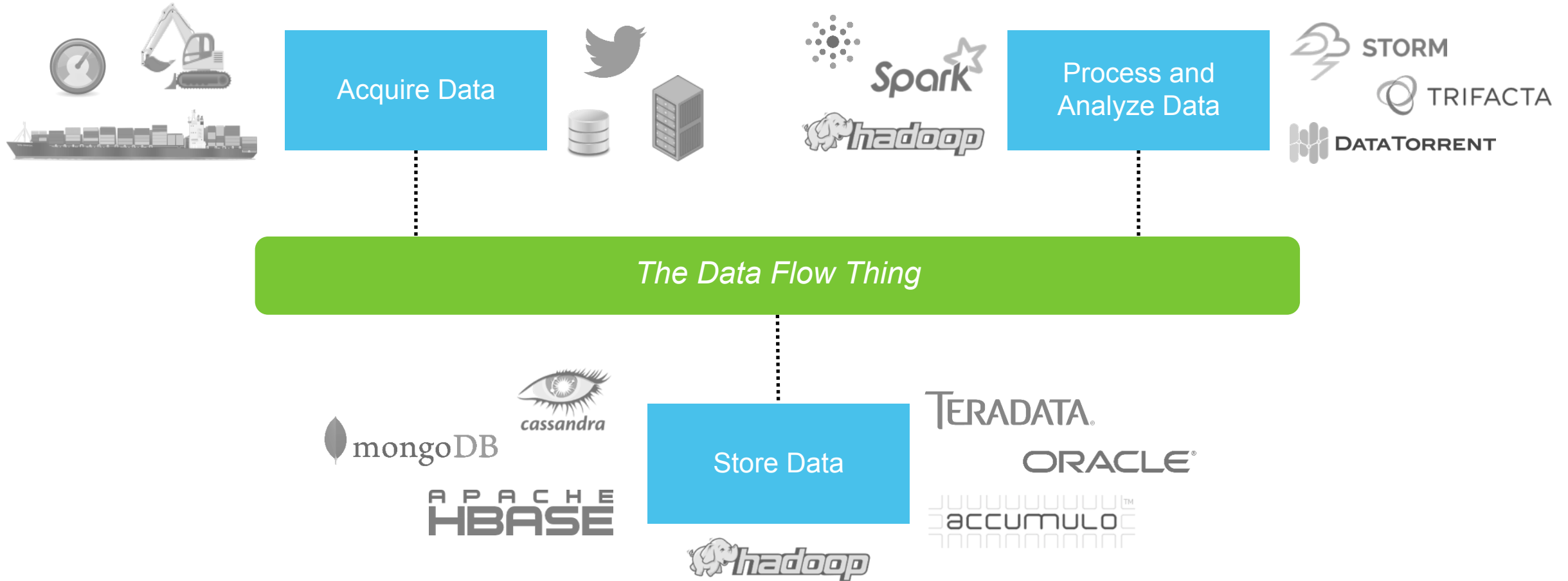# Madison Big Data Meet-Up: Apache NiFi
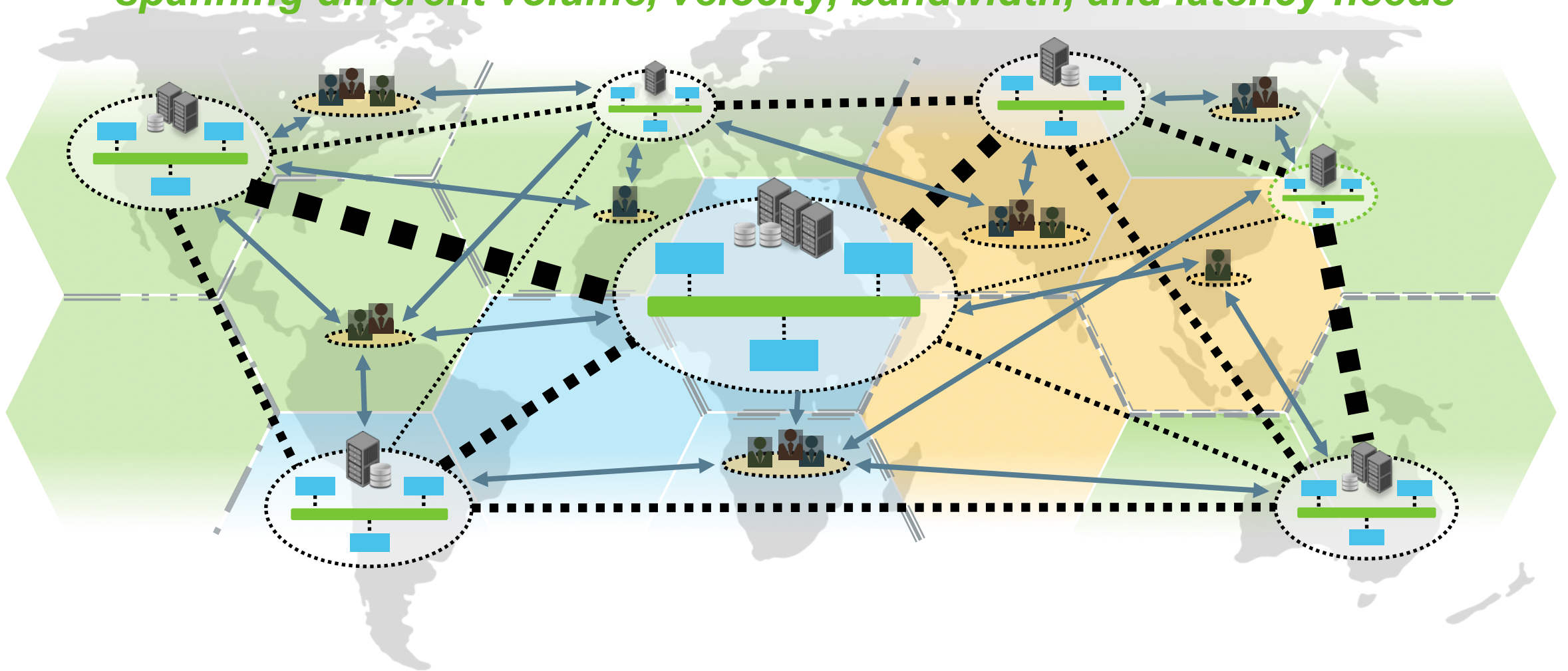
Rob Sader

Data Flow

"It's not just how quickly you move data – it's about how quickly you can change behavior and seize new opportunities"
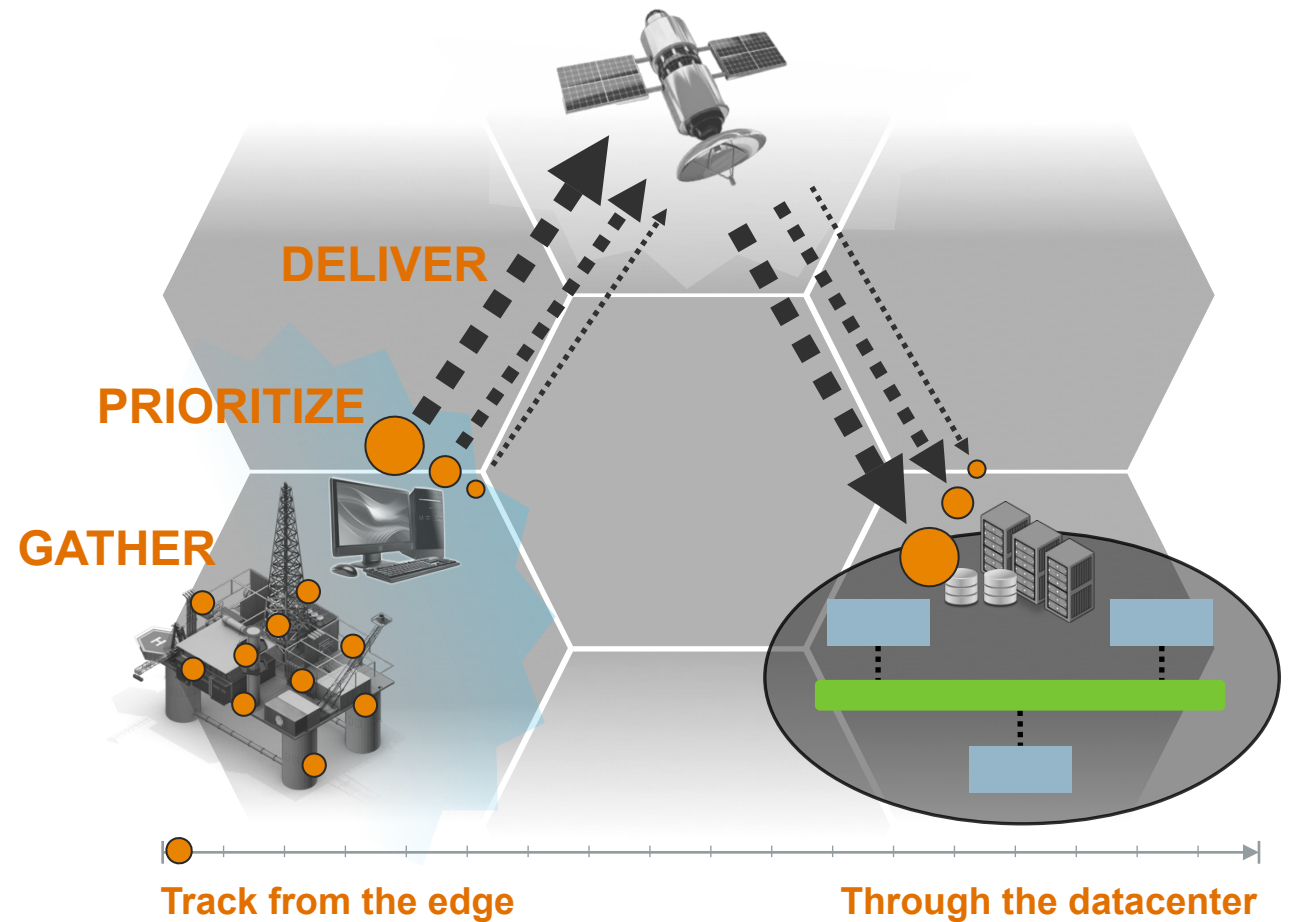
Simplistic View of IoAT & Data Flow

Realistic View of IoAT and Data Flow

Global interactions with customers, business partners, and things spanning different volume, velocity, bandwidth, and latency needs
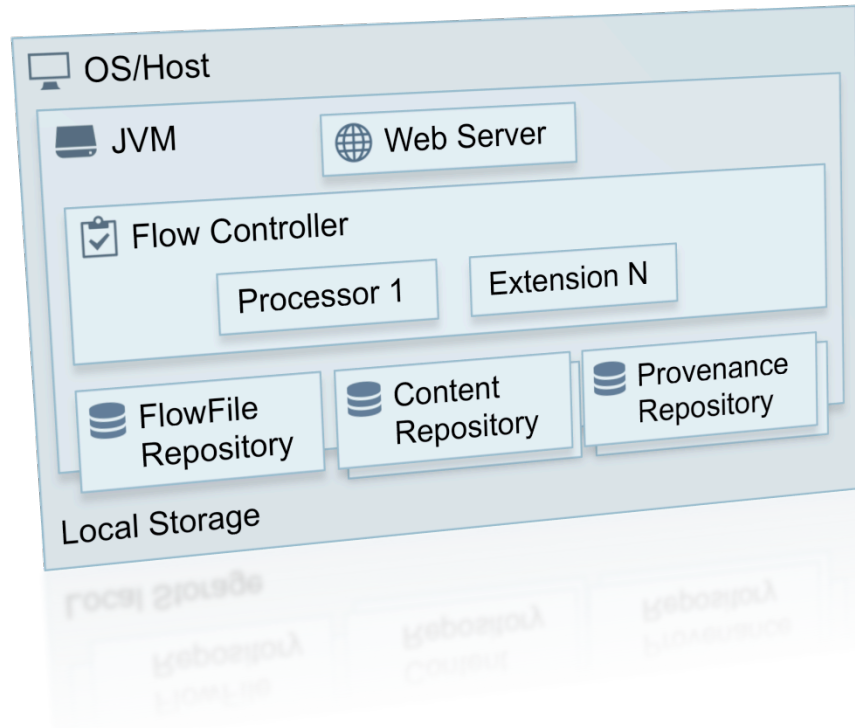
Enterprise DataFlow Challenges

- Variability in Data Protocols, Formats and Schemas

- Data Size and Speed

- Security at Data Plane

- Traceability (Data Lineage)

- Prioritization of Resources

- Multi-Directional Flow

- Recoverability and Replay

- Transparency of DataFlow

- Scaling Down

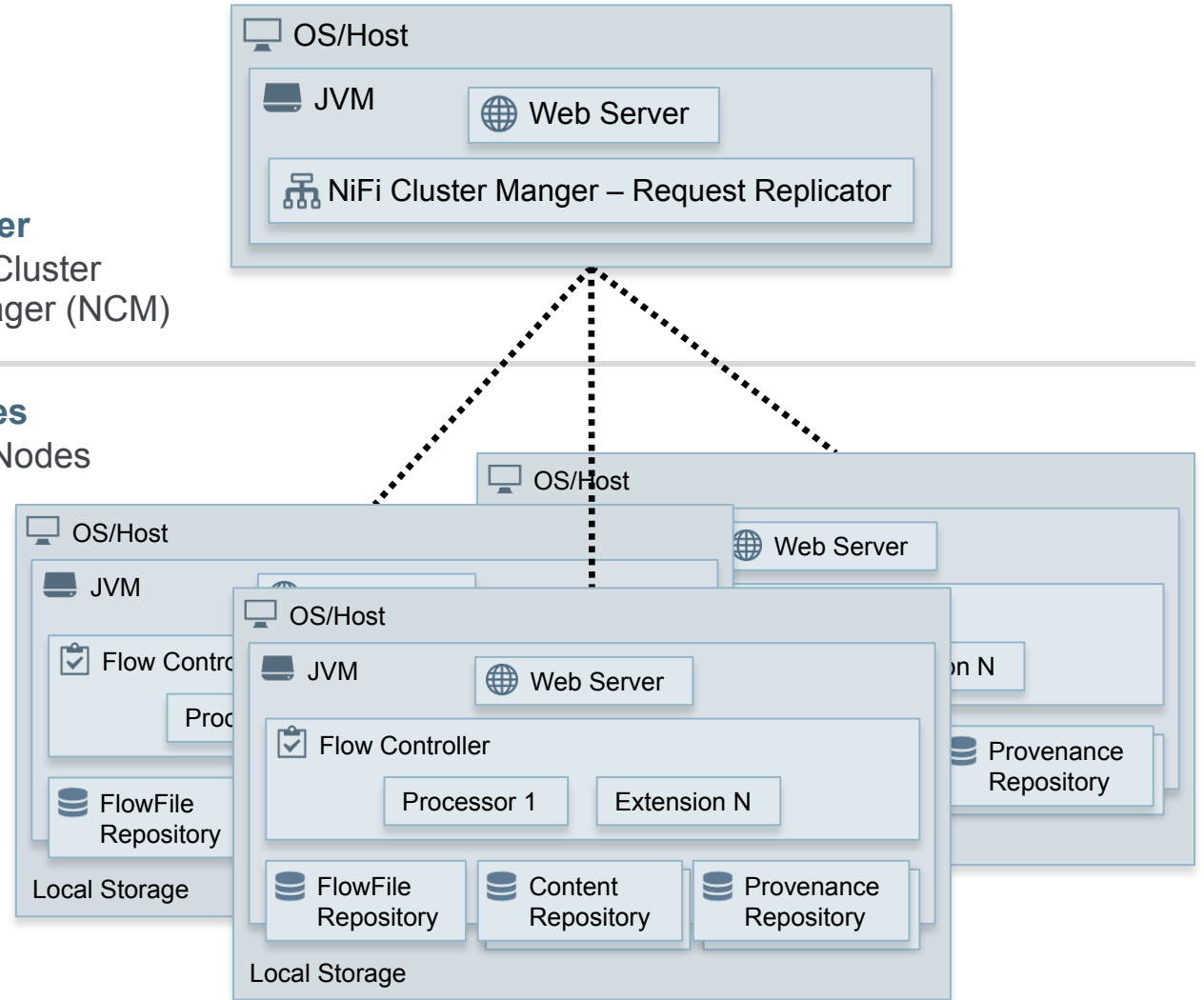- Enrichment/Transformation

- Unreliable Comms

DELIVER

PRIORITIZE

GATHER

Track from the edge   Through the datacenter

Architecture

Flow Based Programming (FBP)

FBP Term	NiFi Term	Description
Information Packet	FlowFile	Each object moving through the system.
Black Box	FlowFile Processor	Performs the work, doing some combination of data routing, transformation, or mediation between systems.
Bounded Buffer	Connection	The linkage between processors, acting as queues and allowing various processes to interact at differing rates.
Scheduler	Flow Controller	Maintains the knowledge of how processes are connected, and manages the threads and allocations thereof which all processes use.
Subnet	Process Group	A set of processes and their connections, which can receive and send data via ports. A process group allows creation of entirely new component simply by composition of its components.

Apache NiFi Demo… Thank You