Microsoft Azure

InstaDeep™ NEURAL CONCEPT NVIDIA

Cloud-based physics-driven AI to accelerate Design & Engineering

Dr. Lukasz Miroslaw – EMEA GBB HPC/AI Sr. Specialist (Microsoft)

Dr. Lukasz Miroslaw – EMEA GBB HPC/AI Sr. Specialist (Microsoft)

NWC23 NAFEMS WORLD CONGRESS

What is a digital twin?

According to the very first definition back in 2003 by Michael Grieves [1][2][3], a Digital Twin is a virtual representation of a physical entity, collecting all the information related to his lifecycle management

Digital model	• Digital representation of the physical entity
Digital shadow	• Digital representation and physical-to-virtual information flow
Digital twin	• Digital shadow with virtual-to-physical information flow

Digital twin requires three elements[1][2]:

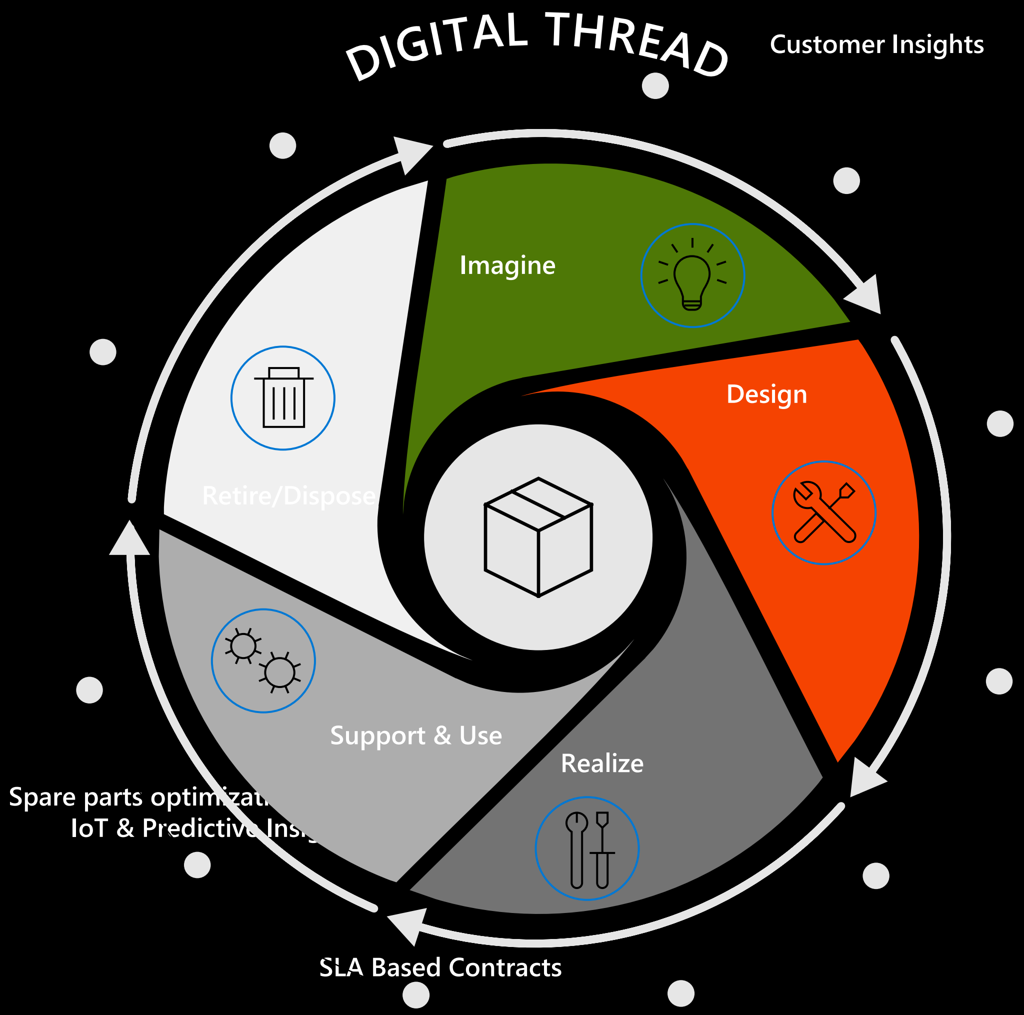
- A physical entity
- A digital representation of the physical entity
- A bi-directional information flow between the two

[1] David Jones, Chris Snider, Aydin Nassehi, Jason Yon, Ben Hicks, **Characterising the Digital Twin: A systematic literature review**, CIRP Journal of Manufacturing Science and Technology, Volume 29, Part A, 2020, Pages 36-52, ISSN 1755-5817

[2] Grieves, Michael. **Digital twin: manufacturing excellence through virtual factory replication**. 2014. *White Paper* (2017).

[3] Mohsen Attaran, Bilge Gokhan Celik, **Digital Twin: Benefits, use cases, challenges, and opportunities**, Decision Analytics Journal, Volume 6, 2023, 100165, ISSN 2772-6622

Digital Twin in Stark's Product Life Cycle



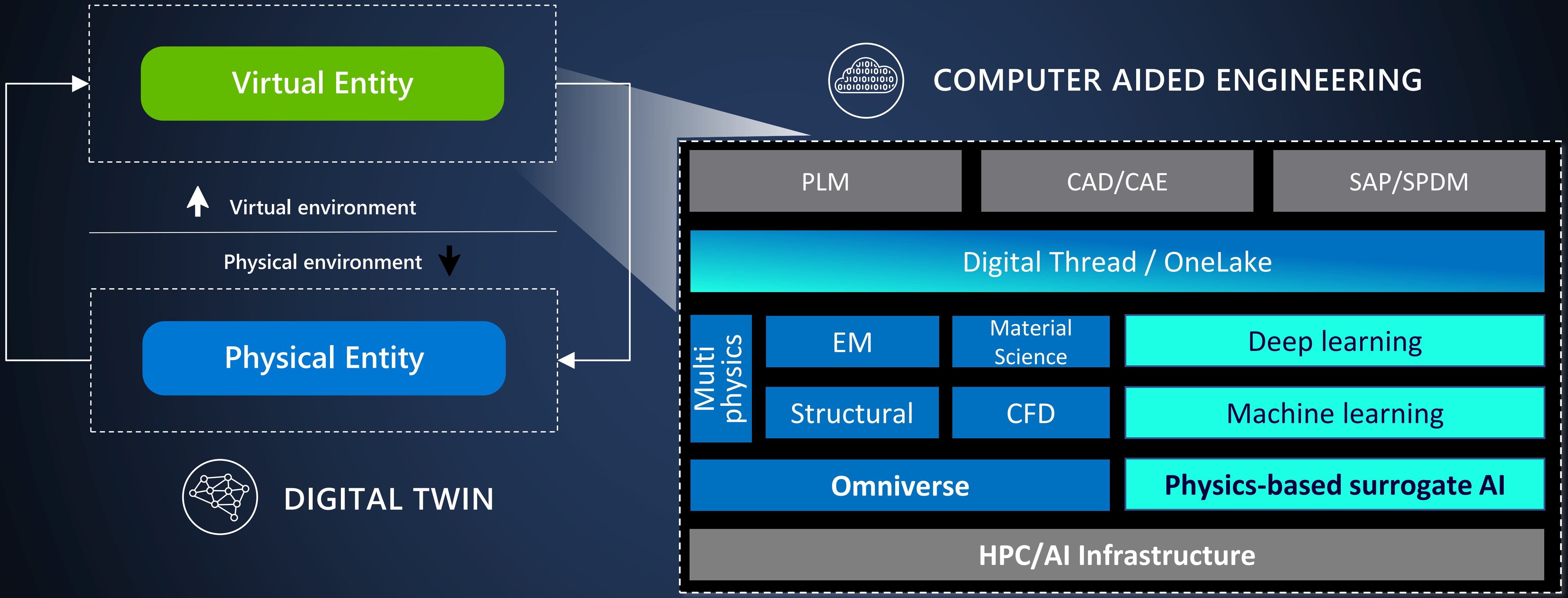
"Digital Thread is a data-driven architecture that links together information generated from across the product lifecycle"

Availability of a Digital Thread architecture is a key enabler to enrich Digital Twin capabilities

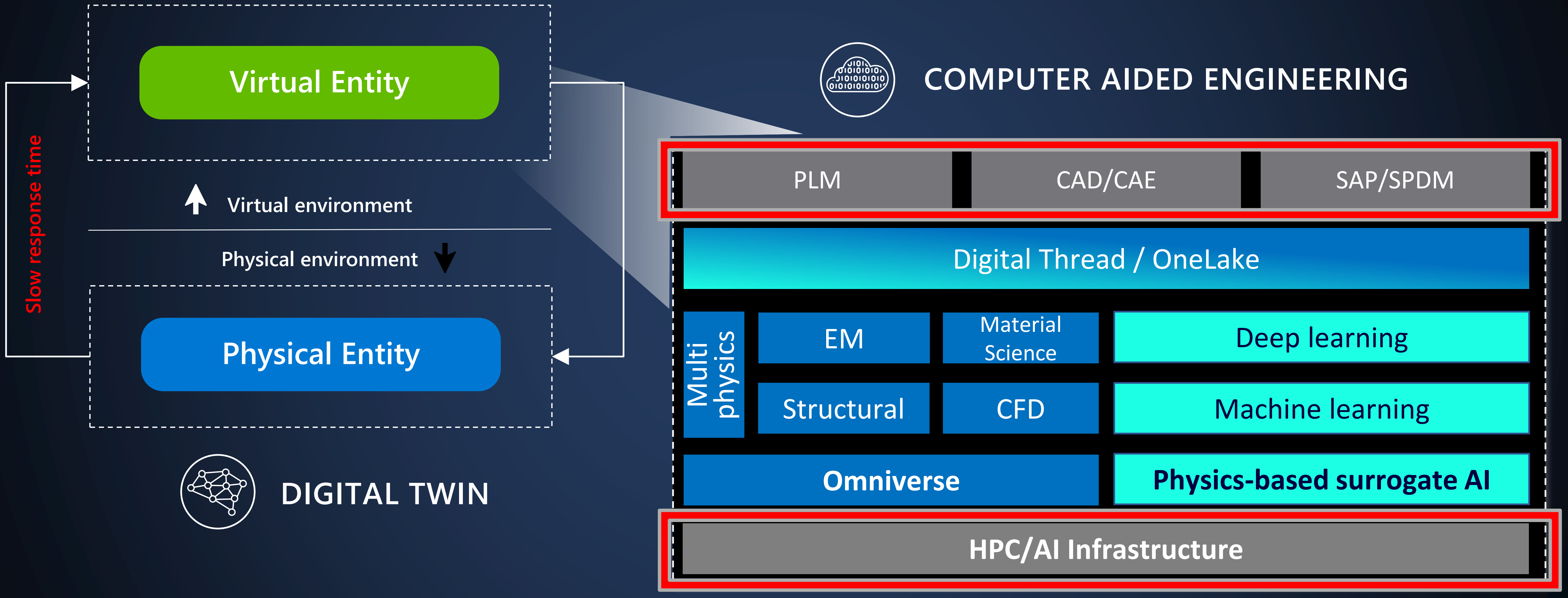
[1] Stark, J. **Product Lifecycle Management (Volume 1)**, 21st Century Paradigm for Product Realisation, Springer International Publishing, 2022

[2] Singh V., Willcox K. E., **Engineering Design with Digital Thread**, AIAA Journal 2018 56:11, 4515-4528

Next Gen CAE: Physics-based Digital Twin



Next Gen CAE: Physics-based Digital Twin



Azure HPC/AI

The universe of computational resources

Compute

SKU	CPU
HC	Intel Xeon Platinum "Skylake"
HB	AMD Epyc "Naples"
HBv2	AMD Epyc "Rome"
HBv3	AMD Epyc "Milan"
HBv4	AMD Epyc "Genoa"

InfiniBand enabled VMs
- EDR: 100 GB/s (HC/ HB)
- HDR: 200 GB/s (HBv2/3)
- NDR: 400 GB/s (HBv4)

Visualization

SKU	GPU
NV	Tesla M60
NVv3	Tesla M60
NVv4	Radeon Instinct MI25
NVads A10 v5	A10 Tensor Core

- NVIDIA GPU coming with NVIDIA Grid license
- Partial GPU available on A10

Machine Learning

SKU	GPU
NC	Tesla K80
NCv2	Tesla P100
NCv3	Tesla V100
NCasT4_v3	Tesla T4
NC A100 v4	A100 Tensor Core

- Multiple GPUs available per node
- NVLink enabled on NCasT4_v3 and NC A100v4

Deep Learning

SKU	GPU
ND	P40
NDv2	Tesla V100
ND A100 v4	A100
NDm A100 v4	A100

- Multiple GPUs available per node
- NVLink enabled and InfiniBand starting from NDv2

ARM-based

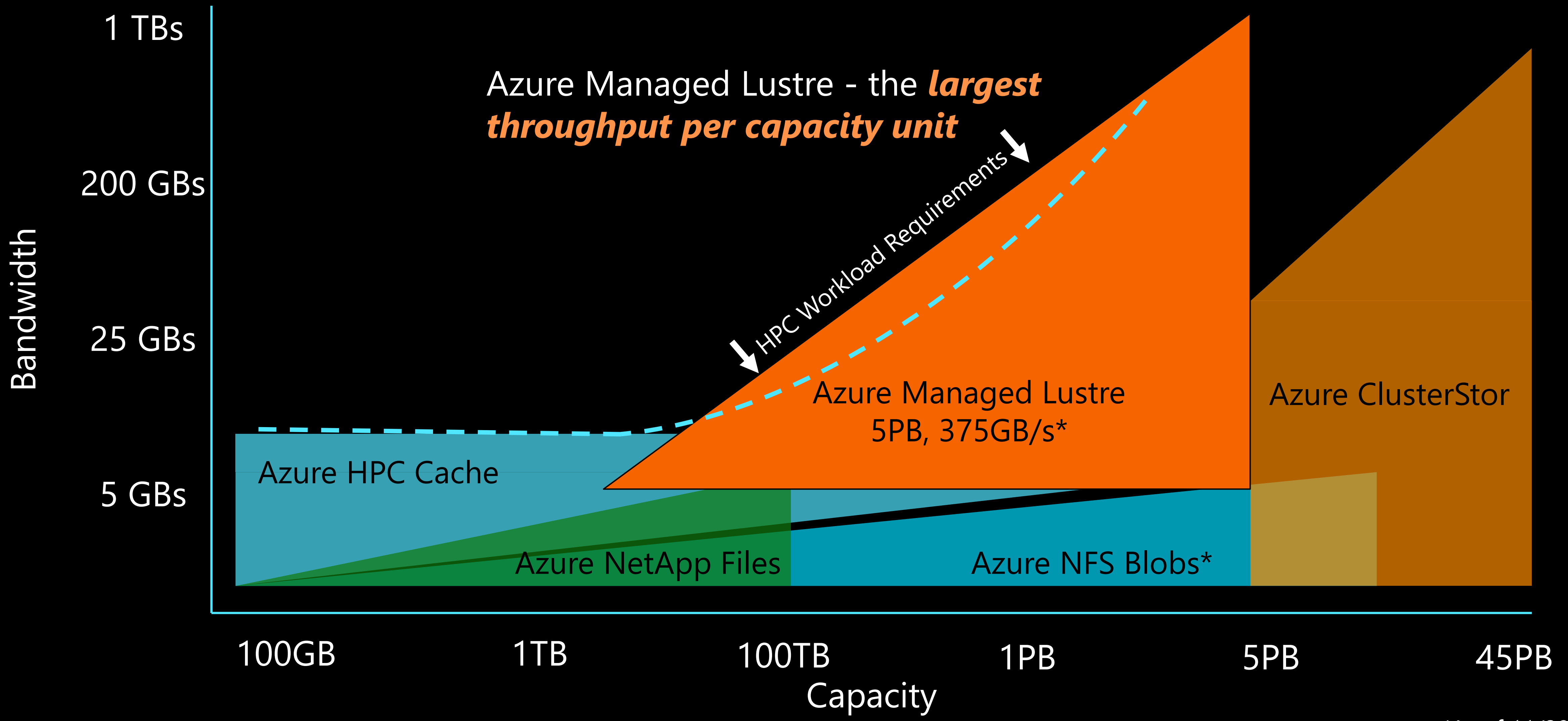
SKU	CPU
DP / EP	Ampere® Altra®

FPGA

SKU	FPGA
NP	Xilinx U250

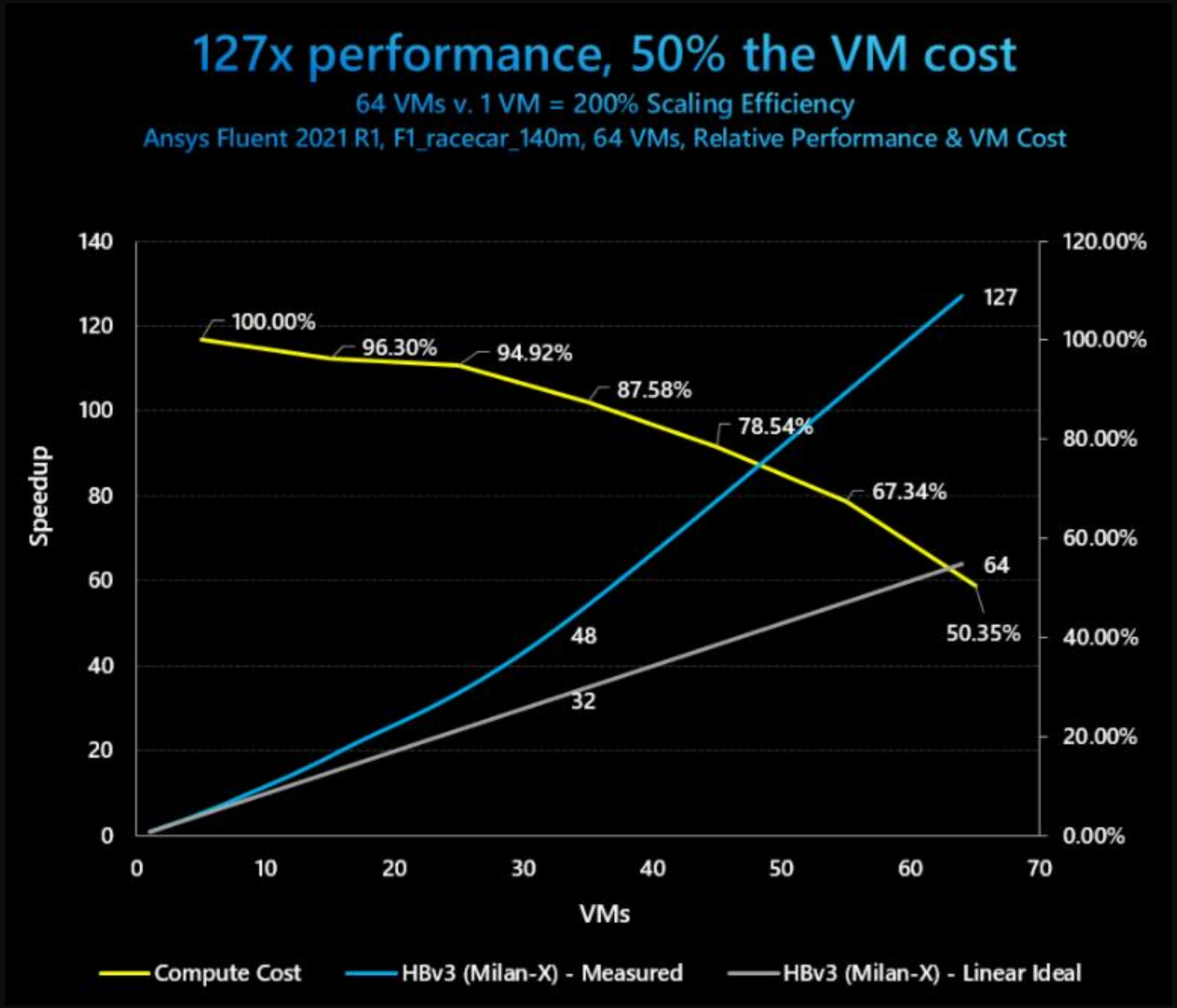
- ARM based SKUs based on Ampere Altra
- FPGAs based on Xilinix U250

Azure HPC File Systems



*As of 11/2022

Faster & Cheaper: Transformation with Cloud



An Example: 127x performance, 50% the VM cost

- Fluid Flow (CFD) simulation study of an F1 race car
- 140 Million Cell model of the car, using ANSYS Fluent

- 64 Virtual Machines VS 1 Virtual Machine
- Using HBv3 (120 cores of AMD Milan-X)
- Strong scaling of the simulation progressively allows a large percentage of active data to fit in memory

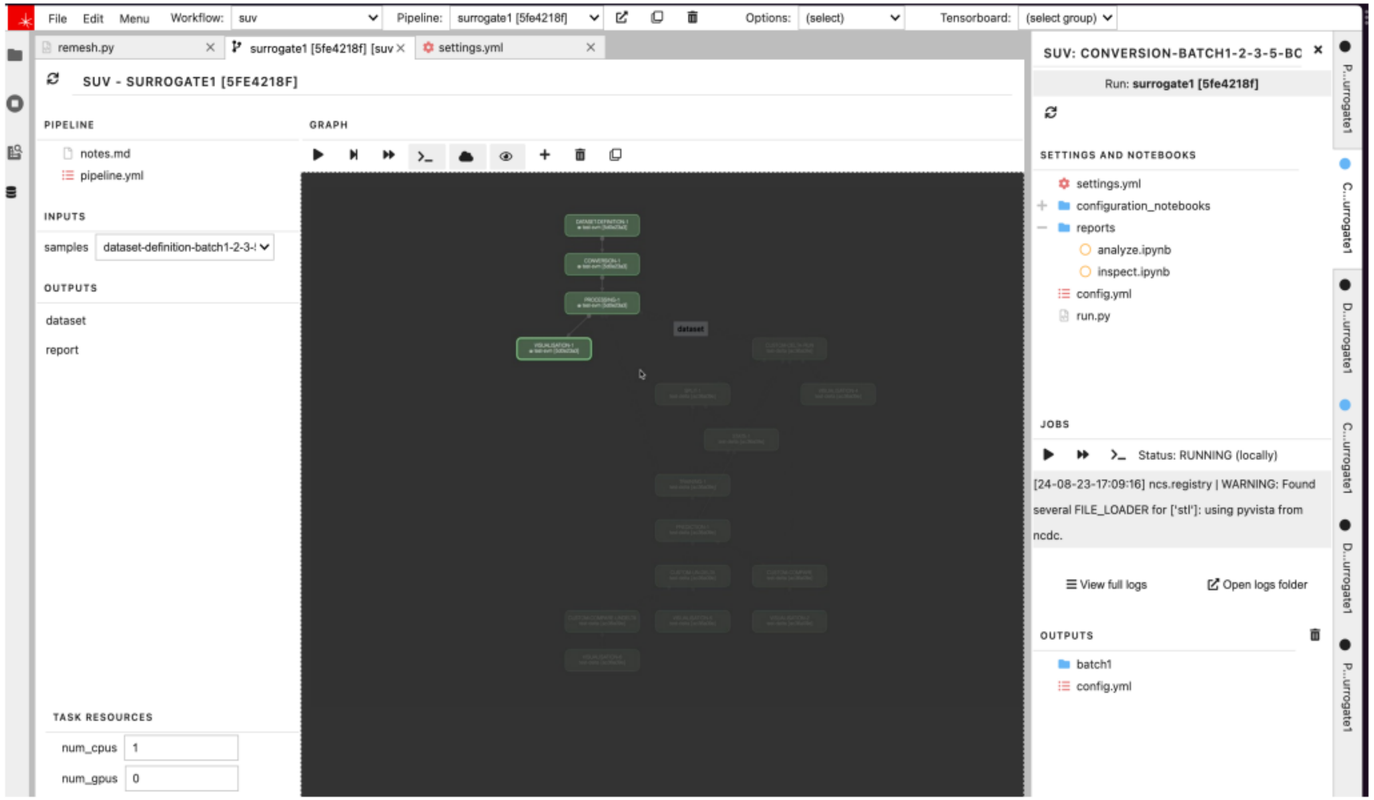
- The 64 VMs can be turned off, half as long as it takes 1 VM to complete the simulation
- Result: 127x less time to complete the simulation, for half the cost

- Example illustrated:
    - 1 VM, taking 100 hrs = **$468**
    - 64 VMs, taking 47 mins = **$234**
    - With AI: 1 GPU: taking 10 seconds = $0.56

AI-infused Digital Twin in Digital Engineering

Neural Concept: Pioneer in AI for Engineering

- Spin-off from EPFL, Computer Vision Lab
- Team of 40+
- 60 + customers worldwide
- Engineering Intelligence at the core



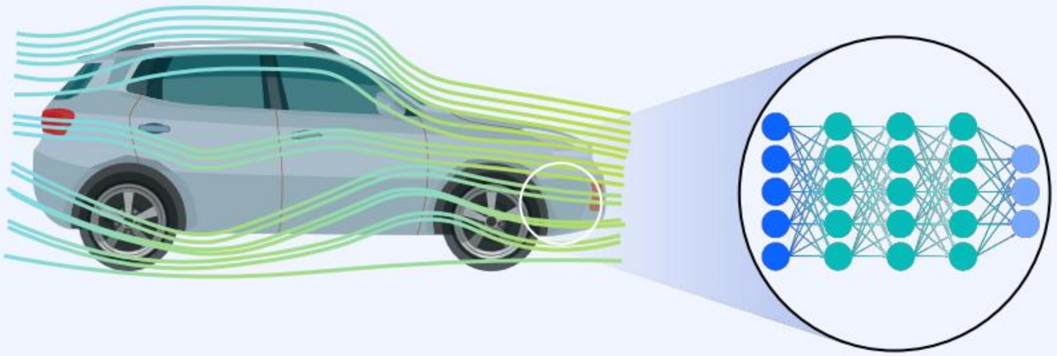
SAFRAN, AIRBUS, BOSCH, PSA GROUPE, NUMECA, mcpp MITSUBISHI CHEMICAL, miniswys, PLASTIC OMNIUM, Parrot, Endress+Hauser, Mubea, GE, SPEO, DSM, esa, LG, BÜHLER, DLR

Joint Engagements FY24

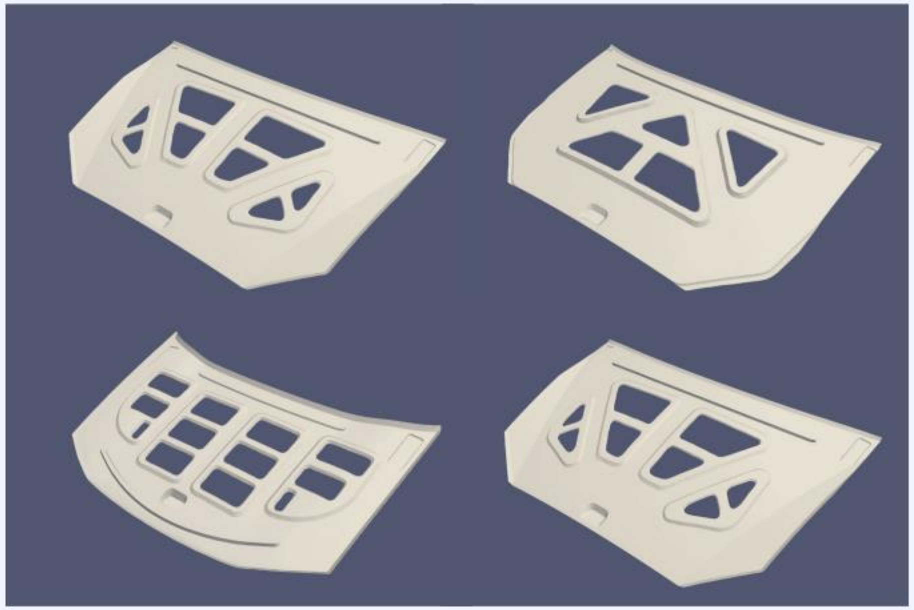
Danfoss, DLR Deutsches Zentrum für Luft- und Raumfahrt German Aerospace Center, Endress+Hauser, LIEBHERR, ABB

AI is redefining Engineering Design software

2 KEY BENEFITS



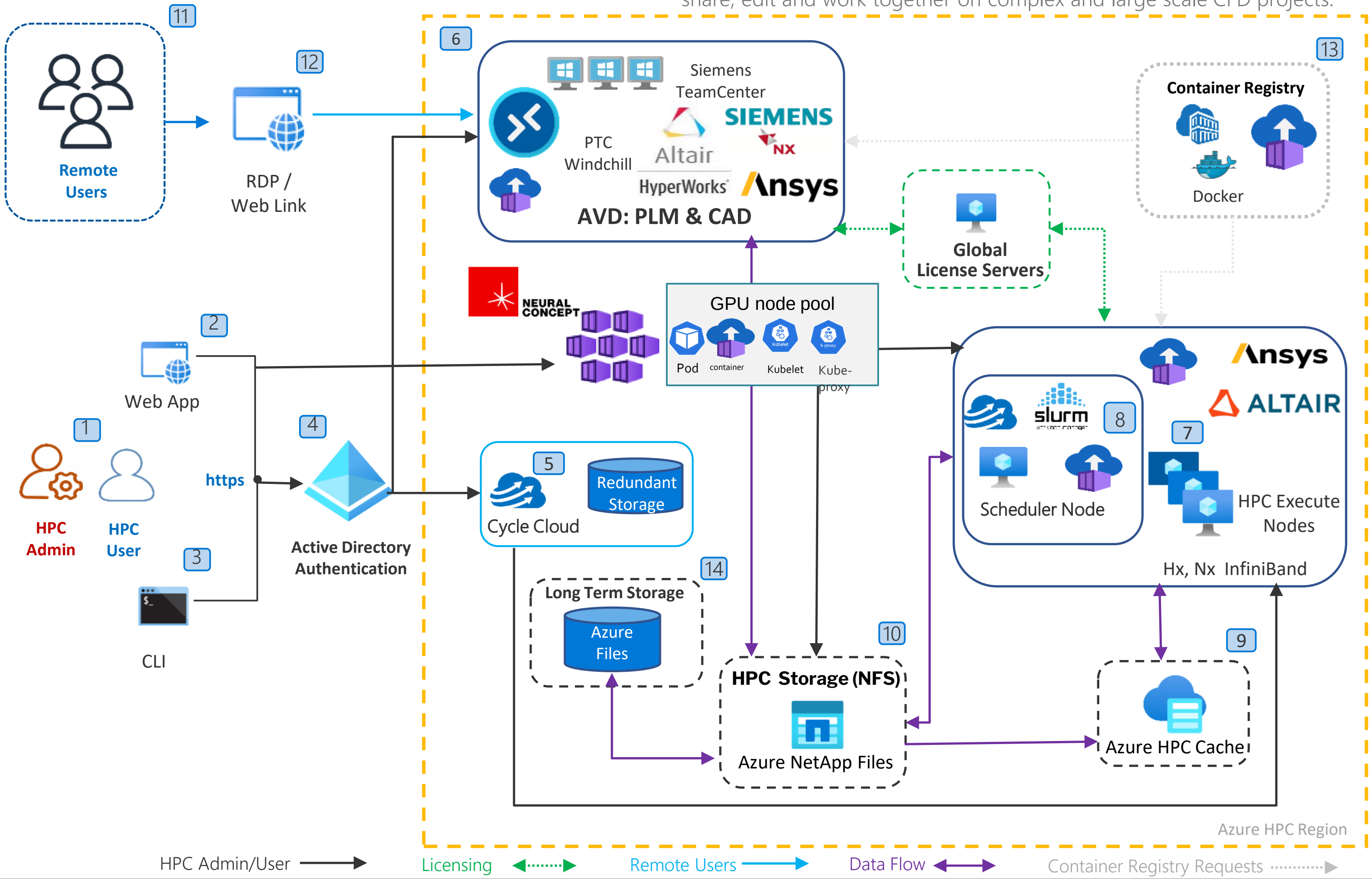
Simulate **10 to 10'000 times faster**



Deep Generative models output innovative and plausible geometries at **100/min rate**

Reference Architecture
HPC & AI integration for CAE

- Supports BYOL: Allows optimization of licenses cost and employee value by leveraging configuration and scalability of the compute nodes.
- Cloud Native Data: Optimizing data transfer costs and GPU acceleration for compute, post-processing and remote high-def visualization.
- Cost effective Remote working: With Azure Virtual desktop interface users can work independently with AVD sessions on pre and post Processing. By this VM resources can be shared effectively between the users.
- Domain Agnostic: Designed for AI-infused CAE/CFD/DEM Workload of Manufacturing Engineering, could be leveraged for Manufacturing, Aerospace.
- Economic and User Experience: High productivity with lower simulation costs, No installation, no maintenance, no overhead, Effective collaboration features to share, edit and work together on complex and large scale CFD projects.



- HPC Admin [1] connects with Cycle-Cloud server [5] via Web [2] or CLI [3] and deploys the HPC execute Nodes [7] with required number and type of VMs . HPC Admin [1] Deploys the Visualization Nodes [6] in AVD Infrastructure and creates the web link [12] with Role based access to Remote set of Users [11].
- Active Directory authentication [4] is responsible for the secure access to services and data.
- The HPC Admin [1] connects with visualization Nodes [6] and installs the respective apps for Pre-processing & Post-processing, also connects to the Scheduler Node [8] and installs the required solver application.
- The Remote Users [11] connects to Visualization Nodes [6] via web link [12] and starts working individually on Pre and Post processing.
- Users can store their permanent data in the long-term data storage [14], Data can be archived from Azure Net App Files [10] to the long-term storage [14].
- The user[1] will now connect to the HPC execute Nodes [7] or Neural Concept Shape and launch the training job using AI or retrain the model with CAE/CFD Simulation jobs. We can make use of Azure HPC Cache [9] for Agility & Caching the large files to improve efficiency.
- After the simulation is completed the output file is saved to Azure NetApp Files [10] for Post-processing.
- The Solver output file from Azure NetApp Files [10] can be accessed from the Visualization Nodes[6] and Post-Processing team who are Remote Users [11] can perform Post-processing using different AVD sessions as required.
- Containerized applications are pulled by Visualization nodes [6] and HPC Nodes[7] from the user's container registry [13].
- The HPC user/Admin [1] stops Visualization Nodes [6] and Cycle-Cloud [5] which deallocates nodes to optimize the cost and spin up again when required.

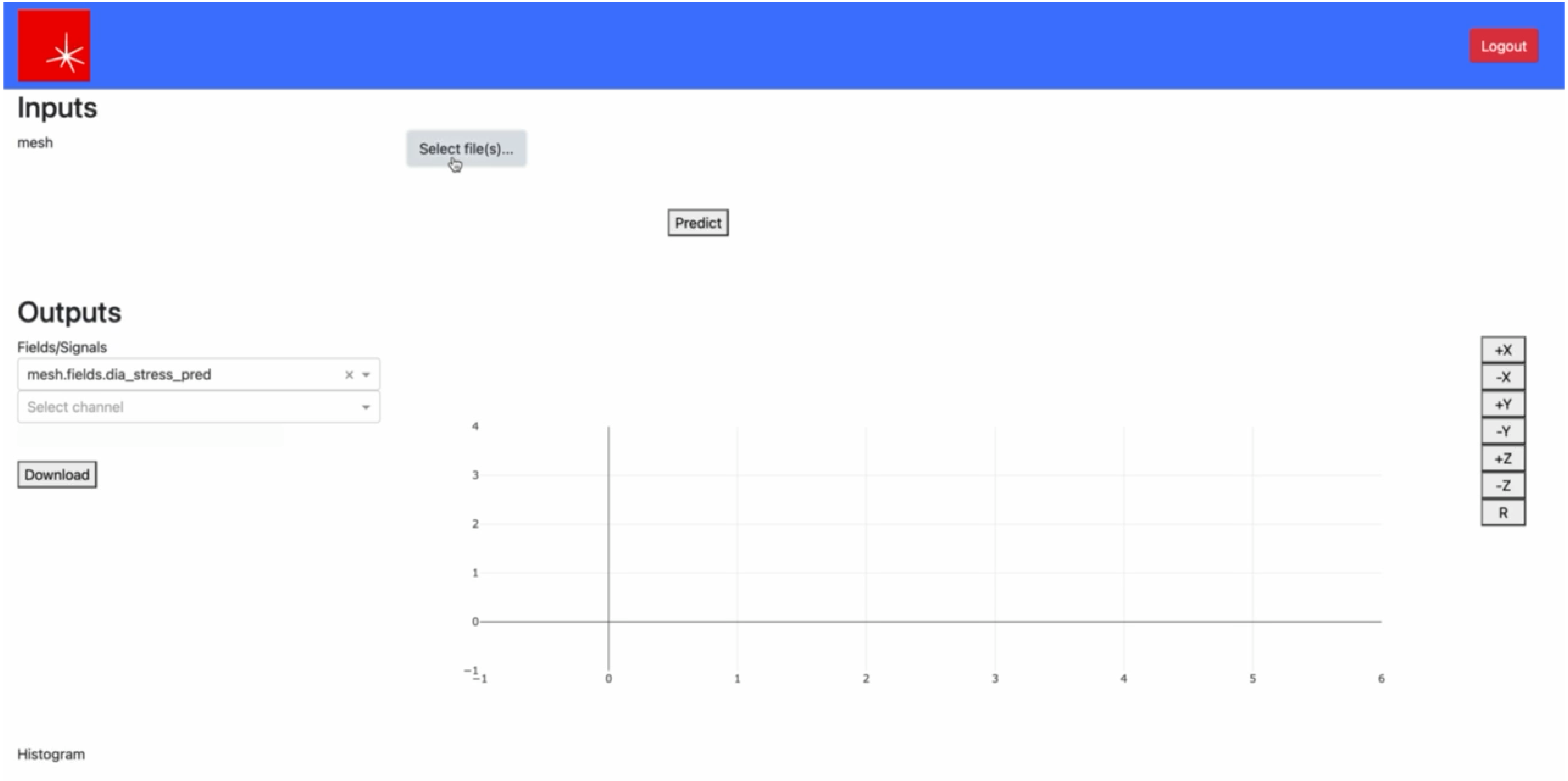
Services	Type	Quantity	Size	Services	Type	Quantity	Size	Services	Type	Quantity	Size
Azure Cycle Cloud	VM	1	D4sv3	Azure NetApp files	Storage	1	1TB	IP Address	Networking		
Scheduler node	VM	1	D4sv4	Azure HPC Cache	Storage	1	3TB	Virtual Network	Networking	1	
HPC Nodes	VM	8	HBv3	Data transfer cost	Bandwidth	1	100GB				
AVD Nodes	VM	2	NVv5	Container registry	Containers	1					
Azure Files	Storage	1	2TB								

CFD Analysis, ~7.6 Million Cells, hexa-mesh, Aerospace Domain

Configuration = 512 cores/8nodes, Compute time: ~0.6 Hrs per job

Cost per job: ~150 USD [Assumed 512 cores / 8 n]

Demo – Brackets analysis



Make better products

∞ DIMENSIONAL LARGE-SCALE OPTIMIZATION, BETTER PRODUCT PERFORMANCES

AN EXAMPLE WITH A MAJOR ACTOR OF HYDRO-ELECTRIC ENERGY

Deep learning models allow design teams to evaluate tens of thousands as many potential designs.

Traditional vs deep learning–based design workflow, turbine design

Traditional workflow (overall time required: months)

- Engineers select relevant parameters based on experience; 10–20% selected manually
- Geometries generated by varying parameters
- Computational fluid dynamics runs, each taking hours
- Engineers filter top 5 geometries based on experience; several geometries to proto-
- Real-life experience determines the best geometry, with one selected, based on

Deep learning–based workflow (overall time required: ~25% le

10,000×

- Deep learning iteratively selects the most relevant set of parameters; up to 100% of parameters selected automatically
- Geometries generated by varying parameters; up to 1 million times as many geometries generated
- Predictions mad through deep le model; all geome generated and v set of constraint tested; seconds model run

McKinsey & Company

25-50% faster design process

ACTUALLY MEASURED BY THE COMPANY

> 2M USD

ANNUAL SAVINGS IN PRODUCTION + HUNDREDS MWATTS OF ENERGY SAVINGS

IGBT Cooling for automotive supplier

EFFICIENT COLD PLATES FOR HIGH POWER ELECTRONICS

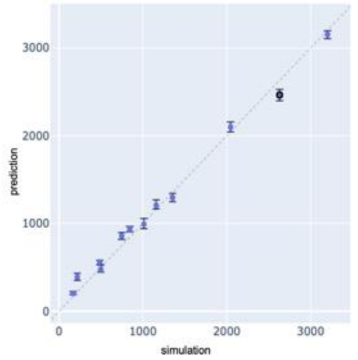
Context

- **Peak temperature needs to be mitigated** as it can reach 200°C
- Using intuition for designing in these Reynolds regime is difficult. **Experience in this domain is not widespread.**
- **Every new design needs to adapt** to specific packaging and boundary condition constraints.

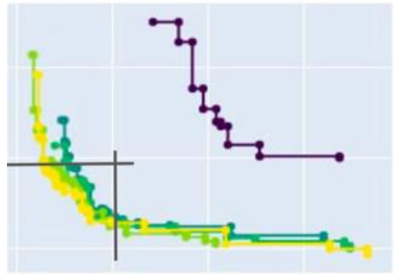
Achievement

- A workflow that allows to **optimize designs in <2h** for new constraints has been deployed.
- Design produced by the company are typically **2 to 3% more efficient** than using the old process.
- **Optimized designs wr.t. multiple objectives** (pressure drop and thermal efficiency)

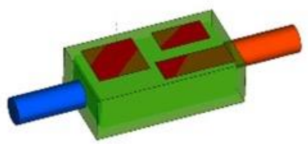
x4 faster	**+2-3%**	**Reduced**
TIME TO REACH FINAL DESIGN	IN THERMAL EFFICIENCY VS PREVIOUS DESIGN WORKFLOW	PEAK TEMPERATURE ACROSS THE FLUID



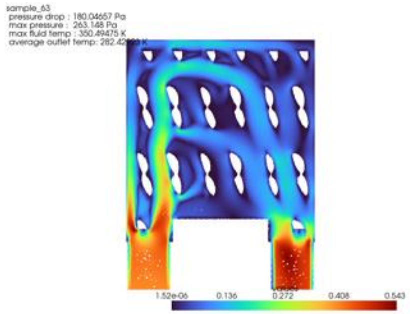
98% correlation between Ansys simulations (3h) and NCS prediction (1s)



1000 designs explored in <1h



Constraints specified via CAD import



One of the geometries on the Pareto-front

Thermal Management of EV Batteries

EFFICIENT COLD PLATES FOR NEXT GENERATION OF EV BATTERIES

Mubea

Context

- The challenge for automotive suppliers is to **quickly adapt a base design concept to varying requirements** – while maintaining optimality.
- EV battery cooling requires to cover a **wide surface with a minimal pressure drop** (energy saving).

Achievement

- **Very accurate predictions** on aggregated values and fields (see Figures).
- For their "channel" product-lines, the engineers are now **able to optimize designs while getting a real time feedback** on performance.

« WE ARE USING NEURAL CONCEPT SHAPE (NCS) TO ACHIEVE OPTIMAL DESIGN OF KNOWN PRODUCTS WAY FASTER. »

NIKLAS KLINKE, TEAM LEAD, TOOLS & METHODS

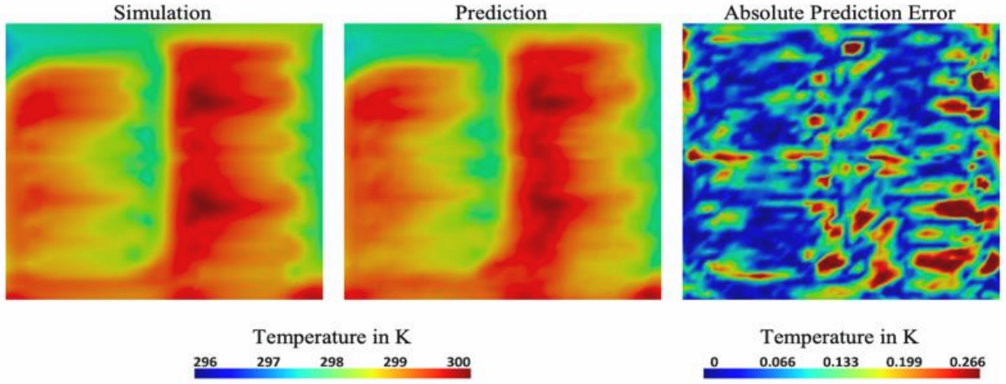


Fig. 6. Prediction of temperature field on the top of the plate T_{top} of Sample B20

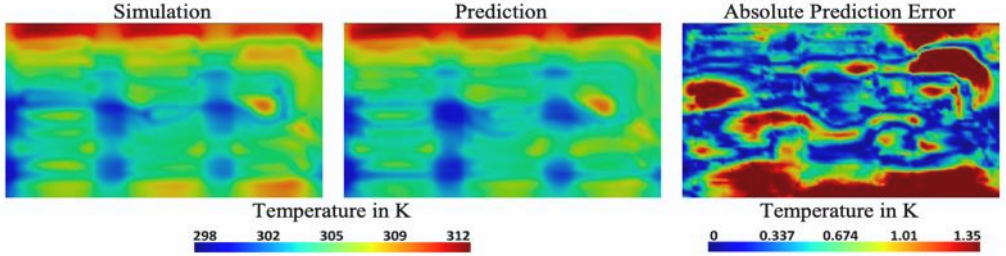


Fig. 7. Prediction of temperature field on the top of the plate T_{top} of sample C14

Image (illustrative): https://www.mubea.com/en/new-body-products
Publication: Dr. Niklas Klinke, Dr. Stefan Buchkremer, Dr. Lutz-Eike Elend, Maksym Kalaidov, Thomas von Tschammer, AI-based performance prediction and its application on the design and simulation of cooling plates for battery electric vehicles, Future Automotive Production Conference Wolfsburg, Germany 17–18th May 2022

Crash simulation of battery housings

Mubea

EVALUATION AND OPTIMIZATION OF CRASH PERFORMANCE FOR BATTERY HOUSING

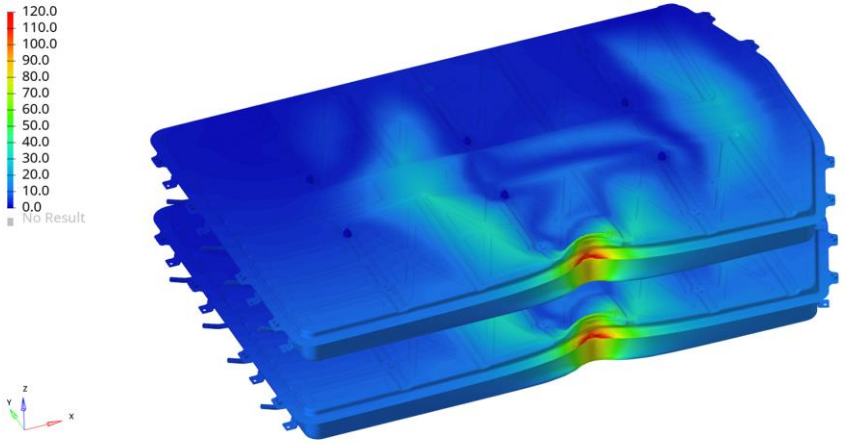
Context

- **Crashworthiness is a key factor for vehicle safety** but is very challenging to characterize efficiently and accurately.
- As it is a highly non-linear problem with many different possible scenarios, **simulation alone cannot provide the full picture** of crashworthiness, let alone optimize it.
- Identifying the most relevant parameters is difficult, which makes it **hard to manually improve the designs**.

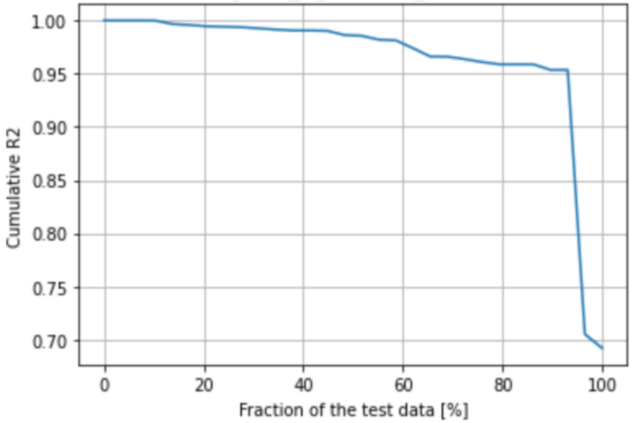
Achievement

- NCS surrogate model is able to predict very accurately the **structural behavior of the housing**, by also predicting **the contact force amplitude and location with the batteries**.
- Engineers can now explore hundreds of design options, to ensure the structural integrity of the battery
- The uncertainty index is used to guide the engineers and improve the accuracy of the model

<1 second	0.98
TIME TO GENERATE A NEW DESIGN AND EVALUATE ITS PERFORMANCE IN NCS	R2 CORRELATION BETWEEN PREDICTED AND LS DYNA SIMULATION



Comparison between the FEA simulation (top) and the NCS prediction (bottom) for the displacement magnitude, on a test geometry. The prediction was done on multiple time steps.



Using the uncertainty feature from Neural Concept Shape, the test samples are sorted by the uncertainty metric given by the model, from lowest to highest.

Aerospace Heat Exchanger Optimization

PHYSICSX

OPTIMIZED HEAT EXCHANGERS FOR **AEROSPACE APPLICATIONS**

Context

- In aerospace applications, the final product performance is the focus. **Efficiency and optimality is key.**
- **Engineers want to evaluate many design concepts** over several months/years and leverage on the experience and data from previous iterations.

Achievement

- **Very accurate predictions** on aggregated values and fields: The engineers can evaluate thousands of designs per day.
- Optimized geometries using morphing of different concepts (pins and fins)

>99%	+1.5%	Constant
ACCURACY FOR MAXIMUM TEMPERATURE AND PRESSURE DROP PREDICTIONS	INCREASE IN THERMAL EFFICIENCY COMPARED TO PREVIOUS DESIGN WORKFLOW	PRESSURE DROP DID NOT DETERIORATE COMPARED TO PREVIOUS DESIGNS

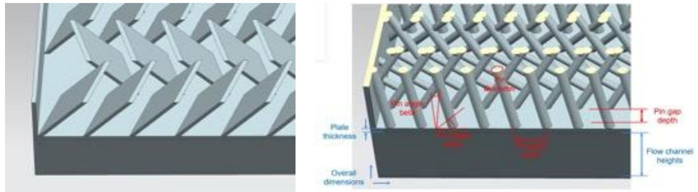
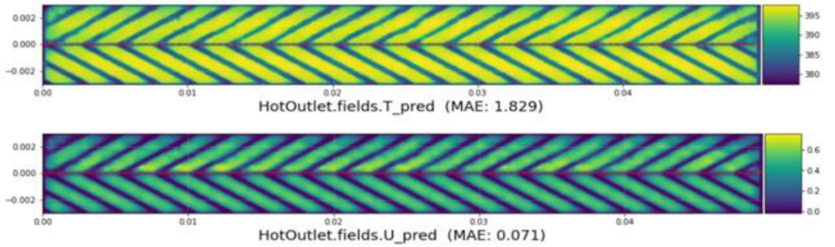


Figure 1: *Examples of heat-exchanger geometries with different topologies*



[Top]: Prediction (top) and CFD (bottom) of the Temperature at the hot (oil) outlet.
[Bottom]: Prediction (top) and CFD (Bottom) of the Velocity magnitude at the hot (oil) outlet.

Figure 2: *Predictions vs GT on the outer surface of the heat-exchanger*

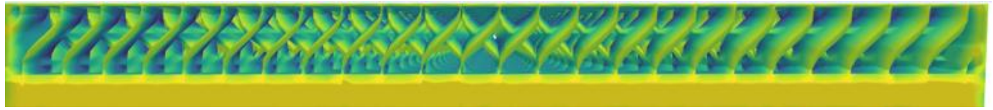


Figure 3: *Optimized geometry using morphing techniques*

Latent Thermal Energy Storage

SCIENTISTS ARE ALSO USING DEEP-LEARNING TO SOLVE THEIR RESEARCH PROBLEMS

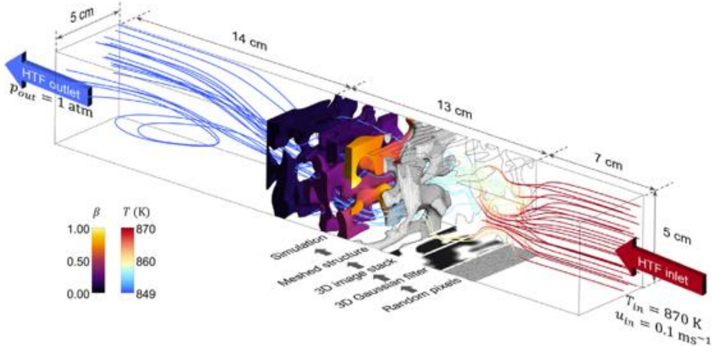
Context

- Latent heat thermal energy storage with metallic alloy phase change is a **new promising technology**.
- Macro-porous latent heat storage can enhance the convective heat transfer.
- Researchers at **EPFL's Renewable Energy Science and Engineering's Lab** are looking for solutions to exploit the potentially very wide design space.

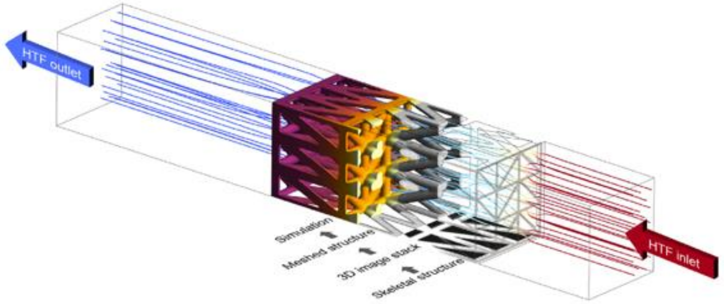
Achievement

- A surrogate model allowed **to precisely assess the performance of designs within an infinite dimensional design space**.
- A paper demonstrating the use Neural Concept's approach in the domain was published: GCNN Characterization of Macro-Porous Latent energy storage – ASME Journal of Heat and Mass Transfer - 2022

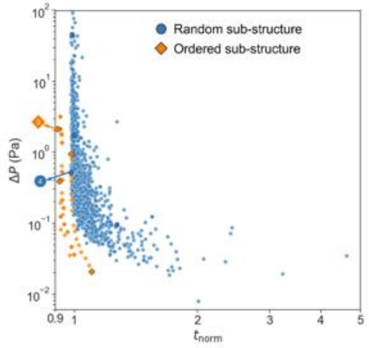
99.1%	75.6%	90.4%
R2 CORRELATION BTW. PREDICTED AND TRUE MELTING TEMPERATURE FIELDS	R2 CORRELATION BTW. PREDICTED AND TRUE NORMALIZED TEMPERATURE FIELDS	R2 CORRELATION BTW. PREDICTED AND TRUE DIFFERENTIAL PRESSURE FIELDS



Prediction of flow for random sub-structures.



Prediction of flow for ordered sub-structures.



Before optimization ordered sub-structures perform better than random ones. Optimization was not performed in this study.

External Aerodynamics (1/2)

USE ENGINEERING INTELLIGENCE TO EXPLORE DIFFERENT DESIGN CONCEPTS

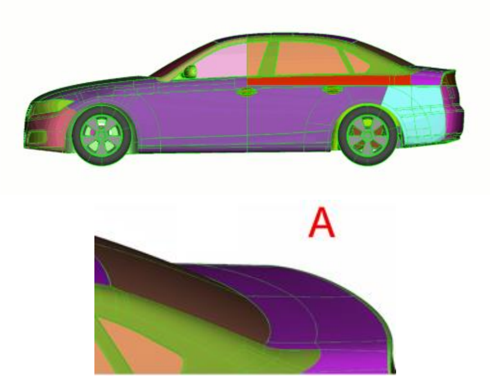
Context

- Aerodynamic performance is **critical due to vehicle energy efficiency**. OEMs must pay enormous fine if they don't meet regulation targets
- As an example, in EU, OEMs must pay €95 per car sold for every gCO2/km above the regulation.
- External Aero **simulations are heavy and expensive**, which mean designers only get very limited feedback on their designs

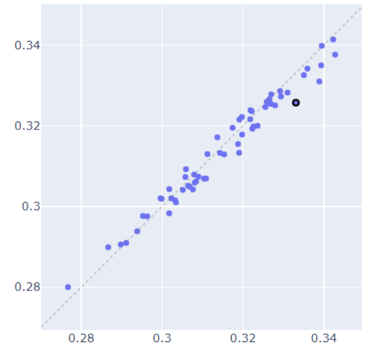
Achievement

- A model was trained to **predict the pressure field and drag coefficient** of different car designs.
- The model was challenged with different design concepts (no spoiler vs spoiler), and it was shown that a **very limited number of new new simulations** was needed to generalize the model to both these concepts.

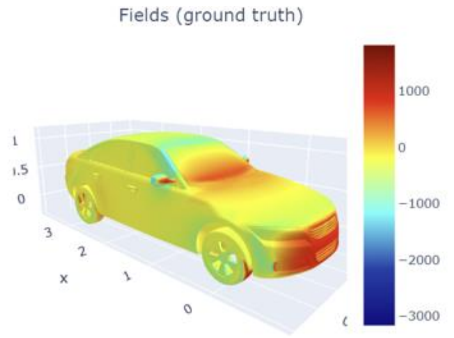
x100	>96%	3D output
INCREASE IN THE NUMBER OF DESIGNS THAT CAN BE EXPLORED	CORRELATION BETWEEN PREDICTED AND SIMULATED DRAG COEFFICIENTS	IN ADDITIOIN TO DRAG COEFFICIENT, THE MODEL CAN PREDICT 3D FIELDS SUCH AS PRESSURE AROUND THE CAR



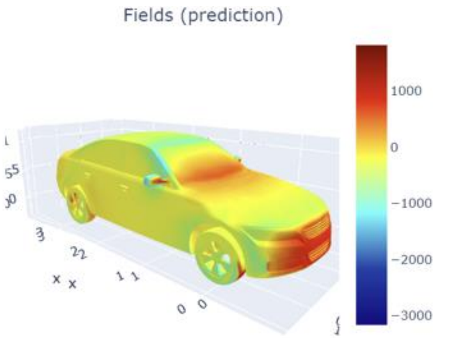
Dataset consists of 400 variants of baseline A (no spoiler)



Predicted vs Simulated drag coefficient on 50 test samples



Simulated pressure field around the car



Predicted pressure field around the car

E-motor optimization

INCREASING THE PERFORMANCE AND DURABILITY OF ELECTRIC POWERTRAIN

Context

- Vehicle electrification comes with many **new engineering challenges**
- E-motor design is a complex process with **strict structural and electromagnetics requirements**.
- **Flux barrier shape** has a significant influence on E-motor performances.

Achievement

- NCS was used to **predict accurately the structural (safety factor) and electromagnetic performances** (torque & ripple).
- NCS design module was used to explore **innovative flux barrier designs**.
- Final design clearly **outperforming standard, parametric geometries**.

PARTNERSHIP WITH MOTORCAD TO ENSURE SMOOTH OPTIMIZATION AND RETRAINING LOOPS

Figure 1: *Prediction of internal stress*

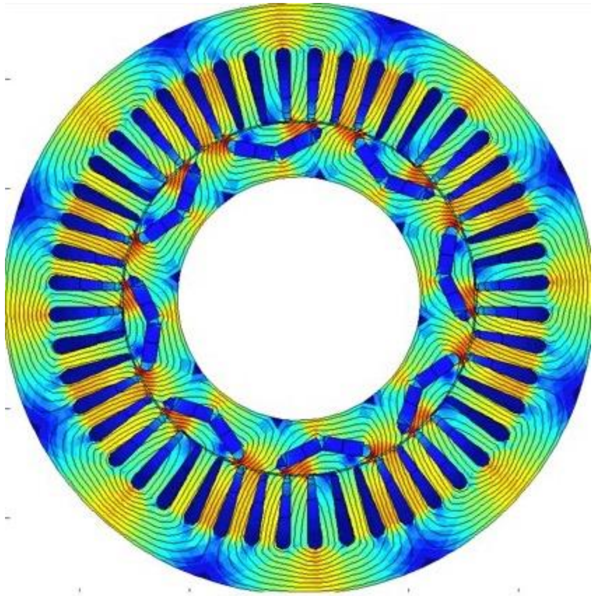
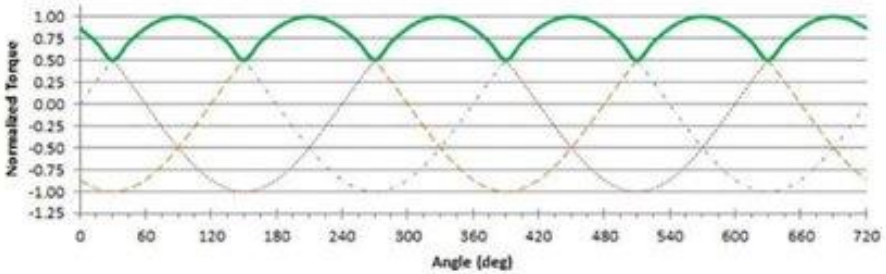


Figure 2: *Torque prediction vs angle*



HVAC Design for automotive OEM

ENHANCING THERMAL SYSTEMS USING FAST AND ACCURATE DEEP LEARNING SURROGATES

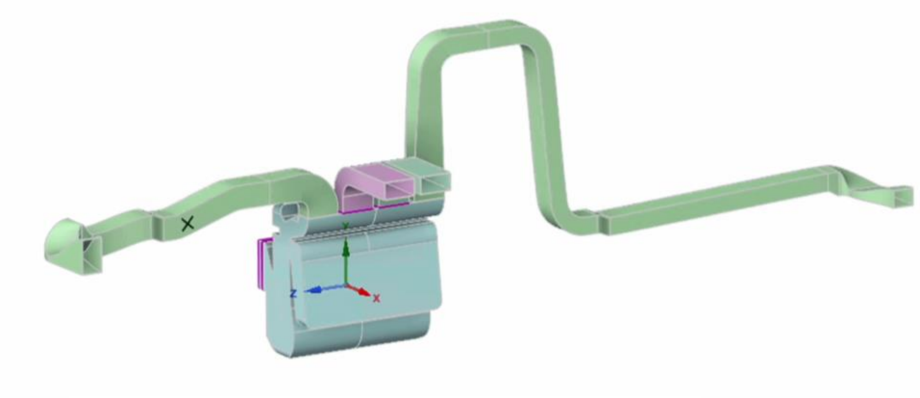
Context

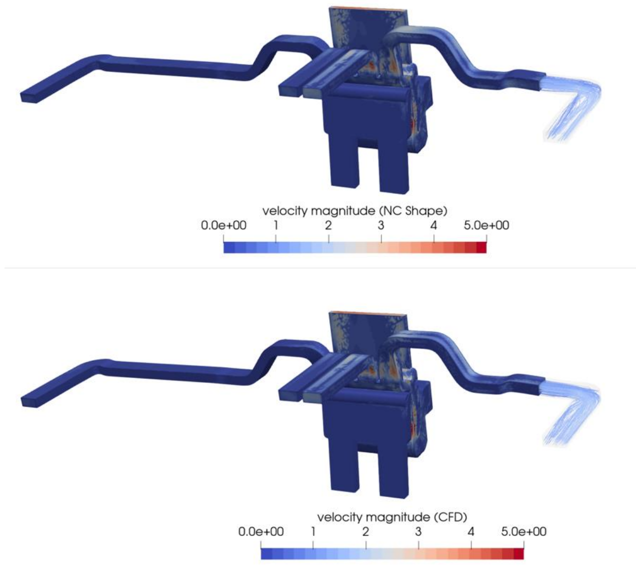
- Thermal systems in vehicles are responsible for **>3% of the total fuel consumptions**
- CFD computation is a major bottleneck in the development of HVACs → **100 to 1000s CFD results for one development.**
- Automotive suppliers are competing already in RFQ phases, doing most of product development pre-order.

Achievement

- Physics performance feedback **directly into the designer's interface**.
- HVAC designers are able **minimize pressure loss** and ensure uniformity on outlets.

100x	2x	1.5x
INCREASE IN THE NUMBER OF DESIGNS THAT CAN BE EXPLORED	FASTER IN REQUEST FOR QUOTATION PHASE	EXPECTED PROJECT WIN-RATE





FEA simulation (bottom) and the NCS prediction (top) for the velocity field on a test geometry.

Hull hydrodynamics

BETTER HYDRODYNAMICS WITH FAST PERFORMANCE PREDICTION

Context

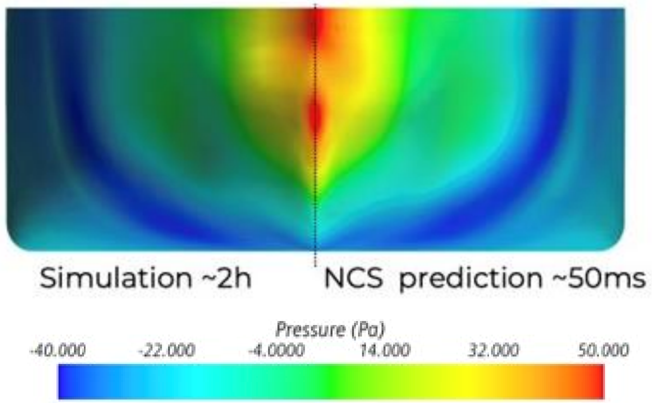
- Several design concepts and many different topologies can be created for the hull, each with **very different behaviors**
- By nature, **hard to parametrize** due to non-conventional designs and large design space

Achievement

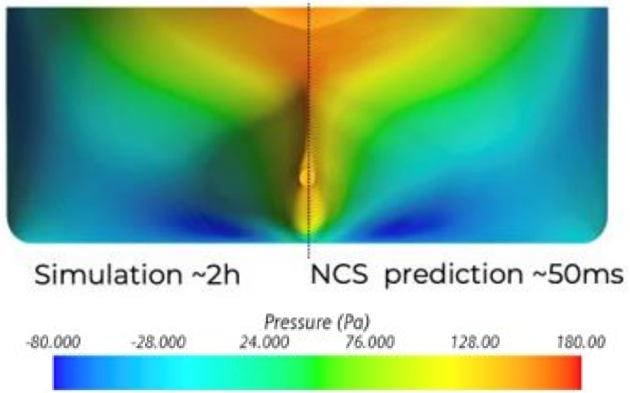
- **Very accurate predictions** on aggregated values and fields
- The engineers are now able to interact in **real time** with the designs
- Slight improvement in the performance is a **game changer**
- **Optimization on the 3D geometry directly** is now used to reach better performing designs

50ms	90.6%	99.4%
TIME TO GENERATE NCS PRESSURE PREDICTION (2H BEFORE)	AVG L1 ERROR ON THE PRESSURE FORCE COEFFICIENT ACROSS THE HULL	AVG L1 ERROR ON THE SHEAR FORCE COEFFICIENT ACROSS THE HULL

Comparison of the pressure distribution between StarCCM+ simulation and NCS prediction (example 1)



Comparison of the pressure distribution between StarCCM+ simulation and NCS prediction (example 2)



Injection molding: warpage predictions

PREDICTING THE MANUFACTURABILITY EARLY IN THE DESIGN PROCESS

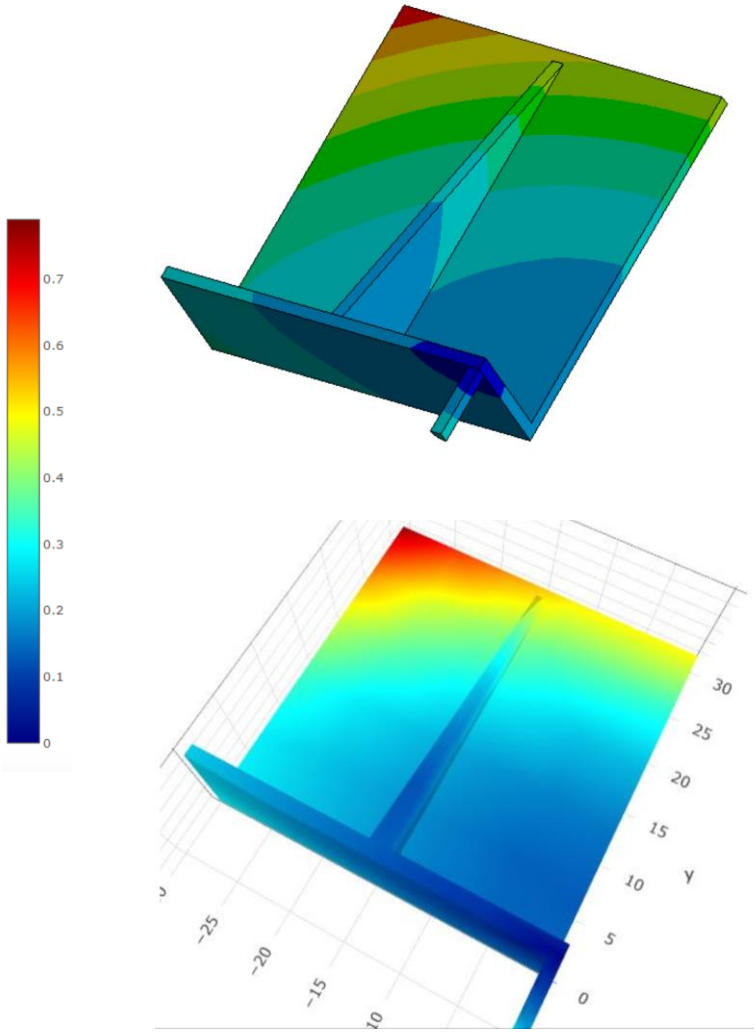
Context

- The process of **warpage deformation is complicated**.
- Fluid shrinkage (varying pressure during process) and coefficient of thermal expansion from eject temperature to room temperature must be considered.
- Simulations with PlanetsX (plugged in Ansys Workbench) are long and are therefore **only performed at the end of the design process**.

Achievement

- **A surrogate model was trained in NCS** to predict the warp deformation of different parts, along with the maximum injection pressure.
- **The accurate surrogate model can be frontloaded to the designer** and used early in the design process to ensure manufacturability criteria

99.8%	99.2%	x10
ACCURACY OF MAX INJECTION PRESSURE PREDICTIONS	ACCURACY OF MAX DEFORMATION PREDICTION	NUMBER OF MANUFACTURABILITY VERIFICATIONS DURING THE DESIGN PROCESS



Groundtruth (FEA simulation) of deformation field (only contour)

NCS prediction of the deformation.

Rubber switch large deformation analysis

PREDICTING HIGHLY NON-LINEAR DEFORMATIONS ACCURATELY USING NCS

Context

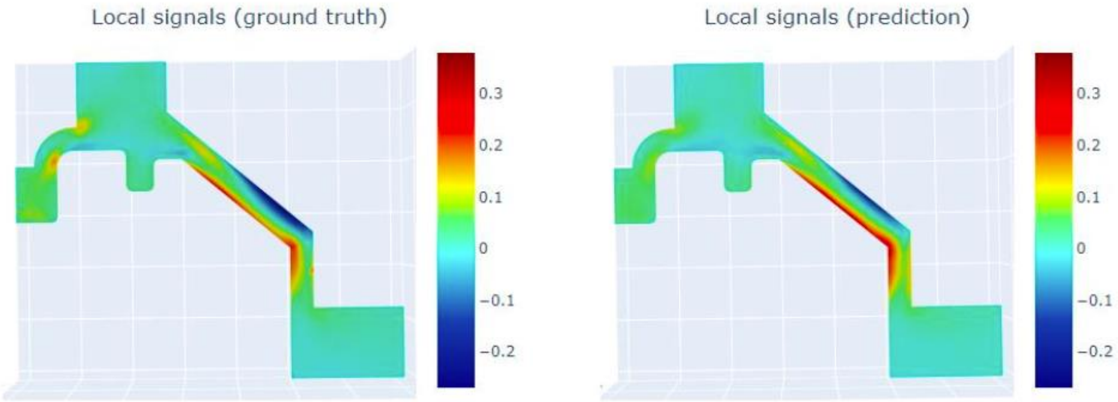
- Rubber switches are the essential element in most keyboards.
- Made of rubber, they sustain **large deformation with non-linear contacts and frictions**.
- Accurately predicting their deformation, internal stress and feeling curve (displacement vs reaction force) is **crucial for the keyboard's durability and usage comfort**.

Achievement

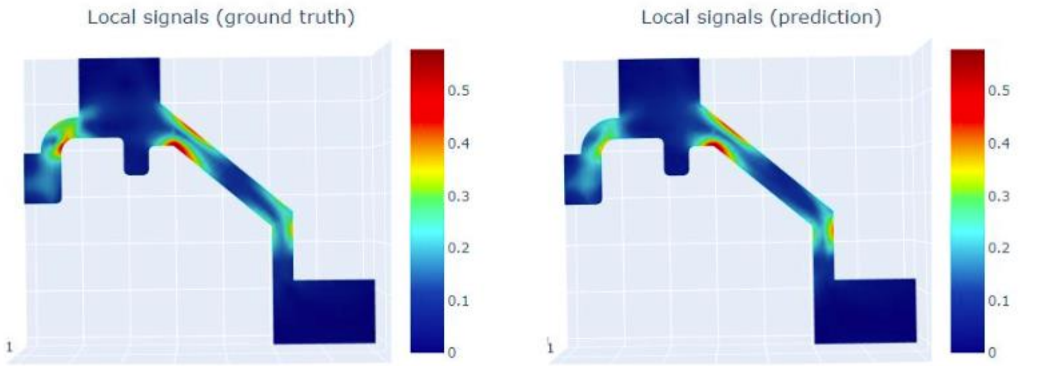
- **A surrogate model was trained in NCS** to predict the large deformation, internal stress and feeling curve of different rubber switch designs.
- **Fast and accurate analysis of non-linear phenomena** such as buckling and contact behaviors can now be performed.

" BY **LEARNING A LARGE AMOUNT OF DATA**, WE WERE ABLE TO CONSTRUCT A PREDICTIVE MODEL FOR **HIGHLY NON-LINEAR PHENOMENA** SUCH AS BUCKLING AND CONTACT ANALYSIS OF RUBBER "

Displacement field inside the stitch: simulation vs prediction



Von Mises stress field inside the stitch: simulation vs prediction



Crashworthiness optimization

EVALUATION AND OPTIMIZATION OF CRASH PERFORMANCE FOR NOVEL VEHICLE CONCEPTS

Context

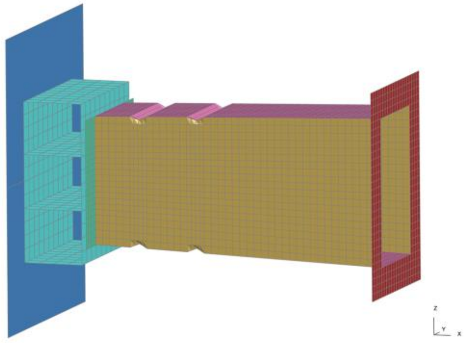
- **Crashworthiness is a key factor for vehicle safety** but is very challenging to characterize efficiently and accurately.
- As it is a highly non-linear problem with many different possible scenarios, **simulation alone cannot provide the full picture** of crashworthiness, let alone optimize it.
- Identifying the most relevant parameters is difficult, which makes it **hard to manually improve the designs**.

Achievement

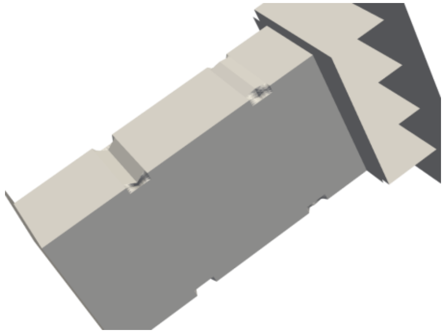
- **A surrogate model was trained in NCS** to predict in real time the main KPIs: specific energy absorption (SEA), intrusion (crushing length) and crushing force efficiency.
- NC Design module used to create instantly new trigger designs.
- **Fast and effective optimization campaigns** resulting in substantial performance improvement (5000 designs evaluated in 6 hours).

<1 second	98.9%	10%
TIME TO GENERATE A NEW DESIGN AND EVALUATE ITS PERFORMANCE IN NCS	R2 CORRELATION BETWEEN PREDICTED AND SIMULATED SEA	PERFORMANCE IMPROVEMENT OBTAINED WITH THE OPTIMIZATION

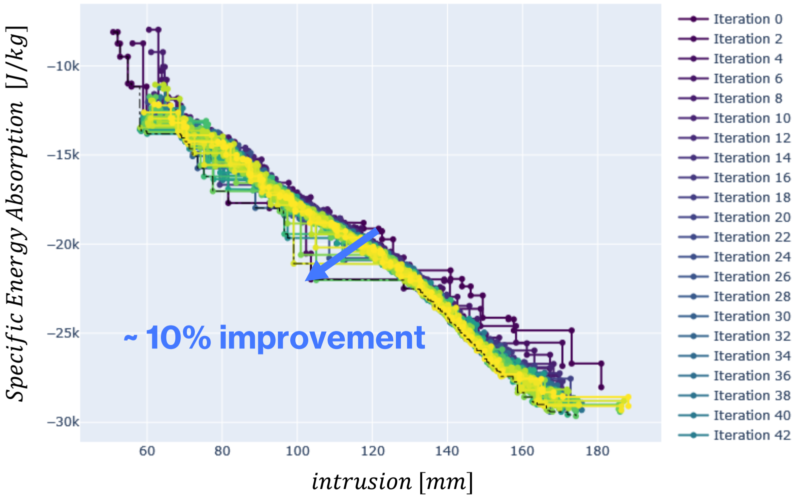
Typical crash-box design



NC design module to create non-parametric innovative designs



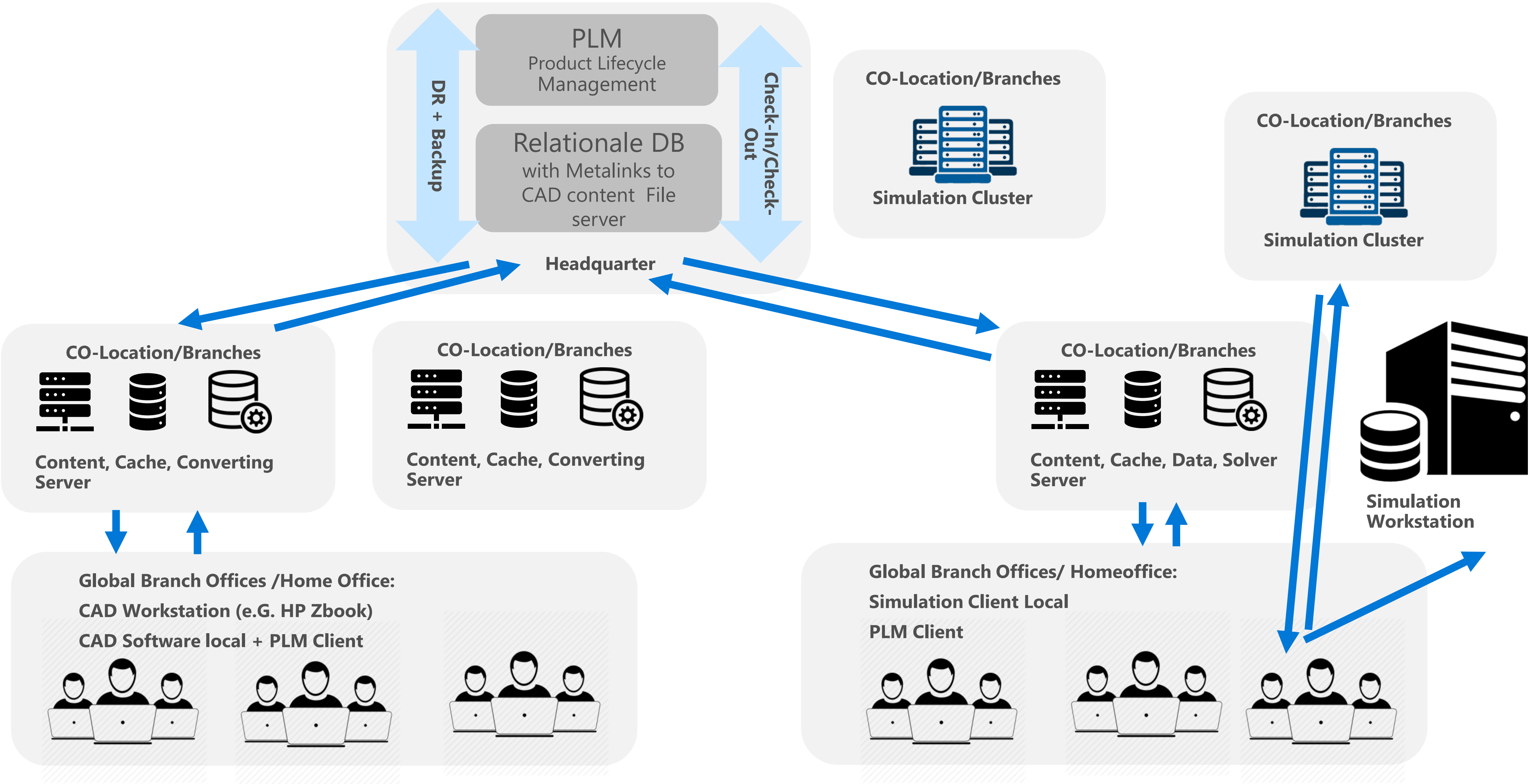
Optimization: Pairwise Pareto fronts at each iteration



PLM & CAD on Azure

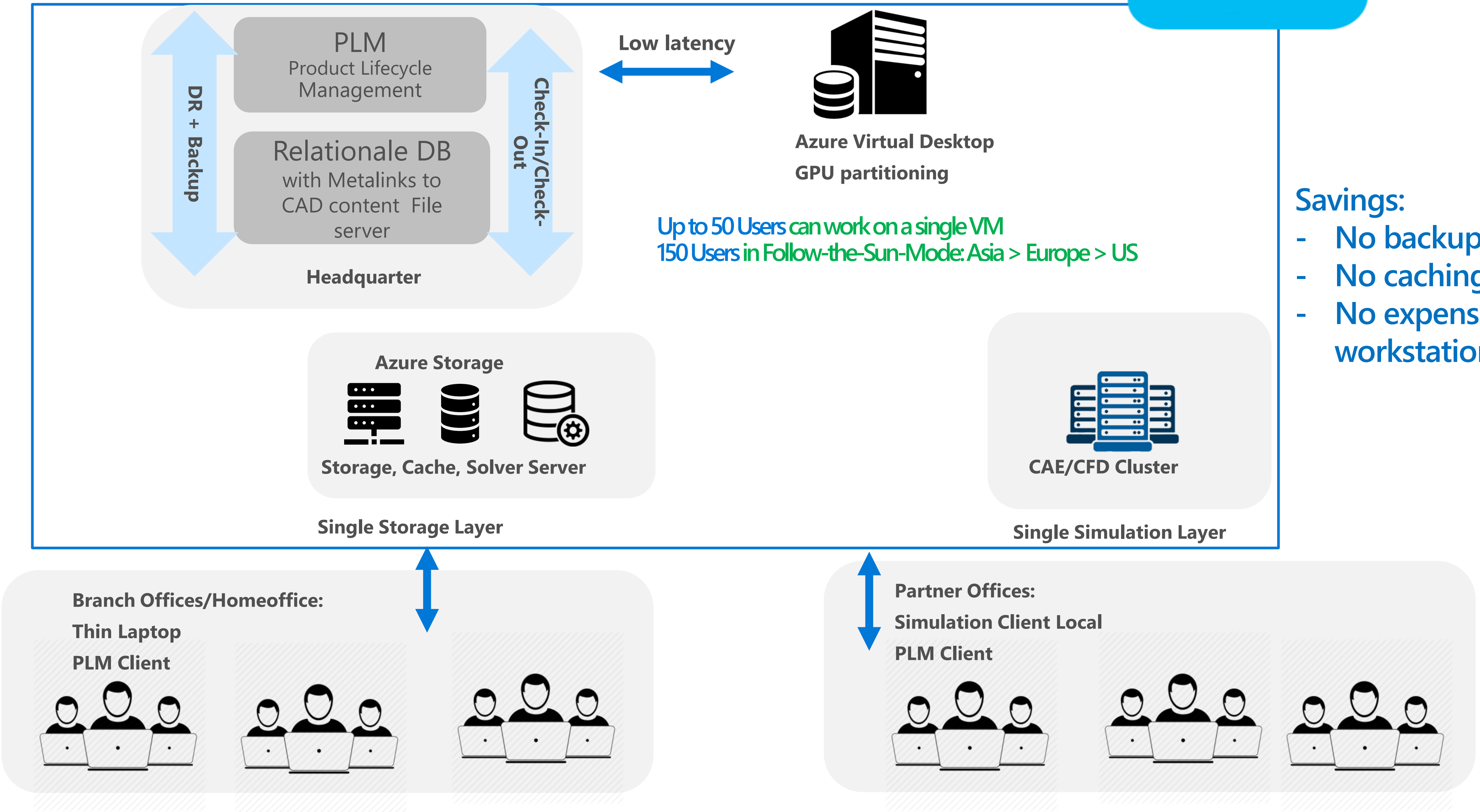
PLM and CAE Architecture in Large Organisations

Architecture today can be complex



PLM and CAE Architecture in Azure

Simpler Architecture provides Saving Potential



Savings:

- No backup servers
- No caching servers
- No expensive workstations

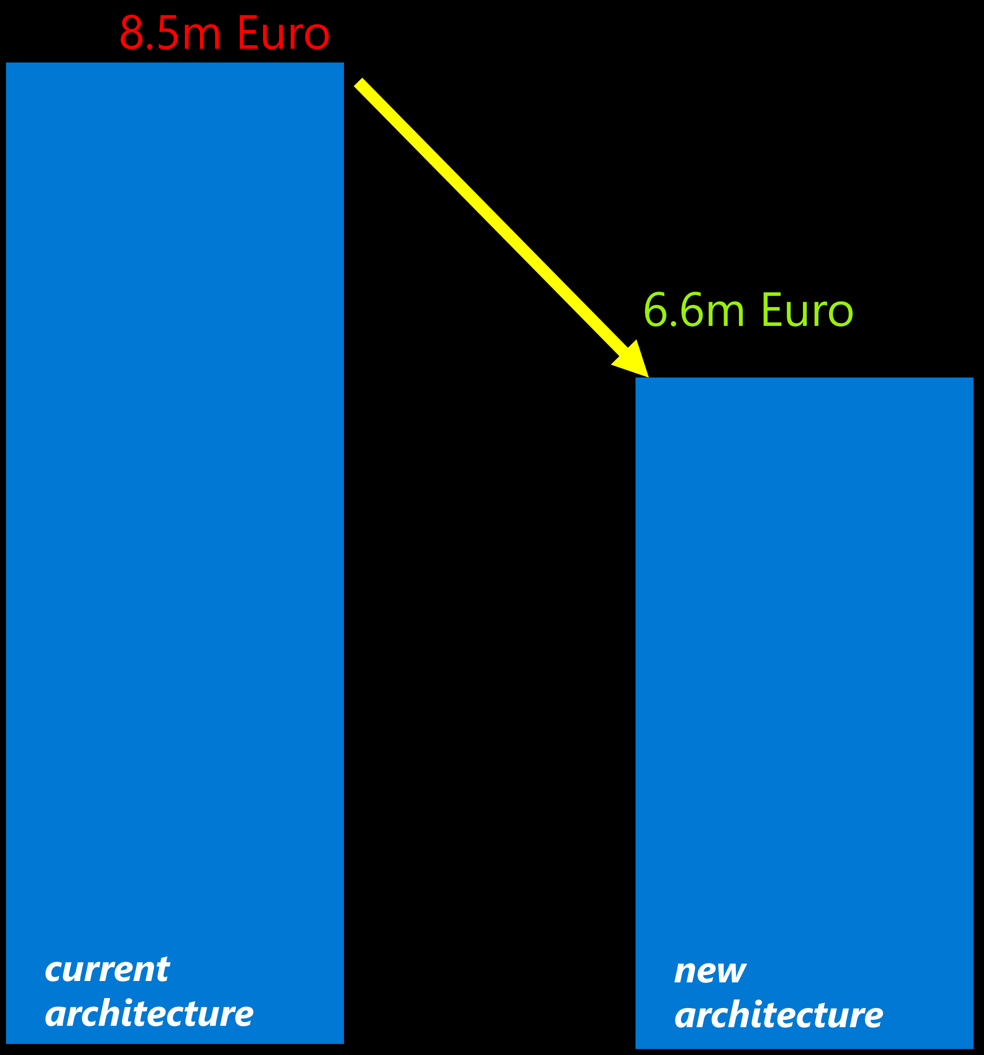
Customer Example: Calculation based on 5 Year Depreciation
Total Cost Comparison

Example 3000 CAD Engineers

€: WAN/SD-WAN
€: Hardware
€: Operational
= 1m Euro (approx.)

7.5m Euro for 3000
CAD Workstations

Hardware over 5
Years
=



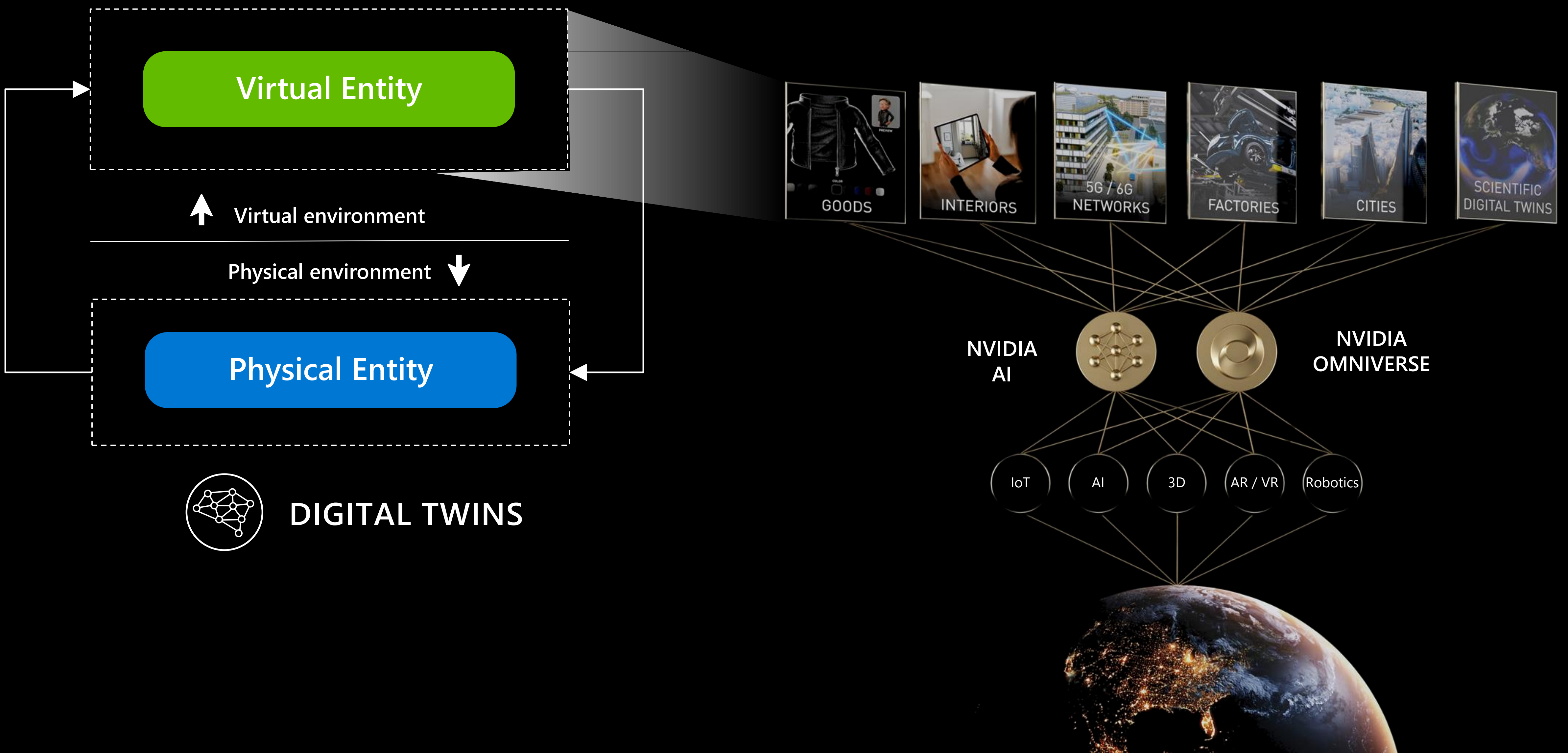
€: 3.6m (ca.) Azure
Infra and Traffic

€ 3m for 3000
CAD thin clients

Hardware over 5
Years

Industrial Metaverse with NVIDIA Omniverse

Metaverse as Digital Twins Virtual Environment



NVIDIA Omniverse Platform

Centralized and collaborative data management

USD as the Open-Source format.

Nucleus is the centralized server exposing data APIs to all users/applications

NVIDIA Omniverse Apps

All NVIDIA Omniverse applications are meant to be client side on our GPU accelerated workstations.

They can be installed manually by users or can be IT managed on user profile

NVIDIA Omniverse

NVIDIA Omniverse is meant to be connected to the standard tools already used in industry, with USD as interface.

It can be interfaced with more than 50 applications

NVIDIA Omniverse Connectors

Compute intensive tasks including rendering / simulation / AI training can be offloaded to a GPU cluster through Omniverse Farm.

NVIDIA Omniverse Farm

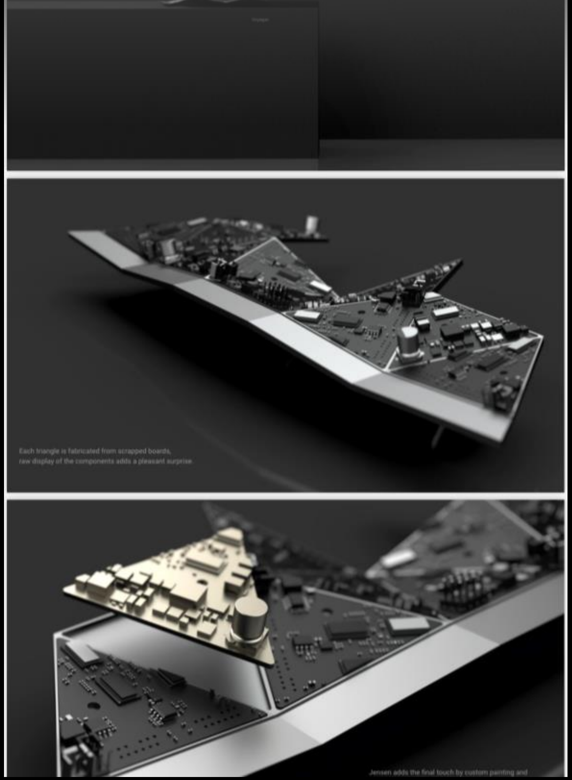
Use cases
Collaborative design

COLLABORATIVE 3D WORKFLOWS

Less Back-and-Forth, Time Lost

FULL FIDELITY CONCEPT REVIEW

Faster Review Cycles, Time to Market

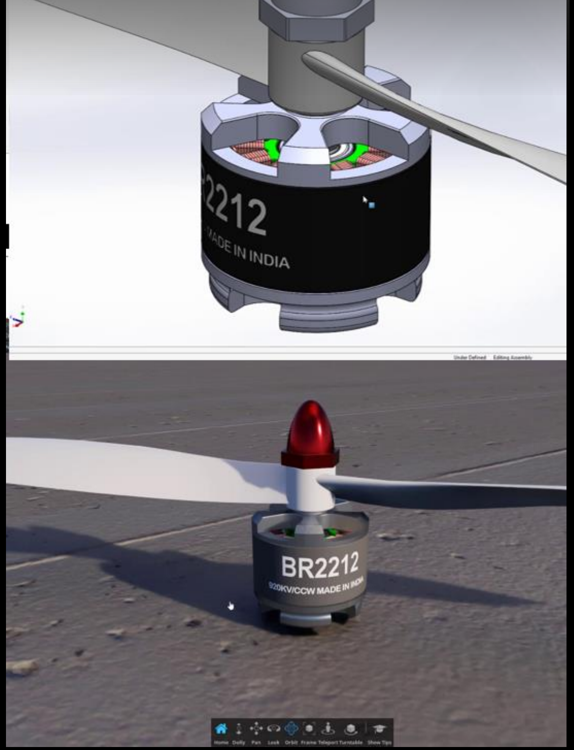


TEST WITHOUT PRODUCING

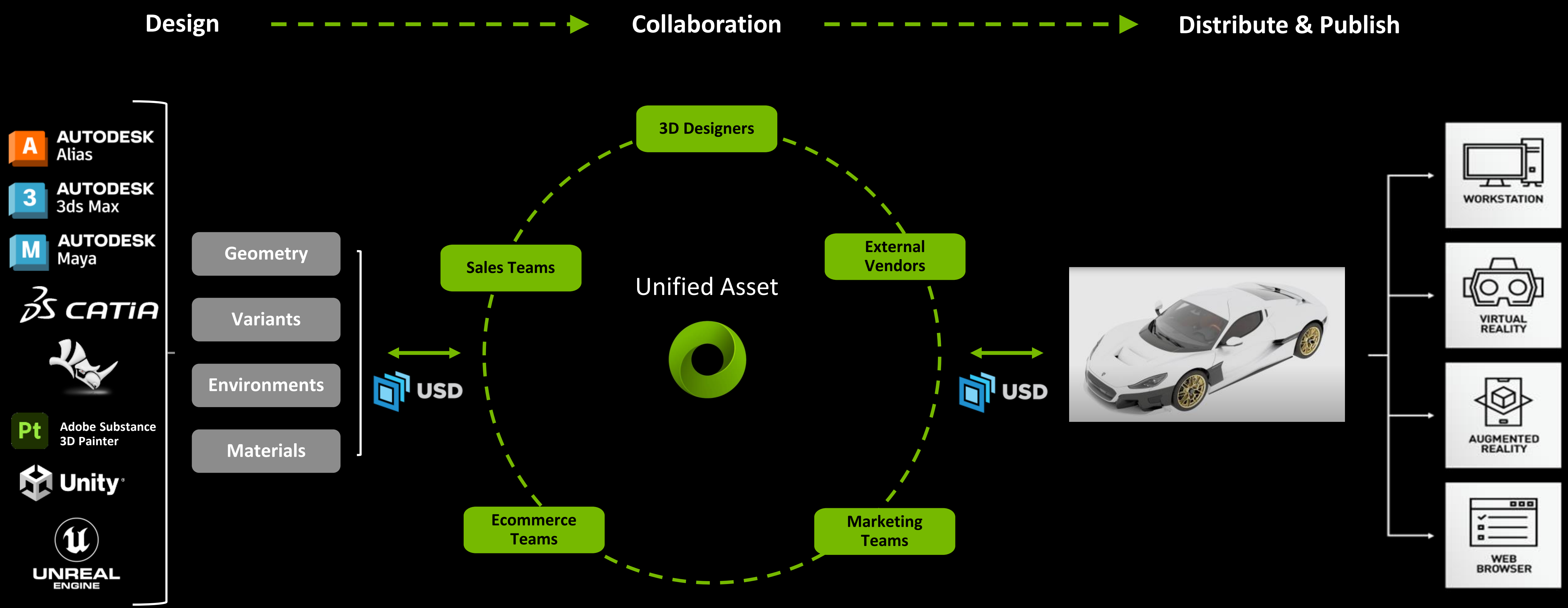
Online Consumer Testing with Physically Accurate Visualizations

REDUCES ITERATIVE WASTE

Single Source of Truth Files, CAD Accurate



Use cases
Collaborative design



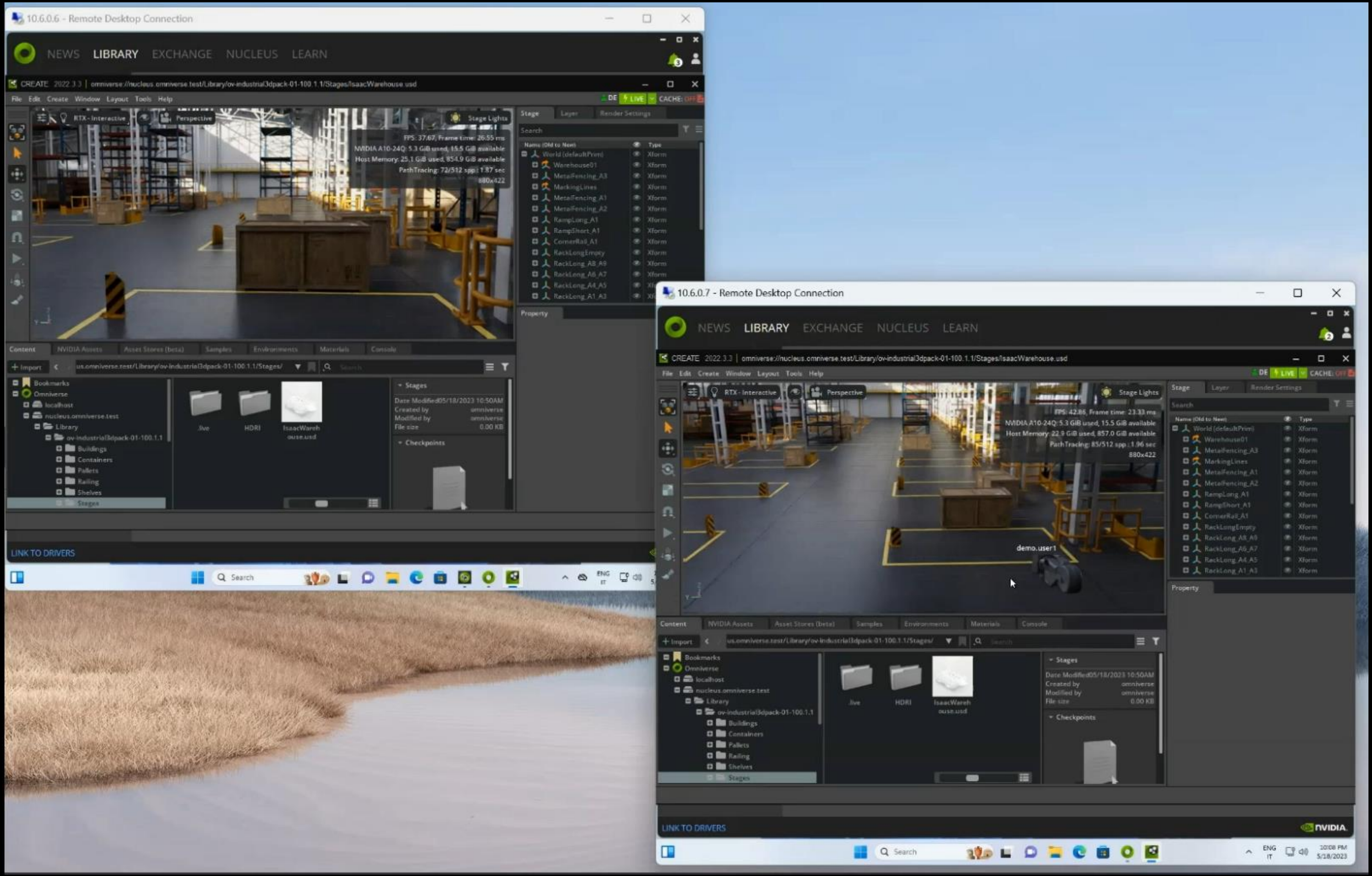
Use cases
Collaborative design

Live High-Fidelity RTX rendering

NVIDIA RTX technology allows to render with high-fidelity in real time the design scene

Visualization through HMD or Mobile

CloudXR allows visualization using VR headsets or mobile devices, targeting also AR applications



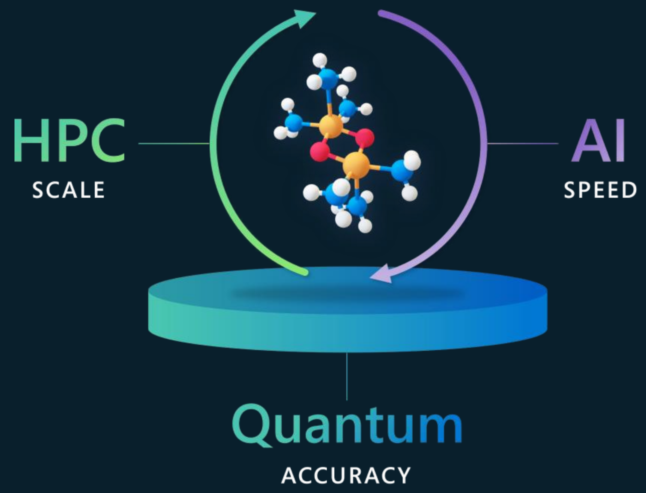
Azure Quantum Elements

Azure Quantum Elements, accelerating scientific discovery

Scale
Hyperscale discovery

Speed
Expand research horizons

Accuracy
Model complex chemistry accurately



Bringing products to market faster

Enabling a faster innovation cycle with Copilot in Azure Quantum

Try the Copilot on quantum.microsoft.com

AZURE QUANTUM

Elements

AkzoNobel

SCGC

aspentech

1910 Genetics

BASF
We create chemistry

JM Johnson Matthey
Inspiring science, enhancing life

Azure Quantum Elements Apps Files Jobs Clusters Interactive Apps Monitoring My Interactive Sessions Logout

AZURE QUANTUM

Elements

Pinned Apps

LAAMPS	QuantumESPRESSO	GROMACS	PySCF
AiiDA	AutoRXN	Heron	AutoCAS

Applied Reinforcement Learning

InstaDeep: EMEA Leader in AI

Founded in 2014, HQ in London

10 Offices (EU, Africa, US)

250+ Employees

Decision-Making products: delivering AI-driven efficiencies for advanced enterprise customers

Three Pillars Differentiate InstaDeep

Solving complex challenges for top tier international customers

BIONTECH, Google, TOTAL, DB BAHN, BNP PARIBAS, salomon

Access to top Talent. Partner with leading Universities

Massachusetts Institute of Technology, University of Michigan, University of Oxford, UCL, Sciences Sorbonne Université, École Polytechnique

Cutting-edge AI Research Joint R&D work with elite partners

DeepMind, Google Research, NVIDIA, intel, ICLR, Neural Information Processing Systems

* 5 InstaDeepers out of 171 Google ML Dev Experts globally

Google Cloud Partner | NVIDIA Elite Partner | AI 100 CBINSIGHTS 2020 | AI 100 CBINSIGHTS 2021 | AI 100 CBINSIGHTS 2022 | intel. AI Builders | InstaDeep raises $100M closes Series B round

InstaDeep is Europe’s AI Leader with Focus on Logistics

Technology

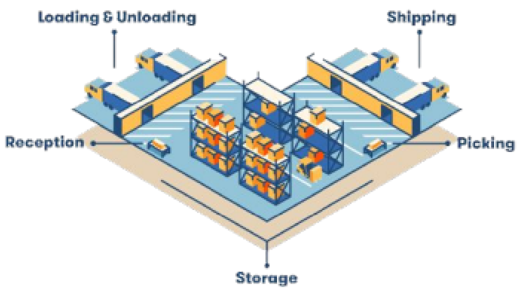
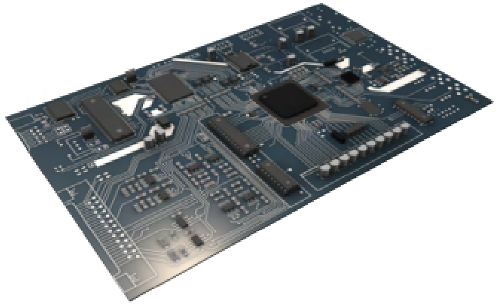
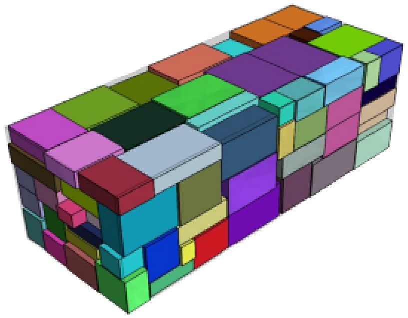
BioNTech to Buy AI Expert InstaDeep for Up to £562 Million

- German company will pay £362 million upfront in cash, shares
- Deal will help BioNTech use AI in areas like manufacturing

Highest Valuation of any pure-play AI Startup Exit in History
Boosting Commitment towards European Manufacturing and Logistics

History of solving complex problems with Decision Making AI

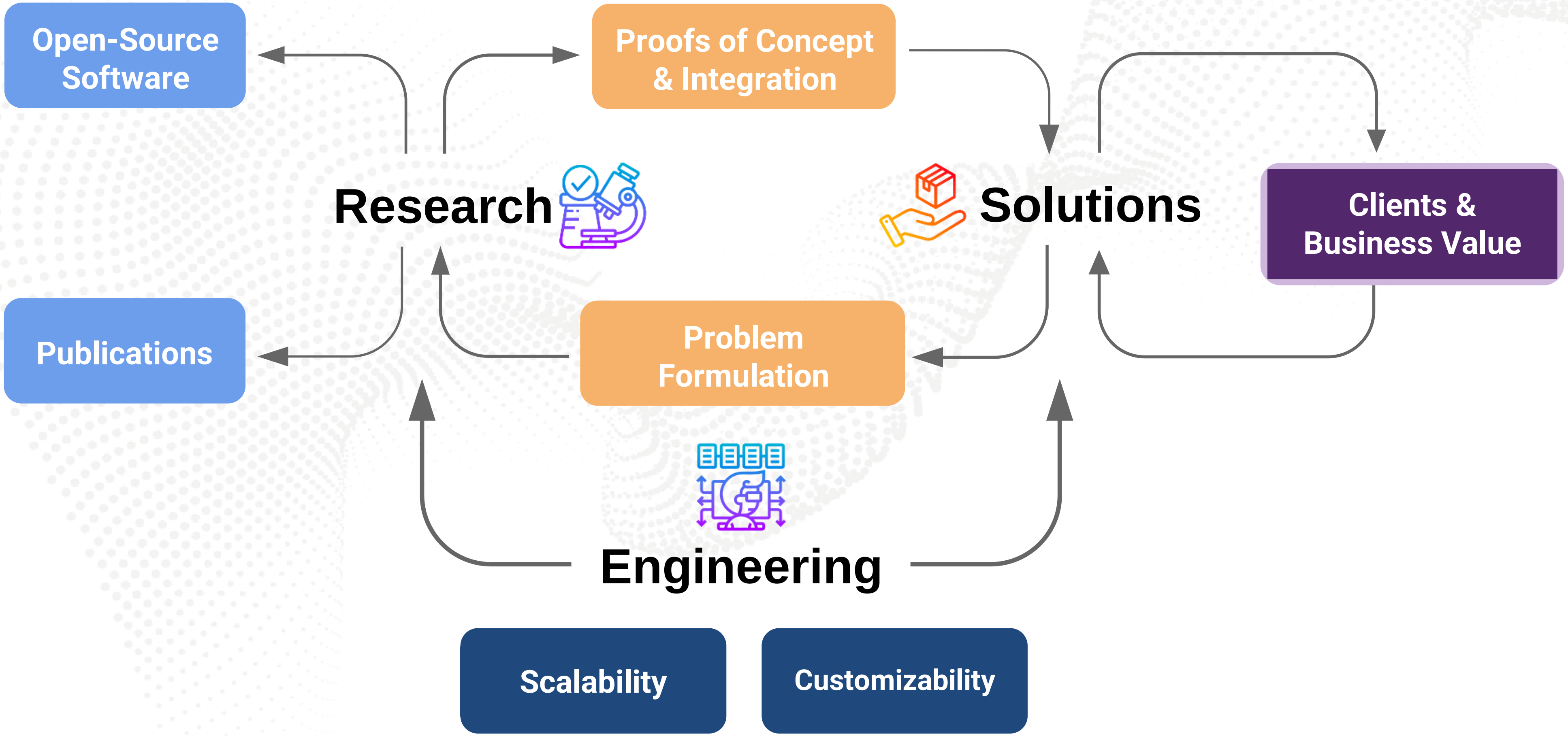
InstaDeep's Decision Making Technology + Domain Expertise = Scalable Products

AI **Warehouse** (Logistics/Supply Chain)	Deep**PCB**™ (Hardware/IOT)	Deep**Pack**™ (Logistics/Supply Chain)	Deep**Rail**™ (Capacity & Traffic Management)
			
Optimize item picking and placement strategies	Route complex printed circuit boards in <24 hours	Pack container logistics items more efficiently	Optimize train scheduling and fleet management
Cut Opex by reducing the overall number of picks and time per pick needed	**Accelerate electronics product development cycles and time-to-market**	**Cut OPEX with faster load planning, while saving load volume and shipment cost**	**Cut OPEX and service delays, reduce CAPEX for infrastructure development**

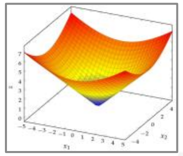
InstaDeep™

End-to-End Differentiated Expertise in AI

A strong **link** between ***fundamental*** and ***applied* research** is at the **core** of InstaDeep's **DNA,** while **AI engineering** and **HPC** skills support scalable and custom deliveries for end clients.

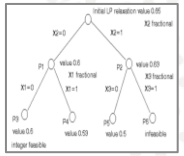


Limitations of Classical Approaches to Optimisation



Optimisation Techniques

- ✔ Guaranteed optimality
- ✖ Poor scalability for large instances



Search Algorithms

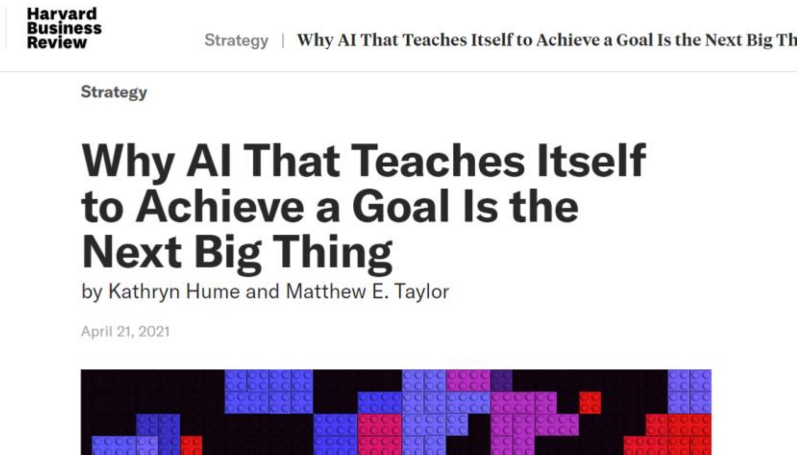
- ✔ Fast and scalable
- ✔ No specific heuristic or evaluation function
- ✖ No guarantee of short-term optimality

Metaheuristics

- ✔ Fast and scalable
- ✖ Tailor-made to a specific problem
- ✖ No guarantees of optimality

Common Flaw: No Learning!

- **Do not** leverage past computations.
- **Unable** to generalize to **unseen** instances
- Might have to be **redesigned** if constraints **change**


Harvard Business Review

Strategy | Why AI That Teaches Itself to Achieve a Goal Is the Next Big Th

Strategy

Why AI That Teaches Itself to Achieve a Goal Is the Next Big Thing

by Kathryn Hume and Matthew E. Taylor

April 21, 2021

AI/RL Optimisations are more Scalable, Flexible & Generalisable

Sequential Decision Optimization

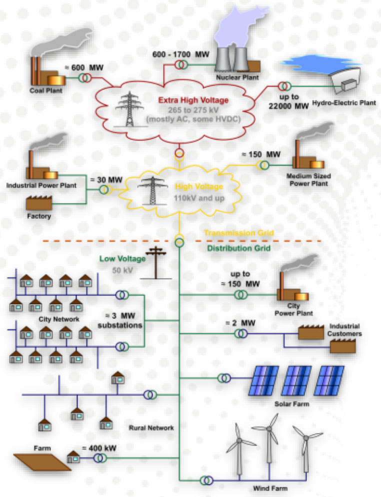
- Your goal is to find an **optimal** behavior
- A sequence of decisions
- Each decision affects future decisions
- Delayed conséquences

RL approach in a nutshell

Learning from **trials** (successes and failures) to optimize a decision making process robust to variations

RL in practice

Agent



Virtual Environment
- simulation
- digital-twin

Benefits of RL for decision-making

- Approximate optimal solutions for complex problems with flexible constraint management
- No need for (optimal) training data
- Logistics ops are subject to change and uncertainty - RL agents can adapt to previously unknown situations
- Real-time, proactive decision-making in highly uncertain and dynamic environments

InstaDeep™

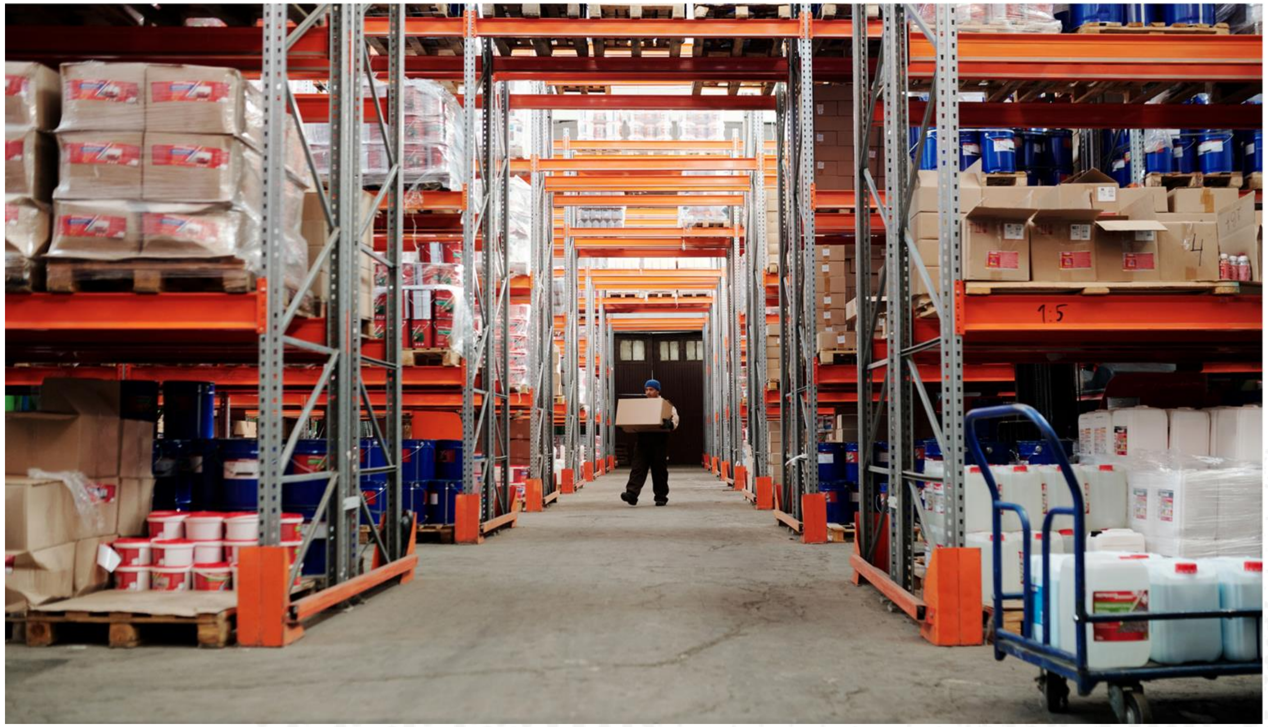
SIEMENS

Microsoft

AI Warehousing:

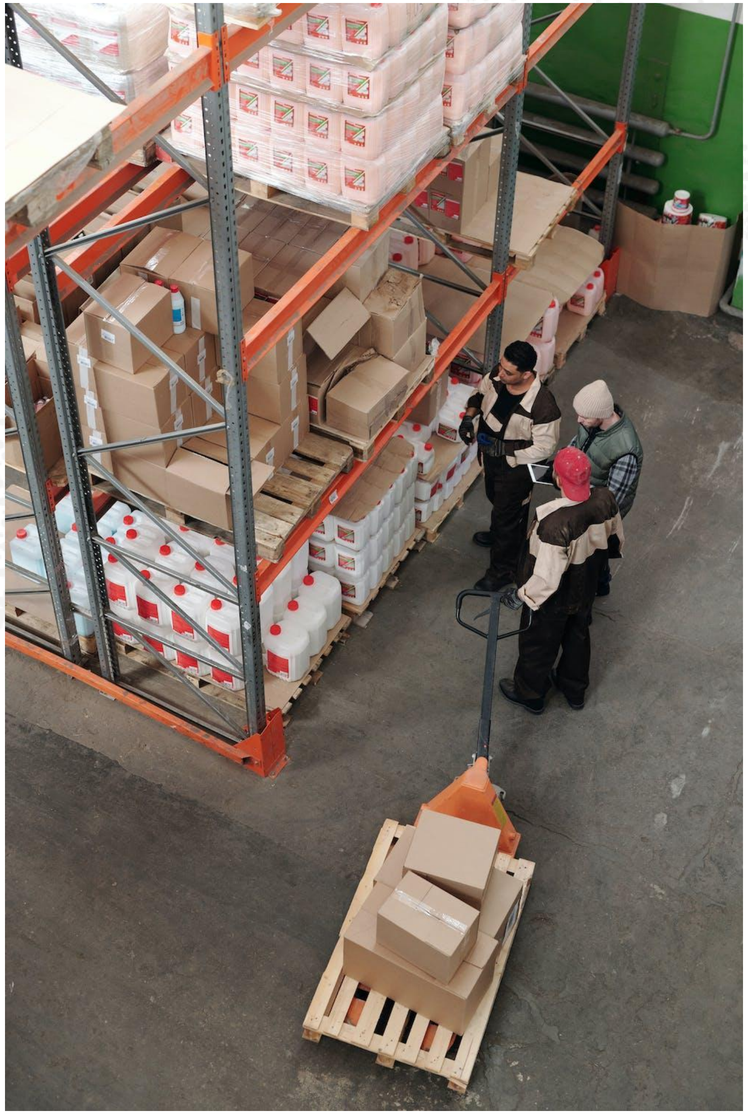
Flexible Warehouse Ops Optimisation using Reinforcement Learning

Problem: Where, When, and What to Pick in a Dynamic World?



Goal Phase 1

A solution that both **minimizes the average picking time to fulfill a delivery request** in a simplified warehouse and it is **quickly adaptable to any additional constraints and complexities**.



Problem: Is Optimizing Picking Enough?

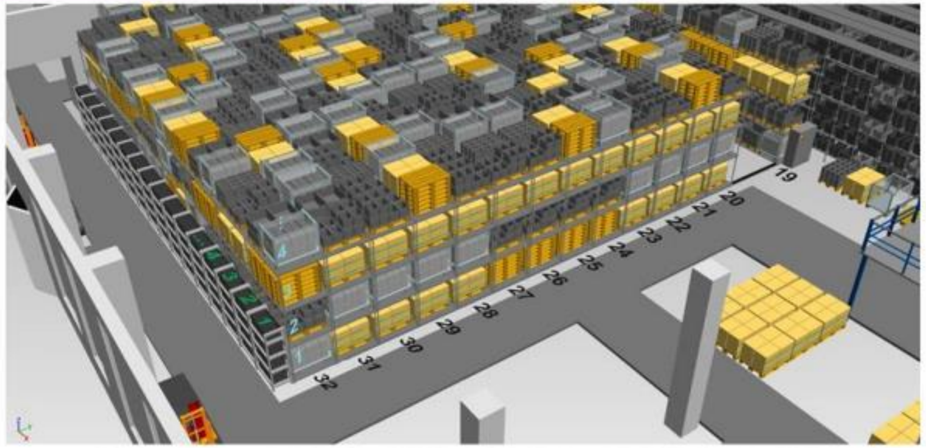
Secondary Goal Phase 1

A solution that also optimizes the **placement of items** inside the warehouse to **minimize the average picking time** it takes to fulfill a delivery request.

InstaDeep™

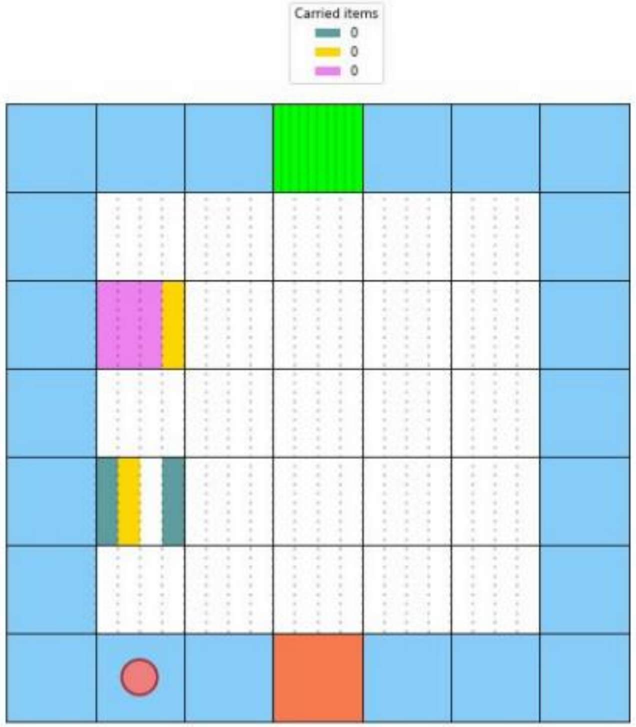
Deep RL can dynamically Optimise Warehouse Operations

Real Warehouse



Real Picker

Simulated Warehouse and Picker



Storage

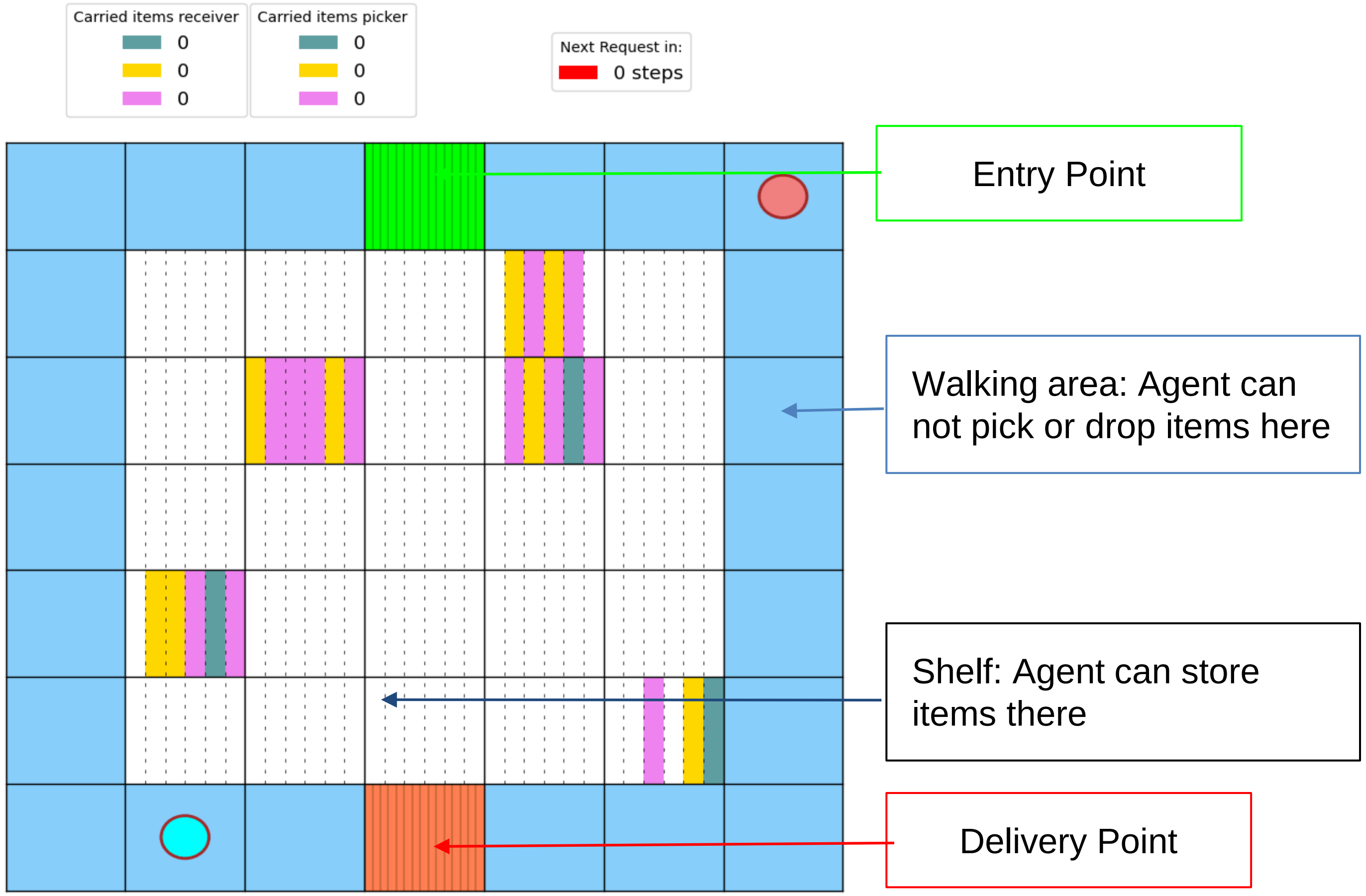
Movement

Delivery

Arrival

InstaDeep™

Deal with unplanned Fluctuations of Loads & Operations



Trained agents on grid environment and 3D model that incl. shelves

Problem and Solution:

Loads fluctuate throughout the week and intra-day **unplanned loads** are also common. Effective **space and labor management** are key to tackle this challenge effectively:

InstaDeep's **multi-agent AI system** optimises picking and storage actions **minimising picking time**. The agents learn continuously to **optimise and adapt storage & picking actions** in real-time.

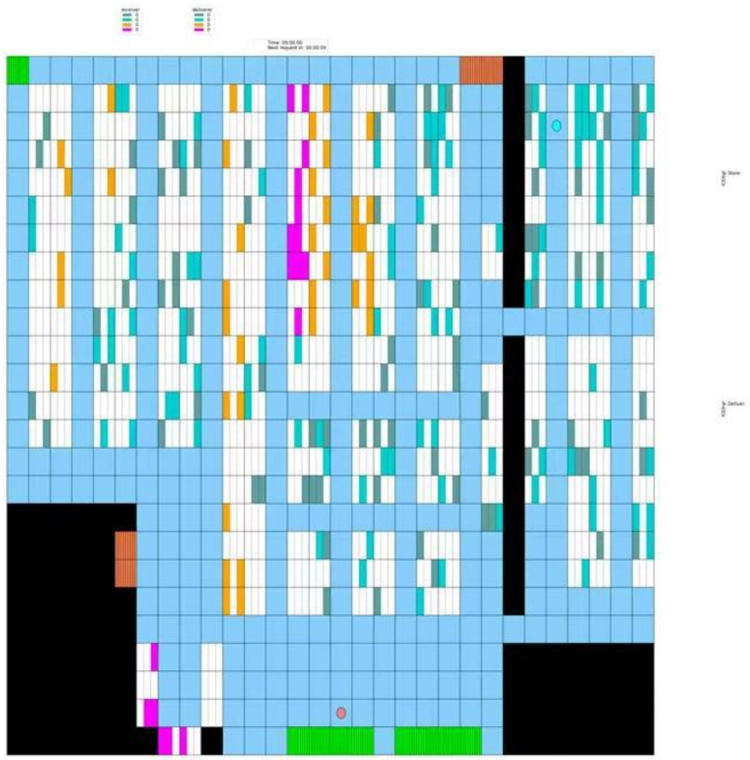
Value-Add:

Opex and time savings thanks to real-time Optimisation of storing and placing actions. **Robust against Load Fluctuations.** Scalable and generalisable.

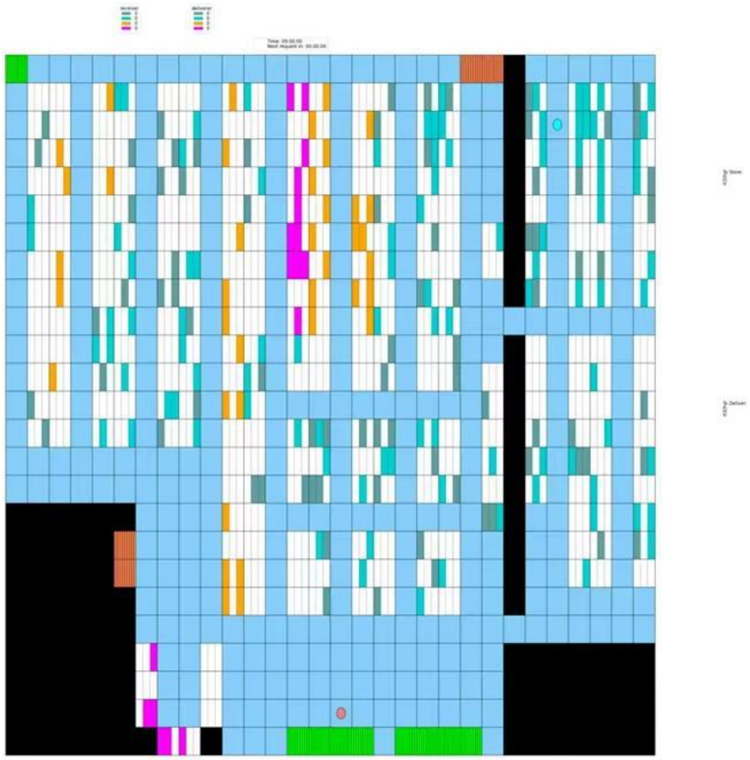
Goal: Optimize warehouse ops with specific constraints and then generalise across other warehouses with minimal extra effort

Results: Strong Improvement Demonstrable for both KPIs

AI Solution



SAP Solution



Results:

- **35% reduction in average picking time for picking jobs.**
- **84% improvement in on-time delivery / "Request Fullfilment Rate".**

InstaDeep™

AI Load Planning:

DeepPack™

DeepPack™: Pack Items more efficiently and save on your logistics costs

DeepPack: First AI-Powered 3D Truck Load Planning

Enhanced Efficiency

Generate optimal load plans fast, with minimal manual effort and planning time. **Maximise space utilisation** and enable workers with clear packing instructions

Improved Safety

Ensure stability by minimizing risks of accidents or cargo shifting. Comply with regulatory requirements and safety guidelines.

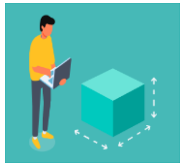
Cost savings

Reduce fuel consumption by cutting weight imbalances. **Cut handling and storage costs** thanks to optimized space allocation.

Scalability

Optimise for all trailer types and constraints as required in operations.

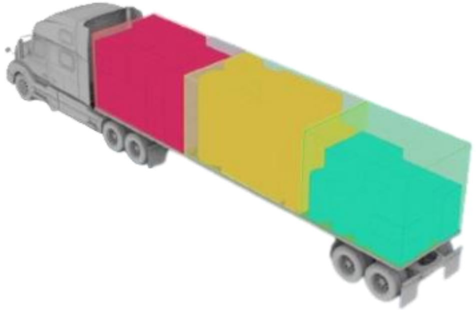
1 Upload cargo

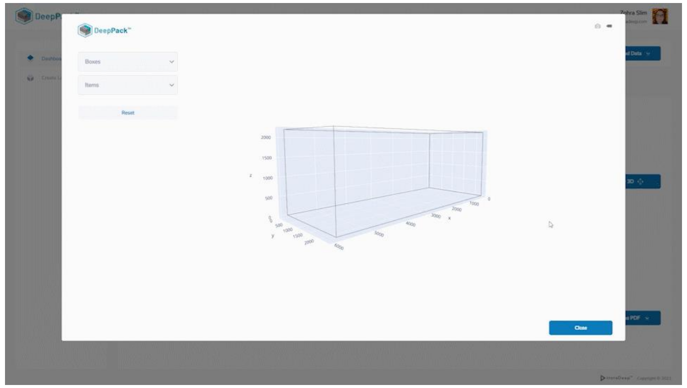


2 Specify load constraints

3 Select bin type

4 3D dynamic visualisation





Smart, self-learning tool using AI

Handle **complex shapes**

API enabled

Collaborative workspace

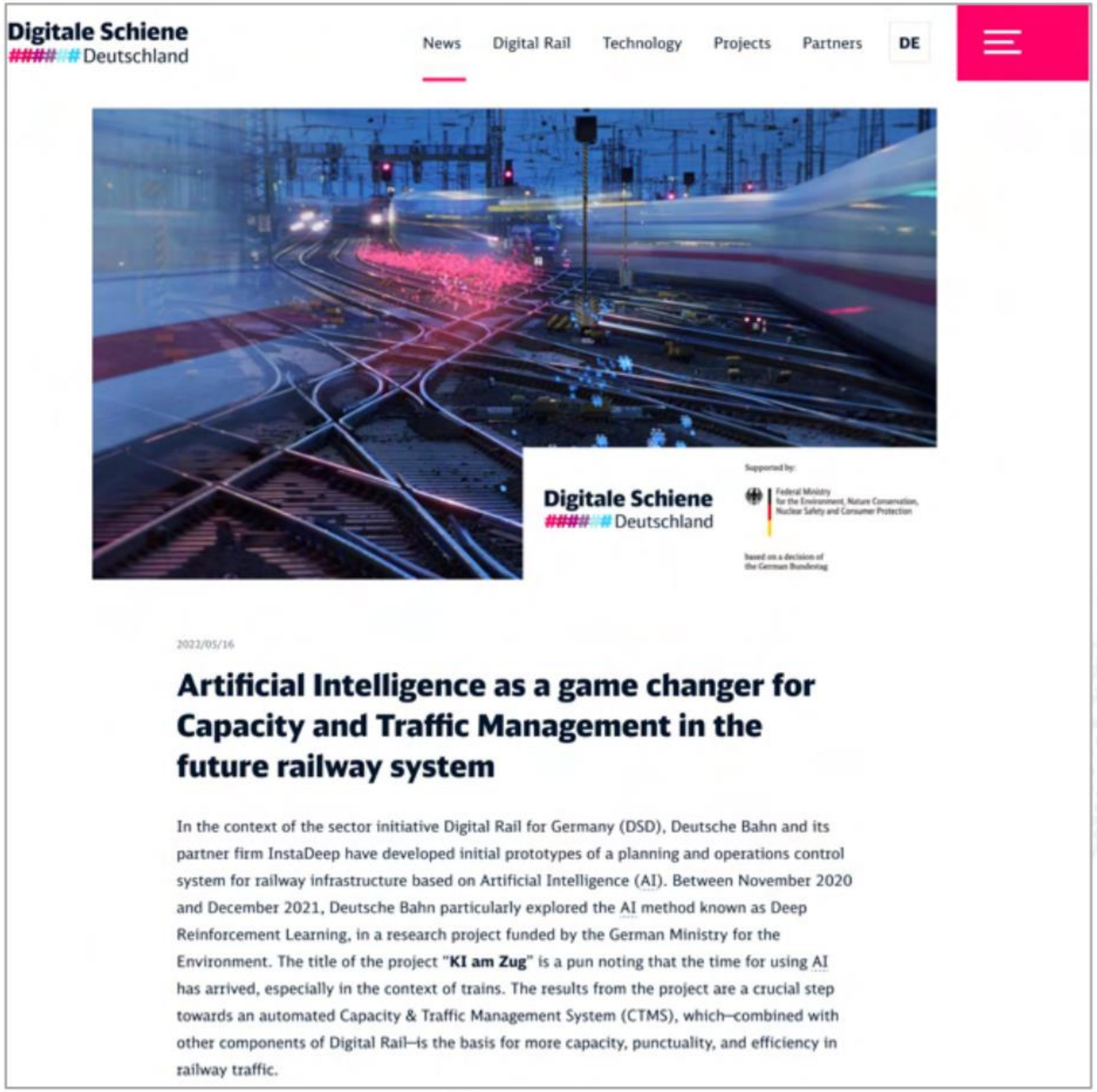
InstaDeep™

Digitale Schiene Deutschland

AI Scheduling & Routing:

DeepRail™: AI-First Capacity & Traffic Management System (CTMS)

AI for CTMS: Strategic Partnership with Deutsche Bahn


Digitale Schiene Deutschland

News Digital Rail Technology Projects Partners DE

Digitale Schiene Deutschland

2022/05/16

Artificial Intelligence as a game changer for Capacity and Traffic Management in the future railway system

In the context of the sector initiative Digital Rail for Germany (DSD), Deutsche Bahn and its partner firm InstaDeep have developed initial prototypes of a planning and operations control system for railway infrastructure based on Artificial Intelligence (AI). Between November 2020 and December 2021, Deutsche Bahn particularly explored the AI method known as Deep Reinforcement Learning, in a research project funded by the German Ministry for the Environment. The title of the project "**KI am Zug**" is a pun noting that the time for using AI has arrived, especially in the context of trains. The results from the project are a crucial step towards an automated Capacity & Traffic Management System (CTMS), which–combined with other components of Digital Rail–is the basis for more capacity, punctuality, and efficiency in railway traffic.

DB's **Digitale Schiene Deutschland** vision: Bring a "revolution into the rail system", increase capacity and reliability using novel tech

Pain Point:

- 10,000's of daily trains in dense traffic need quick decisions on how to adjust operations as disruptions happen

Vision:

- Modernisation and Digitalisation of the infrastructure
- Full ATO based on on-board perception sensors and AI
- AI-powered capacity and traffic management (CTMS) to automate planning, scheduling and dispatching
- Develop interoperability with other operators in EU

InstaDeep is DB's trusted AI partner for the CTMS enabling a real-time, dynamic train schedule at scale **to improve capacity and reduce delays**

AI for CTMS: Running on MS Azure

Use-case 1: schedule construction

Use-case 2: live re-dispatching

Increased demand

Global passenger rail activity expected to **double** and global freight rail activity expected to **triple** by 2050[1].

Major inefficiency

25% of DBs high speed trains were late in 2021[2].

Challenges

Intractable

Optimization solvers **don't scale** to a country like Germany with **10,000+** train **rides/day, 33,000 kms** of railway.

Manual

Time consuming process as it requires human in the loop combined with heuristics.

Inefficient

Significant **inefficiencies**, especially in train rescheduling upon perturbations, leading to **delays** and **costs**.

RL Solution

Scalable

Fast and **scalable** RL-based decision-making engine

Automated

Automatic adjustment of the timetables

No heuristics: learns from first principles with limited/no bias

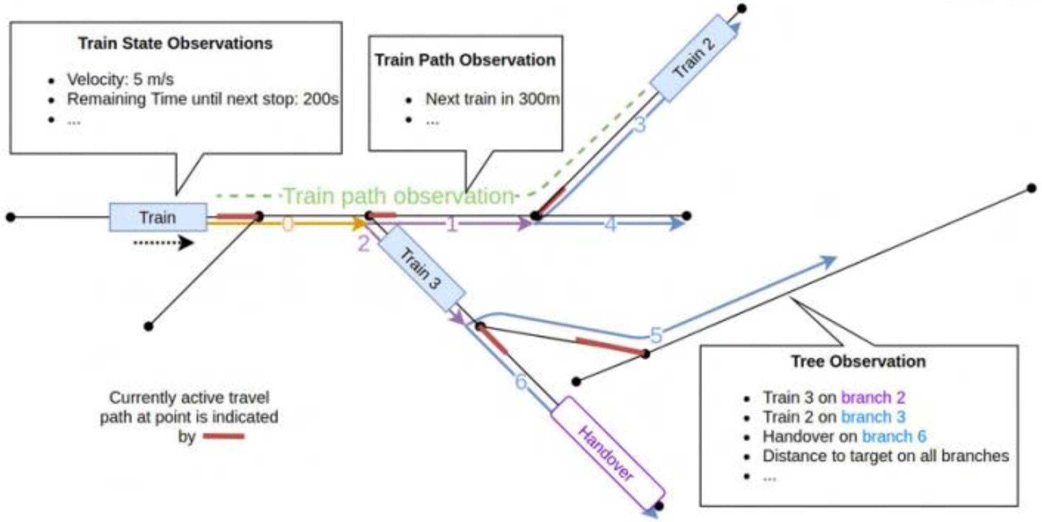
Efficient

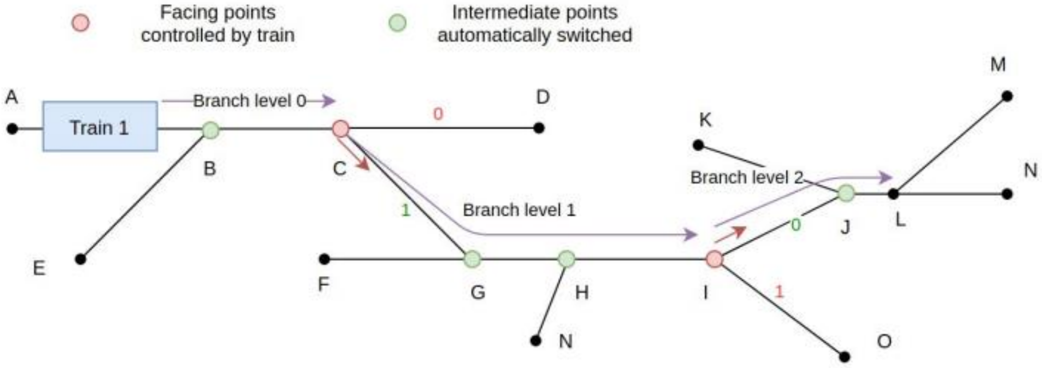
Maximize railway network utilisation, while minimizing delays and saving costs (multi-million euros per week) .

[1] https://iea.blob.core.windows.net/assets/fb7dc9e4-d5ff-4a22-ac07-ef3ca73ac680/The_Future_of_Rail.pdf
[2] https://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/assess_unit_cost_rail_en.pdf

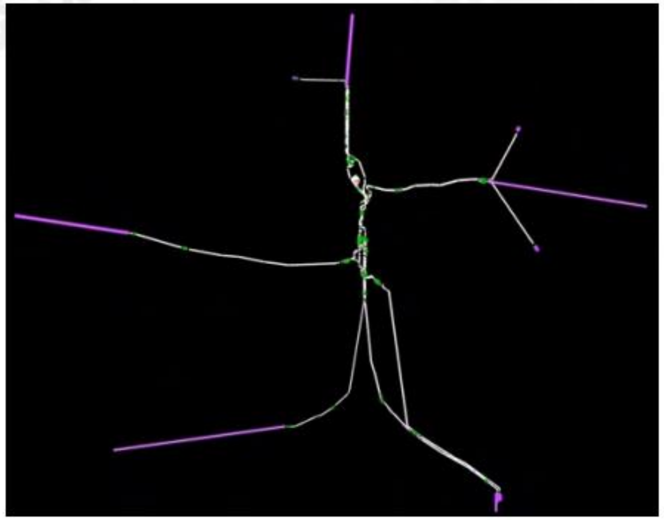
Digital Twin: Realistic environment for 10,000+ rides / day

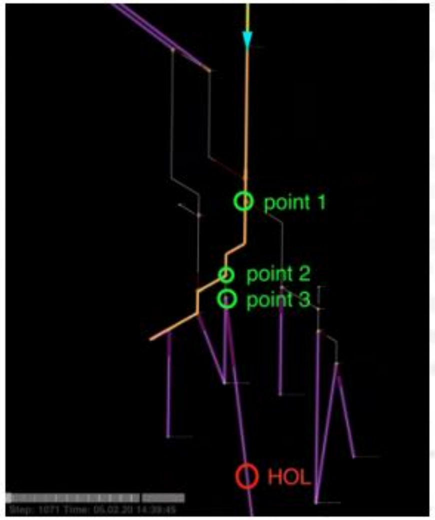
Detailed observations

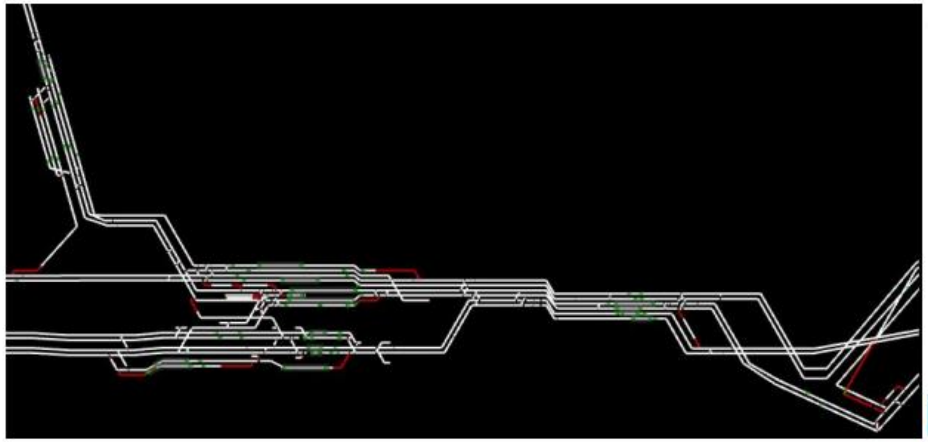




Complex navigation environment







InstaDeep™

Expanding AI Capabilities into more Verticals

Benefits of Optimising with Decision-Making AI

Fraport
Frankfurt Airport Services Worldwide

1

Address highly Complex Business Constraints

Many business challenges are multi-objective, sequential decision-making problems that can benefit from RL

2

Real-Time Decisions in Dynamic Environments

Real-life business operations are subject to uncertainty and change - RL agents can adapt to previously unseen situations

3

No Need for Massive Amounts of Training Data

Digital twin simulators generate required labelled data for RL algorithms to learn from and optimize decision-making

4

Actionable AI

Shift operating model from reactive and prognostic to being proactive by training expert AI systems that take decisions

InstaDeep™

Energy: Forecasting, Planning & Dispatch

elia
Powering a world in progress

1

Power Consumption Forecasts

At the level of a household or a small insulated micro-grid. Or on a national level. Different scales result in different challenges.

2

Power Production Forecasts

E.g. of renewable energy systems like solar panels. Accuracy would depend on data collection available.

3

Optimising production plans of controllable production units

For coupled production fleets (gas, solar, etc.) to cut costs while ensuring supply = demand over a set interval (day ahead, intra-day re-declarations, real-time)

4

Power dispatch

Re-optimisation in real-time in response to hazards and unpredictabilities.

InstaDeep™

Supply Chain: Increase Efficiency and Cost Structure

1

Demand Forecasting

Predict weekly/monthly sales for key product categories months ahead. Reduce prediction error by using RL methods

2

Inventory Management

Optimise planning, cut costs for restocking, holding and fulfilment, and avoid overstocking. Include features for picking optimisation or ABC stratification

3

Raw Material Purchasing Variance

Accurately estimate raw material costs and choose the right vendors. Safeguard against purchasing variance, vendor uncertainty risks and lower margins

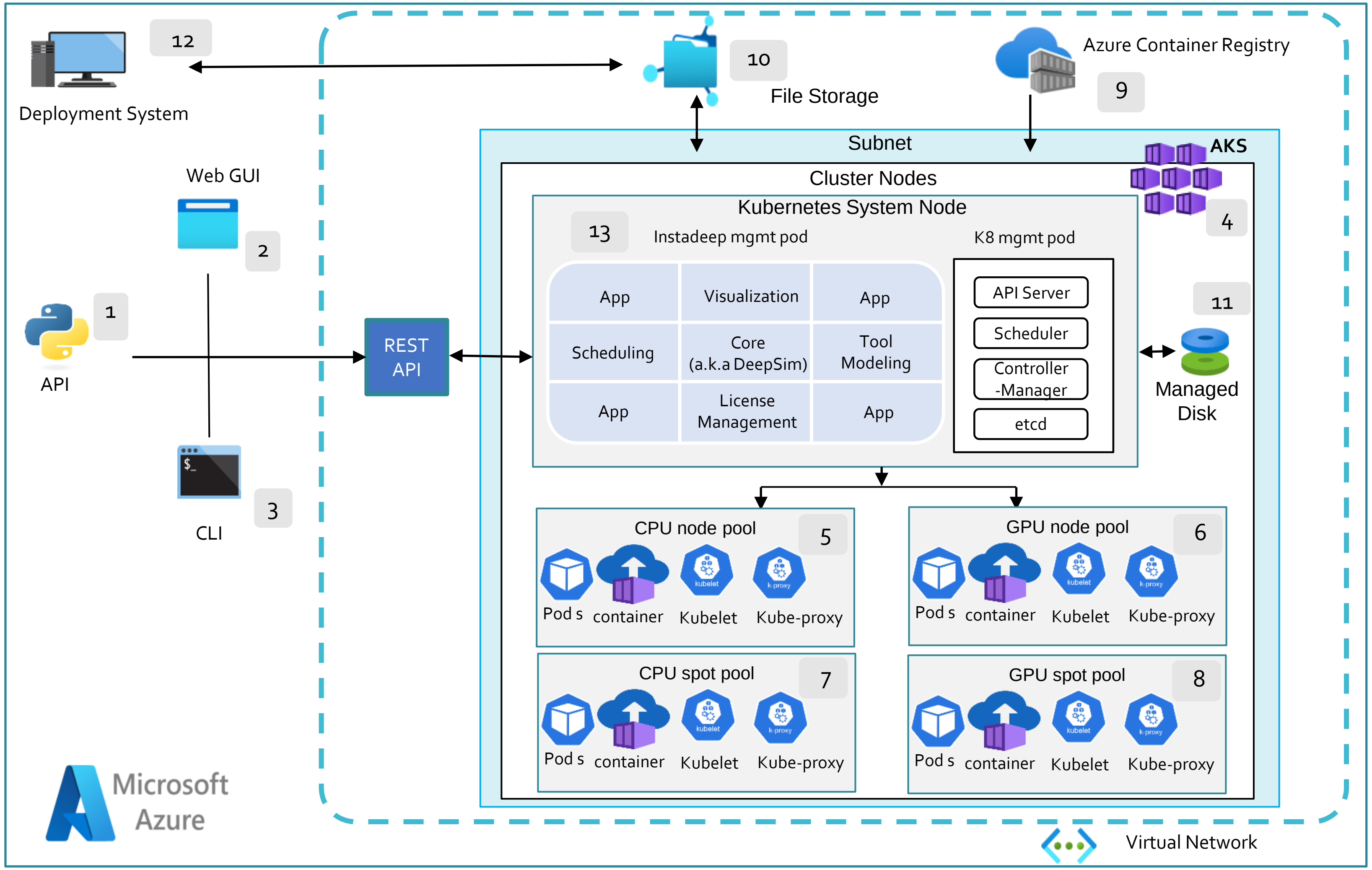
4

Dynamic Pricing

Find customer specific prices for cohorts, taking into account general supply-demand patterns, seasonality, costs and product life cycles

Reference Architecture
Scheduling and dispatching with reinforcement learning

- Instadeep uses deep reinforcement learning (RL) to augment fab scheduling workflows to decrease production costs.
- Cost effective: Fully compatible with spot instances and reserved instances to further reduce the Cost of RL training
- Scalability: Efficiently scale up to tens of thousands of cores and back down as needed for faster time-to-market.
- Scheduler + model agnostic: This architecture may be used for a wide variety of schedulers/simulators, input data, and workflows.



- The whole Architecture is contained within customer's Azure environment.;
- End users interact with InstaDeep™ [13] management system via a REST API that is running on the scalable Azure Kubernetes Service [4] in a variety of ways:
  - Python API [1]
  - Web based User interface [2]
  - Command line client [3]
- User submit the job in Instadeep [13] , which schedules the training jobs on the cluster.
- The cluster will assign pods to the relevant node pools and scales these up if required [5,6,7 or 8].
- The pods get initialized from the containers that are stored in the Azure Container Registry [9].
- During training, the results are stored in the Azure File Storage [10] and the metric tracking system that is part of Instadeep mgmt pods [13] (and backed by addition storage device [11])
- Through InstaDeep ™ [13] tool the user monitors the job progress
- After the training, the agent is pushed to the deployment system [12] from where it can be queried for actions. The deployment server [12] has the option to report back monitoring statistics to the platform[13] for further optimization of the agent via File Storage [10].

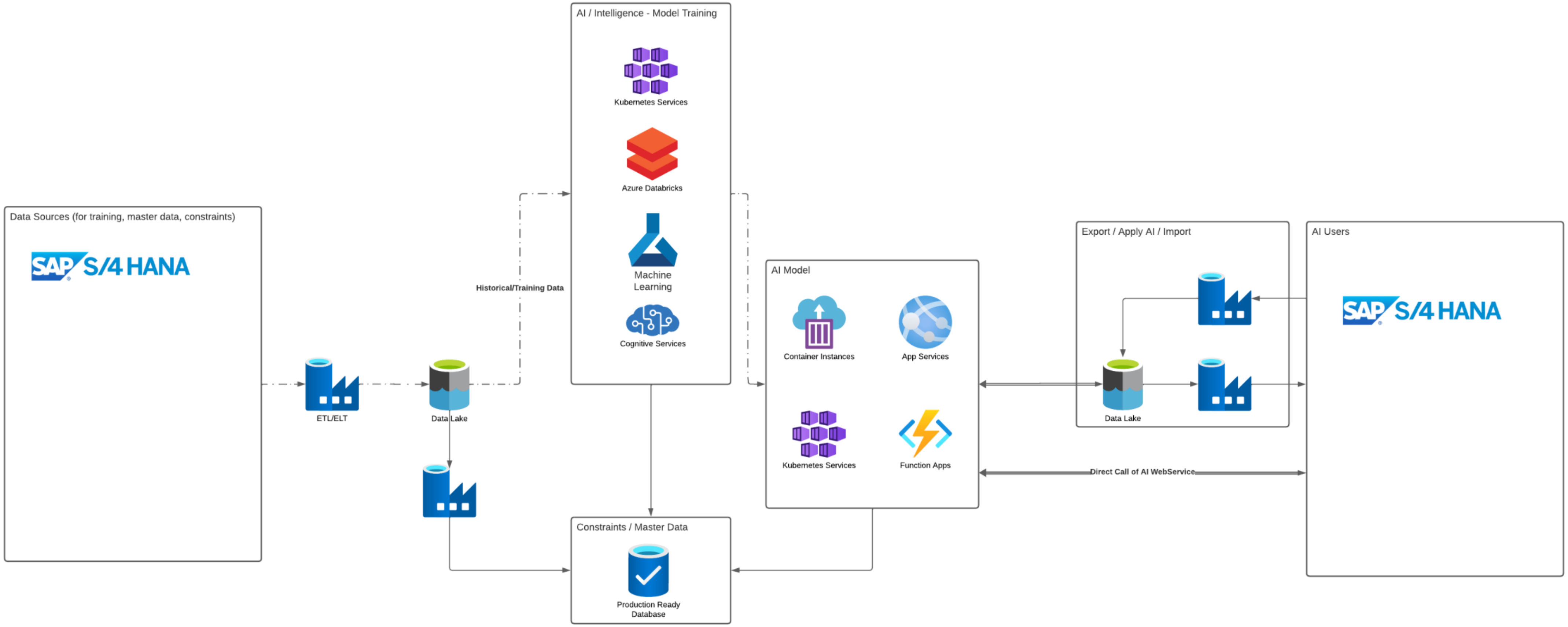
Services	Type	Quantity	Size	Services	Type	Quantity	Size
Standard Node	VM	1	B8ms	Azure Files	Storage	1	1TB
CPU nodes	VM	1	HB120-rsv3	Managed Disk	Storage	1	20GB
CPU spot Nodes	VM	20	HB120-rsv3	Container registry	Containers	1	
GPU Nodes	VM	0	NC6s v3	GPU spot Nodes	VM	5	NC6sv3

Fab scheduling and dispatching with reinforcement learning, Semiconductor

Configuration = CPU: 2400 Cores , 120 cores /node, 20 Nodes, AMD Milan-x

GPU: #5 GPUs, 1GPU/node, 5 Nodes, Nvidia Tesla V100

10 RL Agents training for a month, Compute: 360 Hrs, Cost: ~ $10,030


Data Sources (for training, master data, constraints)
SAP S/4 HANA
ETL/ELT
Data Lake
Historical/Training Data
AI / Intelligence - Model Training
Kubernetes Services
Azure Databricks
Machine Learning
Cognitive Services
Constraints / Master Data
Production Ready Database
AI Model
Container Instances
App Services
Kubernetes Services
Function Apps
Export / Apply AI / Import
Data Lake
AI Users
SAP S/4 HANA
Direct Call of AI WebService

What Microsoft HPC/AI GBB brings to the table

Use Case session

Funding for Projects*

Fast Track for Azure*

Hands-on experience with Hackathon

Technical Skills for Business*
Learning Path (online Course)

Account Team
Cloud Solution Architects

*Subject to Approval after Nomination

How Startups can benefit

Joint projects wih
enterprise cusomers

Funding for Projects*

Fast Track for Azure*

Azure Markeplace
Fast procurement

Technical Skills for Business*
Learning Path (online Course)

Use Case session

*Subject to Approval after Nomination

Concept of Digital Twin in Manufacturing Industry

What is a digital twin?

According to the very first definition back in 2003 by Michael Grieves [1][2][3], a Digital Twin is a virtual representation of a physical entity, collecting all the information related to his lifecycle management

Digital model
- Digital representation of the physical entity

Digital shadow
- Digital representation and physical-to-virtual information flow

Digital twin
- Digital shadow with virtual-to-physical information flow

Digital twin requires three elements[1][2]:

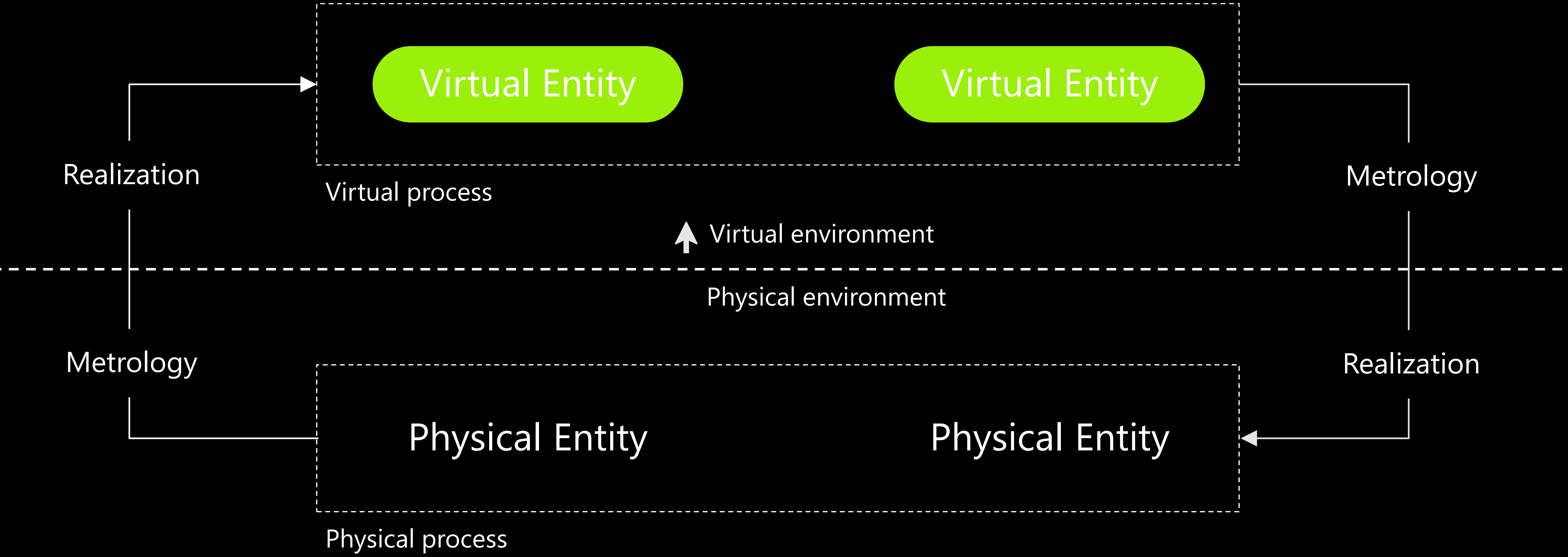
- A physical entity
- A digital representation of the physical entity
- A bi-directional information flow between the two

[1] David Jones, Chris Snider, Aydin Nassehi, Jason Yon, Ben Hicks, **Characterising the Digital Twin: A systematic literature review**, CIRP Journal of Manufacturing Science and Technology, Volume 29, Part A, 2020, Pages 36-52, ISSN 1755-5817

[2] Grieves, Michael. **Digital twin: manufacturing excellence through virtual factory replication**. 2014. *White Paper* (2017).

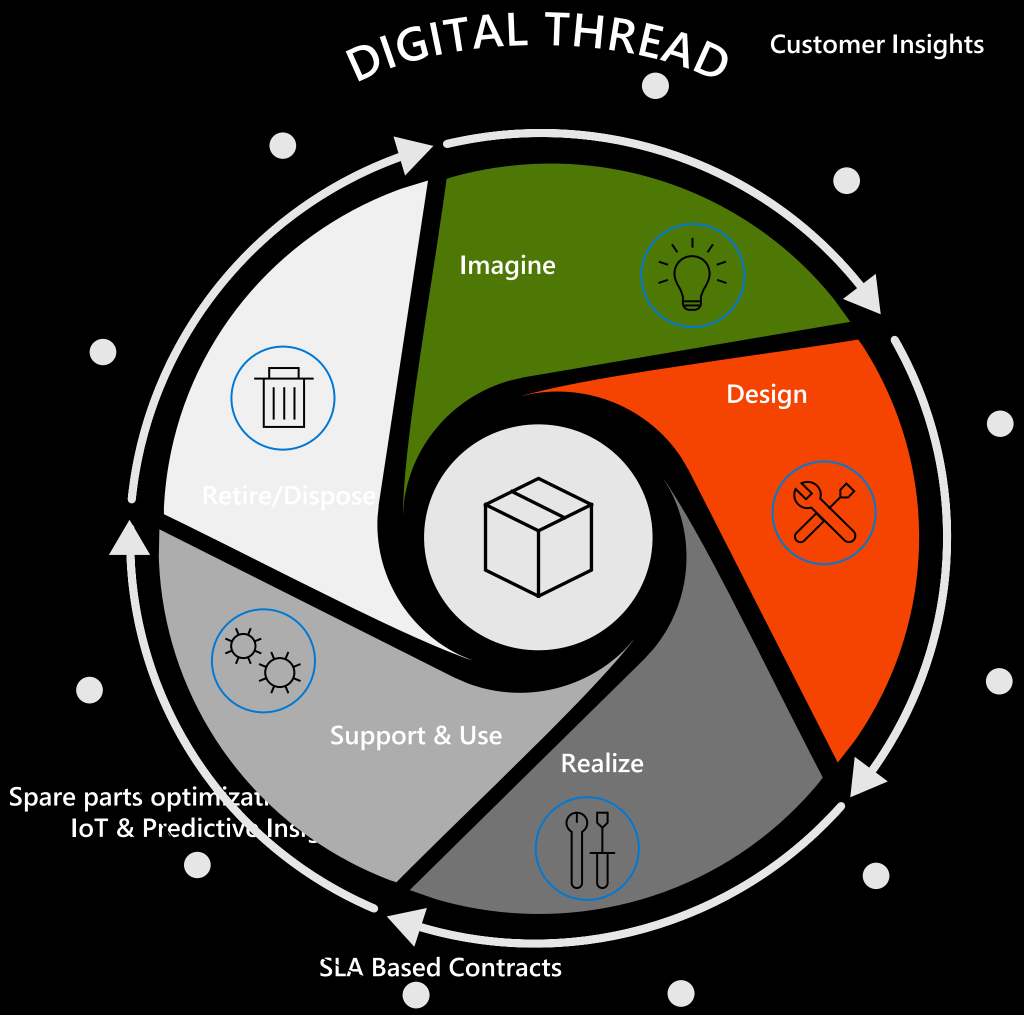
[3] Mohsen Attaran, Bilge Gokhan Celik, **Digital Twin: Benefits, use cases, challenges, and opportunities**, Decision Analytics Journal, Volume 6, 2023, 100165, ISSN 2772-6622

Digital Twin information flow[1]



[1] David Jones, Chris Snider, Aydin Nassehi, Jason Yon, Ben Hicks, **Characterising the Digital Twin: A systematic literature review**, CIRP Journal of Manufacturing Science and Technology, Volume 29, Part A, 2020, Pages 36-52, ISSN 1755-5817

Digital Twin in Stark's Product Life Cycle

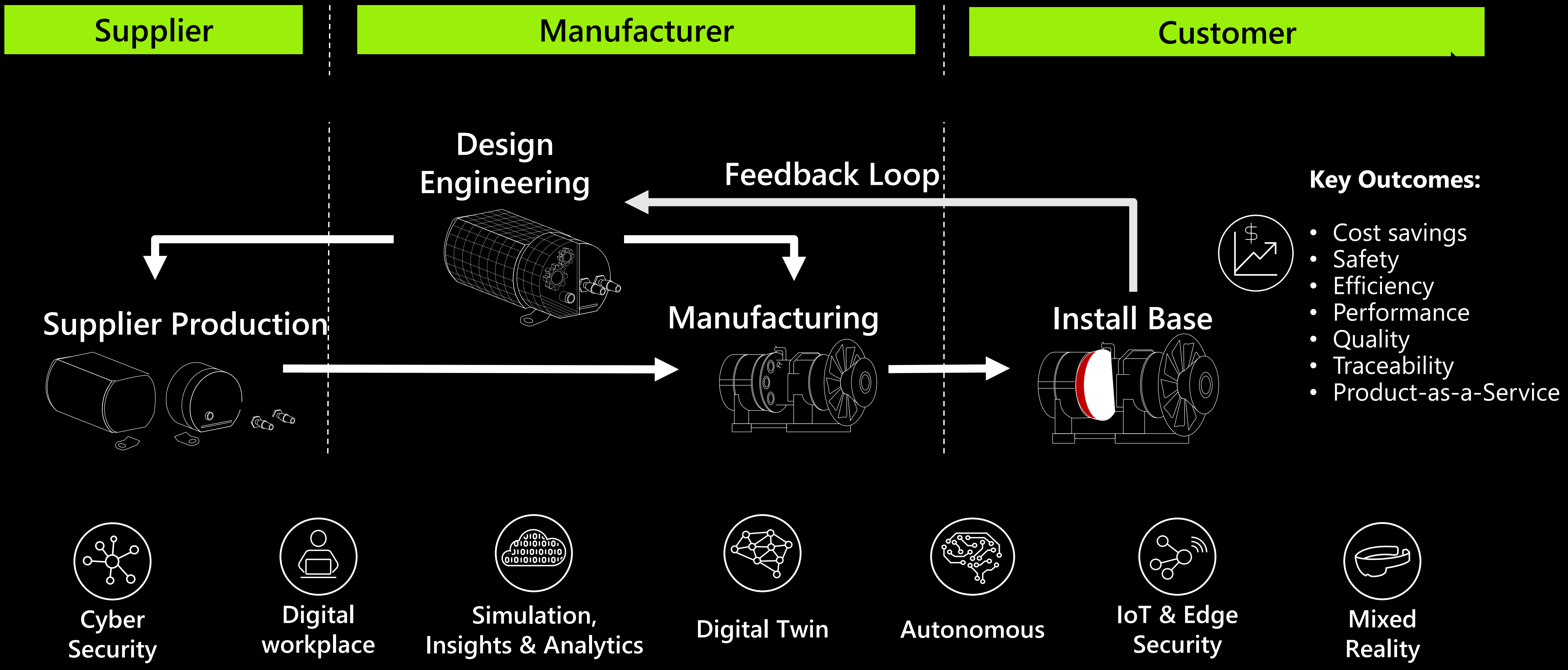


"Digital Thread is a data-driven architecture that links together information generated from across the product lifecycle"

Availability of a Digital Thread architecture is a key enabler to enrich Digital Twin capabilities

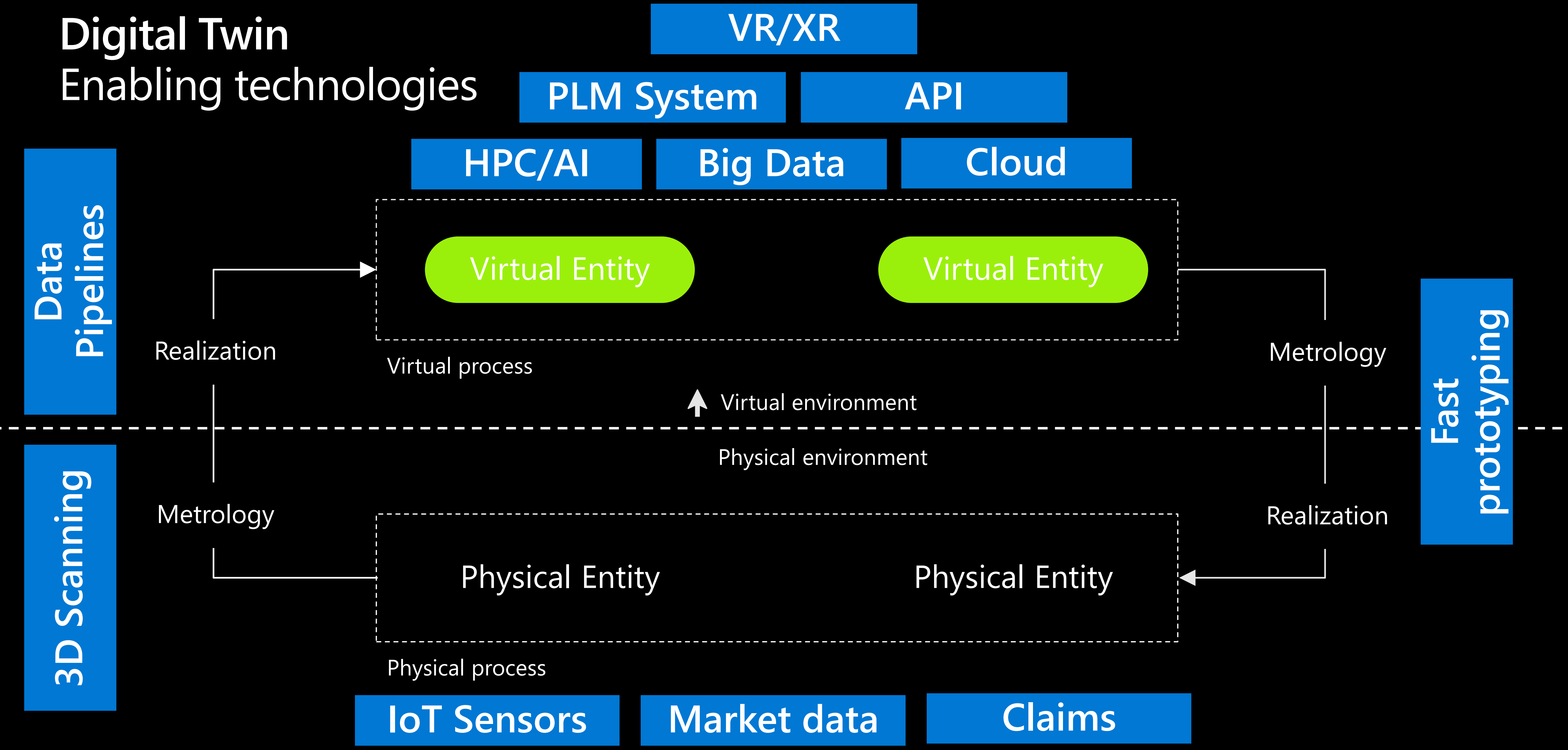
[1] Stark, J. **Product Lifecycle Management (Volume 1)**, 21st Century Paradigm for Product Realisation, Springer International Publishing, 2022
[2] Singh V., Willcox K. E., **Engineering Design with Digital Thread**, AIAA Journal 2018 56:11, 4515-4528

The “Digital Thread” – Empowering Digital Engineering



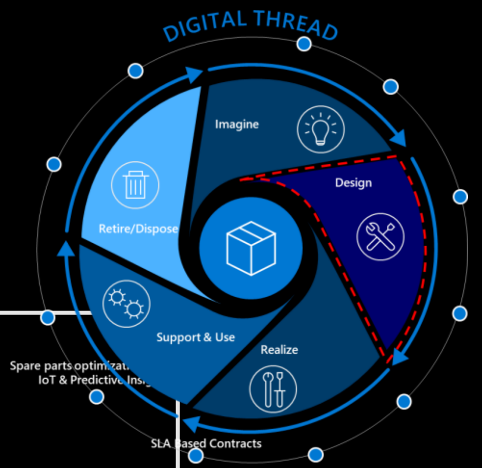
Cloud & Edge | Data & AI | Trust & Security | Sustainability

Digital Twin
Enabling technologies



[1] David Jones, Chris Snider, Aydin Nassehi, Jason Yon, Ben Hicks, **Characterising the Digital Twin: A systematic literature review**, CIRP Journal of Manufacturing Science and Technology, Volume 29, Part A, 2020, Pages 36-52, ISSN 1755-5817

HPC / AI in Manufacturing Digital Thread



Product Data

Market Data

Surrogate models

Multi-physics

Electromagnetic

Chemical

Structural

CFD

Deep learning

Machine learning

Physics-based

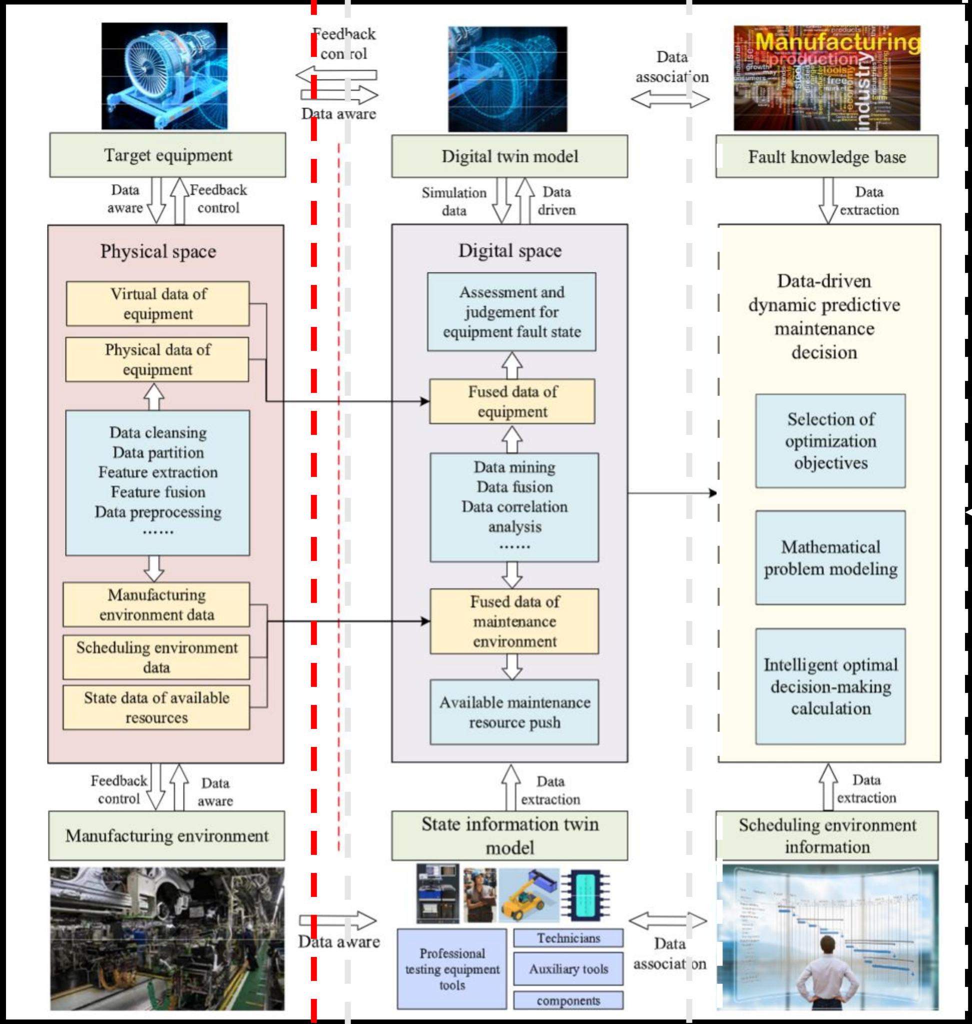
AI-based\

Examples of Digital Twin applications and collaborations

Examples of Digital Twin – Support & Use

[1] Dong Zhong, Zhelei Xia, Yian Zhu, Junhua Duan, **Overview of predictive maintenance based on digital twin technology**, Heliyon, Volume 9, Issue 4, 2023,

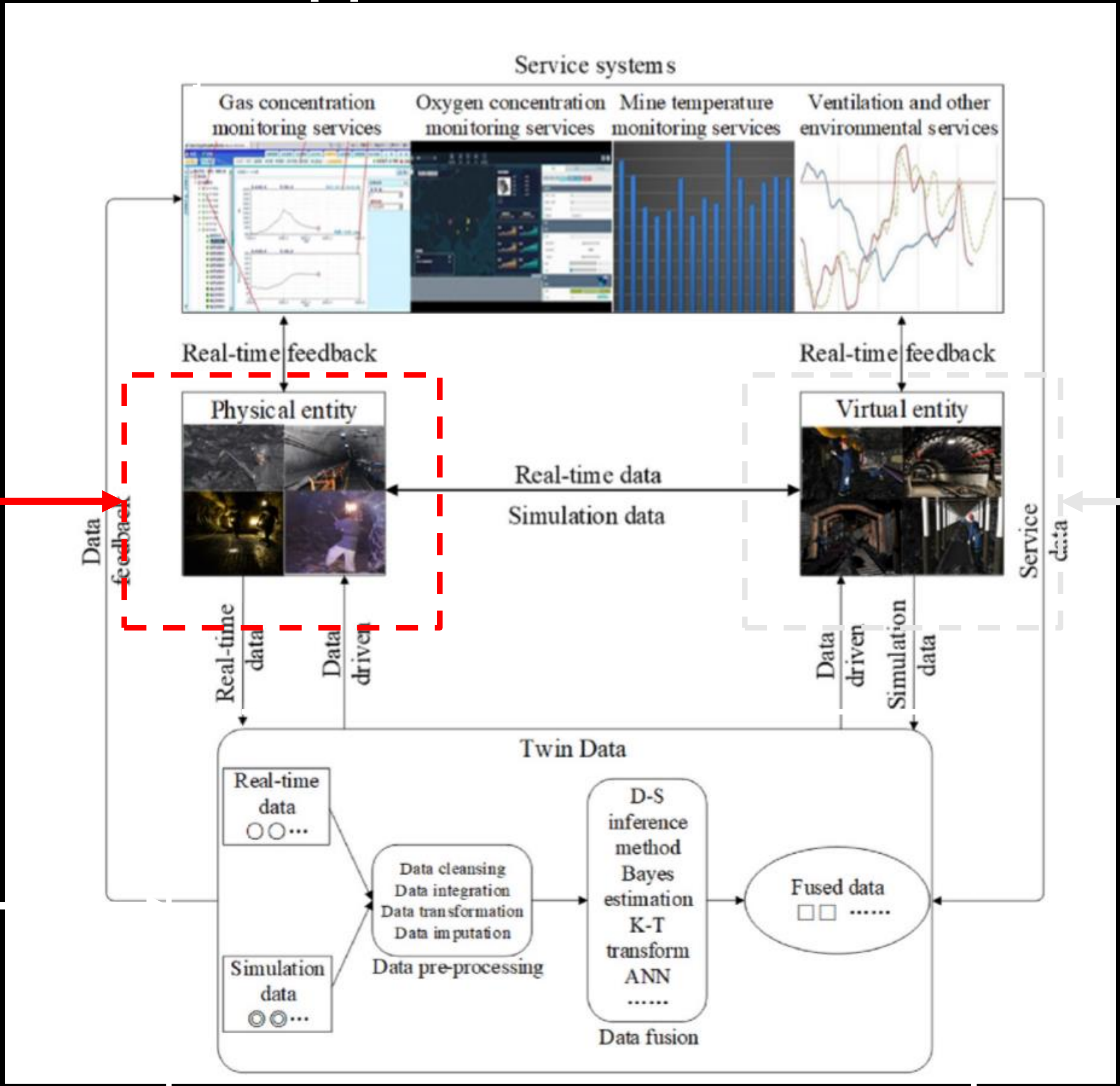
Physical entity



Virtual environment

Virtual entity

Examples of Digital Twin – Support & Use



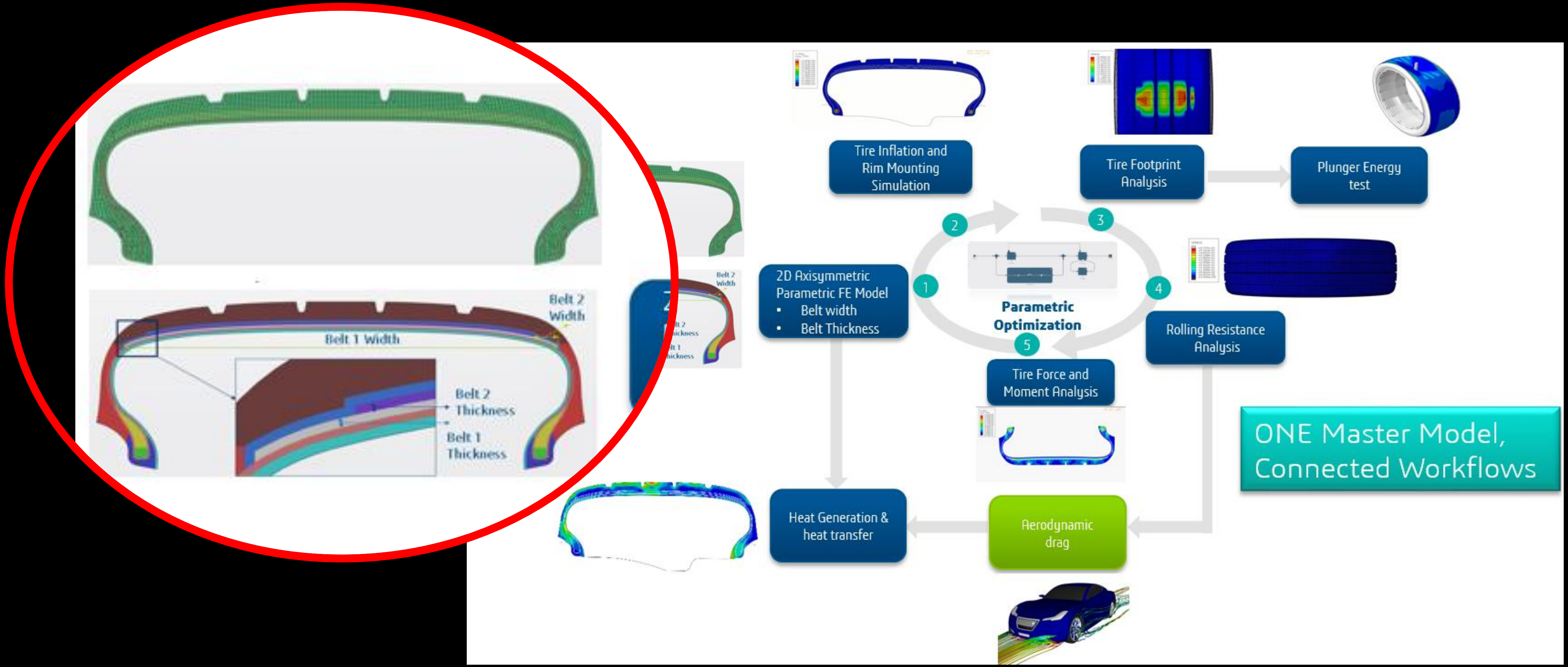
Physical entity

Virtual entity

Virtual environment

[1] Jiaqi Wang, Yanli Huang, Wenrui Zhai, Junmeng Li, Shenyang Ouyang, Huadong Gao, Yahui Liu, Guiyuan Wang, **Research on coal mine safety management based on digital twin**, Heliyon, Volume 9, Issue 3, 2023,

Examples of Digital Twin – Design & Optimization

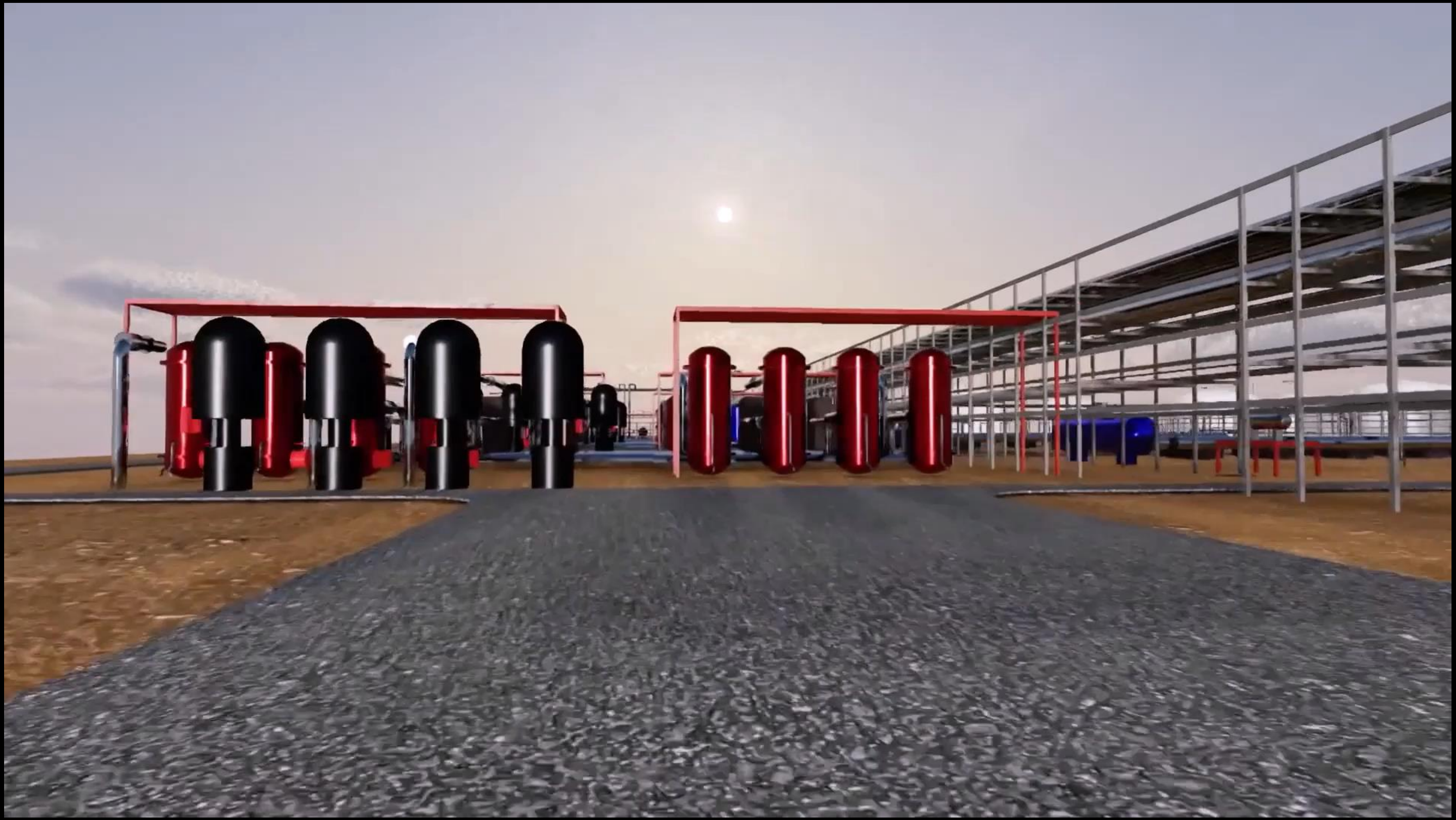


[1] Fast and Furious Tire Design: Dassault Systèmes Blog (3ds.com)

Examples of Digital Twin - Realize

[NVIDIA, BMW Blend Reality, Virtual Worlds to Demonstrate Factory of the Future | NVIDIA Blog]

Examples of Digital Twin - Realize



[Omniverse Accelerates Turning Wind Power Into Clean Hydrogen Fuel | NVIDIA Blog]

Brain storming “What can digital twin bring to you”?

Phase 1

Identify possible Digital Twins inside Bekaert business (**physical entity + virtual entity + information flows**)

Phase 2

Identify boundaries of Digital Twin virtual and physical environments (**Suppliers, Bekaert internal, customers**)

Phase 3

For each Digital Twin identified, vote the one which should be prioritized based on **complexity and impact**