State of C4I : Critical Aspects

Computer/Cyber Technologies and Intelligence

The Data Driven Enterprise

Real-time insights at scale through closed-loop processes
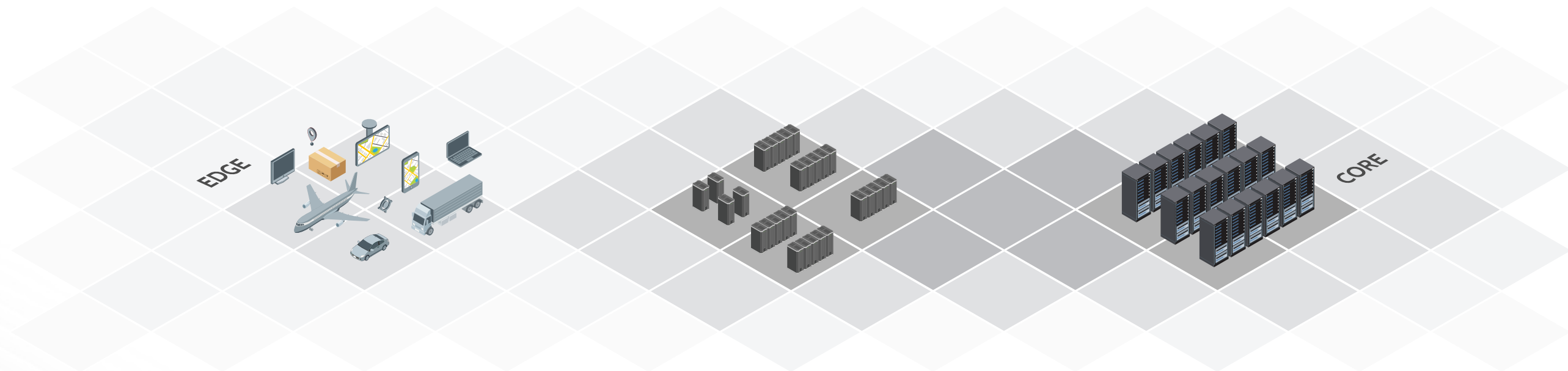
Joe Witt

HORTONWORKS®
POWERING THE FUTURE OF DATA™

The role of 'dataflow' in a Data Driven Enterprise
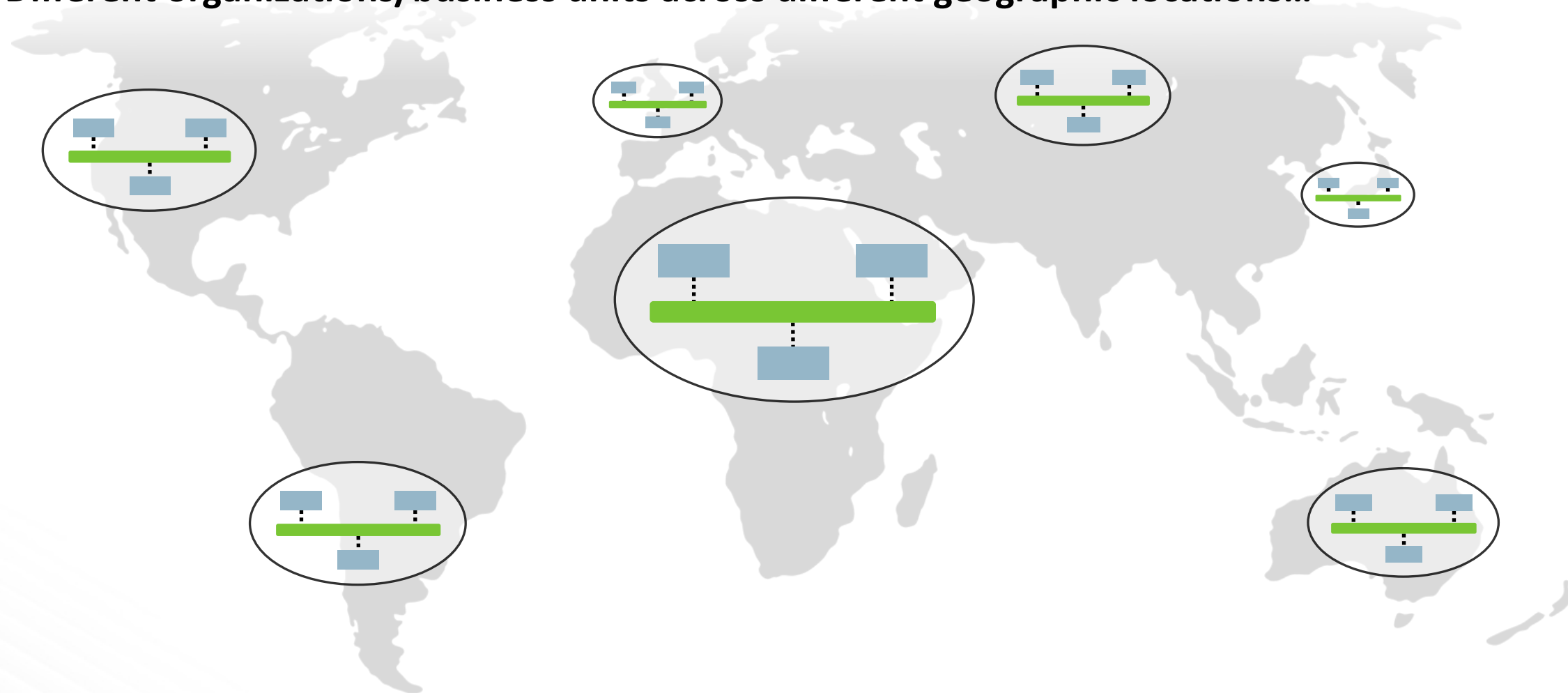
Collect

Process

Store/Query

EDGE

CORE

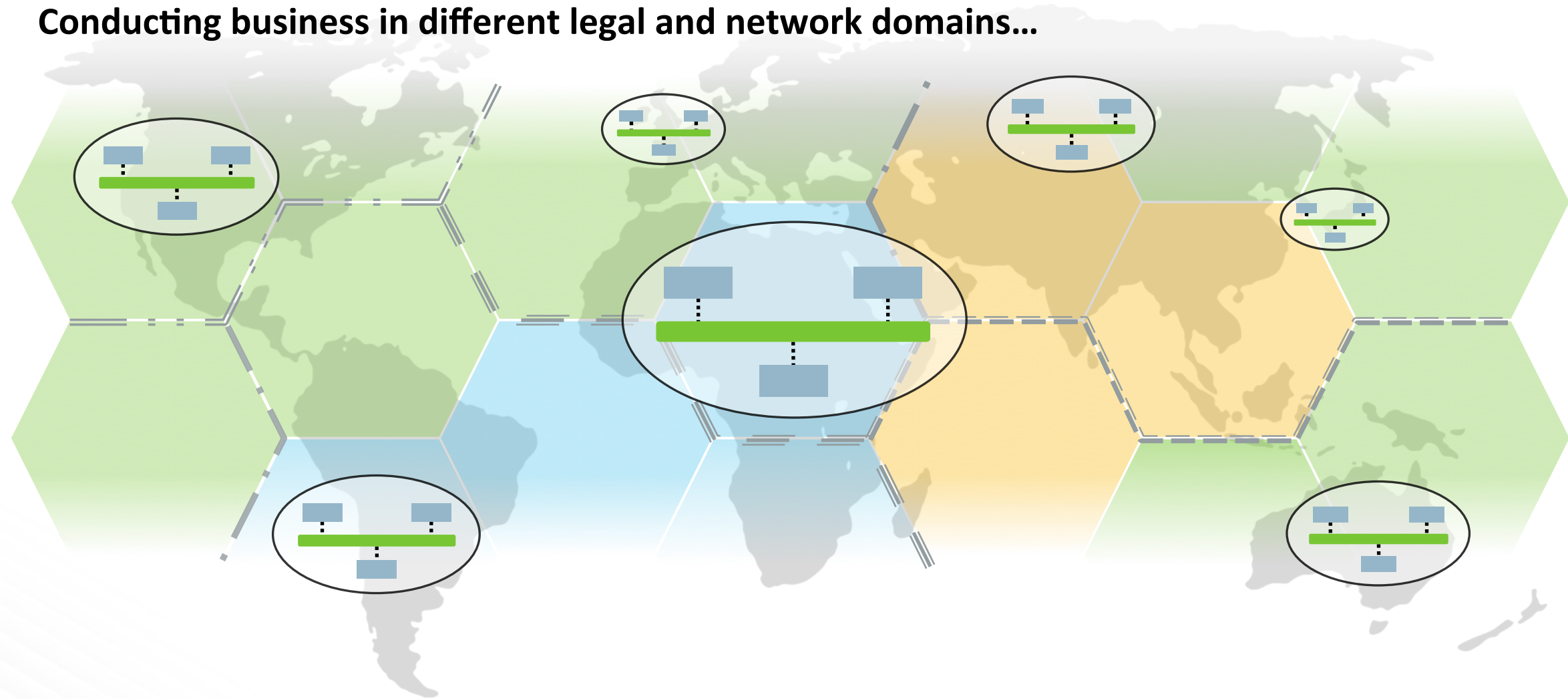Latency : From sensor to repository (far less than half the story)

HORTONWORKS®

Enterprise Dataflow

Different organizations/business units across different geographic locations...
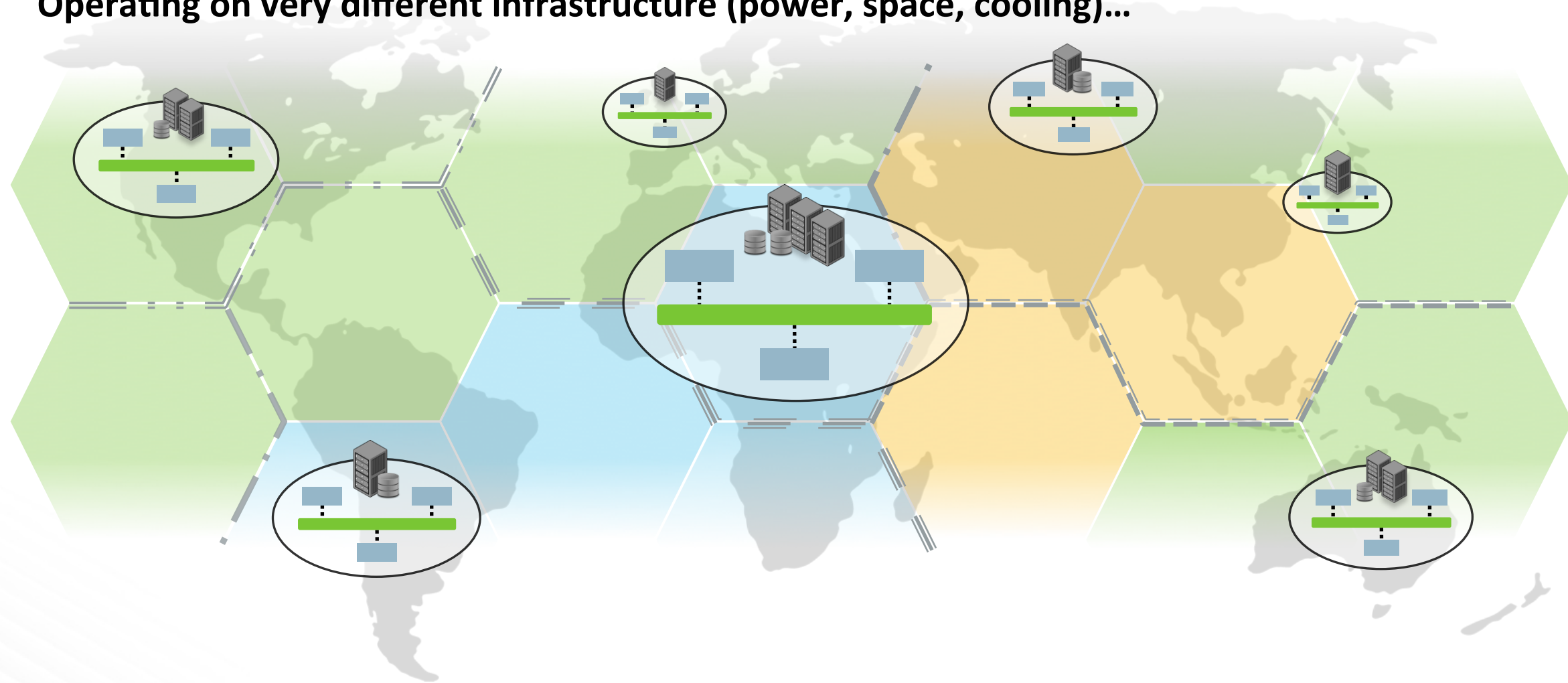
Enterprise Dataflow

Conducting business in different legal and network domains…
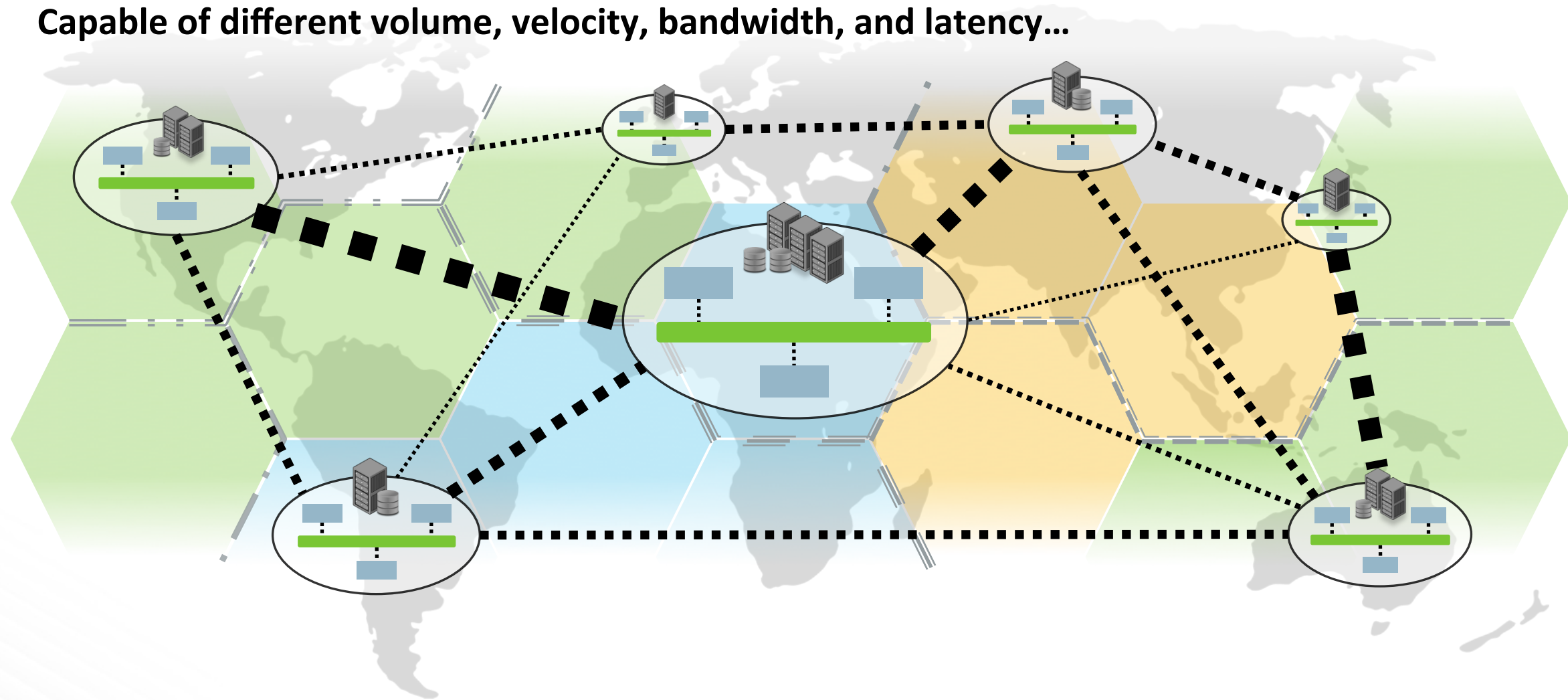
Enterprise Dataflow

Operating on very different infrastructure (power, space, cooling)...

Enterprise Dataflow
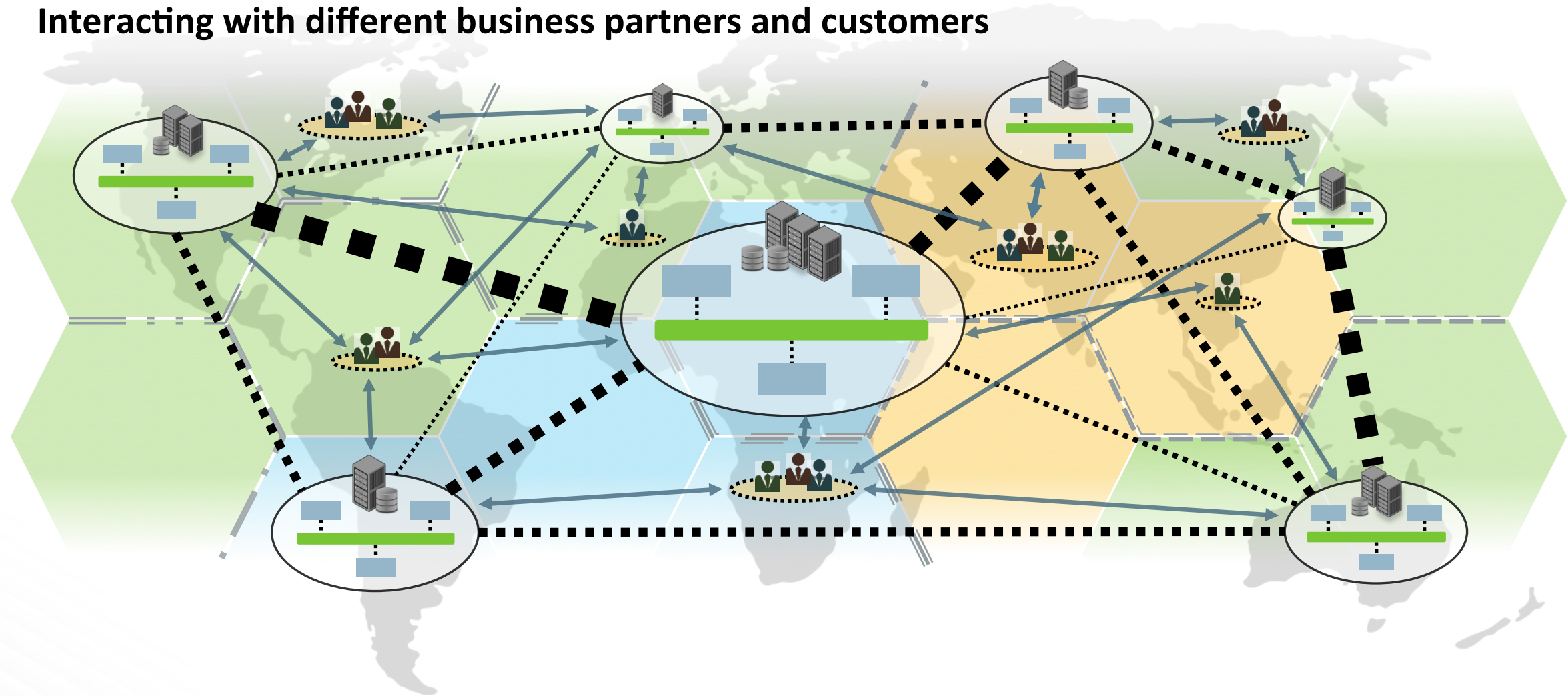
Capable of different volume, velocity, bandwidth, and latency…

Enterprise Dataflow

Interacting with different business partners and customers

Data driven organization or leader...

Wants all the data – now

To understand it all and how it relates

To better inform decisions and decision makers

To outpace adversaries/competitors and maximize value to customers

HORTONWORKS®

Challenges

Wants all the data – now

- Bandwidth: Limited, Error prone, expensive
- Denied access
- Compliance requirements (legal, policy, geographic restrictions)
- Competing priorities (perishable vs delay tolerant)
- Storage is cheap but rate of data sensing is skyrocketing
- Controlling the edge systems

To understand it all and how it relates

To better inform decisions and decision makers

To outpace adversaries/competitors and maximize value to customers

Challenges

Wants all the data – now

To understand it all and how it relates

- Formats/schemas evolve
- Semantics of data evolve
- Processing bandwidth limits
- Signal/noise results in inefficient processing
- Compliance implications of joining data

To better inform decisions and decision makers
To outpace adversaries/competitors and maximize value to customers

HORTONWORKS®

Challenges

Wants all the data – now.
To understand it all and how it relates

To better inform decisions and decision makers

- How to distribute accumulated context
- How to track origin of most/least valuable sources
- Compliance implications of sharing observations/insights

To outpace adversaries/competitors and maximize value to customers

HORTONWORKS®

Challenges

Wants all the data – now.
To understand it all and how it relates
To better inform decisions and decision makers

To outpace adversaries/competitors and maximize value to customers

- How to maximize value of distributed enterprise
- Ability to redo decisions (what if I took a different path)
- Ability to interact with live data and systems to effect change
- Trace actions to decisions to inputs…

HORTONWORKS®

'OODA Loop' – A closed loop system

Wants all the data – now
      OBSERVE

To understand it all and how it relates
      ORIENT

To better inform decisions and decision makers
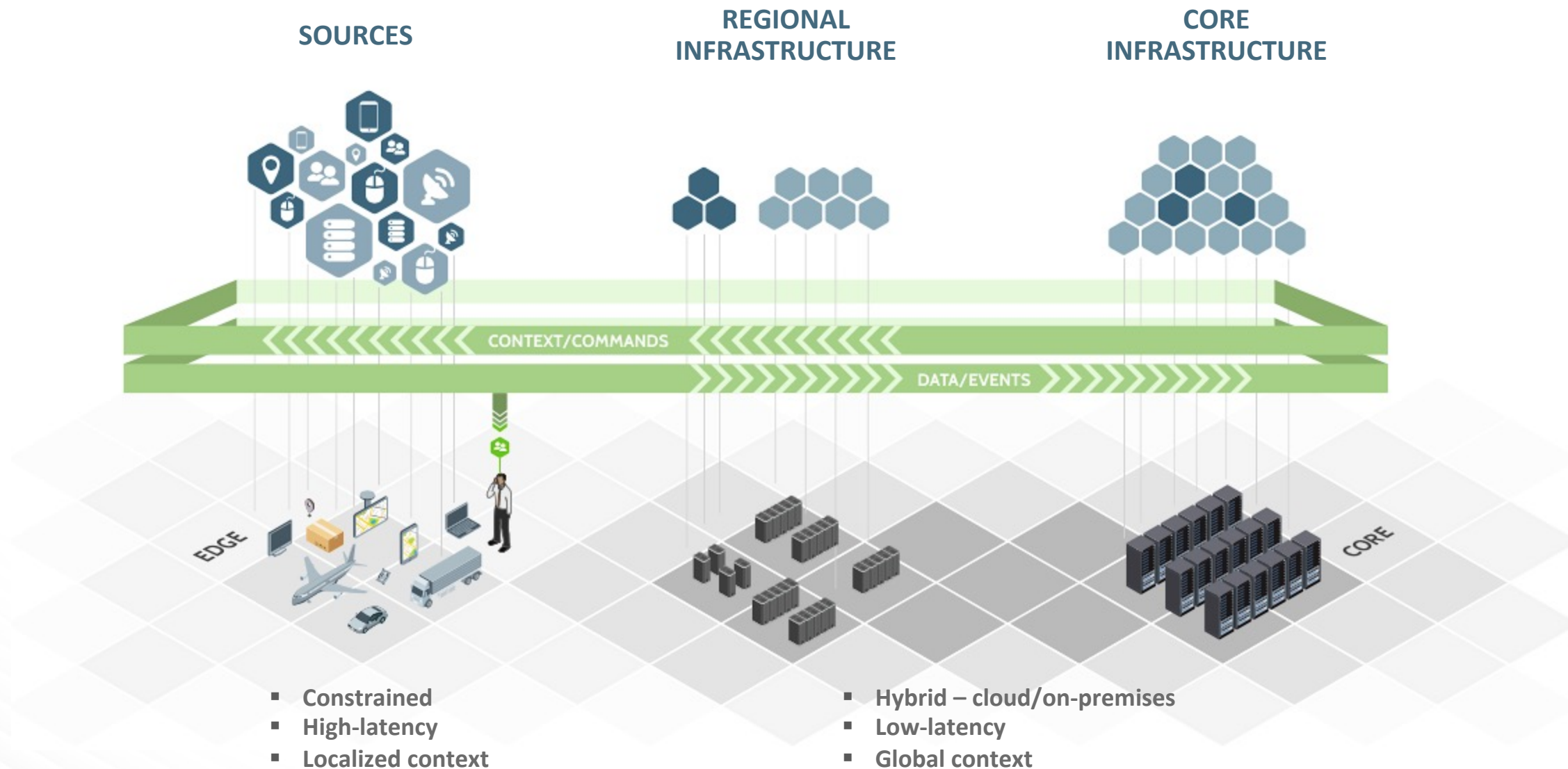   **DECIDE**

To outpace adversaries/competitors and maximize value to customers
      ACT

HORTONWORKS®

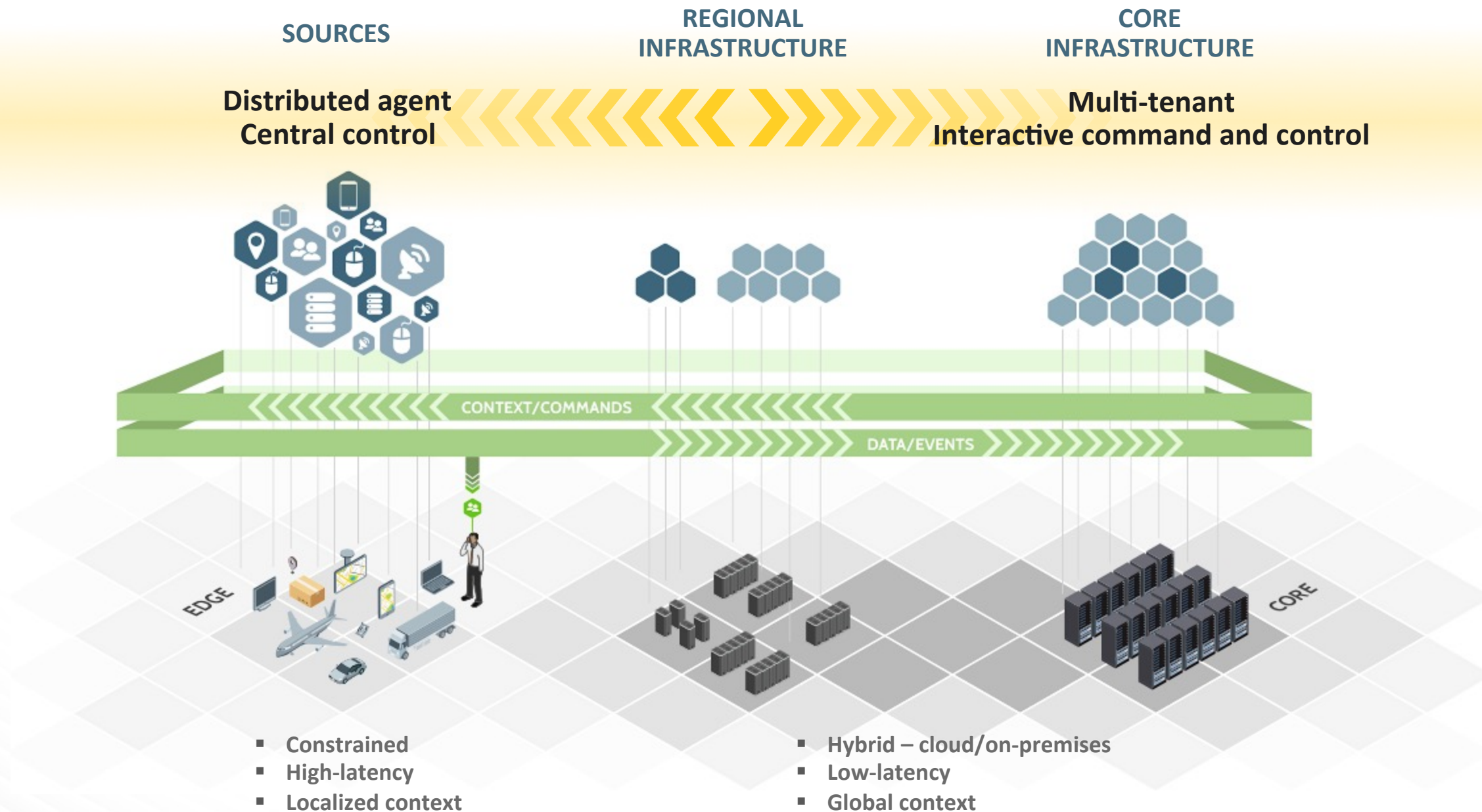It's not just how quickly you move data — it's about how quickly you can change behavior and seize new opportunities

The Data Driven Enterprise – demands a closed loop architecture

SOURCES

REGIONAL INFRASTRUCTURE

CORE INFRASTRUCTURE

CONTEXT/COMMANDS

DATA/EVENTS

EDGE

CORE

- Constrained
- High-latency
- Localized context

- Hybrid – cloud/on-premises
- Low-latency
- Global context

HORTONWORKS®

Command and Control

SOURCES   REGIONAL INFRASTRUCTURE   CORE INFRASTRUCTURE

Distributed agent
Central control

Multi-tenant
Interactive command and control

CONTEXT/COMMANDS

DATA/EVENTS

EDGE

CORE

- Constrained
- High-latency
- Localized context

- Hybrid – cloud/on-premises
- Low-latency
- Global context

HORTONWORKS

Provenance

SOURCES

REGIONAL INFRASTRUCTURE

CORE INFRASTRUCTURE

Origin – attribution
Replay – recovery

Evolution of topologies
Long retention

Types of Lineage
- Event (runtime)
- Configuration (design time)

CONTEXT/COMMANDS

DATA/EVENTS

EDGE

CORE

- Constrained
- High-latency
- Localized context

- Hybrid – cloud/on-premises
- Low-latency
- Global context

HORTONWORKS

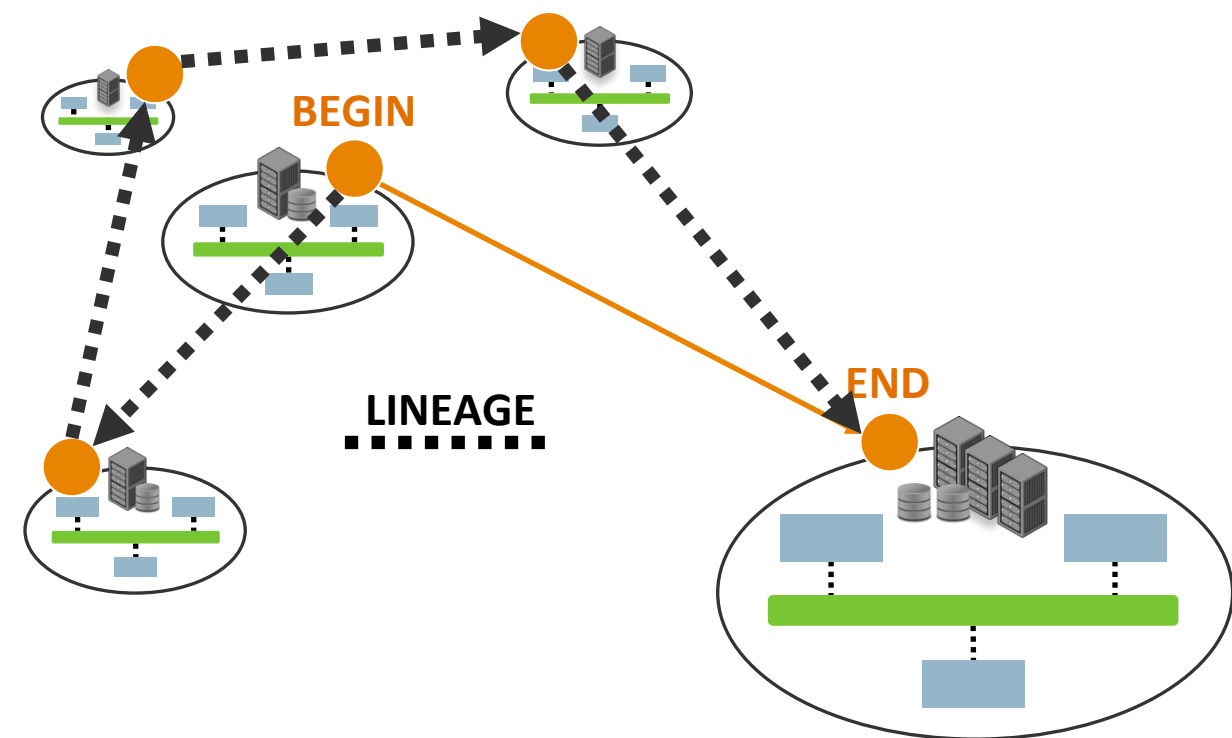# The need for data provenance

For Operators

- Traceability, lineage
- Recovery and replay
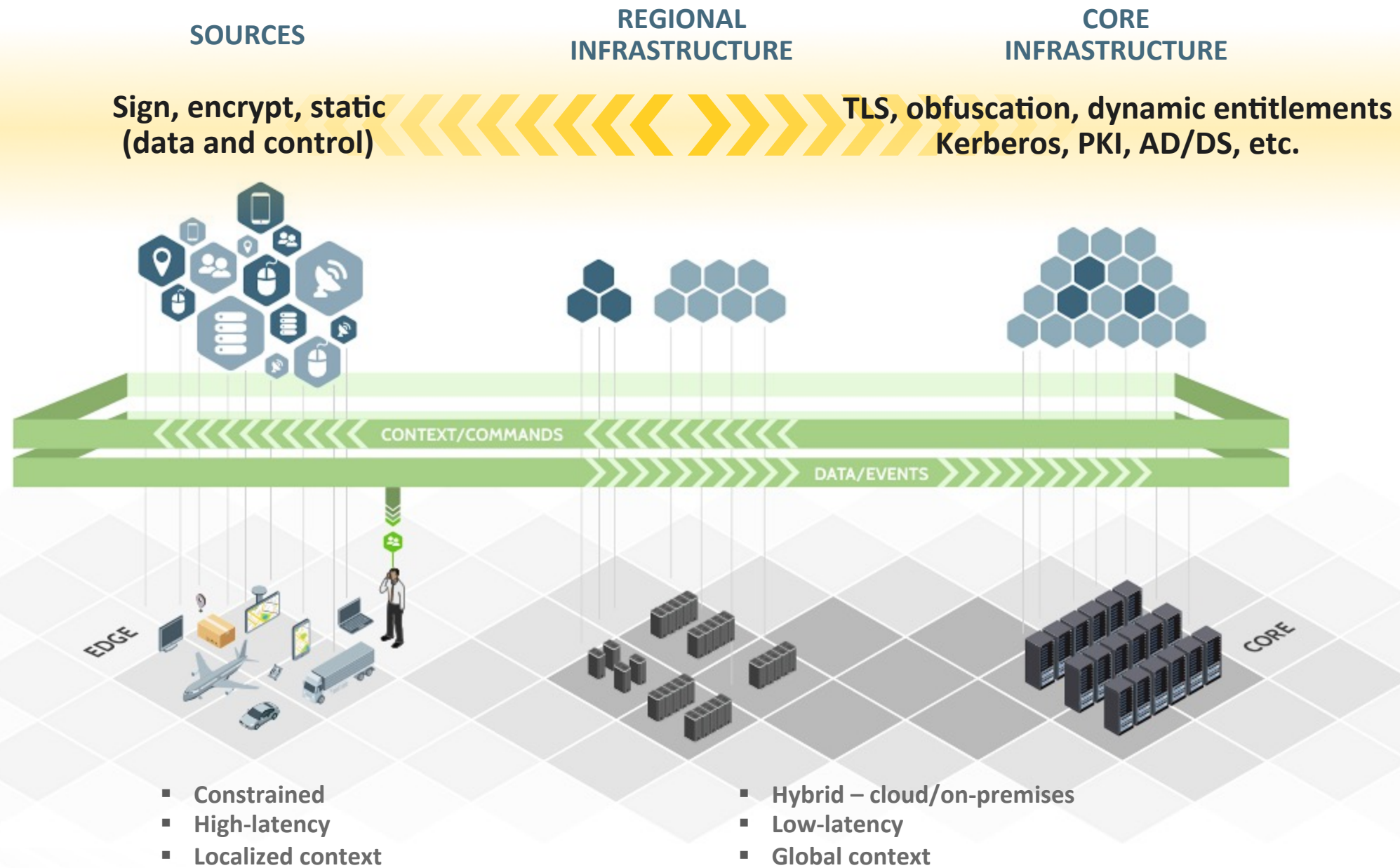
For Compliance

- Audit trail
- Remediation

For Business / Mission

- Value sources
- Value IT investment

Security

SOURCES

REGIONAL
INFRASTRUCTURE

CORE
INFRASTRUCTURE

**Sign, encrypt, static
(data and control)**

**TLS, obfuscation, dynamic entitlements
Kerberos, PKI, AD/DS, etc.**

CONTEXT/COMMANDS

DATA/EVENTS

EDGE

CORE

- Constrained
- High-latency
- Localized context

- Hybrid – cloud/on-premises
- Low-latency
- Global context

HORTONWORKS®
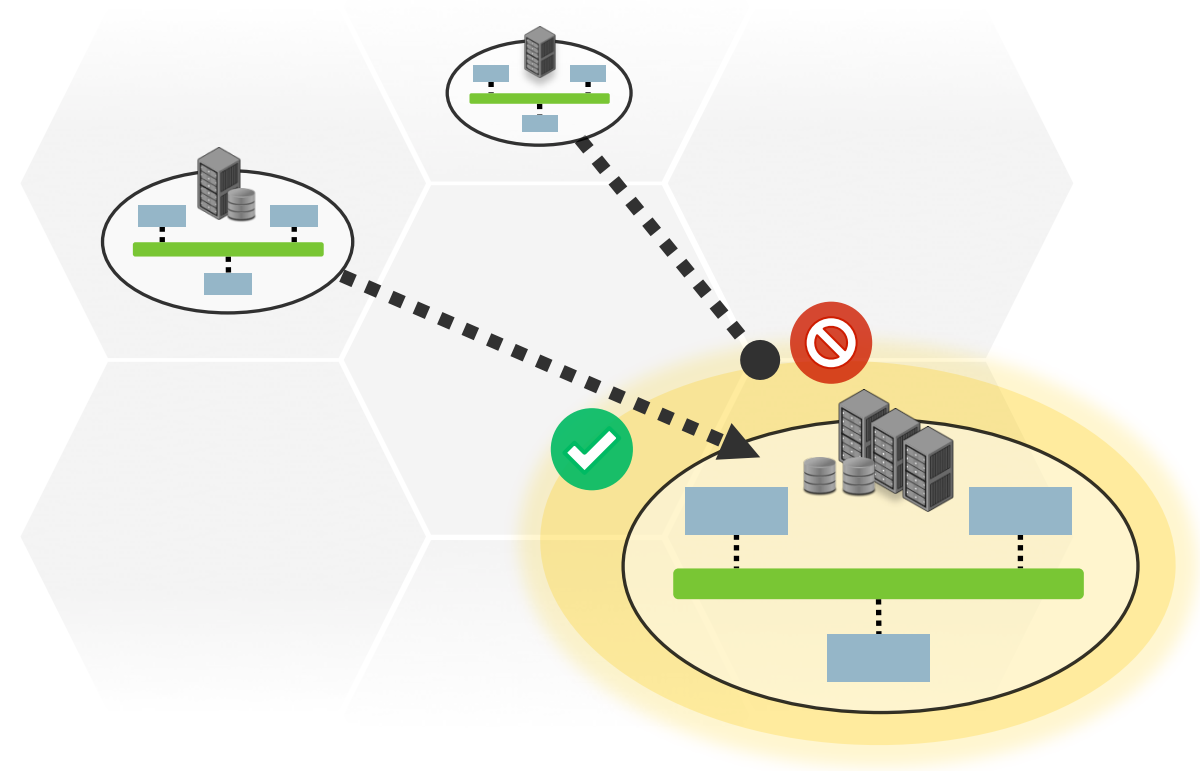
The need for fine-grained security and compliance

It's not enough to say you have encrypted communications

- Enterprise authorization services –entitlements change often

- People and systems with different roles require difference access levels

- Tagged/classified data

Topics to explore further

- Private Information Retrieval (PIR) – Doesn't this break the provenance trail?

- Role of cryptographically verifiable approaches such as Blockchain

- What standards are needed (provenance, security)?