IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

NPTEL

CLOUD COMPUTING

Mobile Cloud Computing - I

Prof. Soumya K Ghosh

Department of Computer Science and Engineering

IIT KHARAGPUR

Motivation

- Growth in the use of Smart phones, apps
- Increased capabilities of mobile devices
- Access of internet using Mobile devices than PCs!

- Resource challenges (battery life, storage, bandwidth etc.) in mobile devices??
- Cloud computing offers advantages to users by allowing them to use infrastructure, platforms and software by cloud providers at low cost and elastically in an on-demand fashion

"Information at your fingertips anywhere anytime.."

https://www.ibm.com/cloud-computing/learn-more/what-is-mobile-cloud-computing/

MobileBackend-as-a-service

What	• Provides mobile application developers a way to connect their application to backend cloud storage and processing
Why	• Abstract away complexities of launching and managing own infrastructure • Focus more on front-end development instead of backend functions
When	• Multiple Apps, Multiple Backends, Multiple Developers • Multiple Mobile Platforms, Multiple Integration, Multiple 3rd Party Systems & Tools
How	• Meaningful resources for app development acceleration – 3rd party API, Device SDK's, Enterprise Connectors, Social integration, Cloud storage

http://www.rapidvaluesolutions.com/whitepapers/How-MBaaS-is-Shaping-up-Enterprise-Mobility-Space.html

Augmenting Mobiles with Cloud Computing

- Amazon Silk browser
  - Split browser
- Apple Siri
  - Speech recognition in cloud
- Apple iCloud
  - Unlimited storage and sync capabilities
- Image recognition apps on smart-phones useful in developing augmented reality apps on mobile devices
  - Augmented reality app using Google Glass

What is Mobile Cloud Computing?

Mobile cloud computing (MCC) is the combination of cloud computing, mobile computing and wireless networks to bring rich computational resources to mobile users.

- **MCC provides mobile users with data storage and processing services in clouds**
  - ✓ *Obviating the need to have a powerful device configuration (e.g. CPU speed, memory ca*
  - ✓ *Al*
- **Moving**
  - ✓ Po
  - ✓ Accessed over the wireless connection based on a thin native client

Mobile Cloud computing is the combination of cloud computing and mobile networks to bring benefits for mobile users, network operators, as well as cloud providers

Why Mobile Cloud Computing?

Speed and flexibility

Mobile cloud applications can be built or revised quickly using cloud services. They can be delivered to many different devices with different operating systems

Shared resources

Mobile apps that run on the cloud are not constrained by a device's storage and processing resources. Data-intensive processes can run in the cloud. User engagement can continue seamlessly from one device to another.

Integrated data

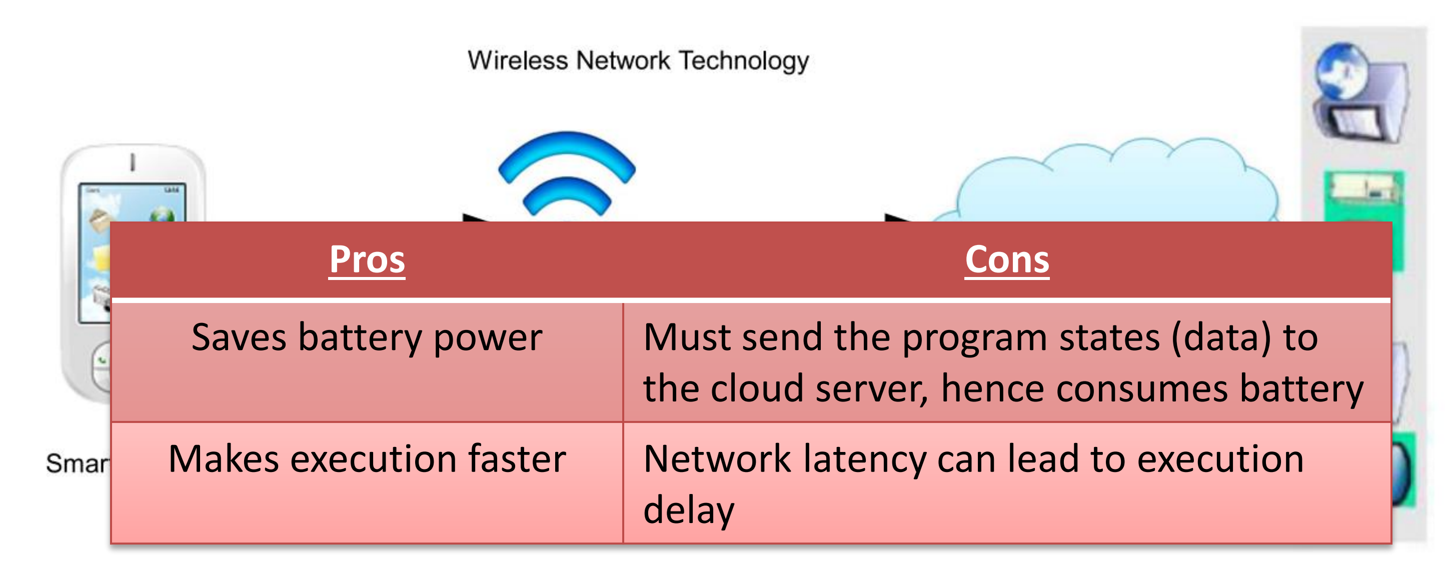
Mobile cloud computing enables users to quickly and securely collect and integrate data from various sources, regardless of where it resides.

Key-features of Mobile Cloud Computing

Mobile cloud computing delivers applications to mobile devices quickly and securely, with capabilities beyond those of local resources

- Facilitates the quick development, delivery and management of mobile apps
- Uses fewer device resources because applications are cloud-supported
- Supports a variety of development approaches and devices
- Mobile devices connect to services delivered through an API architecture
- Improves reliability with information backed up and stored in the cloud

Mobile Cloud Computing

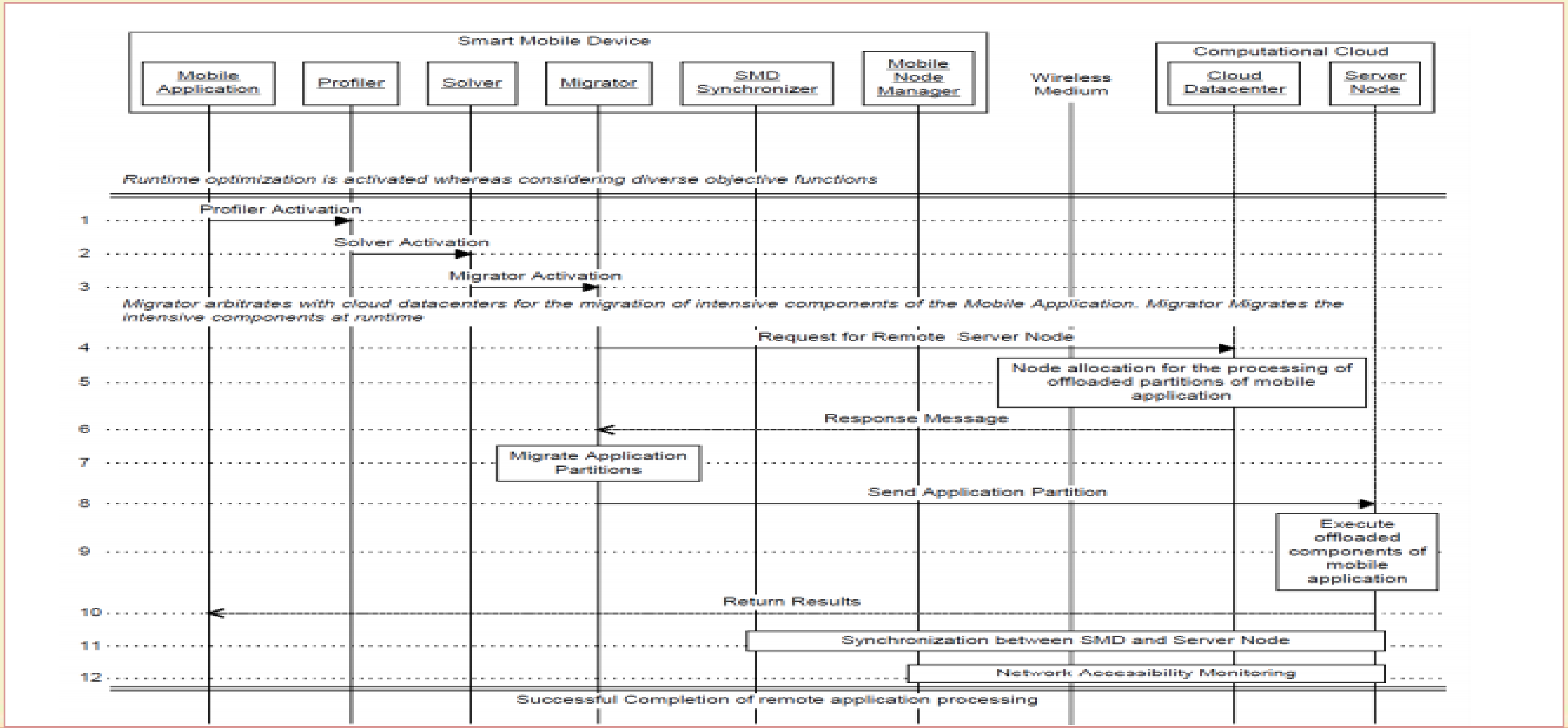


Pros	Cons
Saves battery power	Must send the program states (data) to the cloud server, hence consumes battery
Makes execution faster	Network latency can lead to execution delay

Mobile Cloud Computing is a framework to augment a resource constrained mobile device to execute parts of the program on cloud based servers

Shiraz, Muhammad, et al. "A review on distributed application processing frameworks in smart mobile devices for mobile cloud computing." Communications Surveys & Tutorials, IEEE 15.3 (2013): 1294-1313

Typical MCC Workflow



Shiraz, Muhammad, et al. "A review on distributed application processing frameworks in smart mobile devices for mobile cloud computing." Communications Surveys & Tutorials, IEEE 15.3 (2013): 1294-1313

Dynamic Runtime Offloading

Dynamic runtime offloading involves the issues of

- dynamic application profiling and solver on SMD
- runtime application partitioning
- migration of intensive components
- continuous synchronization for the entire duration of runtime execution platform.

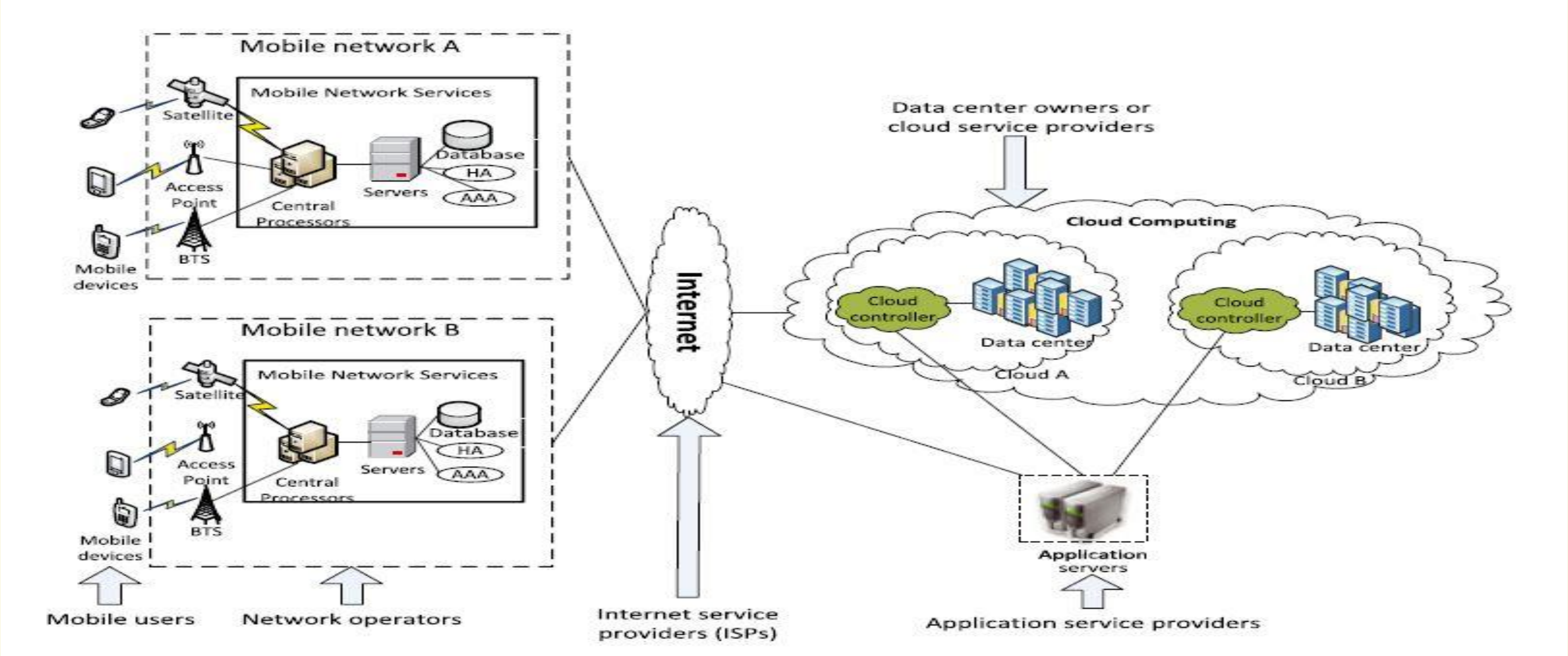
MCC key components

- Profiler
  - Profiler monitors application execution to collect data about the time to execute, power consumption, network traffic
- Solver
  - Solver has the task of selecting which parts of an app runs on mobile and cloud
- Synchronizer
  - Task of synchronizer modules is to collect results of split execution and combine, and make the execution details transparent to the user

Key Requirements for MCC

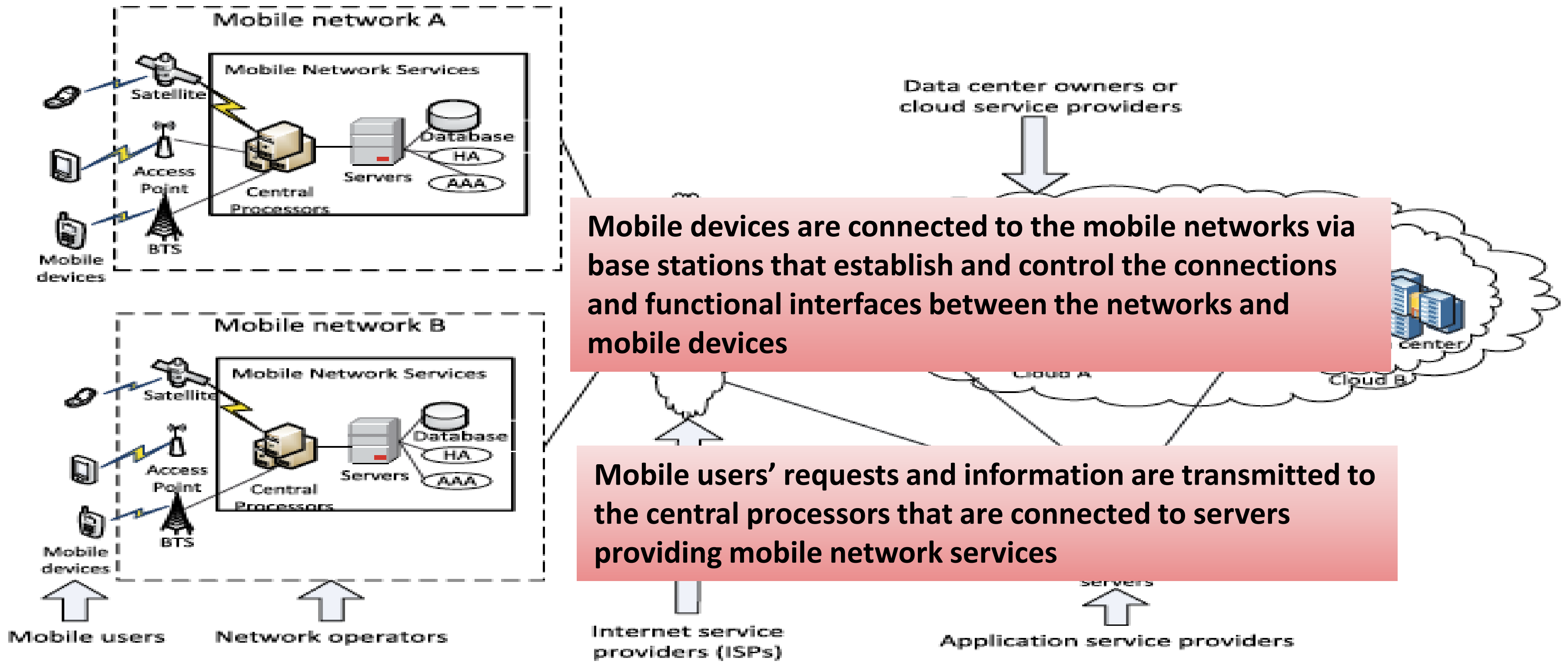
- ***Simple APIs*** offering access to mobile services, and requiring no specific knowledge of underlying network technologies
- ***Web Interface***
- ***Internet access*** to remotely stored applications in the cloud

Mobile Cloud Computing – Typical Architecture



Dinh, Hoang T., et al. "A survey of mobile cloud computing: architecture, applications, and approaches." Wireless communications and mobile computing 13.18 (2013): 1587-1611.

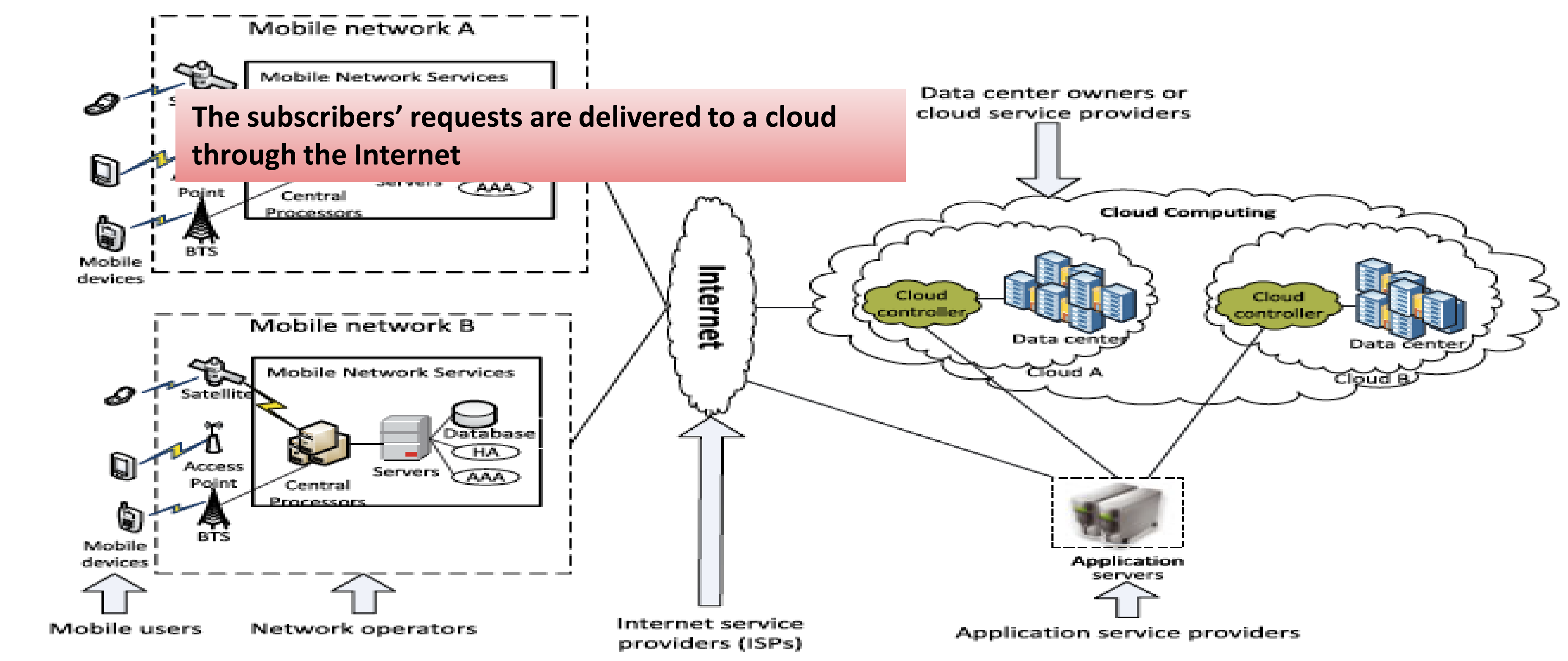
Mobile Cloud Computing - Architecture



Mobile devices are connected to the mobile networks via base stations that establish and control the connections and functional interfaces between the networks and mobile devices

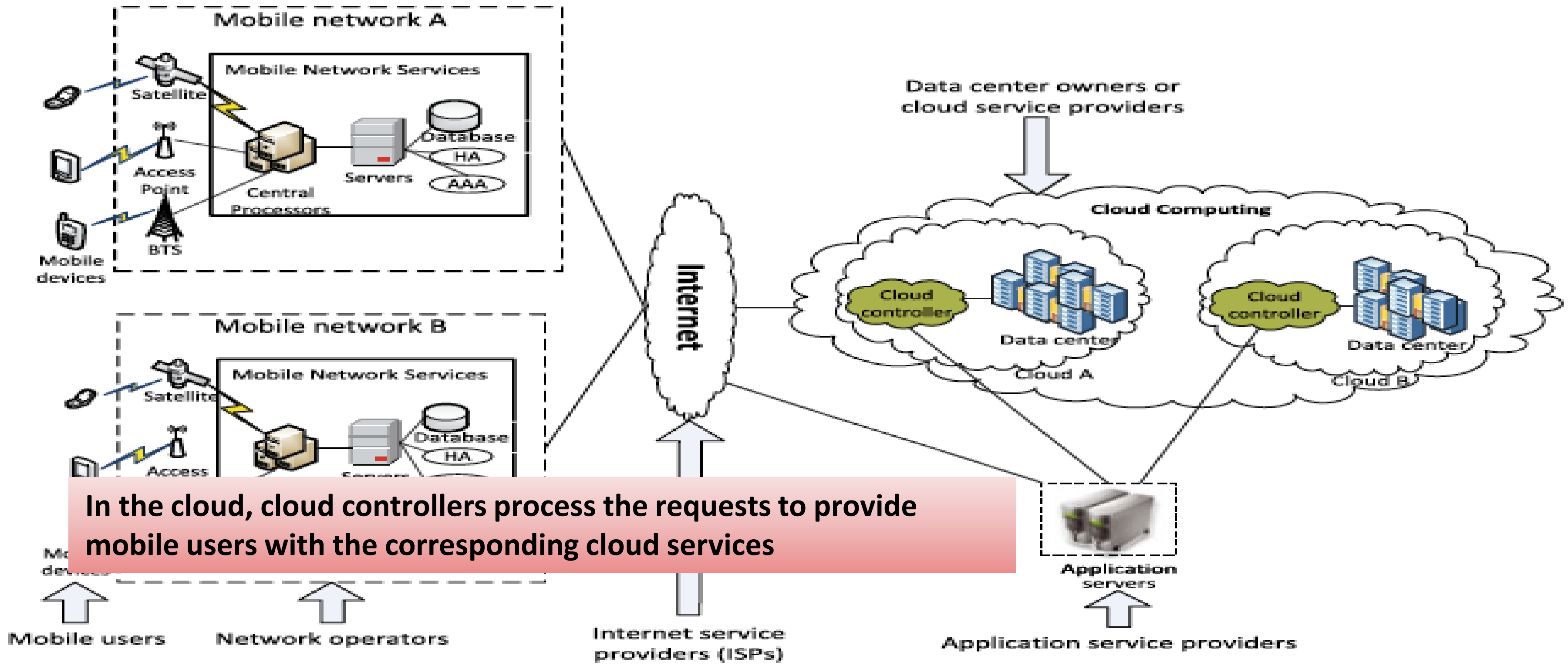
Mobile users' requests and information are transmitted to the central processors that are connected to servers providing mobile network services

Mobile Cloud Computing - Architecture



The subscribers' requests are delivered to a cloud through the Internet

Mobile Cloud Computing - Architecture



In the cloud, cloud controllers process the requests to provide mobile users with the corresponding cloud services

Advantages of MCC

Extending battery lifetime

- Computation offloading migrates large computations and complex processing from resource-limited devices (i.e., mobile devices) to resourceful machines (i.e., servers in clouds).
- Remote application execution can save energy significantly.
- Many mobile applications take advantages from task migration and remote processing

Improving data storage capacity and processing power

- MCC enables mobile users to store/access large data on the cloud.
- MCC helps reduce the running cost for computation intensive applications.
- Mobile applications are not constrained by storage capacity on the devices because their data now is stored on the cloud

Advantages of MCC (contd…)

Improving Reliability and Availability

– Keeping data and application in the clouds reduces the chance of lost on the mobile devices.

– MCC can be designed as a comprehensive data security model for both service providers and users:

  - Protect copyrighted digital contents in clouds.
  - Provide security services such as virus scanning, malicious code detection, authentication for mobile users.

– With data and services in the clouds, then are always(almost) available even when the users are moving.

Advantages of MCC

- Dynamic provisioning
- Scalability
- Multi-tenancy
  - Service providers can share the resources and costs to support a variety of applications and large no. of users.
- Ease of Integration
  - Multiple services from different providers can be integrated easily through the cloud and the Internet to meet the users' demands.

Mobile Cloud Computing – Challenges

MCC Security Issues

Protecting user privacy and data/application secrecy from adversaries is key to establish and maintain consumers' trust in the mobile platform, especially in MCC.

MCC security issues have two main categories:

- Security for mobile users
- Securing data on clouds

Mobile Cloud Computing – Challenges

Security and Privacy for Mobile Users

- Mobile devices are exposed to numerous security threats like malicious codes and their vulnerability.
- GPS can cause privacy issues for subscribers.
- Security for mobile applications:
  - Installing and running security software are the simplest ways to detect security threats.
  - Mobile devices are resource constrained, protecting them from the threats is more difficult than that for resourceful devices.
- Location based services (LBS) faces a privacy issue on mobile users' provide private information such as their current location.
- Problem becomes even worse if an adversary knows user's important information.

Mobile Cloud Computing – Challenges

Security for Mobile Users

- Approaches to move the threat detection capabilities to clouds.
- Host agent runs on mobile devices to inspect the file activity on a system. If an identified file is not available in a cache of previous analyzed files, this file will be sent to the in cloud network service for verification.
- Attack detection for a smartphone is performed on a remote server in the cloud.
- The smartphone records only a minimal execution trace, and transmits it to the security server in the cloud.

Mobile Cloud Computing – Challenges

Context-aware Mobile Cloud Services

- It is important to fulfill mobile users' satisfaction by monitoring their preferences and providing appropriate services to each of the users.
- Context-aware mobile cloud services try to utilize the local contexts (e.g., data types, network status, device environments, and user preferences) to improve the quality of service (QoS).

H. H. La and S. D. Kim, "A Conceptual Framework for Provisioning Context-aware Mobile Cloud Services", in Proceedings of IEEE International Conference on Cloud Computing (CLOUD), pp. 466, August 2010.

Mobile Cloud Computing – Challenges

Network Access Management:

- An efficient network access management not only improves link performance but also optimizes bandwidth usage

Quality of Service:

- How to ensure QoS is still a big issue, especially on network delay.
- CloneCloud and Cloudlets are expected to reduce the network delay.
- The idea is to clone the entire set of data and applications from the smartphone onto the cloud and to selectively execute some operations on the clones, reintegrating the results back into the smartphone

Pricing:

- MCC involves both mobile service provider (MSP) and cloud service provider (CSP) with different services management, customers management, methods of payment and prices.
- Business model including pricing and revenue sharing has to be carefully developed for MCC.

Mobile Cloud Computing – Challenges

Standard Interface:

- Interoperability becomes an important issue when mobile users need to interact with the cloud.
- Compatibility among devices for web interface could be an issue.
- Standard protocol, signaling, and interface between mobile users and cloud would be required.

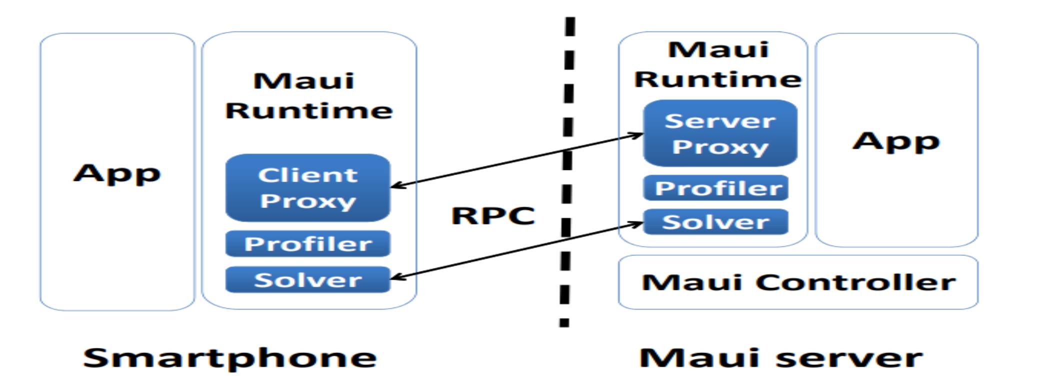
Service Convergence:

- Services will be differentiated according to the types, cost, availability and quality.
- New scheme is needed in which the mobile users can utilize multiple cloud in a unified fashion.
- Automatic discover and compose services for user.
- Sky computing is a model where resources from multiple clouds providers are leveraged to create a large scale distributed infrastructure.
- Service integration (i.e., convergence) would need to be explored.

Key challenges

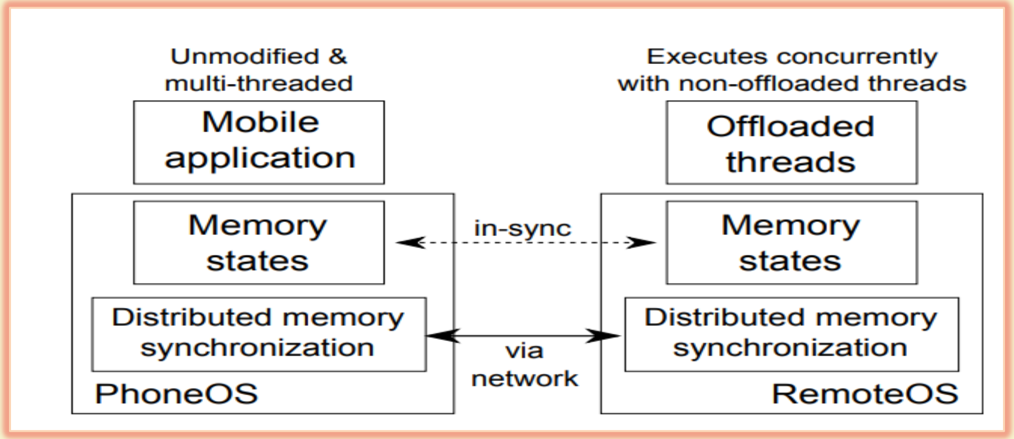
- MCC requires dynamic partitioning of an application to optimize
  - Energy saving
  - Execution time
- Requires a software (middleware) that decides at app launch which parts of the application must execute on the mobile device, and which parts must execute on cloud
  - A classic optimization problem

MCC Systems: MAUI (Mobile Assistance Using Infrastructure)



- ***MAUI enables the programmer to produce an initial partition of the program***
  - Programmer marks each method as "remoteable" or not
  - Native methods cannot be remoteable
- MAUI framework uses the annotation to decide whether a method should be executed on cloud server to save energy and time to execute
- ***MAUI server is the cloud component***. The framework has the necessary software modules required in the workflow.

MCC Systems: COMET



- ***Requires only program binaries Execute multi-threaded programs correctly Improve speed of computation***
- ***Further improvements to data traffic during migration is also possible by sending only the parts of the heap that has been modified***

COMET: Code Offload by Migrating Execution Transparently

- Works on unmodified applications (no source code required)
- Allows threads to migrate between machines depending on workload
- It implements a Distributed Shared Memory (DSM) model for the runtime engine
  - ✓ *DSM allows transparent movement of threads across machines*
  - ✓ *In computer architecture, DSM is a form of memory architecture where the (physically separate) memories can be addressed as one (logically shared) address space*

Gordon, Mark S., et al. "COMET: Code Offload by Migrating Execution Transparently." OSDI. Vol. 12. 2012.

Key Problems to Solve

- At its core, MCC framework must solve how to partition a program for execution on heterogeneous computing resources
- This is a classic “Task Partitioning Problem”
- Widely studied in processor resource scheduling as “job scheduling problem”

Task Partitioning Problem in MCC

Input:

- A call graph representing an application's method call sequence
- Attributes for each node in the graph denotes

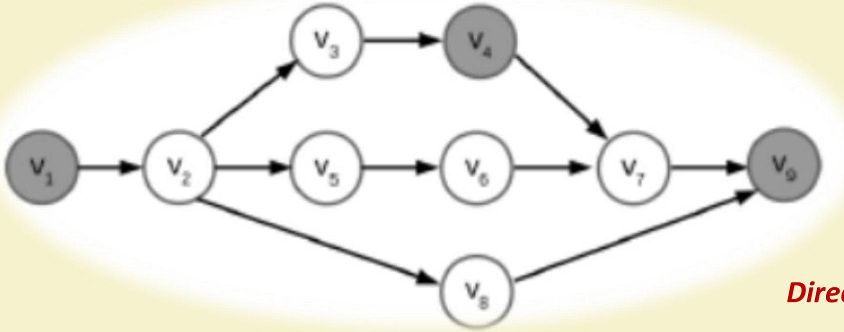
  (a) energy consumed to execute the method on the mobile device,

  (b) energy consumed to transfer the program states to a remote server

Output:

- Partition the methods into two sets – one set marks the methods to execute on the mobile device, and the second set marks the methods to execute on cloud Goals and Constraints:
  1. Energy consumed must be minimized
  2. There is a limit on the execution time of the application
  3. Other constraints could be – some methods must be executed on mobile device, total monetary cost, etc.

Mathematical Formulation



- Highlighted nodes must be executed on the mobile device -> called native tasks (v1, v4, v9)
- Edges represent the sequence of execution - Any non-highlighted node can be executed either locally on the mobile device or on cloud

Directed Acyclic Graph represents an application Call Graph

$$\text{maximize} \sum_{v \in V} I_v \times E_v^l - \sum_{(u,v) \in E} |I_u - I_v| \times C_{u,v}$$

$$\text{such that:} \sum_{v \in V} ((1 - I_v) \times T_v^l) + (I_v \times T_v^r))$$

$$+ \sum_{(u,v) \in E} (|I_u - I_v| \times B_{u,v}) \leq L$$

$$\text{and} \qquad I_v \leq r_v, \ \forall v \in V$$

- 0-1 integer linear program,
  where I_v = 0 if method executed locally,
  = 1 if method executed remotely
- E : Energy cost to execute method v locally
- $C_{u,v}$: Cost of data transfer
- L : Total execution latency
- T : Time to execute the method
- B : Time to transfer program state

Integer Linear Program to solve the Task Partitioning Problem

Mathematical Formulation (Contd..)

- Static Partitioning
  - When an application is launched, invoke an ILP solver which will tell where each method should be executed
  - There are also heuristics to find solutions faster
- Dynamic or Adaptive Partitioning
  - For a long running program, the environmental conditions can vary
  - Depending on the input, the energy consumption of a method can vary

Mobile Cloud Computing – Challenges/ Issues

Mobile communication issues

- Low bandwidth: One of the biggest issues, because the radio resource for wireless networks is much more scarce than wired networks
- Service availability: Mobile users may not be able to connect to the cloud to obtain a service due to traffic congestion, network failures, mobile signal strength problems
- Heterogeneity: Handling wireless connectivity with highly heterogeneous networks to satisfy MCC requirements (always-on connectivity, on-demand scalability, energy efficiency) is a difficult problem

Computing issues (Computation offloading)

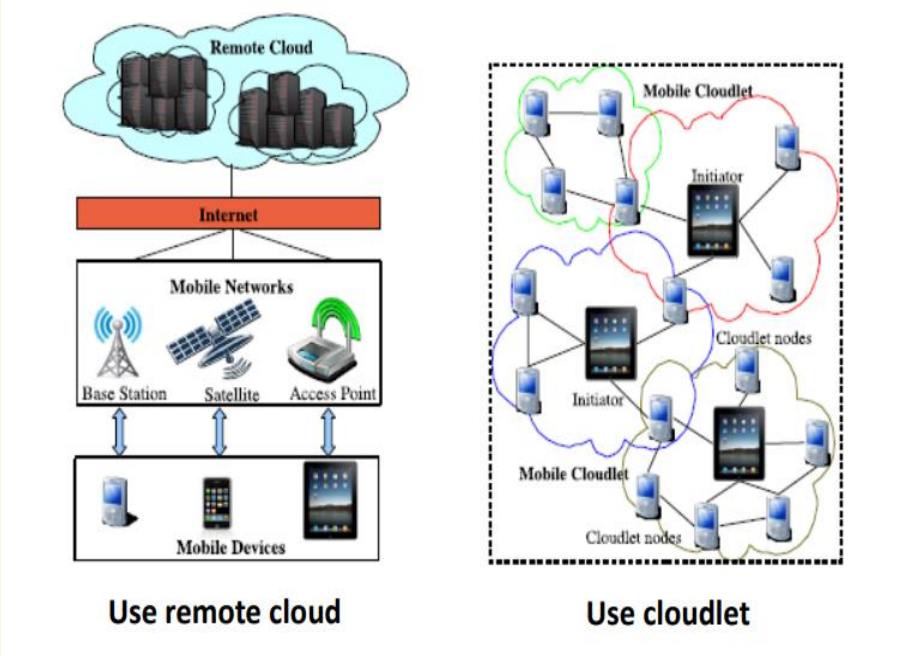
- One of the main features of MCC
- Offloading is not always effective in saving energy
- It is critical to determine whether to offload and which portions of the service codes to offload

CODE OFFLOADING USING CLOUDLET

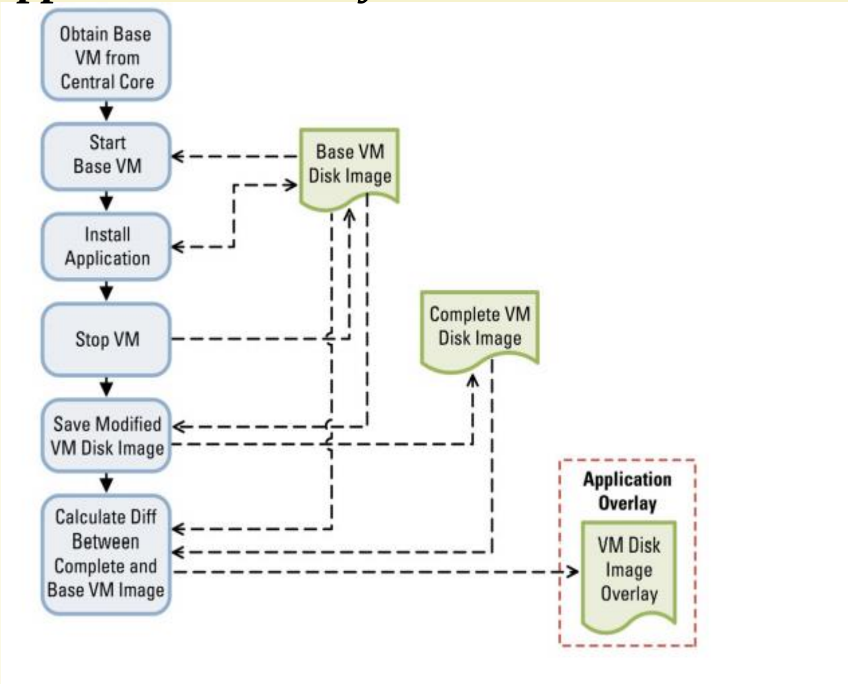
- **CLOUDLET**:
  - ✓ *"a trusted, resource*-rich *computer or cluster of computers that is well-connected to the Internet and is available for use by nearby mobile devices."*
- **Code Offloading** :
  - ✓ Offloading the code to the remote server and executing it.
  - ✓ This architecture decreases latency by using a single-hop network and potentially lowers battery consumption by using Wi-Fi or short-range radio instead of broadband wireless which typically consumes more energy.

CODE OFFLOADING USING CLOUDLET

Cloudlet

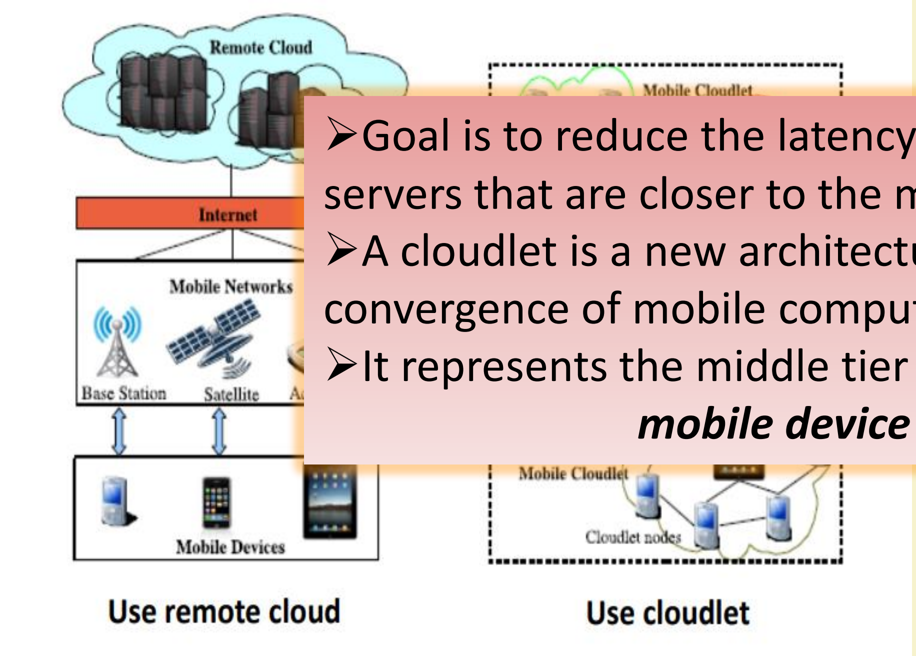


Application Overlay Creation Process

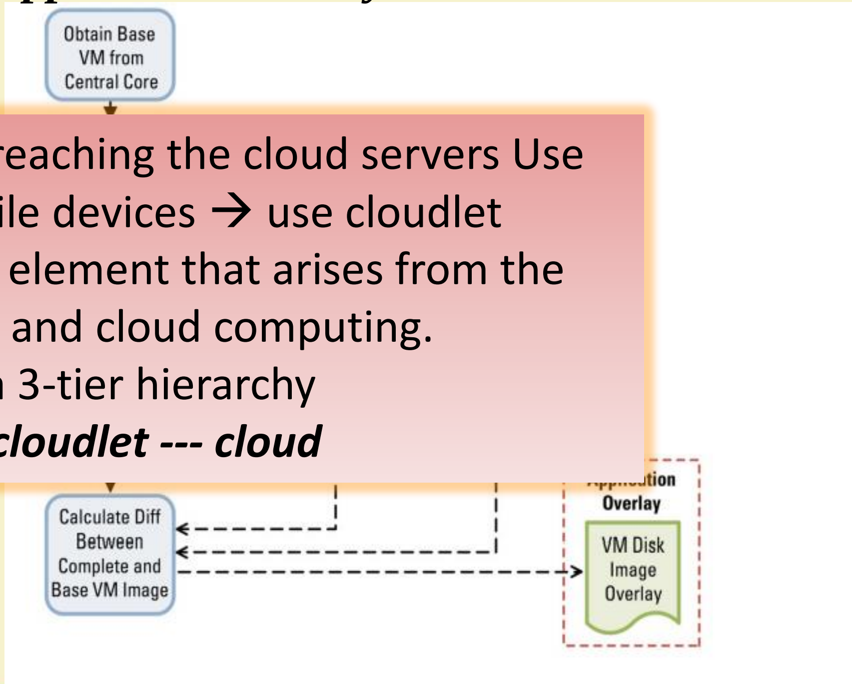


CODE OFFLOADING USING CLOUDLET

Cloudlet



Application Overlay Creation Process



- ➢Goal is to reduce the latency in reaching the cloud servers Use servers that are closer to the mobile devices → use cloudlet
- ➢A cloudlet is a new architectural element that arises from the convergence of mobile computing and cloud computing.
- ➢It represents the middle tier of a 3-tier hierarchy

mobile device --- cloudlet --- cloud

When to Offload??

The amount of energy saved is :

$$P_c \times \frac{C}{M} - P_i \times \frac{C}{S} - P_{tr} \times \frac{D}{B}$$

S: *the speed of cloud to compute* ***C*** *instructions*
M: *the speed of mobile to compute* ***C*** *instructions*
D: *the data need to transmit*
B: *the bandwidth of the wireless Internet*
P_c: *the energy cost per second when the mobile phone is doing computing*
P_i: *the energy cost per second when the mobile phone is idle.*
P_{tr}: *the energy cost per second when the mobile is transmission the data.*

Suppose the server is F times faster—that is, S= F × M.

We can rewrite the formula as

$$\frac{C}{M} \times \left(Pc - \frac{P_i}{F}\right) - P_{tr} \times \frac{D}{B}$$

When to Offload? (contd..)

- Energy is saved when the formula produces a positive number. The formula is positive if D/B is sufficiently small compared with C/M and F is sufficiently large.
- Cloud computing can potentially save energy for mobile users.
- Not all applications are energy efficient when migrated to the cloud.
- Cloud computing services would be significantly different from cloud services for desktops because they must offer energy savings.
- The services should consider the energy overhead for privacy, security, reliability, and data communication before offloading.

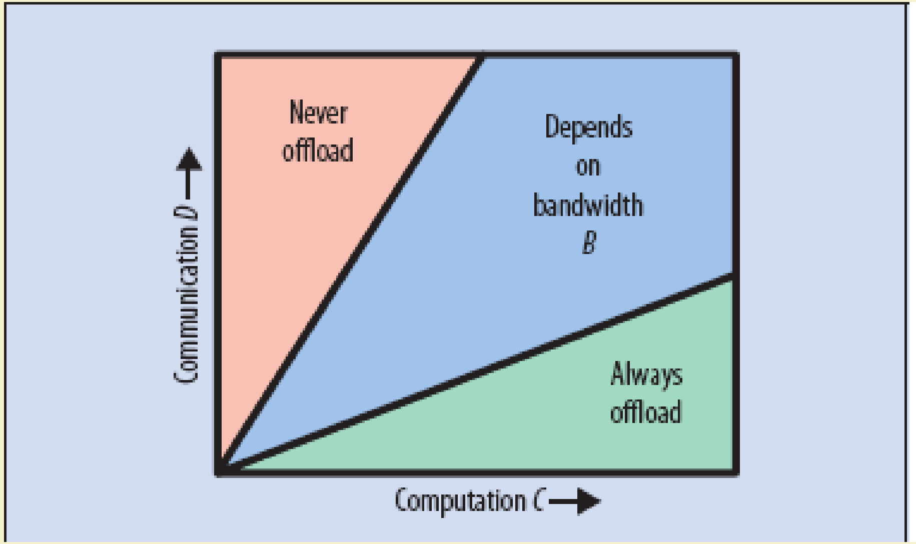
The amount of energy saved is :

$$P_c \times \frac{C}{M} - P_i \times \frac{C}{S} - P_{tr} \times \frac{D}{B}$$

We can rewrite the formula as

$$\frac{C}{M} \times \left(Pc - \frac{P_i}{F}\right) - P_{tr} \times \frac{D}{B}$$

When to Offload?? (contd..)



The amount of energy saved is :

$$P_c \times \frac{C}{M} - P_i \times \frac{C}{S} - P_{tr} \times \frac{D}{B}$$

We can rewrite the formula as

$$\frac{C}{M} \times \left(Pc - \frac{P_i}{F}\right) - P_{tr} \times \frac{D}{B}$$

Offloading is beneficial when large amounts of computation C are needed with relatively small amounts of communication D

Computation Offloading Approaches

- Partition a program based on estimation of energy consumption before execution
- Optimal program partitioning for offloading is dynamically calculated based on the trade-off between the communication and computation costs at run time.
- Offloading scheme based on profiling information about computation time and data sharing at the level of procedure calls.
  - A cost graph is constructed and a branch-and-bound algorithm is applied to minimize the total energy consumption of computation and the total data communication cost.

Z. Li, C. Wang, and R. Xu, “Computation offloading to save energy on handheld devices: a partition scheme,” in Proc 2001 Intl Conf on Compilers, architecture, and synthesis for embedded systems (CASES), pp. 238-246, Nov 2001.
K. Kumar and Y. Lu, “Cloud Computing for Mobile Users: Can Offloading Computation Save Energy,” IEEE Computer, vol. 43, no. 4, April 2010

How to evaluate MCC performance

- Energy Consumption
  - Must reduce energy usage and extend battery life
- Time to Completion
  - Should not take longer to finish the application compared to local execution
- Monetary Cost
  - Cost of network usage and server usage must be optimized
- Security
  - As offloading transfers data to the servers, ensure confidentiality and privacy of data, how to identify methods which process confidential data

Open Questions?

- How can one design a practical and usable MCC framework
  - System as well as partitioning algorithm
- Is there a scalable algorithm for partitioning
  - Optimization formulations are NP-hard
  - Heuristics fail to give any performance guarantee
- Which are the most relevant parameters to consider in the design of MCC systems?

Mobile Cloud Computing – Applications?

Mobile Health-care

Health-Monitoring services, Intelligent emergency management system, Health-aware mobile devices (detect pulse rate, blood pressure, alcohol-level etc.)

Mobile Gaming

It can completely offload game engine requiring large computing resource (e.g., graphic rendering) to the server in the cloud

Mobile Commerce

M-commerce allows business models for commerce using mobile (Mobile financial, mobile advertising, mobile shopping)

Mobile Cloud Computing – Applications?

Pedestrian crossing guide for blind and visually-impaired

Mobile currency reader for blind and visually impaired

Lecture transcription for hearing impaired students

Assistive Technologies

Mobile Learning

- *M-learning combines e-learning and mobility*
- *Traditional m-learning has limitations on high cost of devices/network, low transmission rate, limited educational resources*
- *Cloud-based m-learning can solve these limitations*
- *Enhanced communication quality between students and teachers*
- *Help learners access remote learning resources*
- *A natural environment for collaborative learning*

MuSIC: Mobility-Aware Optimal Service Allocation in Mobile Cloud Computing

- User Mobility introduces new complexities in enabling an optimal decomposition of tasks that can execute cooperatively on mobile clients and the tiered cloud architecture while considering multiple QoS goals such application delay, device power consumption and user cost/price.

- Apart from scalability and access issues with the increased number of users, mobile applications are faced with increased *latencies* and reduced *reliability*

- As a user moves, the physical distance between the user and the cloud resources originally provisioned changes causing additional delays

- Further, the lack of effective handoff mechanisms in WiFi networks as user move rapidly causes an increase in the number of *packet losses*

In other words, user mobility, if not addressed properly, can result in suboptimal resource mapping choices and ultimately in diminished application QoS

Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

MuSIC: Mobility-Aware Optimal Service Allocation in Mobile Cloud Computing

Efficient techniques for *dynamic mapping of resources* in the presence of **mobility**; using a *tiered cloud architecture*, to meet the *multidimensional QoS* needs of mobile users

- Location-time workflow (LTW) as the modeling framework to model mobile applications and capture user mobility. Within this framework, mobile service usage patterns as a function of location and time has been formally modelled

- Given a mobile application execution expressed as a LTW, the framework optimally partitions the execution of the location-time workflow in the 2-tier architecture based on a *utility metric* that combines *service price, power consumption and delay* of the mobile applications

Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

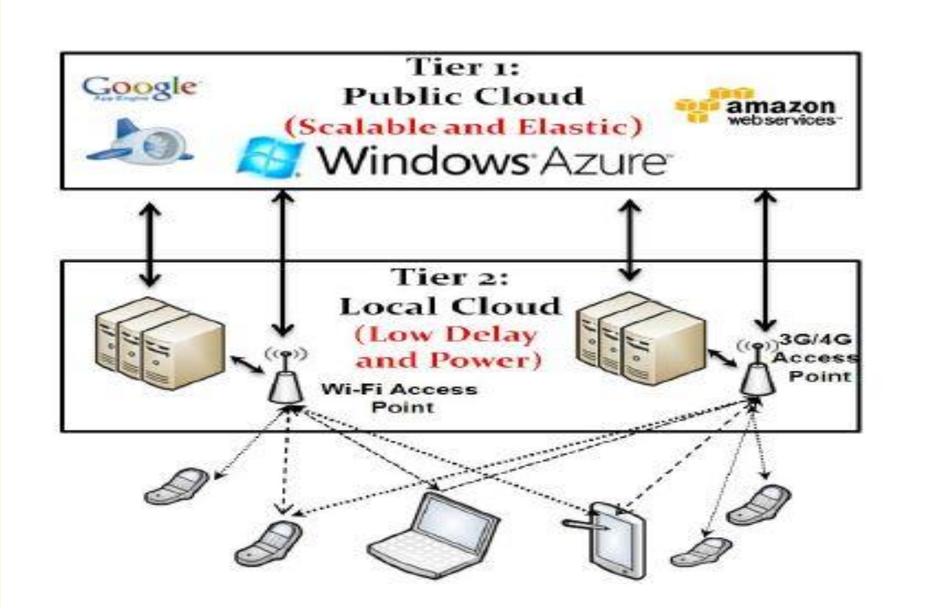
MuSIC: Mobility-Aware Optimal Service Allocation in Mobile Cloud Computing

- ✓ Tier 1 nodes in the system architecture represents *public cloud services* such as Amazon EC2, Microsoft Azure and Google AppEngine. Services provided by these vendors are highly *scalable* and *available*; what they lack is the ability to provide the *fine grain location granularity* required for high performance mobile applications

- ✓ This feature is provided by the second tier local cloud, that consists of nodes that are connected to access points.

- ✓ Location information of these services are available at finer levels of granularity (campus and street level).

- ✓ Mobile users are typically connected to these local clouds through WiFi (via access points) or cellular (via 3G cell towers) connectivity - the aim to intelligently select which local and which public cloud resources to utilize for task offloading.



2-Tier Mobile Cloud Architecture

Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

Mobile Application Modelling

Cloud Service Set:
The set of all services (e.g. compute, storage and software capabilities like multimedia streaming services, content transcoding services, etc) provided by local and public cloud providers

Local Cloud Capacity:
Local cloud services can only accept a limited number of mobile client requests

Location Map:
It is a partition of the 2-D space/region in which mobile hosts and cloud resources are located

User Service Set:
The set of all services that a user has on his own device (e.g. decoders, image editors etc.)

Criteria	Definition
$q_{price}(s_i, u_k^{l_i,t_j})$	The price of using service s_i when user u_k is in location $l_i \in L$ and time t_j.
$q_{power}(s_i, u_k^{l_i,t_j})$	The power consumed on user mobile device using s_i when user u_k is in location $l_i \in L$ and time t_j.
$q_{delay}(s_i, u_k^{l_i,t_j})$	The delay of executing service s_i when user u_k is in location $l_i \in L$ and time t_j.

Mobile User Trajectory:

The trajectory of a mobile user, u_k, is represented as a list of tuples of the form $\{(1; l_k); \ldots; (n; l_m)\}$ where $(i; l_j)$ implies that the mobile user is in location l_j for time duration I

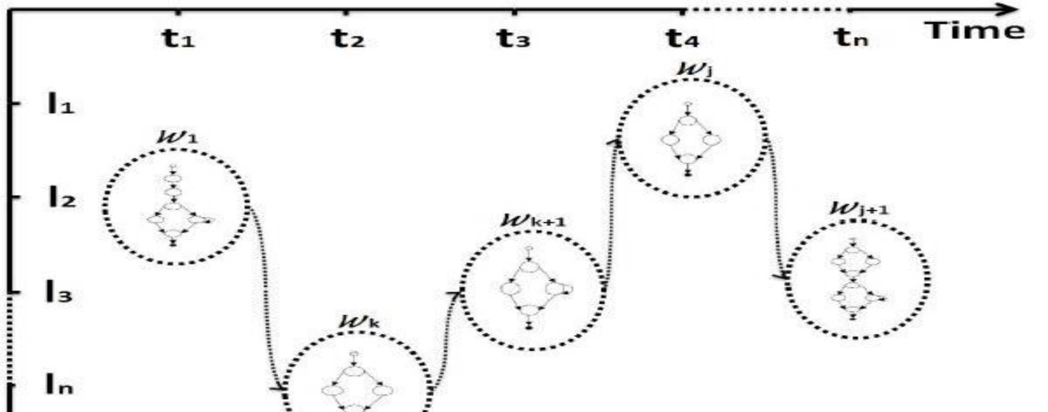
Center of Mobility:
It is the location where (or near where) a mobile user u_k spends most of its time

Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

Mobile Application Modelling

Location-Time Workflow:

Combination of the mobile application workflow concept with a user trajectory to model the mobile users and the requested services in their trajectory.



Location-Time Workflow

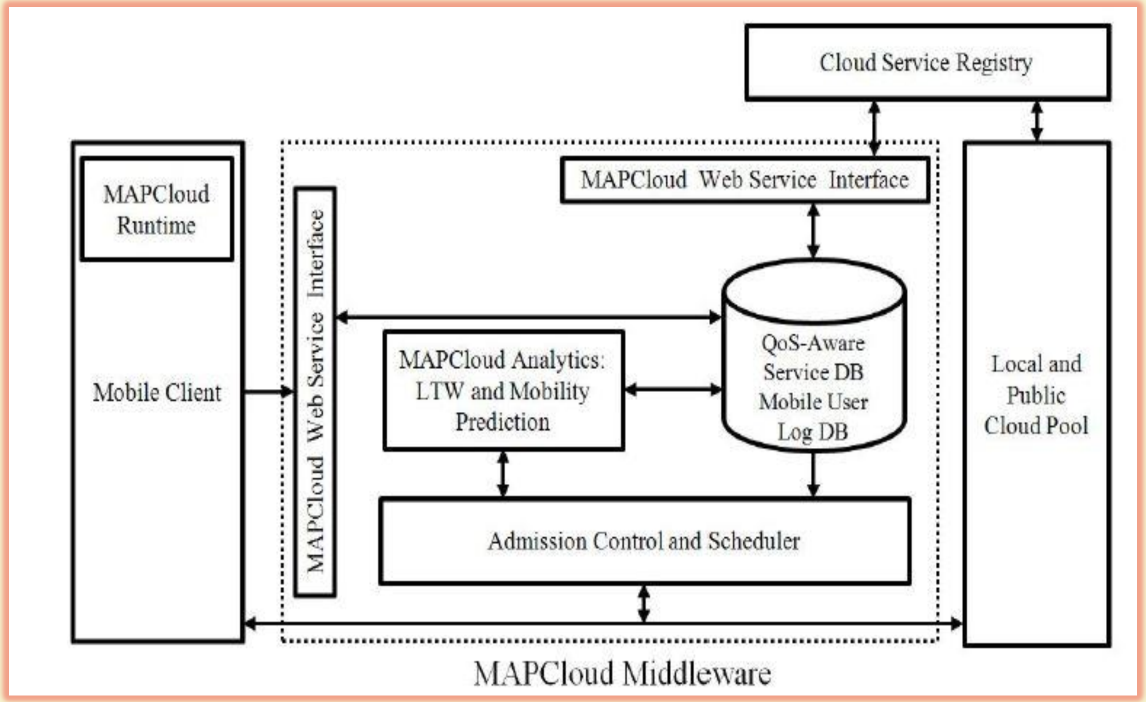
Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

Mobile Application Modelling

Mobile User Log DB and QoS-Aware Service DB:
Unprocessed user data log such as mobile service usage, location of the user, user delay experience of getting the service, energy consumed on user mobile device, etc and service lists on local and public cloud and their QoSes in different locations respectively

MAPCloud Analytic: This module processes mobile user Log DB and updates QoS-aware cloud service DB based on user experience and LTW

Admission Control and Scheduling: This module is responsible for optimally allocate services to admitted mobile users based on MuSIC



Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

A Case Study: Context Aware Dynamic Parking Service

- MCC can provide a flexible method of handling massive computing, storage, and software services in a scalable and virtualized manner.
- The integration of MCC and vehicular networks is expected to promote the development of cost effective, scalable, and data-driven CVC (Context-aware vehicular cyber physical systems)

An application scenario regarding the context-aware dynamic parking services by illuminating the cloud-assisted architecture and logic flow.

- **As the number of vehicles increases, there is an increasing trend of insufficient parking spaces in many large cities, and this problem is gradually getting worse**
- **With the proliferation of wireless sensor networks (WSNs) and cloud computing, there exists strong potential to alleviate this problem using context information (e.g., road conditions and status of parking garages) to provide context-aware dynamic parking services**
- **Cloud Assisted parking services (traditional parking garages and dynamic parking services along the road) and parking reservation service using smart terminals such as smartphones.**

Wan, Jiafu, et al. "Context-aware vehicular cyber-physical systems with cloud support: architecture, challenges, and solutions." *IEEE Communications Magazine* 52.8 (2014): 106-113.

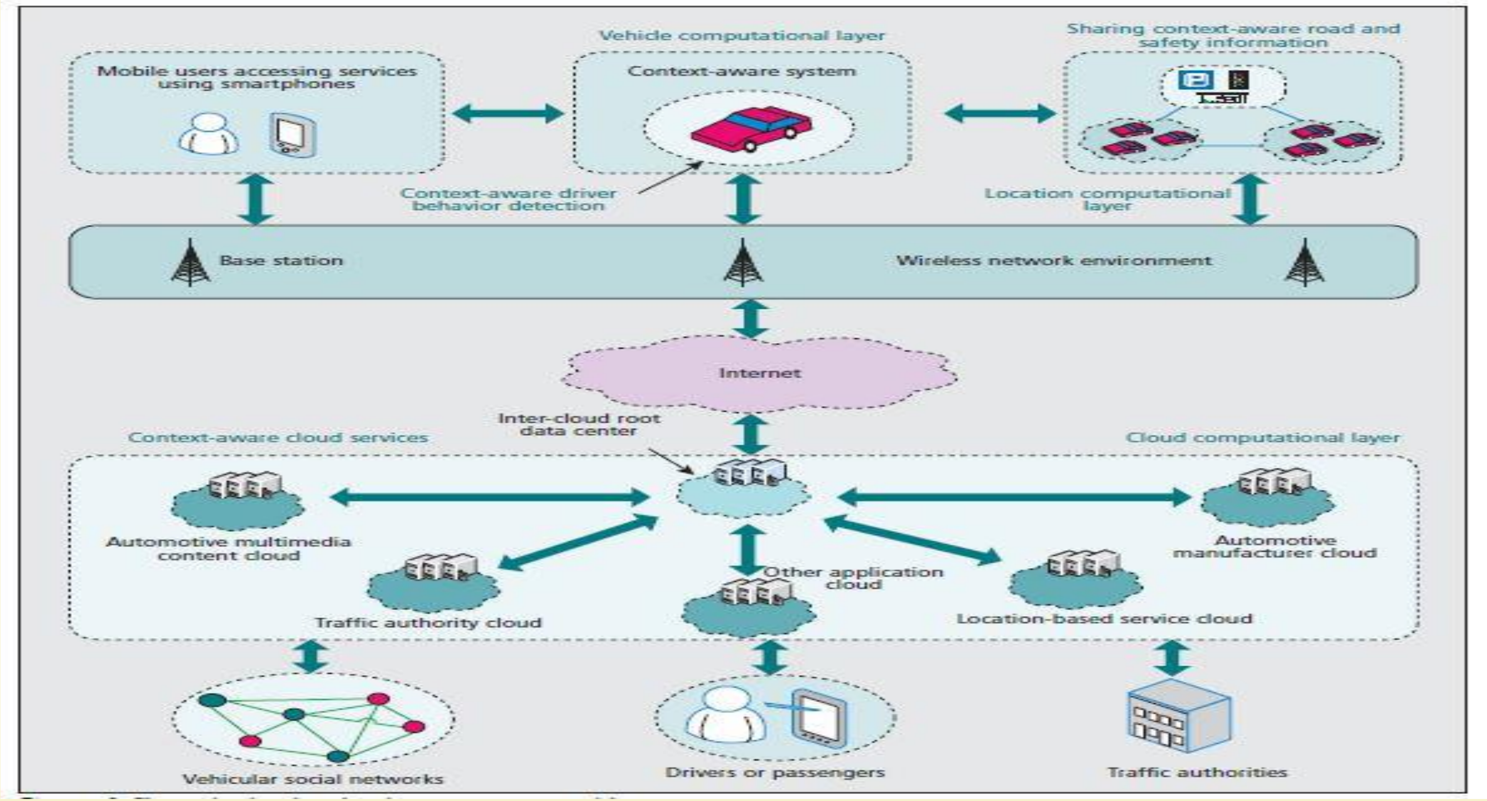
A Case Study: Context Aware Dynamic Parking Service

- MCC can provide a flexible method of handling massive computing, storage, and software services in a scalable and virtualized manner.
- The integration of MCC and vehicular networks is expected to promote the development of cost effective, scalable, and data-driven CVC (Context-aware vehicular cyber physical systems)

An application scenario regarding the context-aware dynamic parking services by illuminating the cloud-assisted architecture and logic flow.

- **As the number of vehicles increases, there is an increasing trend of insufficient parking spaces in many large cities, and this problem is gradually getting worse**
- **With the proliferation of wireless sensor networks (WSNs) and cloud computing, there exists strong potential to alleviate this problem using context information (e.g., road conditions and status of parking garages) to provide context-aware dynamic parking services**
- **Cloud Assisted parking services (traditional parking garages and dynamic parking services along the road) and parking reservation service using smart terminals such as smartphones.**

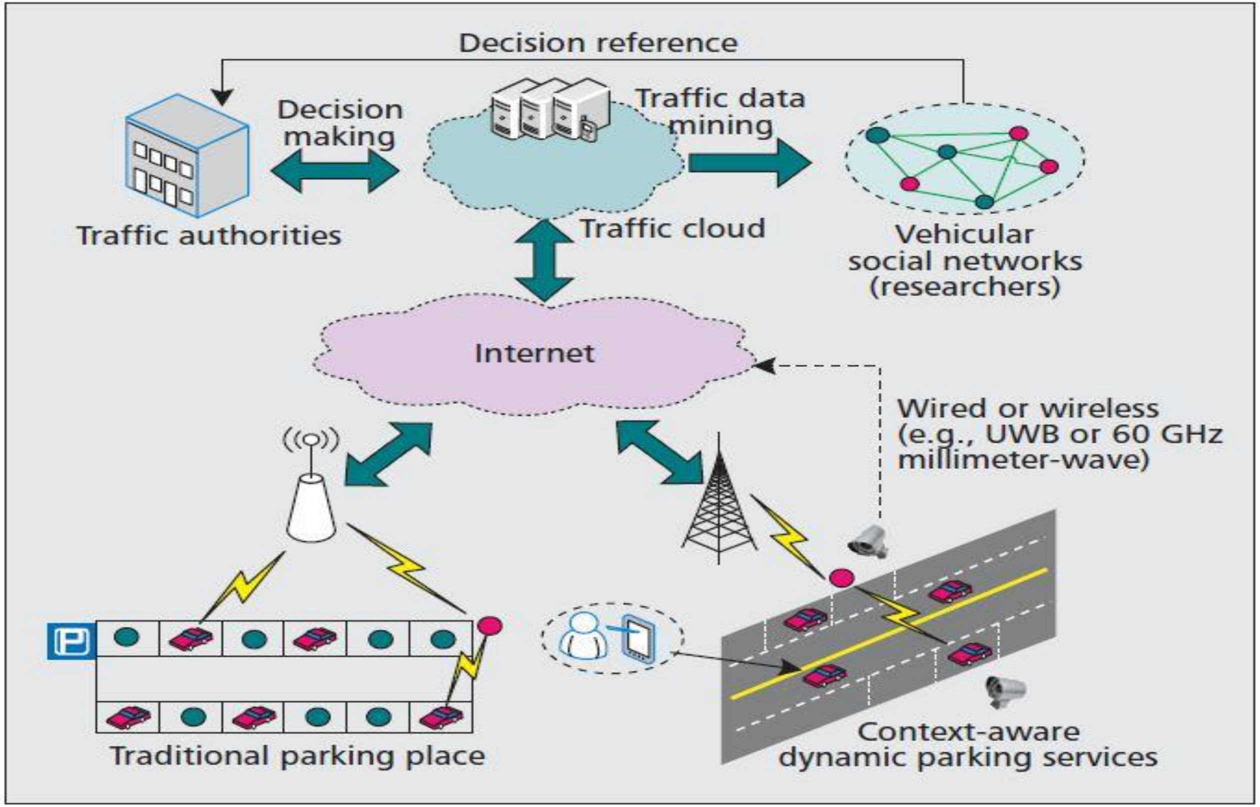
Wan, Jiafu, et al. "Context-aware vehicular cyber-physical systems with cloud support: architecture, challenges, and solutions." *IEEE Communications Magazine* 52.8 (2014): 106-113.



<u>Example cloud-assisted context-aware architecture</u>

Traditional parking garages:

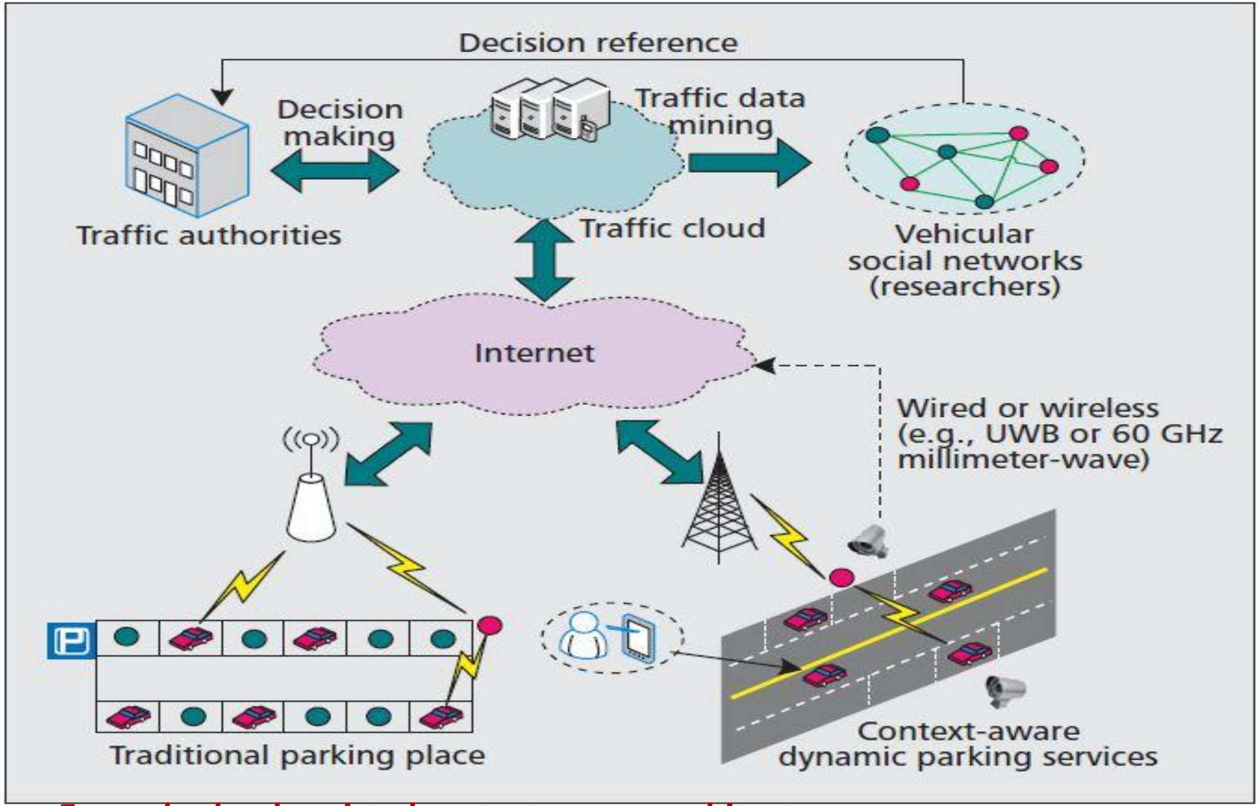
- The context information of each parking space detected by a WSN is forwarded to the traffic cloud by WSNs, third-generation (3G) communications, and the Internet.
- The collected data are processed in the cloud and then selectively transmitted to the users.
- This is helpful for providing more convenience services and evaluating the utilization levels of the parking garage.
- Also, the status of the parking garage may be dynamically published on a nearby billboard to users who have no ability to get the status by smart terminals.



Example cloud-assisted context-aware architecture

Dynamic parking services:

- In this scenario, we consider a situation in which we may temporarily park a vehicle along the road if it does not impede the passage of other vehicles or pedestrians.
- We envision this application scenario based on the common observation that the traffic flow capacity is usually regular for each road. For example, there is usually heavy traffic during morning and evening rush hours.
- Therefore, considering the context information such as rush hours and road conditions, we may dynamically arrange the parking services for a very wide road.
- With the support of many new technologies (e.g., MCC and WSNs), the traffic authorities can carry out the dynamic management of this kind of service.



Example cloud-assisted context-aware architecture

A Case Study: Context Aware Dynamic Parking Service

Three aspects, including service planning of traffic authorities, reservation service process, and context-aware optimization have been studied.

Decision making of traffic authorities

- The decision-making process of the proposed scheme heavily depends on many factors, such as historical traffic flow capacity, road conditions, weather conditions, and traffic flow forecasting

- In order to make an effective prediction, researchers on vehicular social networks carry out traffic data mining to discover useful information and knowledge from collected big data. The prediction process depends on classifying the influence factors and designing a decision tree

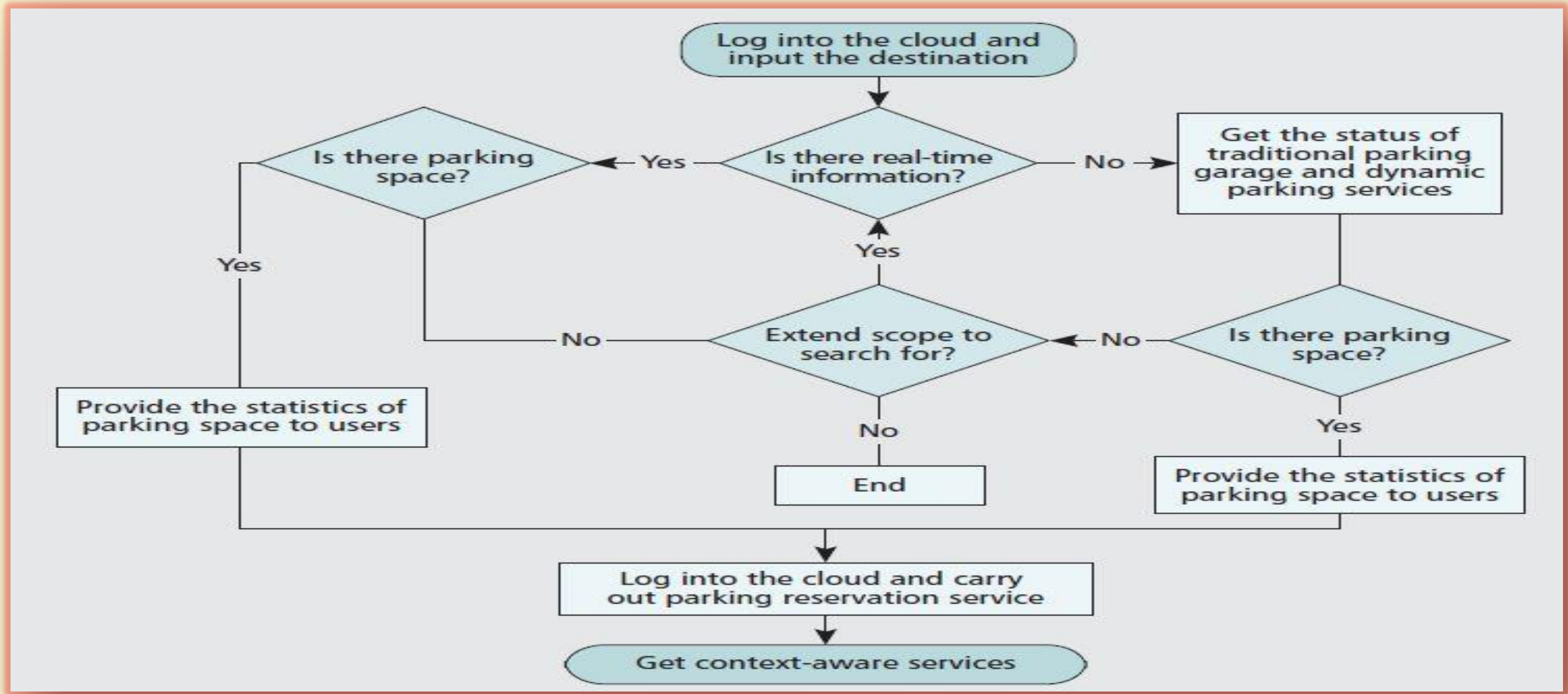
- By the method of probability analysis, the traffic authorities dynamically arrange whether the road can be authorized to provide context-aware parking services. In some particular cases, a fatal factor may directly affect the decision making. For example, when a typhoon is approaching, traffic authorities may immediately terminate services

A Case Study: Context Aware Dynamic Parking Service

Parking reservation services:

- The status of a parking space can be monitored as determined by the corresponding system, and subsequently updated in the traffic cloud.
- The drivers or passengers can quickly obtain the parking space’s information by various smart terminals such as smartphones. If a proper parking space cannot be found, further search scope is extended.
- Within a given time, we may log into the traffic cloud and subscribe to a parking space.

A Case Study: Context Aware Dynamic Parking Service



A Case Study: Context Aware Dynamic Parking Service

Context-aware optimization:

- The context information includes not only road conditions and the status of the parking garage, but also the expected duration of parking as well.
- Since the purpose of a visit to the place in question can determine the expected duration of parking, this context information can be used to optimize the best parking locations for drivers.
- For the parked vehicles, the expected duration of parking can be uploaded to the traffic cloud and shared with potential drivers after analysis.
- In this way, even when the parking garage has no empty parking spaces available, drivers still can inquire as to the status of the parking garage and get the desired service by context-aware optimization.
- The proposed context-aware dynamic parking service is a promising solution for alleviating parking difficulties and improving the QoS of CVC. Many technologies such as WSNs, traffic clouds, and traffic data mining are enabling this application scenario to become a reality

Summary

- Mobile cloud computing is one of the mobile technology trends in the future because it combines the advantages of both MC and CC, thereby providing optimal services for mobile users
- MCC focuses more on user experience : Lower battery consumption , Faster application execution
- MCC architectures design the middleware to partition an application execution transparently between mobile device and cloud servers
- The applications supported by MCC including m-commerce, mlearning, and mobile healthcare show the applicability of the MCC to a wide range.
- The issues and challenges for MCC (i.e., from communication and computing sides) demonstrates future research avenues and directions.

References

- Dinh, Hoang T., et al. “A survey of mobile cloud computing: architecture, applications, and approaches.” Wireless communications and mobile computing 13.18 (2013): 1587-1611
- *Z. Li, C. Wang, and R. Xu, “Computation offloading to save energy on handheld devices: a partition scheme,” in Proc 2001 Intl Conf on Compilers, architecture, and synthesis for embedded systems (CASES), pp. 238-246, Nov 2001.*
- *K. Kumar and Y. Lu, “Cloud Computing for Mobile Users: Can Offloading Computation Save Energy,” IEEE Computer, vol. 43, no. 4, April 2010*
- *H. H. La and S. D. Kim, “A Conceptual Framework for Provisioning Context-aware Mobile Cloud Services,” in Proceedings of IEEE International Conference on Cloud Computing (CLOUD), pp. 466, August 2010*
- Gordon, Mark S., et al. "COMET: Code Offload by Migrating Execution Transparently." OSDI. 2012.
- Yang, Seungjun, et al. "Fast dynamic execution offloading for efficient mobile cloud computing." Pervasive Computing and Communications (PerCom), 2013 IEEE International Conference on. IEEE, 2013
- *Shiraz, Muhammad, et al. "A review on distributed application processing frameworks in smart mobile devices for mobile cloud computing."Communications Surveys & Tutorials, IEEE 15.3 (2013): 1294-1313*
- *https://www.ibm.com/cloud-computing/learn-more/what-is-mobile-cloud-computing/*

Thank You!!

IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

NPTEL

CLOUD COMPUTING

Mobile Cloud Computing - II

Prof. Soumya K Ghosh

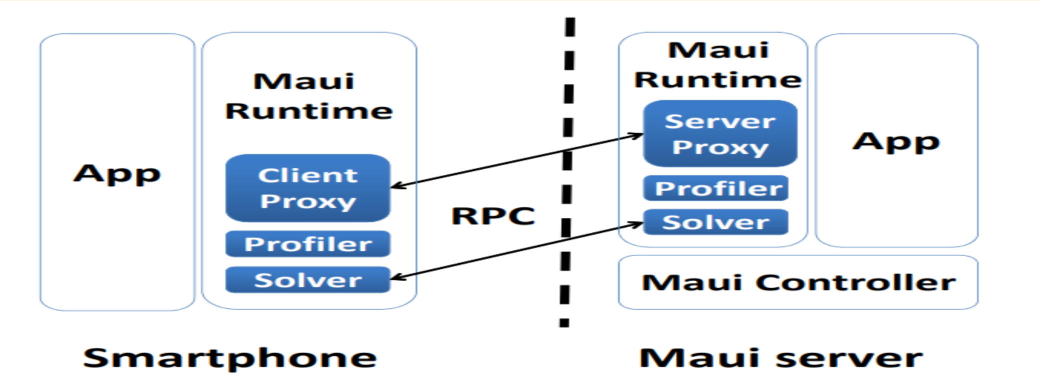
Department of Computer Science and Engineering

IIT KHARAGPUR

Mobile Cloud Computing (MCC) - Key challenges

- MCC requires dynamic partitioning of an application to optimize
  - Energy saving
  - Execution time
- Requires a software (middleware) that decides at app launch which parts of the application must execute on the mobile device, and which parts must execute on cloud
  - A classic optimization problem

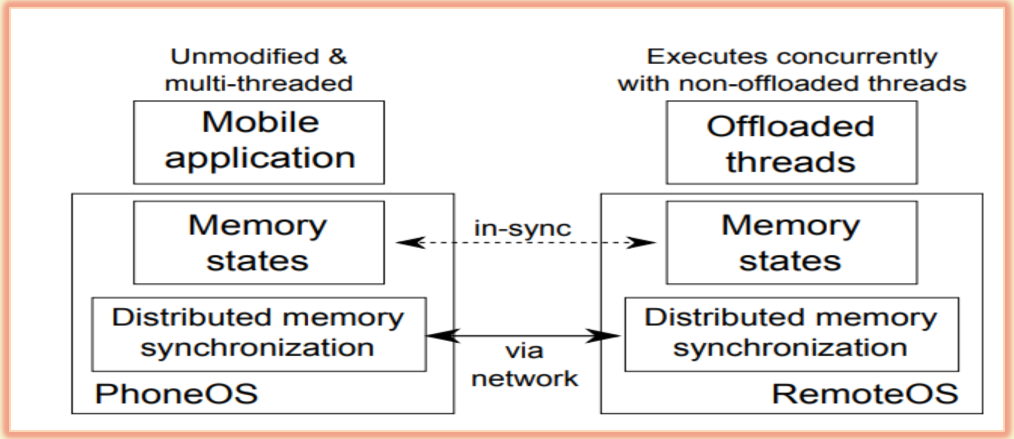
MCC Systems: MAUI (Mobile Assistance Using Infrastructure)



- ***MAUI enables the programmer to produce an initial partition of the program***
  - Programmer marks each method as “remoteable” or not
  - Native methods cannot be remoteable
- MAUI framework uses the annotation to decide whether a method should be executed on cloud server to save energy and time to execute
- ***MAUI server is the cloud component***. The framework has the necessary software modules required in the workflow.

Cuervo, Eduardo, et al. "MAUI: making smartphones last longer with code offload." Proceedings of Mobile systems, applications, and services. ACM, 2010.[MobiSys]

MCC Systems: COMET



- ***Requires only program binaries Execute multi-threaded programs correctly Improve speed of computation***
- ***Further improvements to data traffic during migration is also possible by sending only the parts of the heap that has been modified***

COMET: Code Offload by Migrating Execution Transparently

- Works on unmodified applications (no source code required)
- Allows threads to migrate between machines depending on workload
- It implements a Distributed Shared Memory (DSM) model for the runtime engine
  - ✓ *DSM allows transparent movement of threads across machines*
  - ✓ *In computer architecture, DSM is a form of memory architecture where the (physically separate) memories can be addressed as one (logically shared) address space*

Gordon, Mark S., et al. "COMET: Code Offload by Migrating Execution Transparently." OSDI. Vol. 12. 2012.

Key Problems to Solve

- At its core, MCC framework must solve how to partition a program for execution on heterogeneous computing resources
- This is a classic “Task Partitioning Problem”
- Widely studied in processor resource scheduling as “job scheduling problem”

Task Partitioning Problem in MCC

Input:

- A call graph representing an application's method call sequence
- Attributes for each node in the graph denotes
  - (a) energy consumed to execute the method on the mobile device,
  - (b) energy consumed to transfer the program states to a remote server

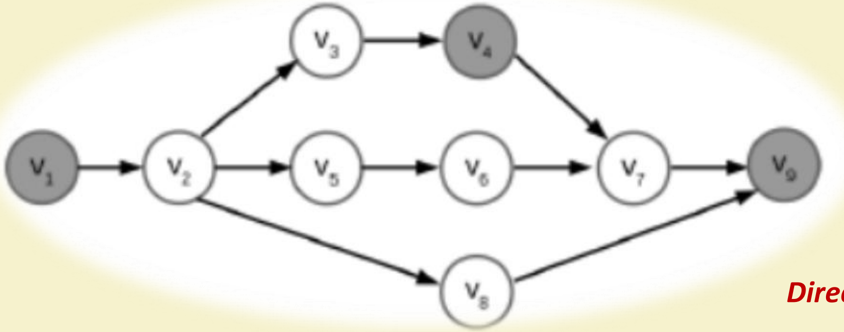
Output:

- Partition the methods into two sets – one set marks the methods to execute on the mobile device, and the second set marks the methods to execute on cloud

Goals and Constraints:

1. Energy consumed must be minimized
2. There is a limit on the execution time of the application
3. Other constraints could be – some methods must be executed on mobile device, total monetary cost, etc.

Mathematical Formulation



- Highlighted nodes must be executed on the mobile device -> called native tasks (v1, v4, v9)
- Edges represent the sequence of execution - Any non-highlighted node can be executed either locally on the mobile device or on cloud

Directed Acyclic Graph represents an application Call Graph

$$\text{maximize} \sum_{v \in V} I_v \times E_v^l - \sum_{(u,v) \in E} |I_u - I_v| \times C_{u,v}$$

$$\text{such that:} \sum_{v \in V} ((1 - I_v) \times T_v^l) + (I_v \times T_v^r))$$

$$+ \sum_{(u,v) \in E} (|I_u - I_v| \times B_{u,v}) \leq L$$

$$\text{and} \qquad I_v \leq r_v, \forall v \in V$$

- 0-1 integer linear program,
  where I_v = 0 if method executed locally,
  = 1 if method executed remotely
- E : Energy cost to execute method v locally
- $C_{u,v}$: Cost of data transfer
- L : Total execution latency
- T : Time to execute the method
- B : Time to transfer program state

Integer Linear Program to solve the Task Partitioning Problem

Static and Dynamic Partitioning

- Static Partitioning
  - When an application is launched, invoke an ILP solver which will tell where each method should be executed
  - There are also heuristics to find solutions faster
- Dynamic or Adaptive Partitioning
  - For a long running program, the environmental conditions can vary
  - Depending on the input, the energy consumption of a method can vary

Mobile Cloud Computing – Challenges/ Issues

Mobile communication issues

- *Low bandwidth*: One of the biggest issues, because the radio resource for wireless networks is much more scarce than wired networks
- *Service availability*: Mobile users may not be able to connect to the cloud to obtain a service due to traffic congestion, network failures, mobile signal strength problems
- *Heterogeneity*: Handling wireless connectivity with highly heterogeneous networks to satisfy MCC requirements (always-on connectivity, on-demand scalability, energy efficiency) is a difficult problem

Computing issues (Computation offloading)

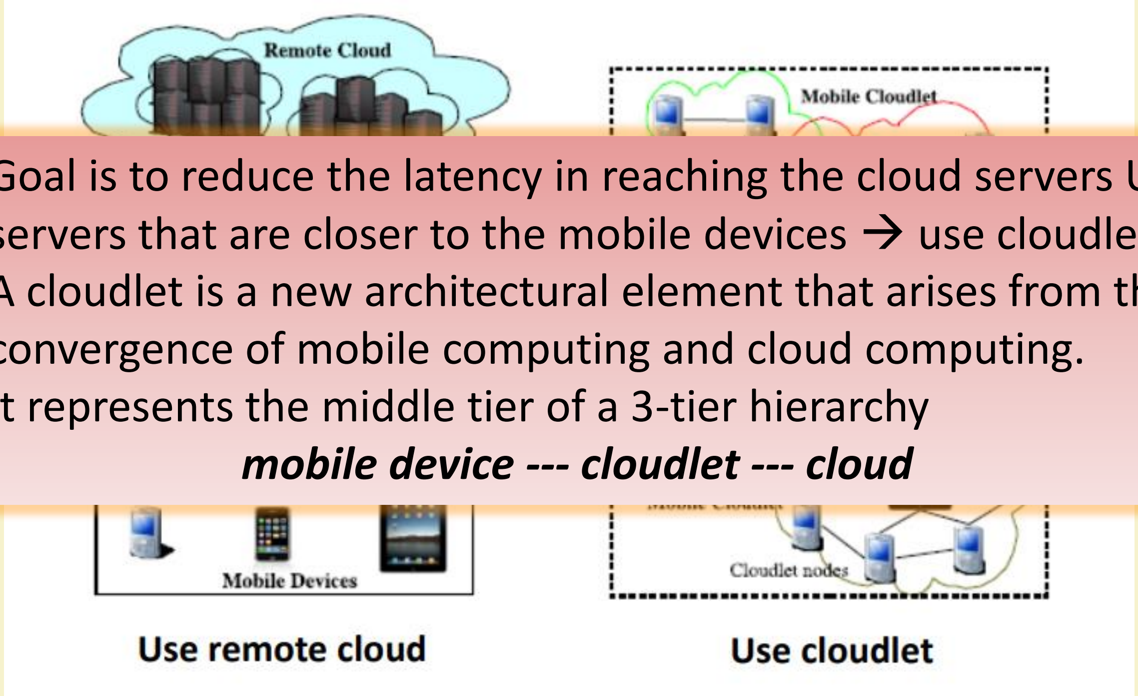
- One of the main features of MCC
- Offloading is not always effective in saving energy
- It is critical to determine whether to offload and which portions of the service codes to offload

CODE OFFLOADING USING CLOUDLET

- **CLOUDLET**:
  - ✓ *"a trusted, resource*-rich *computer or cluster of computers that is well-connected to the Internet and is available for use by nearby mobile devices."*
- **Code Offloading** :
  - ✓ Offloading the code to the remote server and executing it.
  - ✓ This architecture decreases latency by using a single-hop network and potentially lowers battery consumption by using Wi-Fi or short-range radio instead of broadband wireless which typically consumes more energy.

CODE OFFLOADING USING CLOUDLET

Cloudlet



- Goal is to reduce the latency in reaching the cloud servers Use servers that are closer to the mobile devices → use cloudlet
- A cloudlet is a new architectural element that arises from the convergence of mobile computing and cloud computing.
- It represents the middle tier of a 3-tier hierarchy

  mobile device --- cloudlet --- cloud

When to Offload ?

Amount of energy saved is :

$$P_c \times \frac{C}{M} - P_i \times \frac{C}{S} - P_{tr} \times \frac{D}{B}$$

S: *Speed of cloud to compute **C** instructions*
M: *Speed of mobile to compute **C** instructions*
D: *Data need to transmit*
B: *Bandwidth of the wireless Internet*
P_c: *Energy cost per second when the mobile phone is doing computing*
P_i: *Energy cost per second when the mobile phone is idle.*
P_{tr}: *Energy cost per second when the mobile is transmitting the data.*

Suppose the server is F times faster—
S= F × M.

We can rewrite the formula as

$$\frac{C}{M} \times \left(Pc - \frac{P_i}{F}\right) - P_{tr} \times \frac{D}{B}$$

When to Offload? (contd..)

- Energy is saved when the formula produces a positive number. The formula is positive if D/B is sufficiently small compared with C/M and F is sufficiently large.
- Cloud computing can potentially save energy for mobile users.
- Not all applications are energy efficient when migrated to the cloud.
- Cloud computing services would be significantly different from cloud services for desktops because they must offer energy savings.
- The services should consider the energy overhead for privacy, security, reliability, and data communication before offloading.

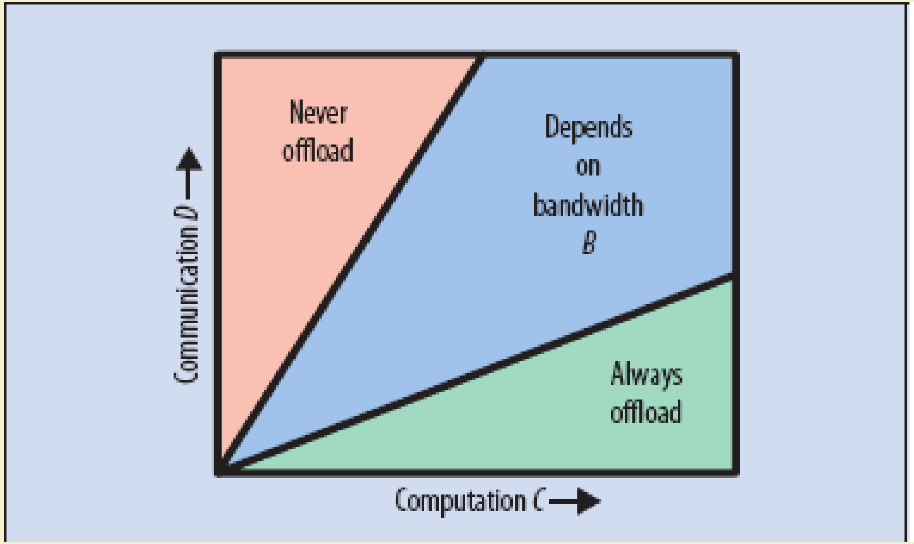
The amount of energy saved is :

$$P_c \times \frac{C}{M} - P_i \times \frac{C}{S} - P_{tr} \times \frac{D}{B}$$

We can rewrite the formula as

$$\frac{C}{M} \times \left(Pc - \frac{P_i}{F}\right) - P_{tr} \times \frac{D}{B}$$

When to Offload?? (contd..)



The amount of energy saved is :

$$P_c \times \frac{C}{M} - P_i \times \frac{C}{S} - P_{tr} \times \frac{D}{B}$$

We can rewrite the formula as

$$\frac{C}{M} \times \left(Pc - \frac{P_i}{F}\right) - P_{tr} \times \frac{D}{B}$$

Offloading is beneficial when large amounts of computation C are needed with relatively small amounts of communication D

Computation Offloading Approaches

- Partition a program based on estimation of energy consumption before execution

- Optimal program partitioning for offloading is dynamically calculated based on the trade-off between the communication and computation costs at run time.

- Offloading scheme based on profiling information about computation time and data sharing at the level of procedure calls.
    - A cost graph is constructed and a branch-and-bound algorithm is applied to minimize the total energy consumption of computation and the total data communication cost.

Z. Li, C. Wang, and R. Xu, "Computation offloading to save energy on handheld devices: a partition scheme," in Proc 2001 Intl Conf on Compilers, architecture, and synthesis for embedded systems (CASES), pp. 238-246, Nov 2001.

K. Kumar and Y. Lu, "Cloud Computing for Mobile Users: Can Offloading Computation Save Energy," IEEE Computer, vol. 43, no. 4, April 2010

How to evaluate MCC performance

- Energy Consumption
  - Must reduce energy usage and extend battery life
- Time to Completion
  - Should not take longer to finish the application compared to local execution
- Monetary Cost
  - Cost of network usage and server usage must be optimized
- • Security
  - As offloading transfers data to the servers, ensure confidentiality and privacy of data, how to identify methods which process confidential data

Challenges

- How can one design a practical and usable MCC framework
  - System as well as partitioning algorithm
- Is there a scalable algorithm for partitioning
  - Optimization formulations are NP-hard
  - Heuristics fail to give any performance guarantee
- Which are the most relevant parameters to consider in the design of MCC systems?

Mobile Cloud Computing – Applications?

Mobile Health-care

Health-Monitoring services, Intelligent emergency management system, Health-aware mobile devices (detect pulse rate, blood pressure, alcohol-level etc.)

Mobile Gaming

It can completely offload game engine requiring large computing resource (e.g., graphic rendering) to the server in the cloud

Mobile Commerce

M-commerce allows business models for commerce using mobile (Mobile financial, mobile advertising, mobile shopping)

Mobile Cloud Computing – Applications?

Pedestrian crossing guide for blind and visually-impaired

Mobile currency reader for blind and visually impaired

Lecture transcription for hearing impaired students

Assistive Technologies

Mobile Learning

- *M-learning combines e-learning and mobility*
- *Traditional m-learning has limitations on high cost of devices/network, low transmission rate, limited educational resources*
- *Cloud-based m-learning can solve these limitations*
- *Enhanced communication quality between students and teachers*
- *Help learners access remote learning resources*
- *A natural environment for collaborative learning*

MuSIC: Mobility-Aware Optimal Service Allocation in Mobile Cloud Computing

- User Mobility introduces new complexities in enabling an optimal decomposition of tasks that can execute cooperatively on mobile clients and the tiered cloud architecture while considering multiple QoS goals such application delay, device power consumption and user cost/price.
- Apart from scalability and access issues with the increased number of users, mobile applications are faced with increased *latencies* and reduced *reliability*
- As a user moves, the physical distance between the user and the cloud resources originally provisioned changes causing additional delays
- Further, the lack of effective handoff mechanisms in WiFi networks as user move rapidly causes an increase in the number of *packet losses*

In other words, user mobility, if not addressed properly, can result in suboptimal resource mapping choices and ultimately in diminished application QoS

Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

MuSIC: Mobility-Aware Optimal Service Allocation in Mobile Cloud Computing

Efficient techniques for *dynamic mapping of resources* in the presence of **mobility**; using a *tiered cloud architecture*, to meet the *multidimensional QoS* needs of mobile users

- Location-time workflow (LTW) as the modeling framework to model mobile applications and capture user mobility. Within this framework, mobile service usage patterns as a function of location and time has been formally modelled

- Given a mobile application execution expressed as a LTW, the framework optimally partitions the execution of the location-time workflow in the 2-tier architecture based on a *utility metric* that combines *service price, power consumption and delay* of the mobile applications

Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

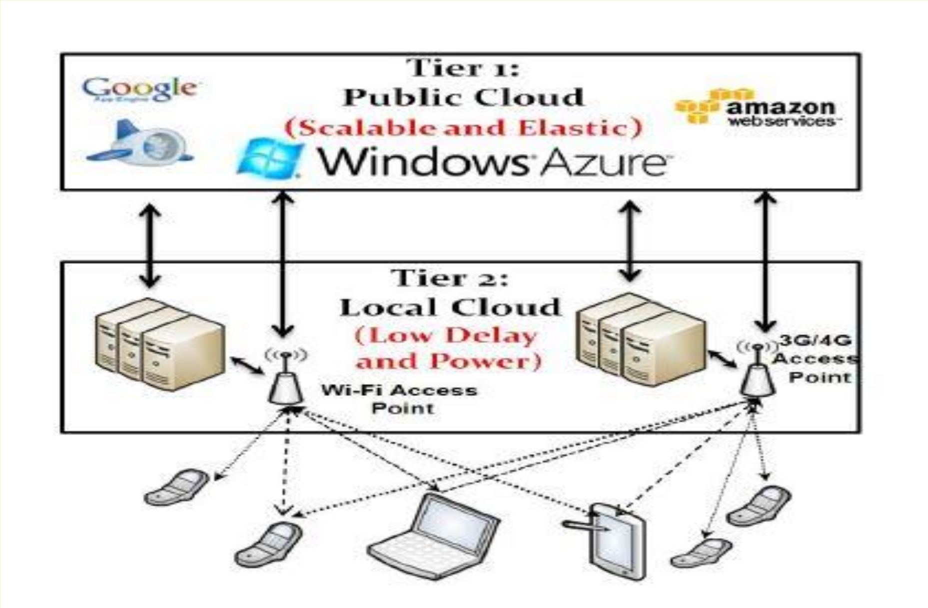
MuSIC: Mobility-Aware Optimal Service Allocation in Mobile Cloud Computing

- ✓ Tier 1 nodes in the system architecture represents *public cloud services* such as Amazon EC2, Microsoft Azure and Google AppEngine. Services provided by these vendors are highly *scalable* and *available*; what they lack is the ability to provide the *fine grain location granularity* required for high performance mobile applications

- ✓ This feature is provided by the second tier local cloud, that consists of nodes that are connected to access points.

- ✓ Location information of these services are available at finer levels of granularity (campus and street level).

- ✓ Mobile users are typically connected to these local clouds through WiFi (via access points) or cellular (via 3G cell towers) connectivity - the aim to intelligently select which local and which public cloud resources to utilize for task offloading.



2-Tier Mobile Cloud Architecture

Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

Mobile Application Modelling

Cloud Service Set:
The set of all services (e.g. compute, storage and software capabilities like multimedia streaming services, content transcoding services, etc) provided by local and public cloud providers

Local Cloud Capacity:
Local cloud services can only accept a limited number of mobile client requests

Location Map:
It is a partition of the 2-D space/region in which mobile hosts and cloud resources are located

User Service Set:
The set of all services that a user has on his own device (e.g. decoders, image editors etc.)

Criteria	Definition
$q_{price}(s_i, u_k^{l_i,t_j})$	The price of using service s_i when user u_k is in location $l_i \in L$ and time t_j.
$q_{power}(s_i, u_k^{l_i,t_j})$	The power consumed on user mobile device using s_i when user u_k is in location $l_i \in L$ and time t_j.
$q_{delay}(s_i, u_k^{l_i,t_j})$	The delay of executing service s_i when user u_k is in location $l_i \in L$ and time t_j.

Mobile User Trajectory:

The trajectory of a mobile user, u_k, is represented as a list of tuples of the form $\{(1; l_k); \ldots; (n; l_m)\}$ where $(i; l_j)$ implies that the mobile user is in location l_j for time duration I

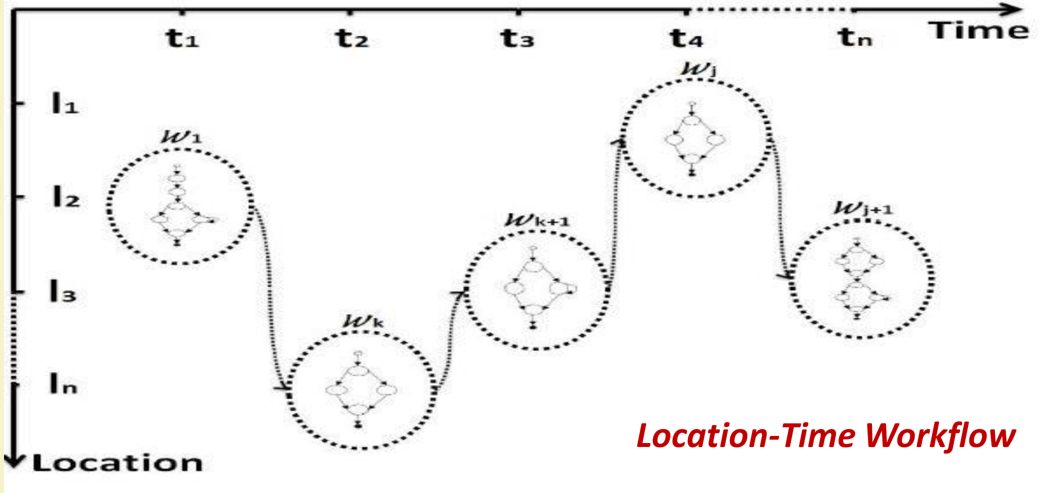
Center of Mobility:
It is the location where (or near where) a mobile user u_k spends most of its time

Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

Mobile Application Modelling

Location-Time Workflow:

Combination of the mobile application workflow concept with a user trajectory to model the mobile users and the requested services in their trajectory.



Location-Time Workflow

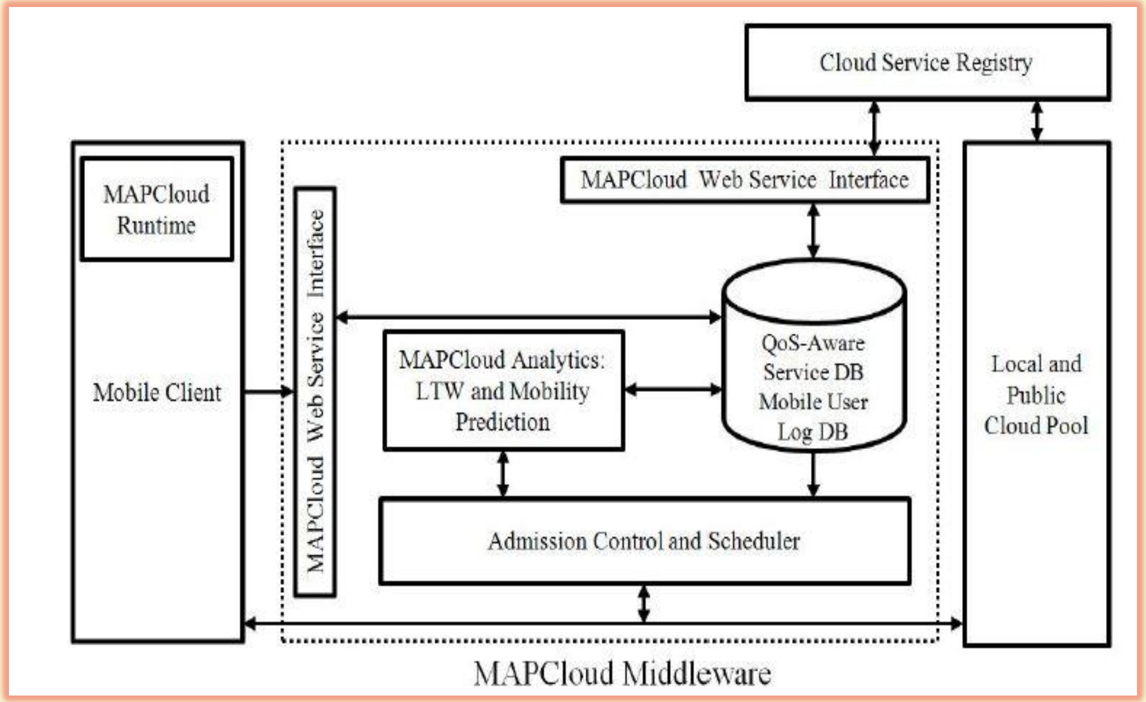
Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

Mobile Application Modelling

Mobile User Log DB and QoS-Aware Service DB:
Unprocessed user data log such as mobile service usage, location of the user, user delay experience of getting
the service, energy consumed on user mobile device, etc and service lists on local and public cloud and their QoSes in different locations respectively

MAPCloud Analytic: This module processes mobile user Log DB and updates QoS-aware cloud service DB based on user experience and LTW

Admission Control and Scheduling: This module is responsible for optimally allocate services to admitted mobile users based on MuSIC



Rahimi, M. Reza et al. "MuSIC: Mobility-aware optimal service allocation in mobile cloud computing." Cloud Computing (CLOUD), 2013 IEEE Sixth International Conference on. IEEE, 2013.

A Case Study: Context Aware Dynamic Parking Service

- MCC can provide a flexible method of handling massive computing, storage, and software services in a scalable and virtualized manner.
- The integration of MCC and vehicular networks is expected to promote the development of cost effective, scalable, and data-driven CVC (Context-aware vehicular cyber physical systems)

An application scenario regarding the context-aware dynamic parking services by illuminating the cloud-assisted architecture and logic flow.

- **As the number of vehicles increases, there is an increasing trend of insufficient parking spaces in many large cities, and this problem is gradually getting worse**
- **With the proliferation of wireless sensor networks (WSNs) and cloud computing, there exists strong potential to alleviate this problem using context information (e.g., road conditions and status of parking garages) to provide context-aware dynamic parking services**
- **Cloud Assisted parking services (traditional parking garages and dynamic parking services along the road) and parking reservation service using smart terminals such as smartphones.**

Wan, Jiafu, et al. "Context-aware vehicular cyber-physical systems with cloud support: architecture, challenges, and solutions." *IEEE Communications Magazine* 52.8 (2014): 106-113.

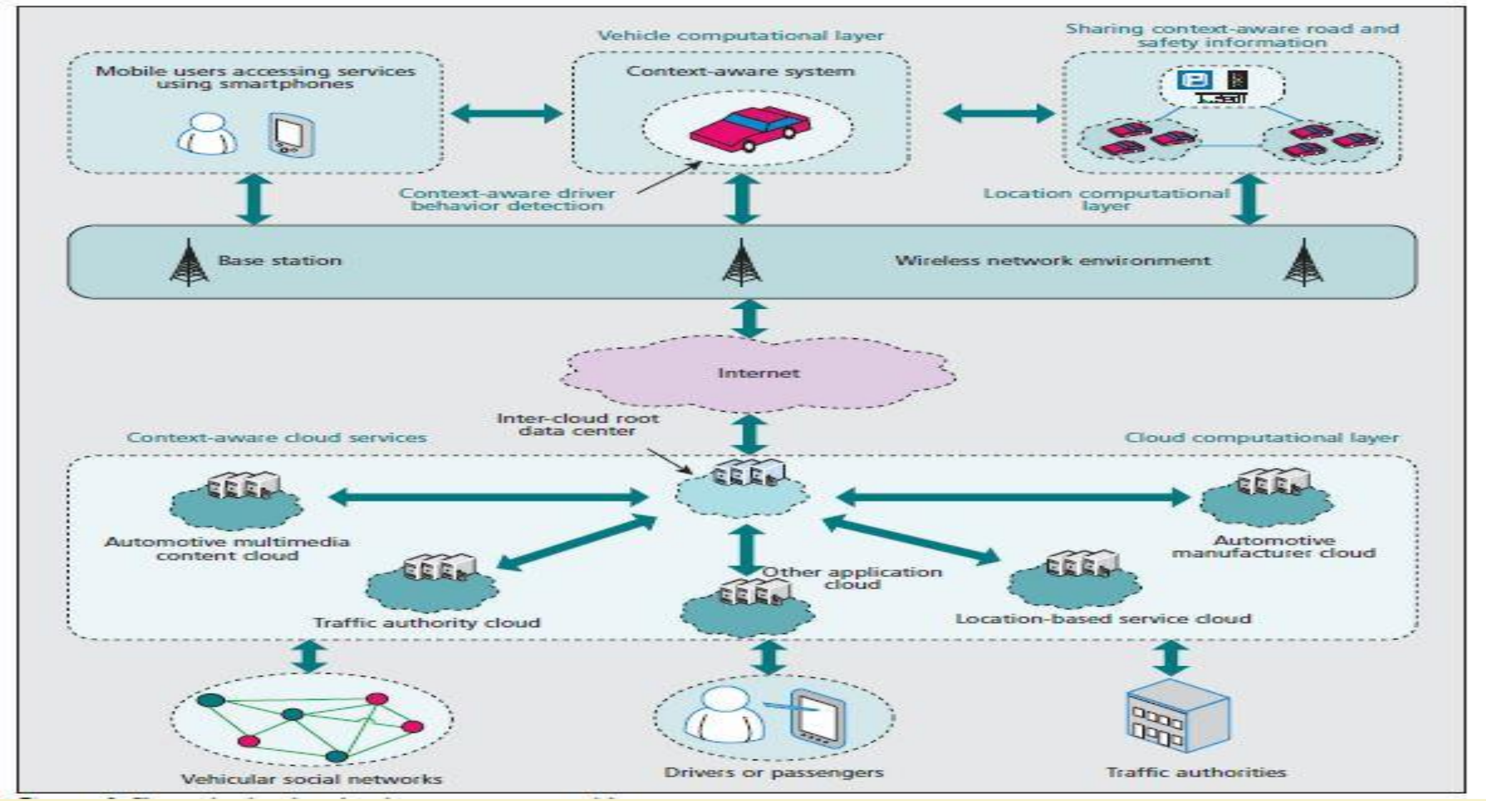
A Case Study: Context Aware Dynamic Parking Service

- MCC can provide a flexible method of handling massive computing, storage, and software services in a scalable and virtualized manner.
- The integration of MCC and vehicular networks is expected to promote the development of cost effective, scalable, and data-driven CVC (Context-aware vehicular cyber physical systems)

An application scenario regarding the context-aware dynamic parking services by illuminating the cloud-assisted architecture and logic flow.

- **As the number of vehicles increases, there is an increasing trend of insufficient parking spaces in many large cities, and this problem is gradually getting worse**
- **With the proliferation of wireless sensor networks (WSNs) and cloud computing, there exists strong potential to alleviate this problem using context information (e.g., road conditions and status of parking garages) to provide context-aware dynamic parking services**
- **Cloud Assisted parking services (traditional parking garages and dynamic parking services along the road) and parking reservation service using smart terminals such as smartphones.**

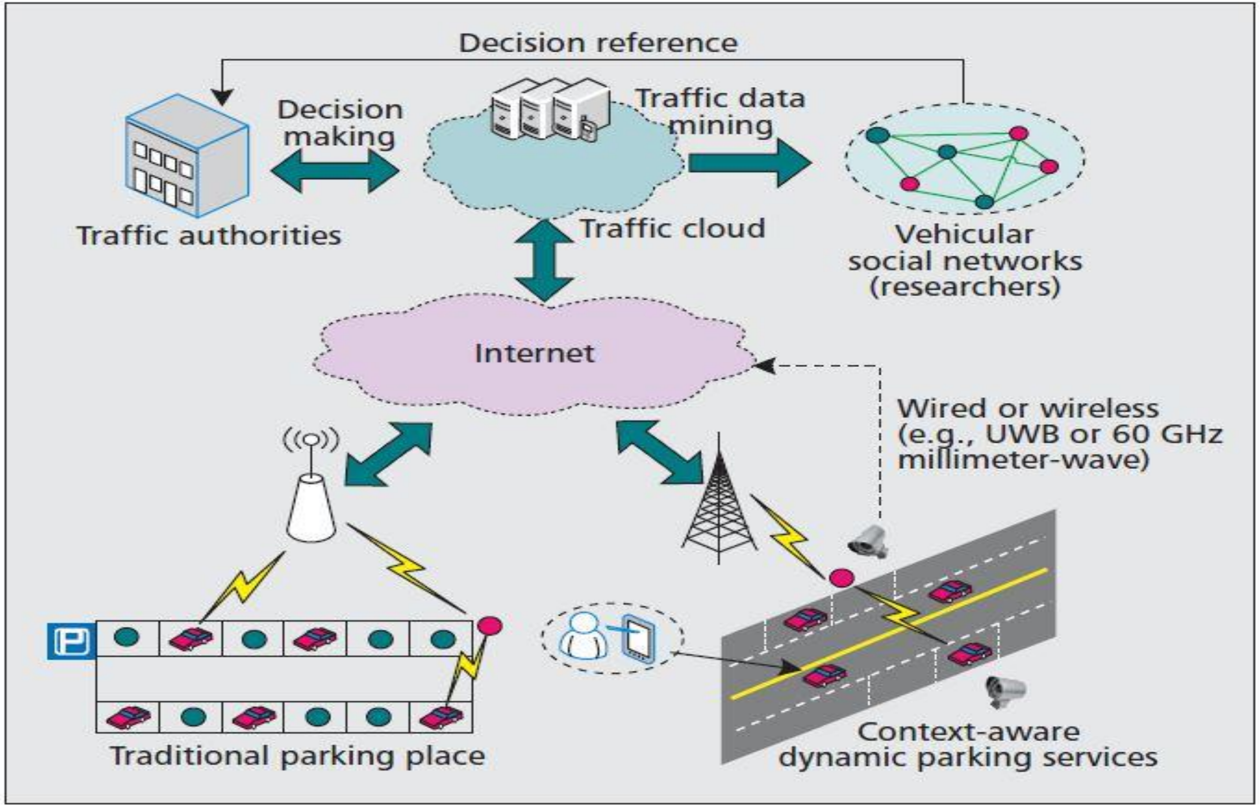
Wan, Jiafu, et al. "Context-aware vehicular cyber-physical systems with cloud support: architecture, challenges, and solutions." *IEEE Communications Magazine* 52.8 (2014): 106-113.



Example cloud-assisted context-aware architecture

Traditional parking garages:

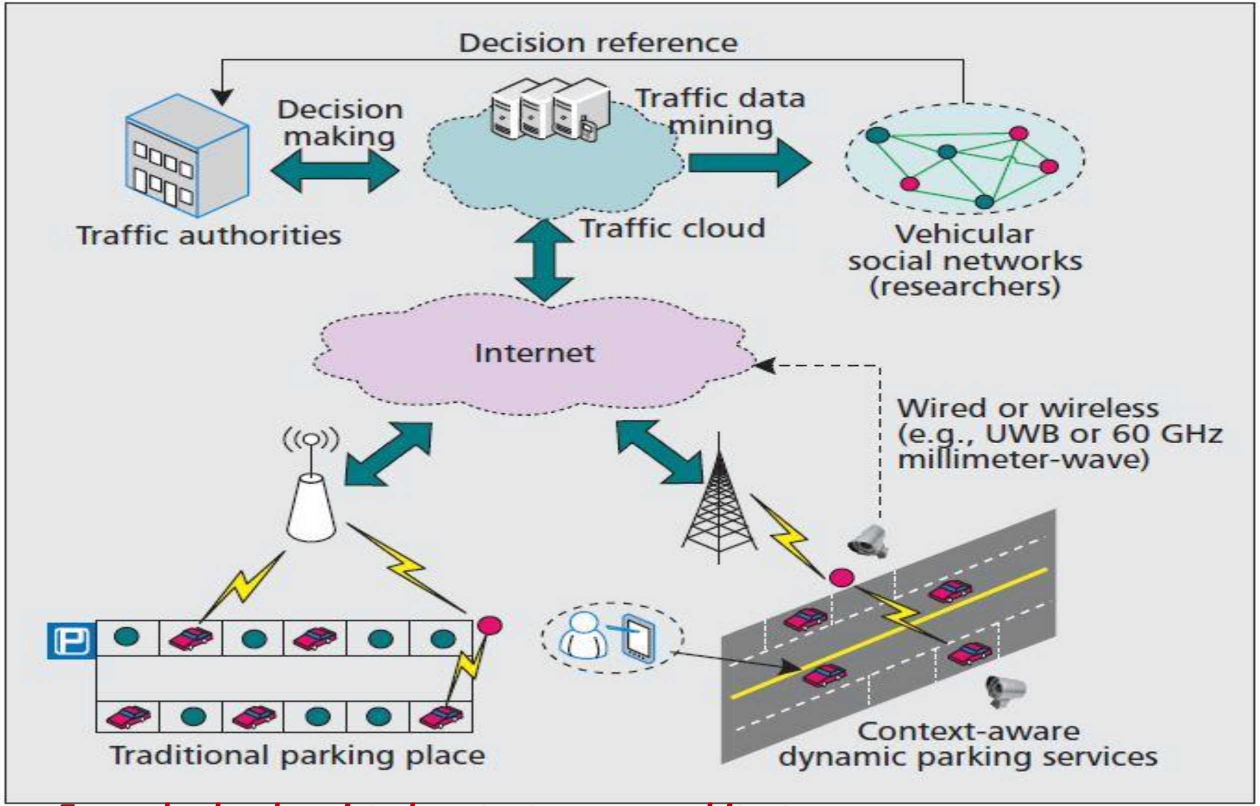
- The context information of each parking space detected by a WSN is forwarded to the traffic cloud by WSNs, third-generation (3G) communications, and the Internet.
- The collected data are processed in the cloud and then selectively transmitted to the users.
- This is helpful for providing more convenience services and evaluating the utilization levels of the parking garage.
- Also, the status of the parking garage may be dynamically published on a nearby billboard to users who have no ability to get the status by smart terminals.



Example cloud-assisted context-aware architecture

Dynamic parking services:

- In this scenario, we consider a situation in which we may temporarily park a vehicle along the road if it does not impede the passage of other vehicles or pedestrians.
- We envision this application scenario based on the common observation that the traffic flow capacity is usually regular for each road. For example, there is usually heavy traffic during morning and evening rush hours.
- Therefore, considering the context information such as rush hours and road conditions, we may dynamically arrange the parking services for a very wide road.
- With the support of many new technologies (e.g., MCC and WSNs), the traffic authorities can carry out the dynamic management of this kind of service.



Example cloud-assisted context-aware architecture

A Case Study: Context Aware Dynamic Parking Service

Three aspects, including service planning of traffic authorities, reservation service process, and context-aware optimization have been studied.

Decision making of traffic authorities

- The decision-making process of the proposed scheme heavily depends on many factors, such as historical traffic flow capacity, road conditions, weather conditions, and traffic flow forecasting

- In order to make an effective prediction, researchers on vehicular social networks carry out traffic data mining to discover useful information and knowledge from collected big data. The prediction process depends on classifying the influence factors and designing a decision tree

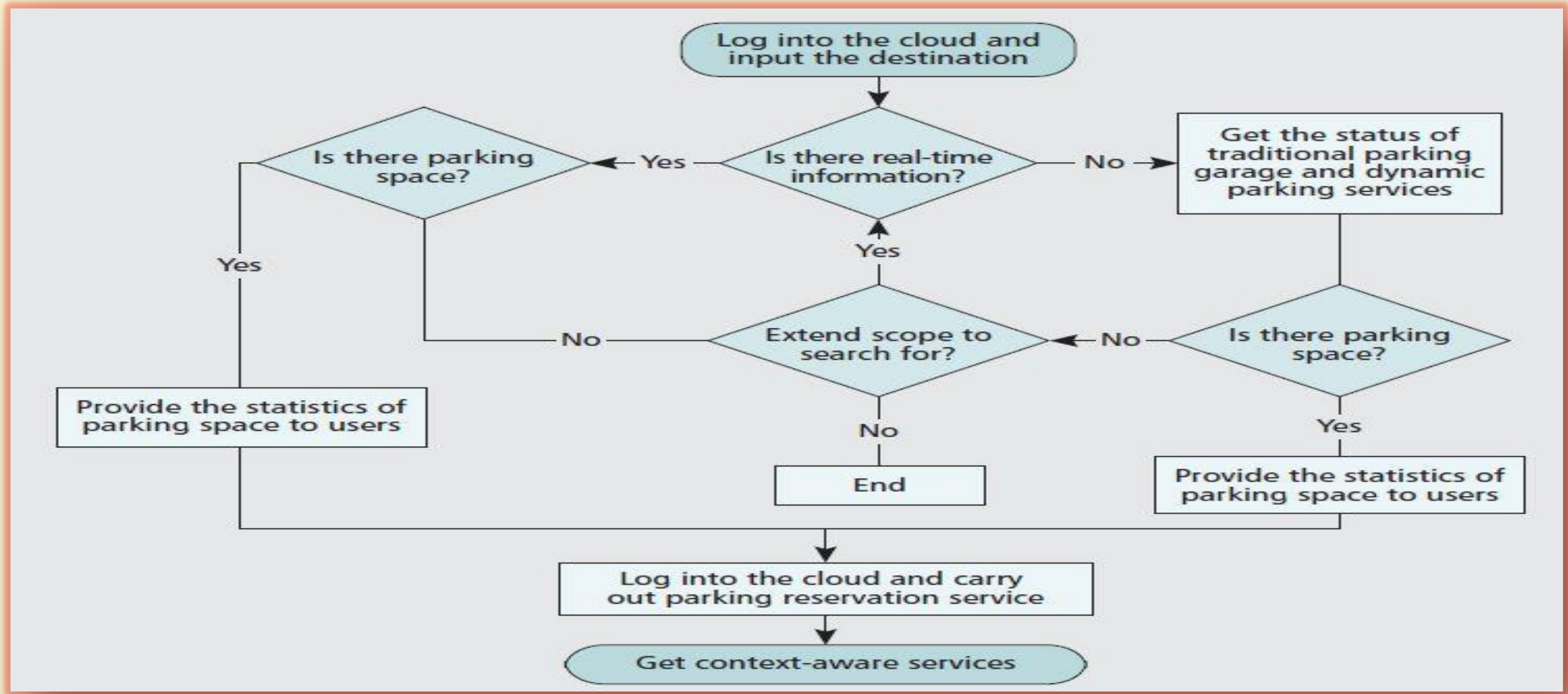
- By the method of probability analysis, the traffic authorities dynamically arrange whether the road can be authorized to provide context-aware parking services. In some particular cases, a fatal factor may directly affect the decision making. For example, when a typhoon is approaching, traffic authorities may immediately terminate services

A Case Study: Context Aware Dynamic Parking Service

Parking reservation services:

- The status of a parking space can be monitored as determined by the corresponding system, and subsequently updated in the traffic cloud.
- The drivers or passengers can quickly obtain the parking space's information by various smart terminals such as smartphones. If a proper parking space cannot be found, further search scope is extended.
- Within a given time, we may log into the traffic cloud and subscribe to a parking space.

A Case Study: Context Aware Dynamic Parking Service



A Case Study: Context Aware Dynamic Parking Service

Context-aware optimization:

- The context information includes not only road conditions and the status of the parking garage, but also the expected duration of parking as well.
- Since the purpose of a visit to the place in question can determine the expected duration of parking, this context information can be used to optimize the best parking locations for drivers.
- For the parked vehicles, the expected duration of parking can be uploaded to the traffic cloud and shared with potential drivers after analysis.
- In this way, even when the parking garage has no empty parking spaces available, drivers still can inquire as to the status of the parking garage and get the desired service by context-aware optimization.
- The proposed context-aware dynamic parking service is a promising solution for alleviating parking difficulties and improving the QoS of CVC. Many technologies such as WSNs, traffic clouds, and traffic data mining are enabling this application scenario to become a reality

Summary

- Mobile cloud computing is one of the mobile technology trends in the future because it combines the advantages of both MC and CC, thereby providing optimal services for mobile users
- MCC focuses more on user experience : Lower battery consumption , Faster application execution
- MCC architectures design the middleware to partition an application execution transparently between mobile device and cloud servers
- The applications supported by MCC including m-commerce, mlearning, and mobile healthcare show the applicability of the MCC to a wide range.
- The issues and challenges for MCC (i.e., from communication and computing sides) demonstrates future research avenues and directions.

References

- Dinh, Hoang T., et al. “A survey of mobile cloud computing: architecture, applications, and approaches.” Wireless communications and mobile computing 13.18 (2013): 1587-1611

- *Z. Li, C. Wang, and R. Xu, “Computation offloading to save energy on handheld devices: a partition scheme,” in Proc 2001 Intl Conf on Compilers, architecture, and synthesis for embedded systems (CASES), pp. 238-246, Nov 2001.*

- *K. Kumar and Y. Lu, “Cloud Computing for Mobile Users: Can Offloading Computation Save Energy,” IEEE Computer, vol. 43, no. 4, April 2010*

- *H. H. La and S. D. Kim, “A Conceptual Framework for Provisioning Context-aware Mobile Cloud Services,” in Proceedings of IEEE International Conference on Cloud Computing (CLOUD), pp. 466, August 2010*
- Gordon, Mark S., et al. "COMET: Code Offload by Migrating Execution Transparently." OSDI. 2012.

- Yang, Seungjun, et al. "Fast dynamic execution offloading for efficient mobile cloud computing." Pervasive Computing and Communications (PerCom), 2013 IEEE International Conference on. IEEE, 2013

- *Shiraz, Muhammad, et al. "A review on distributed application processing frameworks in smart mobile devices for mobile cloud computing."Communications Surveys & Tutorials, IEEE 15.3 (2013): 1294-1313*

- *https://www.ibm.com/cloud-computing/learn-more/what-is-mobile-cloud-computing/*

Thank You!!

IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

NPTEL

CLOUD COMPUTING

Fog Computing - I

PROF. SOUMYA K. GHOSH
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
IIT KHARAGPUR

Cloud Computing : Challenges

- Processing of huge data in a datacenter.
- Datacenter may be privately hosted by the organization (private cloud setup) or publicly available by paying rent (public cloud).
- All the necessary information has to be uploaded to the cloud for processing and extracting knowledge from it.

Cloud Computing – Typical Characteristics

- **Dynamic scalability**: Application can handle increasing load by getting more resources.
- **No Infrastructure Management by User:** Infrastructure is managed by cloud provider, not by end-user or application developer.
- **Metered Service**: Pay-as-you-go model. No capital expenditure for public cloud.

Issues with “Cloud-only” Computing

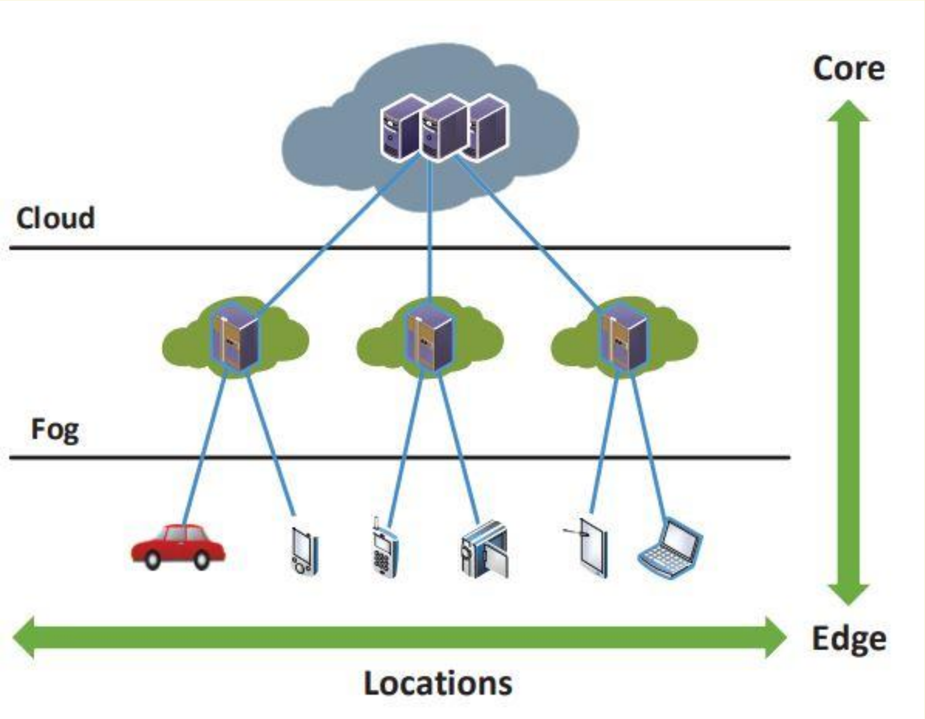
- Communication takes a long time due to human-smartphone interaction.
- Datacenters are centralized, so all the data from different regions can cause congestion in core network.
- Such a task requires very low response time, to prevent further crashes or traffic jam.

Fog Computing

- Fog computing, also known as fogging/edge computing, it is a model in which data, processing and applications are concentrated in devices at the network edge rather than existing almost entirely in the cloud.
- The term "Fog Computing" was introduced by the Cisco Systems as new model to ease wireless data transfer to distributed devices in the Internet of Things (IoT) network paradigm
- CISCO's vision of fog computing is to enable applications on billions of connected devices to run directly at the network edge.
  - Users can develop, manage and run software applications on Cisco framework of networked devices, including hardened routers and switches.
  - Cisco brings the open source Linux and network operating system together in a single networked device

Fog Computing

- Bringing intelligence down from the cloud close to the ground/ end-user.
- Cellular base stations, Network routers, WiFi Gateways will be capable of running applications.
- End devices, like sensors, are able to perform basic data processing.
- Processing close to devices lowers response time, enabling real-time applications.



Source: The Fog Computing Paradigm: Scenarios and Security Issues, Ivan Stojmenovic and Sheng Wen

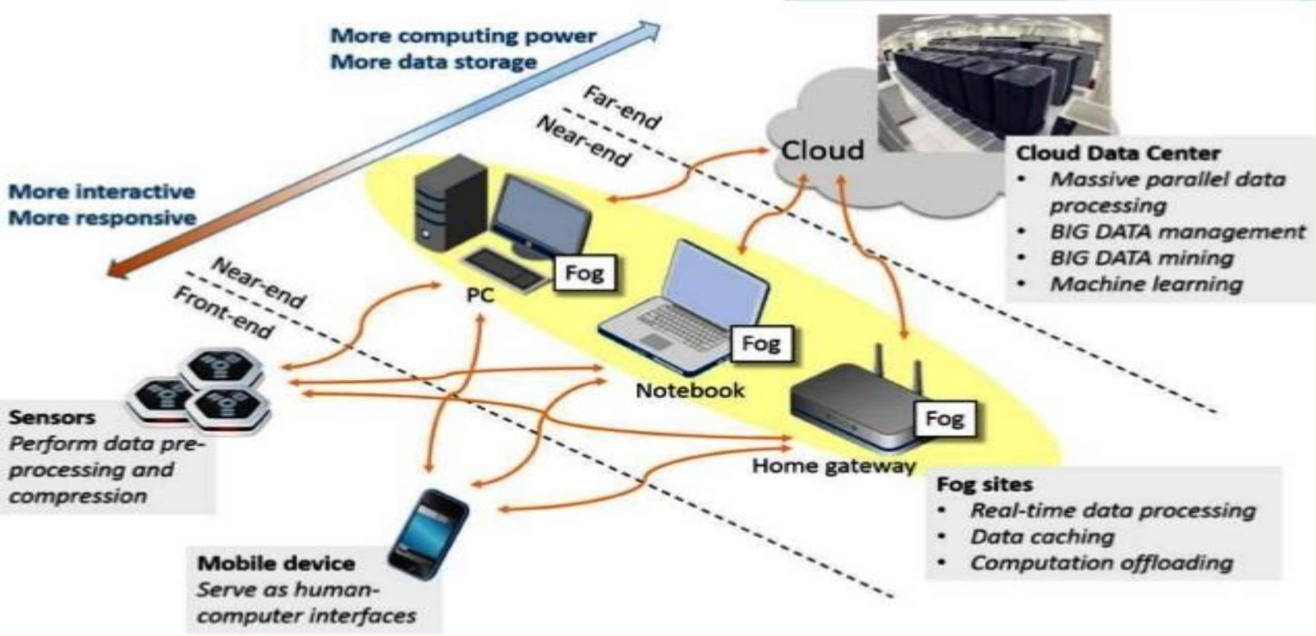
Fog Computing

- Fog computing enables some of transactions and resources at the edge of the cloud, rather than establishing channels for cloud storage and utilization.
- Fog computing reduces the need for bandwidth by not sending every bit of information over cloud channels, and instead aggregating it at certain access points.
- This kind of distributed strategy, may help in lowering cost and improve efficiencies.

Fog Computing - Motivation

- Fog Computing is a paradigm that extends Cloud and its services to the edge of the network
- Fog provides data, compute, storage and application services to the end-user
- Recent developments: Smart Grid, Start Traffic light, Connected Vehicles, Software defined network

Fog Computing



Source: Internet

Fog Computing Enablers

- **Virtualization :** Virtual machines can be used in edge devices.
- **Containers:** Reduces the overhead of resource management by using light-weight virtualizations. Example: *Docker* containers.
- **Service Oriented Architecture:** Service-oriented architecture (SOA) is a style of software design where services are provided to the other components by application components, through a communication protocol over a network.
- **Software Defined Networking:** Software defined networking (SDN) is an approach to using open protocols, such as OpenFlow, to apply globally aware software control at the edges of the network to access network switches and routers that typically would use closed and proprietary firmware.

Fog Computing - *not a replacement* of Cloud Computing

- Fog/edge devices are there to help the Cloud datacenter to better response time for real-time applications. Handshaking among Fog and Cloud computing is needed.
- Broadly, benefits of Fog computing are:
  - Low latency and location awareness
  - Widespread geographical distribution
  - Mobility
  - Very large number of nodes
  - Predominant role of wireless access
  - Strong presence of streaming and real time applications
  - Heterogeneity

FOG Advantages ?

- Fog can be distinguished from Cloud by its proximity to end-users.
- Dense geographical distribution and its support for mobility.
- It provides low latency, location awareness, and improves quality-of- services (QoS) and real time applications.

Security Issues

- Major security issues are authentication at different levels of gateways as well as in the Fog nodes
- Man-in-the-Middle-Attack
- Privacy Issues
- *In case of smart grids, the smart meters installed in the consumer's home. Each smart meter and smart appliance has an IP address. A malicious user can either tamper with its own smart meter, report false readings, or spoof IP addresses.*

Limitations of Cloud Computing

- High capacity(bandwidth) requirement
- Client access link
- High latency
- Security

"Fog" Solution?

- Reduction in data movement across the network resulting in reduced congestion
- Elimination of bottlenecks resulting from centralized computing systems
- Improved security of encrypted data as it stays closer to the end user

Fog Computing and Cloud Computing

Requirement	Cloud computing	Fog computing
Latency	high	low
Delay jitter	High	Very low
Location of server nodes	With in internet	At the edge of local n/w
Distance between the clie nt and server	Multiple hops	One hop
Security	Undefined	Can be defined
Attack on data enrouter	High probability	Very Less probability
Location awareness	No	Yes

Source: Internet

Fog Computing and Cloud Computing

Requirement	Cloud computing	Fog computing
Geographicaldistribution	Centralized	Distributed
No. of server nodes	Few	Very large
Support for Mobility	Limited	Supported
Real time interactions	Supported	Supported
Type of last mile connecti vity	Leased line	Wireless

Source: Internet

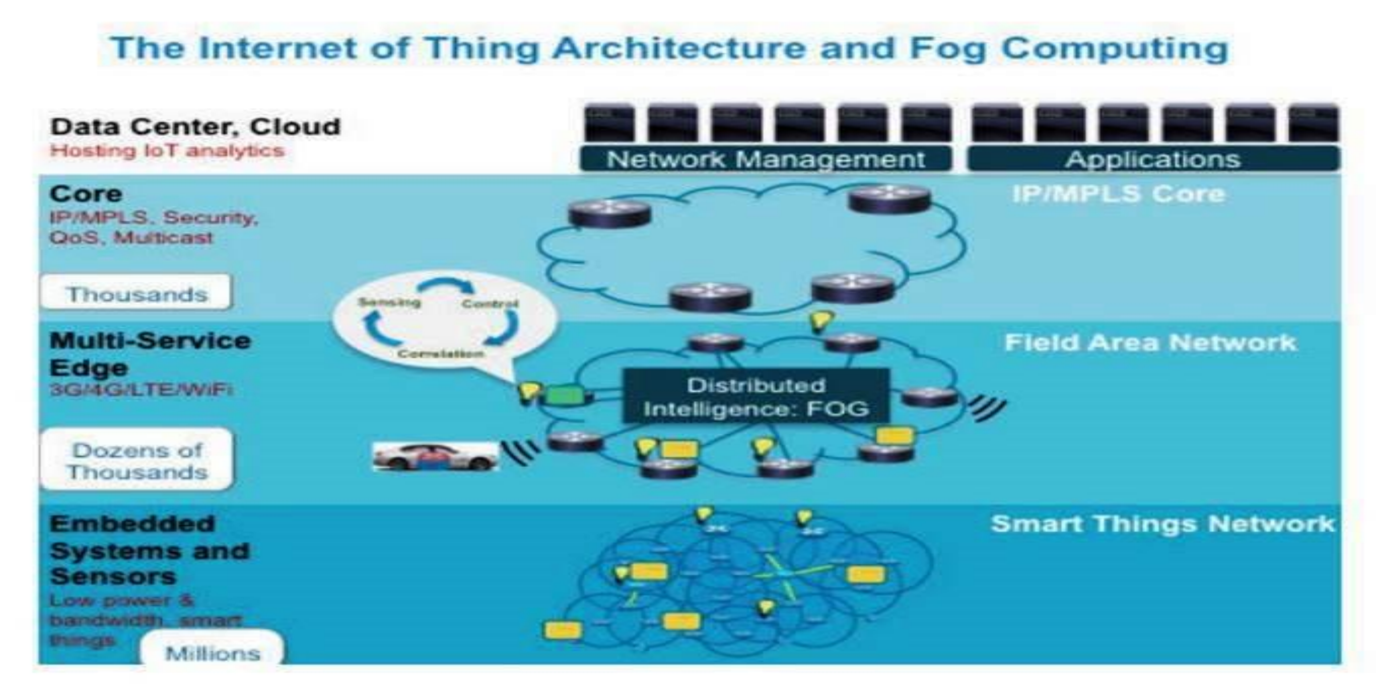
Fog Computing Use-cases

- **Emergency Evacuation Systems:** Real-time information about currently affected areas of building and exit route planning.
- **Natural Disaster Management:** Real-time notification about landslides, flash floods to potentially affected areas.
- Large sensor deployments generate a lot of data, which can be pre-processed, summarized and then sent to the cloud to reduce congestion in network.
- **Internet of Things (IoT)** based big-data applications: Connected Vehicle, Smart Cities, Wireless Sensors and Actuators Networks(WSANs) etc.

Applicability

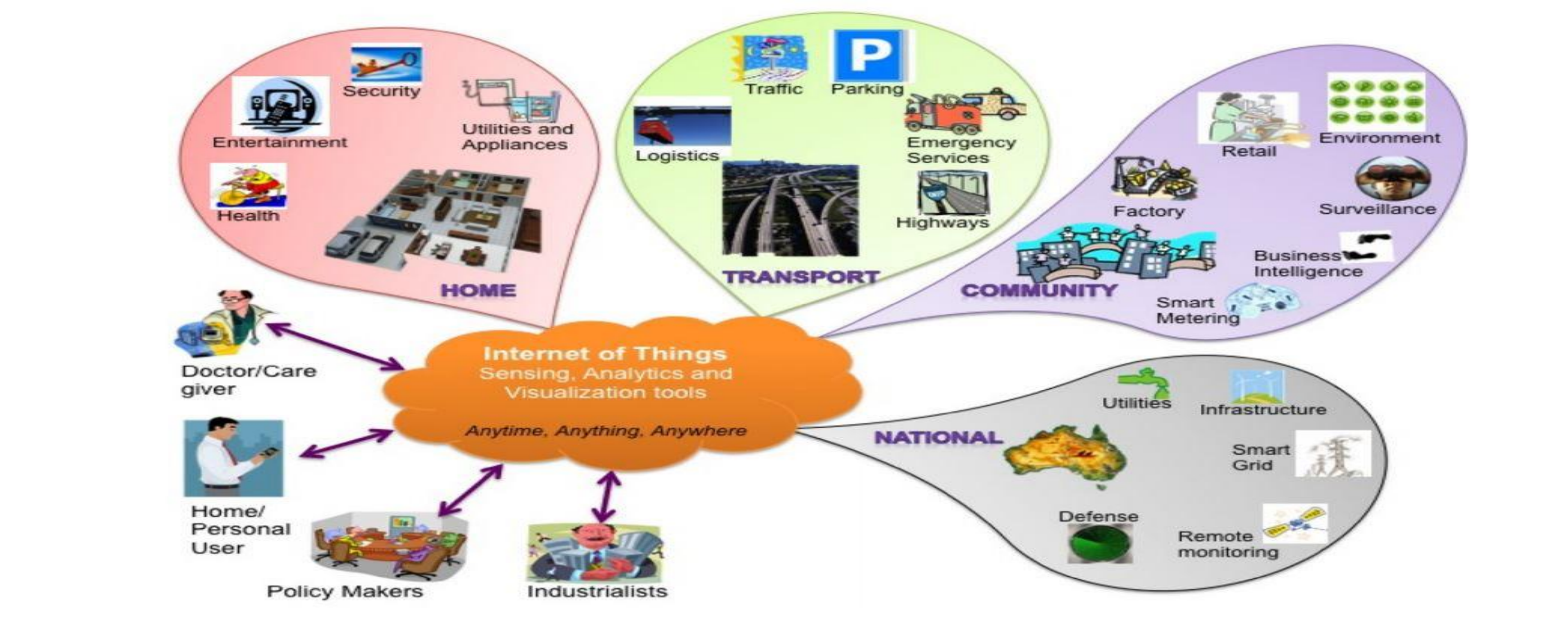
- Smart Grids
- Smart Traffic Lights
- Wireless Sensors
- Internet of Things
- Software Defined Network

Fog Computing and IoT (Internet of Things)



Source: Fog Computing and Its Role in the Internet of Things, Flavio Bonomi, Rodolfo Milito, Jiang Zhu, Sateesh Addepalli

Internet of Things



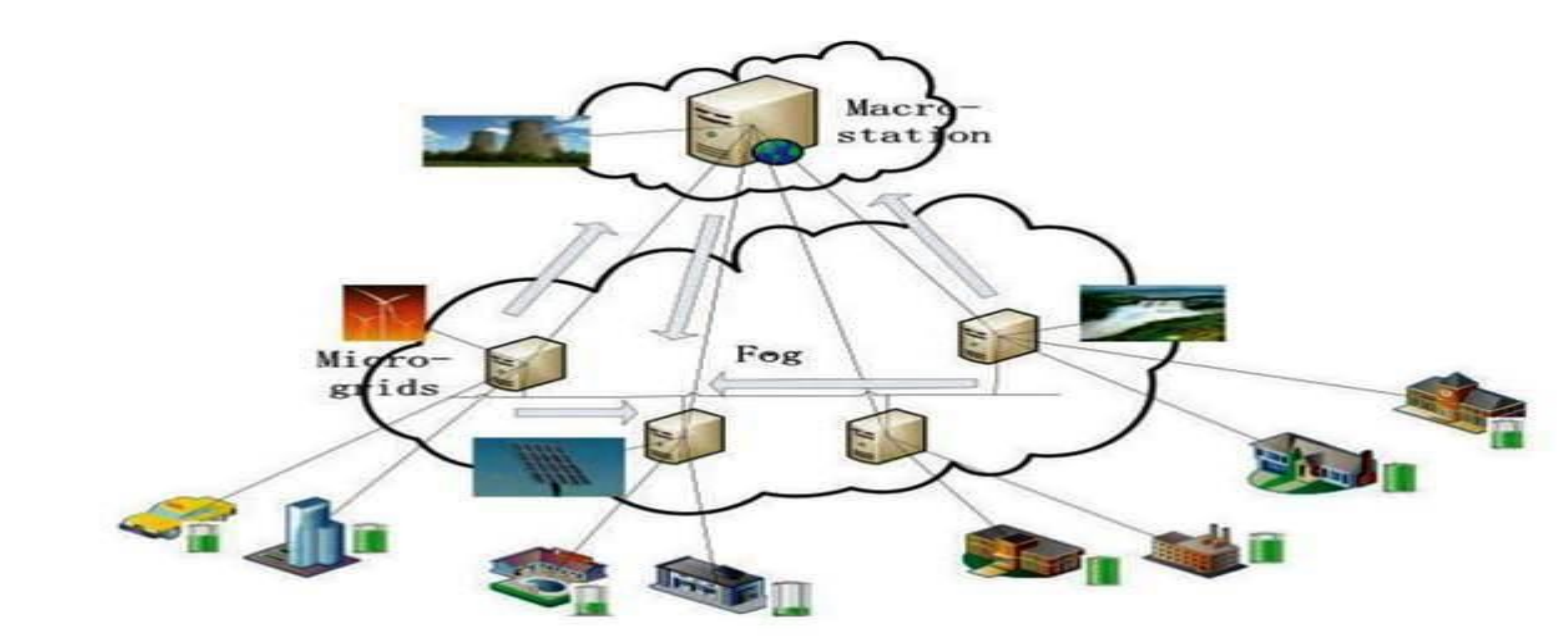
Source: Internet of Things (IoT): A vision, architectural elements, and future directions, Jayavardhana Gubbi, Rajkumar Buyya, Slaven Marusic, Marimuthu Palaniswami

Connected Vehicle (CV)

- The Connected Vehicle deployment displays a rich scenario of connectivity and interactions: cars to cars, cars to access points (Wi-Fi, 3G, LTE, roadside units [RSUs], smart traffic lights), and access points to access points. The Fog has a number of attributes that make it the ideal platform to deliver a rich menu of SCV services in infotainment, safety, traffic support, and analytics: geo-distribution (throughout cities and along roads), mobility and location awareness, low latency, heterogeneity, and support for real-time interactions.

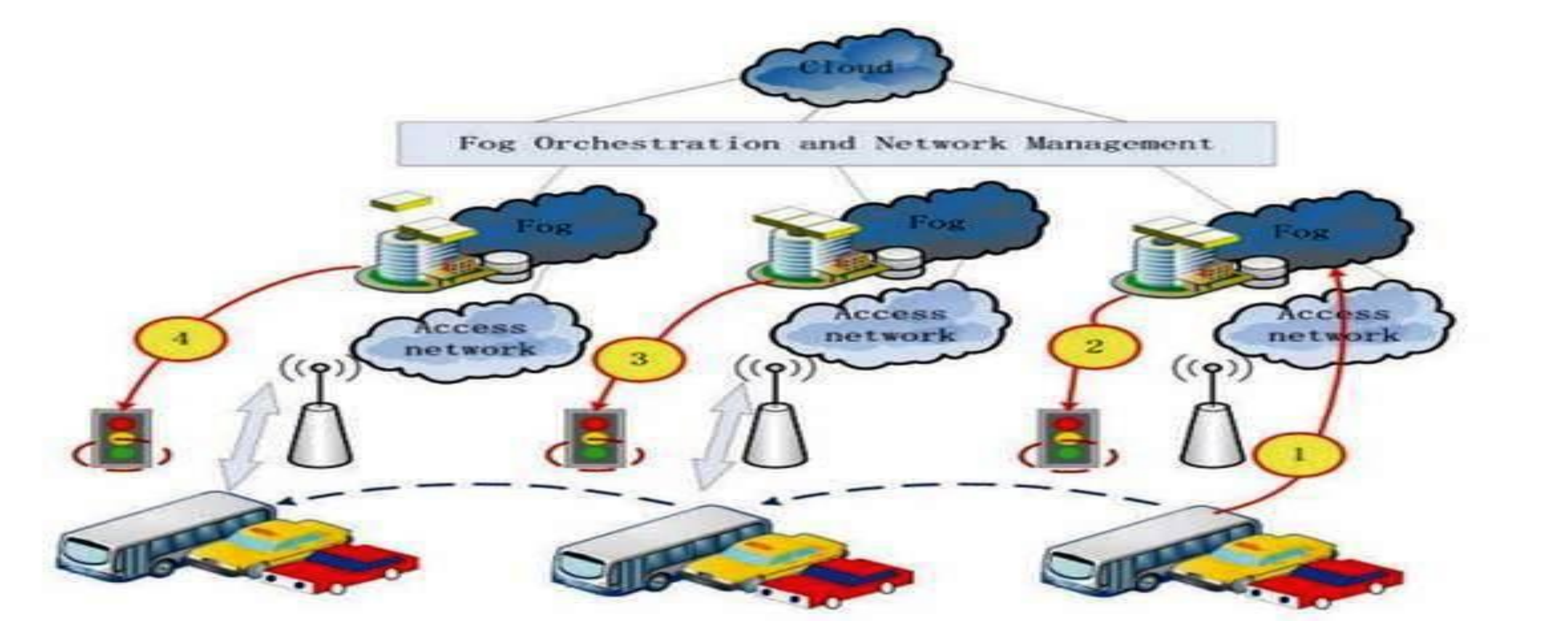
Source: Fog Computing and Its Role in the Internet of Things, Flavio Bonomi, Rodolfo Milito, Jiang Zhu, Sateesh Addepalli

Smart Grid and Fog Computing



Source: Source: The Fog Computing Paradigm: Scenarios and Security Issues, Ivan Stojmenovic and Sheng Wen

Fog computing in Smart Traffic Lights and Connected Vehicles



Source: Source: The Fog Computing Paradigm: Scenarios and Security Issues, Ivan Stojmenovic and Sheng Wen

Thank You!!

IIT KHARAGPUR

NPTEL ONLINE CERTIFICATION COURSES

NPTEL

CLOUD COMPUTING

Fog Computing - II

PROF. SOUMYA K. GHOSH

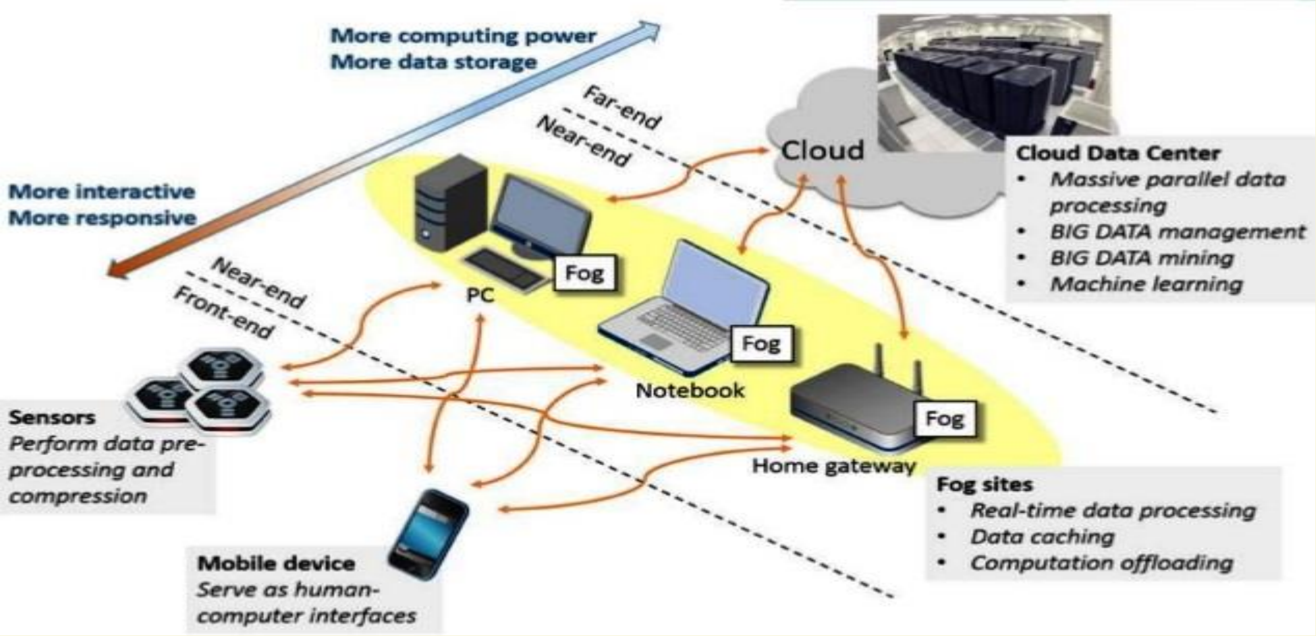
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

IIT KHARAGPUR

FOG Computing

- Cloud computing has been able to help in realizing the potential of IoT devices by providing scalability, resource provisioning as well as providing data intelligence from the large amount of data.
- But, the cloud has few limitations in the context of real-time latency (response required in seconds) sensitive applications.
- Fog computing has been coined in order to serve the real-time latency sensitive applications faster.
- Fog computing leverages the local knowledge of the data that is available to the fog node and draws insights from the data by providing faster response.

Fog Computing



Source: Internet

Fog Computing and Cloud Computing

Requirement	Cloud computing	Fog computing
Latency	high	low
Delay jitter	High	Very low
Location of server nodes	With in internet	At the edge of local n/w
Distance between the clie nt and server	Multiple hops	One hop
Security	Undefined	Can be defined
Attack on data enrouter	High probability	Very Less probability
Location awareness	No	Yes

Source: Internet

Fog Computing and Cloud Computing

Requirement	Cloud computing	Fog computing
Geographicaldistribution	Centralized	Distributed
No. of server nodes	Few	Very large
Support for Mobility	Limited	Supported
Real time interactions	Supported	Supported
Type of last mile connecti vity	Leased line	Wireless

Source: Internet

Fog Computing Use-cases

- **Emergency Evacuation Systems:** Real-time information about currently affected areas of building and exit route planning.
- **Natural Disaster Management:** Real-time notification about landslides, flash floods to potentially affected areas.
- Large sensor deployments generate a lot of data, which can be pre-processed, summarized and then sent to the cloud to reduce congestion in network.
- **Internet of Things (IoT)** based big-data applications: Connected Vehicle, Smart Cities, Wireless Sensors and Actuators Networks(WSANs) etc.

Applicability

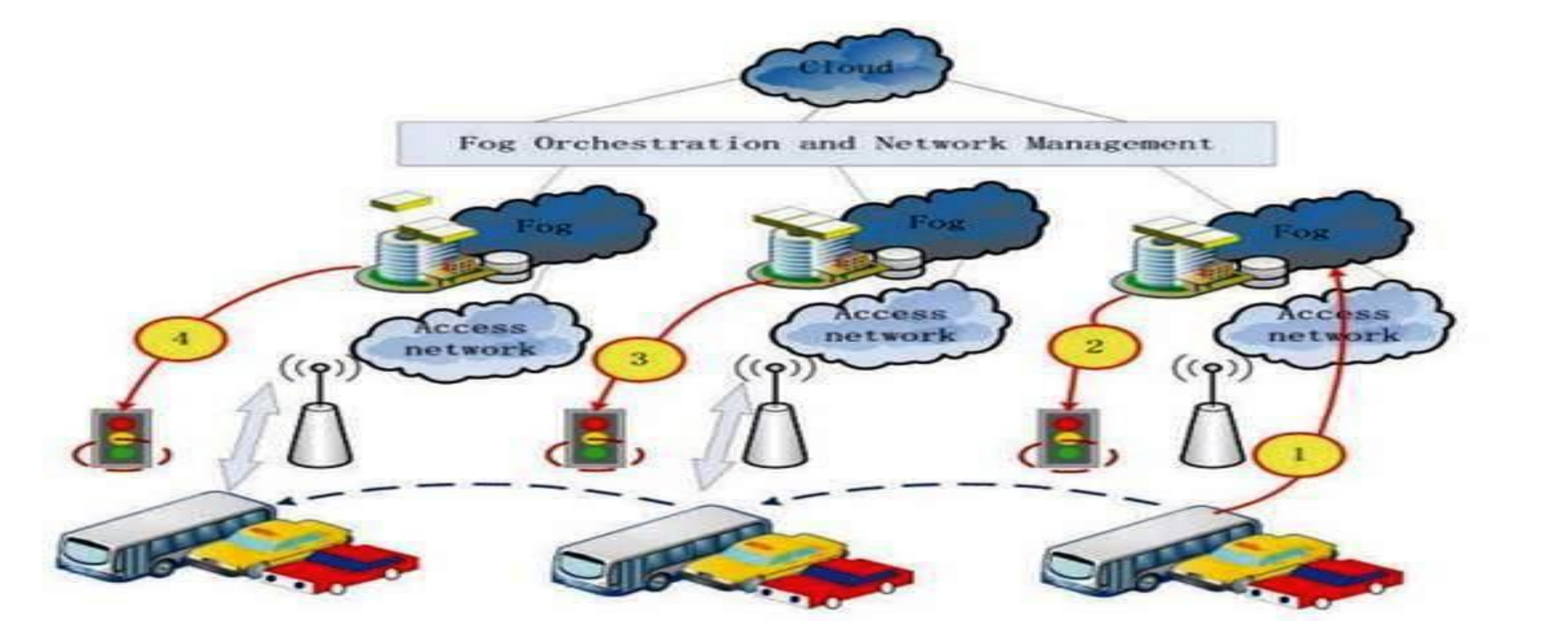
- Smart Traffic Lights
- Connected Vehicles
- Smart Grids
- Wireless Sensors
- Internet of Things
- Software Defined Network

Connected Vehicle (CV)

- The Connected Vehicle deployment displays a rich scenario of connectivity and interactions: cars to cars, cars to access points (Wi-Fi, 3G, LTE, roadside units [RSUs], smart traffic lights), and access points to access points.
- Fog has a number of attributes that make it the ideal platform for CV in providing services, like infotainment, safety, traffic support, and analytics: geo-distribution (throughout cities and along roads), mobility and location awareness, low latency, heterogeneity, and support for real-time interactions.

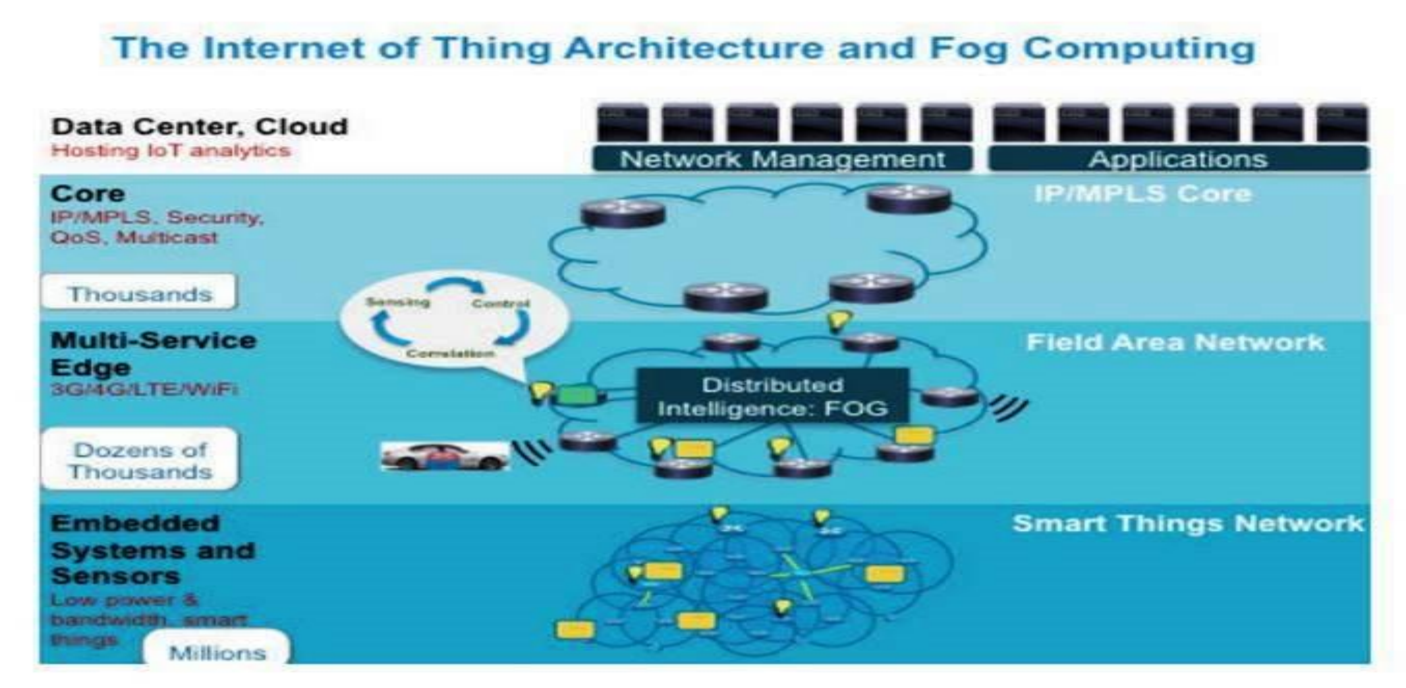
Source: Fog Computing and Its Role in the Internet of Things, Flavio Bonomi, Rodolfo Milito, Jiang Zhu, Sateesh Addepalli

Fog Computing in Smart Traffic Lights and Connected Vehicles



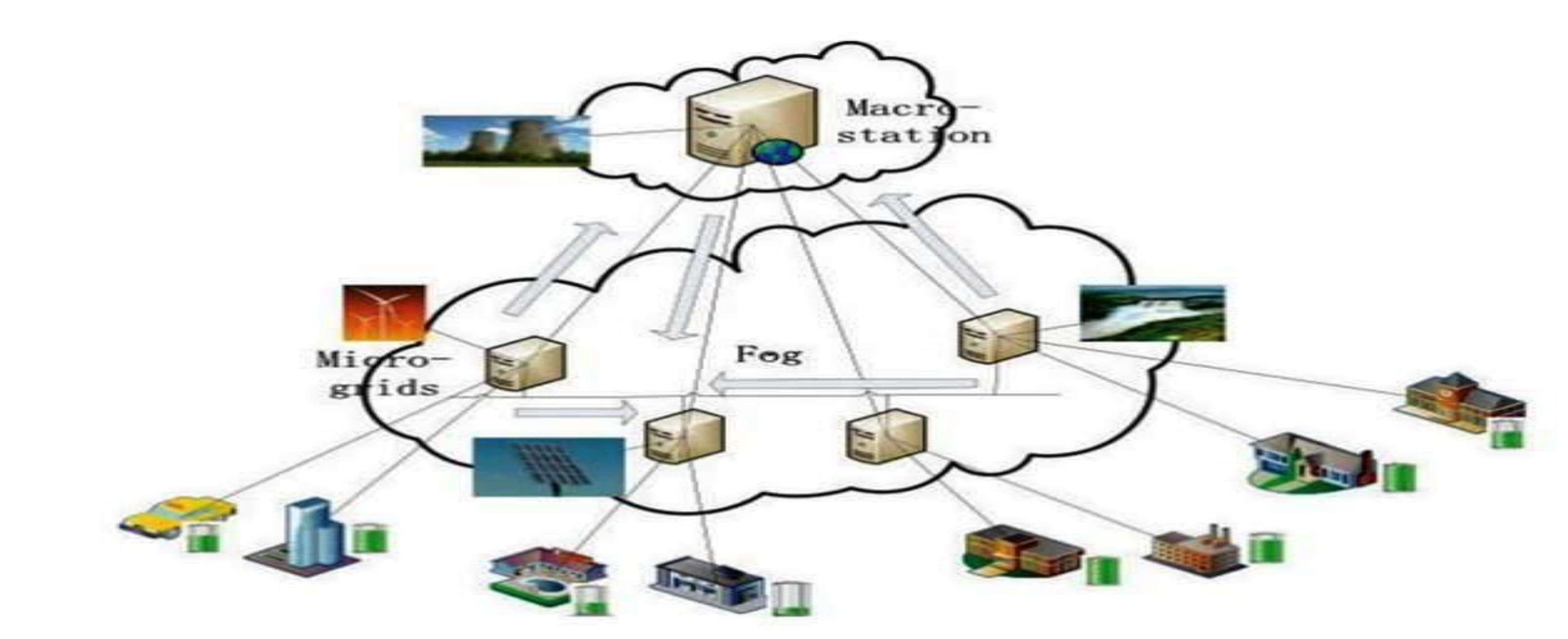
Source: Source: The Fog Computing Paradigm: Scenarios and Security Issues, Ivan Stojmenovic and Sheng Wen

Fog Computing and IoT (Internet of Things)



Source: Fog Computing and Its Role in the Internet of Things, Flavio Bonomi, Rodolfo Milito, Jiang Zhu, Sateesh Addepalli

Fog Computing and Smart Grid



Source: Source: The Fog Computing Paradigm: Scenarios and Security Issues, Ivan Stojmenovic and Sheng Wen

Fog Challenges

- Fog computing systems suffer from the issue of proper resource allocation among the applications while ensuring the end-to-end latency of the services.
- Resource management of the fog computing network has to be addressed so that the system throughput increases ensuring high availability as well as scalability.
- Security of Applications/Services/Data

Resource Management of Fog network

- Utilization of idle fog nodes for better throughput
- More parallel operations
- Handling load balancing
- Meeting the delay requirements of real-time applications
- Provisioning crash fault-tolerance
- More scalable system

Resource Management – Challenges

- Data may not be available at the executing fog node. Therefore, data fetching is needed from the required sensor or data source.
- The executing node might become unresponsive due to heavy workload, which compromises the latency.
- Choosing a new node in case of micro-service execution migration so that the response time gets reduced.
- Due to unavailability of an executing node, there is a need to migrate the partially processed persistent data to a new node. (State migration)

Resource Management – Challenges (contd…)

- Due to unavailability of an executing node, there is a need to migrate the partially processed persistent data to a new node. (State migration)
- Final result has to transferred to the client or actuator within very less amount of time.
- Deploying application components in different fog computing nodes ensuring latency requirement of the components.
- Multiple applications may collocate in the same fog node. Therefore, the data of one application may get compromised by another application. Data security and integrity of individual applications by resource isolation has to be ensured.

Resource Management – Approaches

- Execution migration to the nearest node from the mobile client.
- Minimizing the carbon footprint for video streaming service in fog computing.
- Emphasis on resource prediction, resource estimation and reservation, advance reservation as well as pricing for new and existing IoT customers.
- Docker as an edge computing platform. Docker may facilitate fast deployment, elasticity and good performance over virtual machine based edge computing platform.

Resource Management – Approaches (contd...)

- Resource management based on the fluctuating relinquish probability of the customers, service price, service type and variance of the relinquish probability.

- Studying the base station association, task distribution, and virtual machine placement for cost-efficient fog based medical cyber-physical systems. The problem can be formulated into a mixed-integer non-linear linear program and then they linearize it into a mixed integer linear programming (LP). LP-based two-phase heuristic algorithm has been developed to address the computation complexity.

Fog - Security Issues

- Major security issues are authentication at different levels of gateways as well as in the Fog nodes
- Man-in-the-Middle-Attack
- Privacy Issues
- *In case of smart grids, the smart meters installed in the consumer's home. Each smart meter and smart appliance has an IP address. A malicious user can either tamper with its own smart meter, report false readings, or spoof IP addresses.*

Thank You!!

Cloud Computing

Use Case: **Geospatial Cloud**

Soumya K Ghosh
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur
skg@cse.iitkgp.ernet.in

Broad Agenda

- Geospatial Information
- Geospatial Cloud
- IIT Kharagpur Geo-Cloud

CLOUD ?

- **On-demand self service**
  - Use resources as and when needed
  - Minimal human interaction between user and CSP
- **Ubiquitous Network Access**
  - Services accessible over Internet using Web applications
- **Resource Pooling**
  - Large and flexible resource pooling to meet the consumers' need
  - Allocating resources efficiently and optimally for execution of applications
- **Location Independence**
  - Resources may be located at geographically dispersed locations
- **Rapid Elasticity**
  - Dynamic scaling up and down of resources
- **Measured Services (*pay-as-you-use*)**
  - Customers charged based on measured usage of the cloud resources

Geographic Information

- Information explicitly linked to locations on the earth's surface
- Geographic information can be static or dynamic
  - Static: does not change position
    - Locations, such as city/town, lake, park
  - Dynamic: changes over time
    - Population of a city
- Geographic information vary in scale
  - Information can range from meters to the globe
  - Scale vs. detail and ecological fallacies

Geospatial Information

Legal	(cadastral; zoning laws)
Political	(county lines; school districts)
Cultural	(language; ethnicity; religion)
Climatic	(temperature; precipitation)
Topographic	(elevation; slope angle; slope aspect)
Biotic	(biodiversity; species ranges)
Medical	(disease; birth rate, life expectancy)
Economic	(median income; resource wealth)
Infrastructure	(roads; water; telecommunications)
Social	(education; neighborhood influences)

Geospatial data source

- Social surveys
- Natural surveys (i.e. SOI maps)
- Remotely sensed (air photos, satellite imagery)
- Reporting networks (weather stations)
- Field data collection (GPS data or map marking associated with some attribute of interest)

Geographic Information Systems (GIS)

- A computer system for capturing, storing, querying, analyzing, and displaying geospatial data. (Chang, 2006)

- Geographic information systems are tools that allow for the processing of spatial data into information, generally information tied explicitly to, and used to make decisions about, some portion of the earth (Demers, 2002).

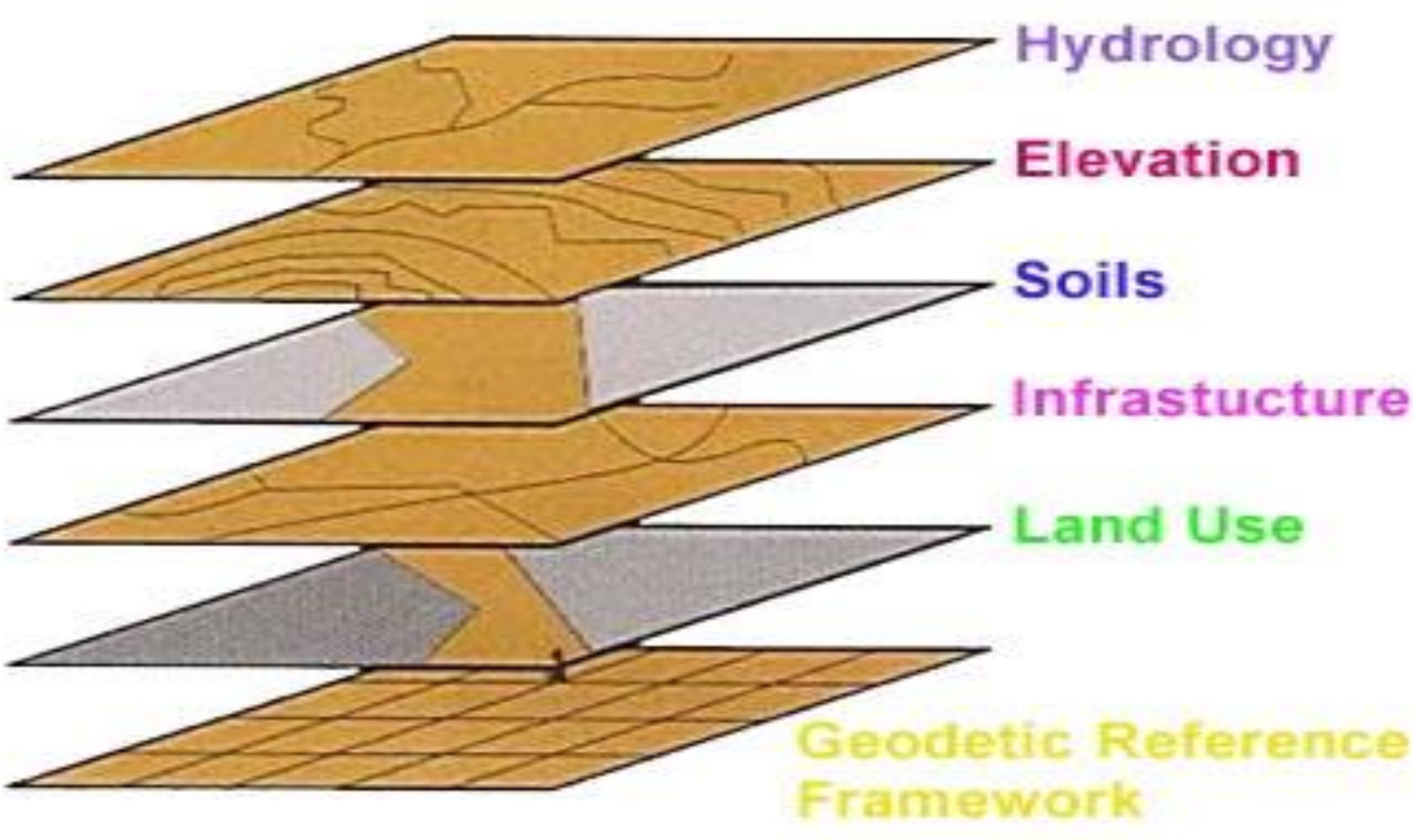
Components of a GIS

- Computer hardware
- Software
- Data management and analysis procedures (this could be considered part of the software)
- Spatial data
- People needed to operate the GIS

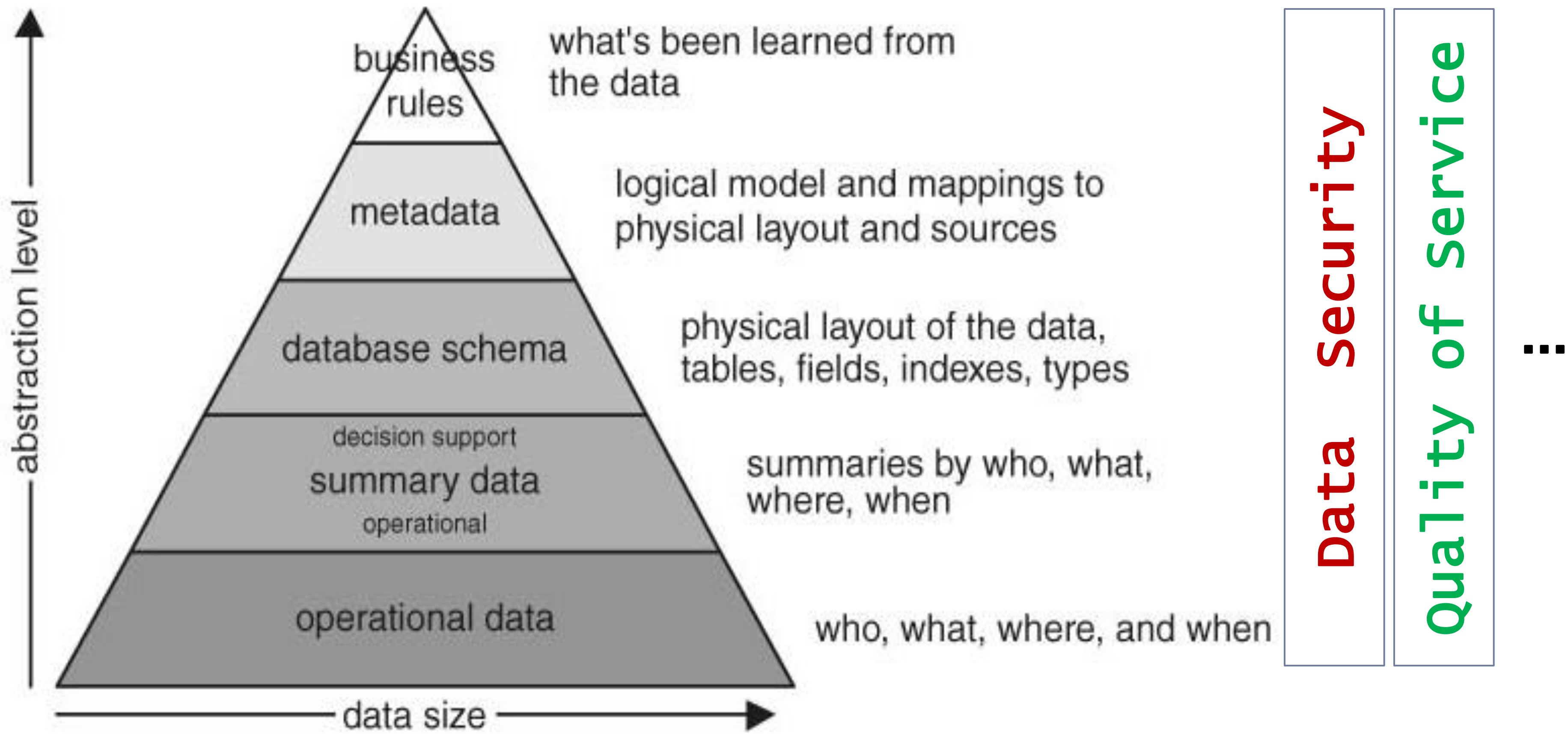
Geospatial Information System - Challenges

- Data intensive
- Computation Intensive
- Variable Load on the GIS server demands dynamic scaling in/out of resources
- GIS requires high level of reliability and performance
- Uses Network intensive web services

Geospatial Layers



Generic Architecture of Data



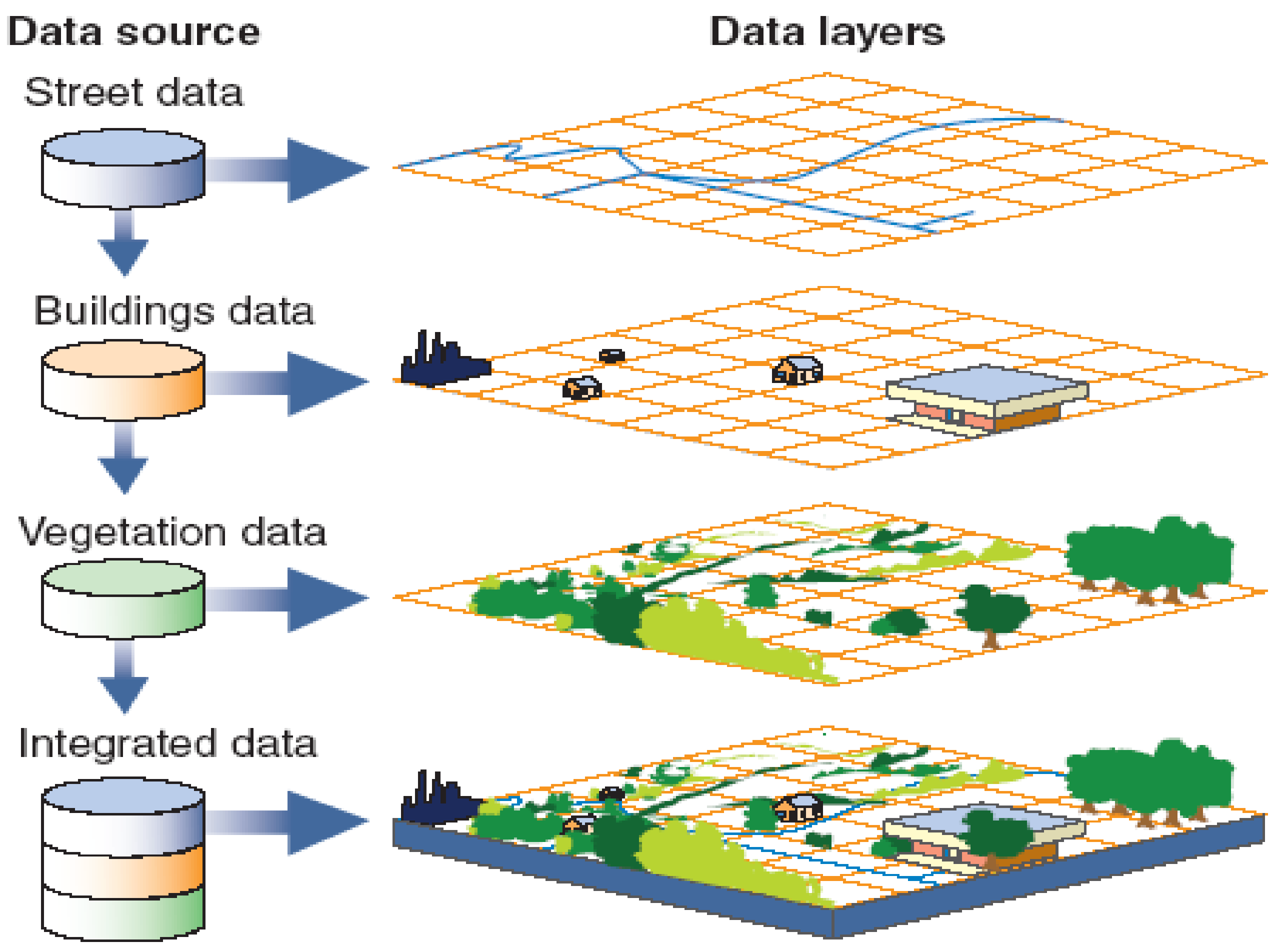
Heterogeneity Issue

- **GIS layers** are often developed by **diverse departments** relying on a mix of software and information systems

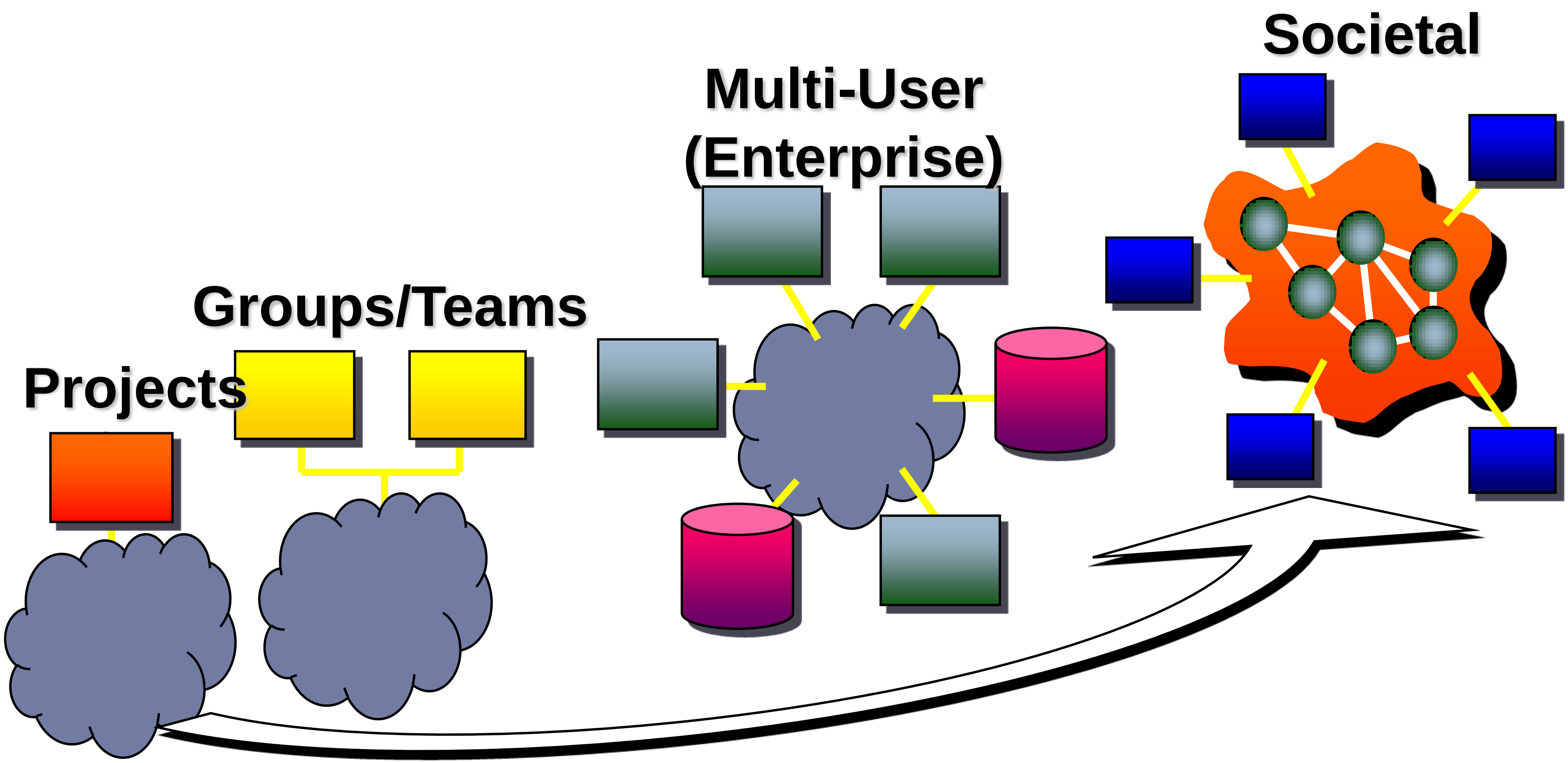
- **Each department** uses its individual system to **increase efficiency**, but sharing data and applications across the enterprise is a near impossible

- Issues to be resolved
  - Making *data description* homogeneous
  - Standard encoding for data
  - Standard mechanism for data sharing

Homogeneity (Needs to be achieved !)



GIS Users - Trend



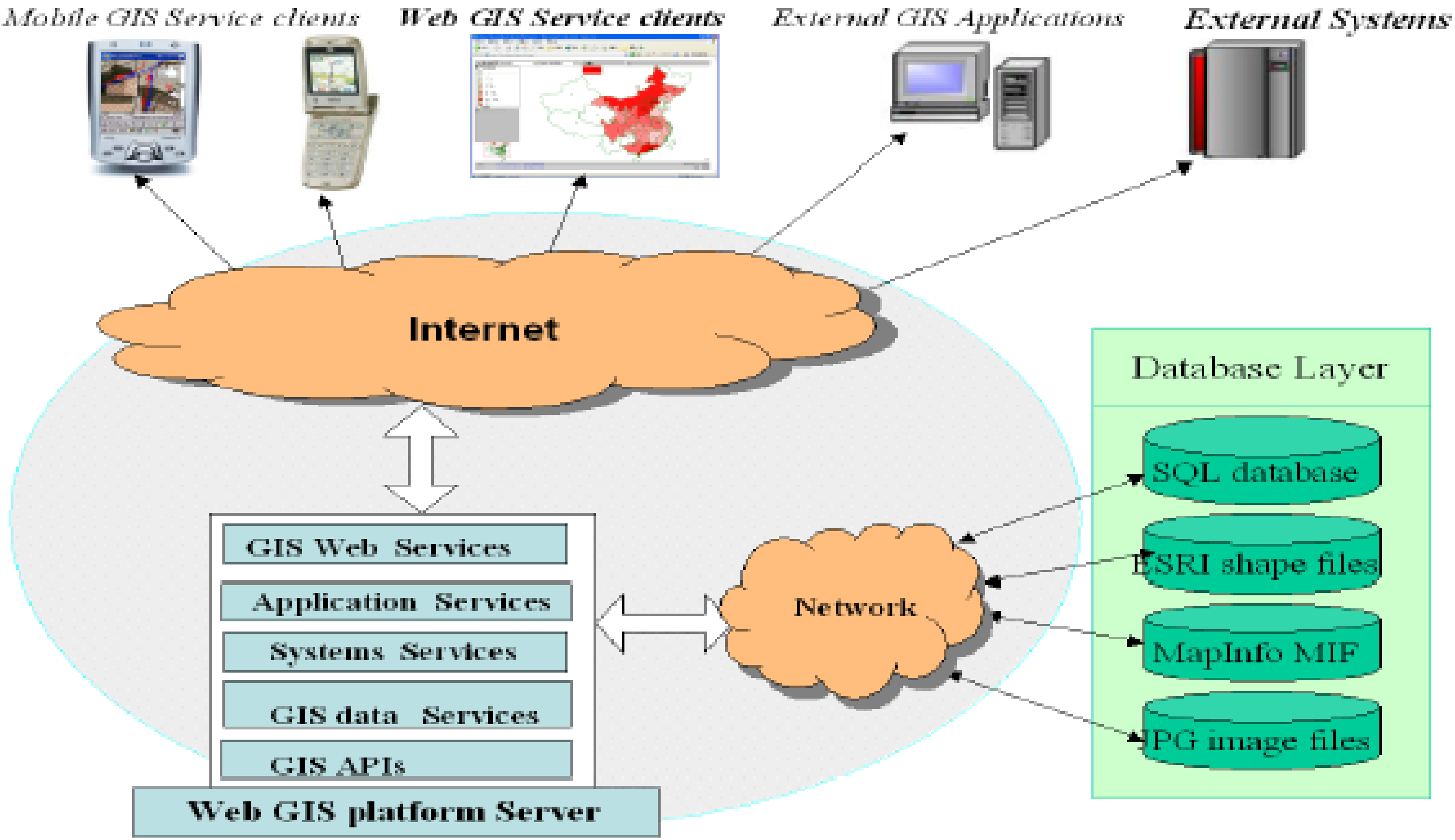
Spatial Data Infrastructure (SDI)

- "Infrastructure" implies that there should be some sort of coordination for policy formulation and implementation

- "The SDI provides a basis for spatial data discovery, evaluation, and application for users and providers within all levels of Government, the Commercial sector, the non-profit sector, Academia and by Citizens in general."

  --The SDI Cookbook

Interoperable GIS – Service driven



Need for Geospatial Cloud

- "Huge" volume of Data and Metadata
- Need of Services and Service Orchestration
- Evolving Standards and Policies
- Need for Geospatial Cloud

Need of Geospatial Cloud

- Private and public organization wants to share their spatial data
- Different requirement of geospatial data space and network bandwidth
- Get benefits by accessing others' spatial services
- Less infrastructure and spatial web service expertise needed
- Easy to port spatial service image to multiple virtual machines
- Organizations lack this type of expertise
- GIS decisions are made easier
  - Integrate latest databases
  - Merge disparate systems
  - Exchange information internally and externally

Need of Geospatial Cloud (contd…)

- It supports shared resource pooling which is useful for participating organizations with common or shared goals
- Choice of various deployment, service and business models to best suit organization goals
- Managed services prevent data and work loss from frequent outages, minimizing financial risks, while increasing efficiency
- Cloud infrastructure provides an efficient platform to share spatial data
- Provide controls in sharing of data with high security provision of cloud.
- Organizations can acquire the web service space as per needed with nominal cost.

Cloud Computing

NIST's (National Institute of Standards and Technology) definition:

- *"Cloud computing is a model for enabling convenient, on demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."*

Cloud Advantage

- Scalability on demand
  - Better resource utilization
- Minimizing IT resource management
  - Managing resources (servers, storage devices, network devices, softwares, applications, IT personnel, etc.) difficult for non-IT companies
  - Outsources to cloud
- Improving business processes
  - Focus on business process
  - Sharing of data between an organization and its clients

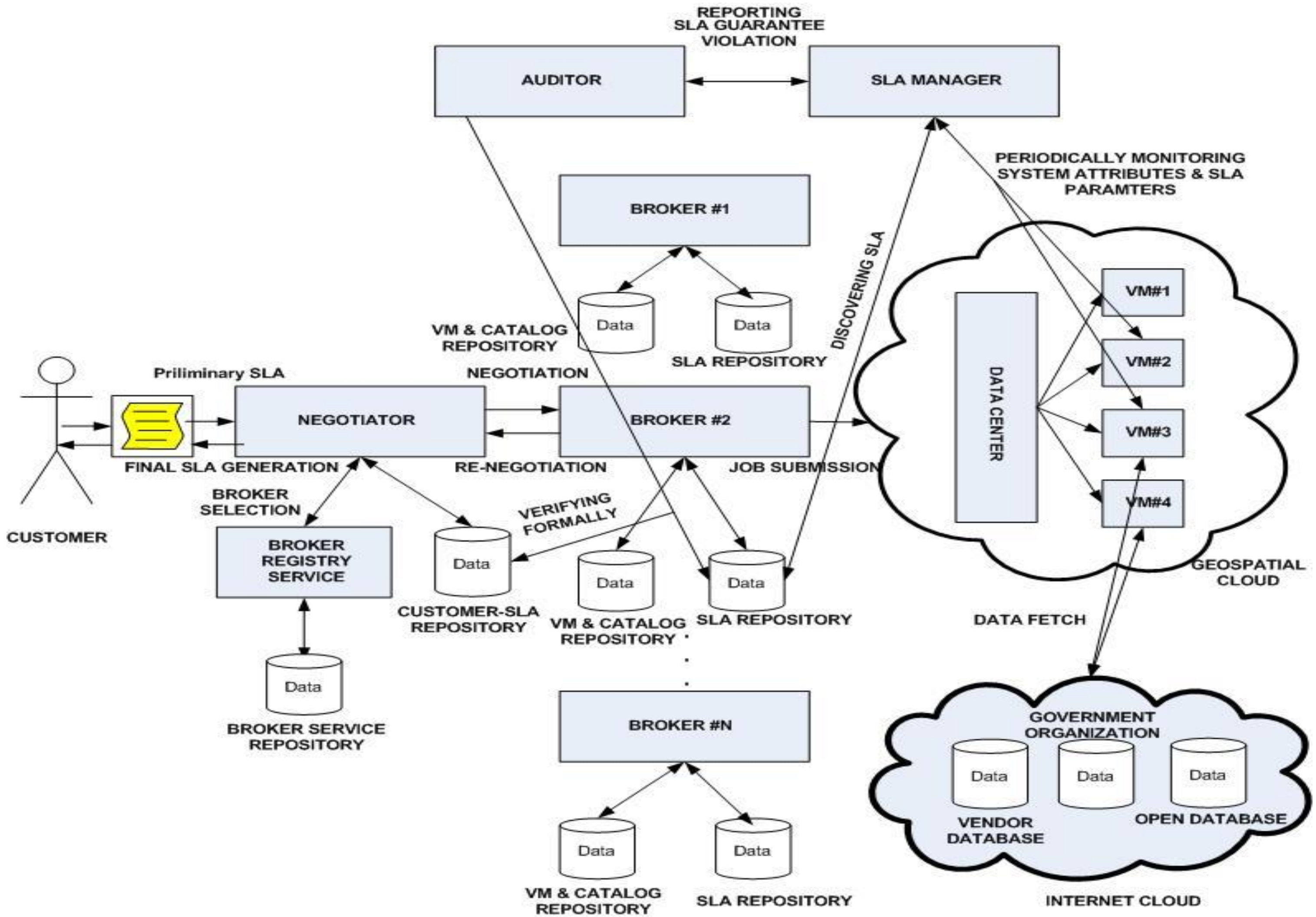
Cloud Advantage (contd)

- Minimizing start-up costs
  - Small scale companies and startups can reduce CAPEX (Capital Expenditure)
- Consumption based billing
  - Pay-as-you-use model
- Economy of scale
  - Multiplexing of same resource among several tenants
- Green computing
  - Reducing carbon footprints

Cloud Actors

- Cloud Service Provider (CSP) or Broker
  - Provides with the infrastructure, or the platform, or the service
- Customer
  - May be a single user or an organization
- Negotiator (optional)
  - Negotiates agreements between a broker and a customer
  - Publishes the services offered on behalf of the broker
- SLA Manager/Security Auditor (Not present in current clouds)

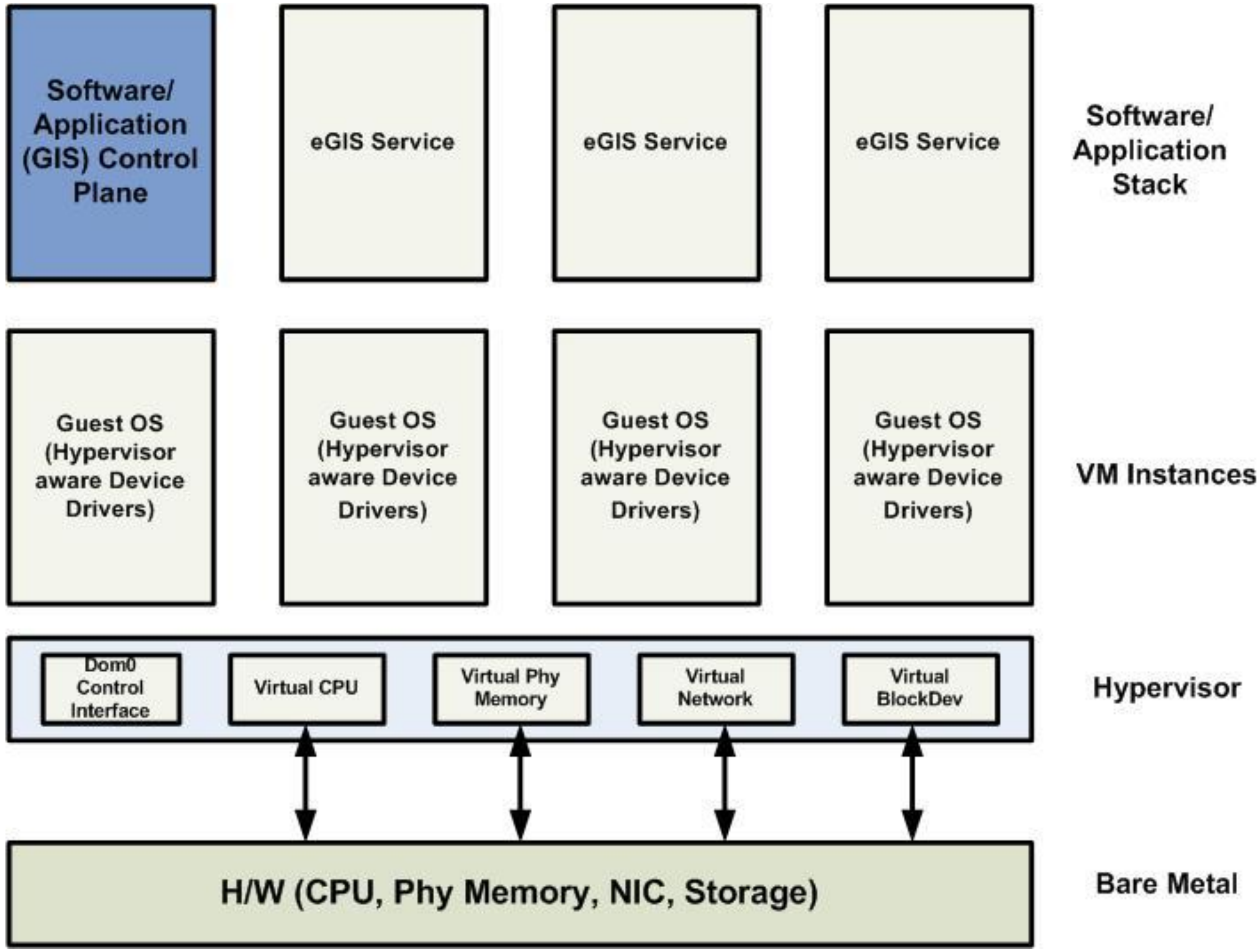
Typical Geospatial Cloud Architecture



Cloud as Service Provider

- Collection of Enterprise GIS (eGIS) Instances
  - **Resource Service** – resource allocation, manipulation of VM and network properties, monitoring of system components and virtual resources
  - **Data Service** – maintains persistent user and system data to provide a configurable user environment
  - **Interface Service** – user visible interfaces, handling authentication and other management tools.

Geospatial Cloud

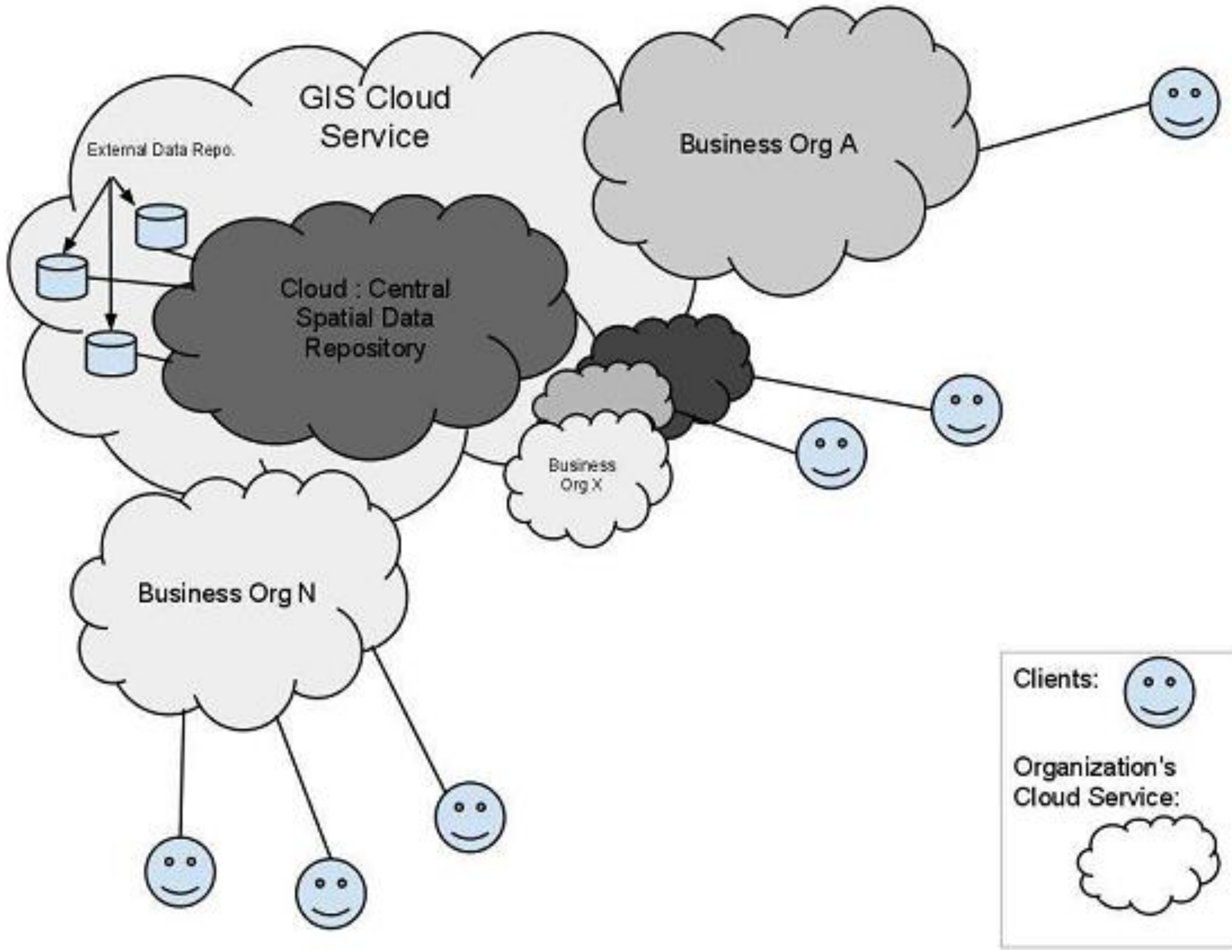


Geospatial Cloud Model

Geospatial Cloud Model

- Web service is the key technology to provide geospatial services.
- Need to integrate data from heterogeneous back-end data services.
- Data services can be inside and/or outside the cloud environment.
- Data services inside cloud can be run through Paas service model.
- Using PaaS makes load balancing, distributed replica and dynamic scaling transparent.

Geospatial Cloud – Typical Scenerio



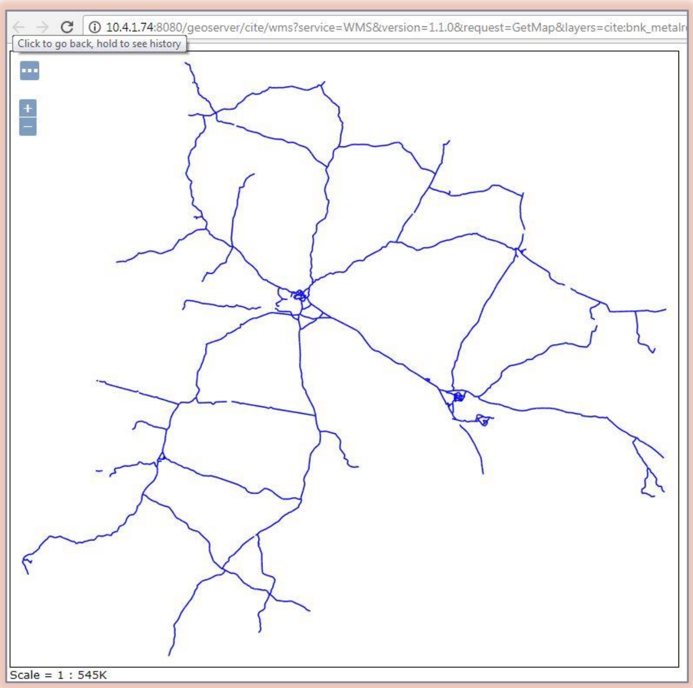
Geospatial Cloud

- Need to integrate data in an unified format.
- Performance Metrics: computation power, network bandwidth.

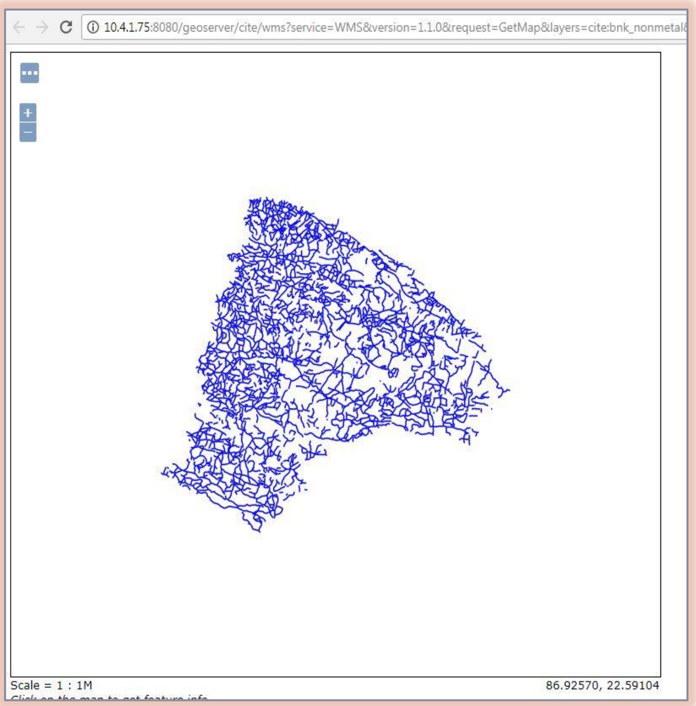
- Data sources:
  - Central Data Repository within the cloud.
  - External Data Repository providing data as WFS, WMS services.

Experimental *GeoSpatial-Cloud @IITKgp*

Service Integration for Query in Cloud (Case Study 1)

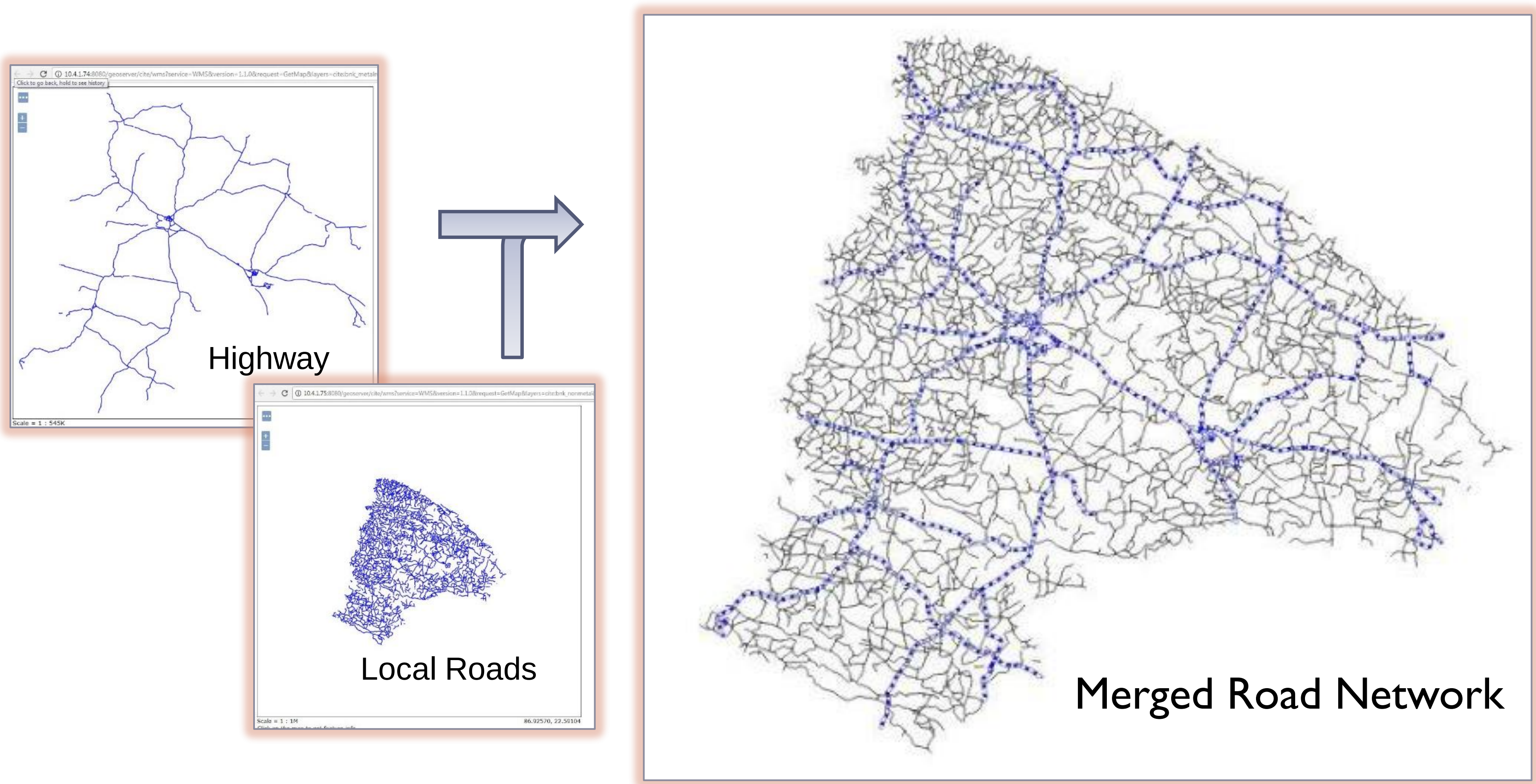


Highway

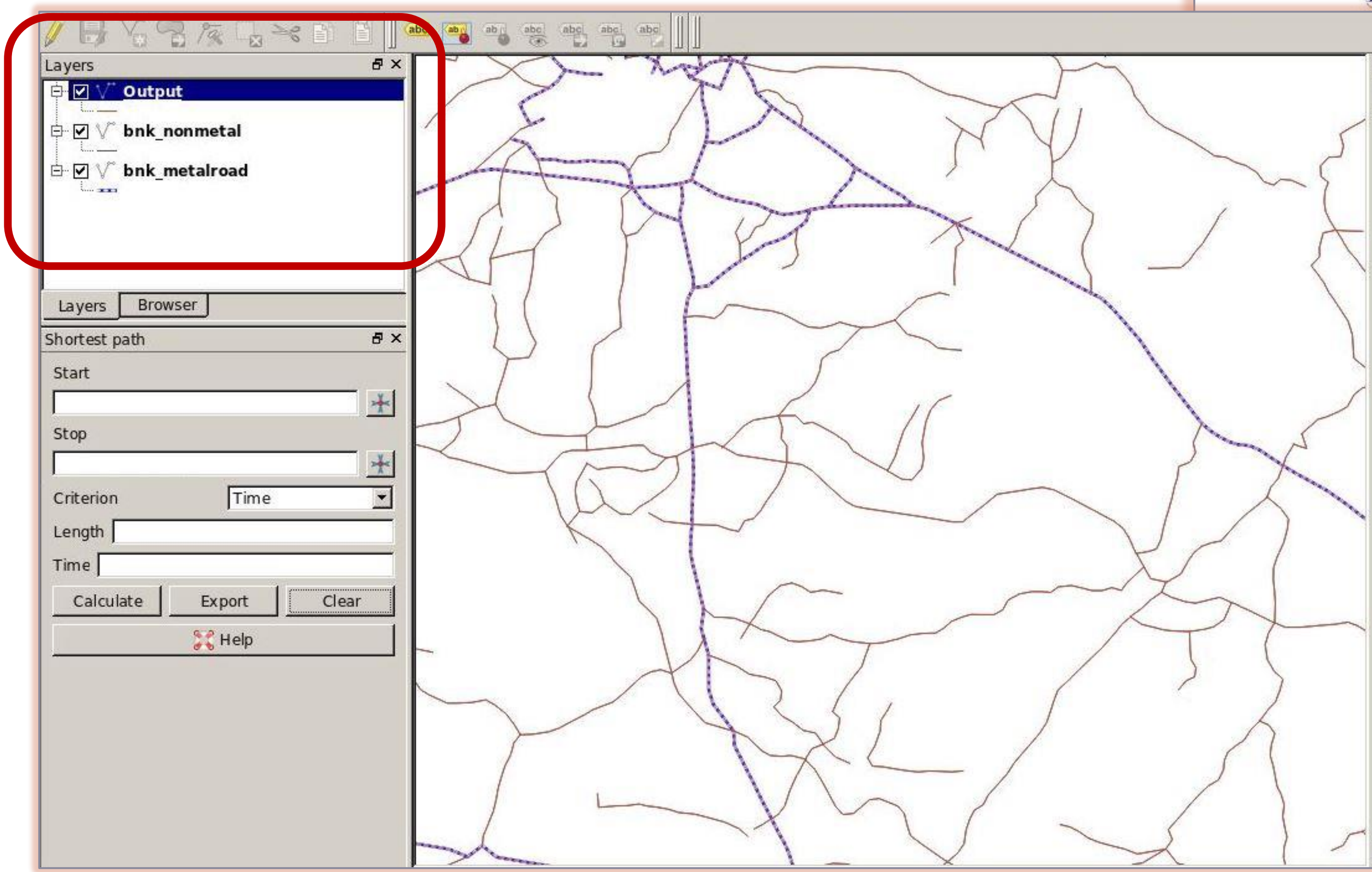


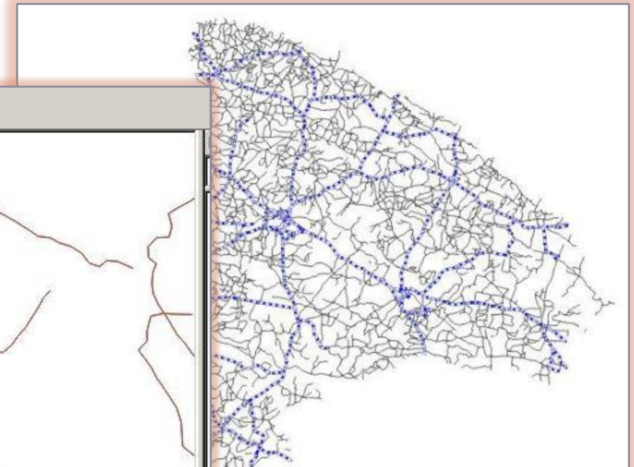
Local Roads

Service Integration for Query in Cloud



Service Integration for Query in Cloud

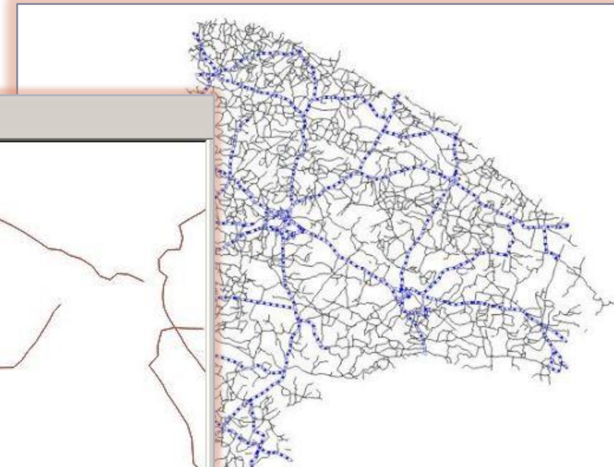




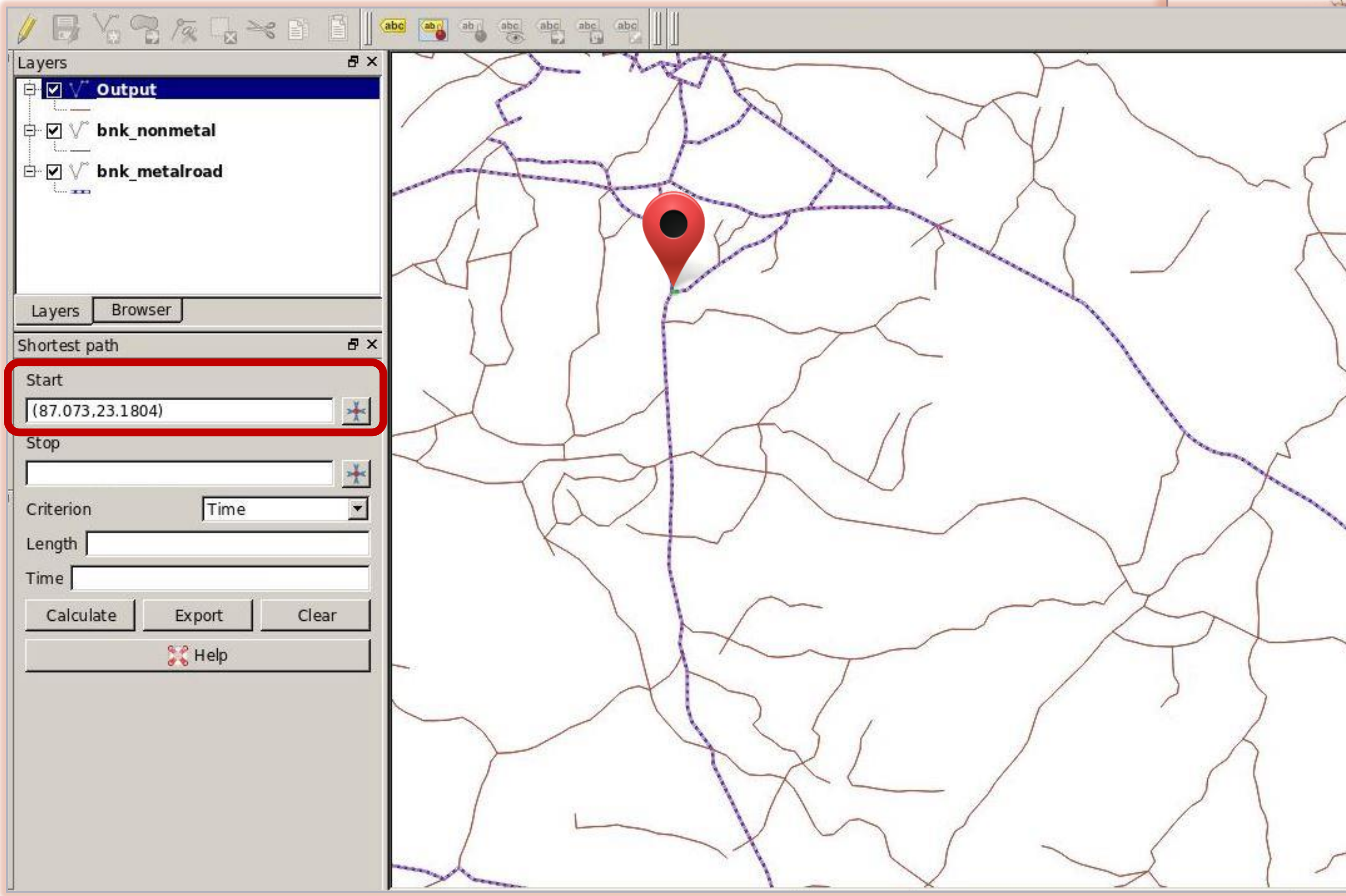
Merged Road Network

Shortest Path Calculation

Service Integration for Query in Cloud

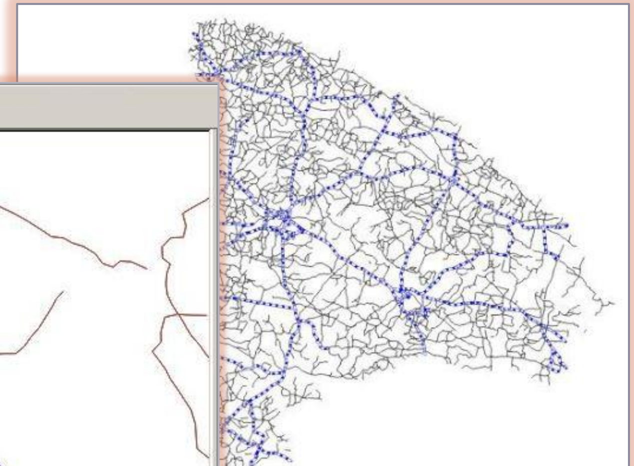


Merged Road Network

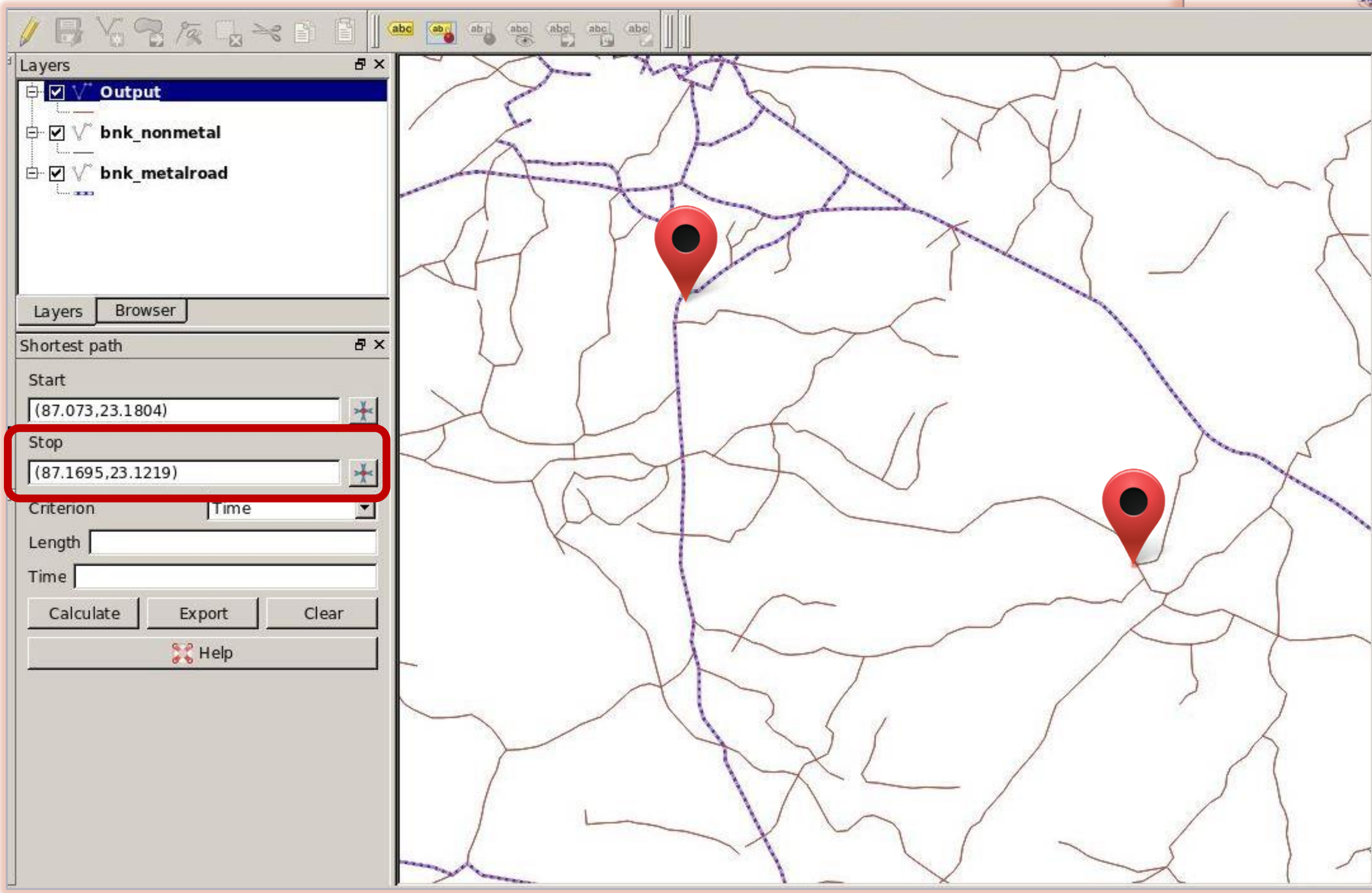


Shortest Path Calculation

Service Integration for Query in Cloud

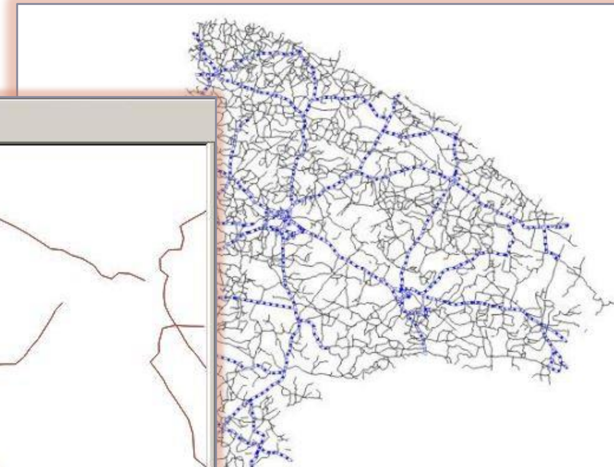


Merged Road Network

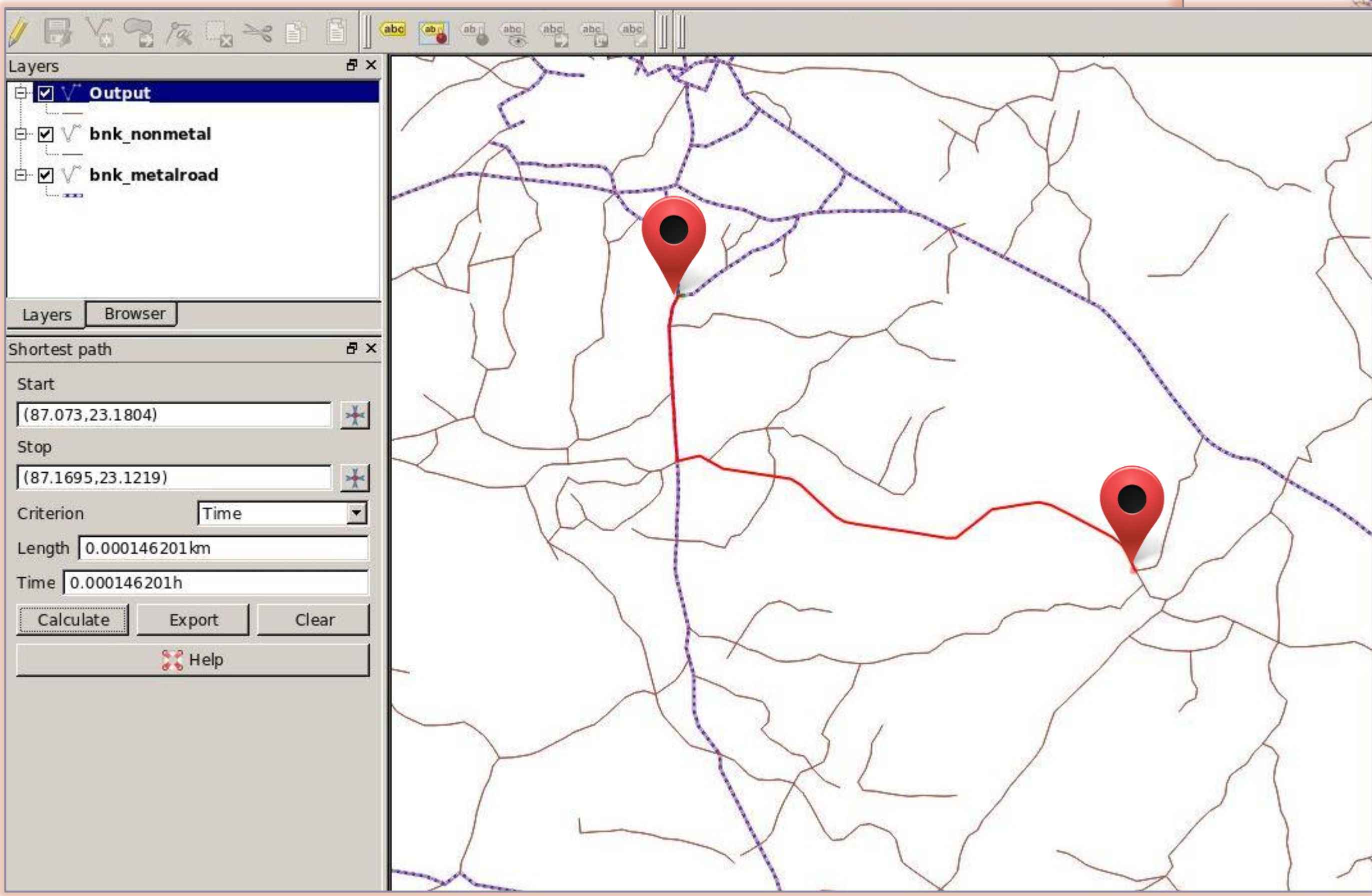


Shortest Path Calculation

Service Integration for Query in Cloud

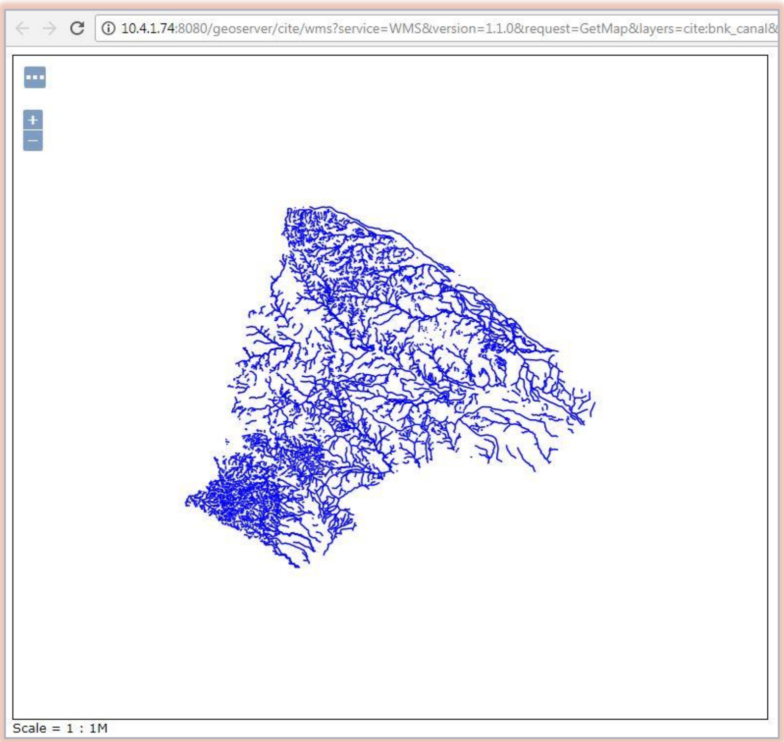


Merged Road Network

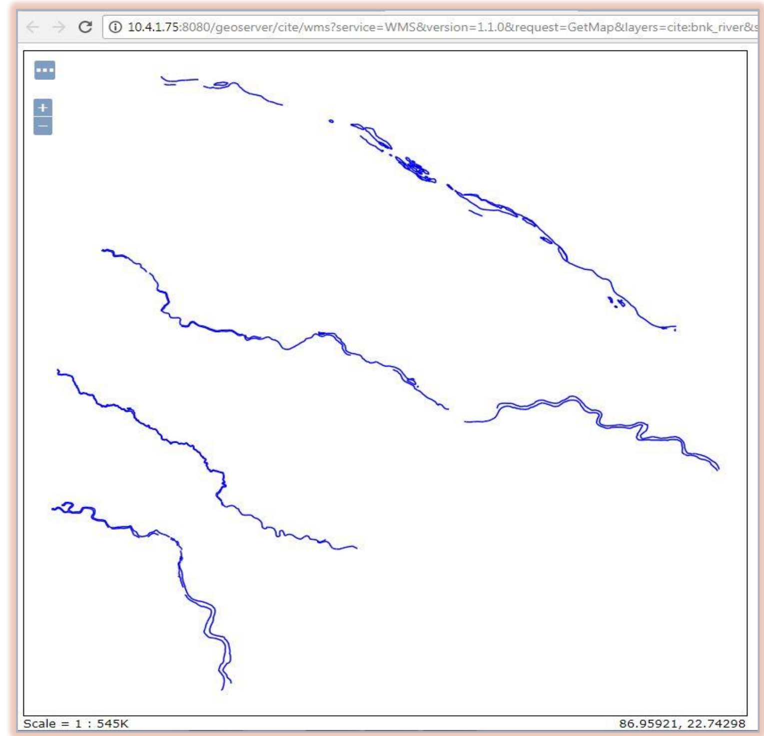


Shortest Path Calculation

Service Integration for Query in Cloud (Case Study 2)

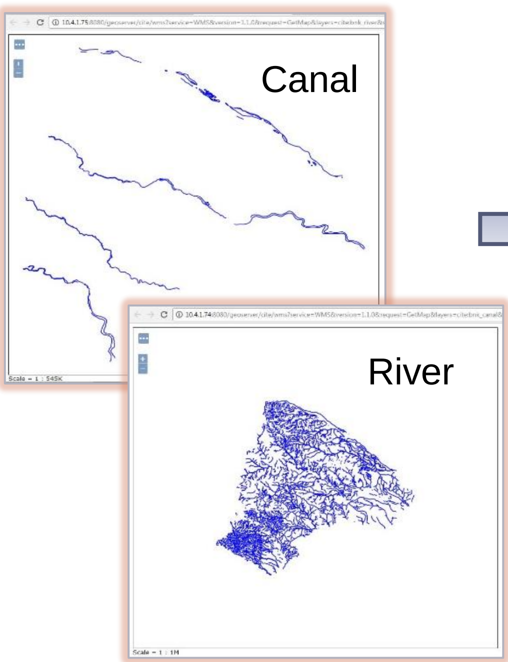


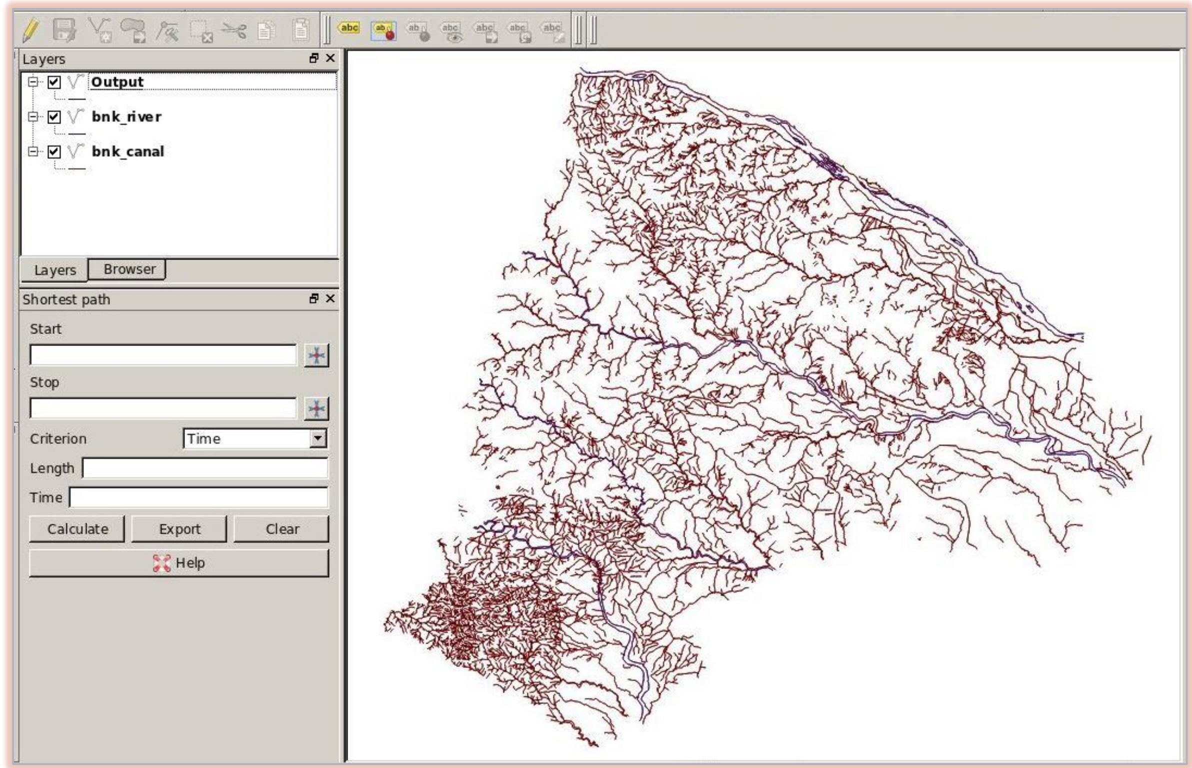
Canal



River

Service Integration for Query in Cloud

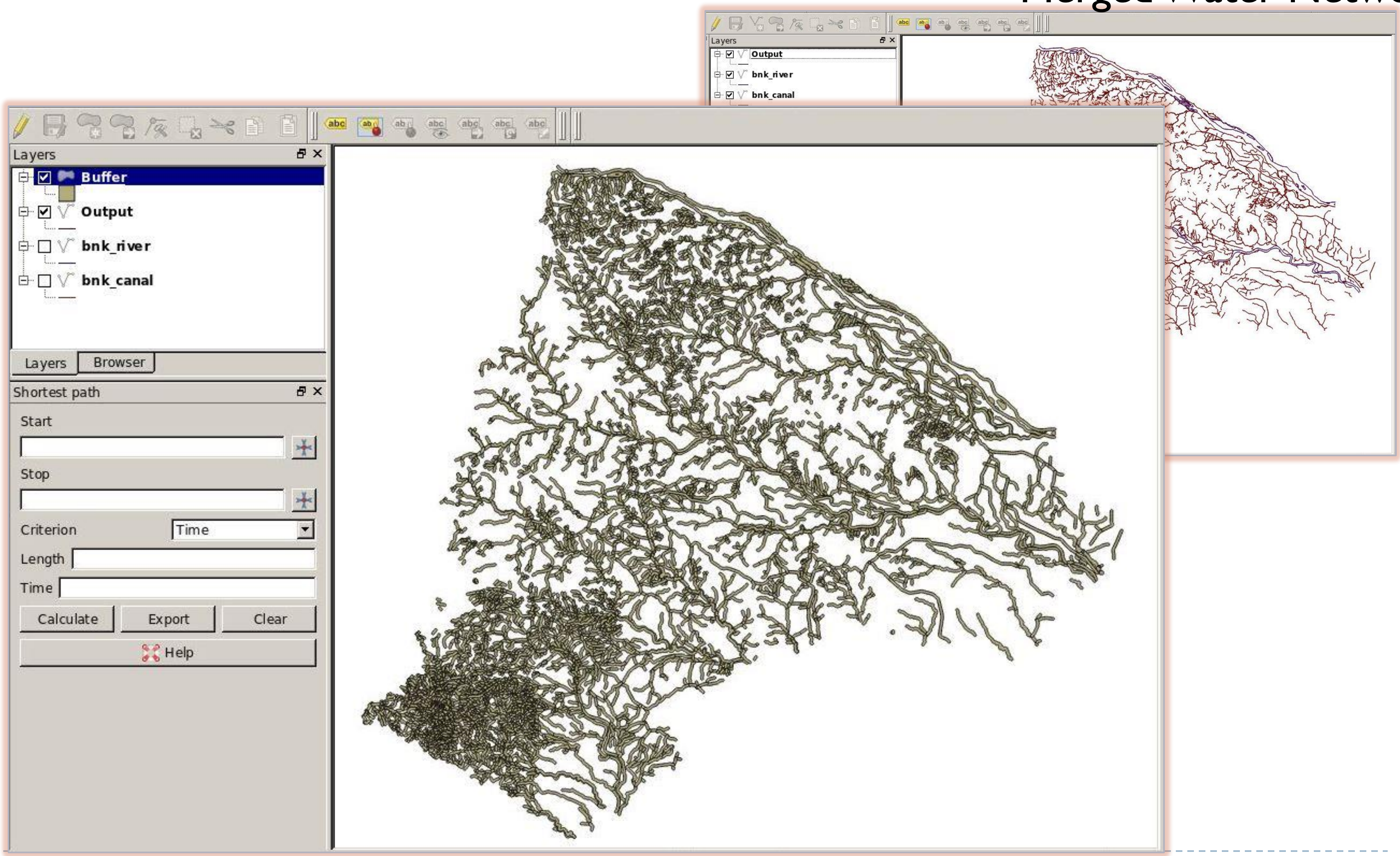




Merged Water Network

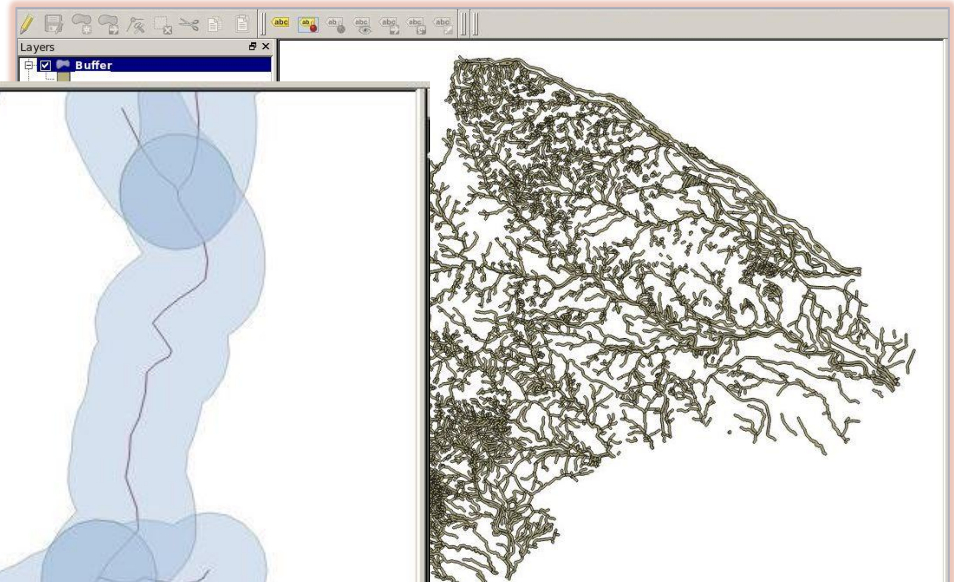
Service Integration for Query in Cloud

Merged Water Network

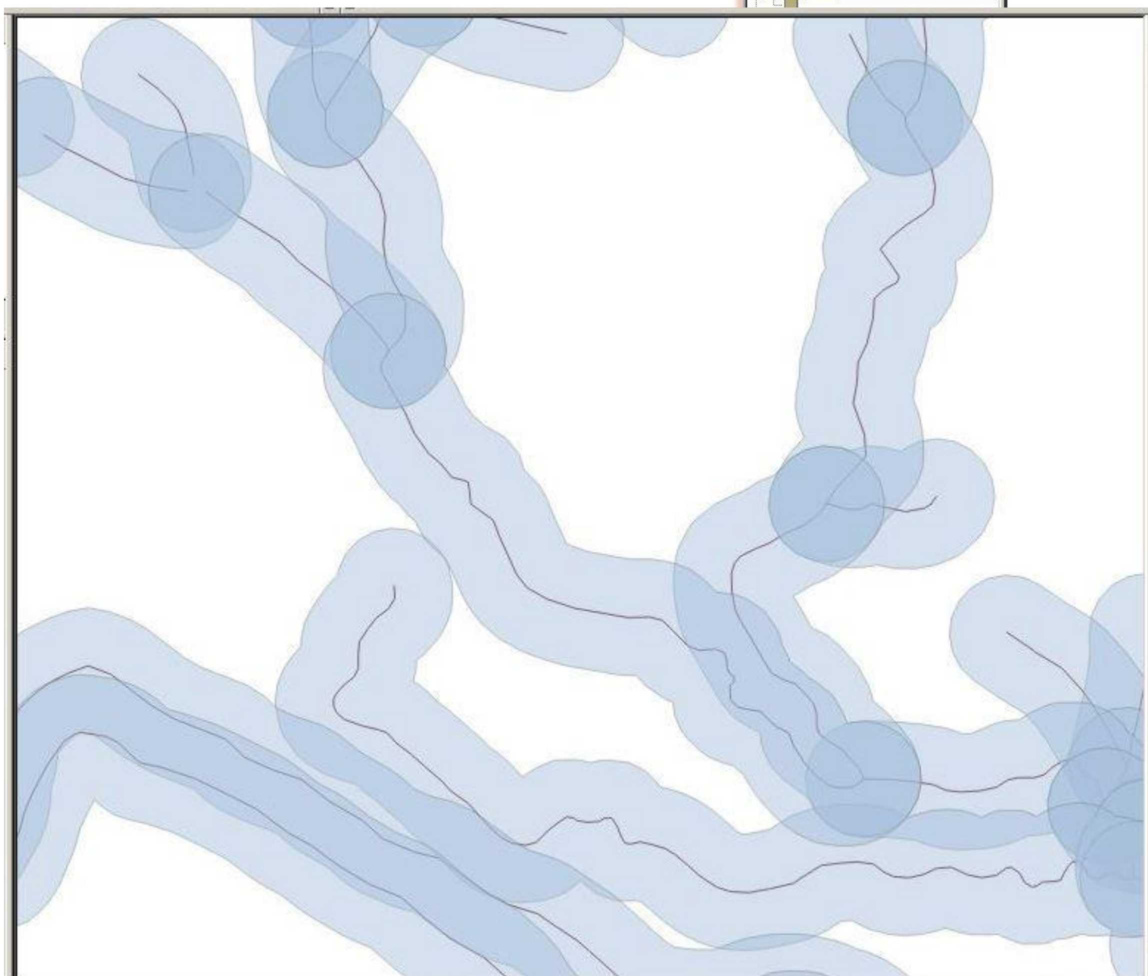


Buffer on Merged Water Network

Service Integration for Query in Cloud



Buffer on Merged Water Network



Buffer on Merged Water Network (Zoomed)

Challenges in Geospatial Cloud

Challenges in Geospatial Cloud

- Implementation of Spatial Databases.
- Scaling of Spatial Databases
- Need to be Multi-Tenant
- Policy management among the tenants.
- Geographically situated Backups
- Security of Data

Interoperability Issue

- Exchanging and processing of geospatial Information requires interoperability on different levels:
  - **Data Level Interoperability** ensures the ability to "consume" the information
  - **Service Level Interoperability** ensures the ability to exchange / obtain the information to be "consumed"
  - **Security Level Interoperability** ensures the ability to the above in a reliable and trustworthy fashion
- Implementation of all levels can be done by using standards from the OGC and other bodies

Geo-Cloud – Major Security Concern

- Multi-tenancy
- Lack of complete control - data, applications, services

Concerns

- Which assets to be deployed in the cloud?
  - Identify: data, applications/functions/processes
- What is the value of these assets?
  - Determine how important the data or function is to the organization
- What are the different ways these assets can be compromised?
  - Becomes widely public & widely distributed
  - An employee of the cloud provider accessed the assets
  - The processes or functions were manipulated by an outsider
  - The info/data was unexpectedly changed
  - The asset were unavailable for a period of time

Thank You !