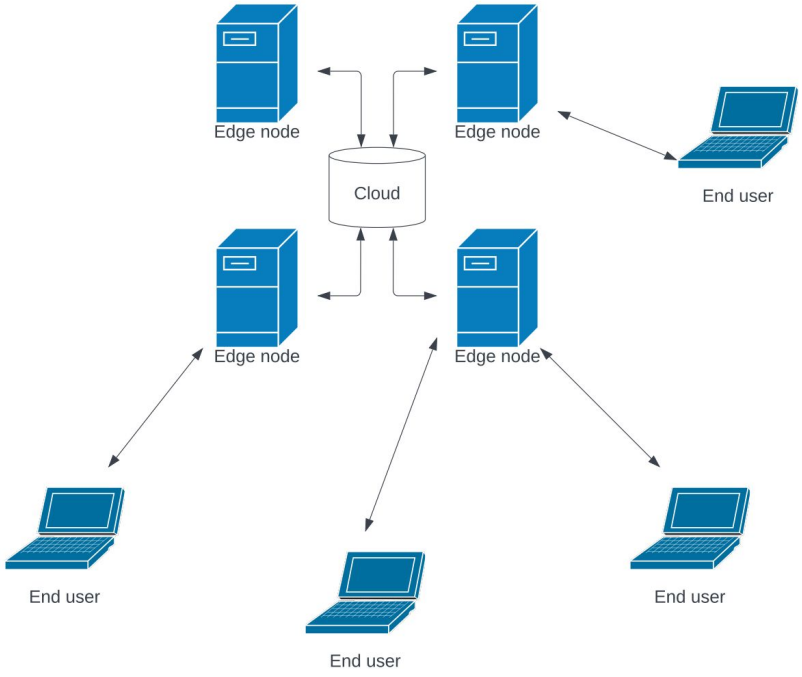
Ubiquitous and Mobile Computing CS 528: *The Emerging Landscape of Edge Computing*

Devesh Bhangale, Luke Foley,
Nick Markou, Saurabh Pande,
Mark Renzi

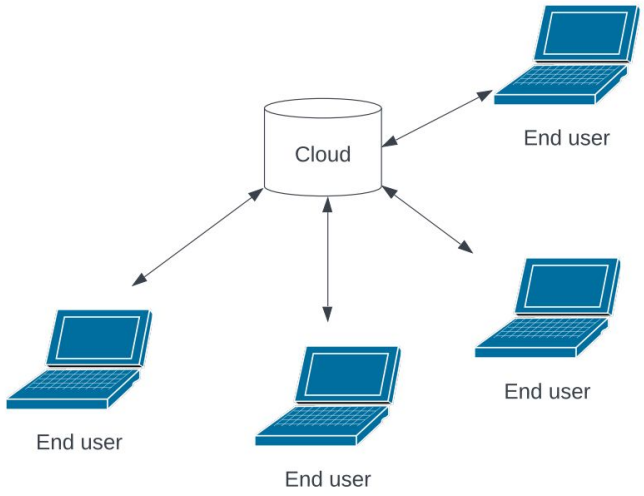
Computer Science Dept.
Worcester Polytechnic Institute (WPI)

Edge vs. Cloud: Latency

Edge Infrastructure Deployment

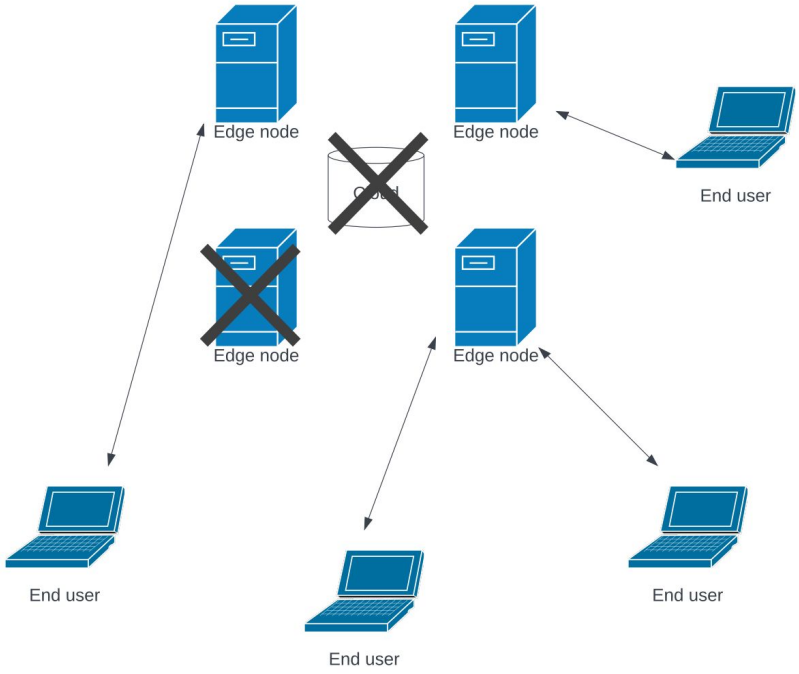


Cloud-only Deployment

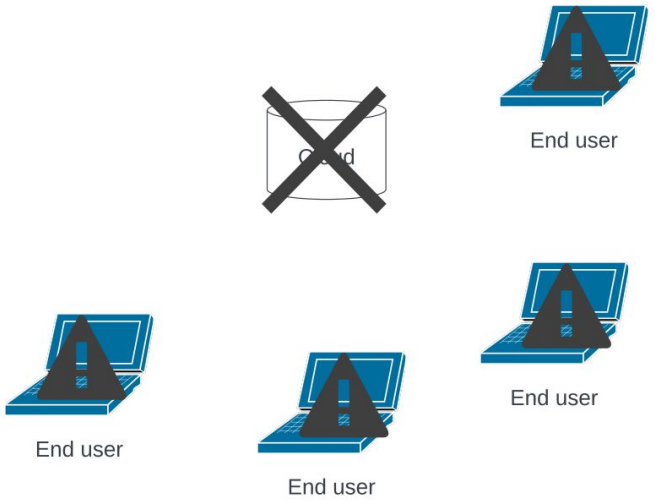


Edge vs. Cloud: Reliability

Edge Infrastructure Deployment



Cloud-only Deployment



What is Edge computing?

- Edge computing refers to a distributed computing framework designed to task other machines with computing problems closer to the user than the central data source
  - Improves response times, saves bandwidth
- Edge deployments localized to users
- Complements cloud computing
  - Processes data closer to where it is generated, reducing latency, then storing in the cloud later

Initial Goal of Edge Computing

- “Cyber Foraging”
  - Short-lived and low-latency jobs
- Mobile Focus
  - Increasing interactivity with more powerful machines
- Latency Reduction
  - Allows for enhancing the user experience

Altered Course of Edge Computing

- Enterprise-Driven Deployments
  - Single-tenant, long-running
- Network Limitations and Reliability
  - Limited bandwidth, unreliable network links to the cloud, and the need to tolerate cloud outages
- Mobile → Industrial and Safety-Critical Applications
  - Industrial sensing, video analytics, anything that requires reliability and continuous operation
- On-Premise Deployments

Edge Applications

1) Business Intelligence
   a) Predicting customer interactions and restocking
2) Smart Cities
   a) Cameras and sensors for accidents and traffic flow

Industry		Company	Use case
Business	Restaurants	Chick-fil-A	– forecast food preparation (e.g., more food needs to be fried)
	Retail	Wal-Mart, Coca Cola (vending machines)	– monitoring (e.g., fridge temperature ensuring produce quality) – tracking customers & improving sales (e.g., customized coupons)
	Gas station	Shell	– detect safety hazards (e.g., a person smoking a cigarette) across their 44,000 gas stations
Smart Cities	Cities	City of Bellevue	– traffic administration (e.g., intelligent control of traffic light) – safety at intersections (e.g., alerting drivers to prevent accidents)
	Construction	PCL, ATF Services	– increase safety, efficiency, and productivity (e.g., detecting a temperature spike or gas leak in a unit) – increase security of construction cites (e.g., protecting equipment overnight)

Edge Applications

3) Intelligent Transportation
   a) Sensing train and airplane issues to prevent accidents and delays

4) Industrial Plants
   a) monitor mechanical equipment, worker safety, and production workflows

Transportation	Aviation	Airbus, Bombardier	– analyze in-flight experience of customers – monitor aircraft operations and maintenance
	Railway	CAF	– monitor train tracks, freight cars, and wheels for problems that lead to derailment
	Road Control	Alaska DOT	– monitor quality of roads and detect roads with need of maintenance (e.g., finding spots that need snow plowing to prevent icing)
Industrial Plants	Oil Refinary	Schneider Electric, ExxonMobil	– predictive maintenance (of the pumps and equipment) – workplace safety
	Manufacturing	GE, CPG, DAIHEN, Airbus	– improve manufacturing yields (e.g., automation or detecting defected products) – monitor equipment & predict need for maintenance
	Manufacturing	BMW	– manage fleet of robots aiding in production pipeline
	Agriculture	Buhler	– control quality of produce at harvest, storage, and processing using imagery (e.g., for grains, processing 20,000 kernels/s).
	Agriculture	DroneWorks, FarmBeats	– observe and monitor agricultural fields using sensors and drone imagery (e.g., detect areas that need water or pesticides)

Deployment Architecture

Input Devices:

- Sensors and/or cameras ranging in number of devices and data stored

Edge Compute:

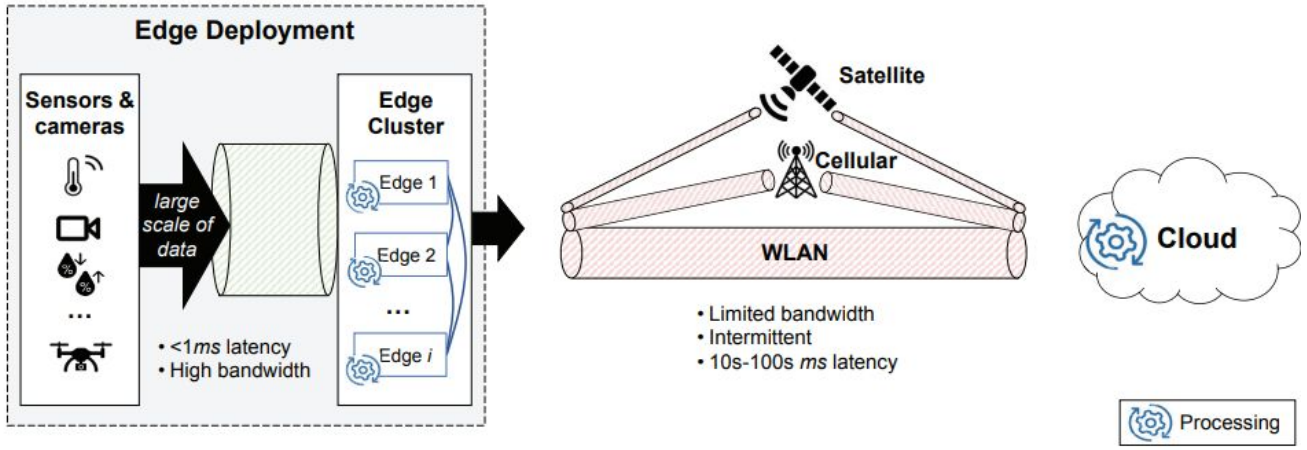
- Owned by one enterprise and perform in hierarchy of computing power

Connectivity to Cloud:

- Outlinks to cloud are shared among every edge

Current edge computing deployments differ from original vision

- Deployed by single entity (not multi-tenant) with enterprises (not consumer applications) driving deployments.



Problem

- How can this altered direction of edge computing be deployed
- Edge computing needs further research to be more applicable and effective for future works
- Current systems poorly support developing and utilizing both edge and cloud computing in an effective way

Motivation

- To propose new methods to use edge computing in mobile deployments
- To identify solutions that can make edge computing more effective for deployment

Related Work

- Adapting app by current network environment
  - Reduced quality media over cellular data or a weak network connection
  - No common set of abstractions to use edge in this scenario
- Bandwidth Optimization
  - Multiple servers producing a response to one client
  - Client with multiple receivers requesting information from servers

Related Work Contd.

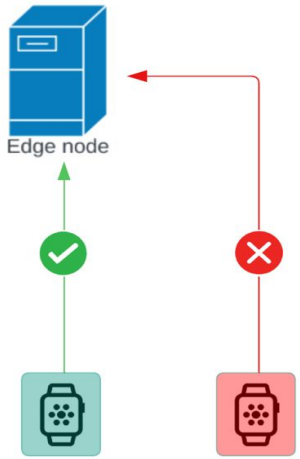
- Odessa: Enabling Interactive Perception Applications on Mobile Devices
  - Designed software that intends to make parallelism/offloading from smartphones dynamic
  - 3x faster than expert-suggested distributed computing strategies for computer vision
  - Easy runtime implementation for distributing the work of any type of computer vision to available edge servers

Methodology: Limitations

- Cyber foraging model
  - This may not be feasible because a lightweight computing device searching for nearby edge computers only works consistently when either:
    - There are edge sites available for repurposing under different tasks/for different devices
    - There are many edge sites deployed across a region only for a specific type of wearable or device, which is expensive

Methodology: Design Tradeoffs

- Security considerations
  - A device operating under the cyber foraging model needs to be resilient to devices designed to mimic a machine that the lightweight device is searching for/intending to connect to



Results

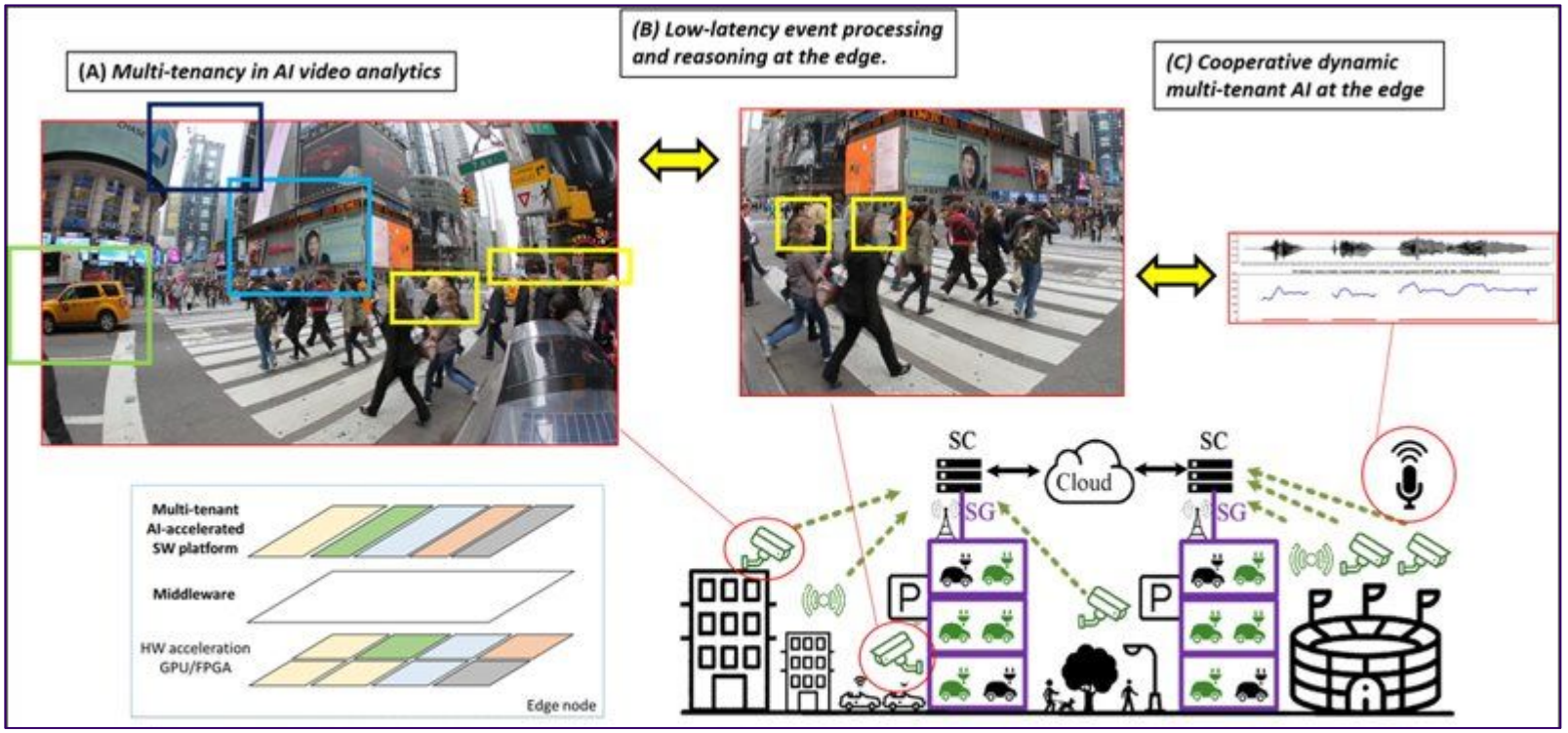
- Edge computing originally needed for additional computation in mobile devices that experience high network latency to the cloud
- Latency has become less of an issue in recent years
  - Use case for edge computing has changed
- Current edge deployments are enterprises driven, and generate high volumes of data for mission-critical applications
  - Used because of lack of reliable connectivity and sufficient bandwidth

Discussion/Conclusion

- Edge has primarily been developed as a solution for situations where data redundancy is critical, and for times where the cloud is inaccessible
- The possibilities of edge computing are extensive, and there are many opportunities for development in the original vision for edge computing
  - Advanced cyber-foraging
  - Better multi-tenant flexibility in deployments
  - More mobile device interactivity

Discussion/Conclusion Contd.

- Use cases of edge computing need to be rethought for maximum effectiveness



Future Research

- Graceful Adaptation of Applications:
  - Research methods to improve switching between edge servers so the end user doesn't notice
- Collaborative & App-aware Network Orchestration
  - Research improved management and prioritization methods
- Test and Verification Frameworks
  - Adding adaptation logic increases application complexity
  - Simplify choosing the right adaptation strategies for developers to decrease complexity

References

- *Shadi A. Noghabi, Landon Cox, Sharad Agarwal, and Ganesh Ananthanarayanan. 2020. The Emerging Landscape of Edge Computing. GetMobile: Mobile Comp. and Comm. 23, 4 (December 2019), 11–20. https://doi.org/10.1145/3400713.3400717*
- *M.-R. Ra, A. Sheth, L. Mummert, P. Pillai, D. Wetherall, and R. Govindan. 2011. Odessa: Enabling interactive perception applications on mobile devices. In Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services (New York, NY, USA, MobiSys '11, ACM, pp. 43–56*