

Federated Analytics in the Dawn of the Digital Era

GLOBAL PRODUCT DATA INTEROPERABILITY **S U M M I T** 2017



ELYSIUM

Parker Aerospace

NORTHROP GRUMMAN

BOEING

ELYSIUM

Parker Aerospace

NORTHROP GRUMMAN

BOEING



Federated Analytics in The Dawn of the Digital Era

Patricia Florissi, PH D
VP and Global CTO for Sales

DELLEMC

There times in life where...
One must pause, think, and...
Try to grasp the moment we are living...

Many factors at play

DELLEMC



Incremental computing
reaching near zero
marginal cost



Internet
of things



Exponential
technology advances
every 5 years



Cloud



Big data
analytics



Social



Mobile

Convergence of these Many factors at play

DELLEMC

Dawn of a new era

DELL EMC

Dawn of the digital era

DELL EMC

What does this digital tide bring?

DELL EMC

8

© Copyright 2017 Dell Inc.

The things the tide brings

1
DATA SIZE
MATTERS

2
DATA
ECOSYSTEM
MATTERS

3
DATA
LOCALITY
MATTERS

DELL EMC

The things the tide brings

1
DATA SIZE
MATTERS

2
DATA
ECOSYSTEM
MATTERS

3
DATA
LOCALITY
MATTERS

DELLEMC

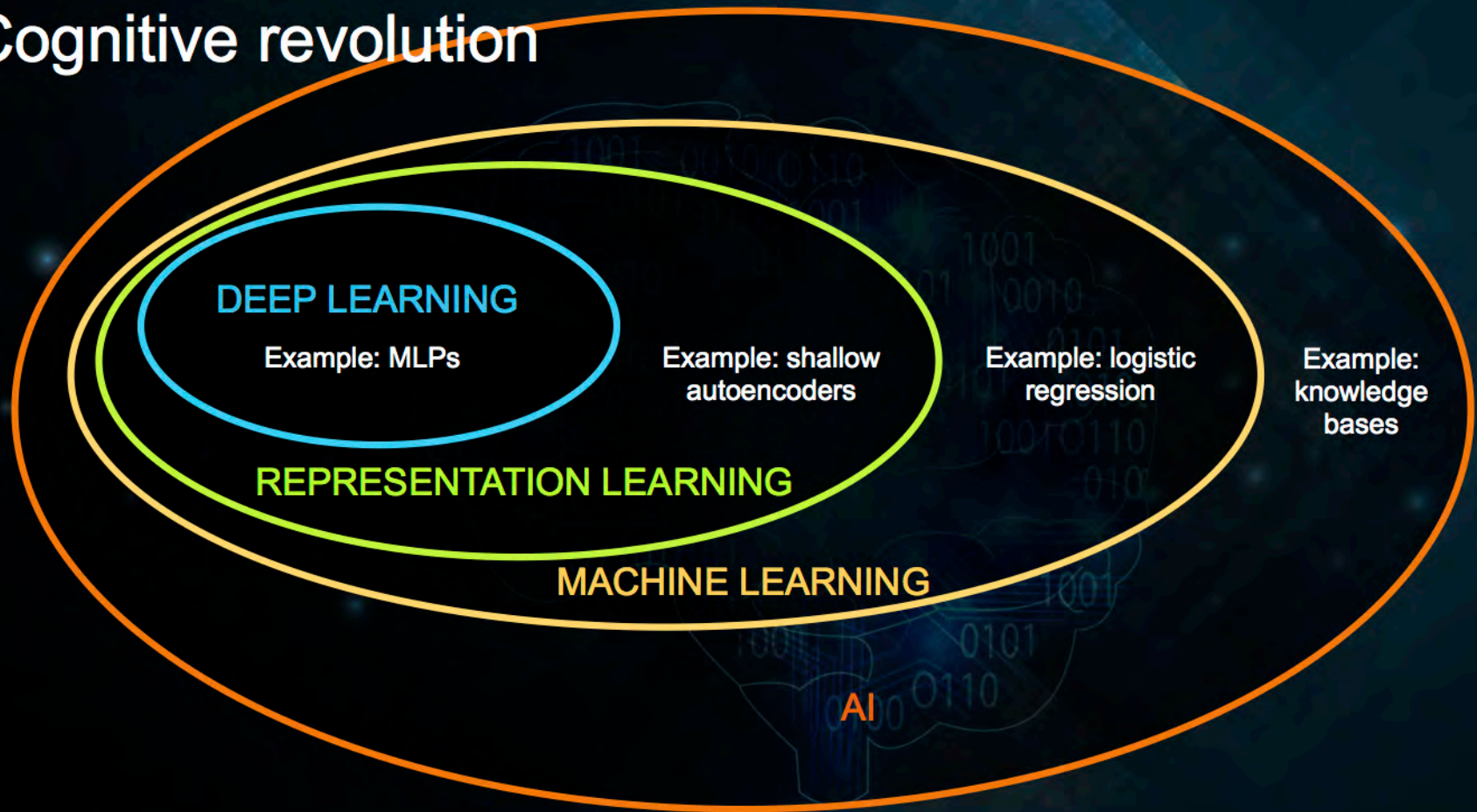
Cognitive revolution 2.0

Artificial intelligence

The theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.

DELL EMC

Cognitive revolution

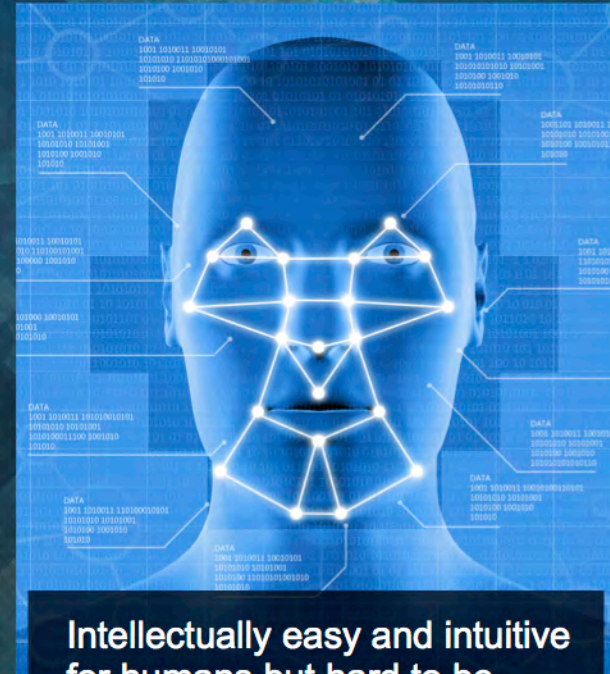


In the early days



Intellectually difficult for human being but relatively straight forward for computers

In the now days

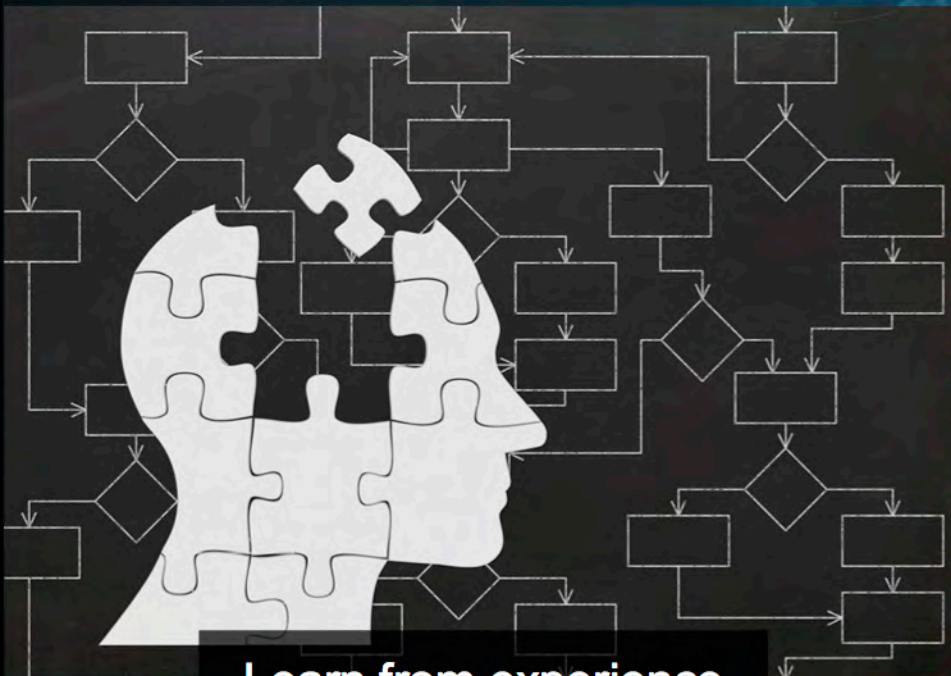


Intellectually easy and intuitive for humans but hard to be described formally for computers

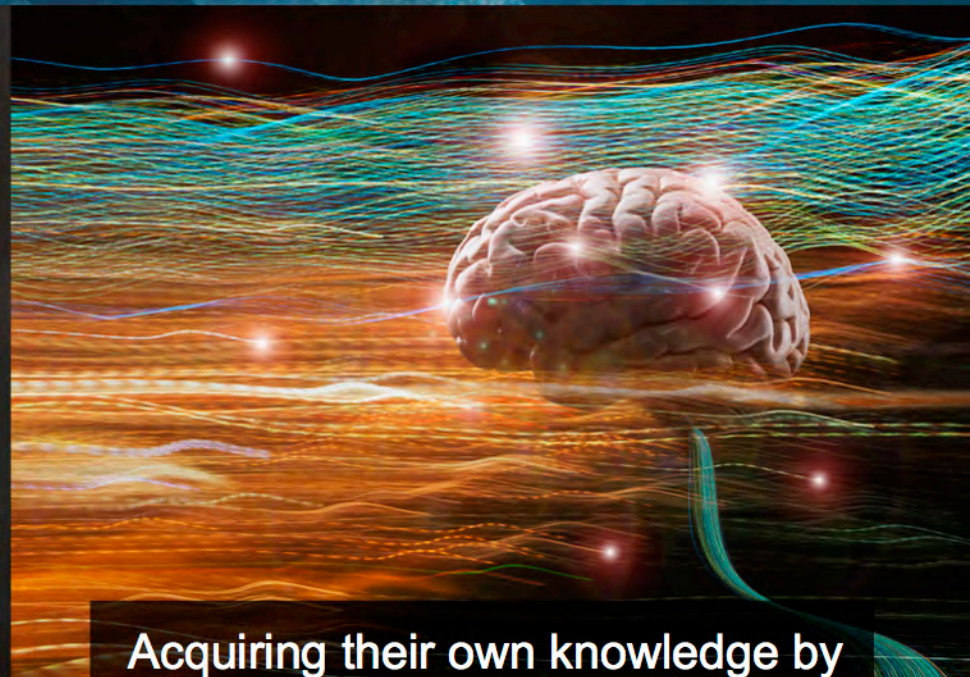
Deep Learning, Ian Goodfellow, Yoshua Bengio, and Aaron Courville

DELL EMC

Machine learning

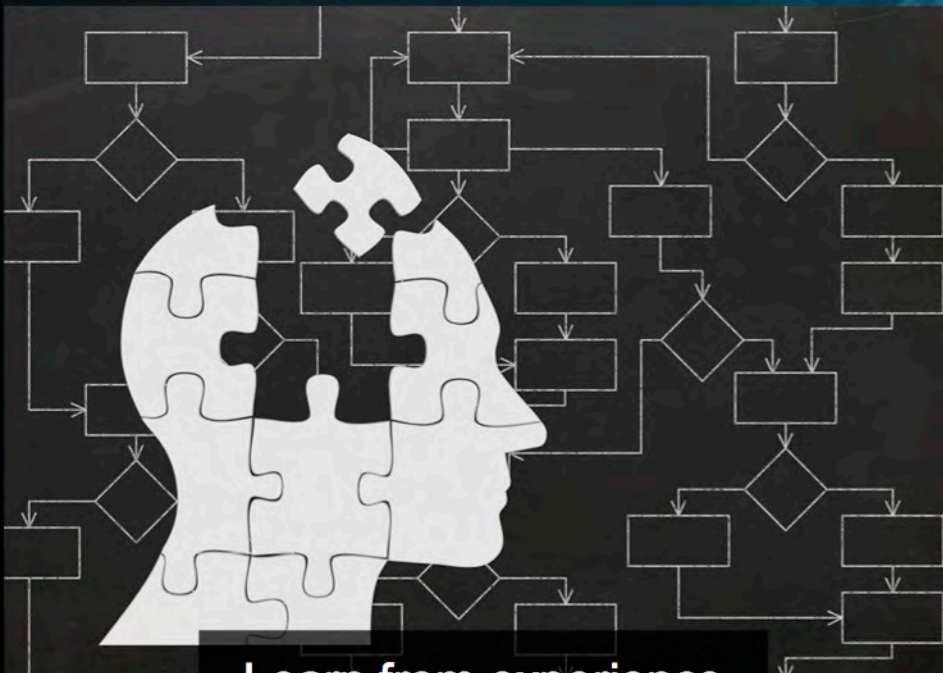


Learn from experience



Acquiring their own knowledge by
extracting patterns from raw data

Deep learning



Learn from experience



Understand the world in terms of a hierarchy of concepts, with each concept defined through its relation to simpler concepts

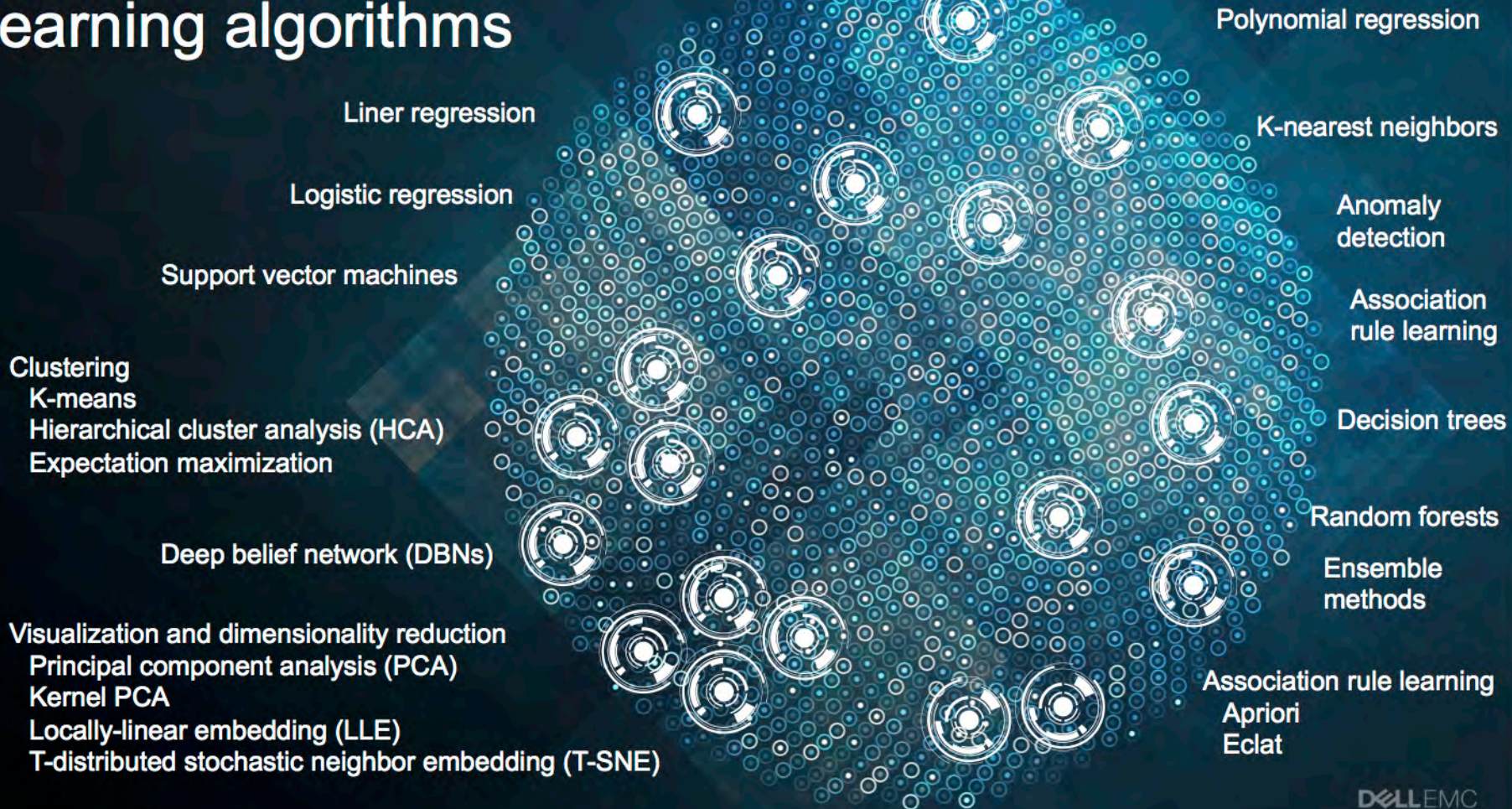
Why is it called deep learning

Represents the world as a nested hierarchy of concepts

The depth of the representational graphs high

Greater amount of composition of learned functions or learned concepts than traditional machine learning does

Learning algorithms



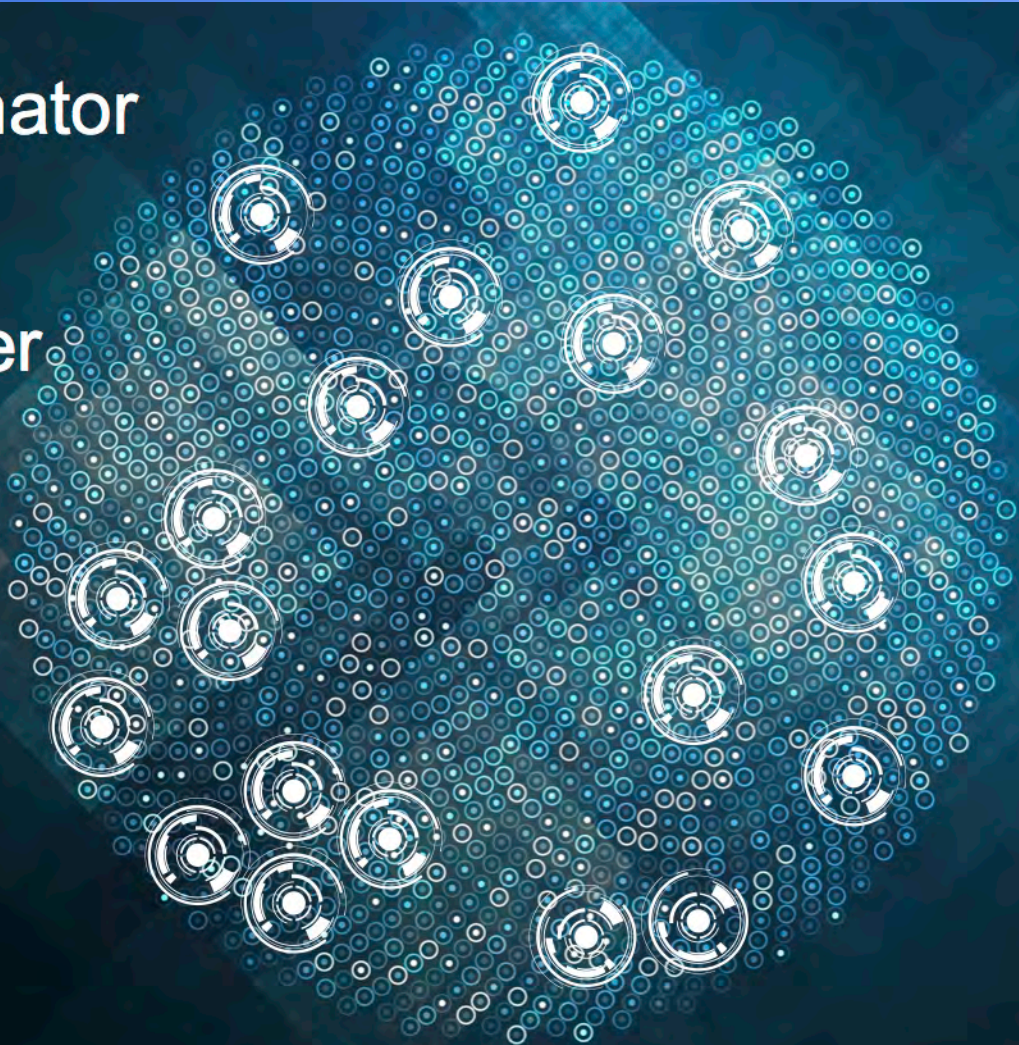
DELLEMC

Common Denominator

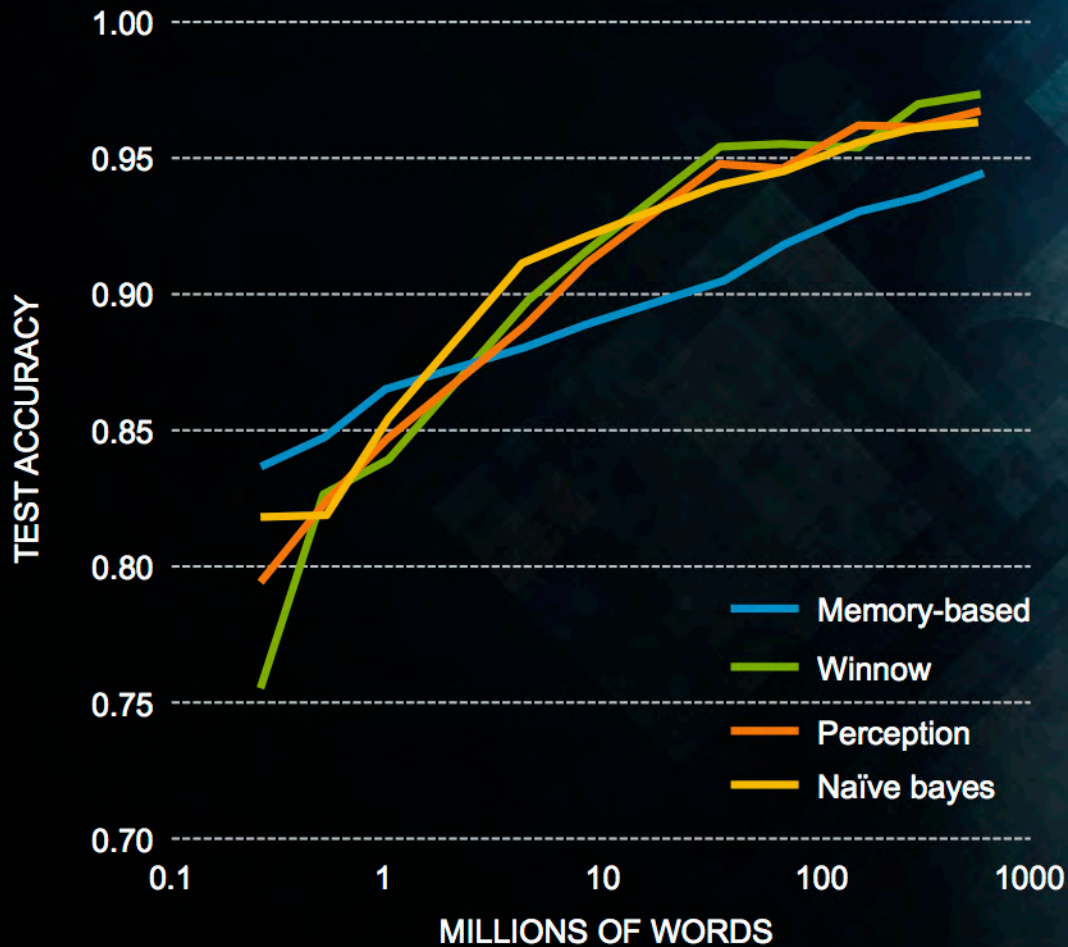
They perform better

The more data


They analyze



DELLEMC



The unreasonable effectiveness of data



Simple algorithms with lots of data
will out perform
Sophisticated algorithms with less data

DELL EMC

The things the tide brings

1
DATA SIZE
MATTERS

2
DATA
ECOSYSTEM
MATTERS

3
DATA
LOCALITY
MATTERS

DELLEMC

The things the tide brings

DATA SIZE
MATTERS

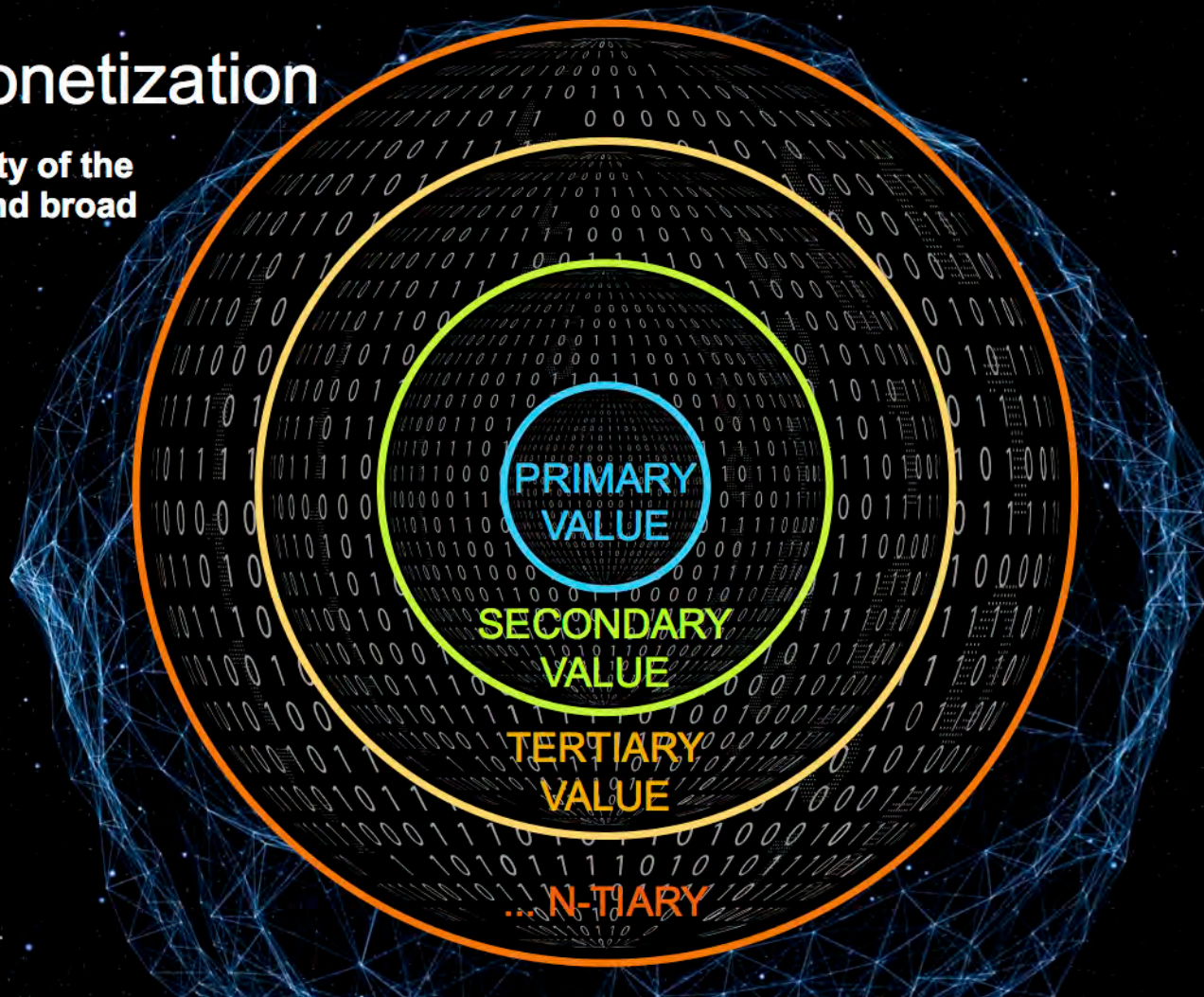
2
DATA
ECOSYSTEM
MATTERS

3
DATA
LOCALITY
MATTERS

DELLEMC

Data monetization

The applicability of the data is large and broad



Source: Glenn Allmendinger

24 © Copyright 2017 Dell Inc.

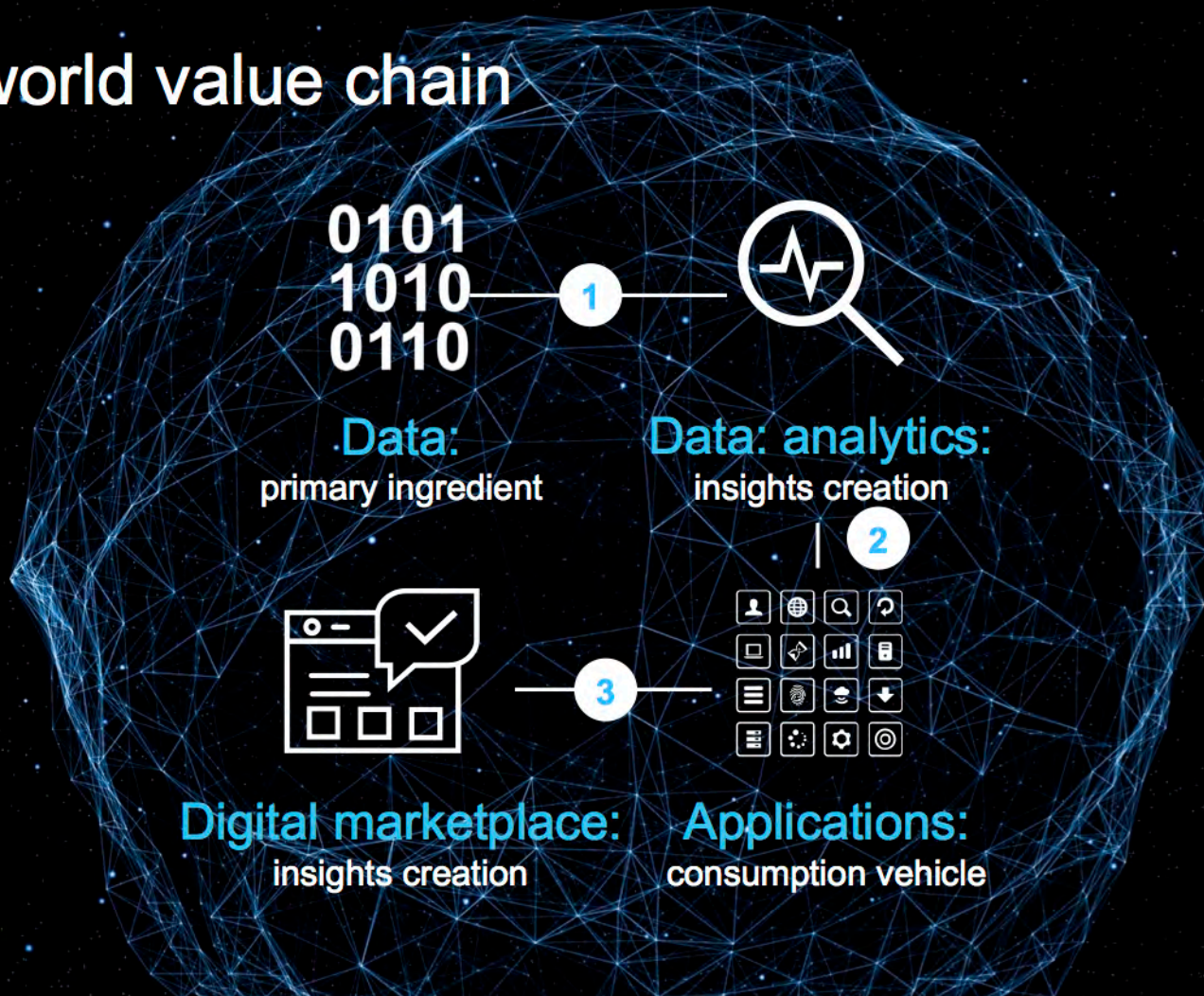
DELLEMC

Data fusion

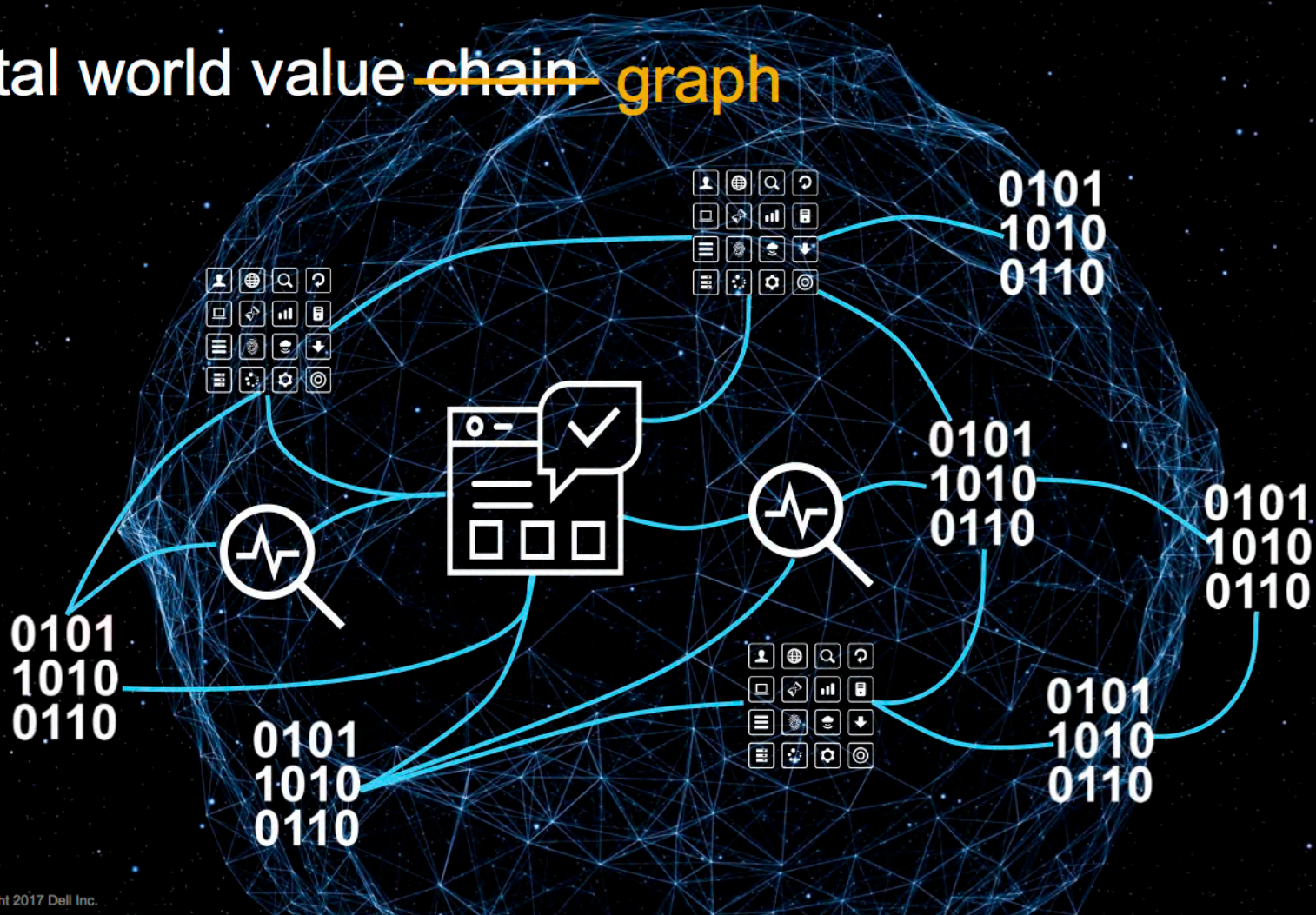
Occurs when different sources are brought into contact and new facts emerge.

DELLEMC

Digital world value chain



Digital world value ~~chain~~ graph



The things the tide brings

DATA SIZE
MATTERS

DATA
ECOSYSTEM
MATTERS

3
DATA
LOCALITY
MATTERS

DELLEMC

Major data analytics trends emerging

And at the beginning
there was one lake...

Centralization

As the traditional approach to gain
analytical insight from data in
disparate sources

Batched data gathered in a single
repository or cloud in preparation
for analysis

DELL EMC

The perfect data in-mobility storm

IoT effect

Bandwidth constraints

Regulatory compliance

Traditional Data Systems

Multi-Cloud

DELLEMG

Challenge imposed by traditional architecture thinking

Understanding the data challenge

DATA HANDLING CHALLENGE

Massive numbers of endpoints for systems

Storage requirement is unknown

Managing massive, heterogeneous data an enormous challenge.



Diverse data sources -> diverse protocols, diverse transport requirements, diverse data

Must normalize and standardize for future data orchestration and integration

Network

DATA USAGE CHALLENGE

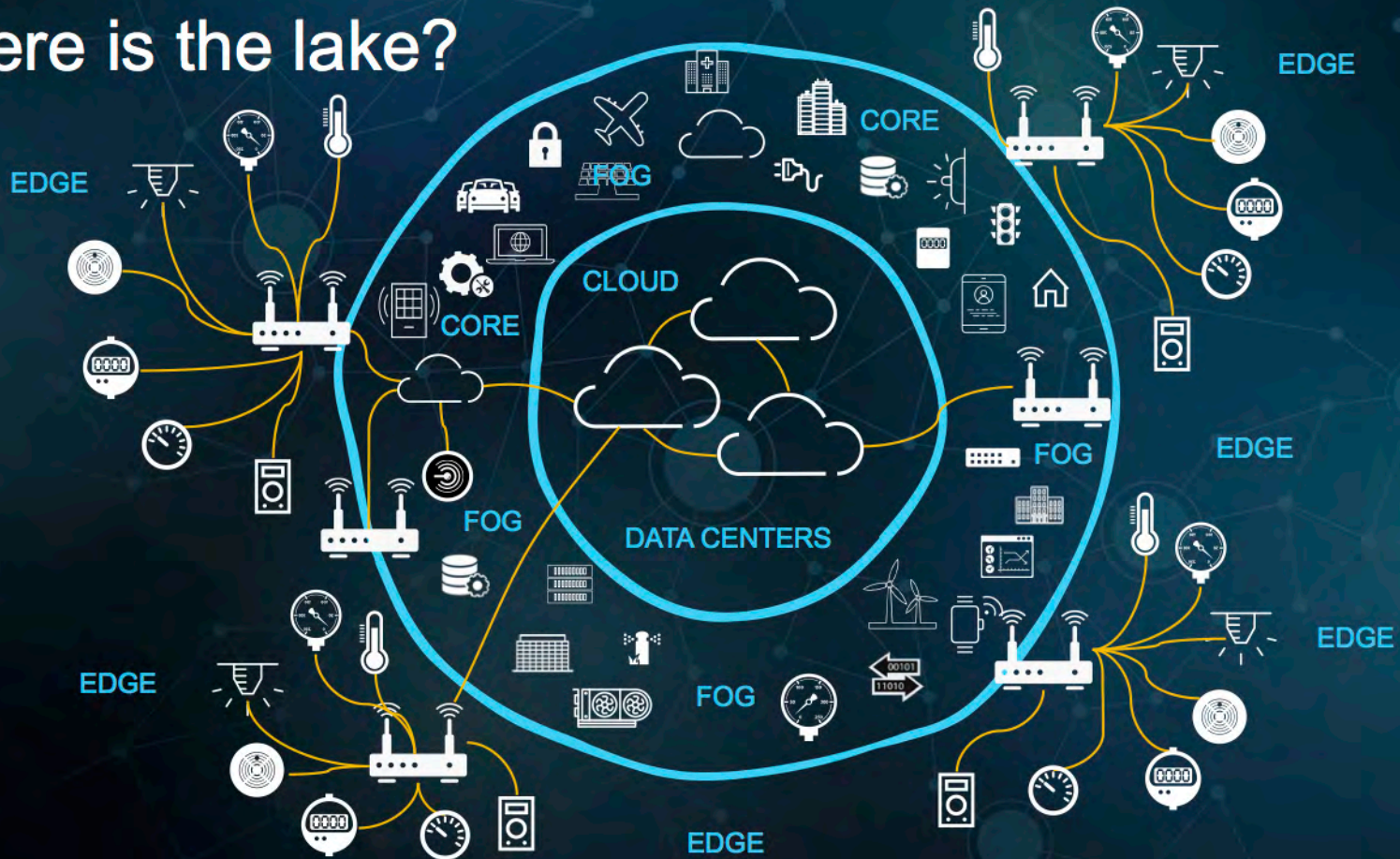


Need to create an innovative business that uses this data

->

creates more value and fully uses current investment

Where is the lake?



The things the tide brings

**DATA SIZE
MATTERS**

**DATA
ECOSYSTEM
MATTERS**

**DATA
LOCALITY
MATTERS**

DELLEMC

The things the tide brings...



Incremental computing
reaching near zero
marginal cost



Internet
of things



Exponential
technology advances
every 5 years



Cloud



Big data
analytics



Social



Mobile



DELLEMC

In the dawn of the digital era ...

DELL EMC

The dawn of federated analytics

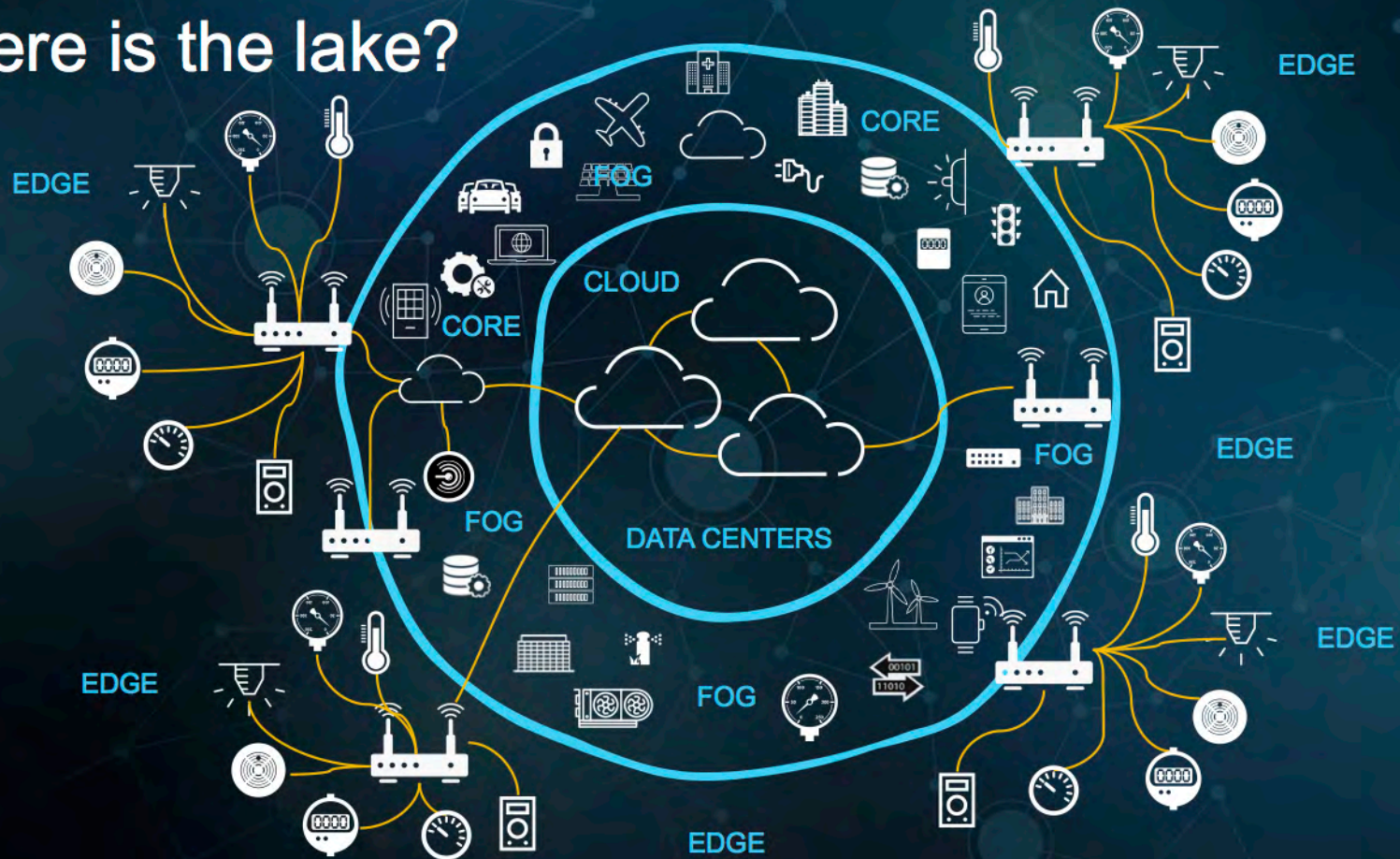
DELL EMC

Federated analytics

Distributed analysis of data in place, as close as possible to the data source, while intermediate results can be shared and can be further analyzed along their path to other locations, enabling higher order learning at scale.

DELLEMC

Where is the lake?



DELLEMC

And at the beginning
there was one lake...

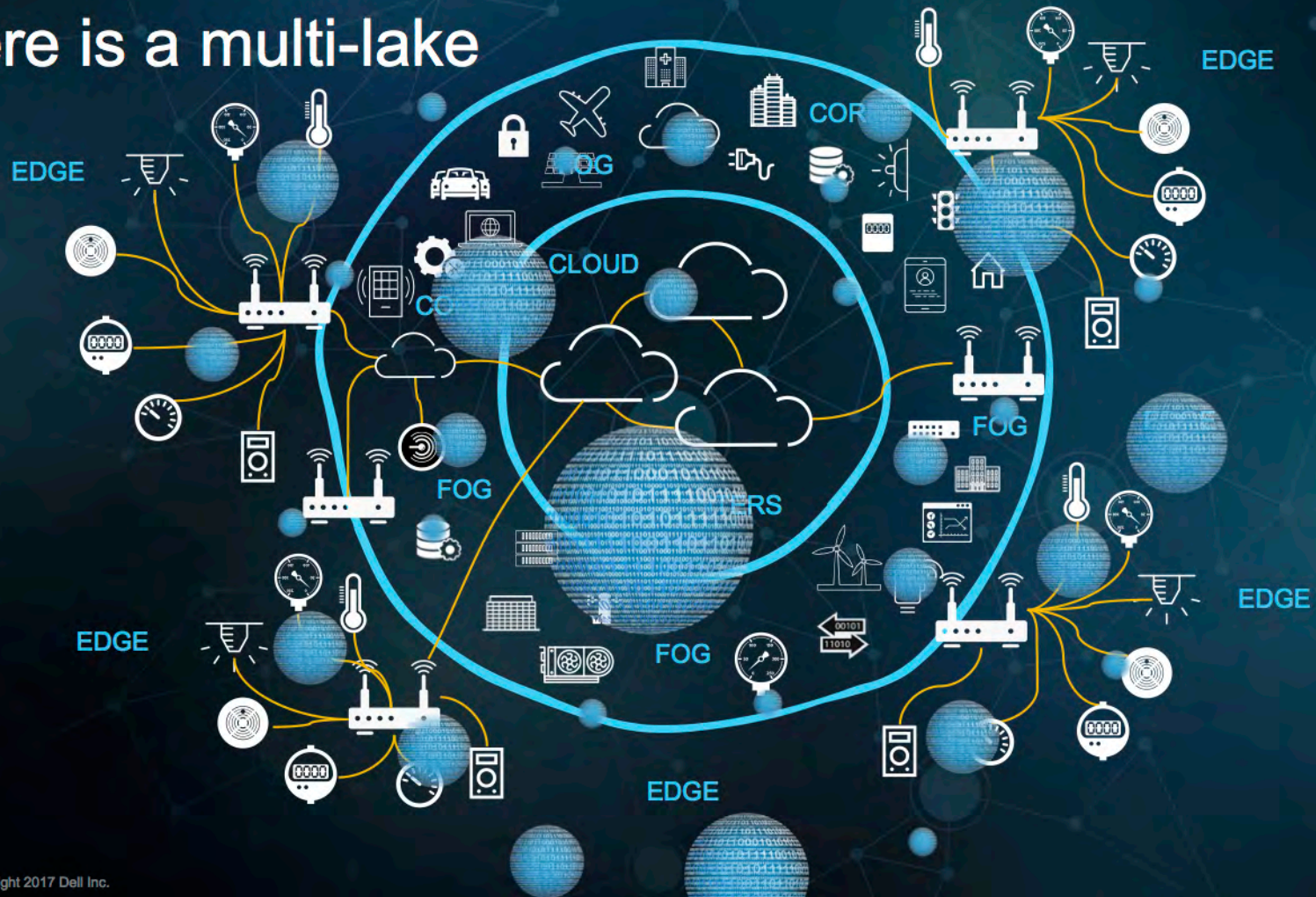
Centralization

As the traditional approach to gain
analytical insight from data in
disparate sources

Batched data gathered in a single
repository or cloud in preparation
for analysis

DELL EMC

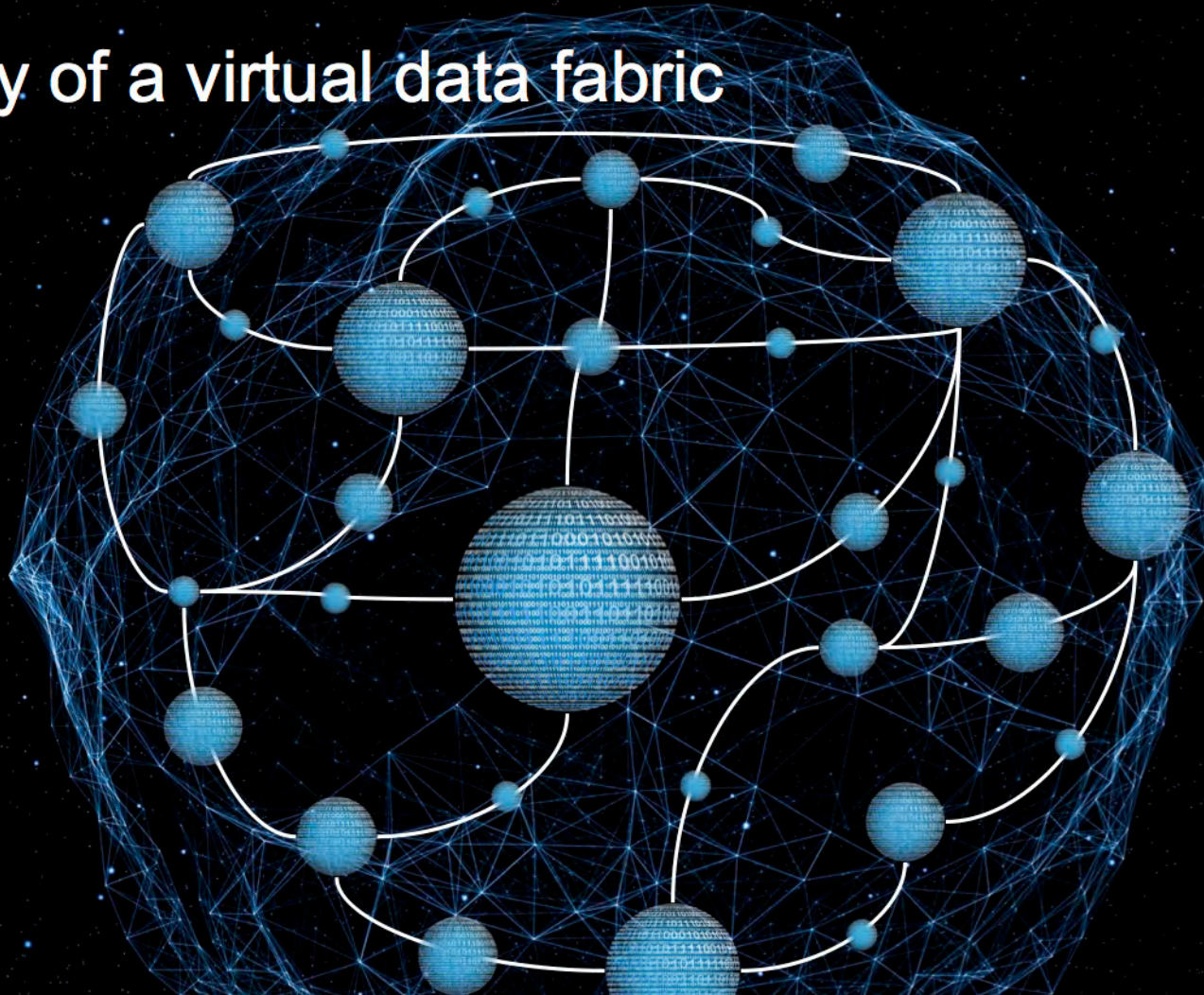
There is a multi-lake



Anatomy of the world's data



Anatomy of a virtual data fabric



Main challenges

how data science algorithms will be re-designed to reason and learn in a federated manner;

how analytics will be distributed close to where the data is collected; and

how the intermediate results will be aggregated and analyzed together to drive higher-order learning at scale. Those are much more challenging problems to solve

DELL EMC

World Wide Herd (WWH)

DELL EMC

Data sources generate data

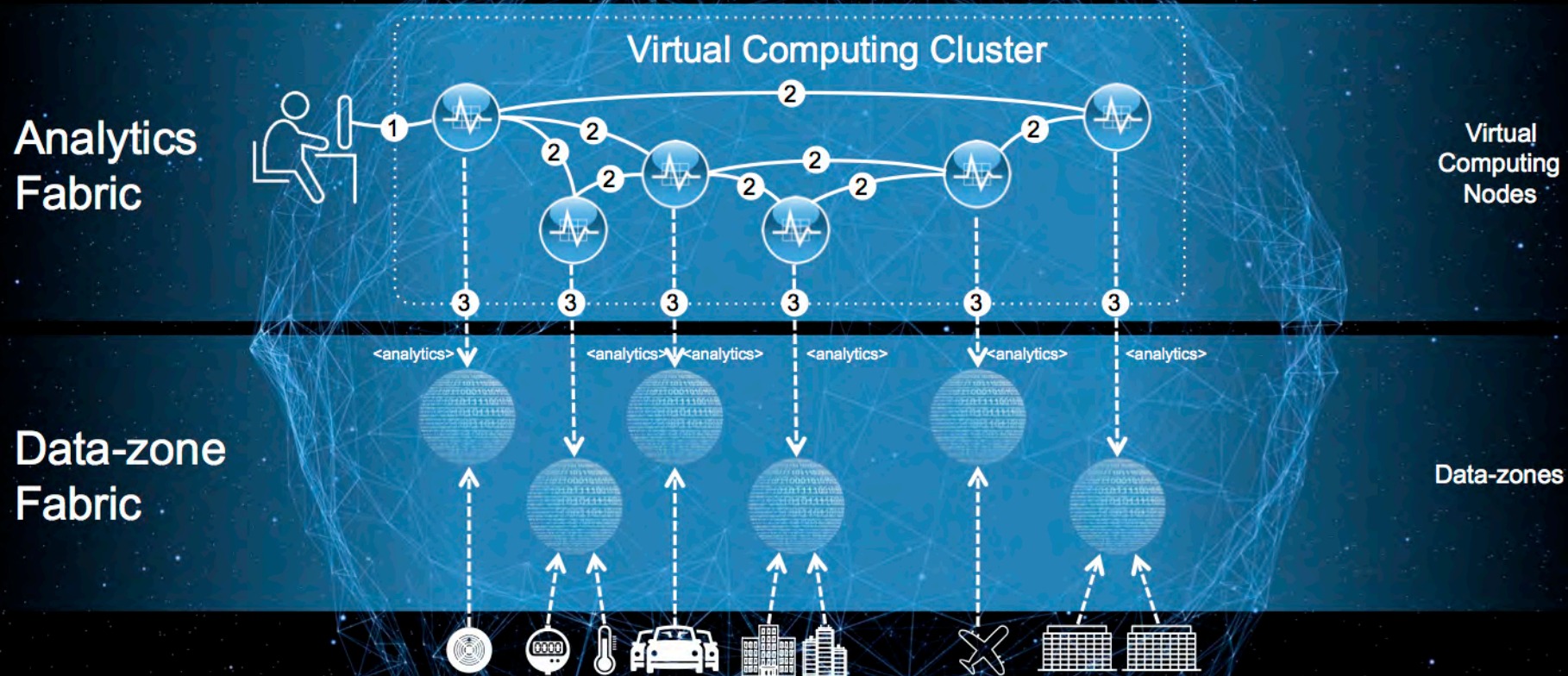
Data-zone
Fabric

Data-zones

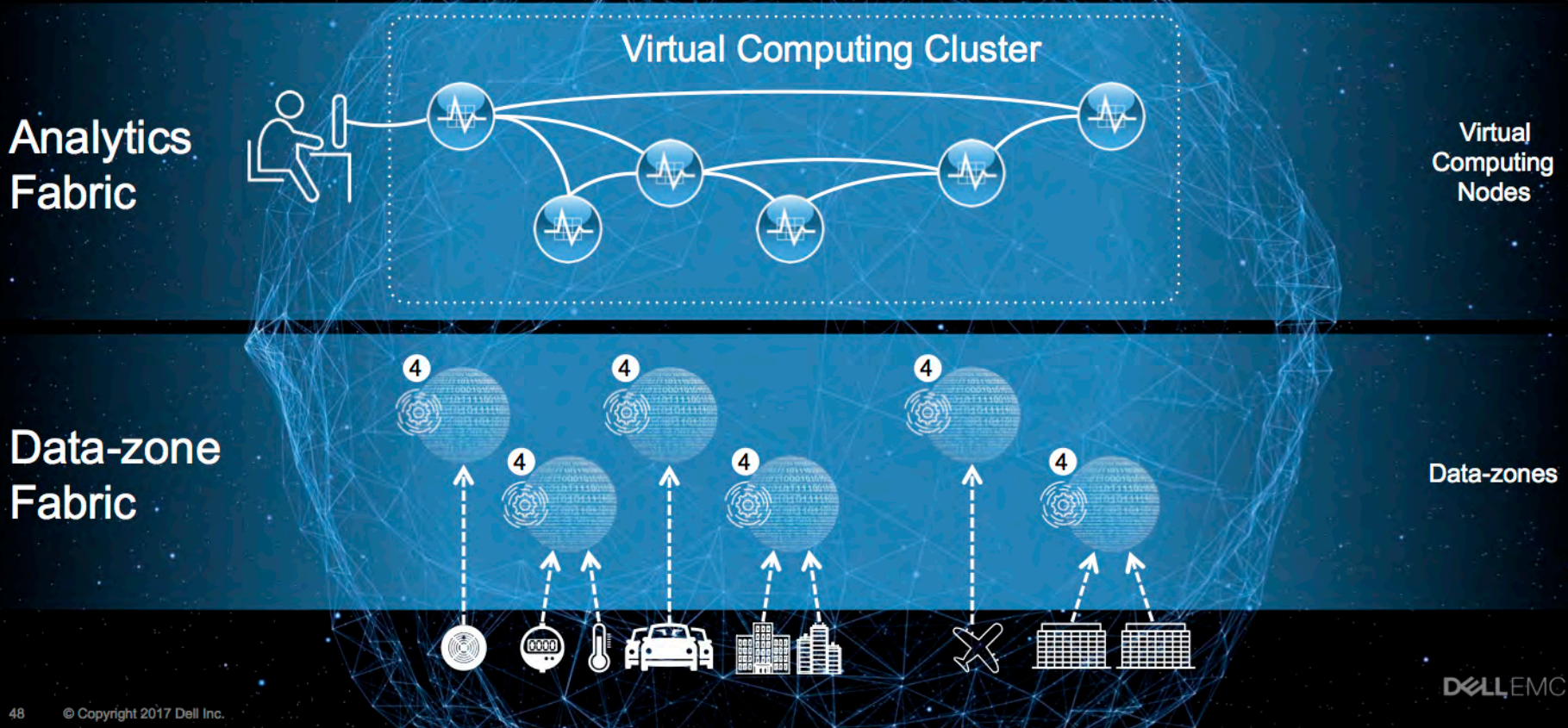


DELL EMC

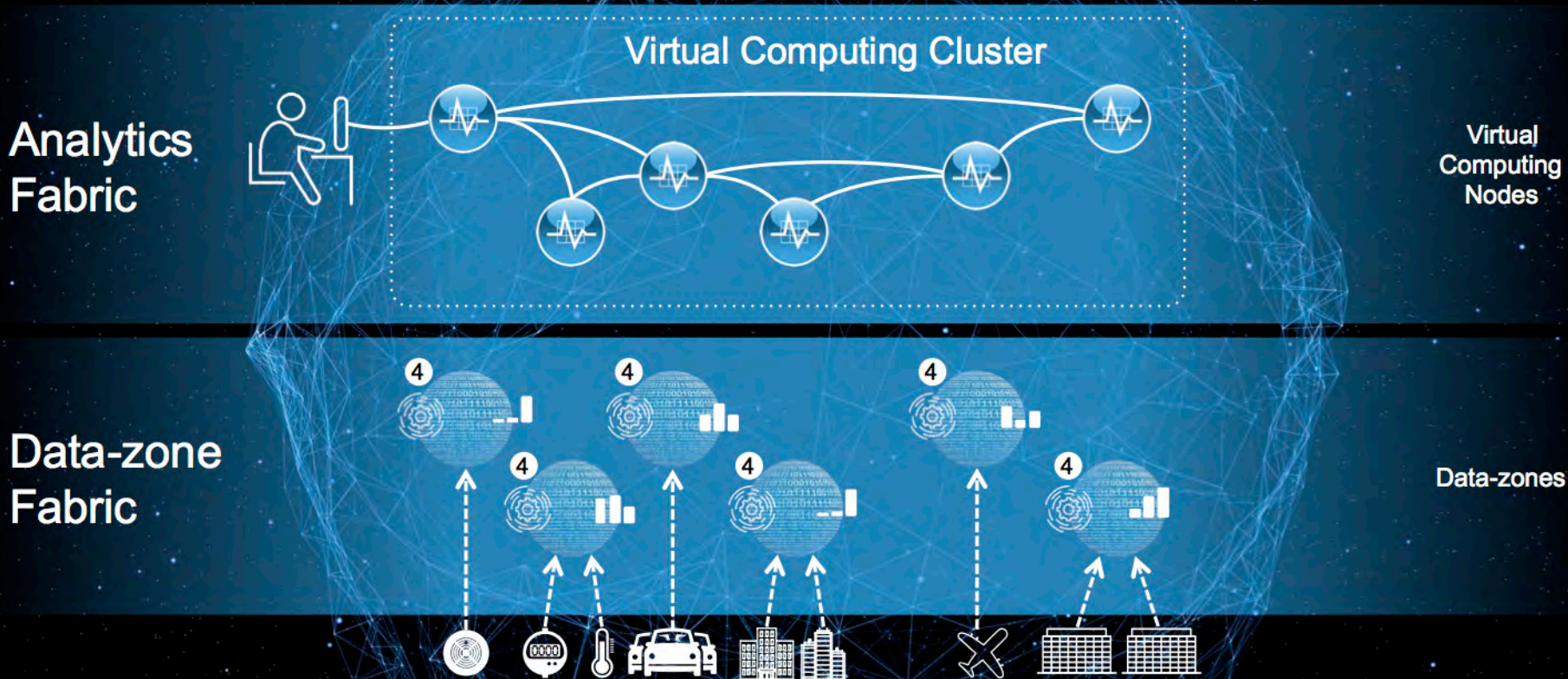
Along comes the analytics explorer



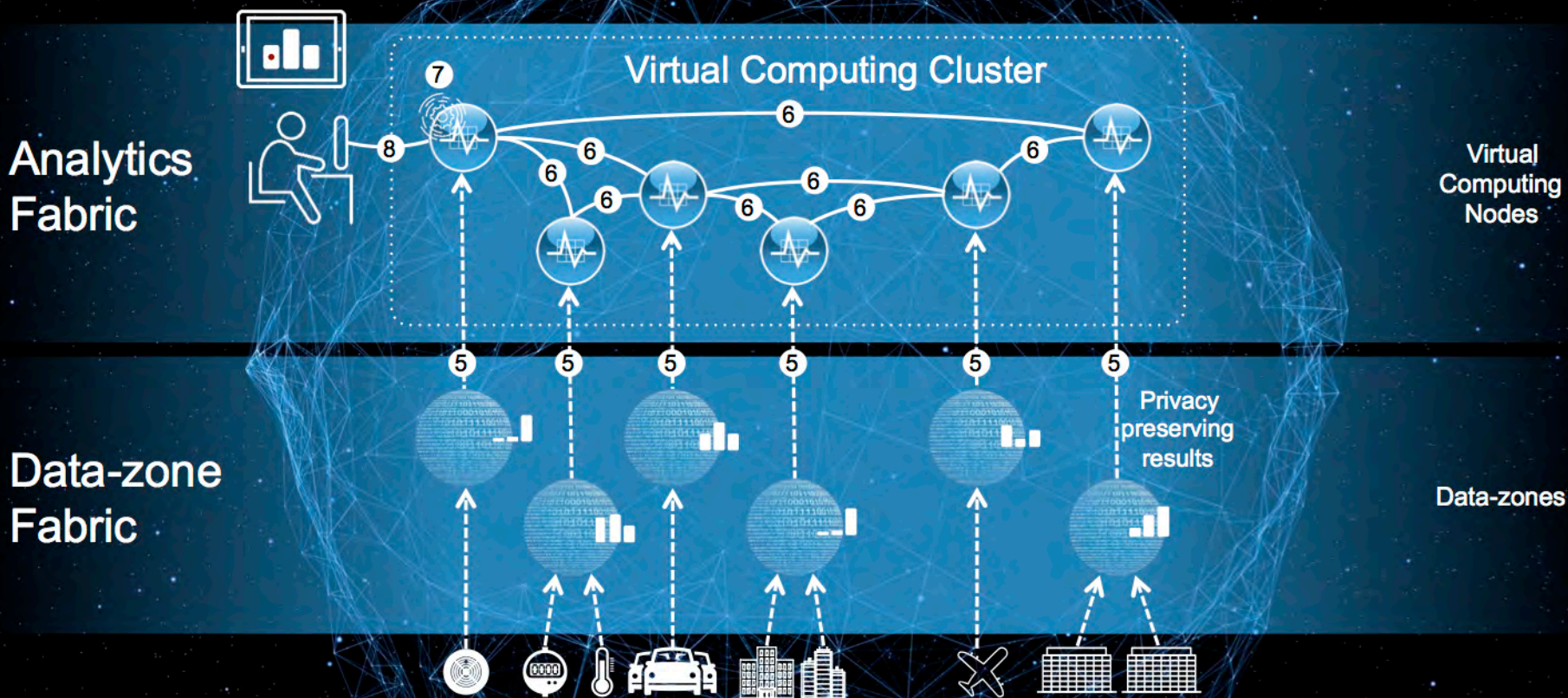
Parallel and distributed computation



Parallel and distributed computation



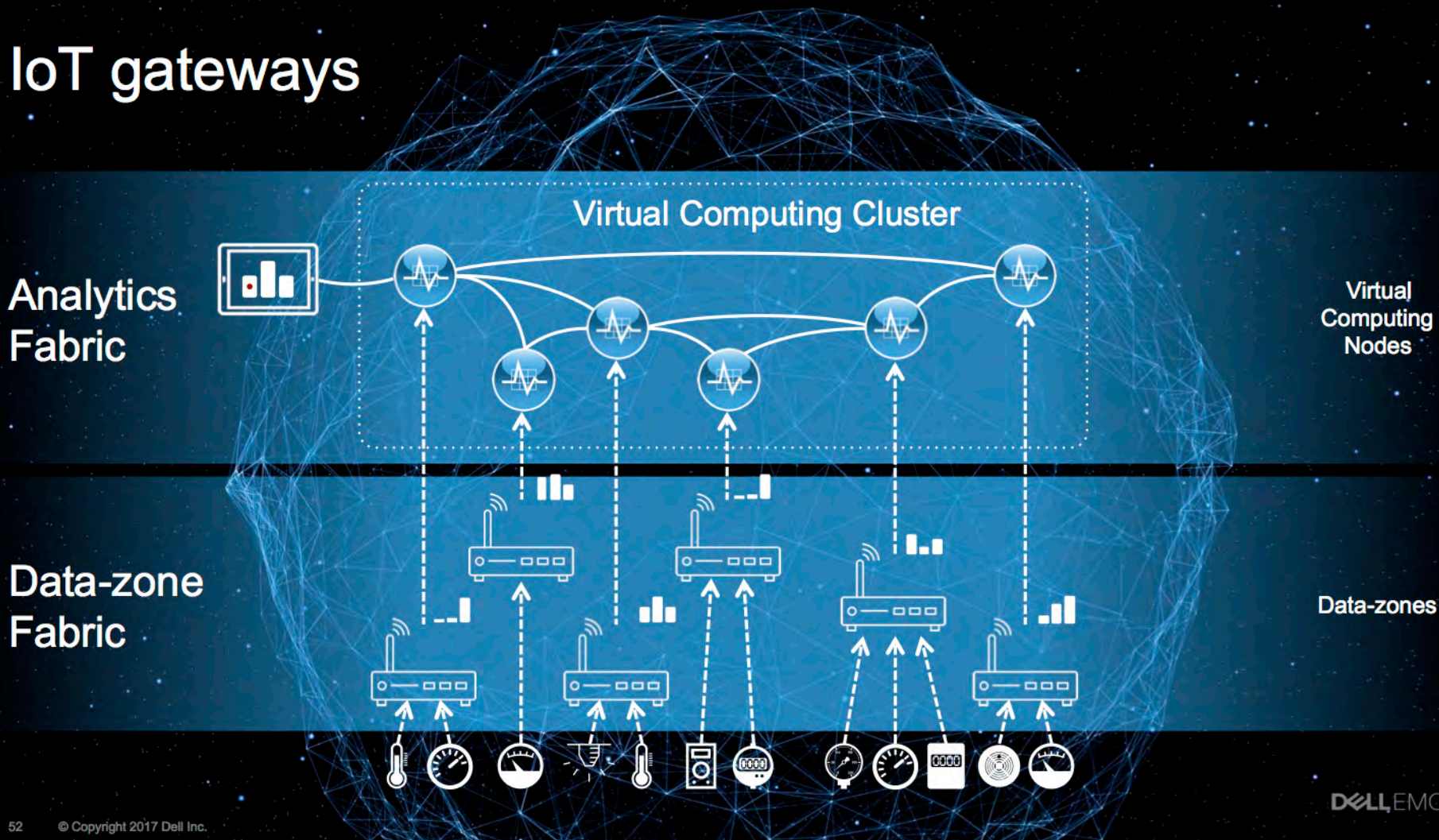
Privacy-preserving results shared



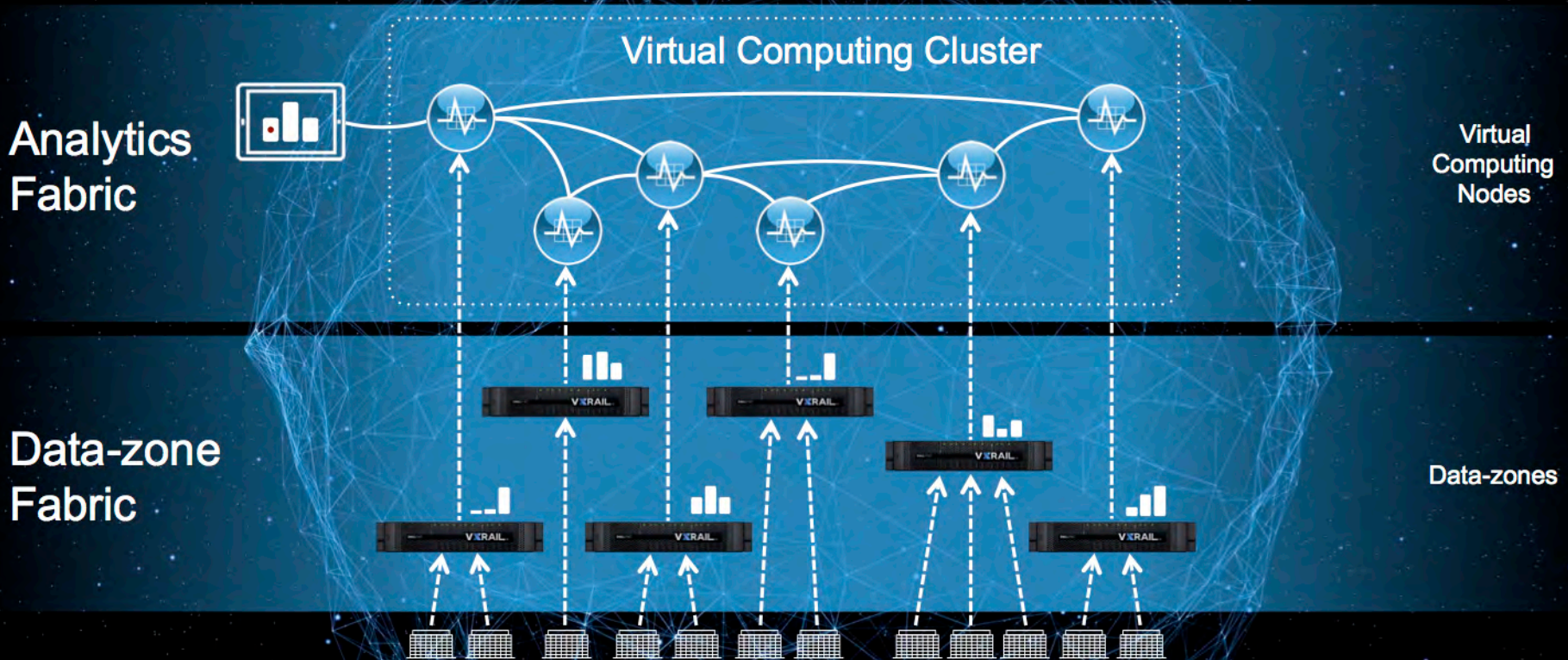
Broad Range of Data-zones

DELL EMC

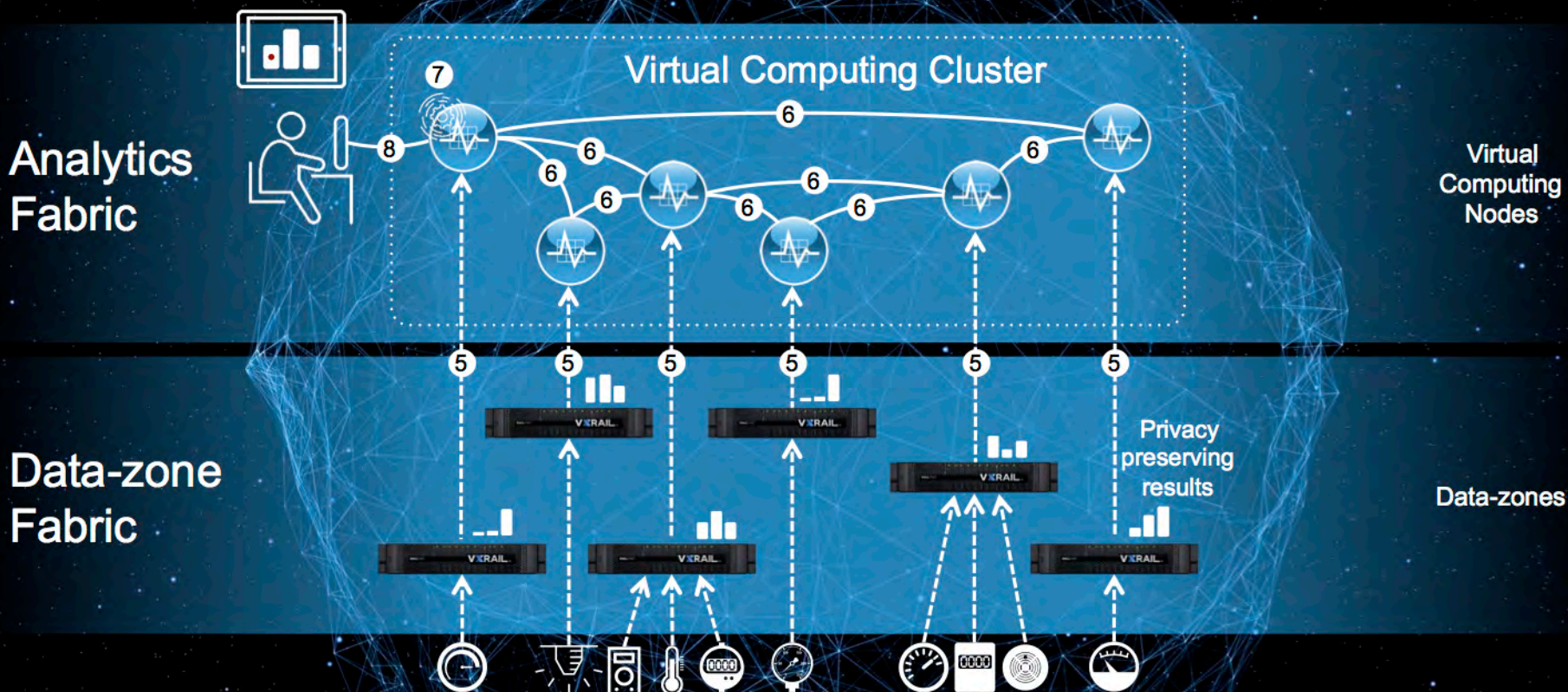
IoT gateways



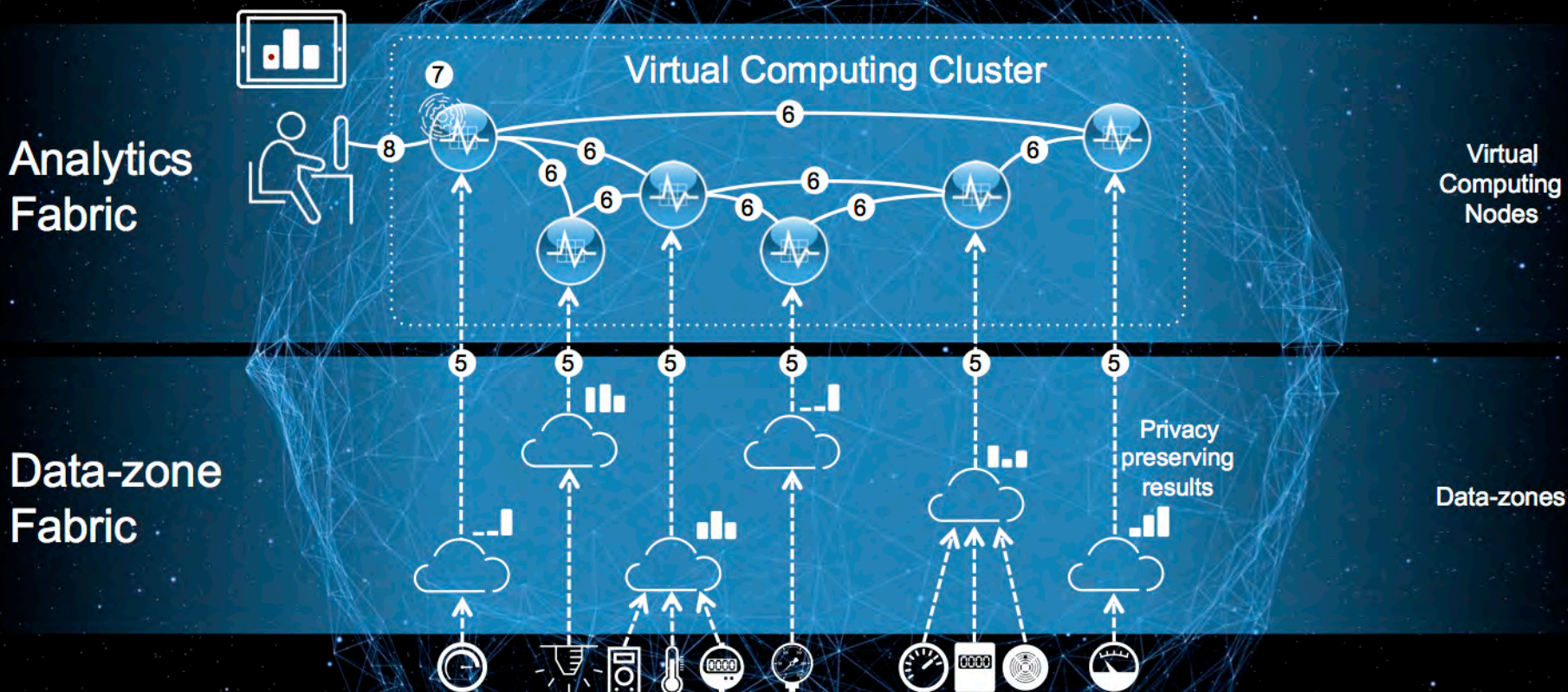
HCI receiving data from digital buildings



HCI receiving data from sensors

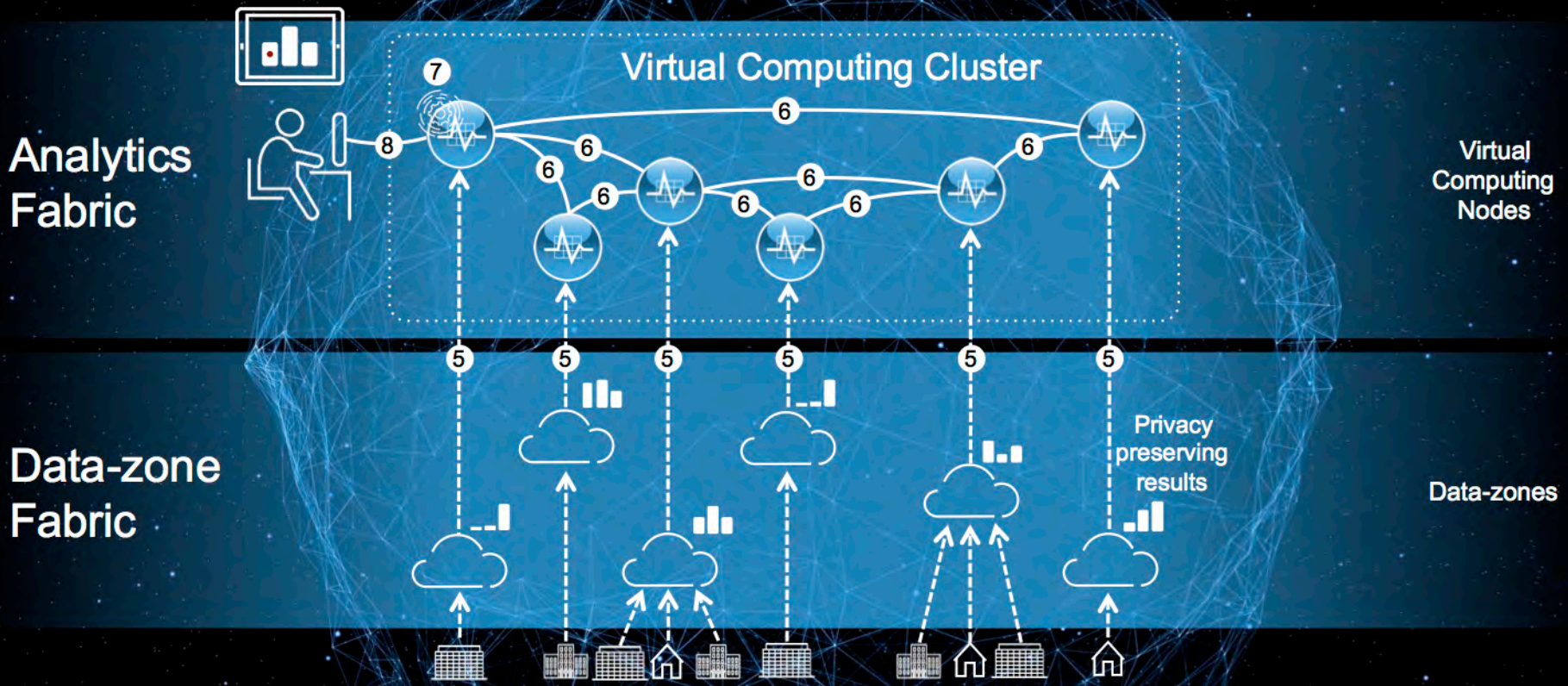


Cloud instances receiving data from sensors



DELL EMC

Cloud Instances receiving data from Digital Buildings

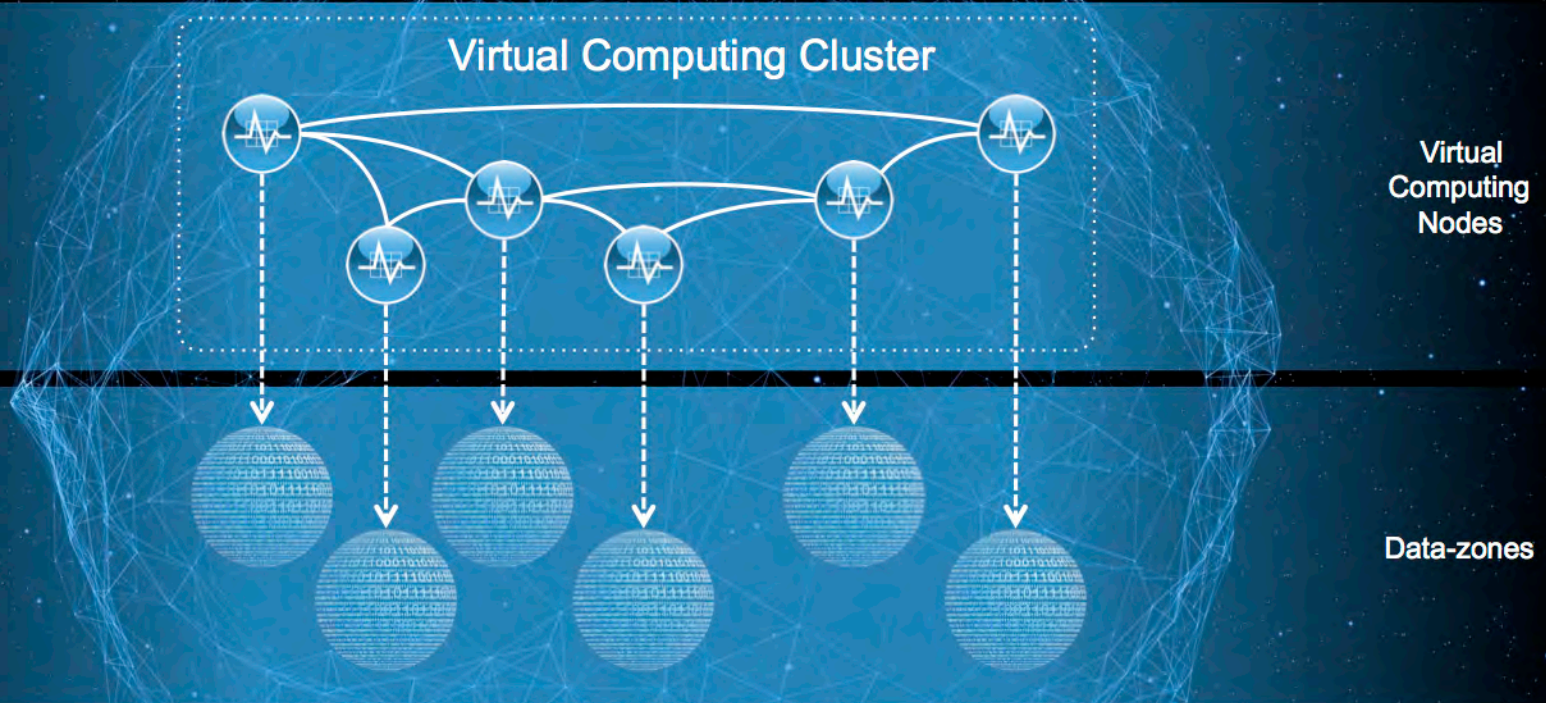


Computing nodes become virtual nodes in a virtual computing cluster at scale

Connecting clusters at the analytics layer

Analytics
Fabric

Data-zone
Fabric



DELL EMC

Rebirth of data science

DELL EMC

Benchmarking

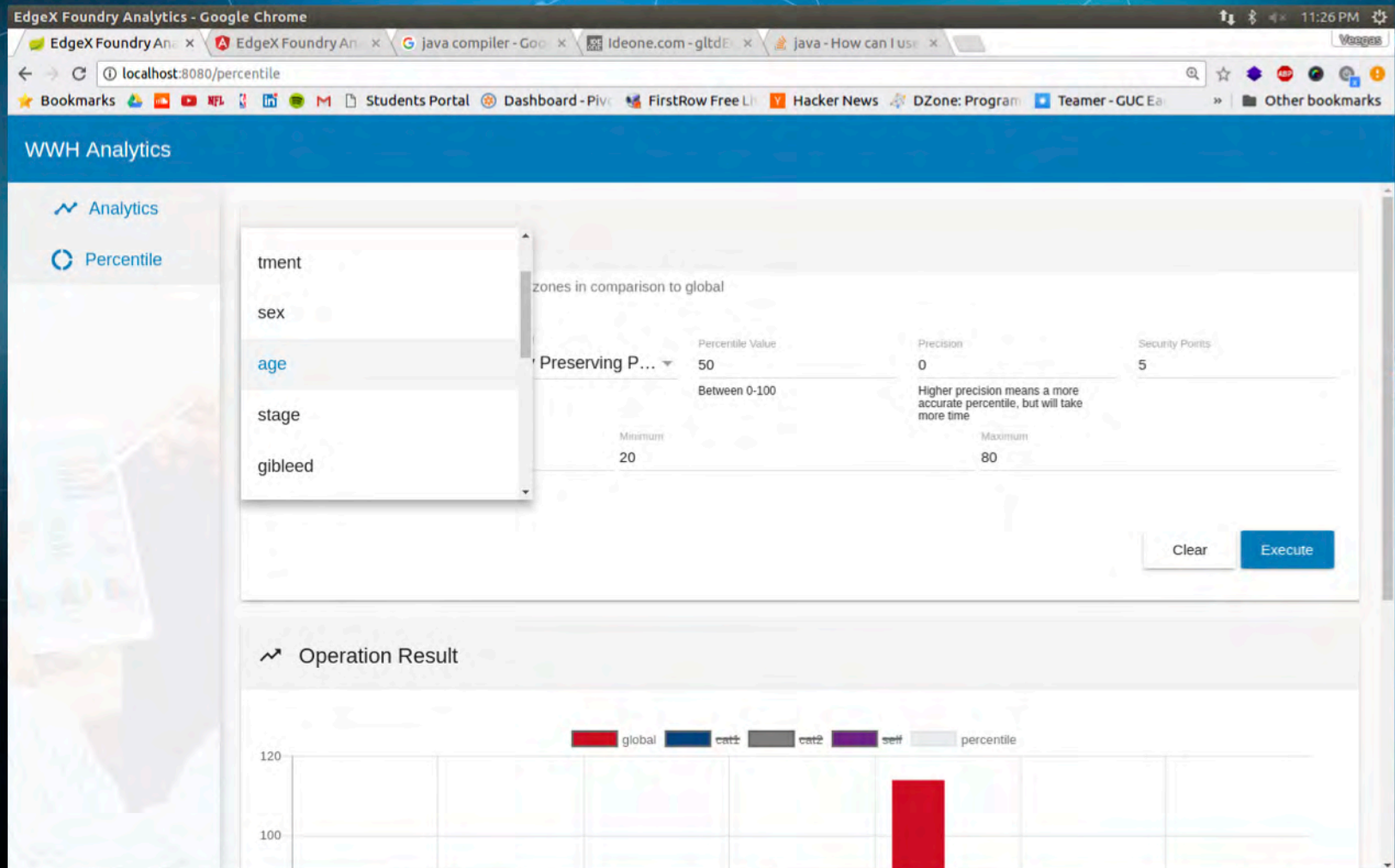
DELL EMC

Parameters for Histogram & Median Use Case

- Attribute selected: age, ranging from 20 to 80
- Number of records: 275 total
- Three data-zones:
 - purple,
 - blue, and
 - gray

DELLEMC

Global Product Data Interoperability Summit | 2017



DELLEMC

Global Product Data Interoperability Summit | 2017

EdgeX Foundry Analytics - Google Chrome

localhost:8080/percentile

WWH Analytics

- Analytics
- Percentile

Percentile Job Input

The operation result will be all the local database in comparison to global.

Columns: age

Value: [dropdown menu showing Percentile and Privacy Preserving Percentile]

Precision: 0

Security Points: 5

Higher precision means a more accurate percentile, but will take more time.

Slices: 5


Minimum: 20

Maximum: 80

Clear Execute

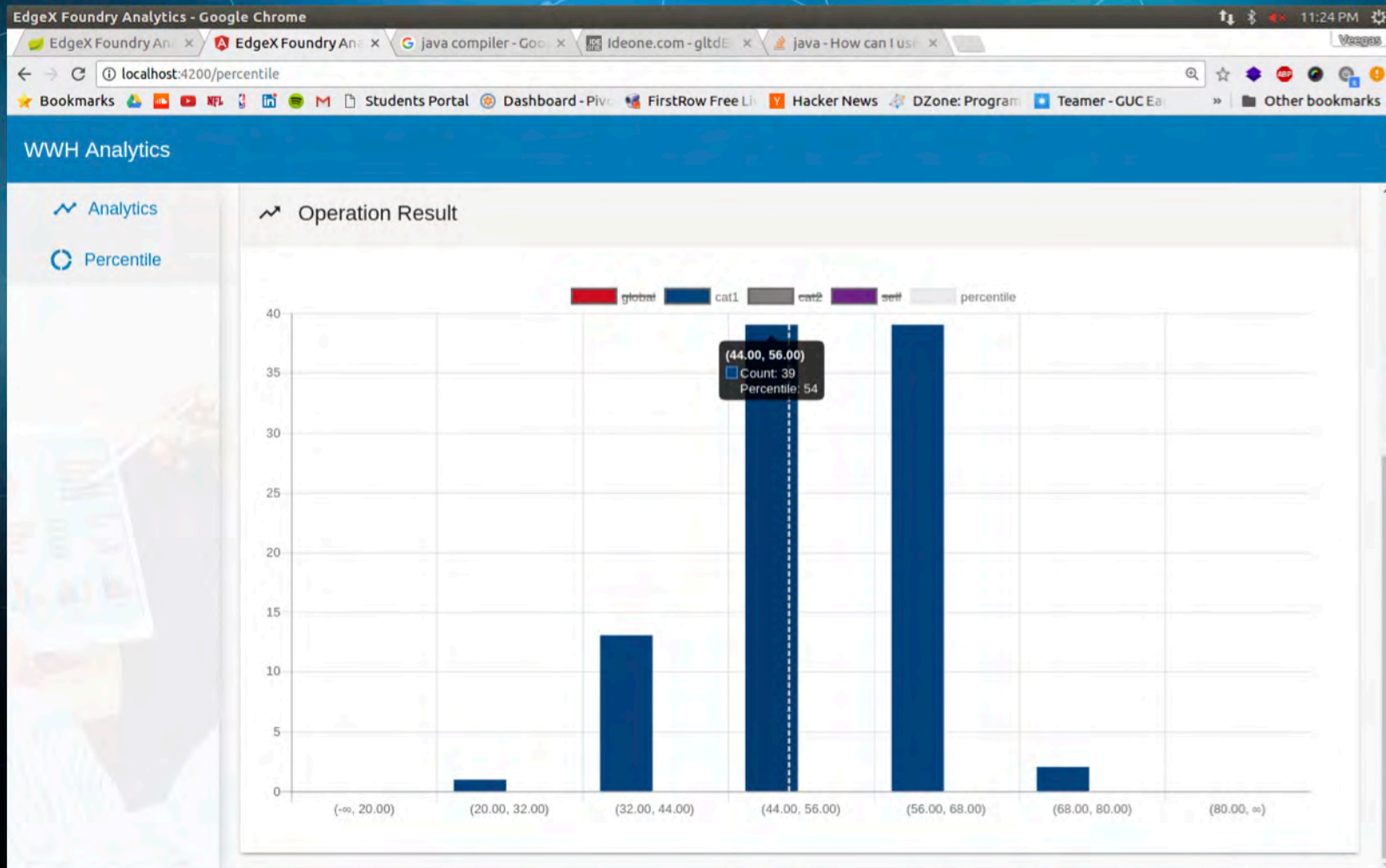
Operation Result

Legend: global (red), cat1 (blue), cat2 (grey), cat3 (purple), percentile (light grey)



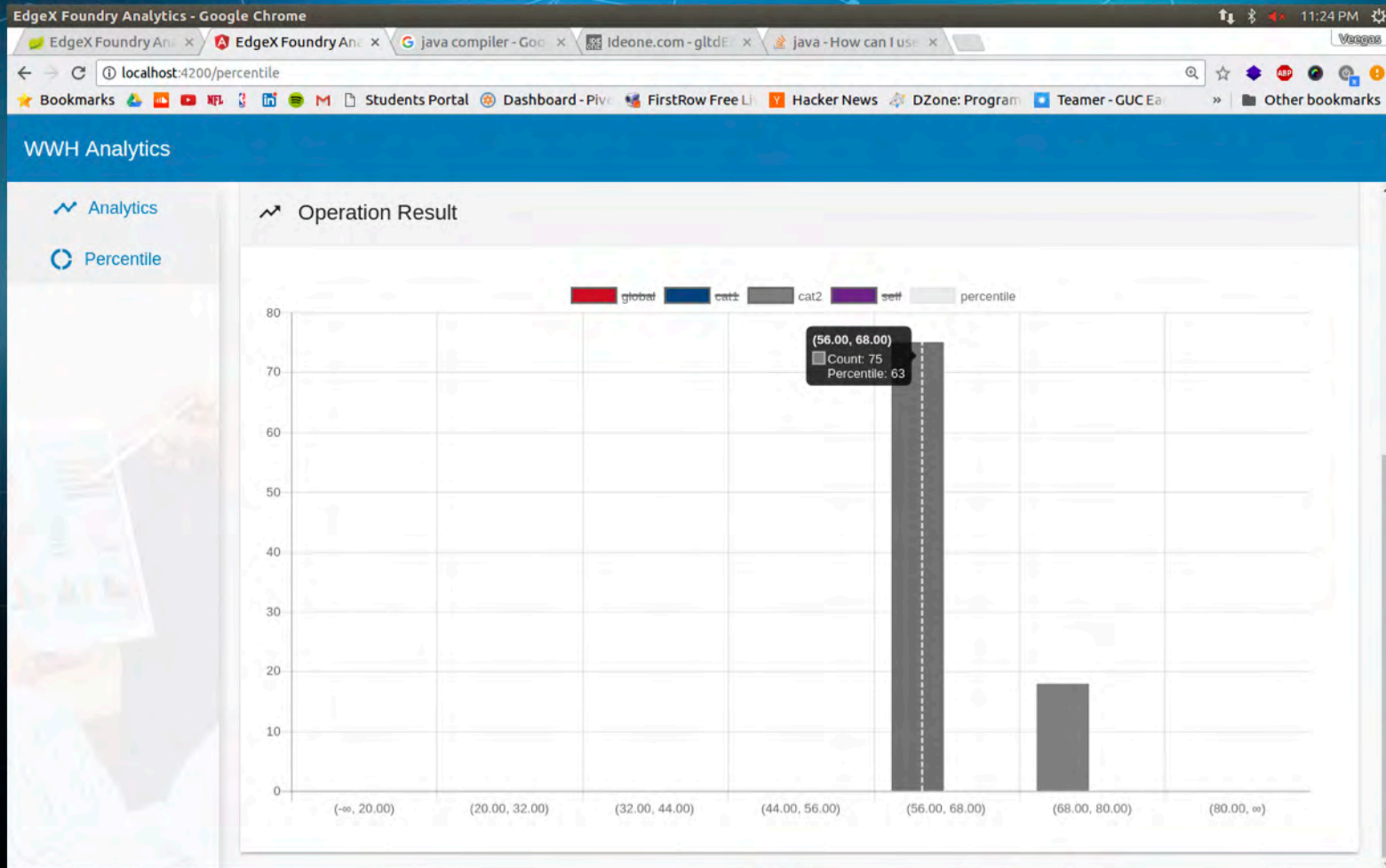
DELLEMC

Global Product Data Interoperability Summit | 2017



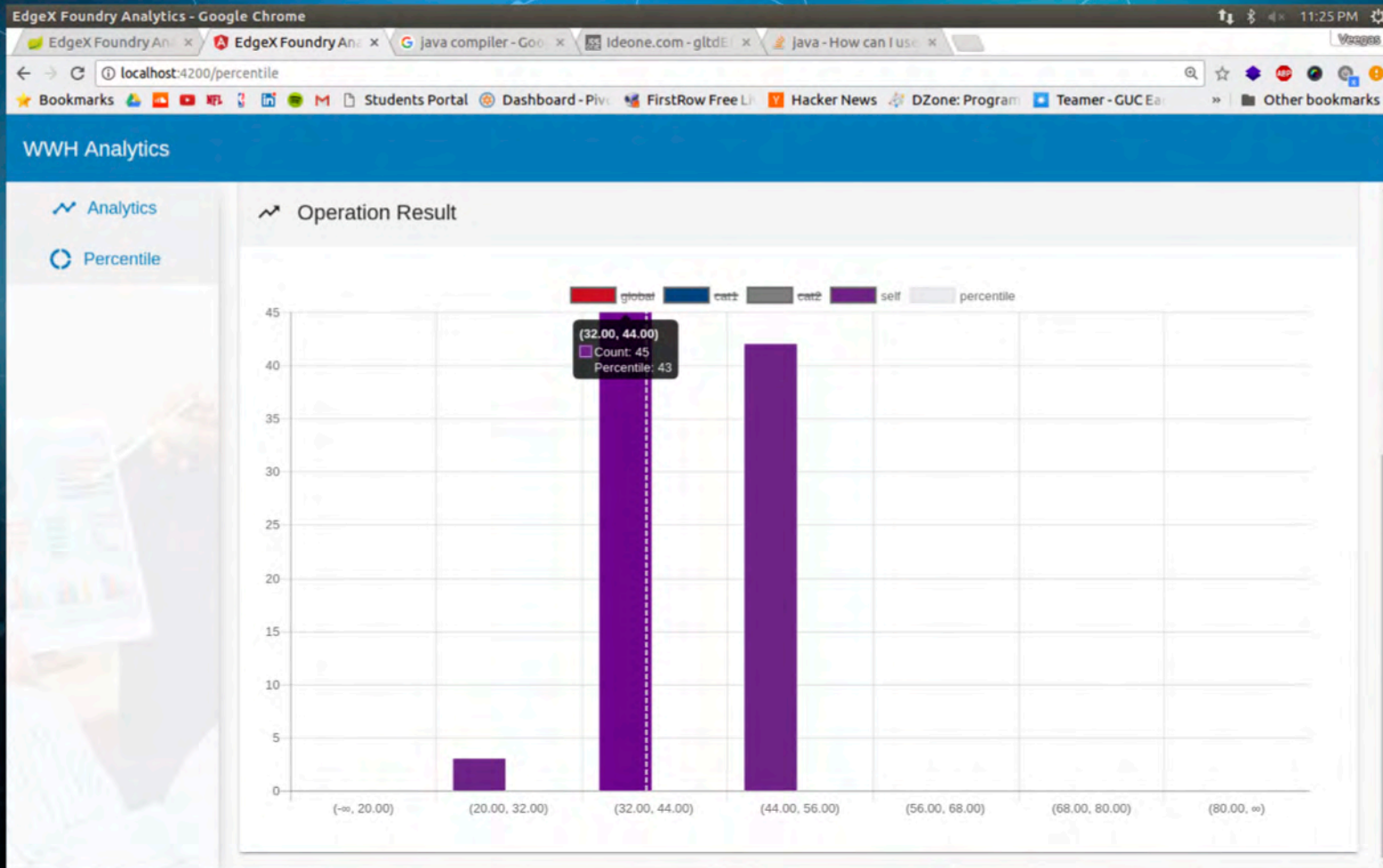
DELLEMC

Global Product Data Interoperability Summit | 2017



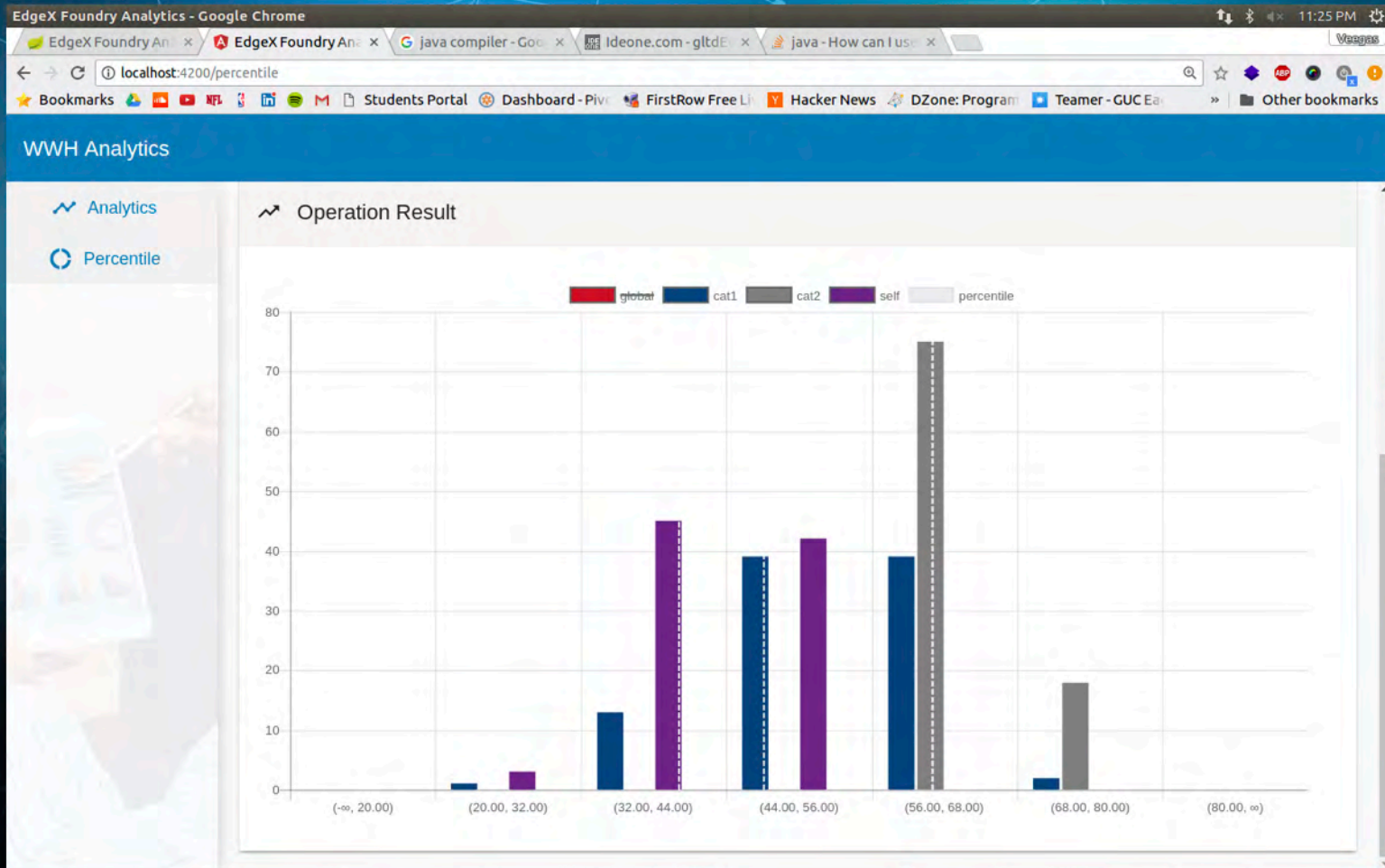
DELLEMC

Global Product Data Interoperability Summit | 2017



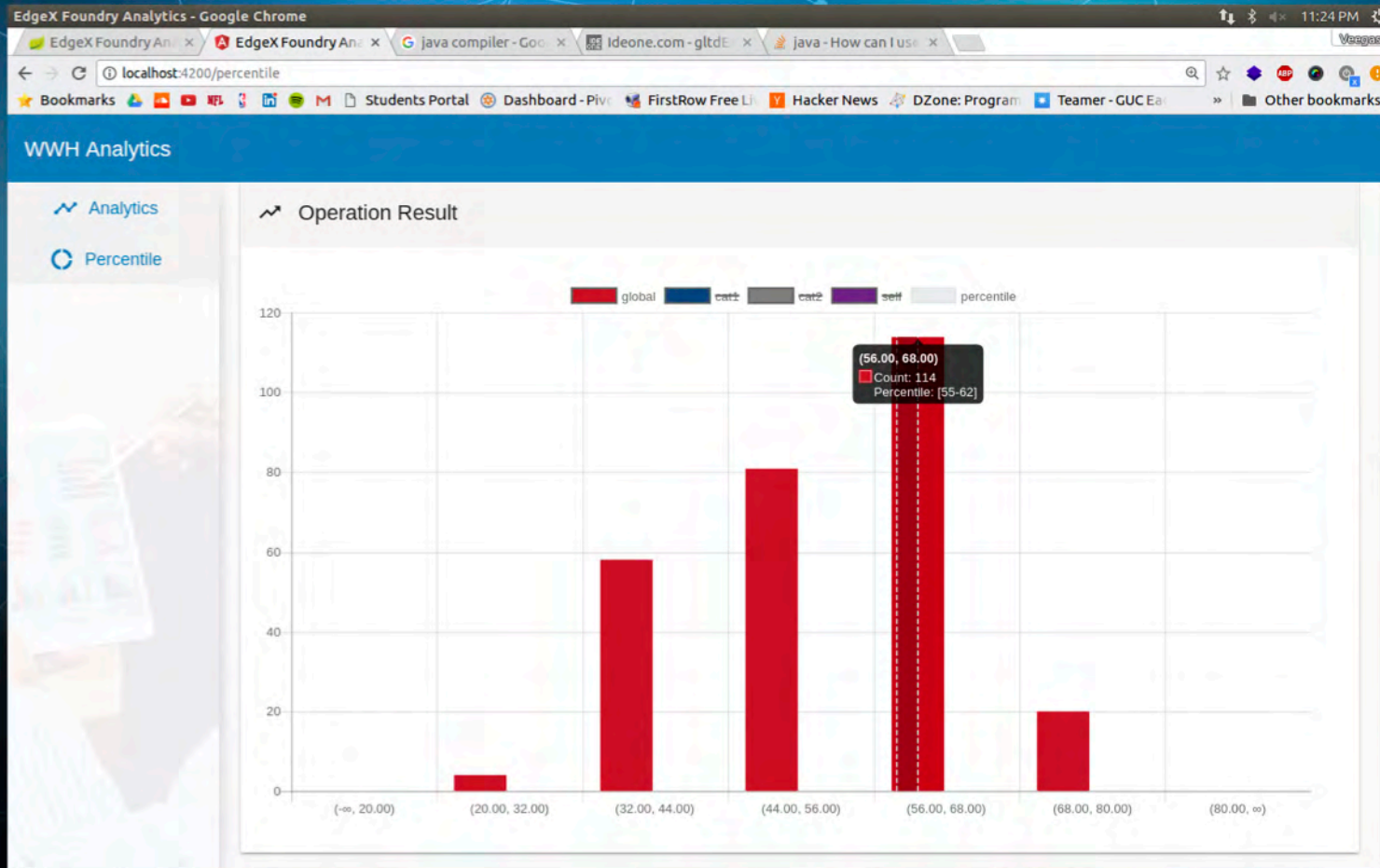
DELLEMC

Global Product Data Interoperability Summit | 2017



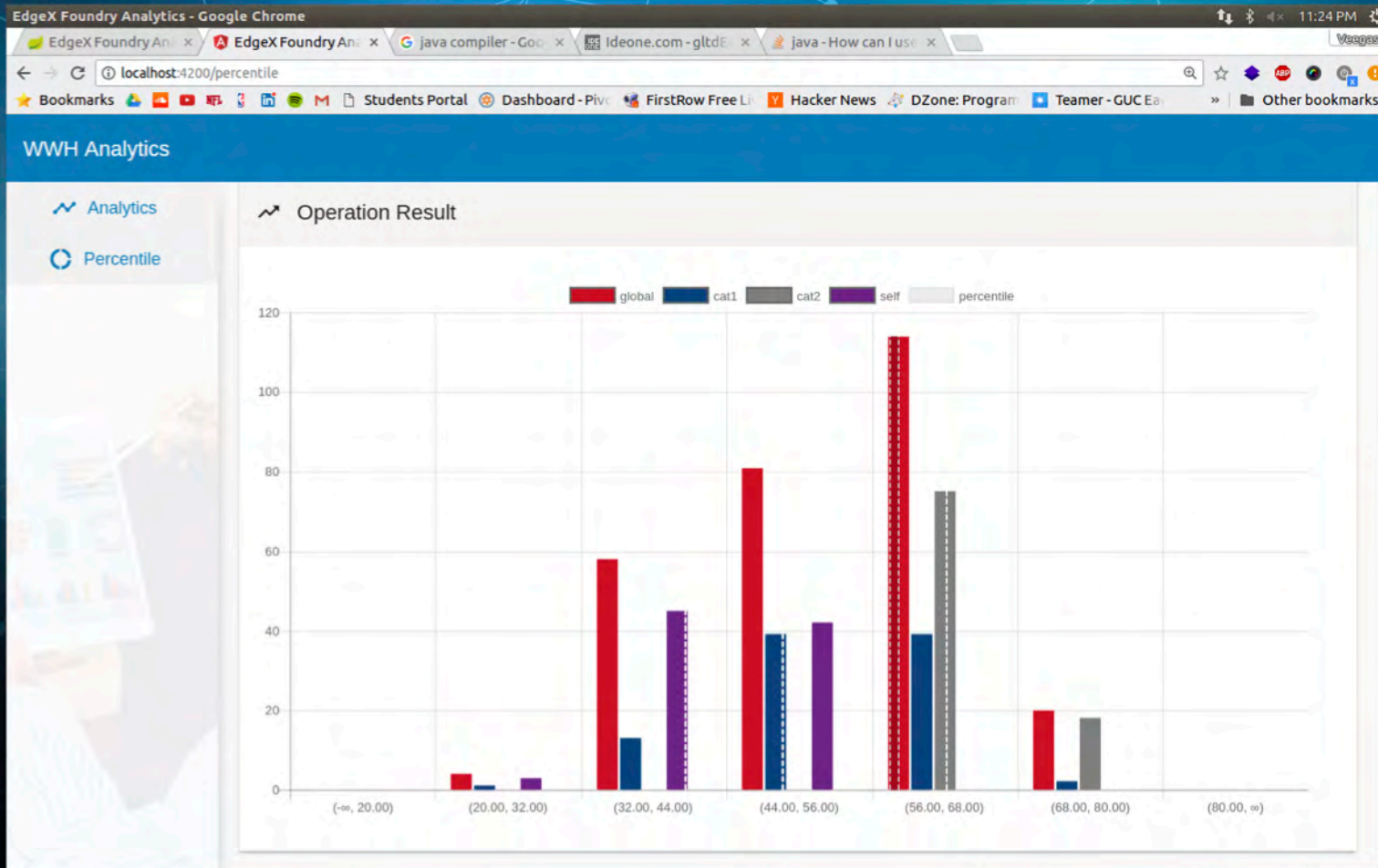
DELLEMC

Global Product Data Interoperability Summit | 2017



DELLEMC

Global Product Data Interoperability Summit | 2017



DELLEMC

Summary of Results

- Main Cluster: Percentile value: 43 (2nd slice)
- Remote cluster 1: Percentile value: 54 (3rd slice)
- Remote cluster 2: Percentile value: 63 (4th slice)
- Global Percentile Range: 55 – 56 (4th slice)

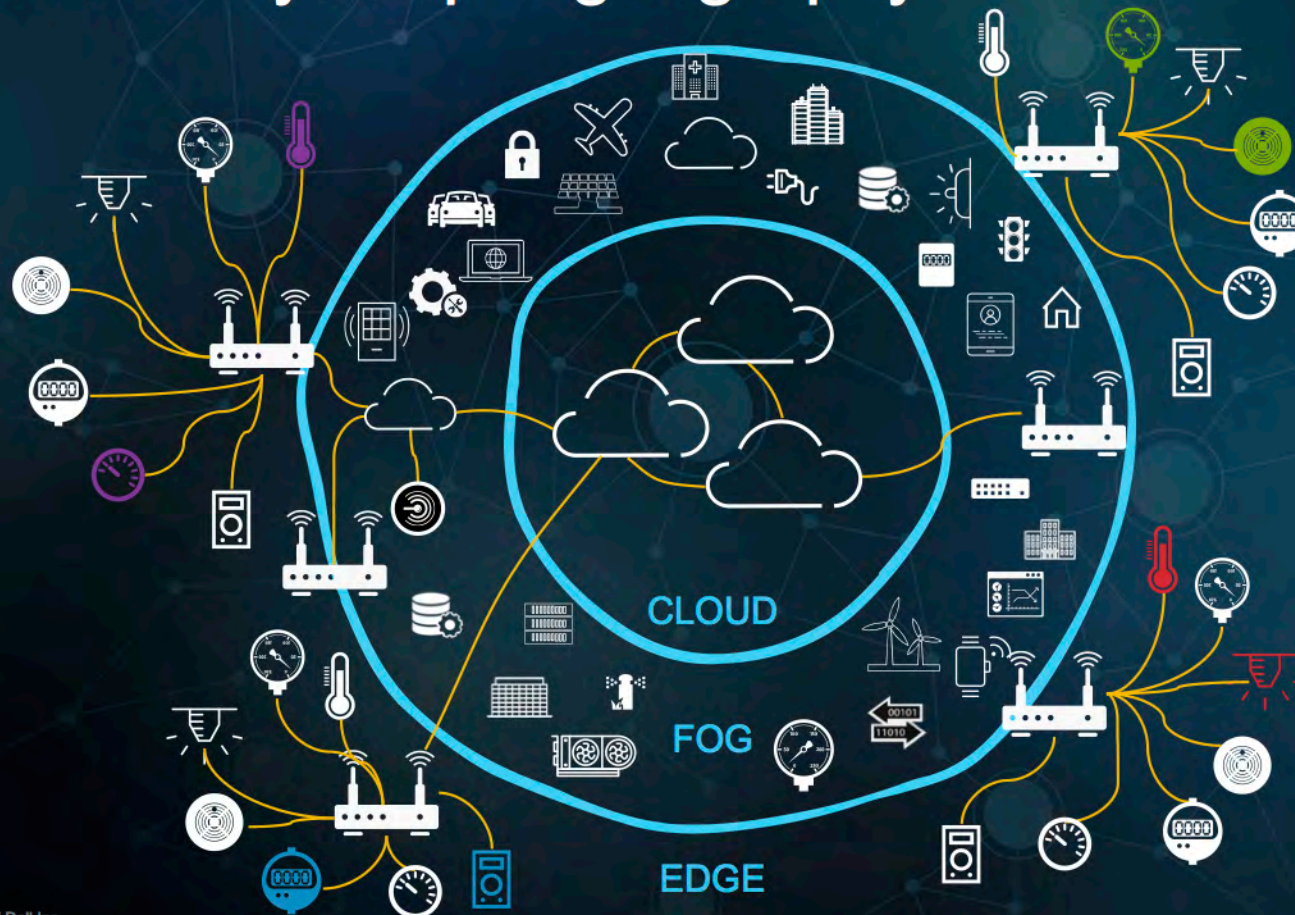
DELLEMC

WWH Fabrics

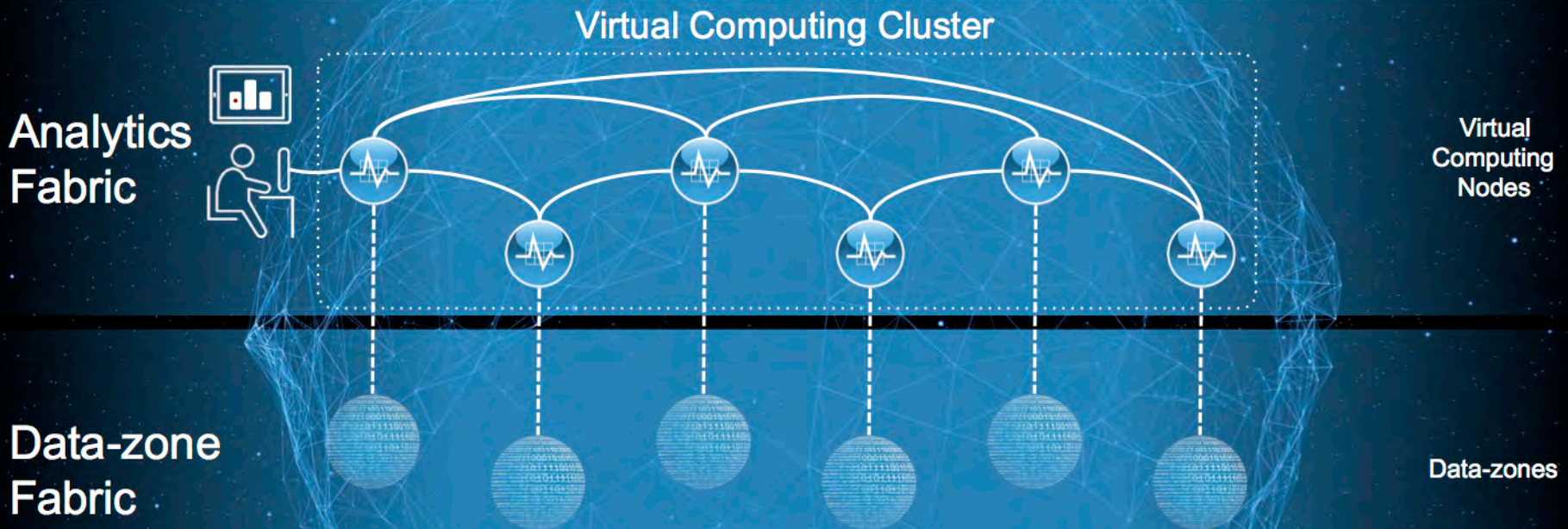
Because “the future is federated”

DELL EMC

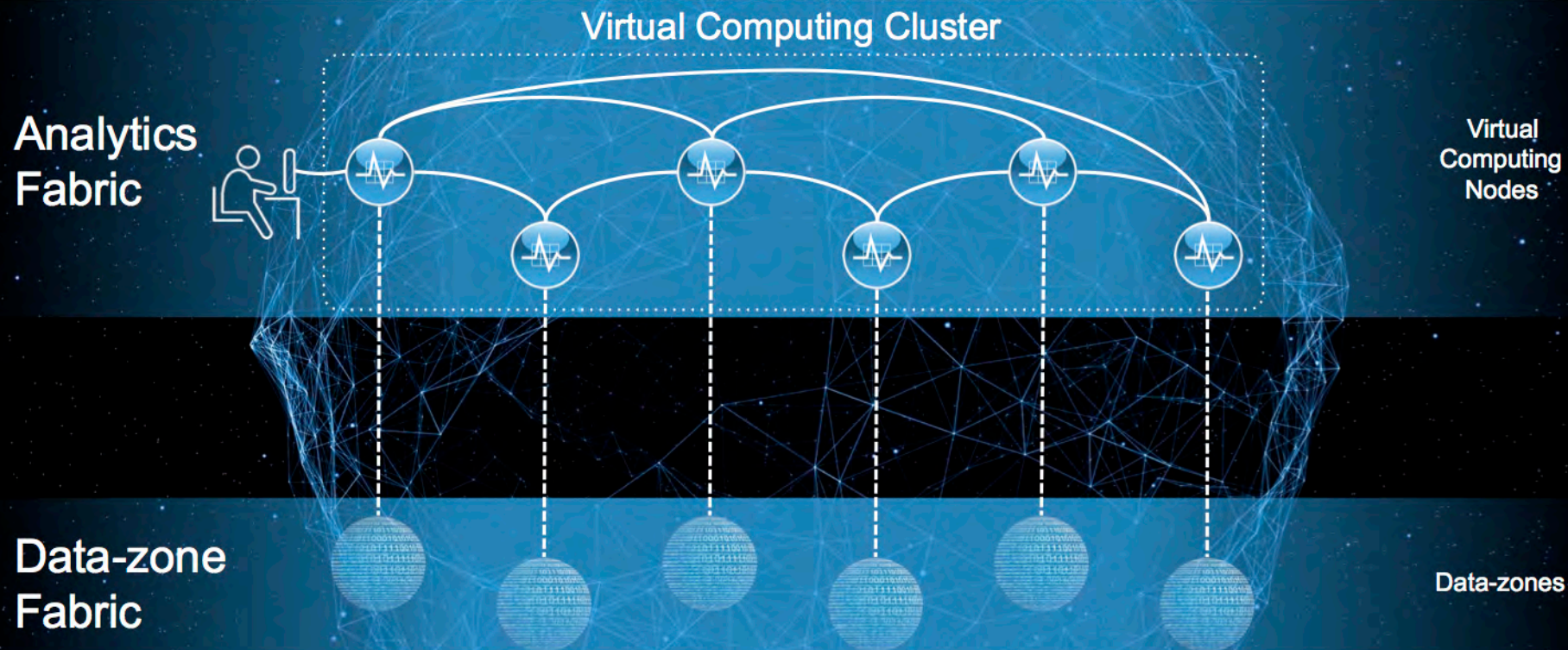
Federate analytics per geography



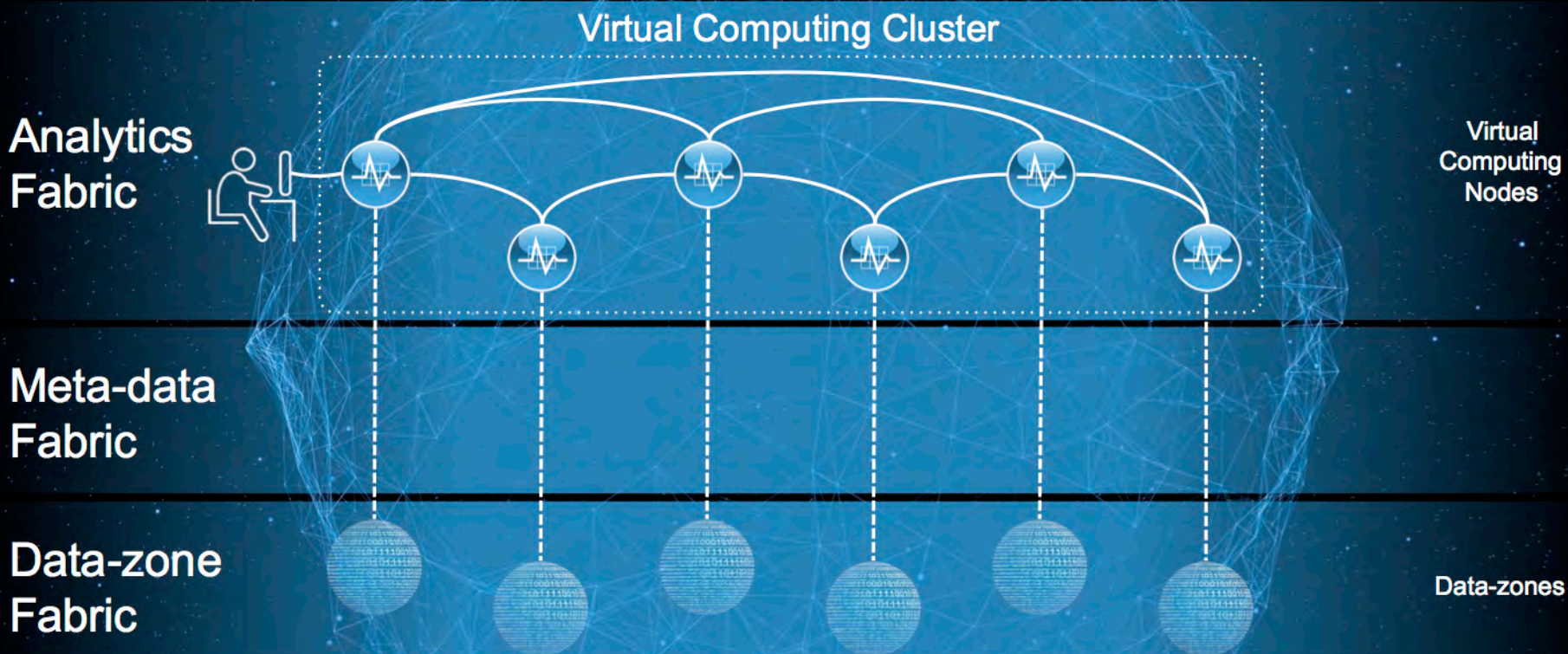
How all these connections get established?



Yet another fabric

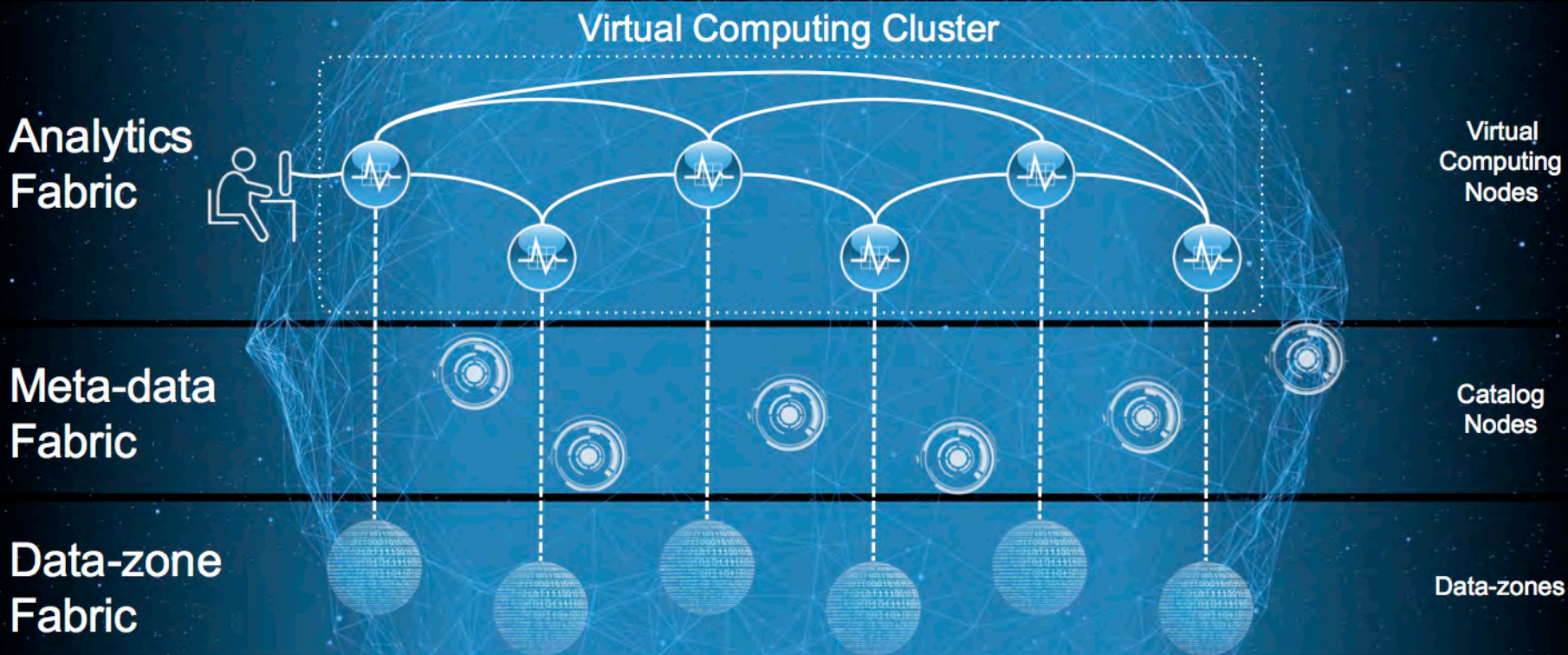


The meta-data fabric



DELL EMC

Catalog nodes



DELL EMC

Persona

A specific, fictional representation of a person who is very likely to represent your ideal customer. A persona is a prototype of your user.

<https://content.pivotal.io/blog/how-to-define-your-persona-part-2>

DELLEMC

Business value creator

Analytics
Fabric

Virtual Computing Cluster

Virtual
Computing
Nodes



Data Scientist

Business value harvester

Analytics
Fabric

Virtual Computing Cluster

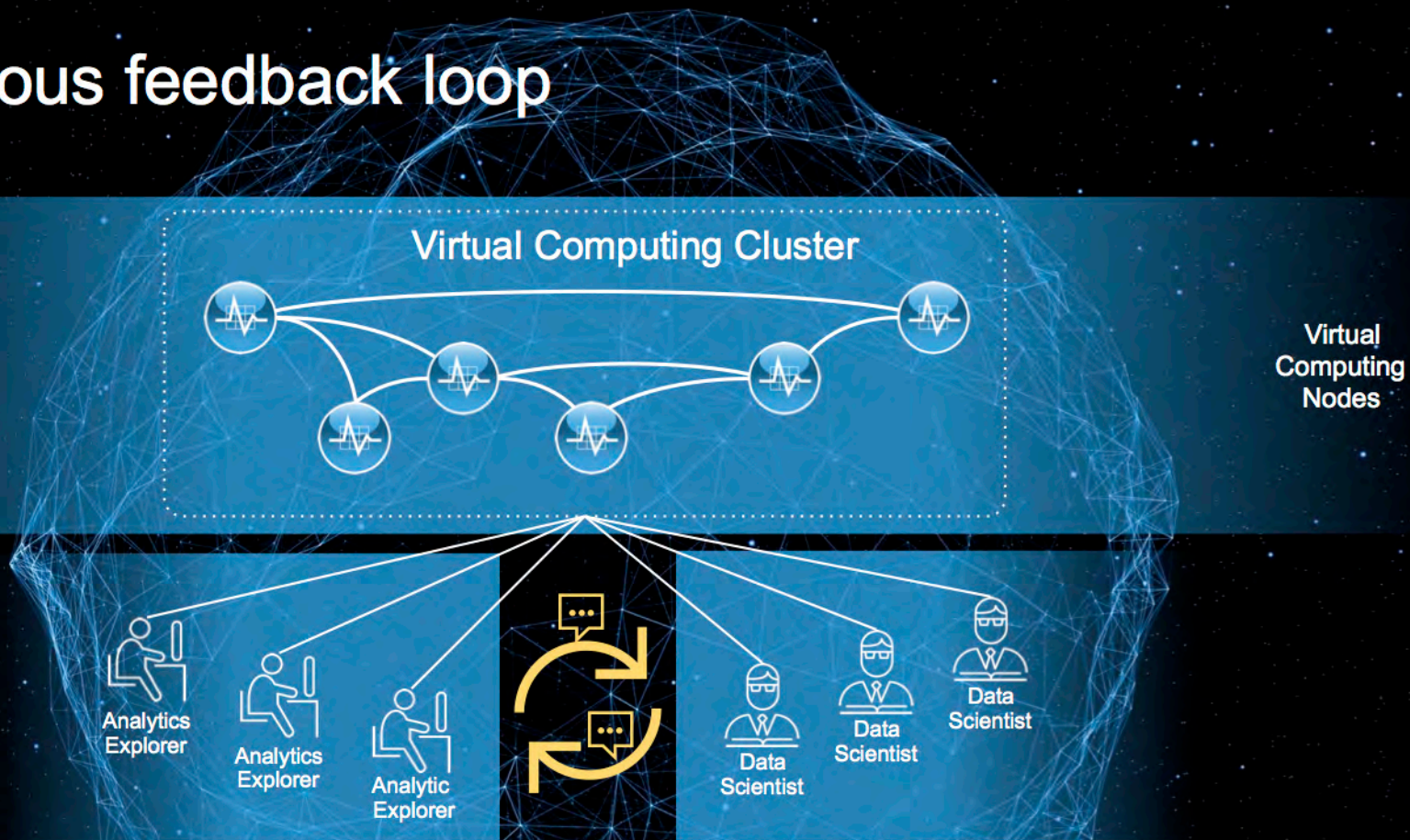
Virtual
Computing
Nodes

Analytics Explorer

DELL EMC

Continuous feedback loop

Analytics
Fabric



DELL EMC

Together, business transformation happens

Analytics
Fabric

Virtual Computing Cluster

Virtual
Computing
Nodes

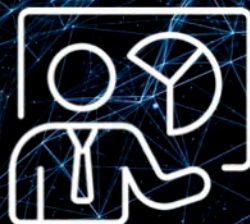


Analytics requires data and processing capacity

Data-zone
Fabric

Data-zones

Data architects manage data



Data
Architect

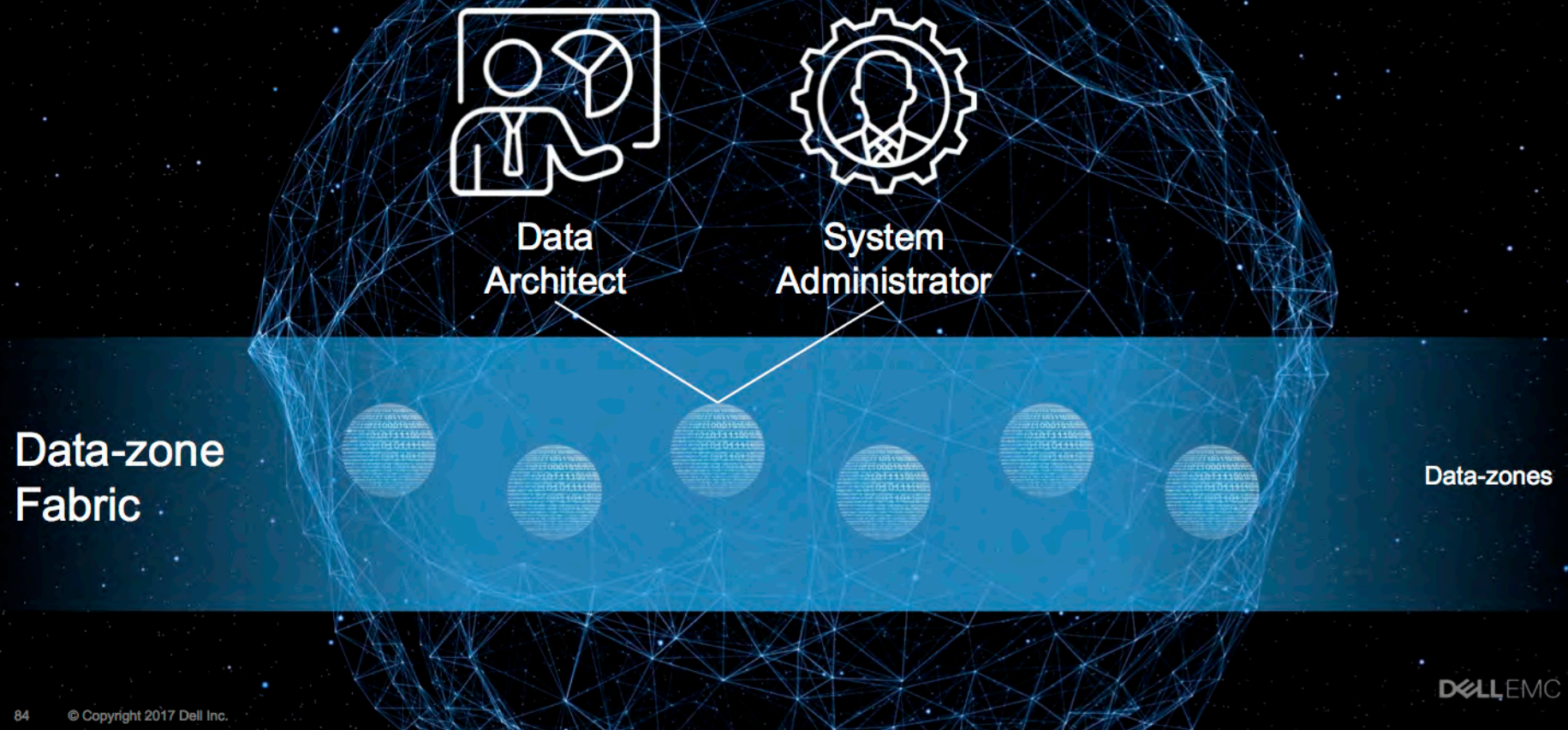
Data-zone
Fabric

Data-zones

Administrators provision processing nodes



Together, data and processing capacity are created

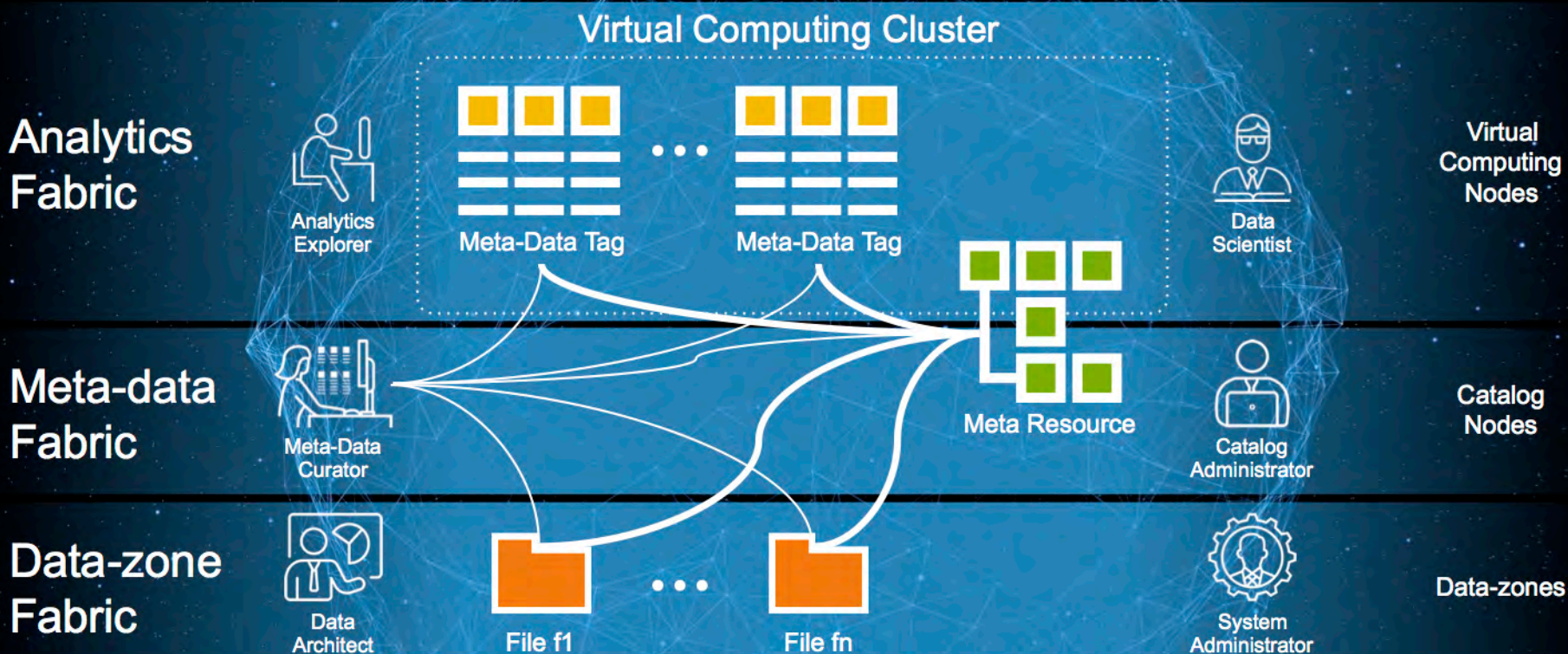


Different data-zones, different set of owners

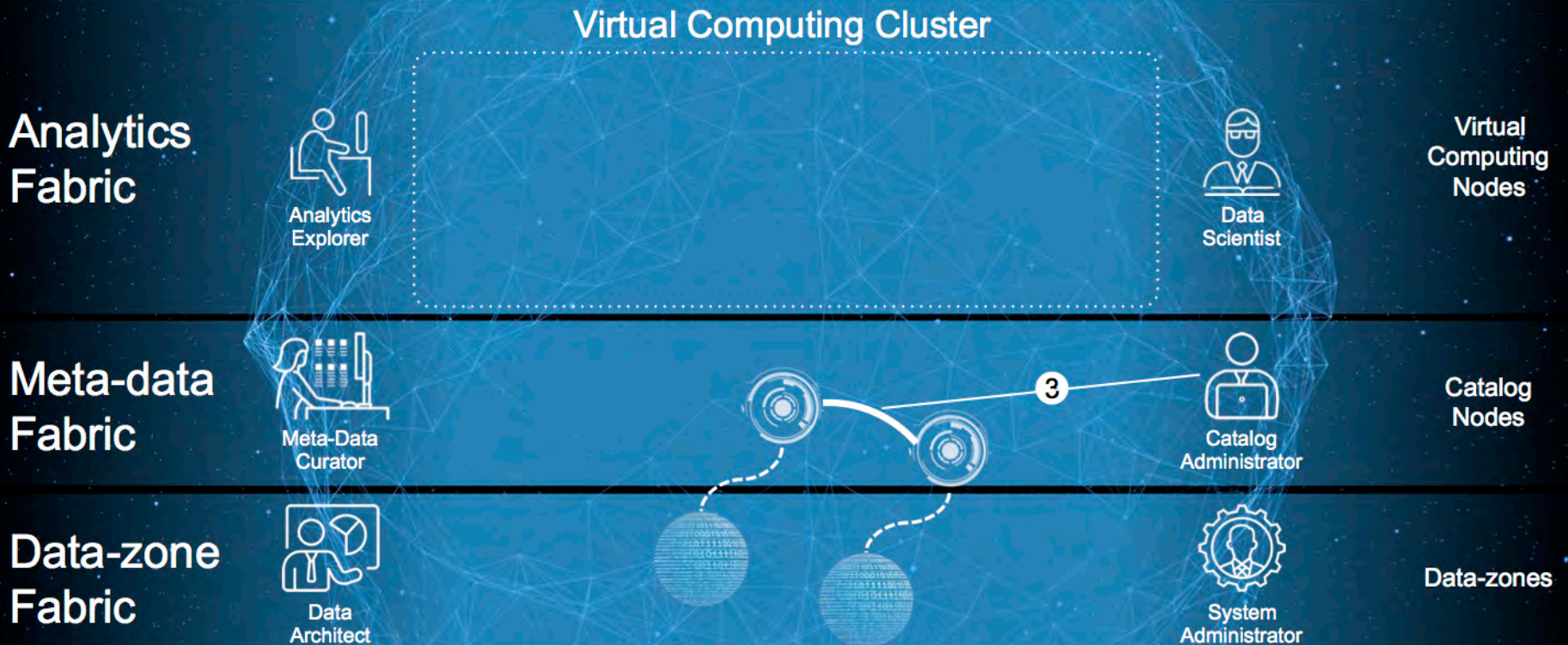
Data-zone
Fabric

Data-zones

Meta-data curator catalogs it all

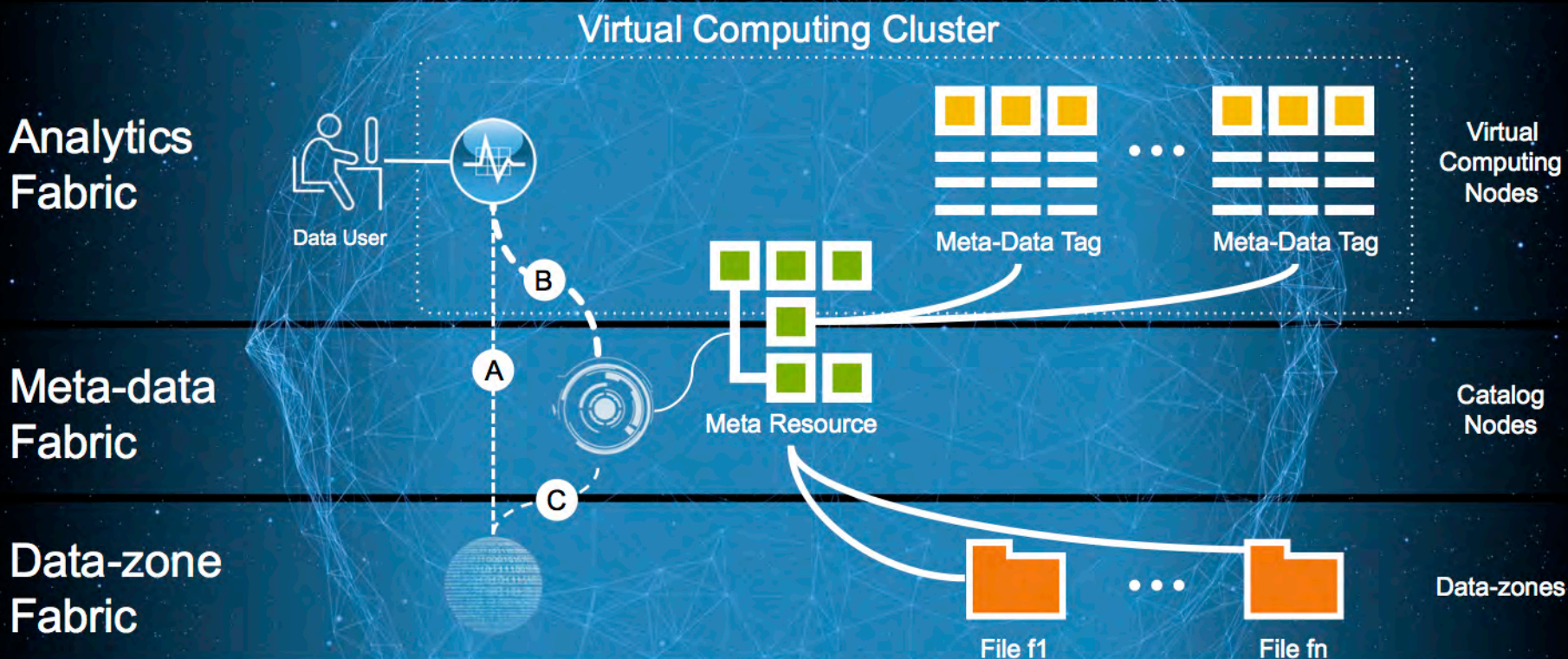



Creating the meta-data fabric



DELL EMC

Computing nodes become virtual nodes in a virtual computing cluster at scale





Some people make it happen
Some people watch it happen
Some people do not know what happened

DELL EMC

Think bigger data...

Think distributed bigger data...

Think world wide scale,
distributed bigger data...

DELL EMC

Dawn of a new era

DELLEMC

91 © Copyright 2017 Dell Inc.

Technology leadership is a team sport


We are better together



Better together because...

only together we can do
things no one else can

DELL EMC

A full-page background image showing two mountaineers standing on a snowy mountain peak at sunrise. The sun is low on the horizon, creating a warm, golden glow. The mountaineer on the left is holding an ice axe, and the one on the right is waving. They are both wearing heavy winter gear and helmets. The background shows a vast, snow-covered mountain range under a clear sky.

The time is now
To change the world
...one challenge at a time

DELL EMC

Let your transformation begin!

Thank you!

DELLEMC

DELL Technologies



DELL EMC

Pivotal.

RSA

SecureWorks®

virtustream.

vmware®



