# Towards a Secure and Flexible Personal Data Platform on the Edge

Tamer Nadeem

Virginia Commonwealth University
USA

tnadeem@vcu.edu

in
collaboration
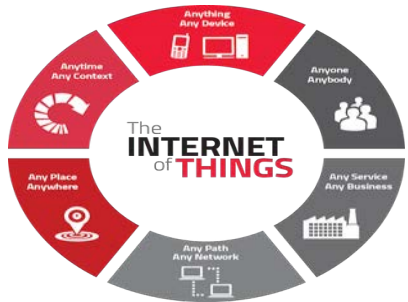with

Hamed Haddadi

Imperial College London
UK

h.haddadi@imperial.ac.uk

**ICE-T PI Meeting, Coimbra, Portugal, Nov. 8, 2019**

# Rapid growth of mobile data traffic

- Number of smart device users expected to exceed 6 billion by 2020
- IoT connected objects are expected to reach 18 billion by 2022

**Mobile devices runs numerous and wide variety of applications**

**High volume of personal traffic**

- By 2020, it's estimated that **1.7MB** of data will be created **every second for every person** on earth.

# Growth of Sensitive Apps

- **Sensitive applications communicate sensitive data over internet**



**Medical Information**:
Blood Pressure Monitoring , Diabetes.



**Activity Tracking:**
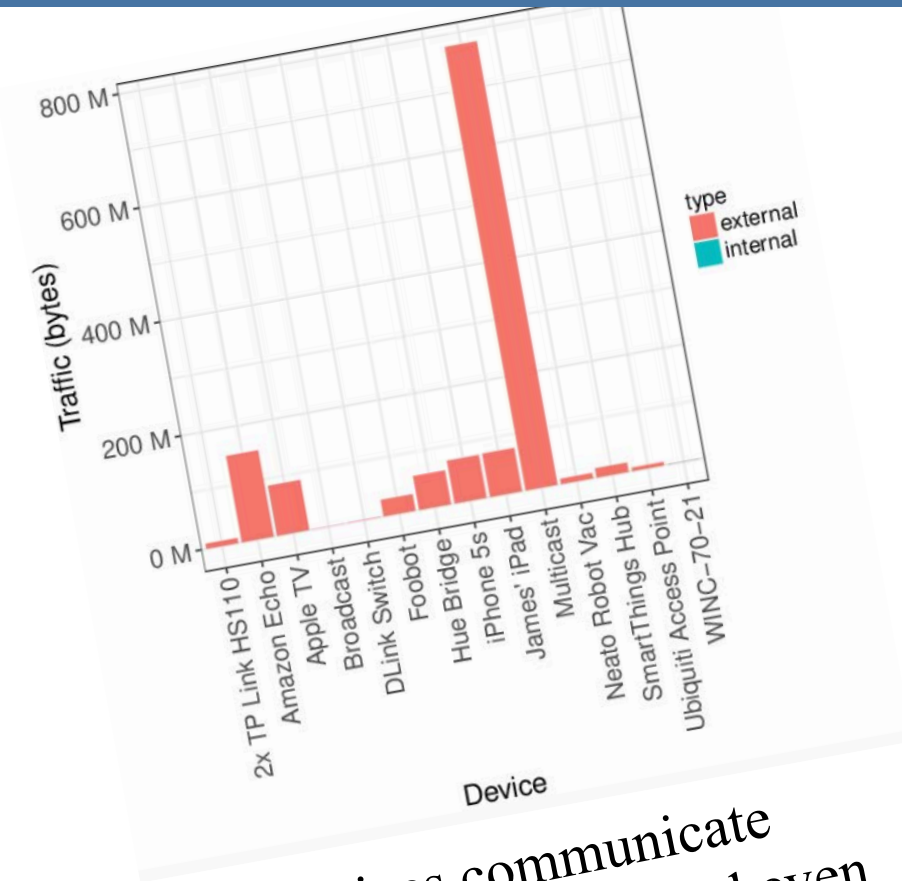Sleeping Patterns , Exercise Routines.

# Growth of Leaking Data



Cambridge Analytica acquired access to private data on millions of Facebook users



Fitness tracking app Strava gives away location of secret US army bases



IoT devices communicate significant data to the cloud even when there is no interaction

4

# Challenges

- Our ability to collect and process data has overwhelmed our ability to protect that information.

- Concerns over privacy, trust, and security are becoming increasingly important as different stakeholders attempt to take advantage of such rich data resources.

**Challenge:** how to provide technical means that enable and assist individuals in managing access to their data by others while retaining personal control over such uses and understanding the implications of any data release.

# Solution: Databox

- A platform for managing secure access to data and enabling authorized third parties to provide the owner authenticated control and accountability.
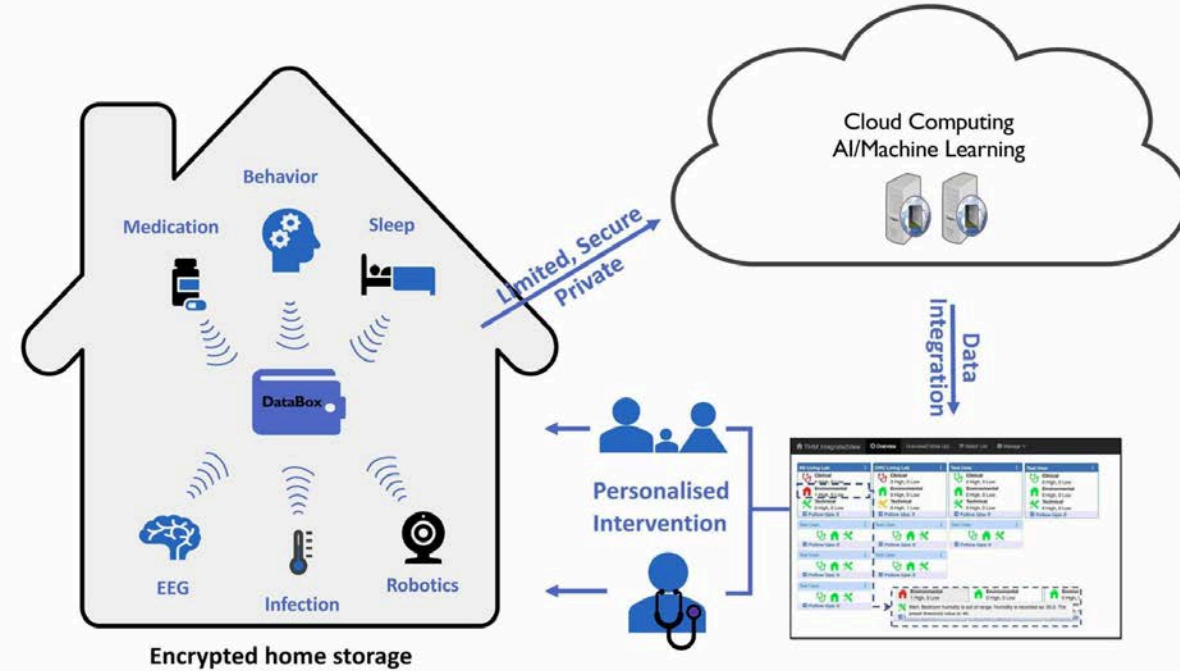
- Databox is a multi-partner research project funded by a £1.2 million Engineering and Physical Sciences Research Council (EPSRC) grant.

# Challeneges

- The various applications support various domains and services, generate and access different data patterns such as periodic, event-based, realtime and continuous data.
  - Sensor-based applications like motion detection generate and access data on periodic bases.
  - Event-based applications generate and access data only when the device state changes.
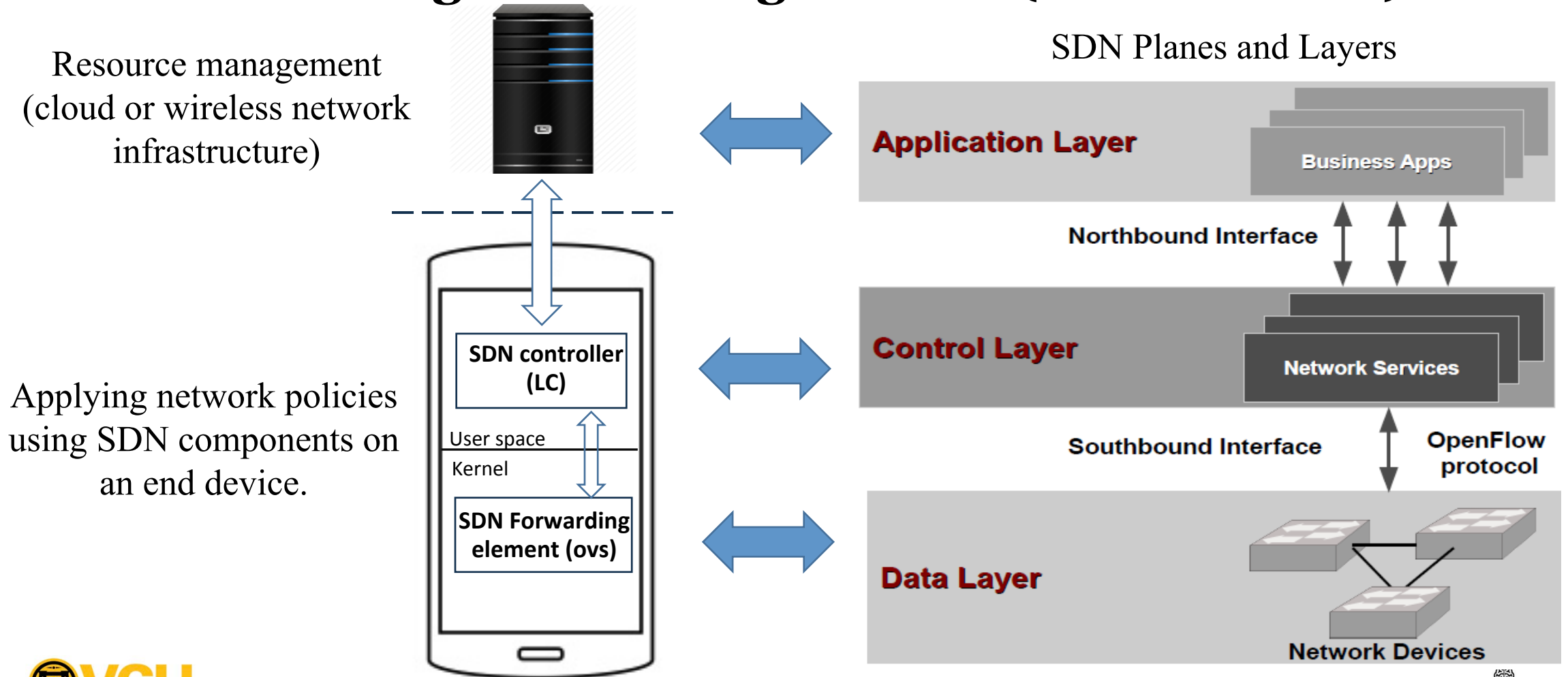  - Video/audio monitoring applications continuously generate and access streams of data.

AR/VR
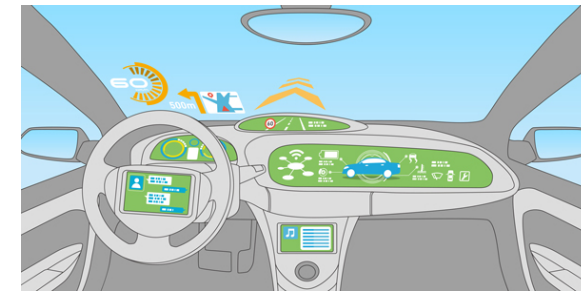
autonomous vehicle

**Challenge:** have greater visibility and control over the traffic generated from smart and IoT devices in order to guarantee an optimized performance of smart and IoT applications as well as high quality of experience to users.

# Solution: SMILE - SMart and Intelligent wireLess Edge

## Utilizing SDN on Edge Device (*extreme SDN*)

Resource management (cloud or wireless network infrastructure)

SDN Planes and Layers

Applying network policies using SDN components on an end device.

SDN controller (LC)

User space

Kernel

SDN Forwarding element (ovs)

**Application Layer**

Business Apps

Northbound Interface

**Control Layer**

Network Services

Southbound Interface

OpenFlow protocol

**Data Layer**

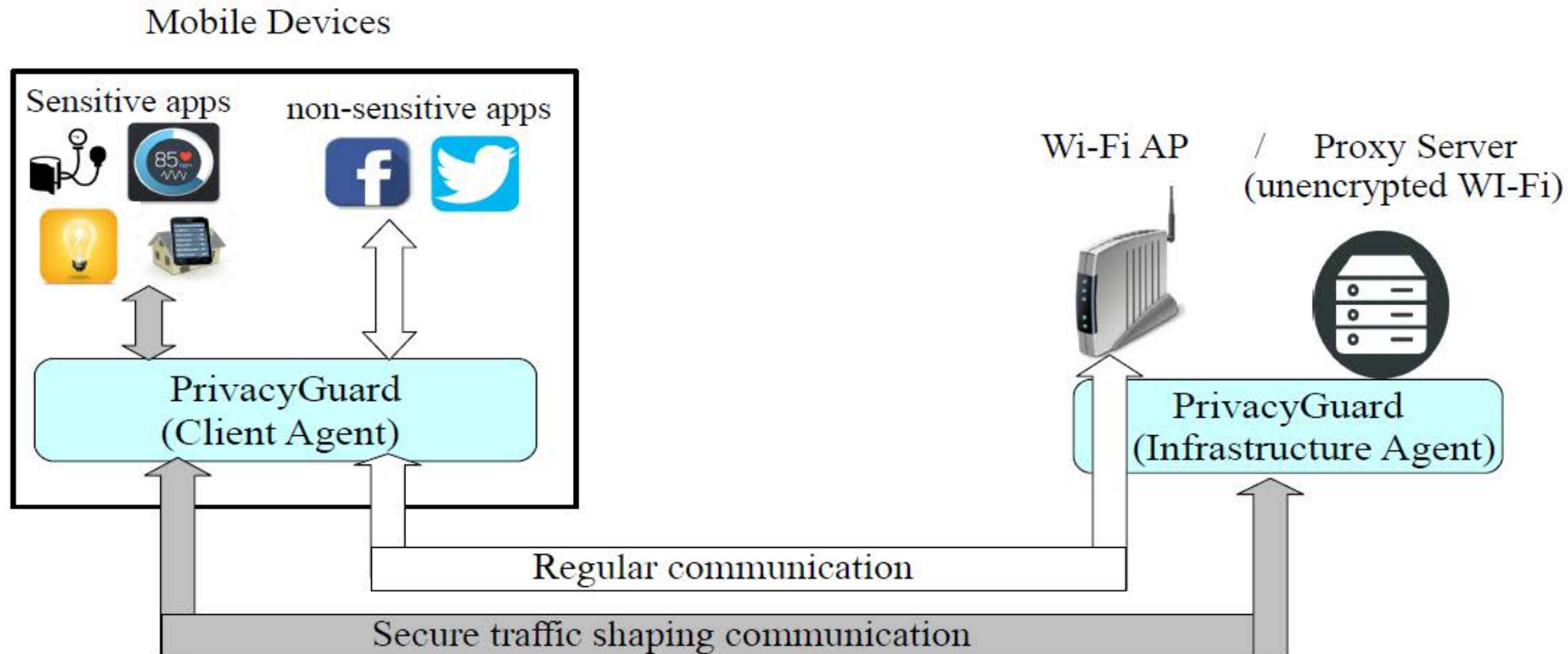Network Devices

# Solution: SMILE - SMart and Intelligent wireLess Edge

- Utilizing SDN on Edge Device (*extreme SDN*)
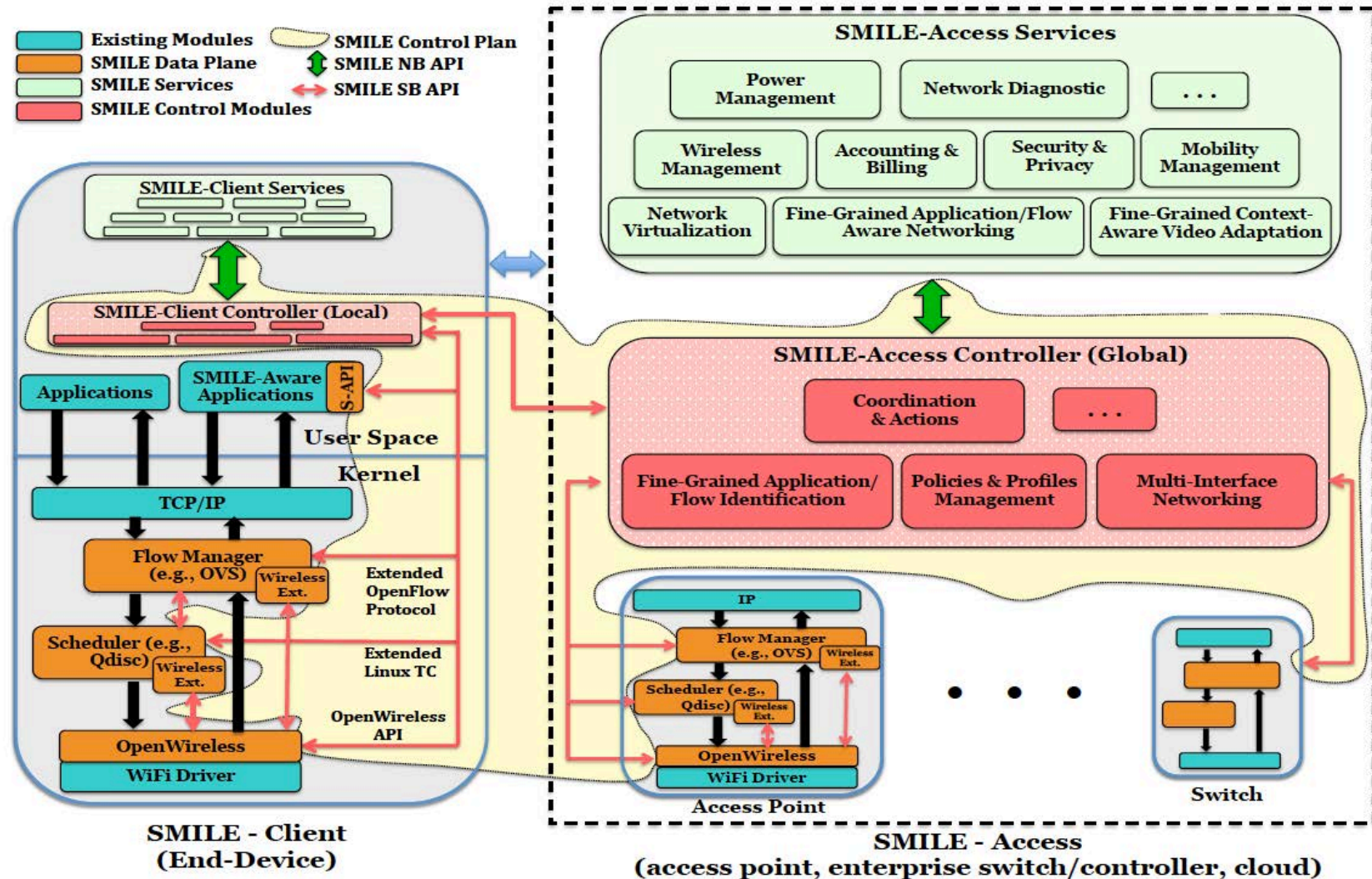
# SMILE App: PrivacyGuard

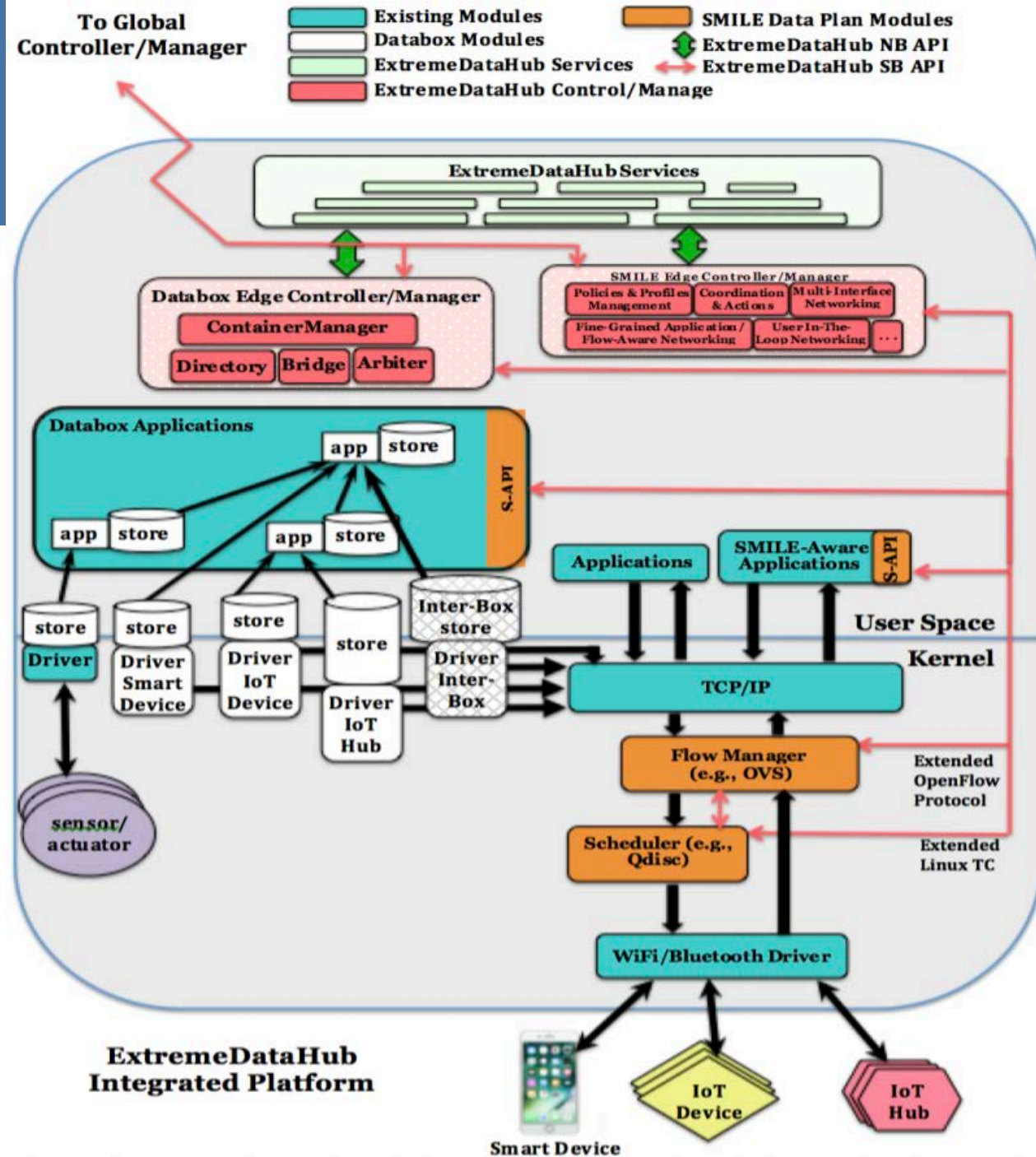# ExtemeDataHub

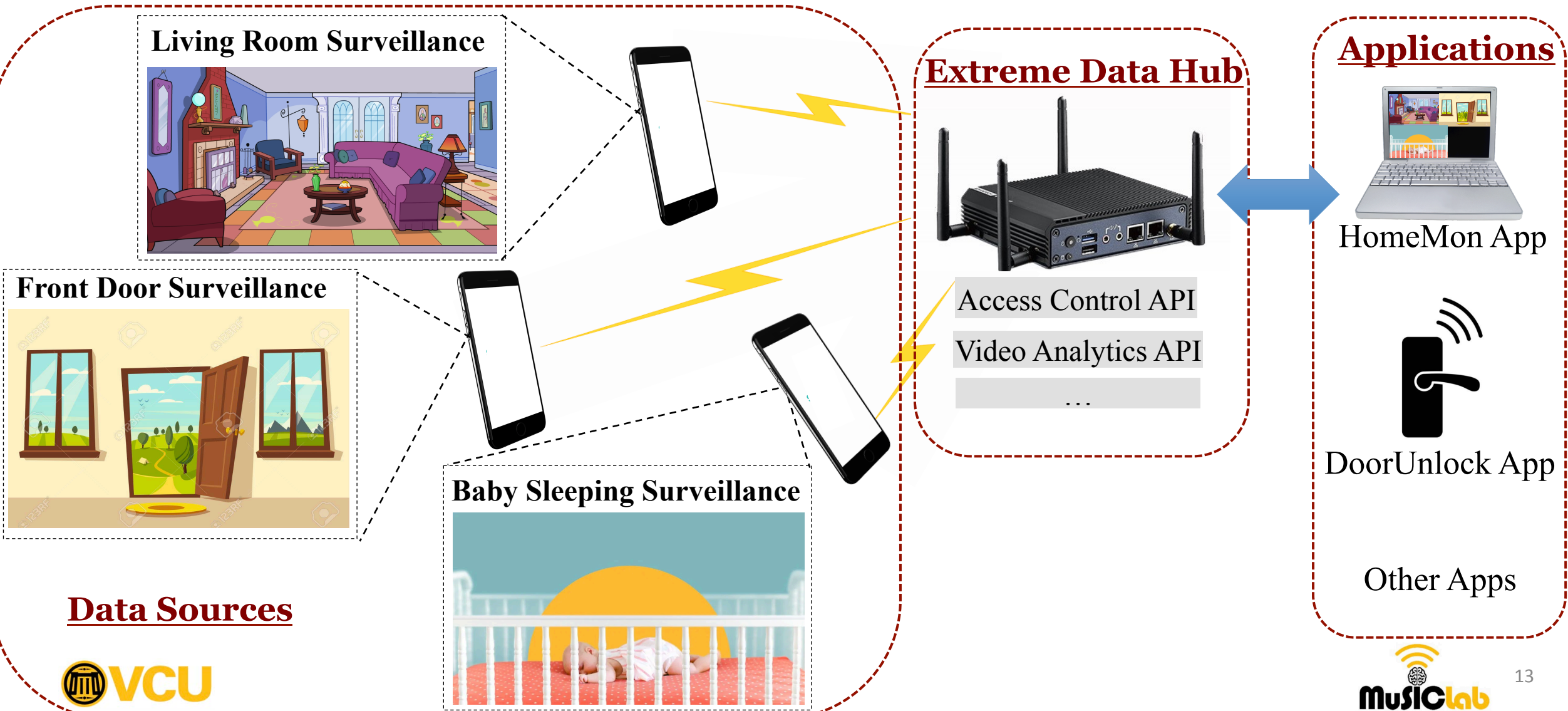Need for programmable networked edge device that:

1. Mediates access to sensitive and personal data.

2. Allows for fine-grained and intelligent management of bandwidth based on real time context awareness and specified policy.

3. Implements network policies.

4. Offers universal approach to work across network technologies, WiFi and cellular.

5. Transparent to applications.

# Use Case - HomeMon

- Streaming video feeds from multiple smartphone devices acting as video camera sensors to the ExtremeDataHub

- Build different home applications such as:
  - HomeMon app: Monitoring different rooms and places in home
  - DoorUnlock app: Unlock the door based on face identification

- Different real-time video feeds have different significance with different priorities. For example the baby-monitor video feed has a higher priority than the living-room video feed.

# Use Case - HomeMon

**Living Room Surveillance**

**Front Door Surveillance**

**Baby Sleeping Surveillance**

**Data Sources**

**Extreme Data Hub**

Access Control API

Video Analytics API

…

**Applications**

HomeMon App

DoorUnlock App

Other Apps

13

# HomeMon - Challenge

- Different real-time video feeds have different significance with different priorities. For example the baby-monitor video feed has a higher priority than the living-room video feed.

- With basic Databox service, several concurrent video feeds will cause network congestion that will results in significant quality loss in all video feeds

- Integrating with SMILE SDN framework, we are able to develop a video streaming differentiation service based on the feed priorities that guarantee certain quality of the high priority video feeds

# HomeMon – Architecture Details

Databox version 0.5.1
Open vSwitch version 2.12
Arch linux (Version 2019.03.01 - Kernel 4.20.13)

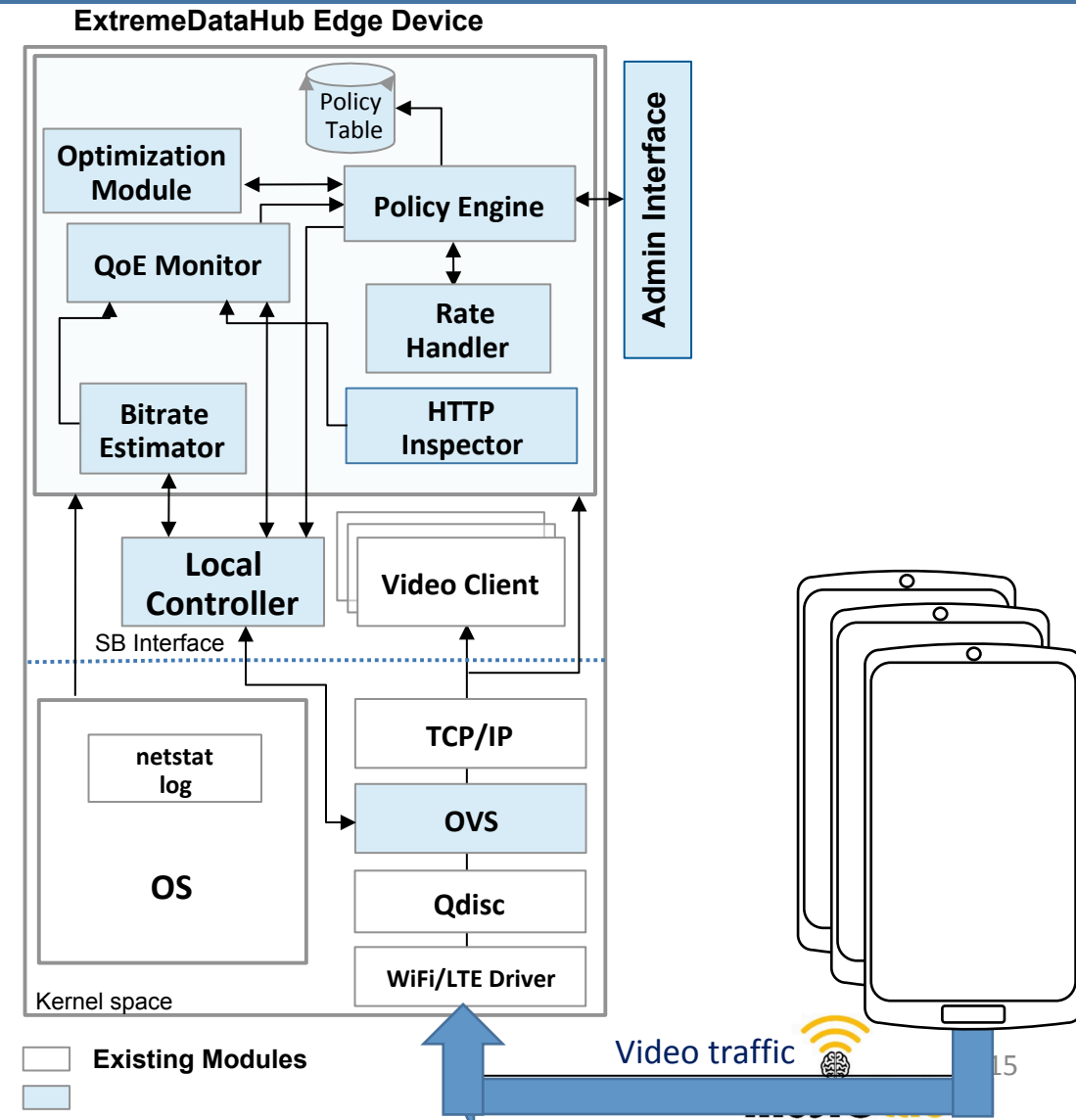Phones used:
   Samsung S9
   Samsung S5
   ZTE Axon 7

PC used to host Databox:
   CPU: AMD 2700X (8 Core, 16 Threads @4.0Ghz)
   GPU: RTX 2080
   RAM: 16gb
   SSD: Samsung EVO 850 240gb



**ExtremeDataHub Edge Device**

# HomeMon – Optimization Module

**Optimization Problem**

$$\max_{x_{ij}} \sum_{i=1}^{N} \sum_{j=1}^{K_i} (u_{ij} - \mu\delta_{ij})x_{ij}$$

$$\text{subject to} \sum_{i=1}^{N} \sum_{j=1}^{K_i} (\epsilon r_{ij})x_{ij} \leq B$$

$$\sum_{j=1}^{K_i} x_{ij} = 1, \; x_{ij} \in 0, 1 \; \forall i$$

**Utility Function**
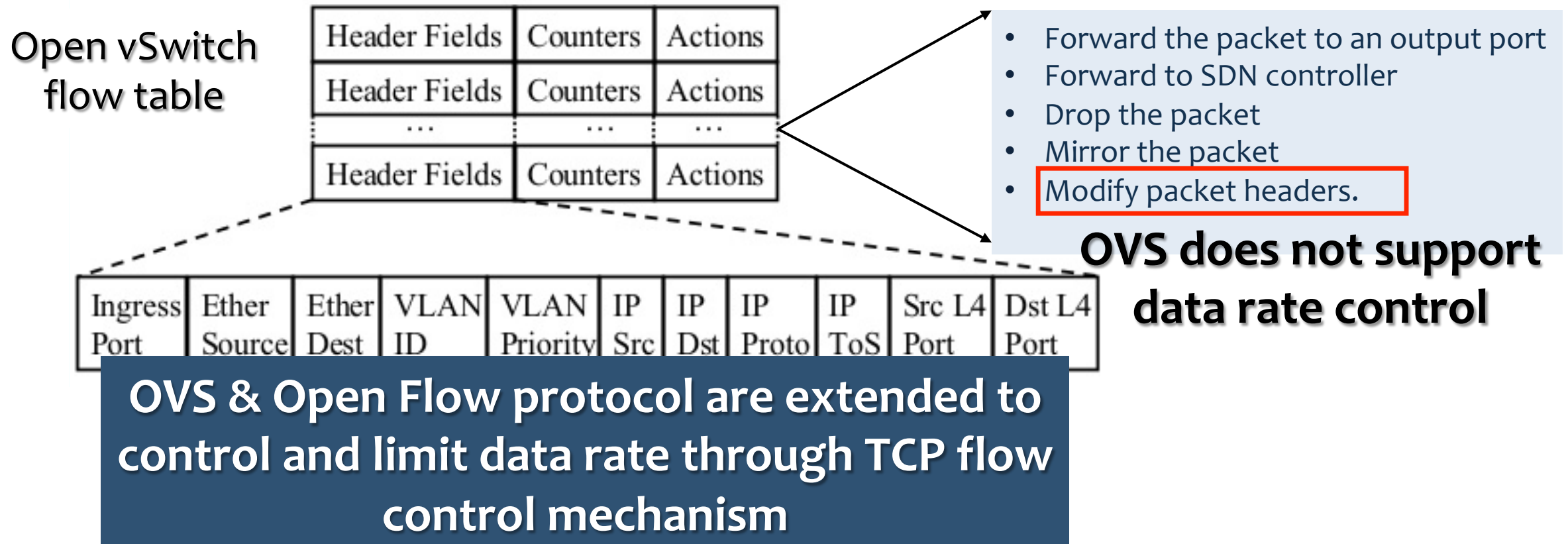
$$u_{ij} = \prod_{l=1}^{a} \beta_{il} . \log(r_{ij})$$

**Penalty Function**

$$\delta_{ij} = \begin{cases} |r_{ij} - r_{ic}|s_i + (m - \lceil\frac{t_i}{k}\rceil), & t < t_{thresh} \\ |r_{ij} - r_{ic}|s_i, & t \geq t_{thresh} \end{cases}$$

VCU

MuSICLab

# HomeMon - Implementation

- Extending SDN planes to enable controlling the data rate on the end device.

Open vSwitch flow table

| Header Fields | Counters | Actions |
|---|---|---|
| Header Fields | Counters | Actions |
| ... | ... | ... |
| Header Fields | Counters | Actions |

- Forward the packet to an output port
- Forward to SDN controller
- Drop the packet
- Mirror the packet
- Modify packet headers.

**OVS does not support data rate control**

| Ingress Port | Ether Source | Ether Dest | VLAN ID | VLAN Priority | IP Src | IP Dst | IP Proto | IP ToS | Src L4 Port | Dst L4 Port |
|---|---|---|---|---|---|---|---|---|---|---|

**OVS & Open Flow protocol are extended to control and limit data rate through TCP flow control mechanism**

# HomeMon - Demo

| Normal Operation | Congested | Controlled |
|---|---|---|
| • 2 1080p video each streaming @ 5Mbps<br>• Network bandwidth needed ~10Mbps | • Incoming traffic to the hub is capped @8Mbps<br>• Traffic from both devices cannot be accommodated<br>• Both streams lowered to 240p streams, with each needing 500Kbps | • The lower priority device is forced to lower its quality<br>• This allows for high priority video to continue its stream @1080p<br>• Combined needed bandwidth is 5.5Mbps |



Bandwidth Utilization Over Time
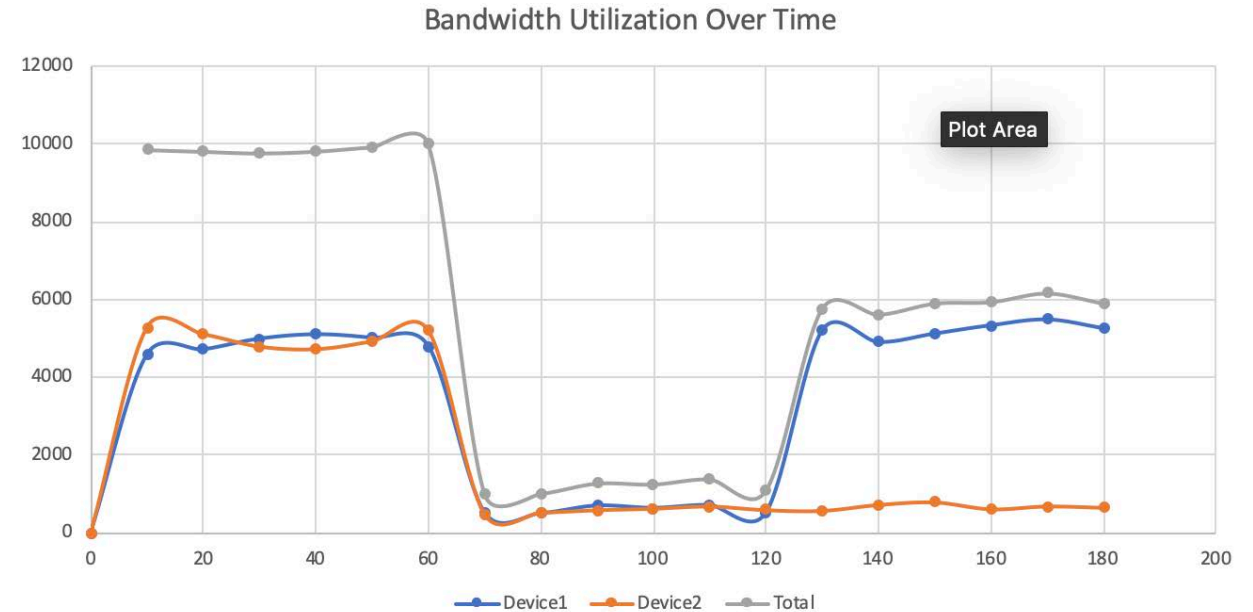
# Future Work



Explore other high-reach sensitive data applications
- ❑ Smart homes
- ❑ Smart healthcare/hospitals system
- ❑ Smart cities (e.g., transportation traffic management)

Large-Scale Distributed ExtremeDataHub Network
- ❑ Networked system of multiple devices that are either geographically located within the same network or distributed across multiple network environments.
- ❑ Application trying to access data should not be aware about the physical location of the data and multiple ExtremeDataHub devices should seamlessly inter-connect together to enable accessing to the required data regardless of the physical location of these data.
- ❑ Asymmetric, unpredictable link characteristics, and non transitive reachability

ExtremeDataHub API
- ❑ To be utilized by application developers.
- ❑ During low battery level, application developer could configure the app to drop less useful functional flows (advertising data).
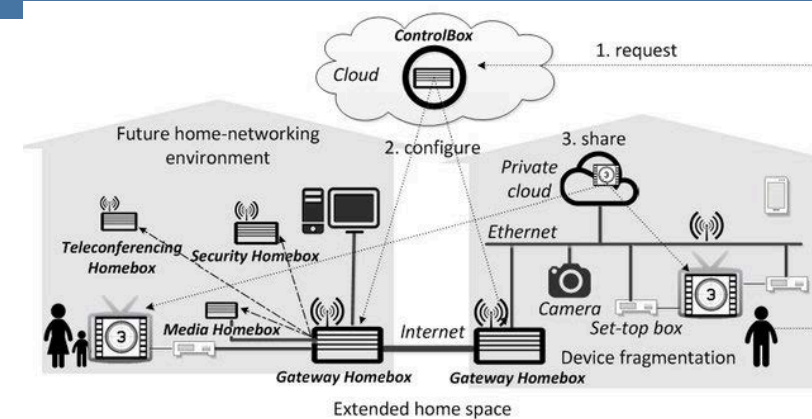
# Joint Students and Education Activities

- Hana Pasandi, female PhD student at VCU, visited Prof. Hamed's research group at Imperial College London for one week in Spring 2019.
  - Got introduced to PhD students of Prof. Hamed's group. Shared their research activities that helped the VCU PhD to refine the objectives of her PhD research direction.
- Had a conference call every two months between PI Nadeem and Prof. Hamed to discuss the progress of the project.
- PI Nadeem and Prof. Hamed arranged to meet during attending a couple of international conferences.
- PI Nadeem allocated another new PhD student to work on specific research activities aligned with project objectives.
- PI Nadeem and Prof. Hamed are continuing discussion on the future direction for this project and we are currently working on a publication paper

# Thank You!

# QUESTIONS

tnadeem@vcu.edu

MuSIClab

https://music.lab.vcu.edu/