

Untangling the Cloud from Edge Computing for IoT

Nabeel Nasir

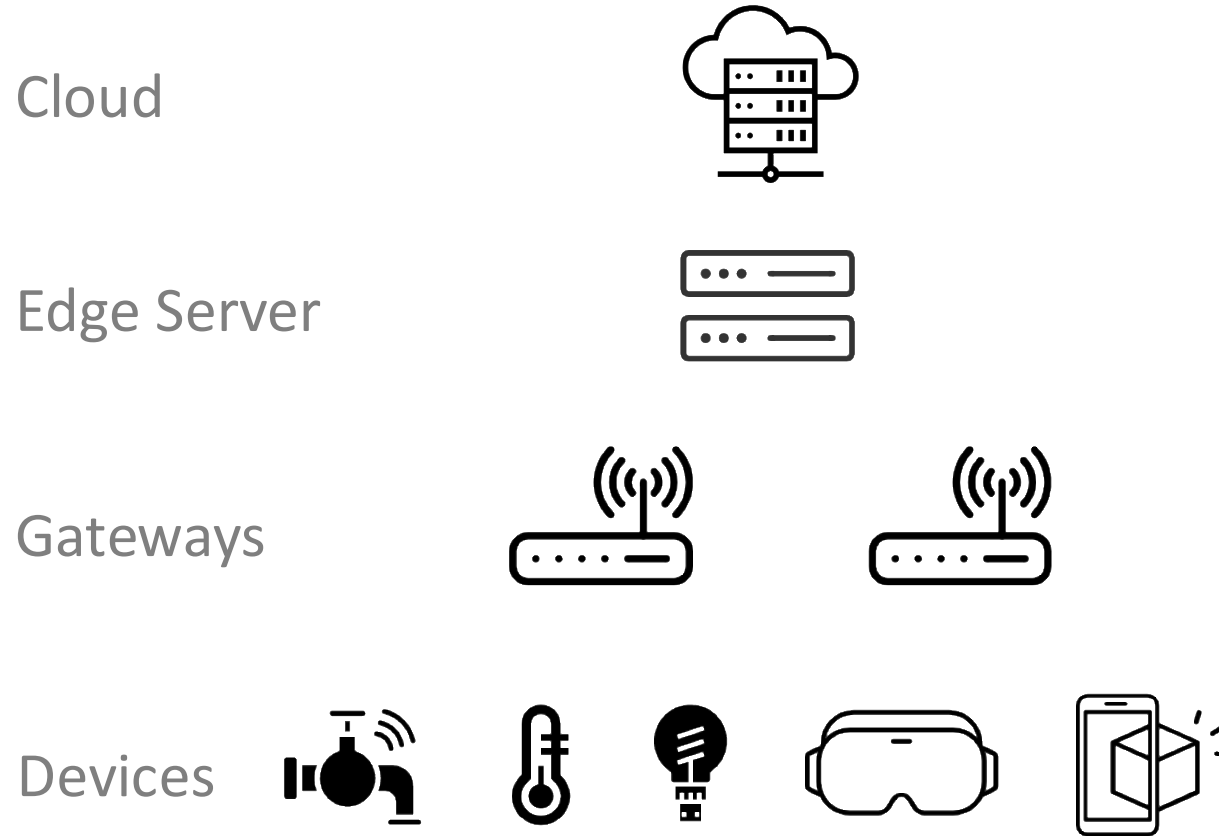
Advised by: Prof. Bradford Campbell

By 2025

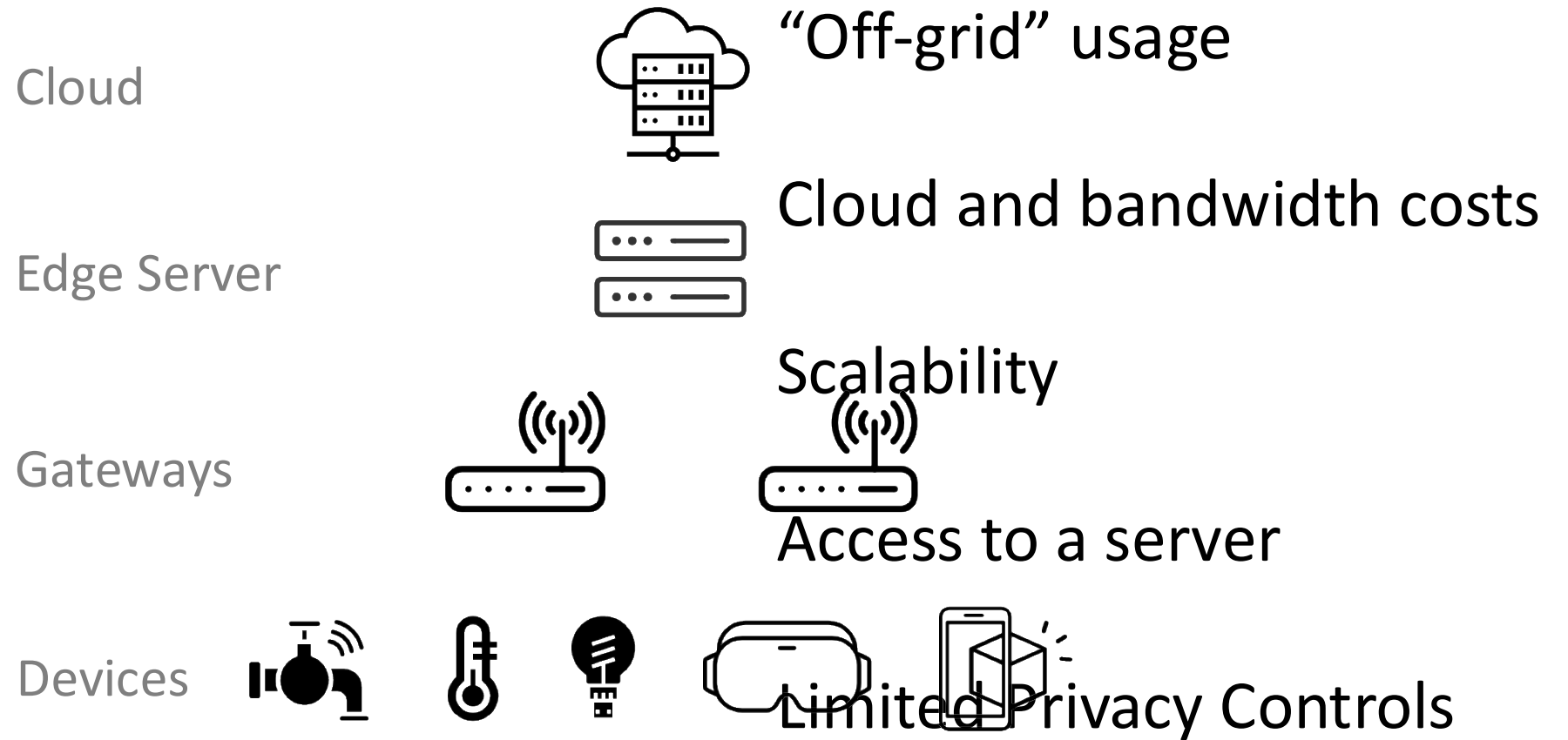
 55 Billion IoT Devices

 73 Zettabytes of Data

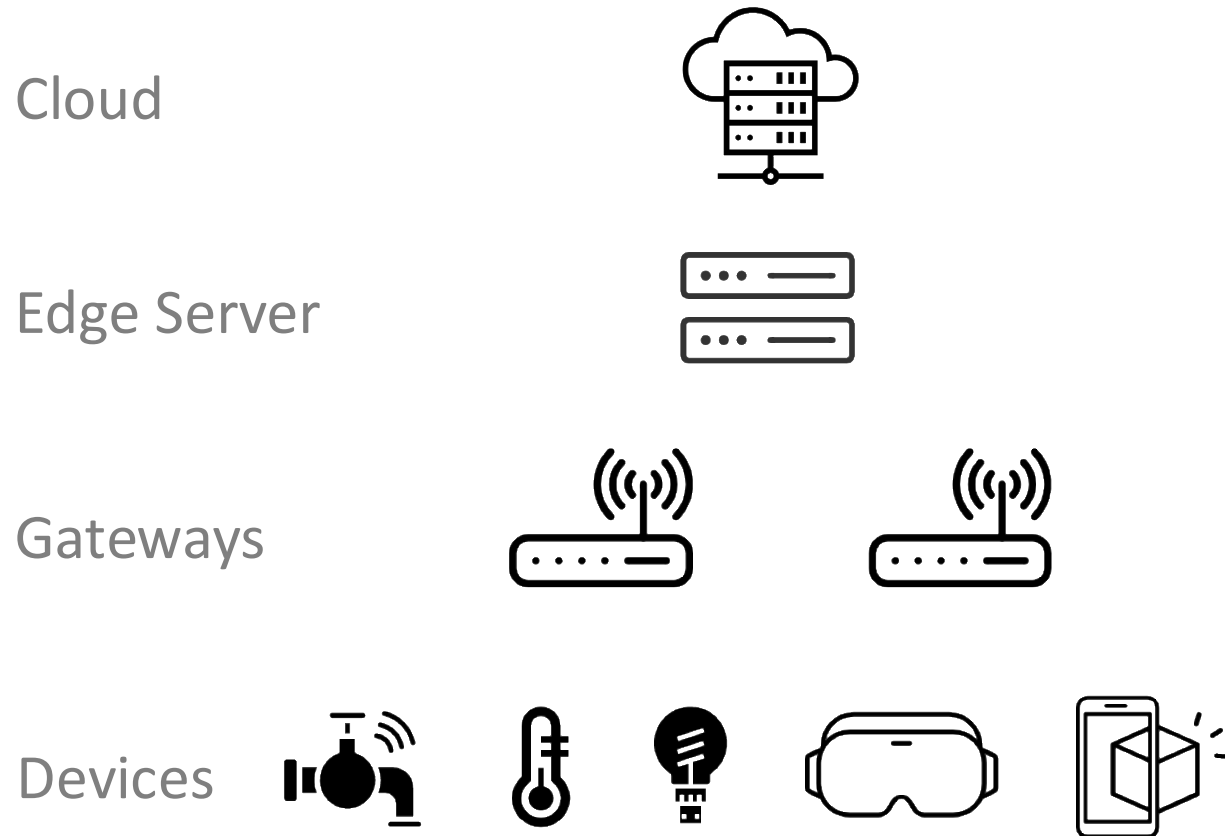
Edge Computing to Handle Massive Influx of IoT Data



Cloud and Edge Server Dependence an Obstacle



Leverage Gateways for Edge Computing



Gateways indispensable for IoT

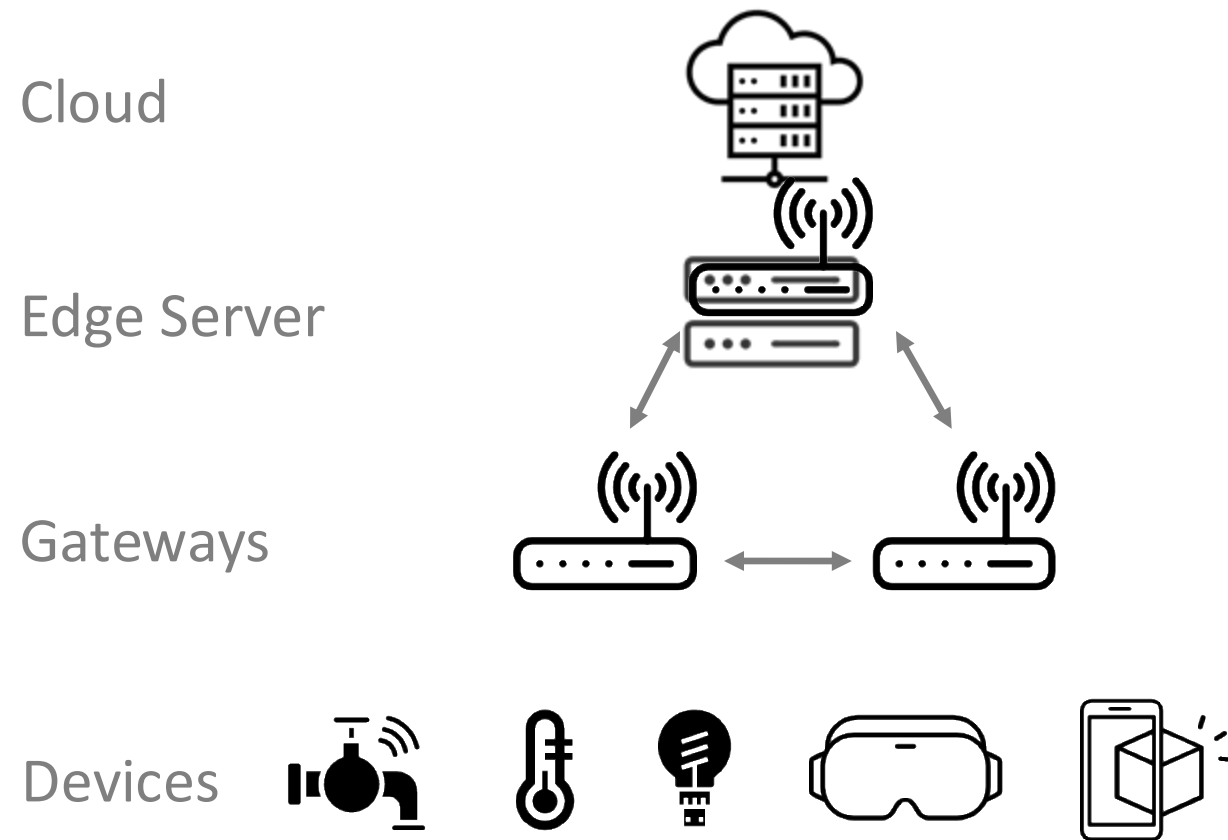
Increasingly capable

Relatively Inexpensive

Key Ideas to Remedy Cloud & Edge Server Issues

1. Distributed Gateway Network instead of an Edge Server

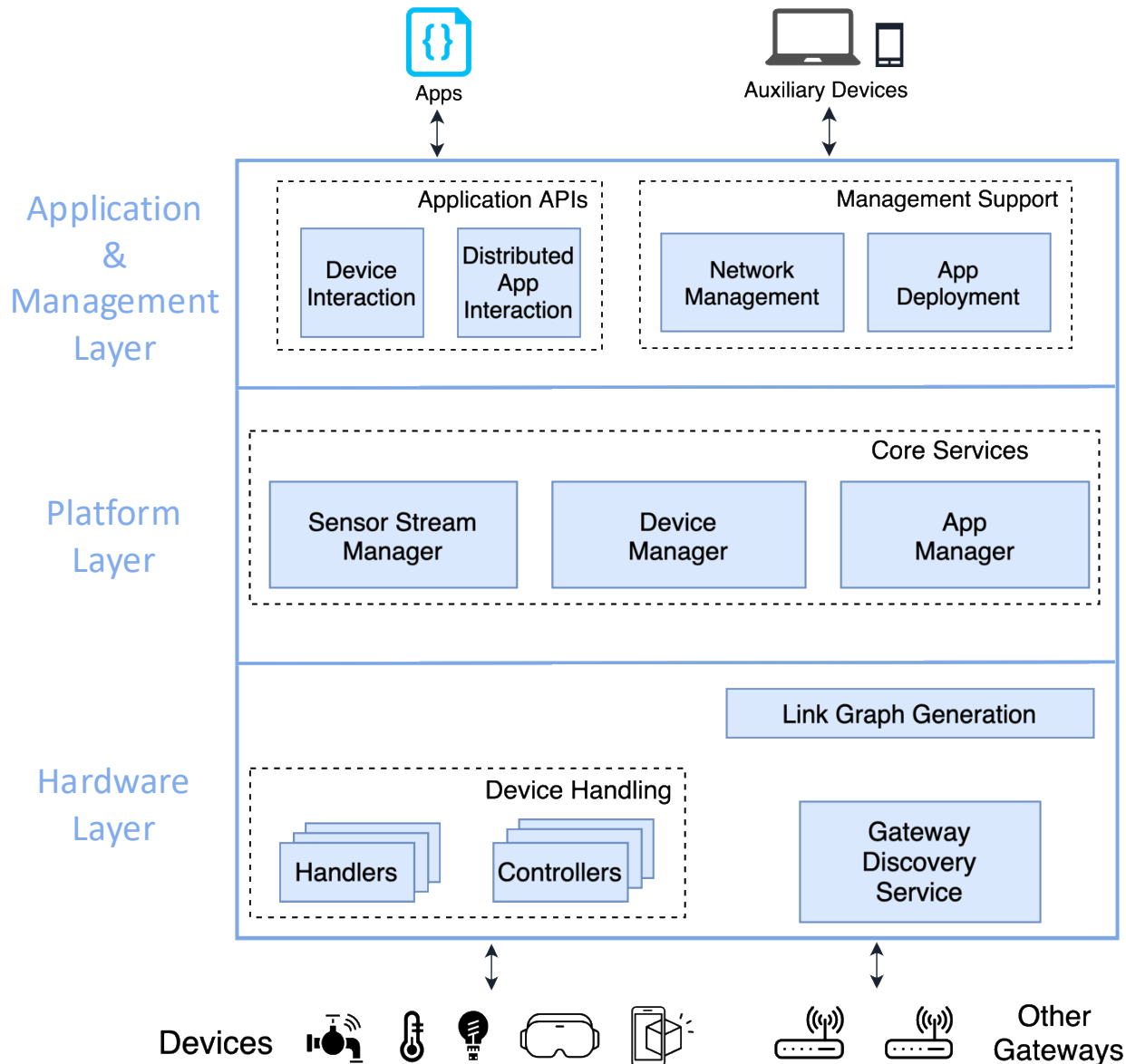
Distributed Gateway Network Instead of Edge Server



Shifting From an Edge Server Poses Some Challenges

1. Overheads with Decentralization
 - Easier application code with simple centralized model
 - No central point to manage platform
2. Scalability
 - Gateway configuration could pose overhead in scaling
3. Resource Constraints
 - Limited resources like storage, GPUs etc.
4. Resilience
 - Gateways not as sturdy as a server machine

Middleware on Gateways to Facilitate this Shift



Gateway discovery for scalability

Supports IoT device handling

Encapsulates network topology

Distributed services to handle devices, applications, and data streams

Device interaction API hides underlying network complexity from apps

Provides remote management

Key Ideas to Remedy Cloud & Edge Server Issues

1. Distributed Gateway Network instead of an Edge Server
2. Specialized Edge Hardware for Elasticity

Edge Gateways Becoming Increasingly Capable and Specialized



Raspberry Pi



NVIDIA Jetson
Xavier NX



Google Coral
with Edge TPU

Substantially cheaper

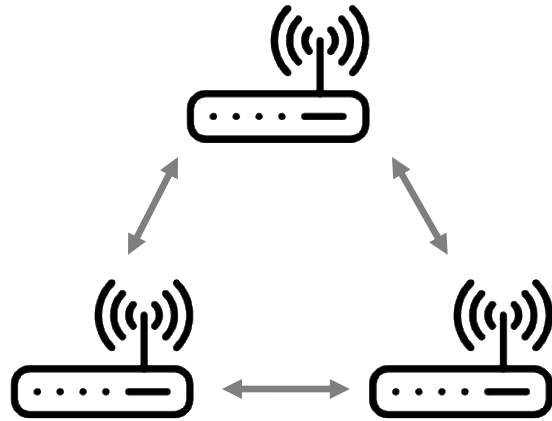
Can match app requirements

Directly interact with IoT devices

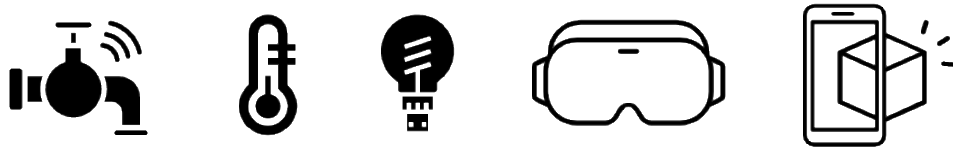
Better deployment flexibility

Utilize Specialized Edge Hardware for Elasticity

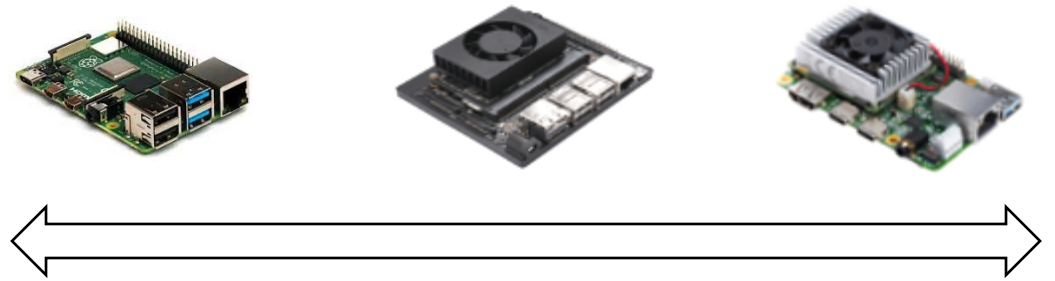
Gateways



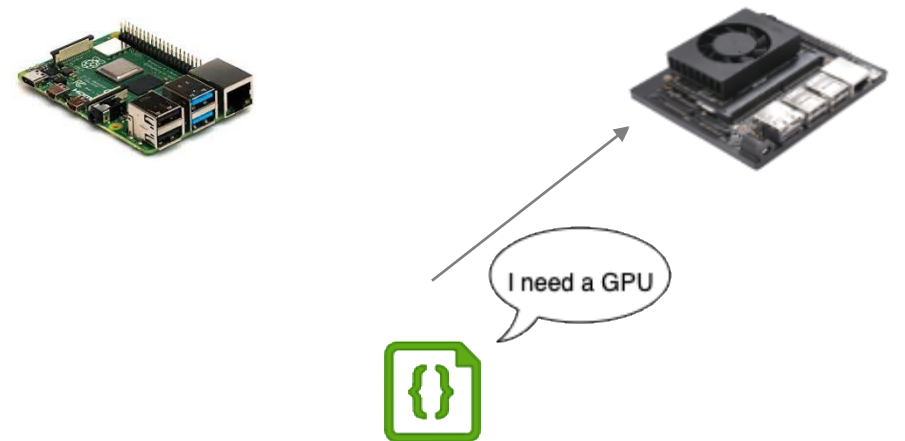
Devices



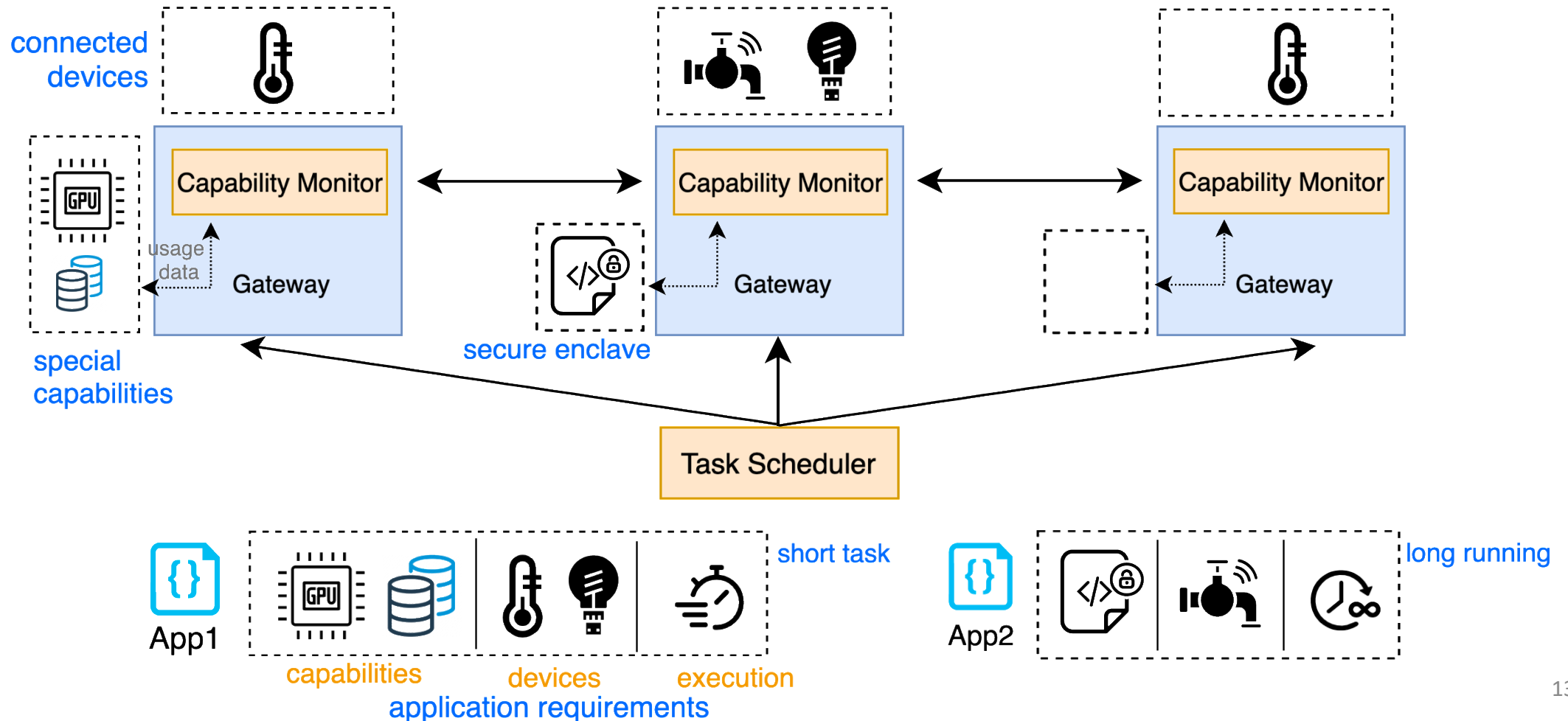
Scale out by adding a new gateway



Schedule based on requirements



Scheduler uses Requirements, Capabilities and Resource Usages



Scheduling the Key Challenge, and Different from Other Domains

- Optimal scheduling algorithm based on
 - Application requirements
 - Gateway capabilities
 - Gateway resource usages
- Different from Task Schedulers in other domains:
 - **Grid Computing, Cluster Computing:** Mostly homogeneous machines
 - **Real-time scheduling:** Requires execution times and deadlines
 - **Function-as-a-Service Platforms:** App's requirements or machine's capabilities not considered

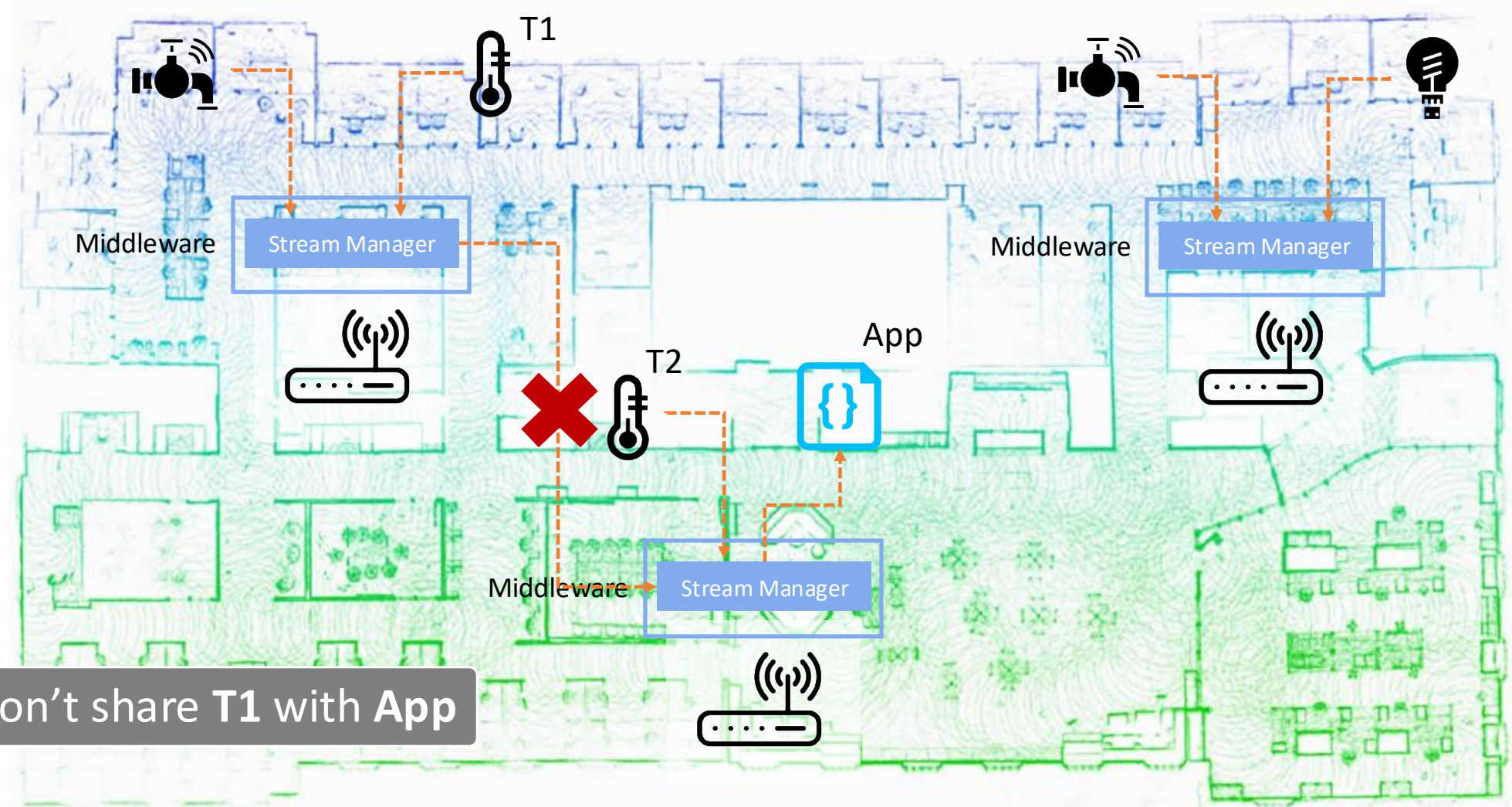
Key Ideas to Remedy Cloud & Edge Server Issues

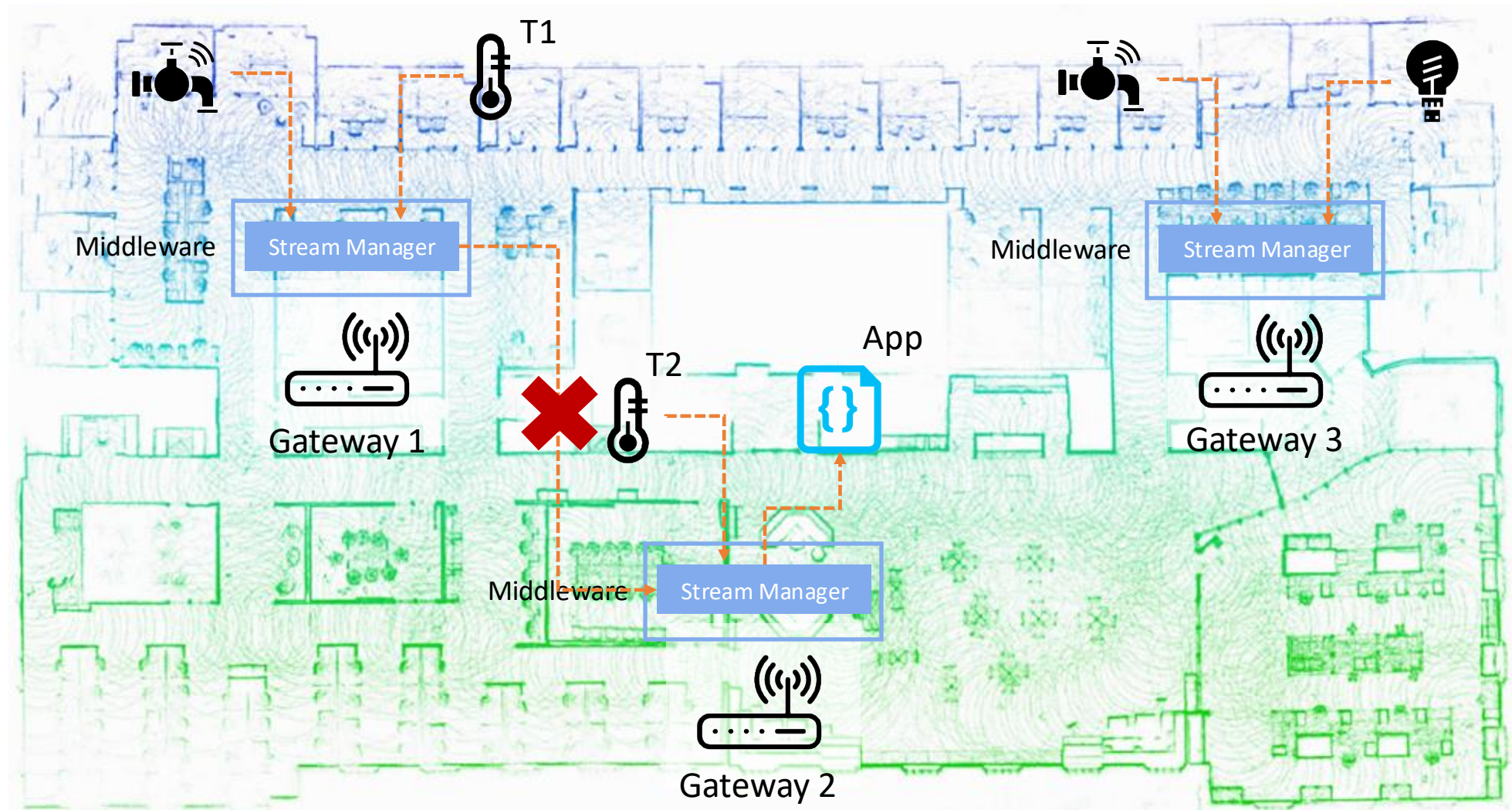
1. Distributed Gateway Network instead of an Edge Server
2. Specialized Edge Hardware for Elasticity
3. Privacy Policies enforced on Gateways to empower Users

People in Shared Spaces have No Idea About Their Sensed Data



Utilize Existing Sensor Stream Service to Enforce Policies





User-driven Privacy Policies to Empower People

- Users can specify policies like:
 - Allow access to **occupancy data** to **all apps**, but only from **9AM – 5PM**
 - Share **occupancy data** to an **emergency evacuation app**, but not to **room scheduling app**
- Challenge:
 - Simple user interface to convey this complex information
- Novelty:
 - Some privacy policy works exist, but novel way of policy enforcement

Conclusion

- Democratize Edge Computing by untangling it from cloud and edge server dependencies
- Utilize inexpensive gateway network to execute applications
- Empower users by providing transparency and control over data

Thank you!

Questions?

- You can find me on:
 - Website: <http://www.virginia.edu/~nn5rh>
 - Email: nabeeln@virginia.edu
 - LinkedIn: <https://www.linkedin.com/in/nabeel-nasir/>
- I am seeking summer internship opportunities!