Cloud and Edge Computing:
Why sustainability matters

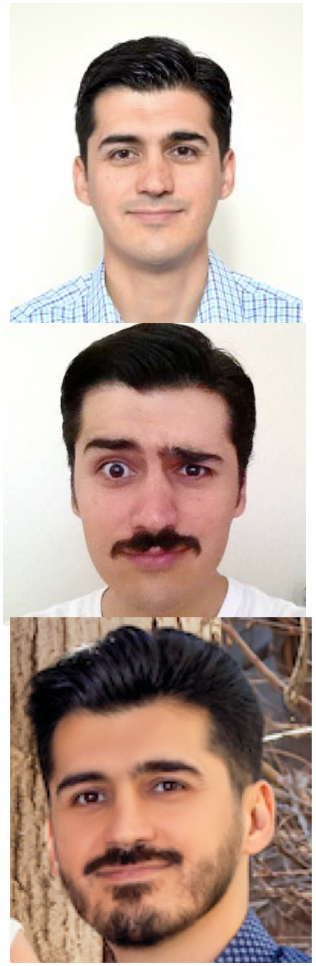
Adel N. Toosi

***Dis**tributed Systems and **Net**work Applications Laboratory (DisNet lab.)*
Department of Software Systems and Cybersecurity,
Faculty of Information Technology
Monash University

Email: adel.n.toosi@monash.edu
Homepage: https://adelnadjarantoosi.info/

Biography

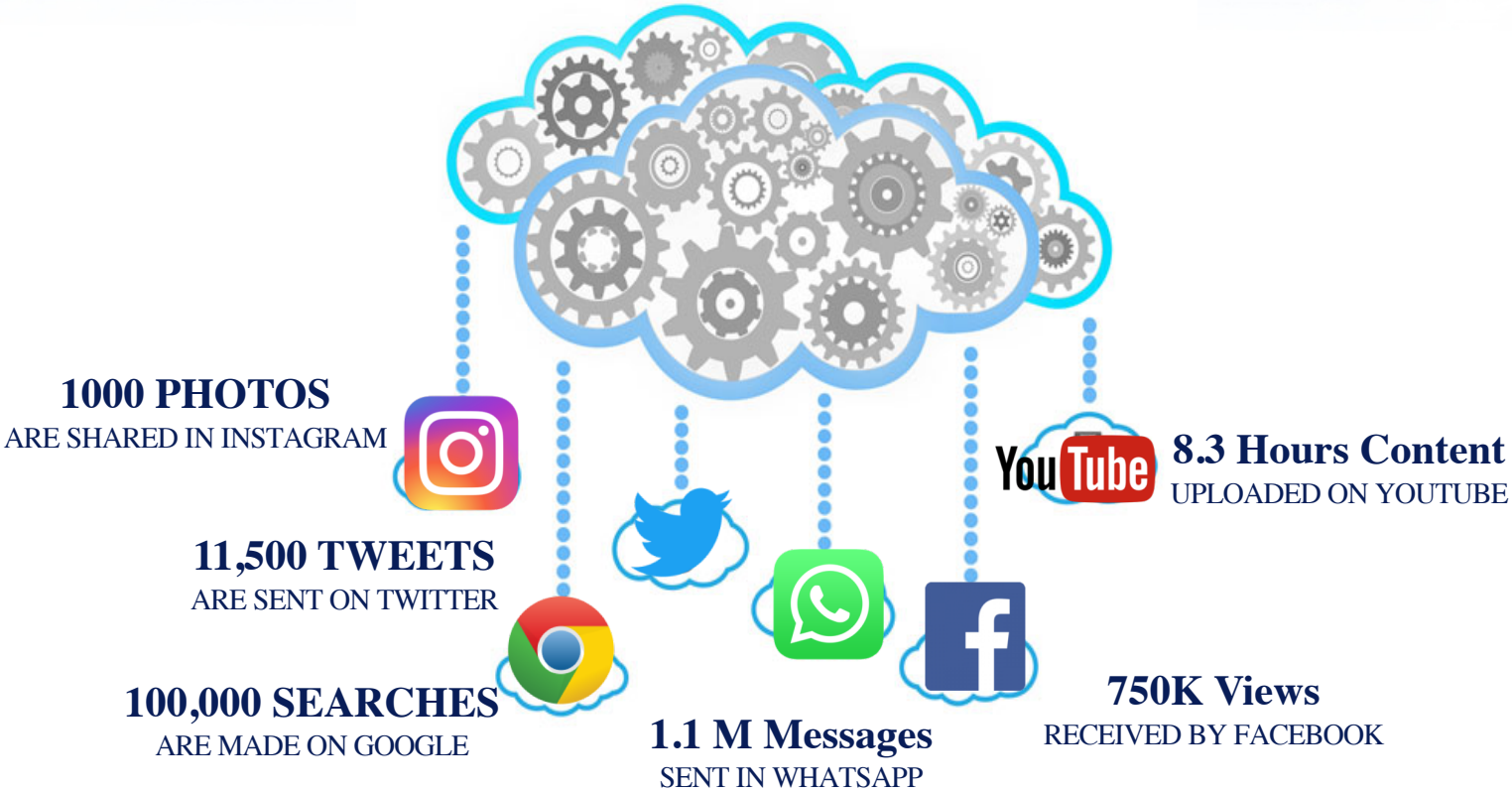
- Lecturer (Assistant Professor), Monash University, 2018-
- Postdoctoral Research Fellow, University of Melbourne, 2015-2018
- PhD, Computer Science and Software Engineering, 2015
  - Thesis: "*On the Economics of Infrastructure as a Service Cloud Providers: Pricing, Markets, and Profit Maximization*"
- Research Interests
  - Cloud/Fog/Edge Computing, Software-Defined Networking (SDN), Sustainable IT, Energy Efficiency and Green Computing
  - Focused on Resource Provisioning and Scheduling in Distributed Systems
- Publications
  - Over **50** publications with >3300 citations (*Src: Google Scholar*)



Outline

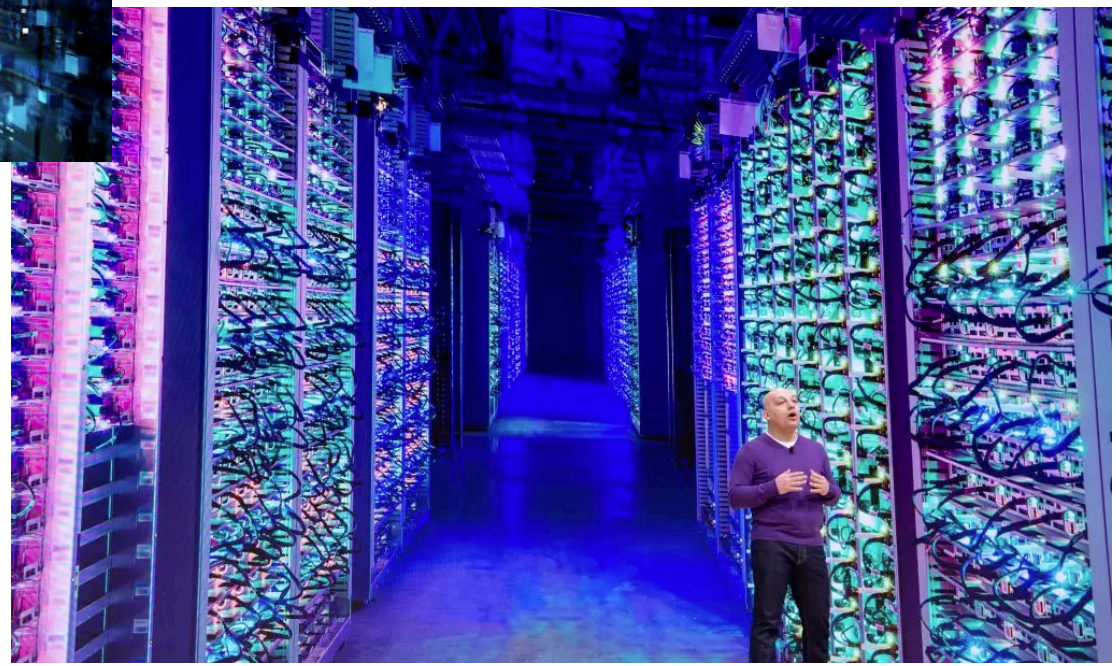
- Introduction
  - High energy consumption of IT sector
- Using renewable energy for sustainable IT
- Challenges of using renewable
- Example solutions for data centers
  - Geographical load balancing
  - Load shaping through resource scheduling
- Edge computing
- Challenges of sustainable edge computing
- Examples: Con-pi, Performance Evaluation, Wattedge, faasHouse
- Summary

Online Activities Every Second



Data Centers in Clouds

https://www.crn.com/news/data-center/alibaba-aims-to-speed-up-covid-19-recovery-with-28-billion-data-center-investment



https://time.com/5814276/google-data-centers-water/

Power Hungry Clouds

- Data centers used for hosting cloud applications consume large amounts of electricity
  - High **operational cost** for the cloud providers
  - High **carbon footprint** on the environment
- Data centers consume over 1000 W/m2 of power
  - more than 10 times of that required by a typical commercial office space - *United States Data Center Energy Usage Report*
- In 2020, US data centers alone consumed 73 billion kilowatt-hours of electricity
  - = Two-year power consumption of all households in **New York**
  - = 100 to 200 times of electricity consumed by standard office space
  - This is projected to nearly **50 million tons of carbon** pollution per annum in 2020.

– *Source: US Natural Resources Defense Council (NRDC)*

Renewable Energy

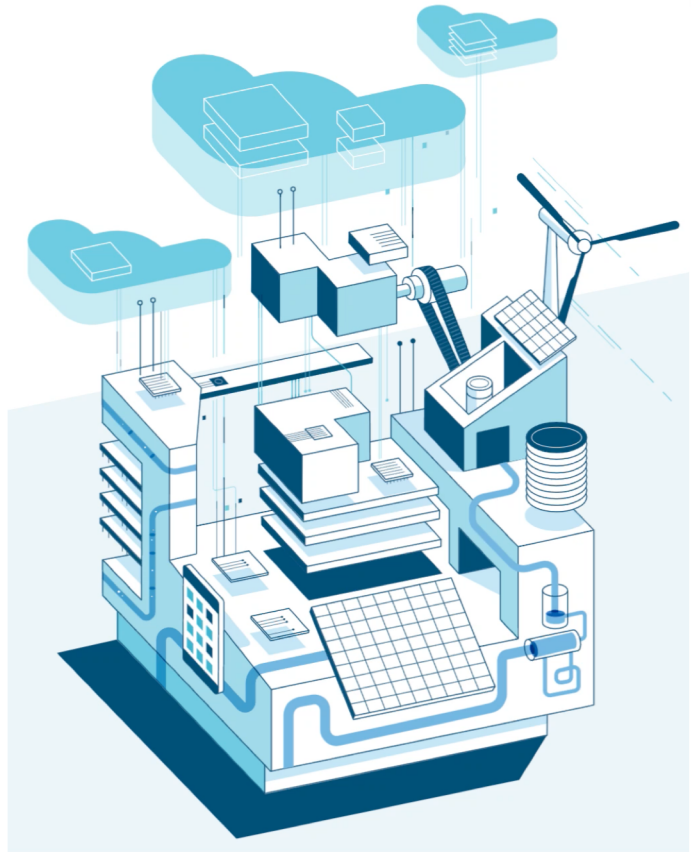
- Cloud providers aims
  - Reducing energy consumption
  - Dependence on brown energy
- Using renewable energy
  - On-site green power generation
  - Google, Microsoft and Amazon

"Amazon Web Services (AWS) is aiming at 100% renewable energy by 2025"

Source: https://aws.amazon.com/about-aws/sustainability/

Challenges

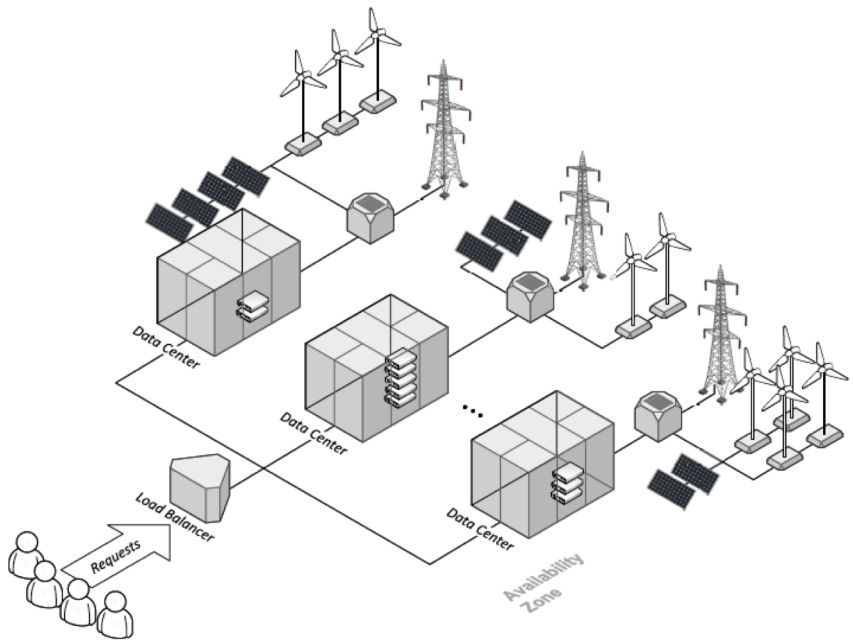
- ***Intermittency*** and ***unpredictability*** of renewable energy sources
  - Variability of Wind and Solar, etc.
- Objectives
  - Minimizing brown energy usage
  - Maximizing renewable energy utilization



https://sustainability.aboutamazon.com/

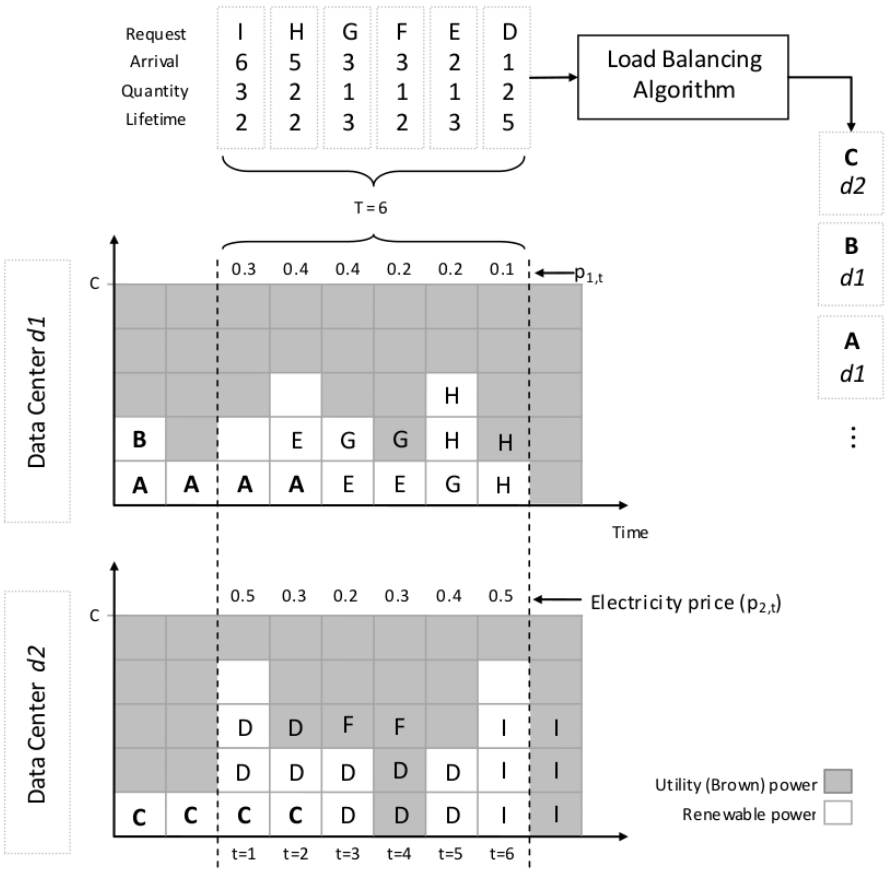
1. Geographical Load Balancing (GLB)

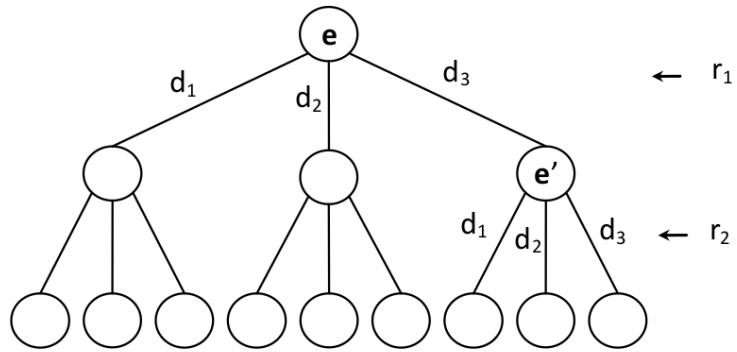
- Geographical load balancing (GLB) potentials:
  - **Follow-the-renewables**
- GLB approach benefits cloud providers but it raises an interesting, and challenging question:



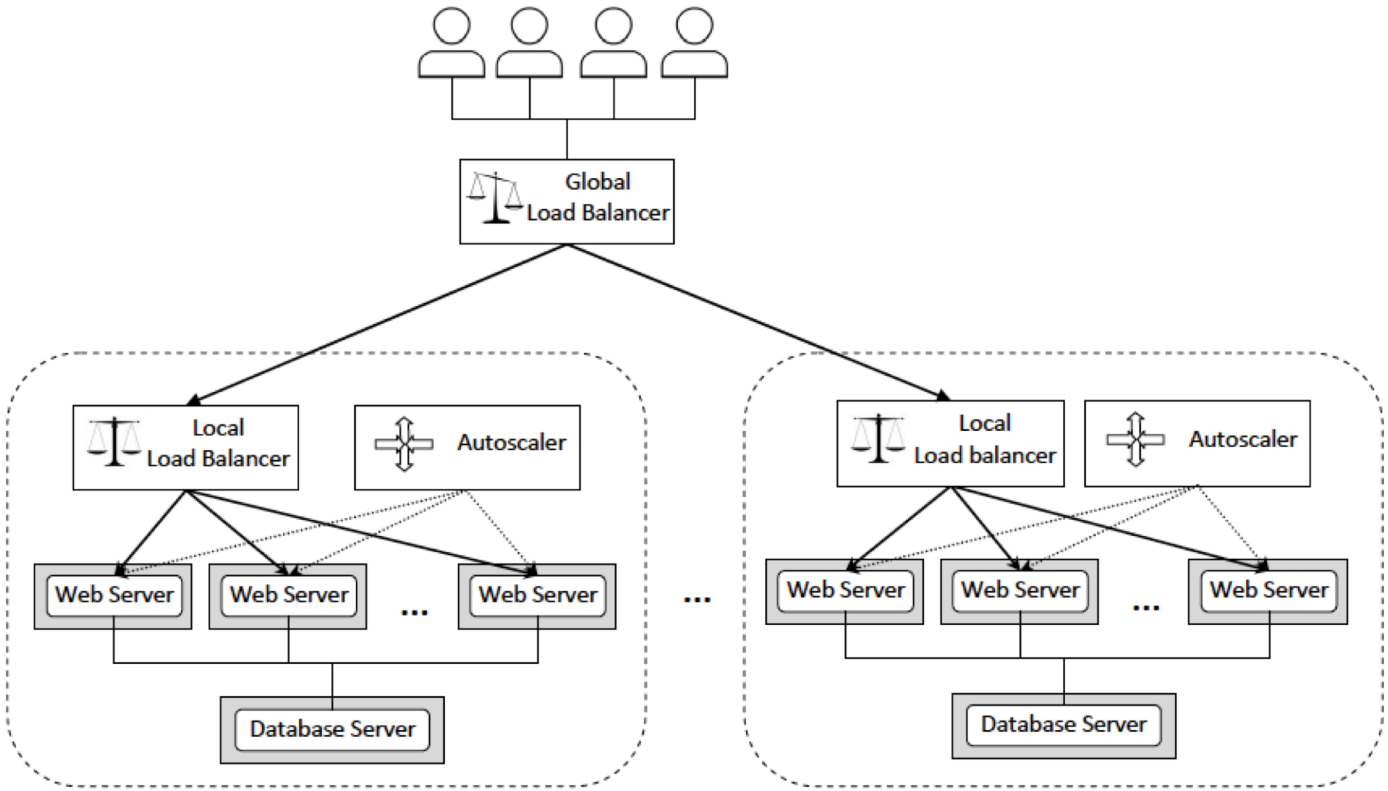
Adel Nadjaran Toosi, Chenhao Qu, Marcos Dias de Assuncao, and Rajkumar Buyya, **Renewable-aware Geographical Load Balancing of Web Applications for Sustainable Data Centers**, *Journal of Network and Computer Applications (JNCA)*, Vol. 83, pp. 155-168, Apr. 2017.

Example: Offline GLB Problem

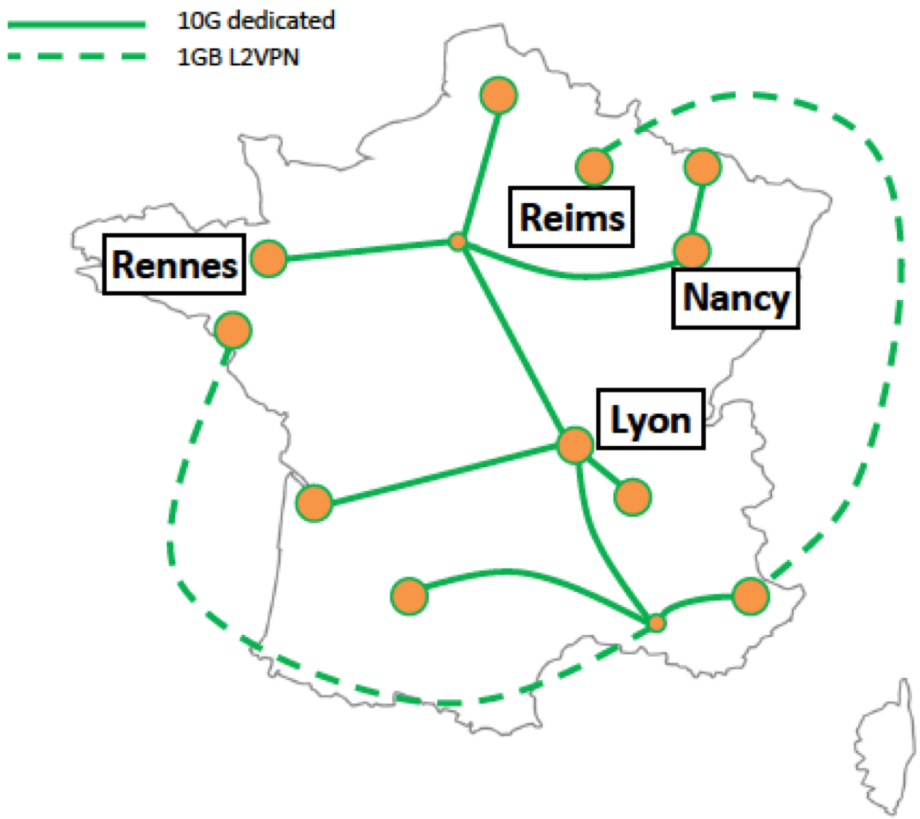




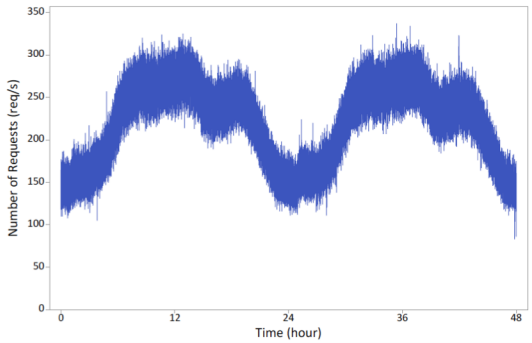
Overall System Architecture



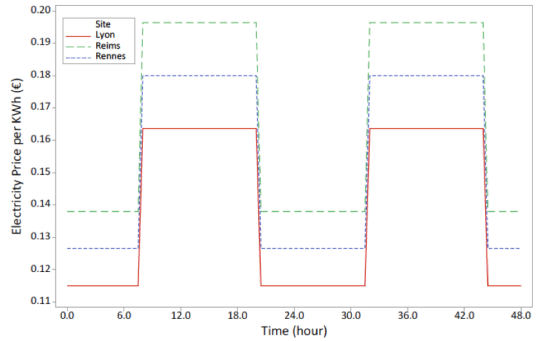
Grid'5000 Testbed



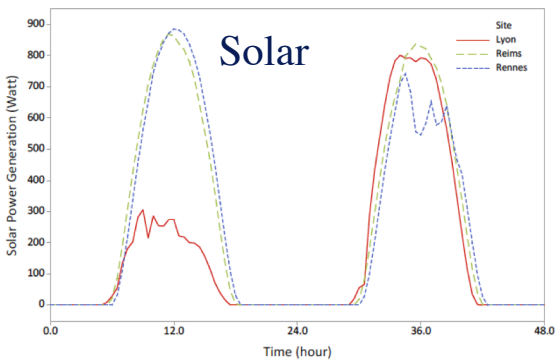
Workload Traces

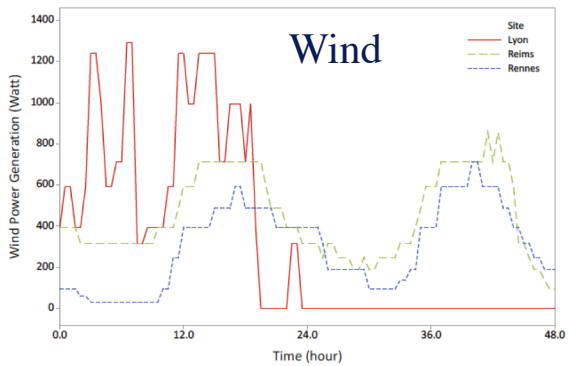


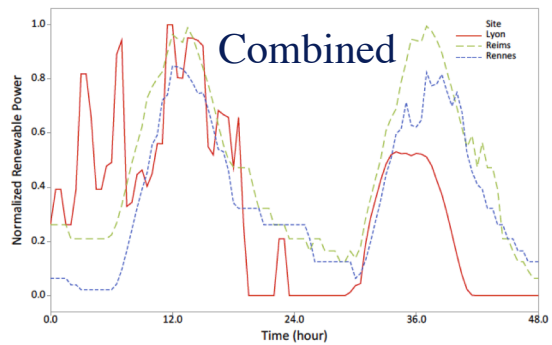
The 5% of English Wikipedia Requests for 19th and 20th of September 2007



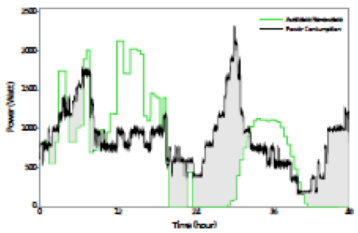
Electricity Prices



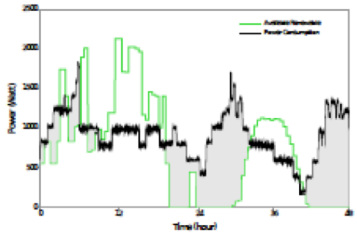




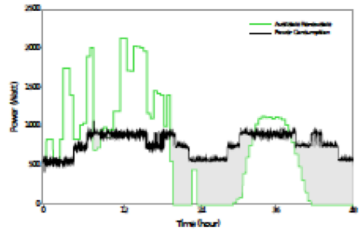
Results



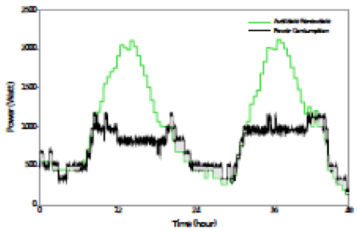
(a) Lyon



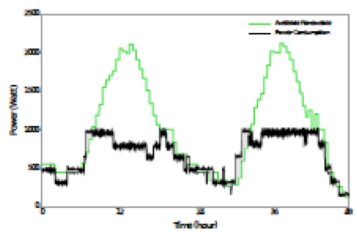
(a) Lyon



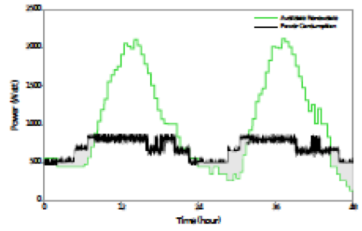
(a) Lyon



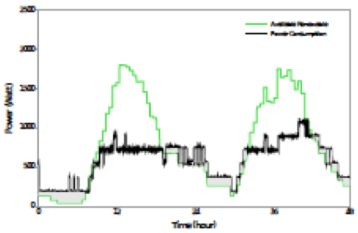
(b) Reims



(b) Reims

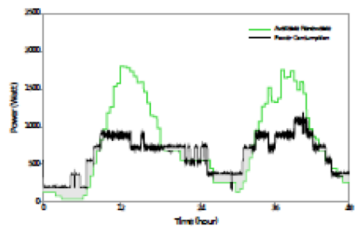


(b) Reims



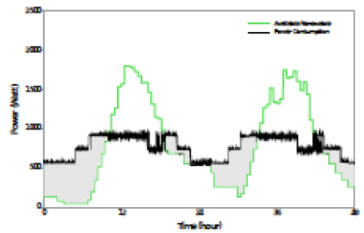
(c) Rennes

GreenLB



(c) Rennes

Capping



(c) Rennes

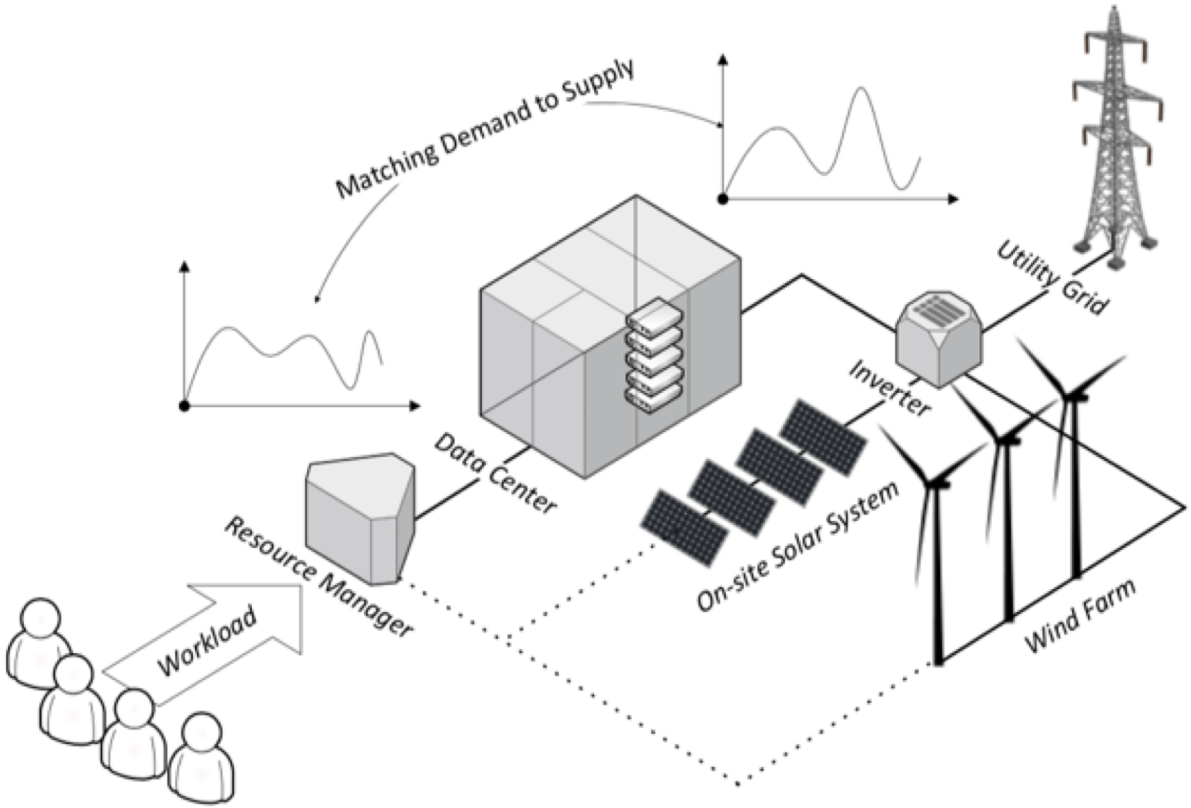
Round Robin

Results

Site	Metric	RR	Capping	GreenLB
Lyon	Power Consumption (kWh)	36.3	42.9	**41.2**
	Brown Consumption (kWh)	13.3	19.0	**16.9**
	Cost (€)	1.71	2.31	**2.01**
Reims	Power Consumption (kWh)	32.5		
	Brown Consumption (kWh)	2.1		
	Cost (€)	0.42		
Rennes	Power Consumption (kWh)	36.4		
	Brown Consumption (kWh)	9.3	2.9	
	Cost (€)	1.23	0.39	**0.35**
Total	**Power Consumption (kWh)**	**105**	**105**	**105**
	Brown Consumption (kWh)	**25.7**	**23.0**	**21.4**
	Cost (€)	**3.36**	**2.85**	**2.63**

Brown Energy:
17% and 7%
Cost Saving:
22% and 8%

2. Renewable-aware resource management and scheduling



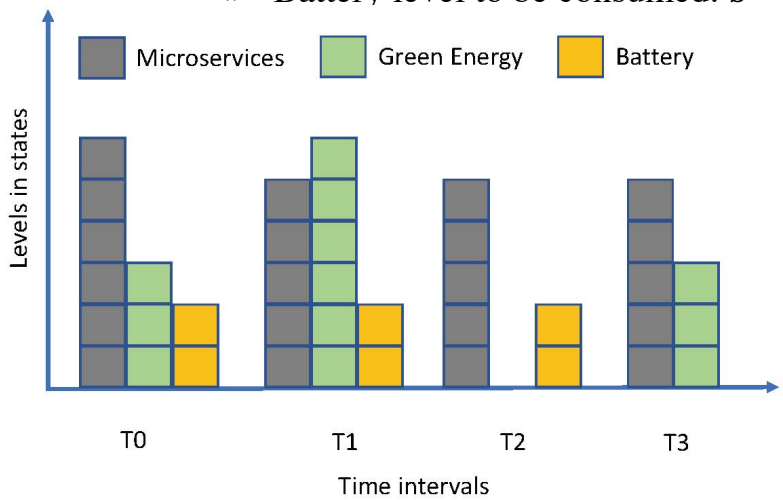
Minxian Xu, Adel N. Toosi, Behrooz Bahrani, Reza Razzaghi, and Martin Singh, **Optimized Renewable Energy Use in Green Cloud Data Centers**, *In Proceedings of the 17th International Conference on Service-Oriented Computing (ICSOC'19)*, 28-31, 2019. Toulouse, France, pp. 314-330, doi:10.1007/978-3-030-33702-5_24.

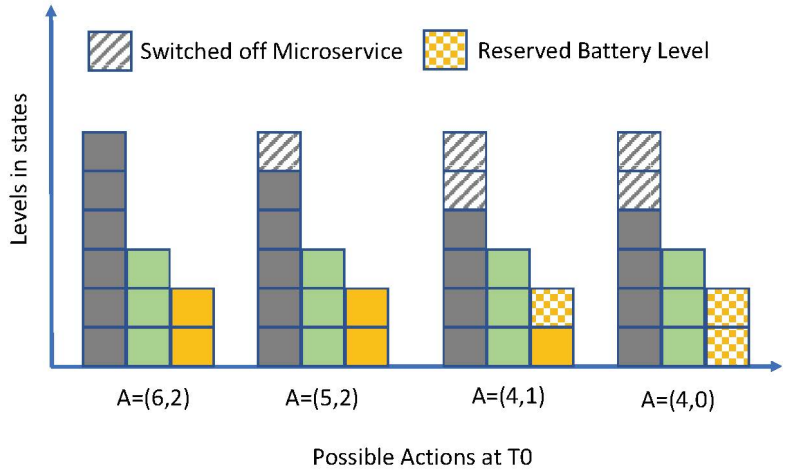
Objectives and Solution

- Aim
  - Maximizing the number of executed microservices while minimizing brown energy (energy from polluting sources) usage
- Solutions
  - An approach based on Markov Decision Process (MDP) to dynamically switch on/off microservices and schedule battery usage to achieve the optimized results

States and Actions

- Example: Possible States and Actions
  - States:
    - » Demanded microservices: $W(t)$
    - » Available renewable energy: $G(t)$
    - » Level of battery: $B(t)$
  - Actions: $A(a, b)$
    - » Number of executed microservices: a
    - » Battery level to be consumed: b



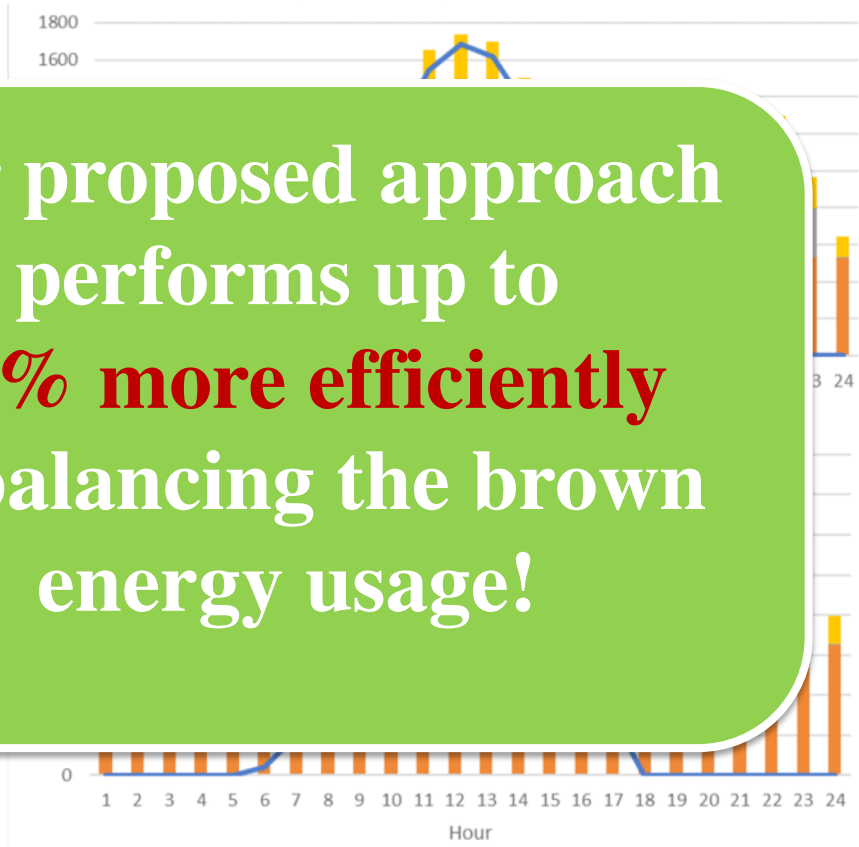


Testbed / Prototype

Some Results

Our Approa

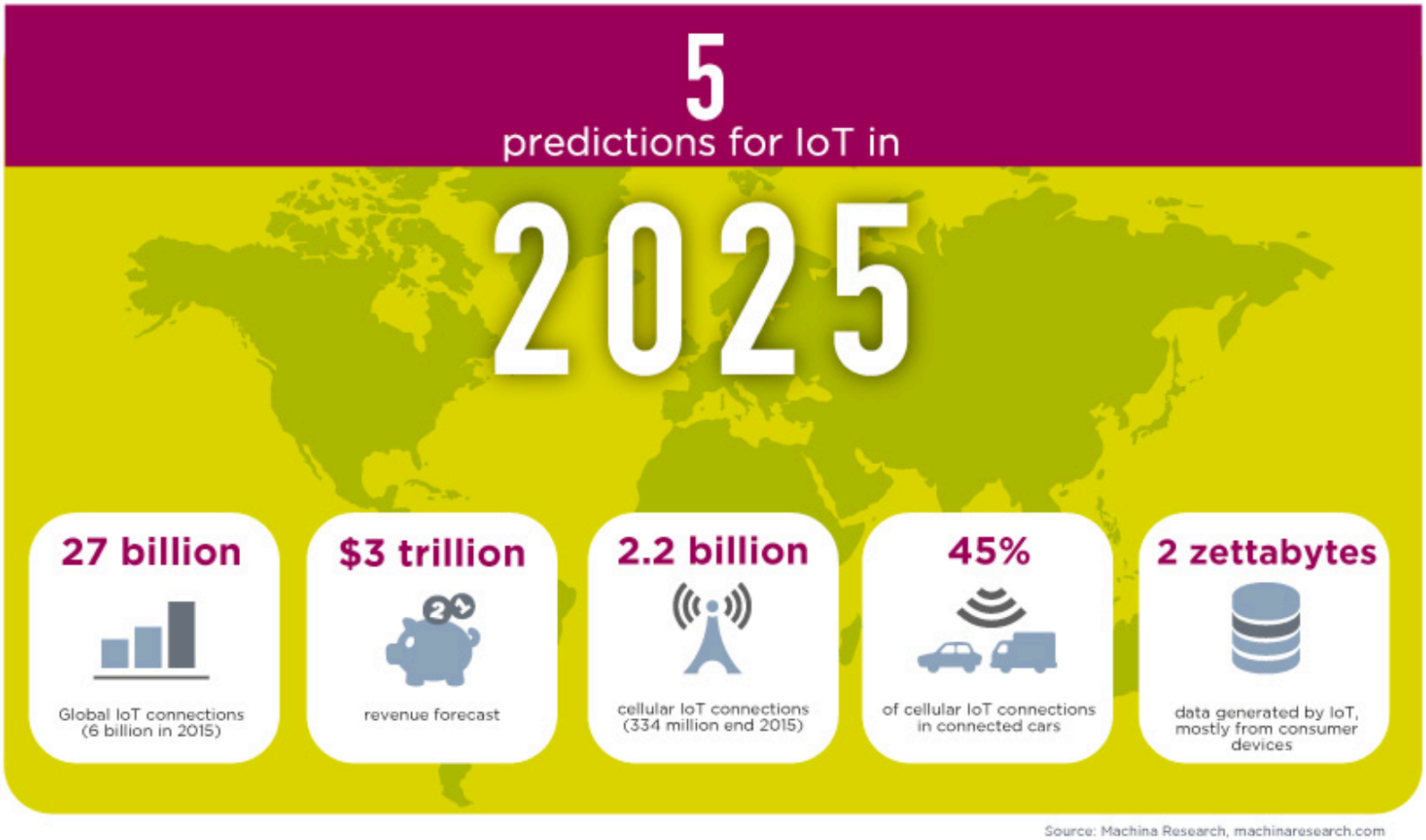
Benchmar



Our proposed approach performs up to 30% more efficiently in balancing the brown energy usage!

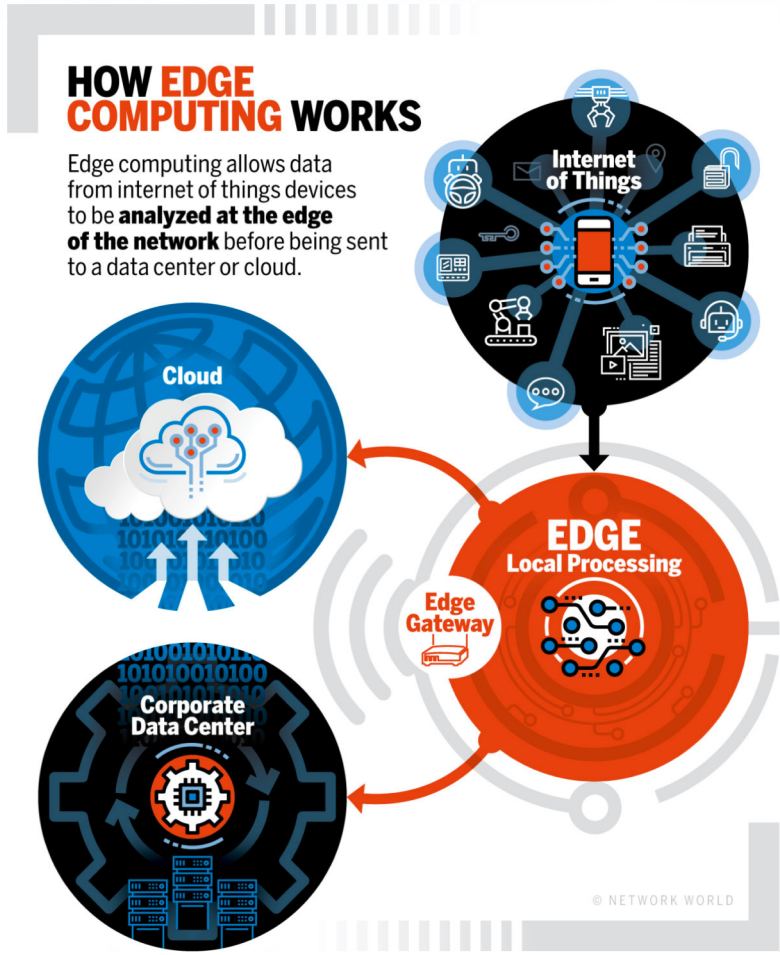
System Prototype

Digital Transformation and IoT



Edge Computing

- cloud data centers reside at a multi-hop distance from the sensors and devices
  - data propagation and transmission can cause significant delays and affect the responsiveness of real-time applications such as autonomous vehicles.
- Edge computing:
  - a **distributed computing** paradigm that brings **computation** and data storage closer to the edge of the network, where it is needed.
- Its ultimate goal is to minimize deficiencies in data throughput and the latency for smart applications



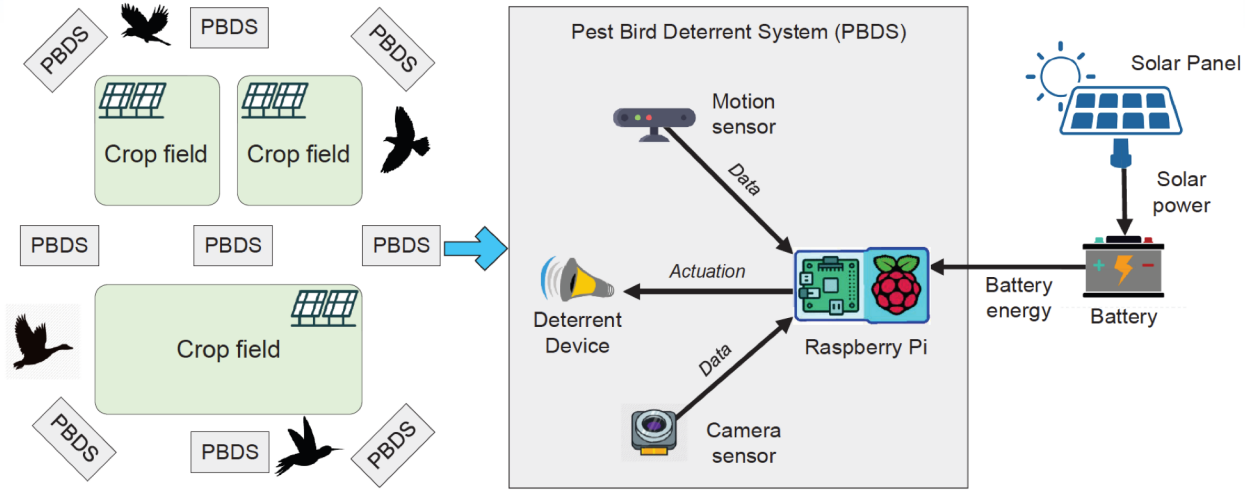
Challenges of Edge Computing

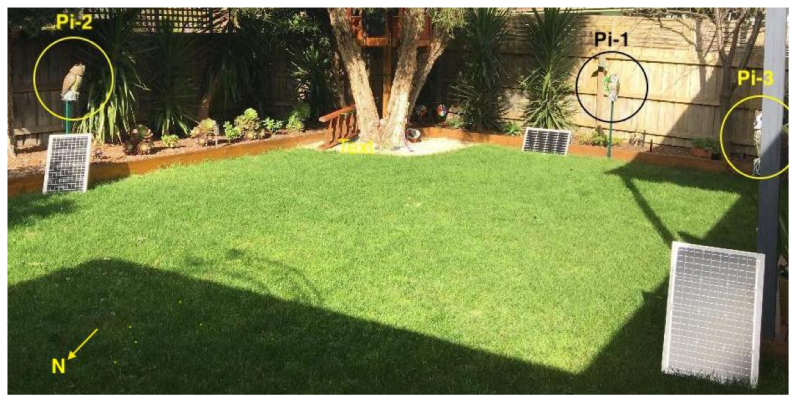
- In remote area application, e.g., smart farming and forestry
  - electricity arrangements to integrate sensory/actuation systems into the edge computing infrastructure is tedious and costly.
- Solution:
  - Battery and energy harvesting (e.g., solar panels)
- Challenge:
  - Edge devices rely on renewable energy sources that are subject to **variability** which can create an imbalance in their **operational availability**.
- Solution:
  - Resource Sharing and task offloading

Con-Pi: Self-Sustained Edge Computing Framework

- An automated bird repellent system
- Edge computing platform with a network of devices (Raspberry Pis)
- Fully off-grid by the deployment of Battery-Solar-powered devices.

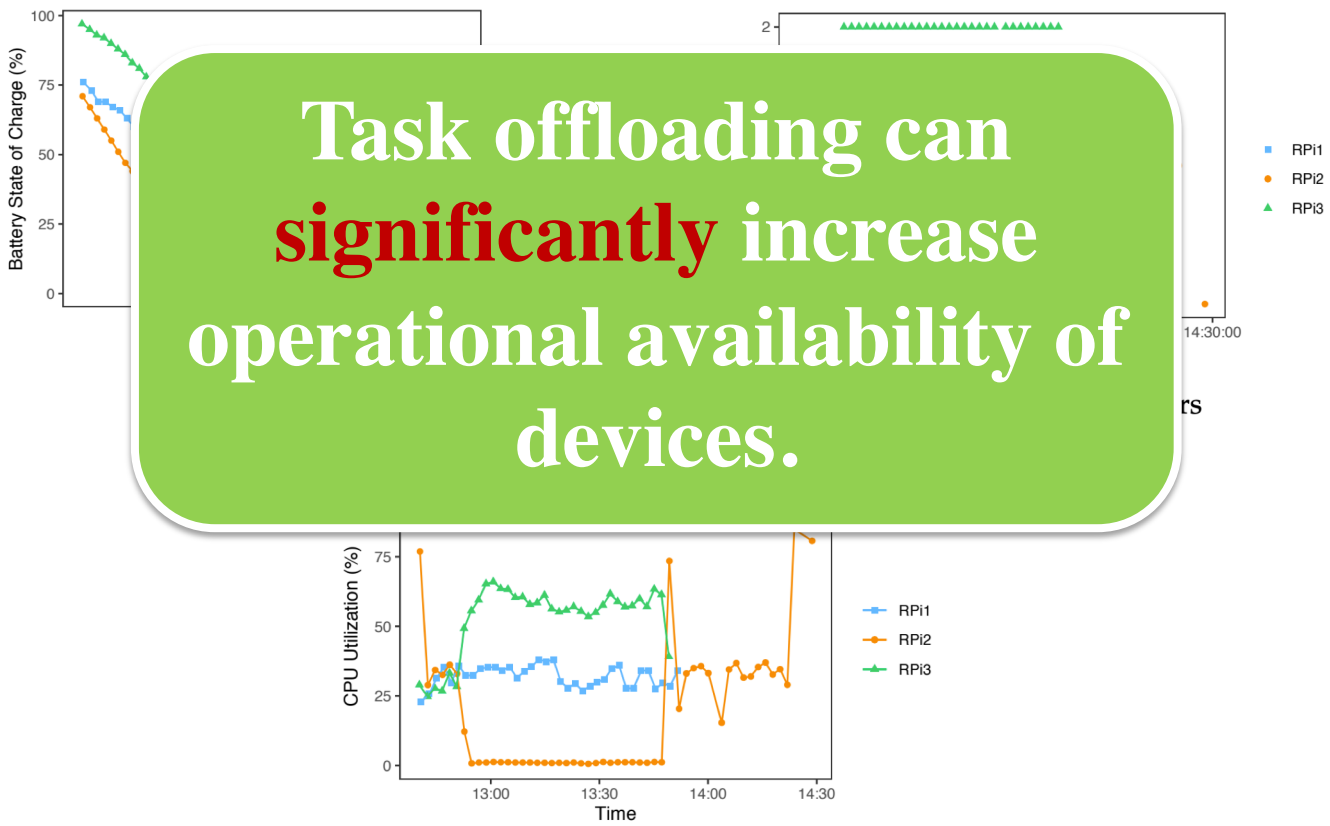
3. Con-Pi: Self-Sustained Edge Computing Framework





R. Mahmud and A. N. Toosi, "**Con-Pi: A Distributed Container-based Edge and Fog Computing Framework**," in *IEEE Internet of Things Journal*, doi: 10.1109/JIOT.2021.3103053.

Some Results



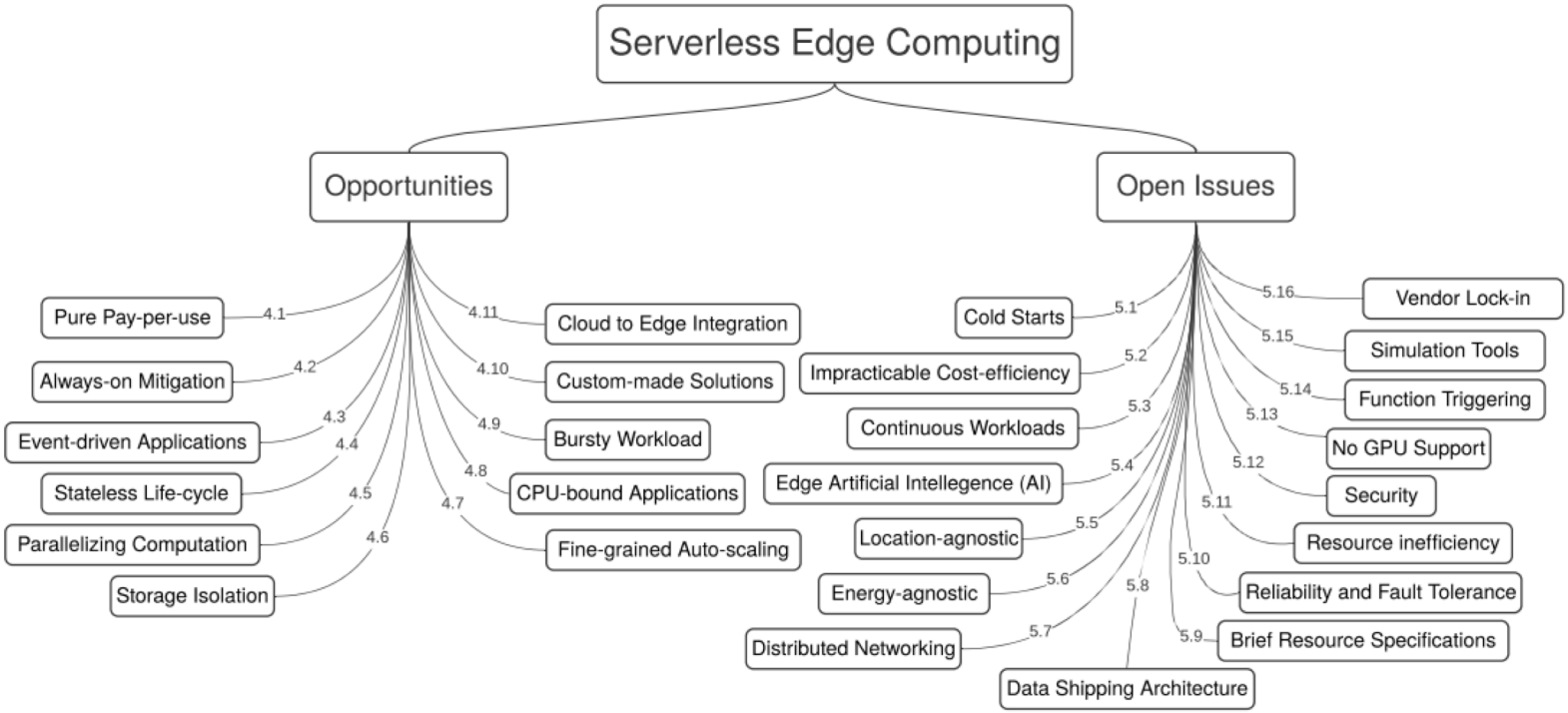
Task offloading can significantly increase operational availability of devices.

(b) CPU utilization

Sustainable Serverless Edge Computing

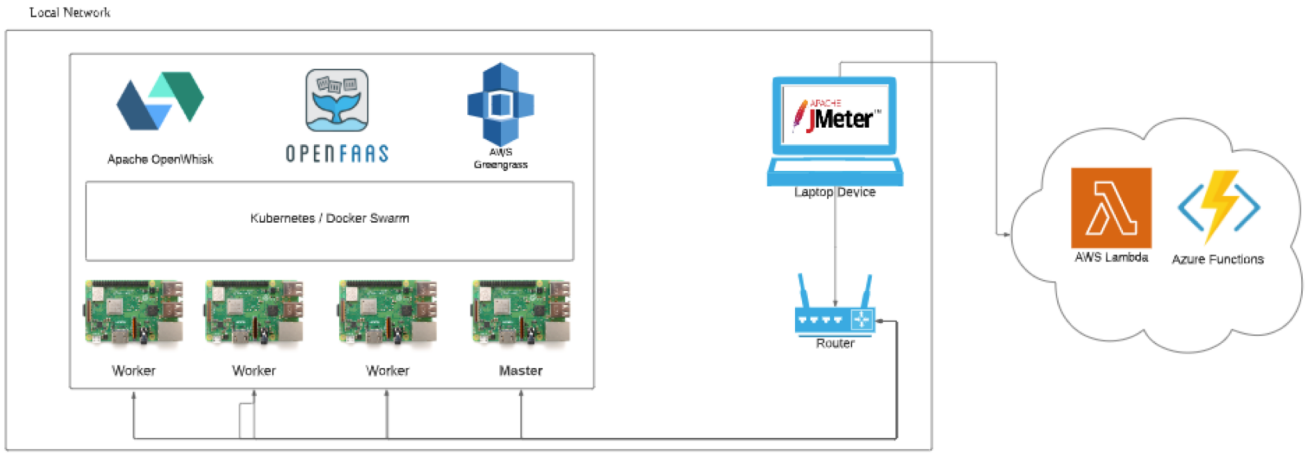
- What is Serverless?
  - Serverless enables us to build applications and services without thinking about the underlying servers.
  - In short, Serverless is an abstraction which hide details of the **underlying infrastructure** and resource management
  - To enable a model in which developers only focus on pure application code (application logic) where codes run in response to different events.
  - The idea was first developed by AWS and a service called **Lambda**, the “poster child” for Serverless, is the most widely known Serverless Service in form of the Function as a Service (FaaS).

Our vision on Serverless Edge Computing



Mohammad Sadegh Aslanpour, Adel N. Toosi, Claudio Cicconetti, Bahman Javadi, Peter Sbarski, Davide Taibi, Marcos Assuncao, Sukhpal Gill, Raj Gaire, Schahram Dustdar, **Serverless Edge Computing: Vision and Challenges**, *In Australasian Computer Science Week Multiconference (ACSW'21)*, article no 10, isbn = 9781450389563, Dunedin, New Zealand, 2021, pp. 1-10 doi:10.1145/3437378.3444367.

4. Performance Evaluation of Serverless Frameworks on the Edge



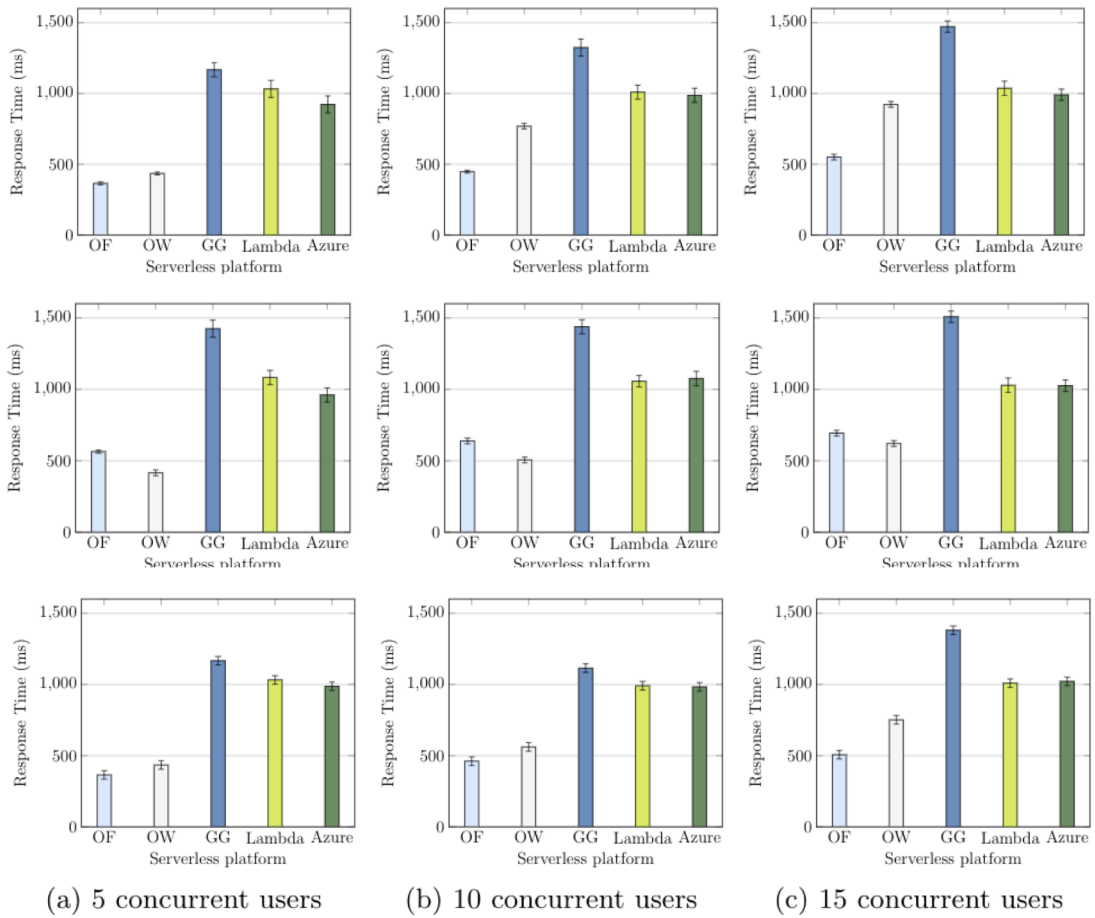
OpenFaaS **OpenWhisk** **AWS Lambda** **AWS GreenGrass** **Azure Functions**

Performance Evaluation Results

CPU-Intensive

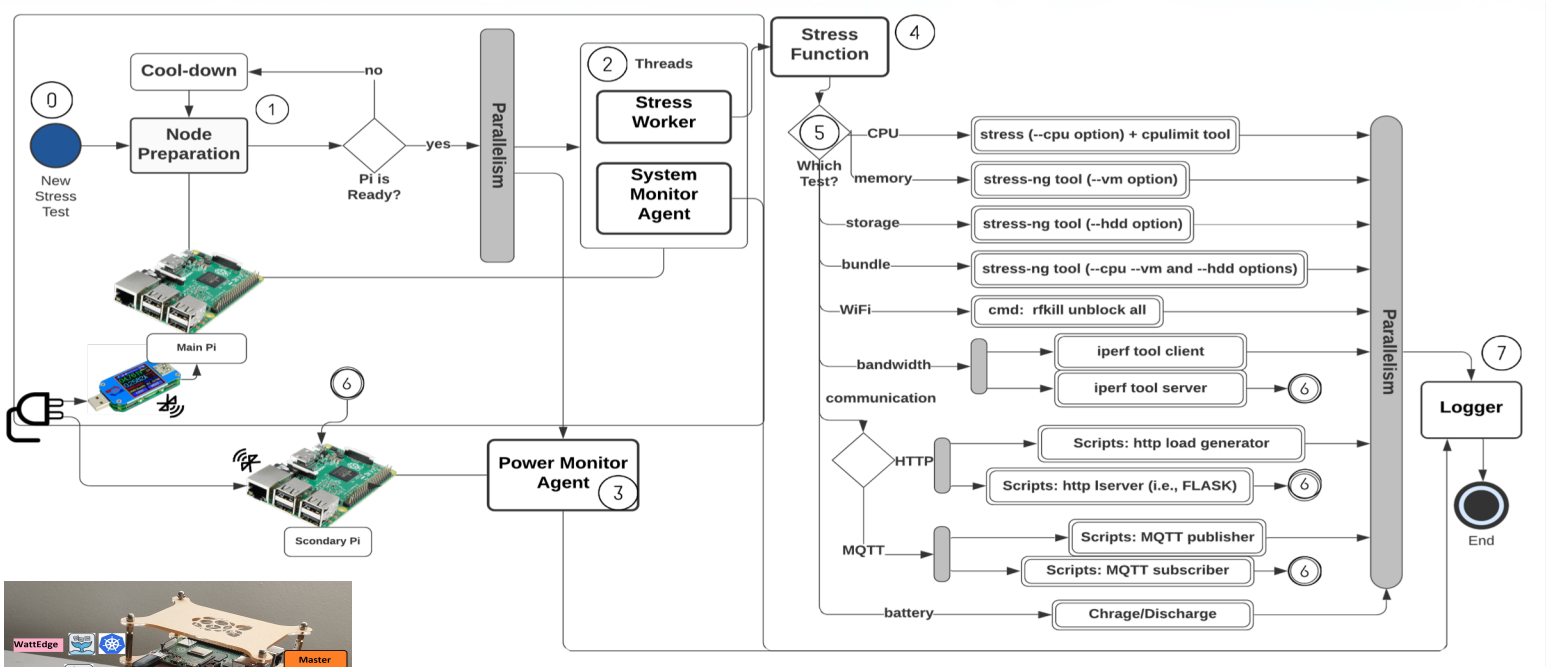
Memory-Intensive

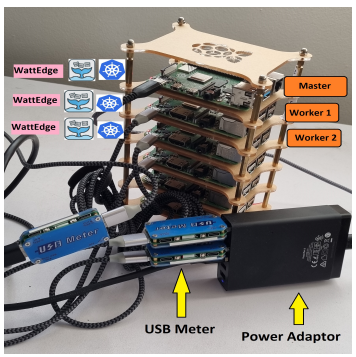
I/O (Disk)-Intensive



(a) 5 concurrent users (b) 10 concurrent users (c) 15 concurrent users

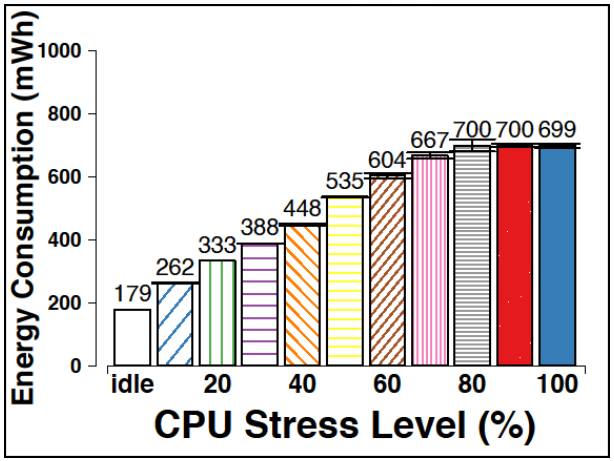
WattEdge

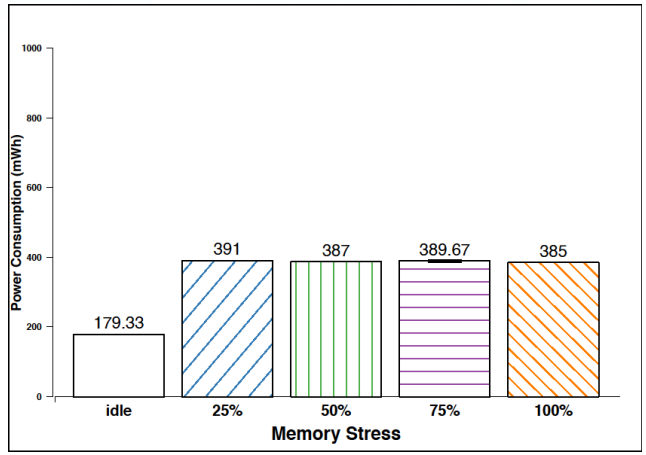


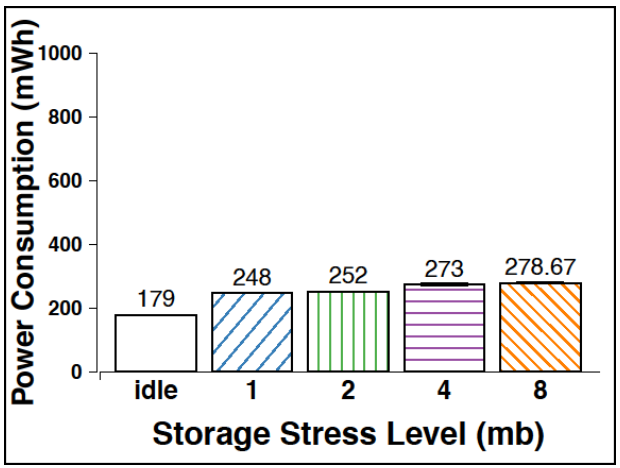


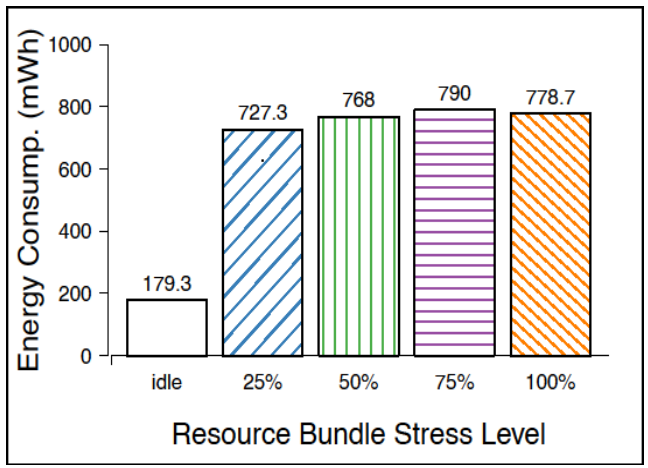
Aslanpour M.S., Toosi A.N., Gaire R., Cheema M.A. (2021) WattEdge: **A Holistic Approach for Empirical Energy Measurements in Edge Computing**. In: Hacid H., Kao O., Mecella M., Moha N., Paik H. (eds) *Service-Oriented Computing*. ICSOC 2021. Lecture Notes in Computer Science, vol. 13121. Springer, Cham.

CPU, Memory and Disk

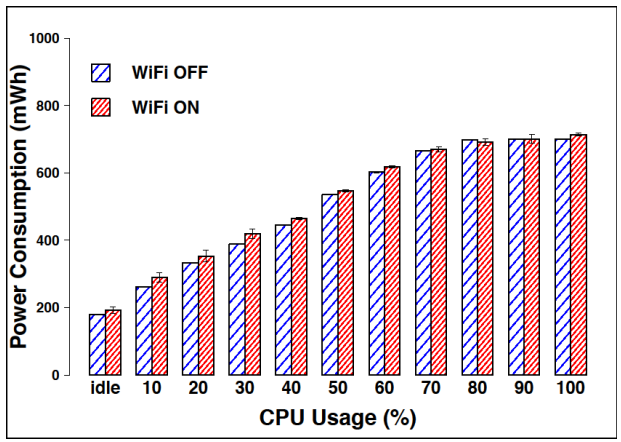


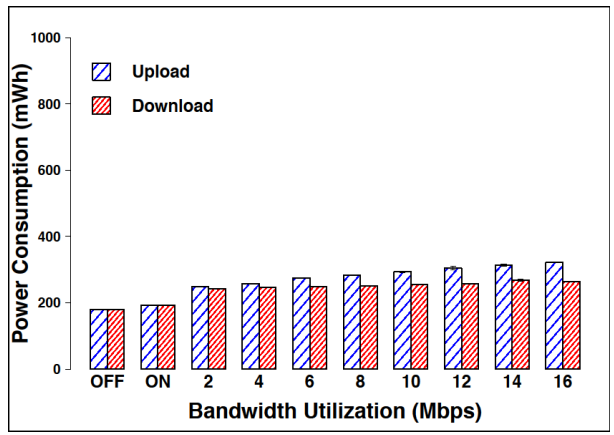


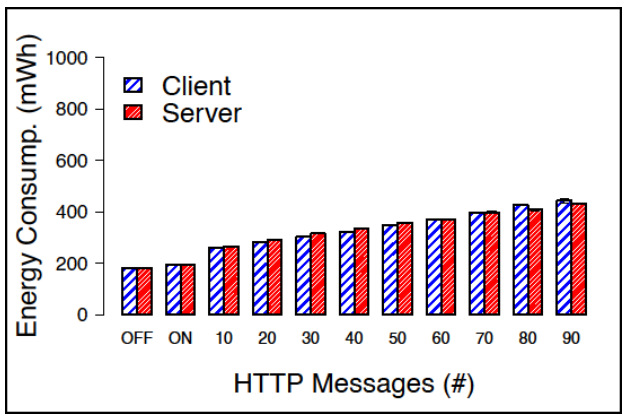


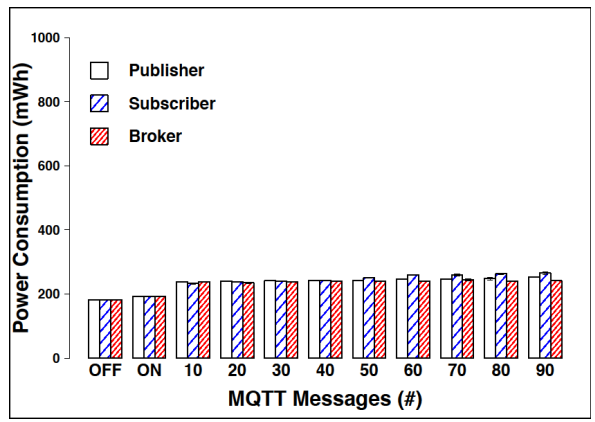


Connectivity and network protocols





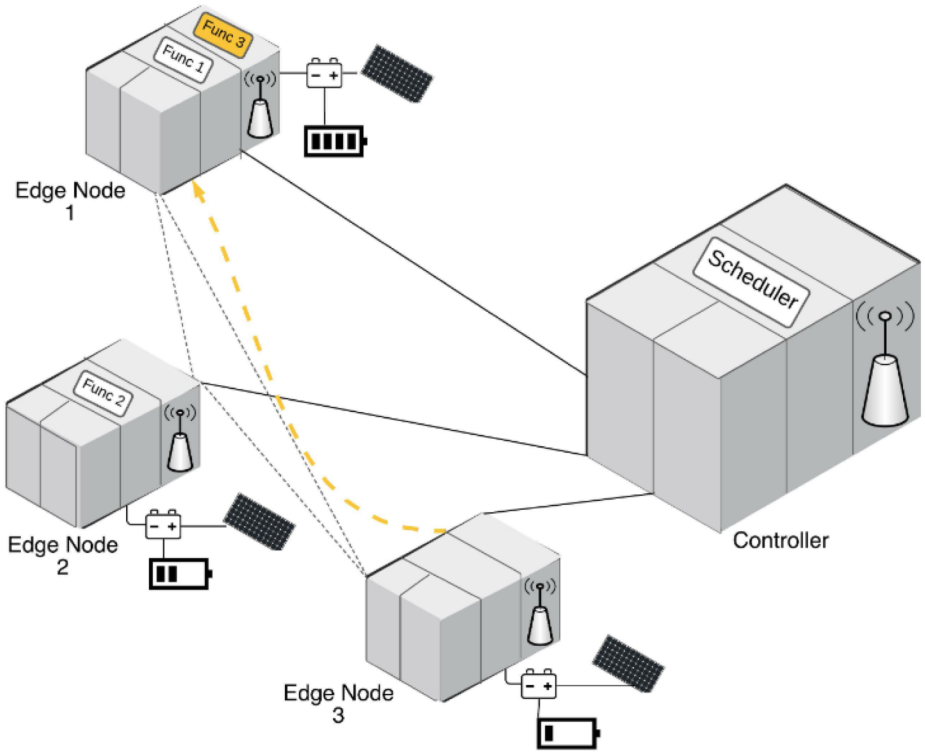




Battery

Major factors: Apart from CPU and idle state that are major energy consumption factors, connectivity prompts to be a major factor, neglected to a large extent by the literature.
Moderate factors: Communication protocols and resource bundle are found to be moderate factors.
Minor factors: Impact of the memory and storage utilization appeared to be less significant.

faasHouse: Sustainable Serverless Edge Computing through Resource Scheduling



faasHouse

- Scoring
  - Energy, Locality, and Stickiness
- Assignment
  - House Allocation Problem: the problem of assigning houses (nodes) to people (functions) considering people's preferences

Benchmarks

- Evaluated against to the following benchmarks:
- **Optimal**: This is an offline optimal algorithm which requiresthe future knowledge of renewable energy input and incomingworkload for each time slot (constrained optimisation problem)
- **Local**: This baseline algorithm always deploys functions locally. This is worth evaluating to understand the impact of offloading.
- **Default**: This is the default performance-aware scheduler in Kubernetes.
- **Random**: This randomly places functions across the cluster.

Some Results

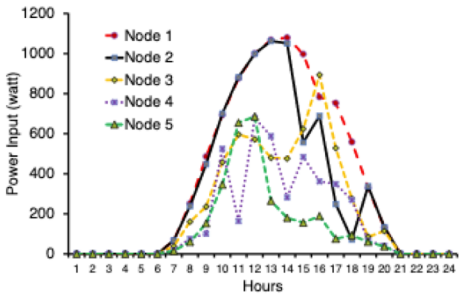


Figure 4: Solar energy input to nodes

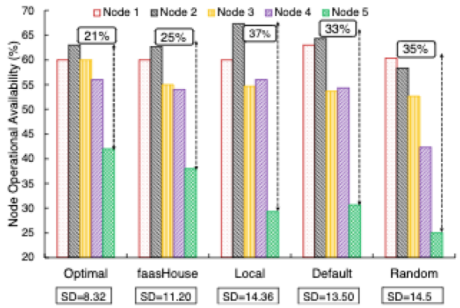


Figure 5: Operationally per scheduler.

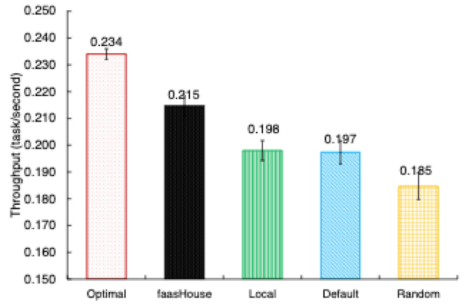
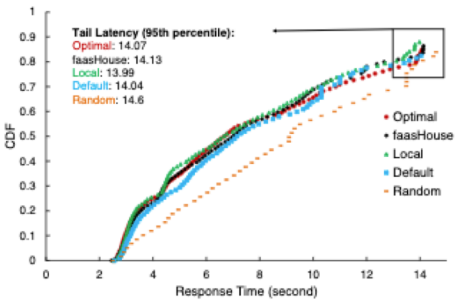
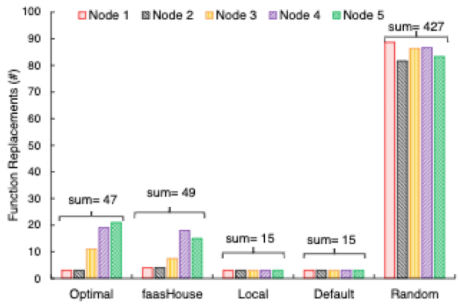
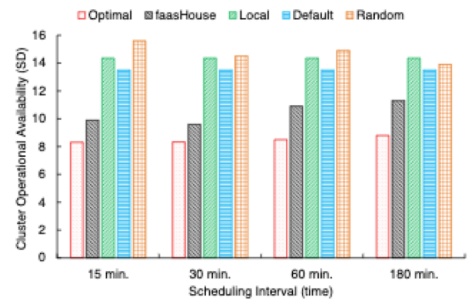


Figure 6: Throughput per scheduler







Summary

- High energy consumptions by cloud data centers
- A few examples of using renewable energy to power data centers
- Issues and challenges of edge computing
- Example of sustainable edge computing and energy discussion
- Sustainability and reduction of energy is an important matter and timely issue, if they are **not** attended to they **may become** severe issue .

THANK YOU

Questions?