AI to the Edge

Cloud AI map-reduce:
• Raw data sampling => insights
• High EMR/Spark concurrency
• Spine-leaf Scatter-gather workloads

Edge Network Aggregation:
• Sub-sec SLA use-cases
• Scale heavy fresh upload
• Leverage GPU economics
• Regulatory and commercial

Edge Alternatives to Map IOT Samples Reduce to AI Insights
Auto Case Study

Resolution 15 ~1 (meter$^2$) => HID  Resolution 9 (7$^6$ meter$^2$) => EID

EID

2001:0db8:85a3:0000:0000:8a2e:0370:7334
Cyber Case Study

5-Tuple Matrix

RAN MAN
MEC Tbs

Mapped Samples

Reduced insights

Cynamics AI Cloud
Insights Mbps

<table>
<thead>
<tr>
<th>Sample point</th>
<th>Full BW</th>
<th>Sampled BW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40Gbps</td>
<td>40Mbps</td>
</tr>
<tr>
<td>2</td>
<td>40Gbps</td>
<td>40Mbps</td>
</tr>
<tr>
<td>..</td>
<td>40Gbps</td>
<td>40Mbps</td>
</tr>
<tr>
<td>n</td>
<td>40Gbps</td>
<td>40Mbps</td>
</tr>
<tr>
<td>TOTAL</td>
<td>40*N Gbps</td>
<td>40 Gbps</td>
</tr>
</tbody>
</table>

Real time insights Mbps feeds
Virtual Routing 101: IP over UDP

Basic Overlay Standards Specify Encapsulations, NOT How to Route in the Overlays

Use Tunnels to route two sets of addresses:
(1) Logical-Overlay (2) Topological-Underlay
LISP Overlay

Mapping System

AAA:1.2
BBB:1.3

LISP - Location Identity Separation
EID - Endpoint Identifier (Logical)
RLOC - Route Locator (Topological)
XTR - Ingress/Egress Tunnel Router
RTR - Retunneling XTR (NAT/Mobile..)

* Overlay route-peering => limits dynamics
* Existing routing => topological constraints

IP Underlay Scales Mapping System: **logical** EIDs => **any** RLOC
EIDs functions as => Map-Reduce Context and Routed Data index
We can use mapping systems for Signal Free multicast. Scales to millions map-reduced scoped feeds to thousands of clients each.
Communications Matrix EIDs

Sflow/IPFIX Sampling

EID = dotGov X marketing.XYZ

f(..)

5 Tupple masks EIDs

EID: 2001:0db8:85a3:0000:0000:8a2e:0370:7334

Attacks and Stats Feeds
The Pattern

Heavy Varying IOT Upload >>

>> Signal Free Subscribed Feeds
Automotive

Vehicle-to-Cloud (V2Cloud)

100 million connected cars globally

Figure 1 Problems of existing technologies’ deployment

Heavy Fresh Vision & Sensory Continues Upload
Physical World EIDs

Resolution 15 ~1 (meter$^2$) => HID  Resolution 9 ($7^6$ meter$^2$) => EID

2001:0db8:85a3:0000:0000:8a2e:0370:7334
Map-Reduced to Feeds

H3EiD.Functions() {
Detection 1, 2, 3, ... n
Localization
Clustering
Propagation
Aging ..
}
Scheduling

Anchoring

gNB

UPF

Internet

Per 10K Tiles (blocks) x MTU ~ Mbps

Per 10K Cars x 4Mbps ~ 40Gbps

Fleet APN-VPN

gNB

Scheduling

Fleet APN-VPN

APN-VPN
Functional Reduction: EID.Context::AI( .. )

Context: DBScan, Simplex coalescing, Homography, Visual localization
Change detection, Lane number, Traffic direction, neighboring tiles …
Interoperable Off-The-Shelf

XTR Routing => Deployment Scaling
Balancing, H/A, Protections

Where/What Reduced Feed Per Context EID

Street Vision & Sensory Sampling
Routed HPC Edge

- Engage high concurrency GPU machines
- Choose best upload aggregation points
- Choose best GPU economics for the task
- Comply with private premise constraints

Portable AI Network Nodes
Each node aggregates IoT vision & sensory and generates geospatial feeds. Feeds consist of verified current state and latest filtered frames per geo-location. Feeds are available to mobile clients, driving & parking apps, command & control. Feed can be uploaded to cloud databases for history, trends, and deeper queries. Each node facilitates protected sandboxed hosted feed-taps for enterprise specific data derivatives.
EID Routing for COIN AI Edge

1. **EID Context**: Natively source-routable logical data-index

2. **XTR Map**: Edge aggregation steers raw uploads to EIDs

3. **Lambda AI Reduce**: Apply EID.context::functions(raw data)

4. **MLD Subscribe**: to portable [Source, Theme] EID feeds

5. **Scales**: Ms of feeds via standard Signal-Free Multicast

RFC6830 RFC 8378 draft-ietf-lisp-nexagon-06
draft-ietf-lisp-rfc6830bis-35 draft-ietf-lisp-rfc6833bis-29
Thank You