## Ultra Low Latency & Backhaul Optimization

<table>
<thead>
<tr>
<th>NFV Edge Infrastructure</th>
<th>Wireless (vRAN,vEPC)</th>
<th>Wireline (PON)</th>
<th>uCPE (SD-WAN)</th>
<th>IP Enterprise Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autonomous Devices</td>
<td>Drones</td>
<td>Autonomous Vehicles</td>
<td>Industry Robots</td>
<td>Medical</td>
</tr>
<tr>
<td>Immersive Experiences</td>
<td>Virtual Reality</td>
<td>Augmented Reality</td>
<td>360 Video</td>
<td>Wearable Cognitive Assistance</td>
</tr>
<tr>
<td>IoT &amp; Analytics</td>
<td>Industrial Sensors</td>
<td>Home Devices</td>
<td>Retail</td>
<td>Healthcare</td>
</tr>
<tr>
<td>On-Demand NFV</td>
<td>Hardware Acceleration</td>
<td>A.I.</td>
<td>Microservices</td>
<td>5G</td>
</tr>
</tbody>
</table>

---

Akraino Blueprint Proposal: **Real-time Edge Media Processing**
Use Case Overview: Edge Media Processing & Analytics

Ultra Low Latency & Backhaul Optimization

Real-time Media Applications

1. RAN Aware Bandwidth Optimization
2. Edge Video Conferencing
3. Media Transcoding & Adaptation
4. Video Analytics / Meta Data
5. Video Quality Metrics KPIs
6. User Generated Content Upload Optimization
7. GPU Acceleration (4K, 8K) & AI / ML Analytics
Ref Use Case #1: Media Traffic Steering/Optimization @ Edge (Surveillance)

- Local Breakout @ Edge - **Backhaul savings** Continuous Surveillance Feeds
- **Real time traffic steering** on Intrusion Detection (eg: face anomaly detection) @ Feed to Core
Ref Use Case #2: Media Traffic Steering/Optimization @ Edge (AR/VR)

WebRTC based AR server

Traffic rule to Offload inbound media stream

Offloaded media stream traffic

UE running WebRTC client

Media stream generated by UE camera

Media stream with cap applied over head of all faces detected

Backhaul Savings and URLLC
Ref Use Case #3: Media Analytics @ edge (Differentiated Media Processing)

- **Differentiated media processing** (trans-rating) applied to ongoing user data stream based on **real time channel quality** reported by the MEC platform.
### Business Drivers

<table>
<thead>
<tr>
<th>Use Case Characteristics</th>
<th>Network Cloud (Real Time Edge Media Applications)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Need</td>
<td>Enabling new applications requiring distributed edge deployments. Low Latency &amp; Bandwidth Optimized Real Time Media Processing and Edge Media AI Analytics.</td>
</tr>
<tr>
<td>User Experience</td>
<td>HD and UHD real-time media content, contextual media processing, low latency Zero touch provisioning and automated remote deployments</td>
</tr>
<tr>
<td>Cost Of Solution</td>
<td>Low Cost with Virtualization and Open HW Range from small footprint SW solution to large scale deployments with HW acceleration for advanced high density AI based media processing</td>
</tr>
<tr>
<td>Scale</td>
<td>Highly Scalable Fully virtualized with low footprint (&lt;4 compute servers) to high density for supporting large scale user services (100s of sites with optionally HW accelerated compute nodes)</td>
</tr>
<tr>
<td>Applications</td>
<td>Within the real time edge media processing and analytics (audio / video) domain Real time HD video processing / transcoding Augmented and Virtual Reality (BW optimized) IoT initiated audio and video stream processing / recording Media Analytics (AI and ML based), speech &amp; video</td>
</tr>
<tr>
<td>Power restrictions</td>
<td>TBD</td>
</tr>
</tbody>
</table>
## Use Case: Edge Media Processing

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>New</td>
</tr>
<tr>
<td>Industry Sector</td>
<td>Telco Carrier Networks and Enterprises</td>
</tr>
<tr>
<td>Business driver</td>
<td>Vast amounts of mobile/wireline data (predominantly video) is expected to continue to grow, particularly with 5G and IoT. Low latency, backhaul bandwidth restrictions/cost, and real time edge media analytics require media processing at network edges versus transporting all media to network core. Without the ability to process real time media at the network edges a number of new advanced applications would not be possible nor economically viable.</td>
</tr>
</tbody>
</table>

### Business use cases

1. Edge deployments at enterprises, entertainment venues, factory automation plants, public facilities where real time media processing required  
2. Edge media applications include multi-party conferencing, gaming, surveillance, IoT generated content, AR and VR applications  
3. Edge media applications requiring low latency and to overcome backhaul BW availability and costs being prohibitive  
4. Real time media analytics with AI and ML based applications for high value and media monetization applications

### Business Cost - Initial Build Cost Target Objective

Initial build requires a small footprint POD with minimal fabric and management switch, 4+ compute nodes with optional GPU acceleration, local storage node(s), PSUs, rack, typically under $100K with SW

### Business Cost – Target Operational Objective

1. Low operation cost, with support for remote FCAPS management, and ONAP based zero-touch resource and service orchestration  
2. Typical 16U height OCP rack with similar power consumption, with minimal footprint of 2 compute nodes.  
3. Edge Media solution shall support POD level consolidated management (OSAM) and service level orchestration and LCM via ONAP.  
4. Zero touch provisioning, upgrades, fault and performance management KPI, and auto-scaling and auto-healing capabilities

### Security need

POD platform SW and application level security vulnerability scanning and automated patching capabilities required  
Media content security and user access authentication capabilities required

### Regulations

Depending on type of Edge Media application GDPR or other regulatory requirements may be applicable. NEBS may be required depending on deployment location and carrier network requirements

### Other restrictions

Depending on deployment location, a single half-height rack to multiple full-height racks at Edge DC or Edge CO locations may drive power and cooling requirements

### Additional details

Edge Media solution shall enable support for high density media processing via GPU or FPGA acceleration for advanced high density AI and ML applications and shall scale from single site to 100s in regional deployments to 1000s globally  
Additional details on architecture and use cases documented in supplementary PPT
<table>
<thead>
<tr>
<th>Case Attributes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
<td>New</td>
</tr>
<tr>
<td><strong>Blueprint Family - Proposed Name</strong></td>
<td>Network Cloud, RT Cloud</td>
</tr>
<tr>
<td><strong>Use Case</strong></td>
<td>Real Time Edge Media Processing</td>
</tr>
</tbody>
</table>
| **Blueprint proposed**          | 1. Unicycle POD (4-6 servers, single 16U rack configurations)  
2. Tricycle POD (16U or 42U rack configurations, multi-rack)  
3. Cruiser POD (Multi-rack Core Network Configurations, with spine leaf fabric and ToR switch) |
| **Initial POD Cost (capex)**    | Estimates (TBD)  
1. Unicycle POD (< 100K)  
2. Tricycle POD (< 200K)  
3. Cruiser POD (< 300K) |
| **Scale**                       | 1. Unicycle POD – 1 rack with < 6 servers  
2. Tricycle POD – Multiple racks, each with < 24 servers  
3. Cruiser POD – Multiple racks, each with < 96 servers |
| **Applications**                | Edge Virtual Function Applications (reference)  
1. Edge deployments at enterprises, entertainment venues, factory automation plants, public facilities where real time media processing required  
2. Edge media applications include multi-party conferencing, gaming, surveillance, IoT generated content, AR and VR applications  
3. Edge media applications requiring low latency and to overcome backhaul BW availability and costs being prohibitive  
4. Real time media analytics with AI and ML based applications for high value and media monetization applications |
| **Power Restrictions**          | TBD                                                                                                                                                                                                        |
| **Preferred Infrastructure Orchestration** | OS – CentOS or similar Linux, KVM  
Under Cloud – Airship  
OpenStack – VM Orchestration  
Docker + K8S - Container Orchestration  
VNF Orchestration - ONAP |
| **SDN**                         | OVS-DPDK, SR-IOV                                                                                                                                                                                             |
| **Workload Type**               | VMs, Containers                                                                                                                                                                                              |
| **Additional Details**          | Edge Media solution shall enable support for high density media processing via GPU or FPGA acceleration for advanced high density AI and ML applications. |
Thank You