Future-proofing BI:
an unexpected journey to leverage 'In-Chip' analytics in IoT and AI

Ani Manian
Head of Product Strategy

SIMPLIFYING Business Analytics for COMPLEX Data

"The key strength of Sisense is the platform's capability to **easily handle and manage large and diverse datasets**, and analyze them in dashboards based on its proprietary In-Chip technology."

- Gartner Magic Quadrant

HOW IT ALL STARTED

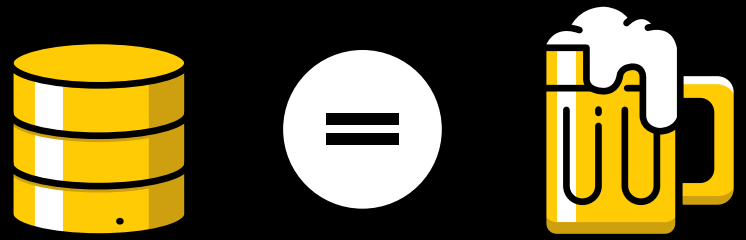
WHAT DO FIVE DATA GEEK STUDENTS DREAM ABOUT?

WELL, BELIEVING THEY'RE BADASS… THEY'RE DREAMING OF…

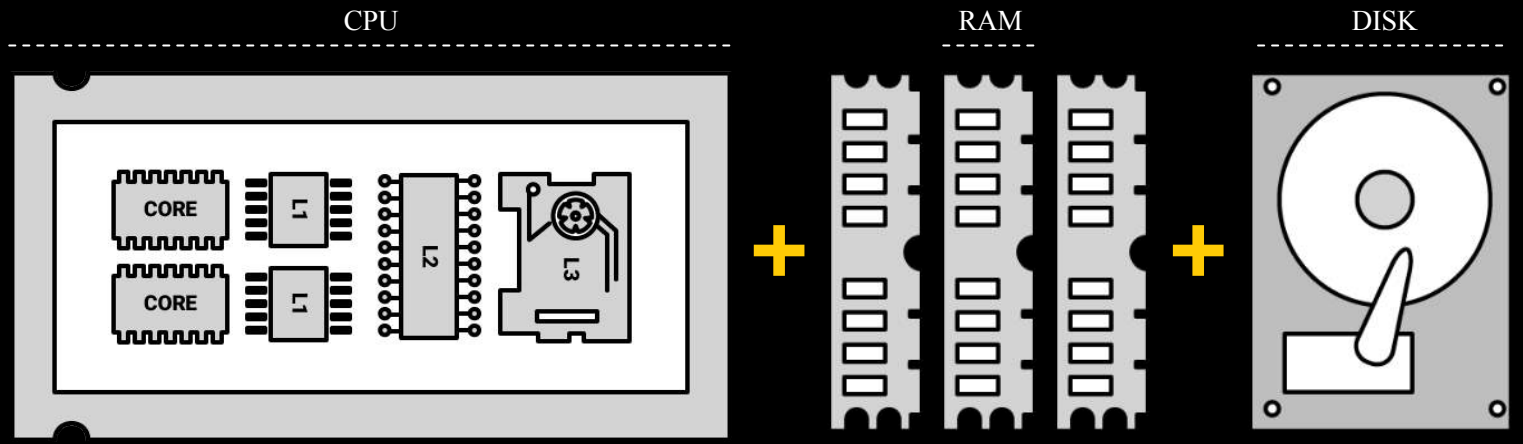
BEER & CHIPS

IN ORDER TO UNDERSTAND IN-CHIP ANALYTICS

LET'S ASSUME THAT:



MEMORY HIERARCHY IN MODERN CPUS



L1 Cache	L2 Cache	L3 Cache	Main Memory	Disk
Capacity: **64KB-128KB**	Capacity: **256KB-1MB**	Capacity: **6MB-20MB**	Capacity: **GBs-TBs**	Capacity: **Unlimited**
Latency: **3 Cycles**	Latency: **10 Cycles**	Latency: **35 Cycles**	Latency: **150-450 Cycles**	Latency: **1M Cycles**

SO, WHY SHOULD WE EVEN CARE?

Slowdown when fetching new data to the CPU

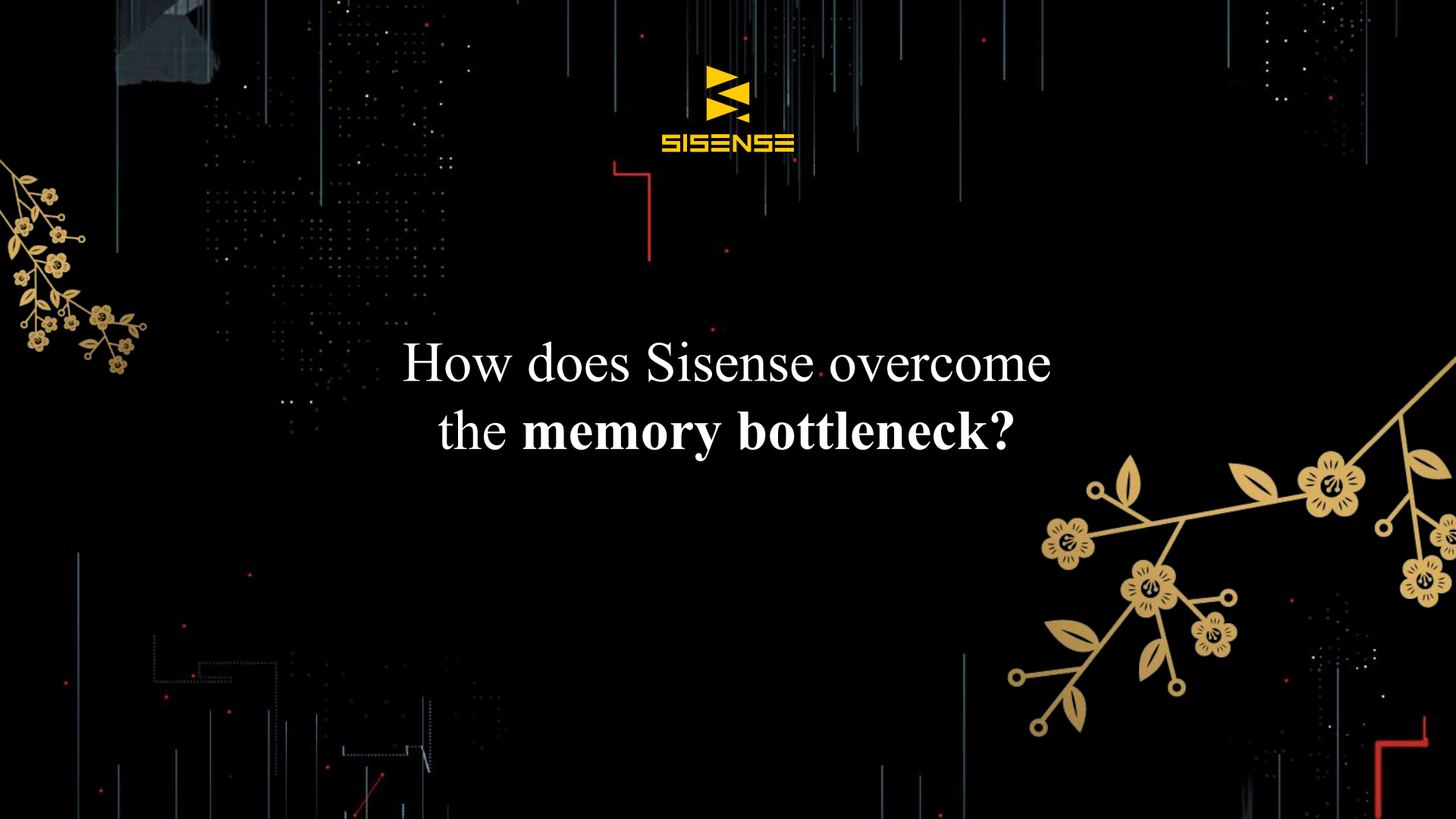
MEMORY BANDWIDTH

If data equals beer then data storage units equal all the places beer is kept!

L1 cache	Home fridge	x1 Distance	Immediate	Customer
L2 / l3 cache	Shop	x10 Distance	Bicycle	Customer
Ram	Supermarket	x50 Distance	Car	Customer
Disk	Brewery	Distance	Airplane	Customer

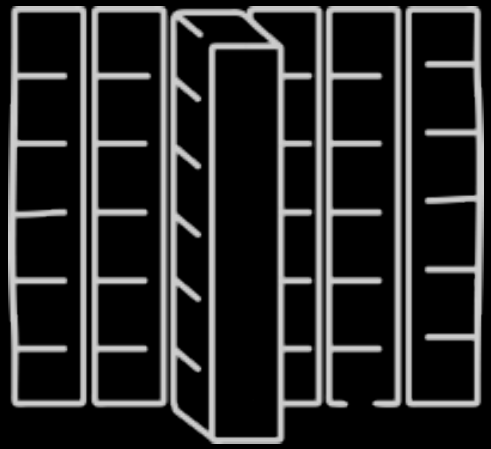
THERE SHOULD HAVE BEEN A SLIDE HERE..

(it’s the beer’s fault…)


SISENSE
How does Sisense overcome
the **memory bottleneck?**

HOW DOES SISENSE OVERCOME THE MEMORY BOTTLENECK?

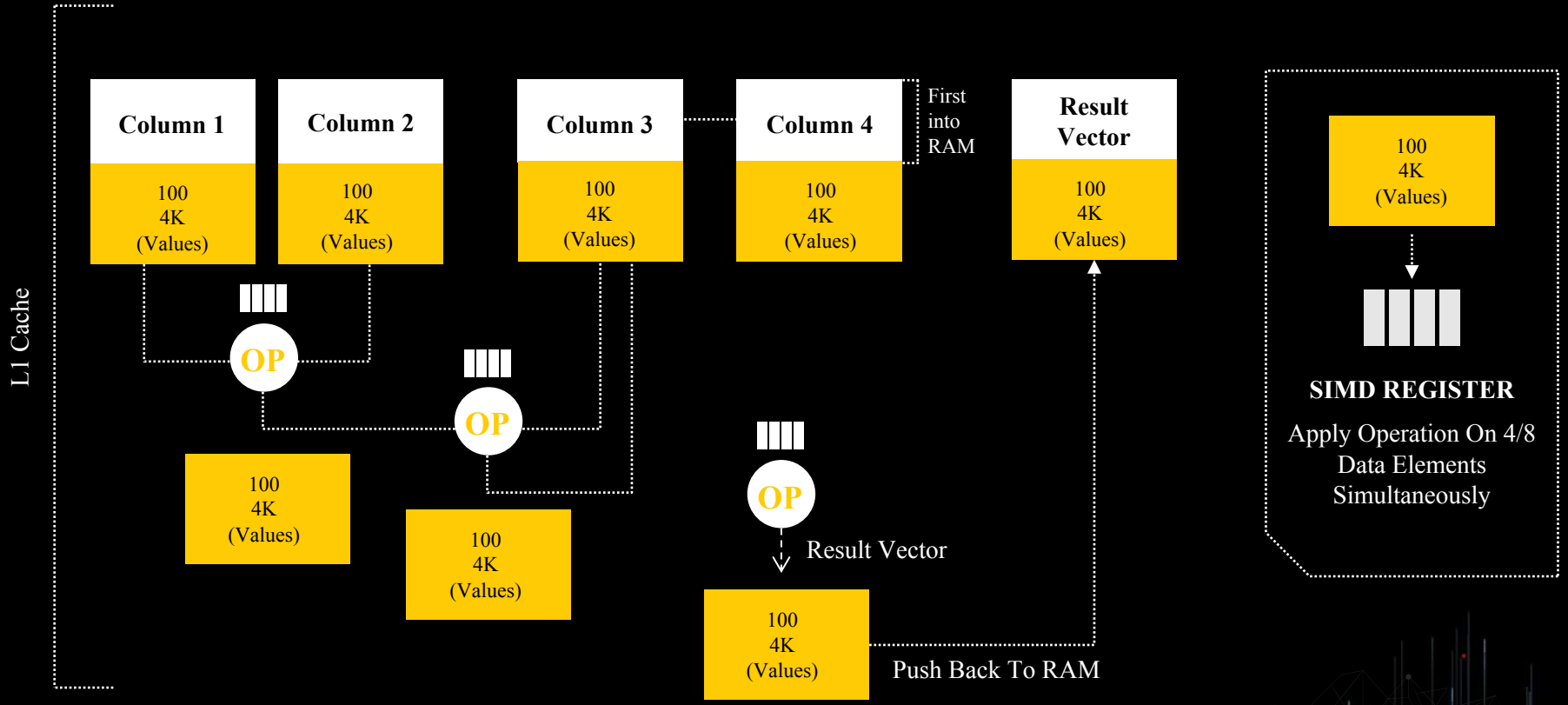
1. Store all data on the Disk
2. Only Use RAM When a Query Runs
3. Load Only the Relevant Columns in RAM



VECTORIZATION

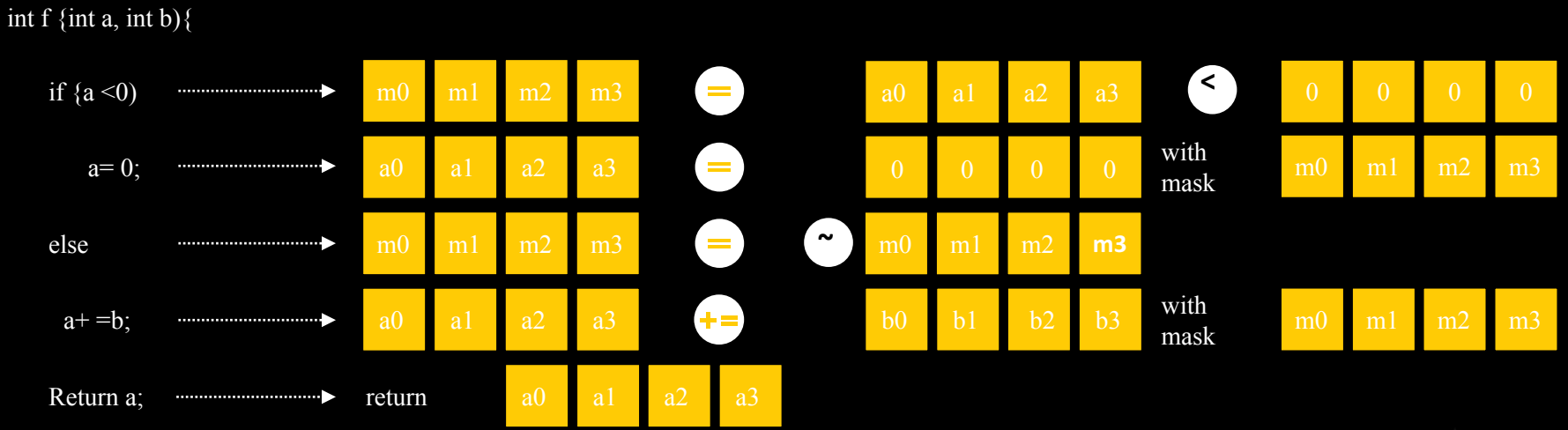
JIT LLVM & SIMD

VECTORIZATION & CACHE AWARENESS



JIT LLVM COMPILATION WITH SIMD SUPPORT

"SIMD" (Single Instruction, Multiple Data) is the process of rewriting a loop so that instead of processing a single element of an array N times, it processes (say) 4 elements of the array simultaneously N/4 times.



SELECT (f_1 = “beer1” **OR** f1 = “beer2”) **AND**

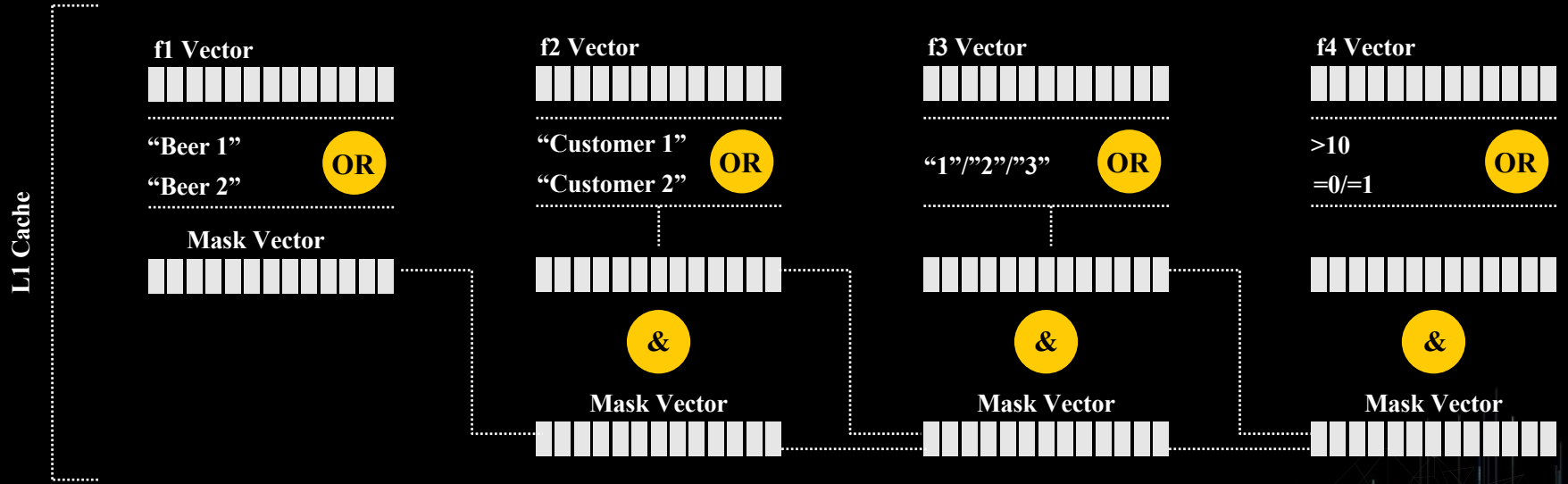
FROM T1 (f2 = “customer1” **OR** f2 = “customer2”) **AND**

WHERE (f3 = “1” OR f3 = “2” **OR** f3 = “3”) **AND**

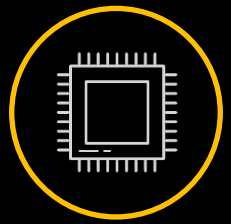
(f4>”10” **OR** f4 = “0” **OR** f4 = “1”)

Field Vector = Value

Mask Vector = True / False



NEXT: PERFORMANCE TUNING FOR MANY USERS



ADD HARDWARE

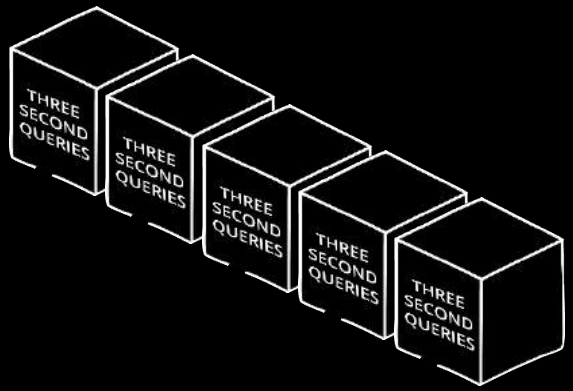
ADD INSTANCES

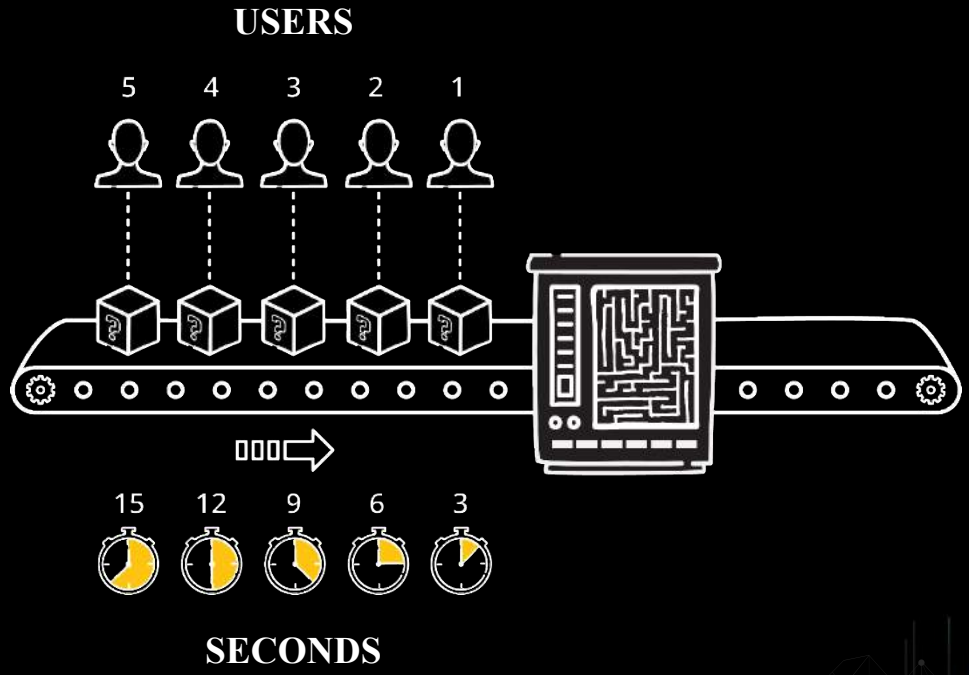
OPTIMIZE DATA MODEL

HOW CAN YOU DELAY USING THESE OPTIONS?

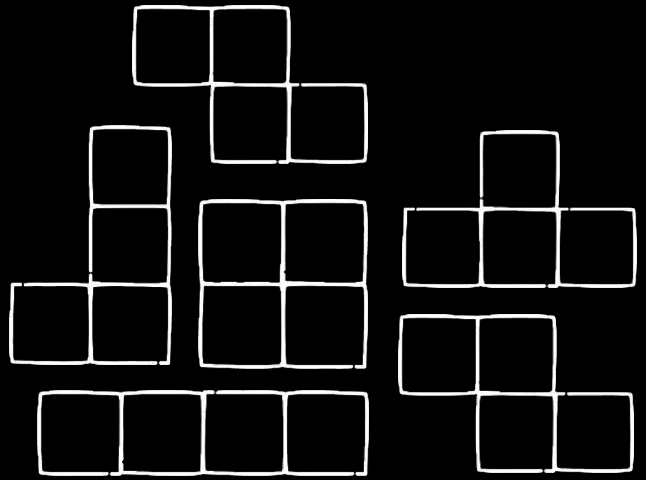
PROBLEM: THE WAITING LINE TO QUERY DATA

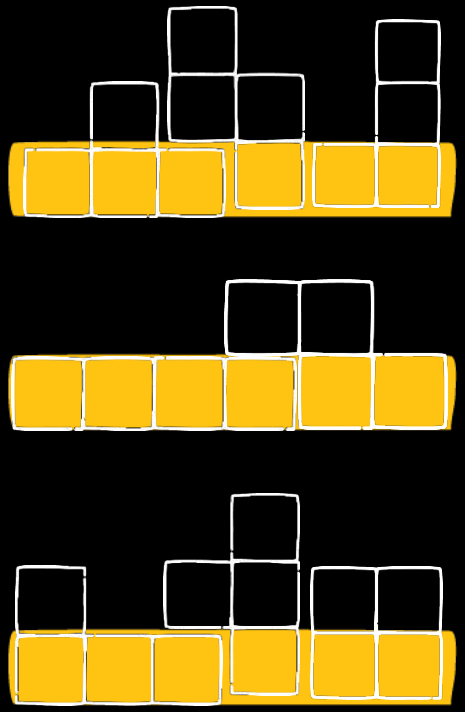
The queue means a user wait is extended by each user in front of them





QUERY'S BUILDING BLOCKS: **THE INSTRUCTION SETS**

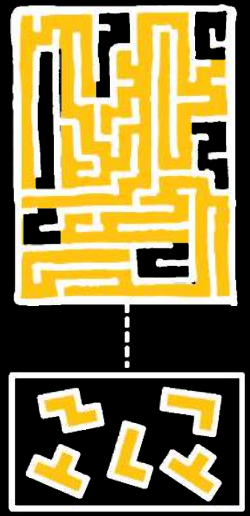




CROWD SPEED: MACHINE LEARNING ARCHITECTURE

1. Break each query into parts
2. Store each ‘query part' and learn
3. Build new queries with matching parts to boost performance

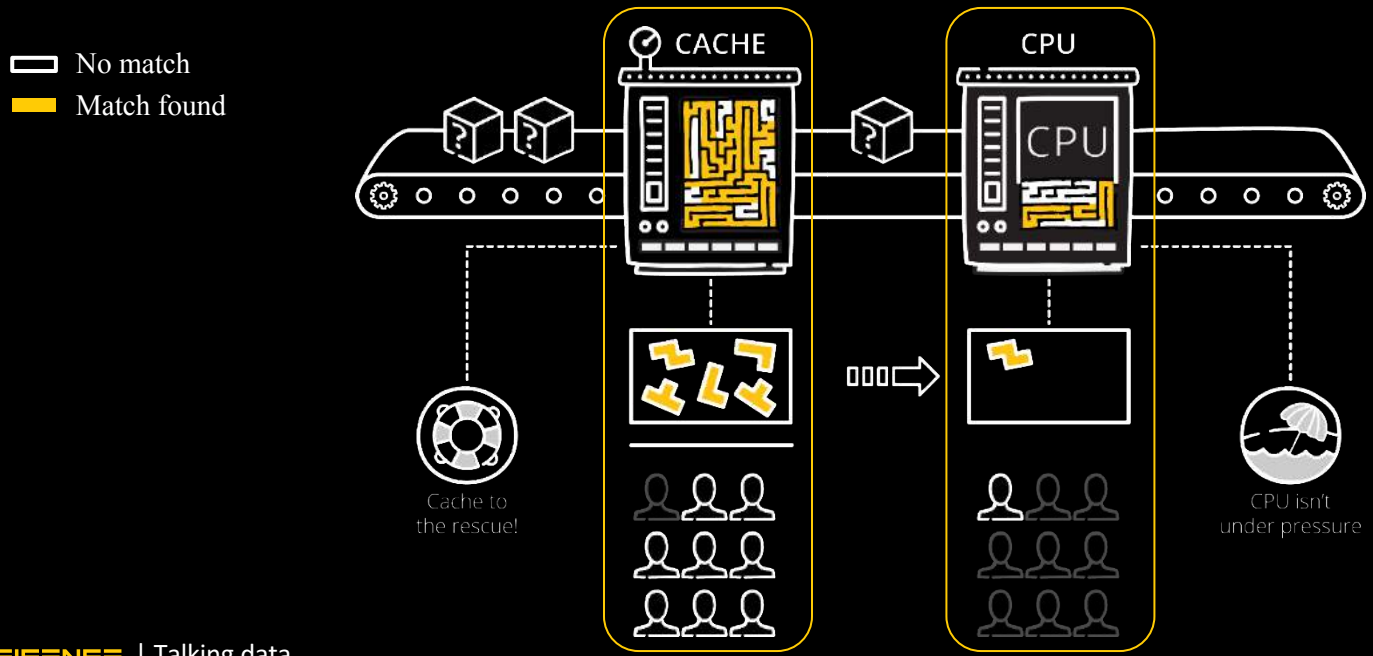
QUERY EXECUTION SPEED



RE-USE REPEATING INSTRUCTION SETS ACROSS QUERIES

MACHINE LEARNING BI

With Machine Learning BI, analytics get faster even when queries are not identical.
The more questions you throw at it - the more efficient it gets!
More users = more queries = faster results



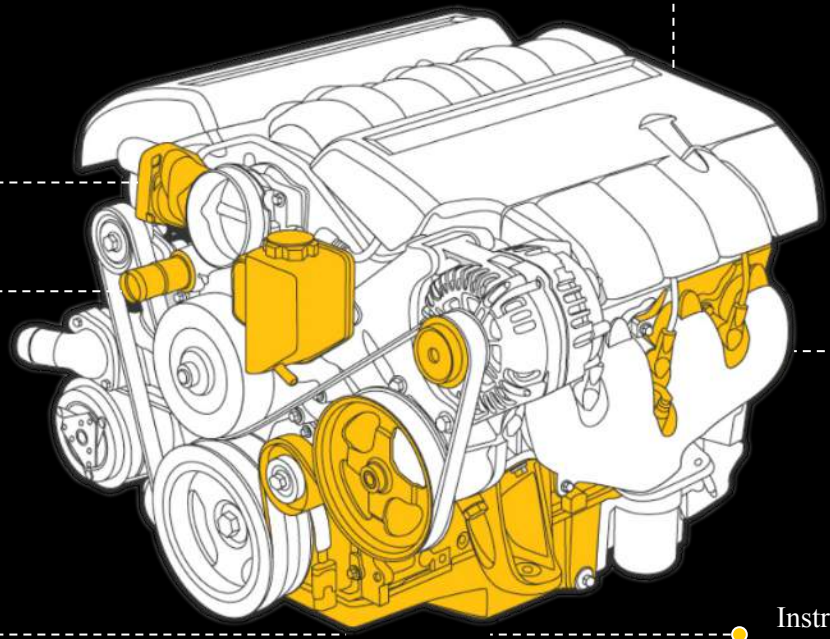
IN-CHIP = POWER + MACHINE LEARNING

- Leverage the unique in-chip cache memory to perform faster than in-memory
  - Without the limitation of having to load the entire model into RAM
- In-Chip recognizes the CPU specs and applies its unique code to organize the query data in the CPU
  - When needed again, **that piece of data exists in the CPU cache**, which is much faster than RAM
- In-Chip **machine-learns** to fetch the associated compressed result sets in advance
  - Sub-query results **pre-loaded into L1 cache** as compressed data
- Decompressed images of that same data can be moved to the larger, but slower, L2 and L3 caches
  - **Decompression operations** (read from and write to cache) **are extremely fast**

IN-CHIP **TECHNOLOGY**

The best engine beer can buy

In memory columnar execution mode



Columnar storage

Full 64BIT support

CACHE aware query kernel

CACHE aware decompression

LLVM based compiler with SIMD support

Instruction recycling & learning algorithms

EMPOWERING **GROWTH**, **ANYWHERE**, **EVERYWHERE**, ON **AFFORDABLE** HW

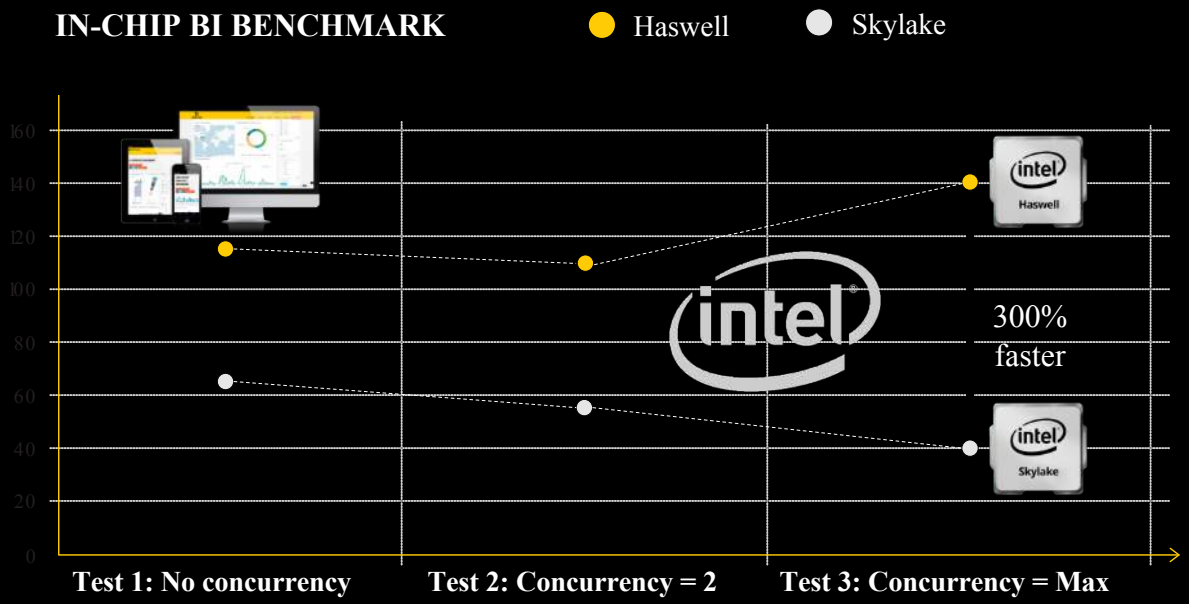
BENCHMARK SETTINGS

Dataset:
120M rows 28GB

8 Analytical queries X 50 cycles

- Aggregations
- Grouping
- Top Ranking
- Large intermediate results

IN-CHIP BI BENCHMARK



Working On
-$900-
workstation

50
Translates into business insights and analytics dashboard:

Analyzing
-200M-
Data Points

Serving 280 Users
In
ONE SEC
Response

SISENSE | Talking data

SPEED! STRATA AWARD

Analyzing 10TB of data in 10 seconds

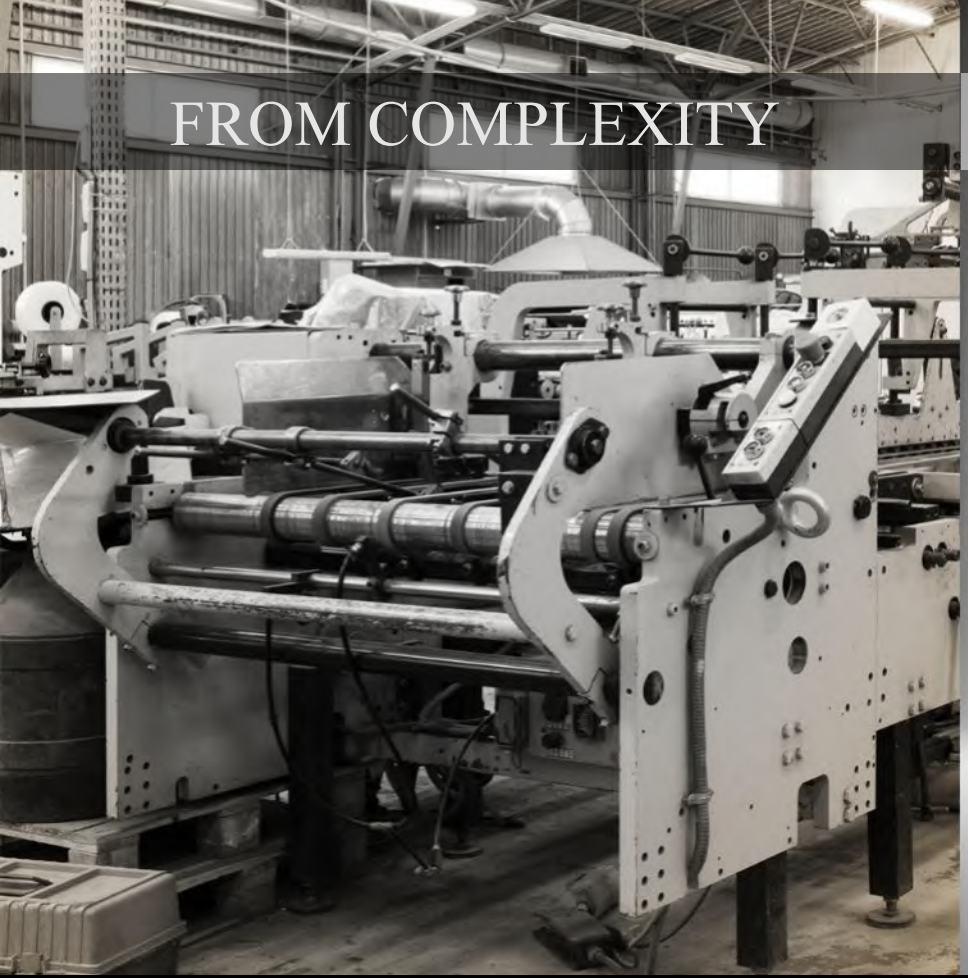
On a single node on a standard Dell Server

REVOLUTION: SCALE-OUT VS IN-CHIP

Architecture	Scale-Out	In-Chip
Users	Data Scientists, IT, Developers	Business Users
Use Cases	ETL, Batch Reports, Machine Learning	Ad-Hoc Analytics
Interface	JAVA, R, C, SQL	Interactive Dashboards, SQL
Time to Implement	Long	Short
Available Resources	Big	Small
Outcome	**Big Data Infrastructure**	**Agile Big Data Analytics**

SO…WHAT IS IT GOOD FOR?

FROM COMPLEXITY



TO SIMPLICITY

TECHNOLOGY HAS NO MEANING IF IT HAS NO IMPACT ON HUMAN LIFE

"If a tree falls in a forest and no one is around to hear it, does it make a sound?"

OVERHYPE OF BUZZWORDS

THE PERSONAL, INTELLIGENT AND CONTEXTUAL WEB

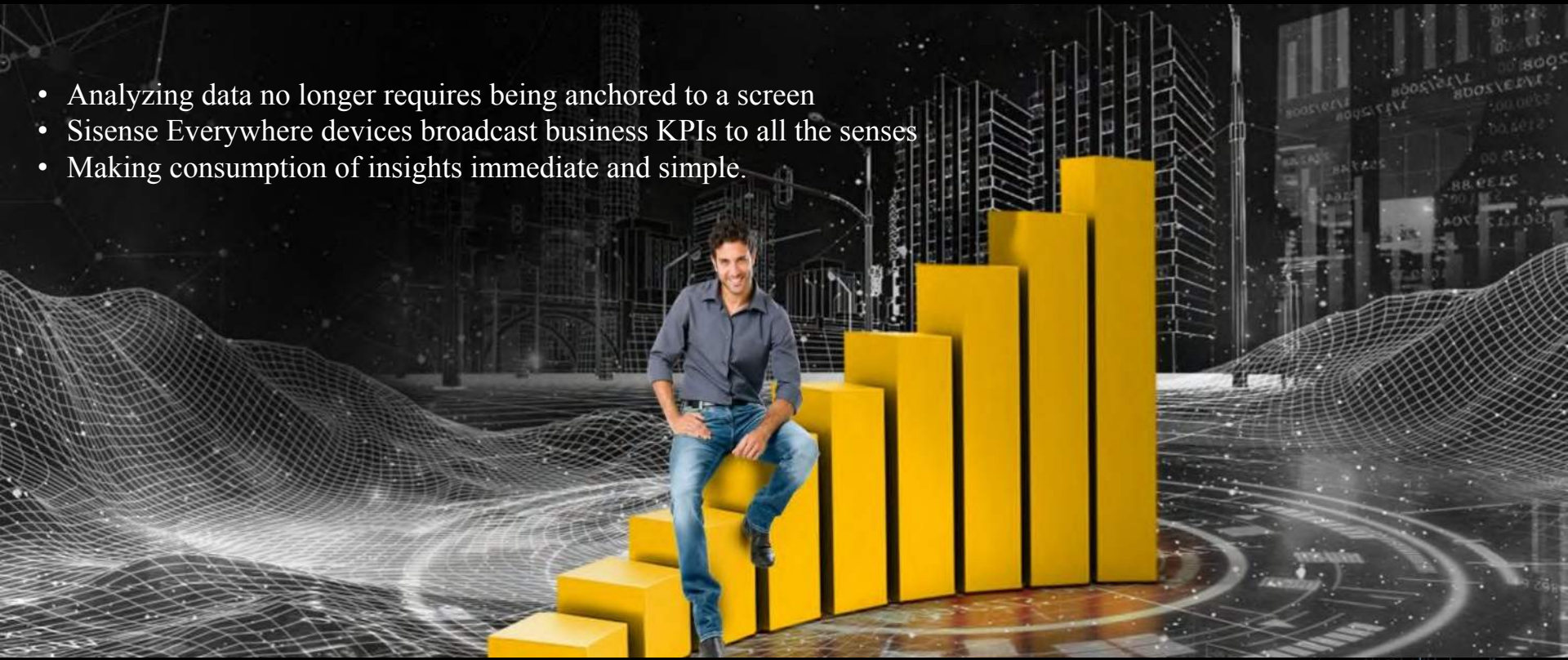
THE INTERNET OF ME

TRANSFORMATION OF BIG DATA ANALYTICS FOR IOM

WE ARE ALL **UNIQUE**

LET THERE BE **LIGHTS**

THE NEXT REALM OF **BUSINESS ANALYTICS**



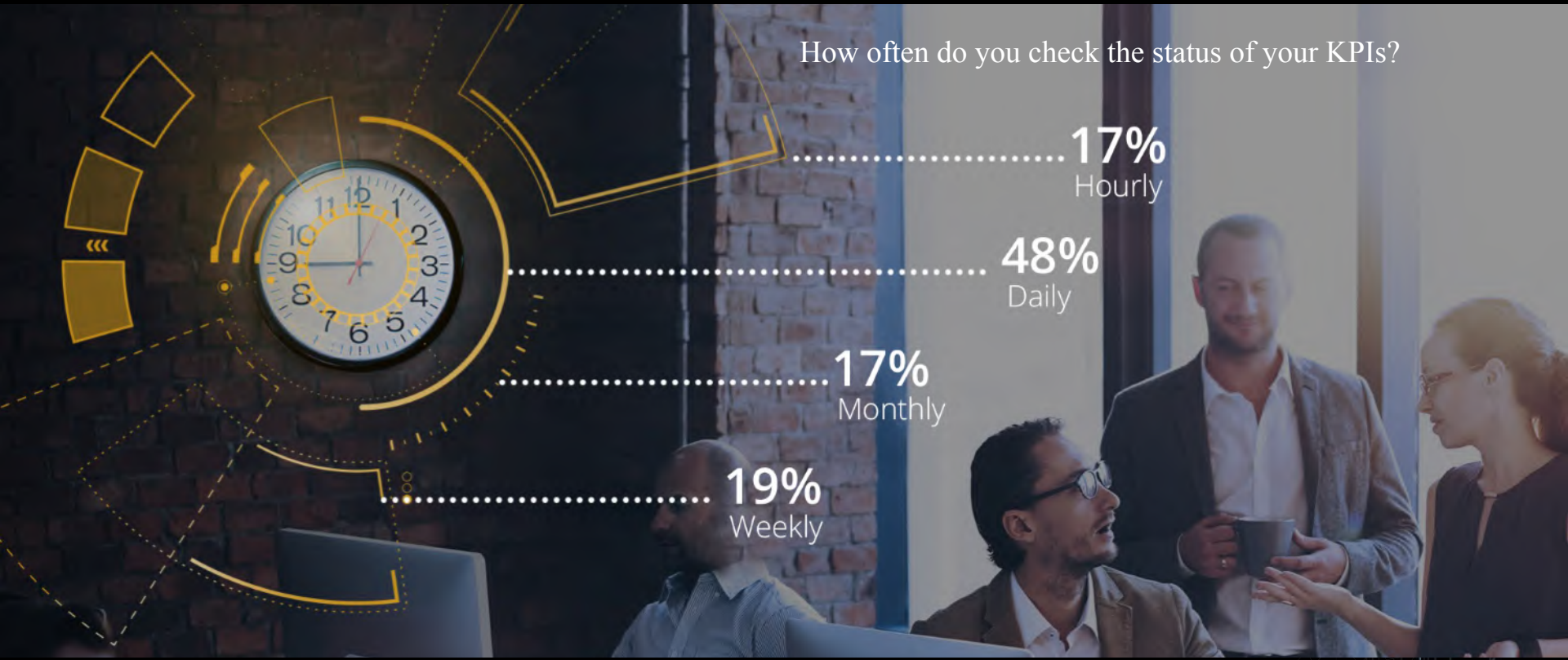
- Analyzing data no longer requires being anchored to a screen
- Sisense Everywhere devices broadcast business KPIs to all the senses
- Making consumption of insights immediate and simple.

HOW IT ALL **STARTED**

The relationship between business professionals and their KPIs

(We asked hundreds of business professionals how they interact with their data and KPIs)

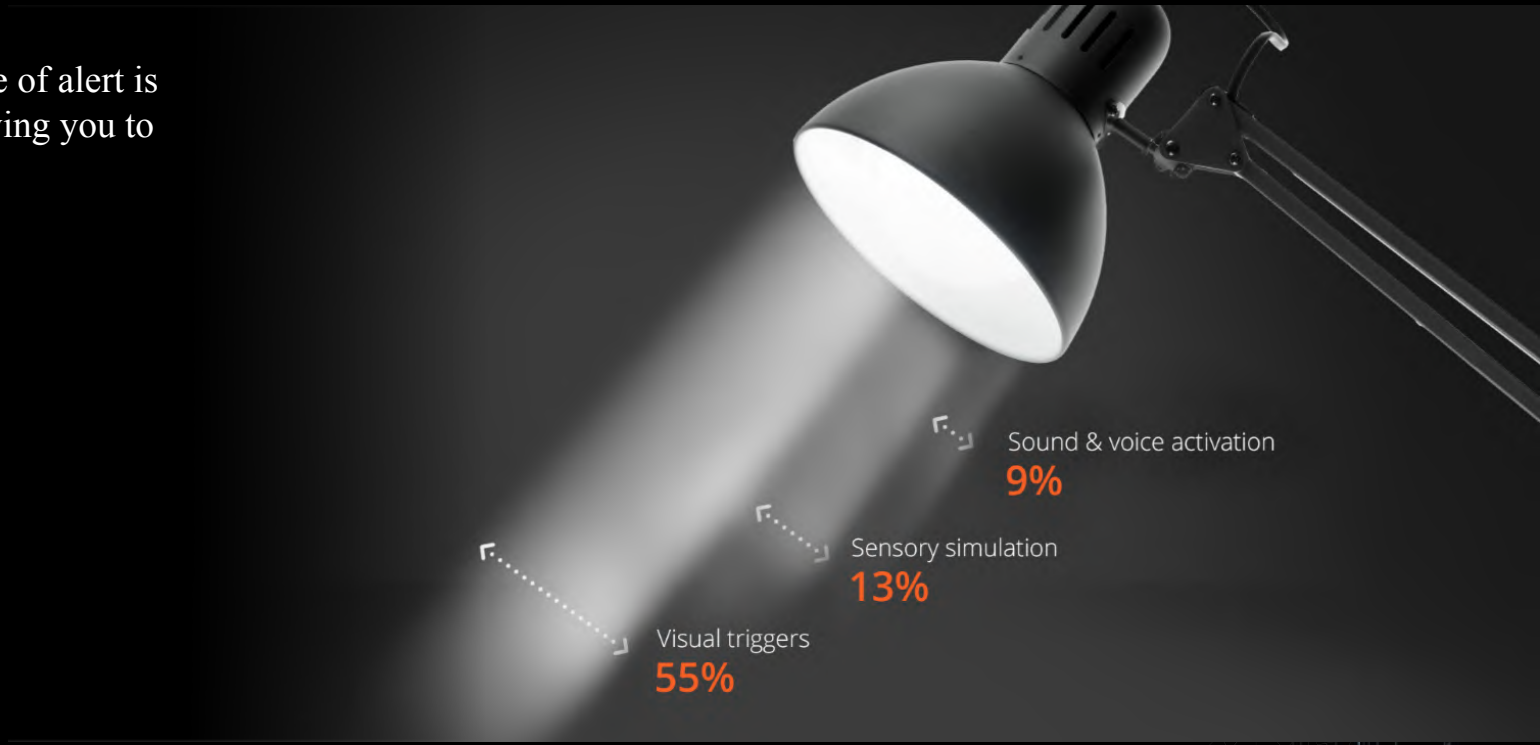
ALMOST HALF OF ALL RESPONDENTS CHECK KPIS DAILY



83% OF RESPONDENTS USE OR WANT TO USE COLOR CODING

VISUAL ALERTS ARE THE MOST EFFECTIVE, ACCORDING TO MORE THAN HALF OF RESPONDENTS

Which type of alert is best in driving you to action?



ACCORDING TO RESPONDENTS THE FUTURE OF BI CONSUMPTION **IS EVERYWHERE**

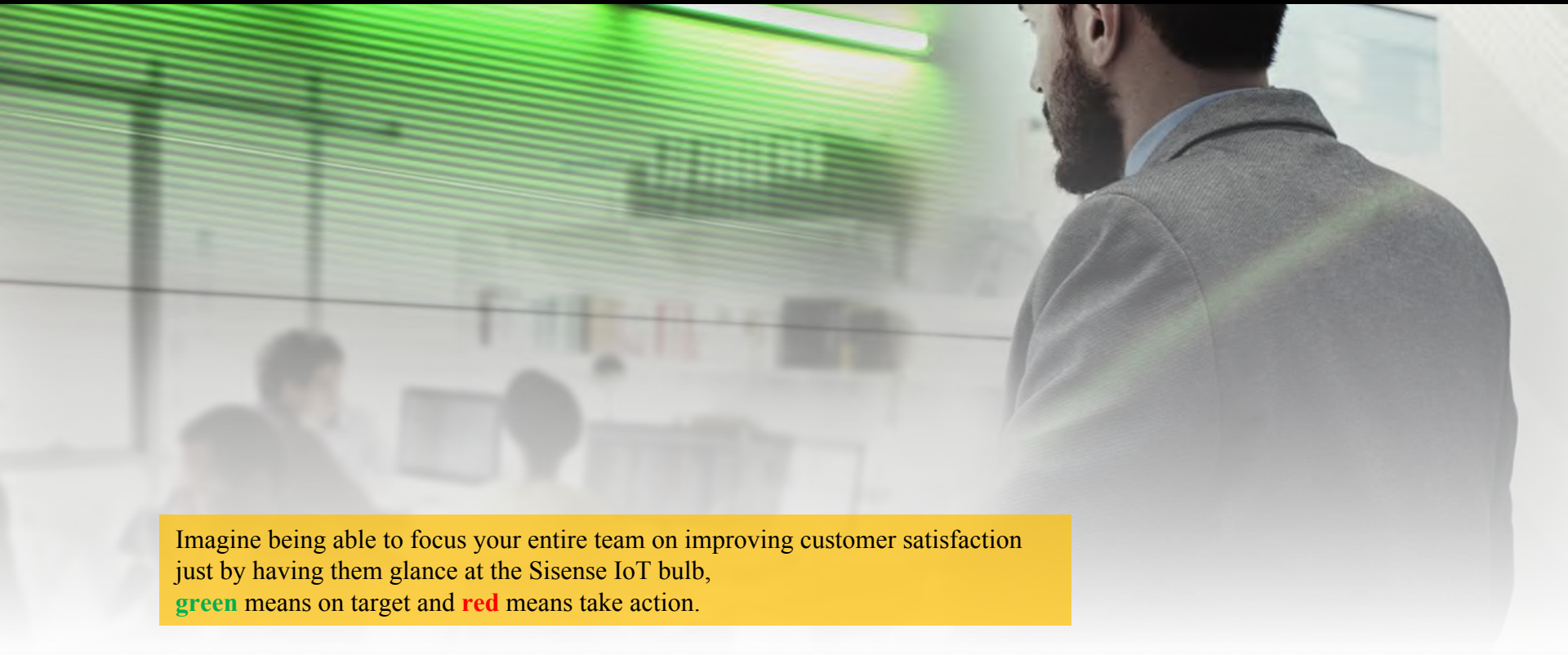
BI EVERYWHERE

Revolutionizing The Way Business Users Consume Data

IMAGINE A **SISENSE** WORLD

Imagine you're driving to work and can ask your voice-operated BI assistant: "What is my sales target for today?"

IMAGINE A SISENSE WORLD



Imagine being able to focus your entire team on improving customer satisfaction just by having them glance at the Sisense IoT bulb,
green means on target and **red** means take action.

IMAGINE A **SISENSE** WORLD

SISENSE BRINGS IMAGINATION TO LIFE

Sisense-Enabled is a new line of devices that present data unlike any dashboard environment

SISENSE LAMP

SISENSE ENABLED ECHO

REDEFINING HOW WE INTERACT WITH DATA

"When I see that bulb change, I get a real sense of satisfaction. It's provided a direct way for us to see how data is changing. The bulb gives me peace of mind because I can see a light change rather than monitoring a screen."

CTSI

RESPOND TO CHANGES IN REAL-TIME

"I think I find it easier to relate to color and sound than a dashboard.
I have seen a change in my behavior using these tools, specifically around time to react - understand when something is changing and going to look
at metrics to find out why."

act|on

LIVE CLOSER TO YOUR DATA

"Bulb is the KPI that you don't need to load up on one of your screen, it's not just another browser window. It's this physical piece that's simply part of your life. It's a simple product with a powerful way of telling you whether things are going well."

Skullcandy

SIMPLIFYING COMPLEX DATA CONSUMPTION

MAINTAIN FOCUS

Keep teams focused on
a common goal and in touch with
your business.

GAIN CONSTANT VISIBILITY

Know what's happening,
wherever you are,
in an instant.

STAY CONNECTED

Keep your finger on
the pulse and act on what's
important.

THE FUTURE OF ENHANCED HUMANISM

SISENSE

T11

2017

THANKS

TalkingData