Mobile Computing is Changing Cloud Computing

Li Erran Li

Collaborators: Victor Bahl (MSR), Rick Han (Colorado),
M. Satyanarayanan (CMU)

Alcatel·Lucent

Mobile Cloud Computing (mCloud) today

- Apple iCloud
  - Store content in cloud and sync to all registered devices
  - Hosted by Windows Azure and Amazon AWS
  - iCloud Storage APIs support third-party app document syncing

Mobile Cloud Computing today (Cont'd)

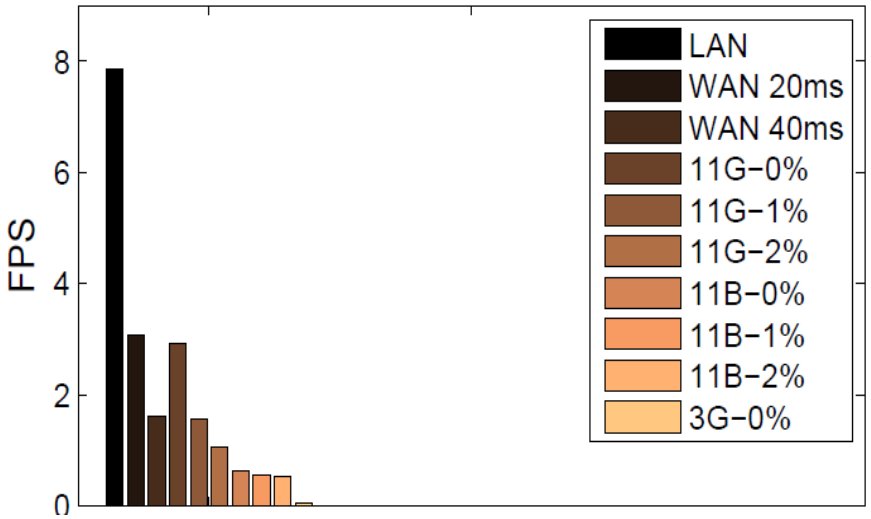
- Amazon Silk browser
  - Accelerates web access
  - Learns user behavior and precache
  - Intelligently partition work between local and Amazon cloud

mCloud Fundamental Challenges

- What architecture best supports mCloud?
- What programming model best enables client to tap mCloud resources?
- What are basic services or building blocks for mCloud?
- What best supports service interaction?

mCloud Architecture

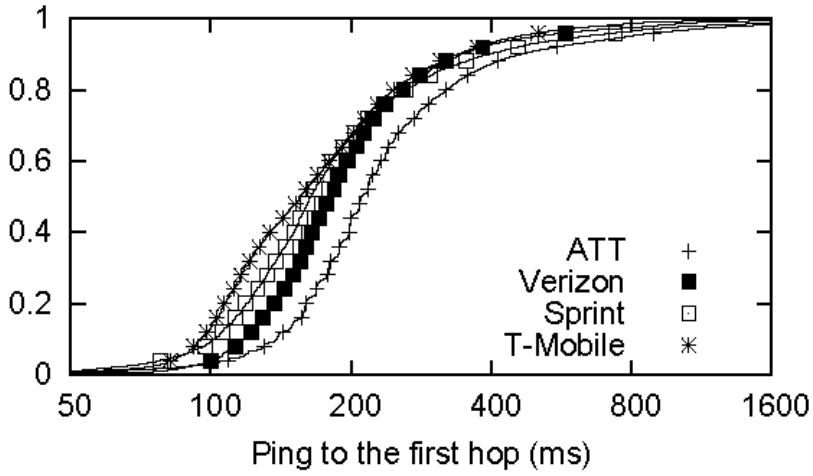
- Resource-intensive mobile applications
  - Face recognition for social networking app
  - Gesture recognition for control media app
  - Object and post recognition for augmented reality app
- End-to-end latency and througput matters for crisp interaction
  - Augmented reality need to display results within 1 sec
  - Need high data rate processing capability, low frame rate can miss gesture



Delay, loss on frame rate of video stream transfer [Odessa, 2011]

mCloud Architecture (Cont'd)

- WAN performance
  - First hop latency in 3G is 200ms
  - Verizon LTE :128ms, 6.44 Mbps downlink, 5Mbps uplink [Pcworld, March 2011]
- There is a need for a middle tier
- cloudlet = (compute cluster + wireless access point + wired Internet access + no battery limitations)
- → *"data center in a box"*

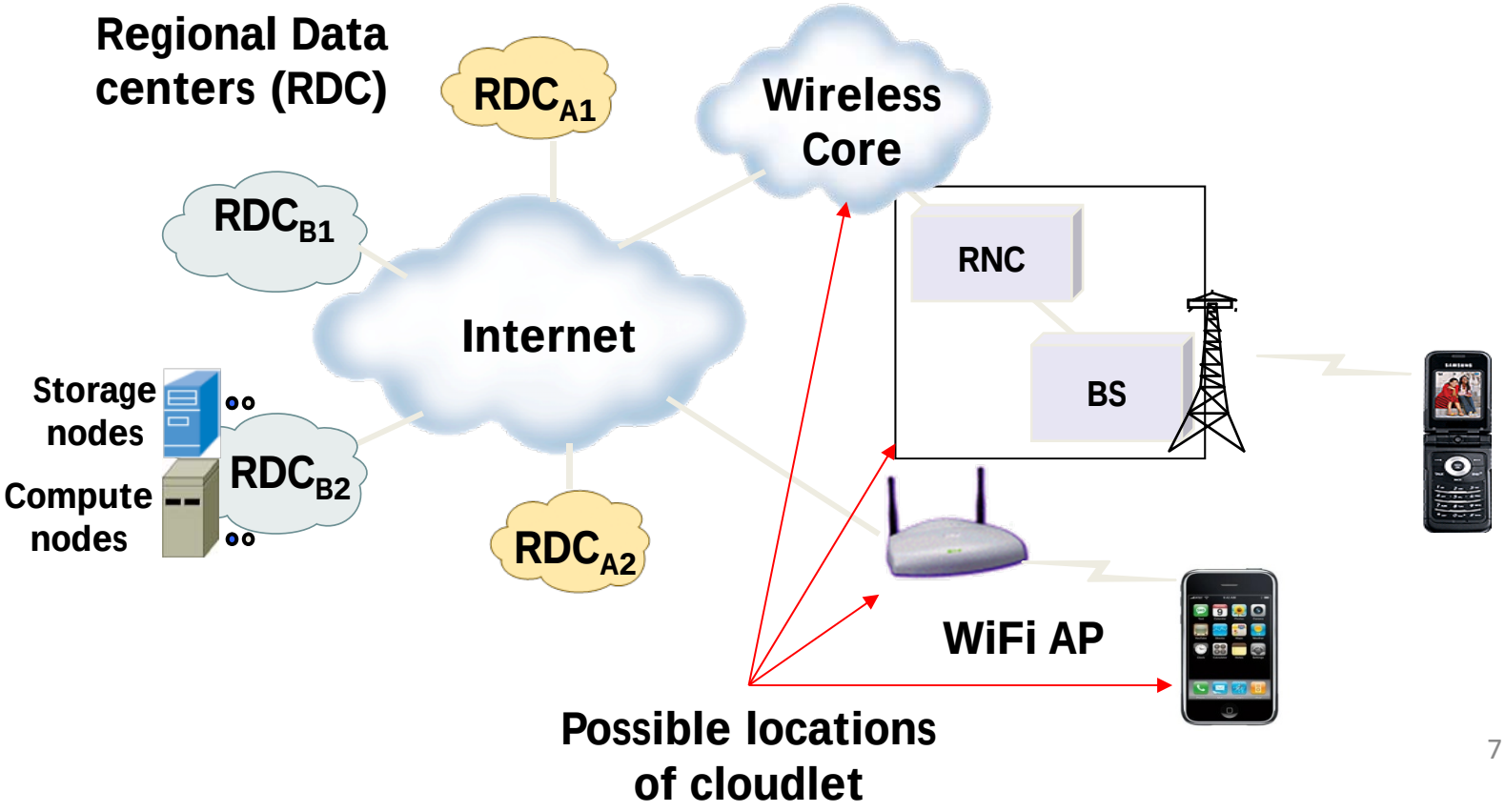


(f) CDF of Ping latency to the first hop

"Anatomizing Application Performance Differences on Smartphones", MobiSys 2010 (Huang et al)

mCloud Architecture (Cont'd)

- Cloudlet possible locations
- Cellular providers has a unique advantage



mCloud Programming Model

- MAUI: RPC based offloading architecture
- Odessa: data-flow graph to exploit parallelism in perception applications
- CloneCloud: tight synchronization between cloud and phone
- Orelans: a new programming model based on grains
- MAUI, CloneCloud , Odessa all have profiler, solver

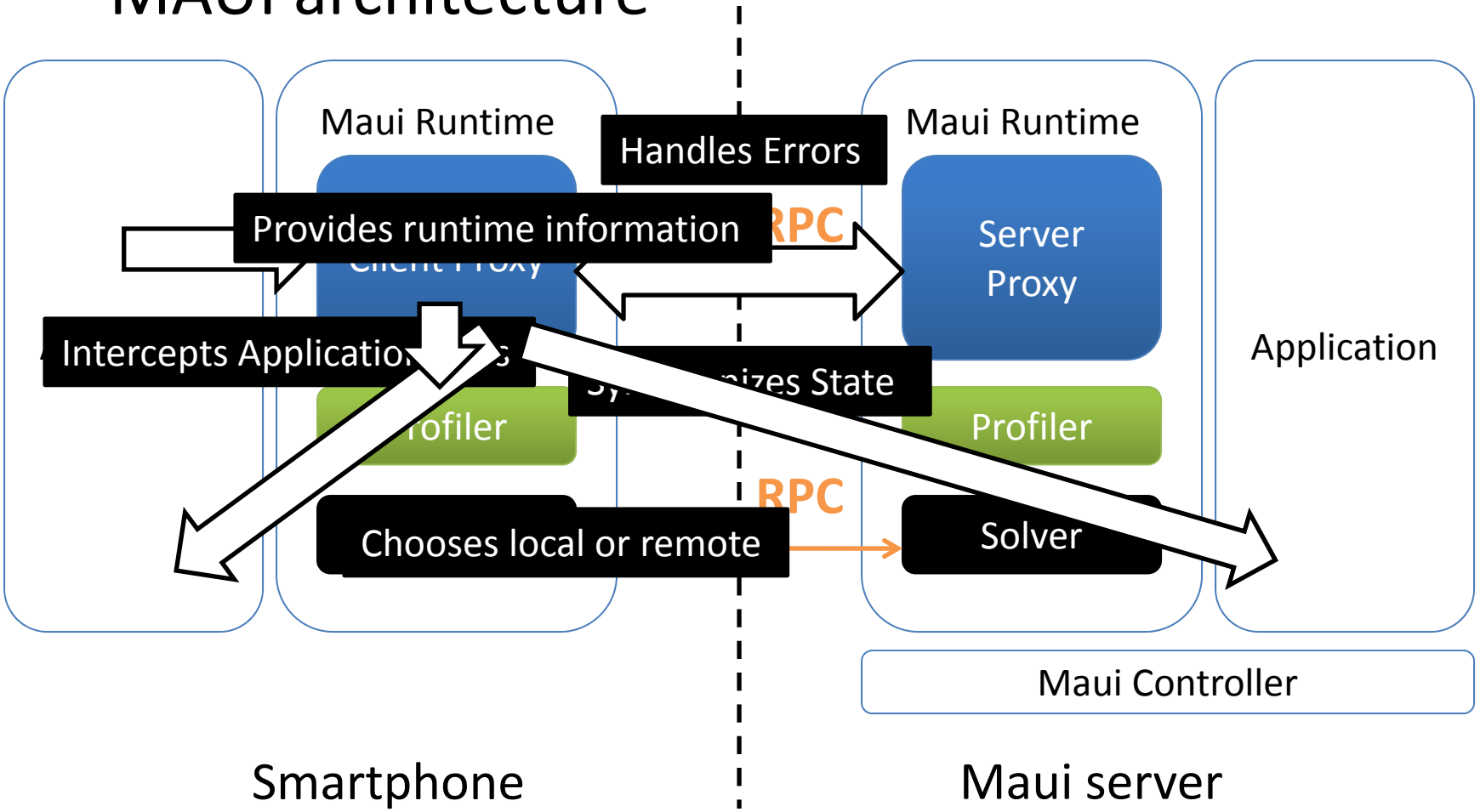
	MAUI	CloneCloud	Odessa	Orleans
Remote execution unit	Methods (RMI)	Threads	Tasks	Grains

mCloud Programming Model: MAUI

- Combine extensive profiling with an ILP solver
  - Makes dynamic offload decisions
  - Optimize for energy reduction
  - Profile: device, network, application
- Leverage modern language runtime (.NET CLR)
  - To simplify program partitioning
  - Reflection, serialization, strong typing
    - Identifies methods with [Remoteable] tag
    - Automates generation of RPC stubs

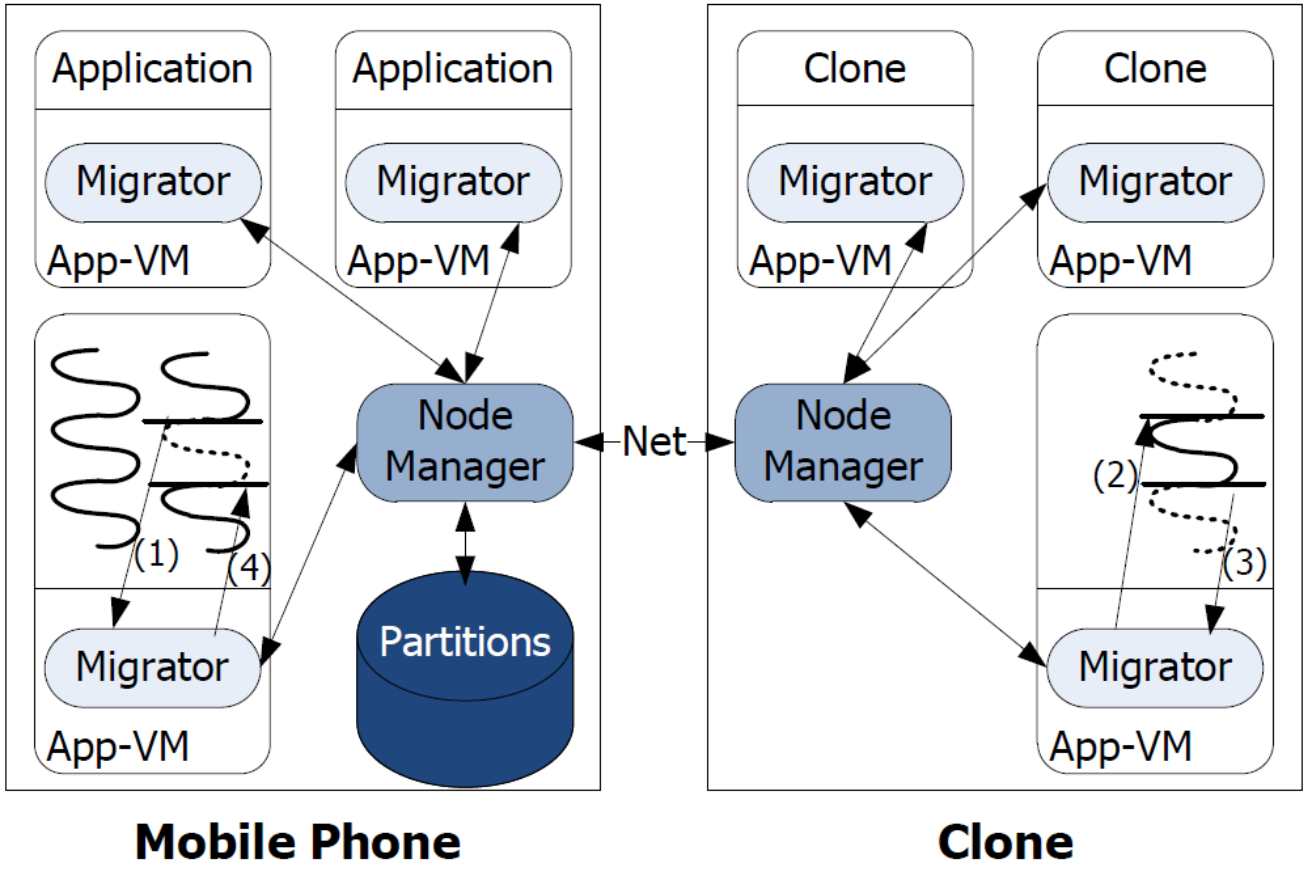
mCloud Programming Model: MAUI (Cont'd)

- MAUI architecture



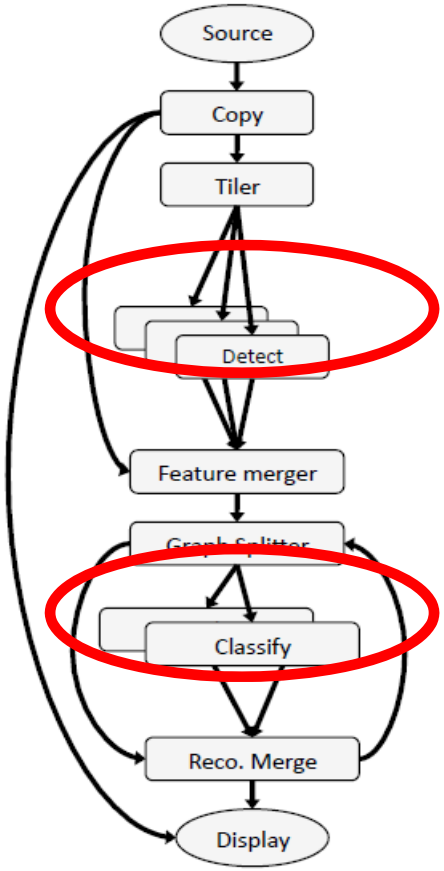
mCloud Programming Model: CloneCloud

- Offloading decision done at beginning of execution

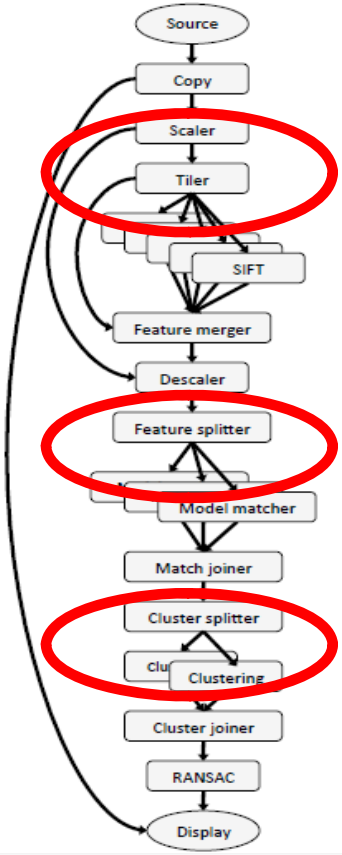


mCloud Programming Model: Odessa

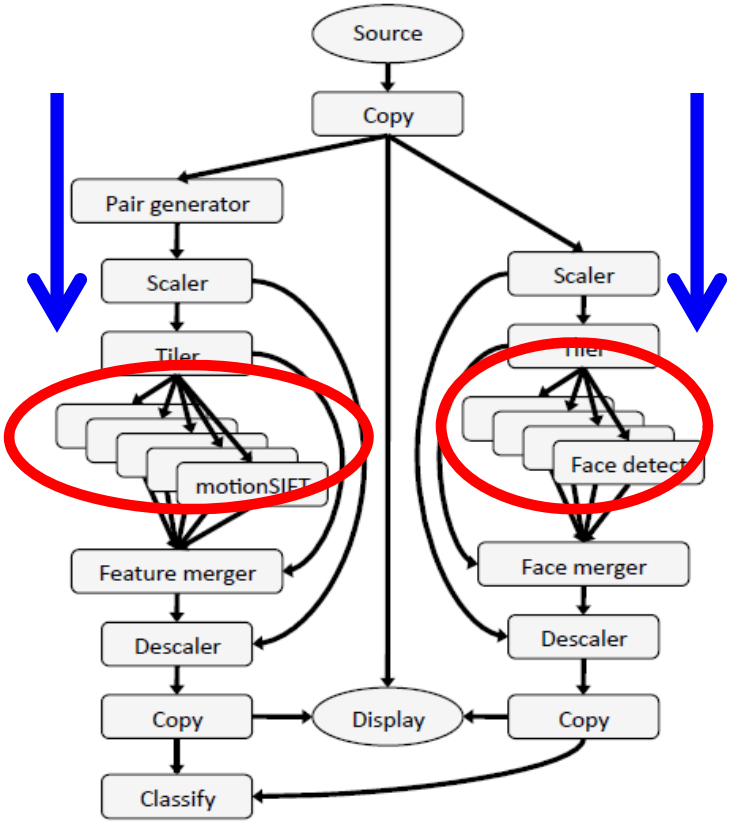
- Data flow graph: vertices are stages; edges are connectors; stages share nothing



Face Recognition



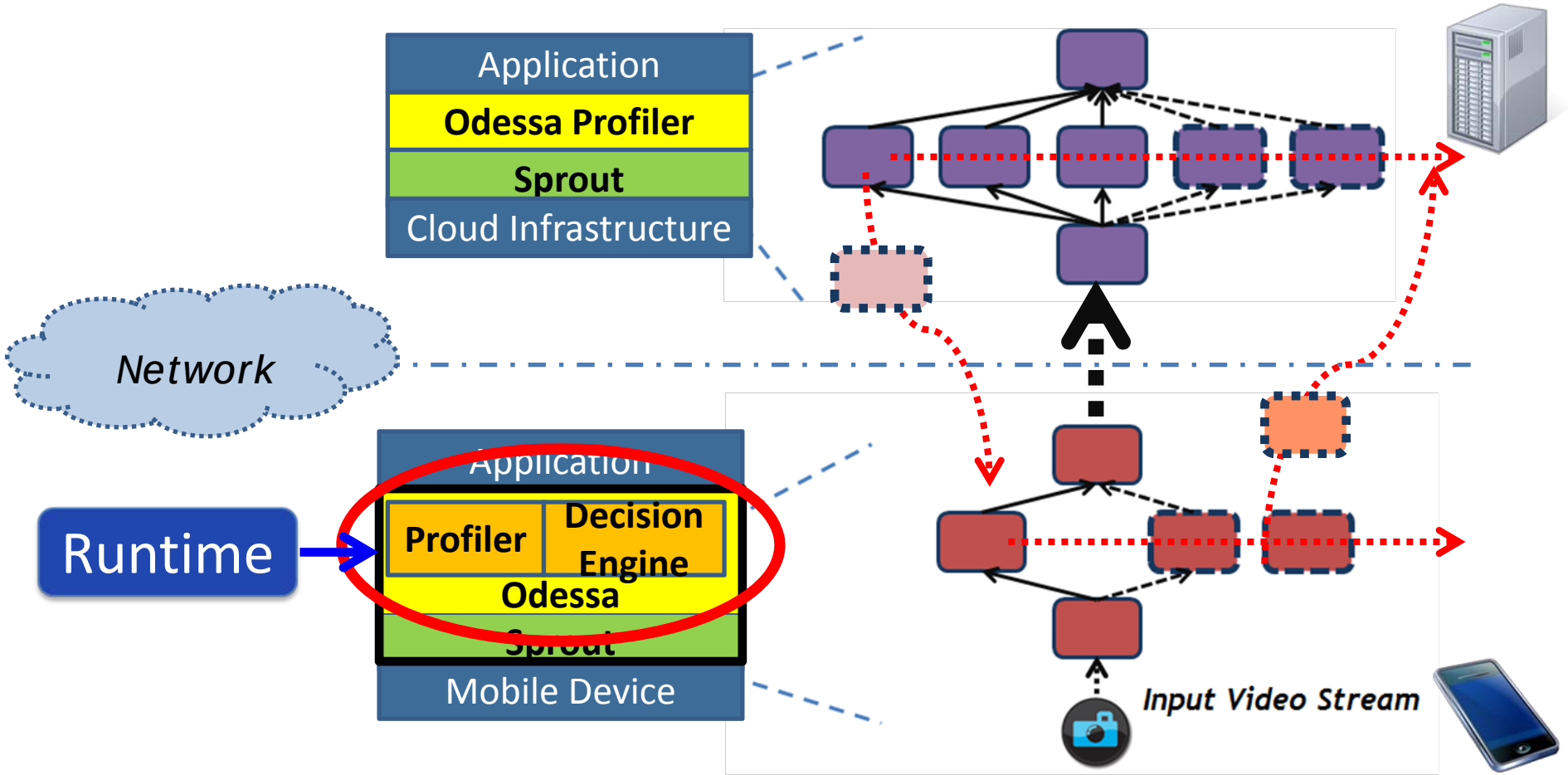
Object Pose Estimation



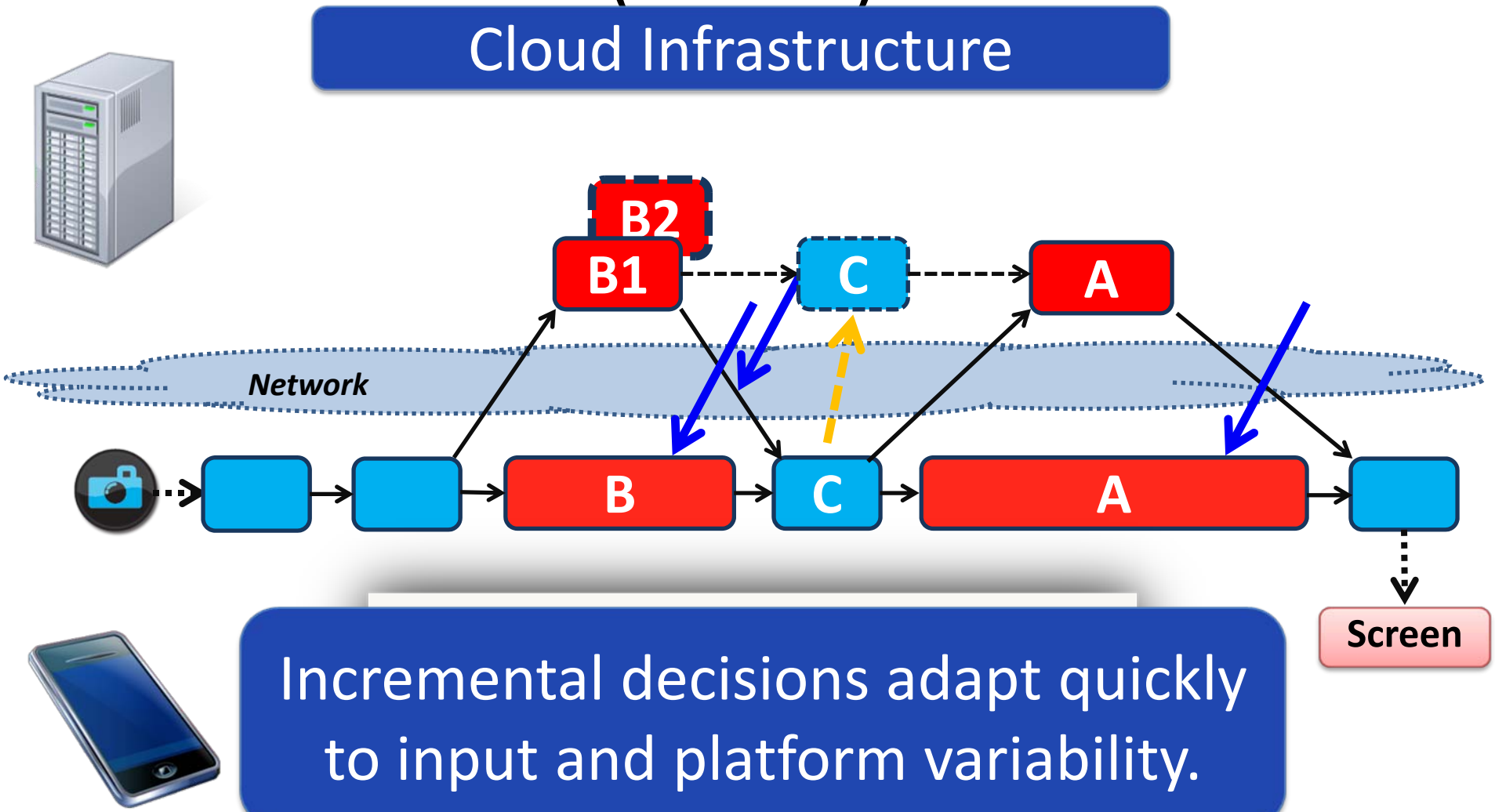
Gesture Recognition

mCloud Programming Model: Odessa (Cont'd)

- Offloading DEcision System for Streaming Applications

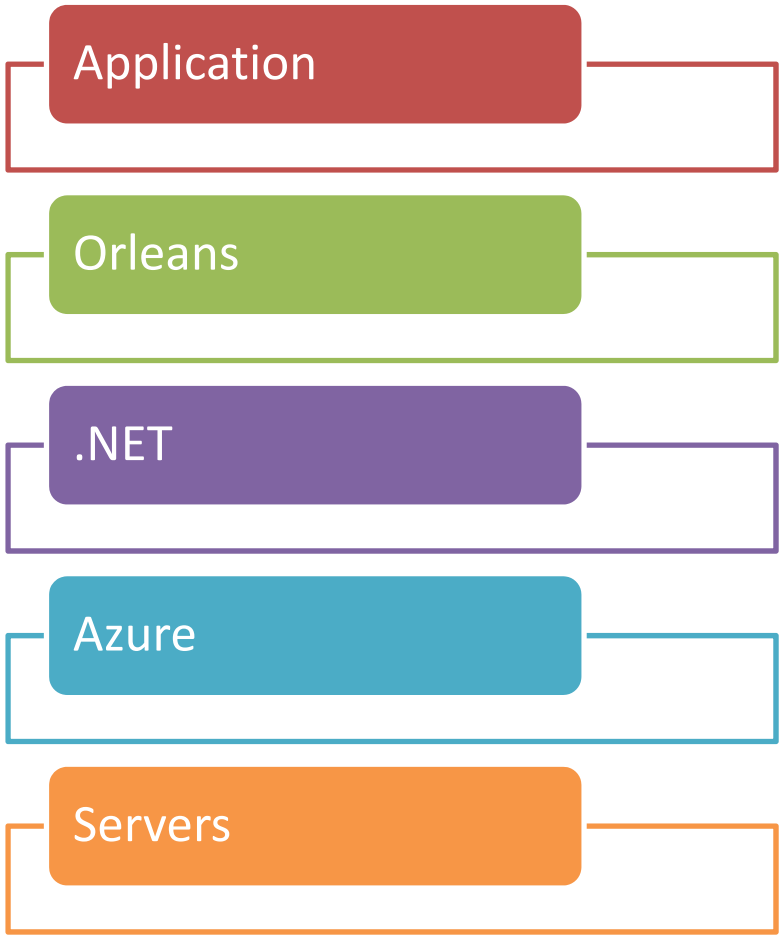


mCloud Programming Model: Odessa (Cont'd)



mCloud Programming Model: Orleans

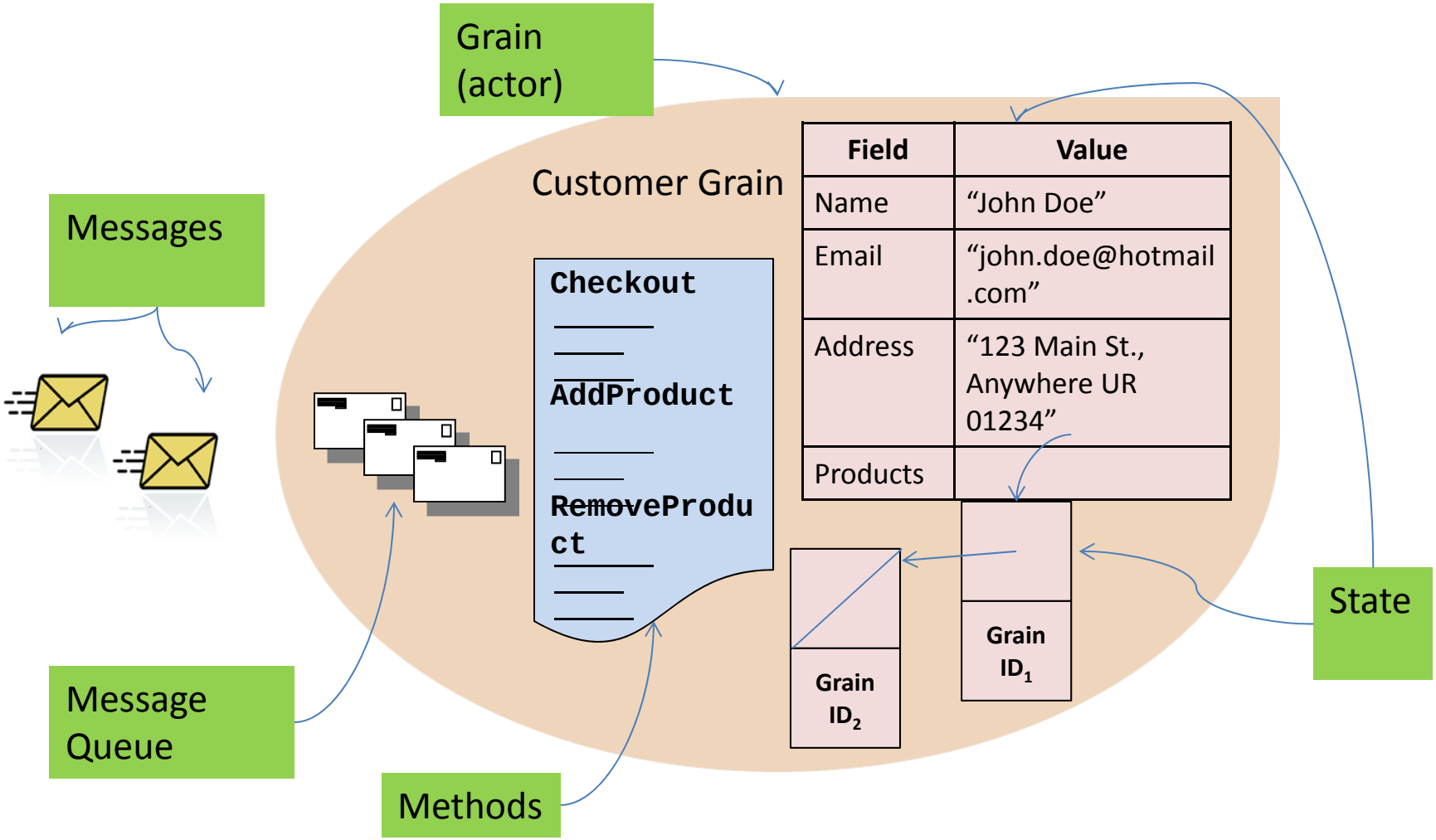
- Software framework and runtime to make cloud programming easier and more productive
- Shift burden of correctness and performance from developer to Orleans system
- Experimental system from Microsoft Research
- Radically simplified, prescriptive programming model
  - Actors
  - Asynchronous messaging
  - Lightweight transactions
  - Persistence
  - Adaptive performance management



mCloud Programming Model: Orleans (Cont'd)

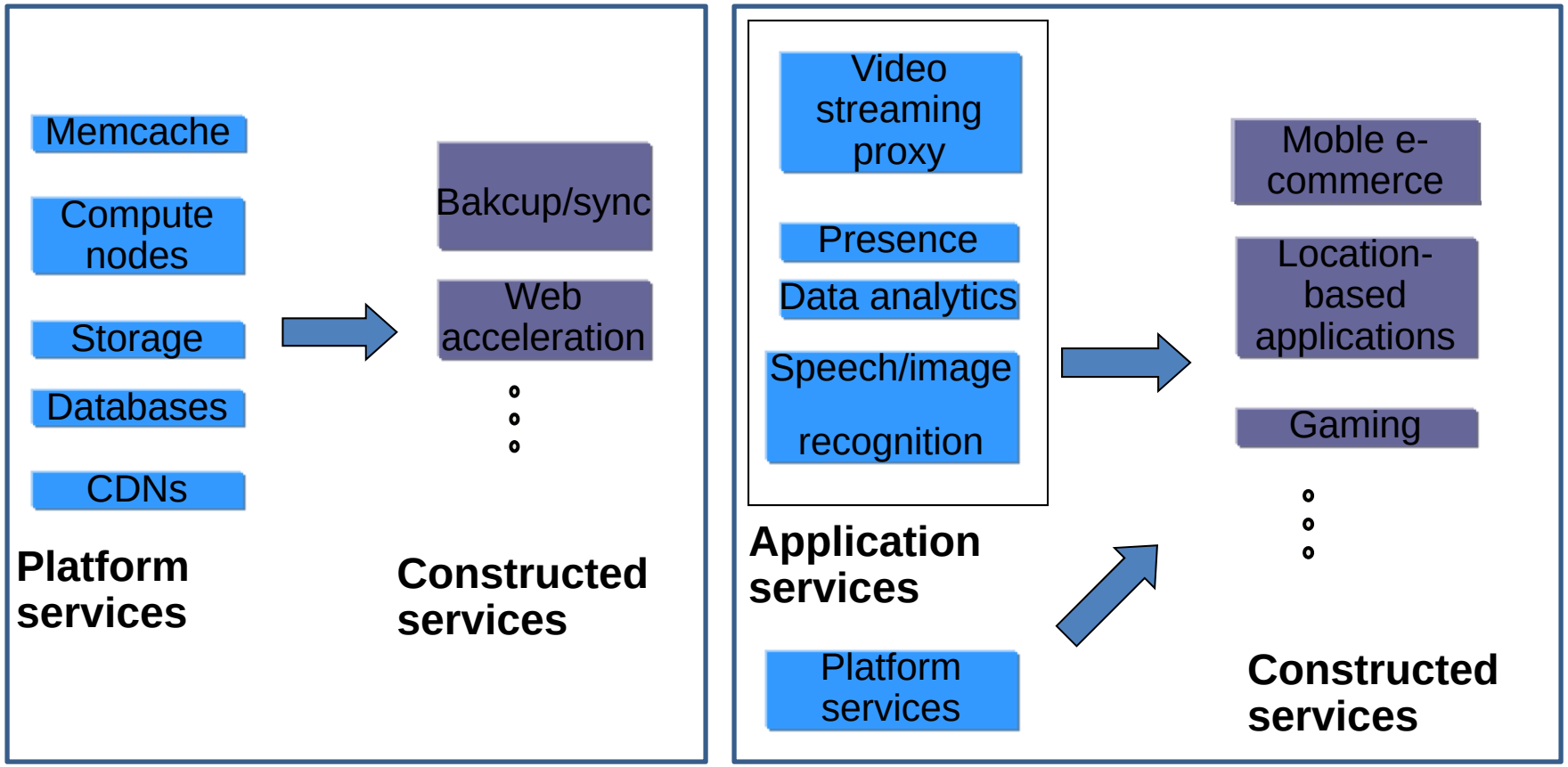
- Actor based: fine-grain distributed objects are natural abstraction
- Grains partition data
- Secure and isolated computation with clear points of communications
  - Enable computation replication
- Natural integration with persistent storage
  - Grain resides on disk until activated

mCloud Programming Model: Orleans (Cont'd)



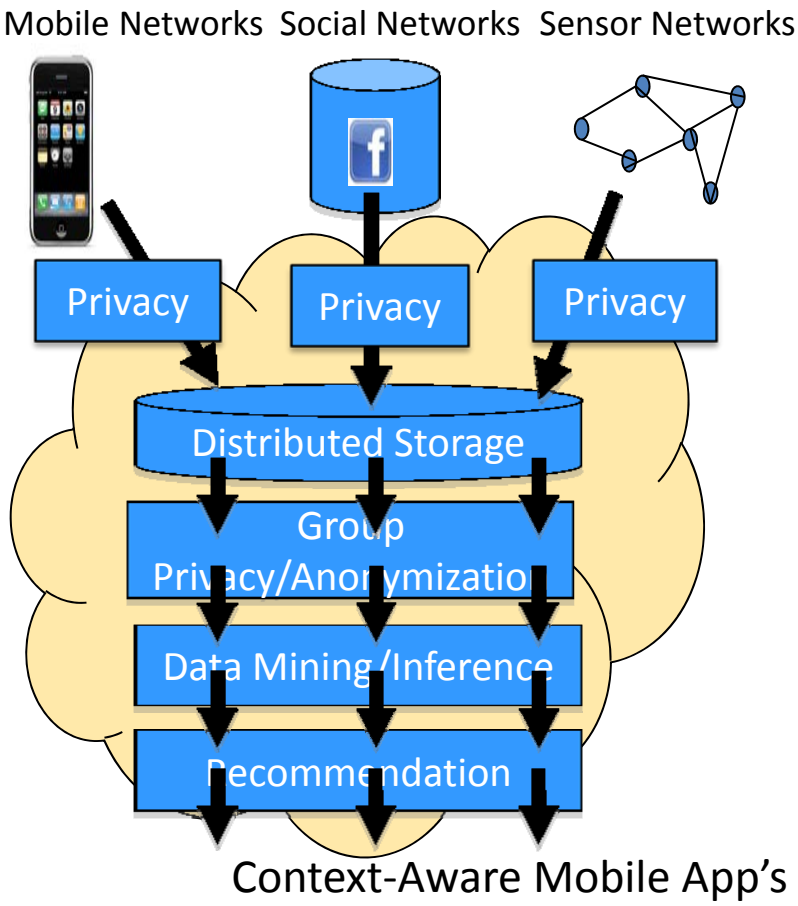
mCloud Building Blocks

- Basic services: platform services and application services



mCloud Building Blocks (Cont'd)

- Construct more sophisticated services
  - Context-aware mCloud services
  - Crowd-sourced mCloud services



Support for Service Interaction

- Synergy of mCloud services
  - Shared basic services enables more efficient usage of cloud resources(e.g. shared memcache eliminates data duplication of individual services)
  - Co-location makes mashed-up applications to achieve native performance (file transfer becomes an object reference)
- mCloud offers shared platform services
- mCloud optimizes service interaction through active VM migration

Closing Thoughts

- Native phone VMs in the cloud to enable better interaction between cloud and phone
- Cloud computing needs to advance in architecture, programming model, platform services to revolutionize mobile computing