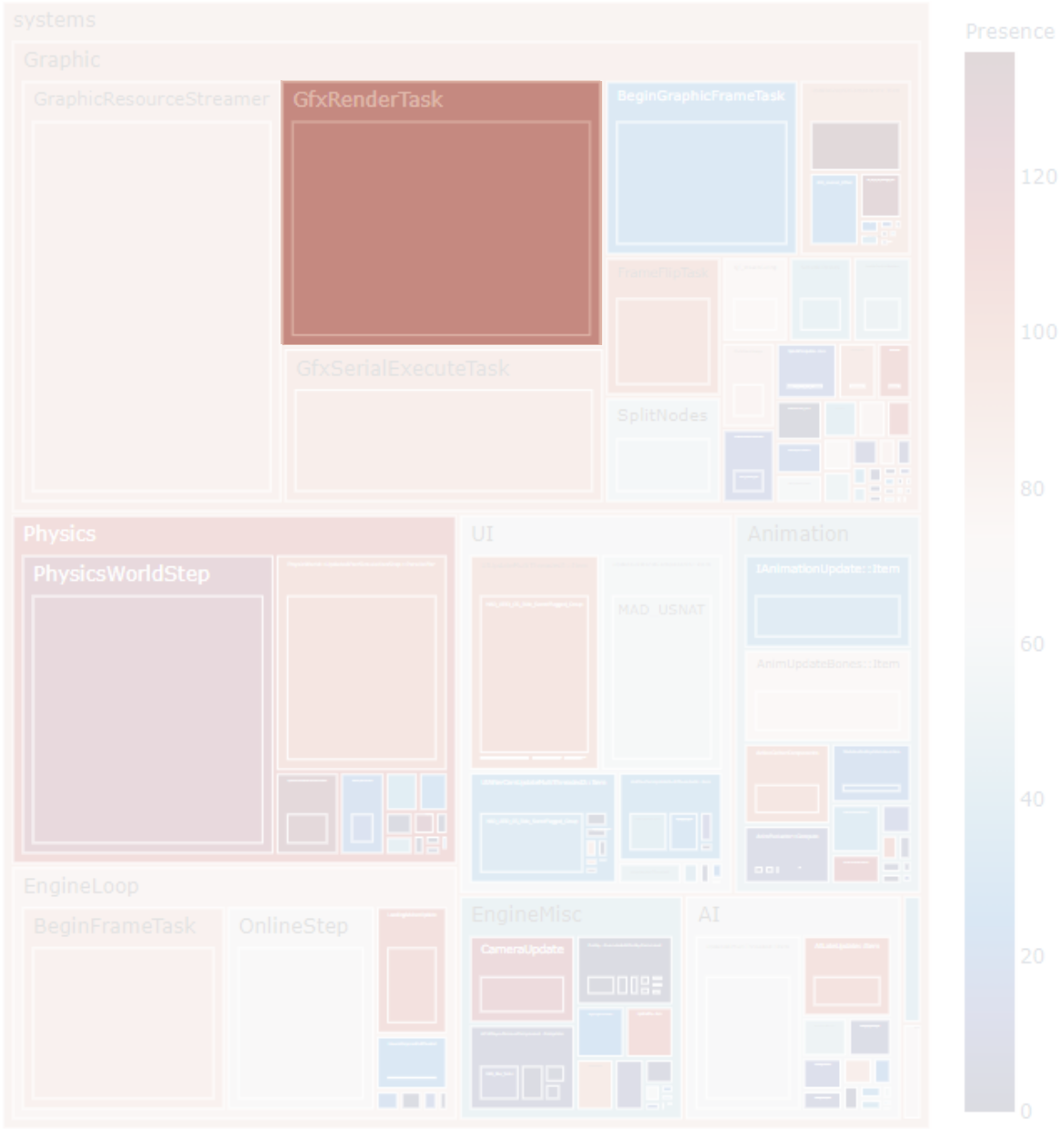
SERENITY

About CPU performance topic

TO BE IMPROVED

RELEVANCE
of some results

PRECISION
of data



POSSIBLE SOLUTION

UX Post processing

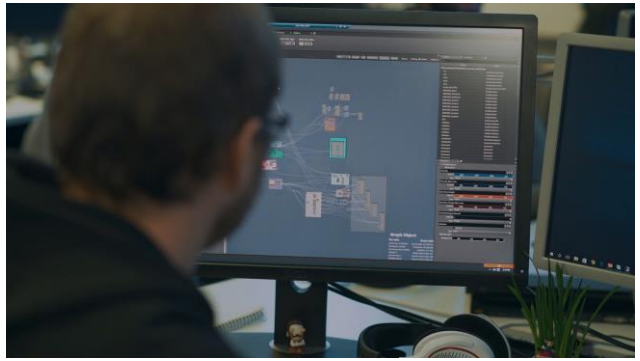
DO NOT:

- Remove completely
- Force assumptions into the model

WHAT CHANGED

BETTER GLOBAL VISIBILITY

on project health



BETTER EXPERTS' WORKFLOW

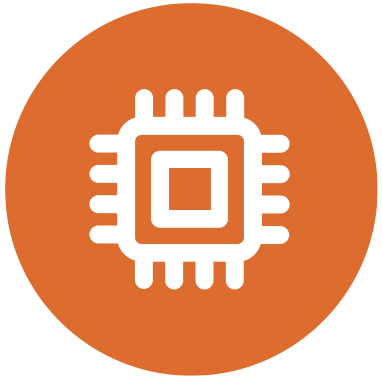
They are less bottleneck

RELIEF

On CPU performance topic

But we can still improve knowledge democratization!

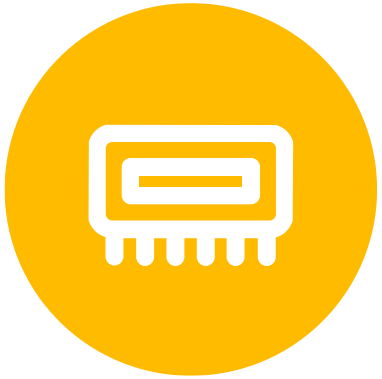
WHERE TO GO NEXT?



CPU

Locks, context switches & More game context

GPU



MEMORY

BANDWIDTH

SPECIAL THANKS

JEAN-PHILIPPE BARETTE
Senior Technical Architect

MATTHIEU FOURNAISON
Lead Programming Architect

FRANÇOIS DETEMMERMAN
Gameplay Architect

THOMAS BIRON
Product Manager

MAXIME MERY
Lead Production Analytics

PIERRE-THOMAS MEISELS
Data Engineer

MATTHEW CHARLTON
Senior Web Developer

CARL LÉTOURNEAU
Data Developer

SPECIAL THANKS

BAPTISTE DECAIN
Lead Programmer

GUILLAUME NEWTON
Programmer

GUILLAUME GRODWOHL
Senior Programmer

XAVIER BILLAULT
Senior Tech Lead Programmer

STEPHANE MARQUIS
Technical Lead

SIMON POLIQUIN
IT Project Manager

MICHEL PERRAULT
IT Team Lead

SIVA VELOUMOUROUGANE
Data Scientist Director

THOMAS GUILLEMBET
Data Visualization specialist

GAUTHIER TANGUY
Lead Programmer

NORBERT REDON
Senior Programmer

ELIA WRZESNIEWSKI
Associate Producer

ALEXANDRE NOVAK
QA Analyst

ARTHUR BLIN
Data Analyst

RESSOURCES

CHARLIE CURTSINGER, EMERY D. BERGER
Coz: Finding Code that Counts with Causal Profiling

BRENDAN GREGG
Flamegraphs visualization

Q&A

March 20–24, 2023 | San Francisco, CA