

DYNAMIC VM SYNTHESIS FOR CLOUDLET

Kiryong Ha, Padmanabhan S Pillai, Mahadev Satyanarayanan (CMU)

MOTIVATION

- Rich, interactive applications using video/voice are emerging in mobile contexts, but are too expensive to run on clients alone
- They do not work well with current, centralized cloud computing infrastructures → high latency and limited
- Let's bring the cloud to the mobile users → Cloudlet!



Key Challenges

- **Management**
 - Decentralization requires extreme standardization
- **Personalization**
 - Much customization needed, even for commercial apps
 - E.g., preferences, speech tuning, domain-specific vocabulary

Our Approach

- Use a personal VM and take it with you everywhere

PROTOTYPE SYSTEM



EXPERIMENT

Overlay Size

Application	Program size (MB)	OS	Base Disk (GB)*	Base Memory (MB)**	Compressed Overlay (MB)
Gimp	55.2	Ubuntu	2.5	476	142
MOPED	27.5	Ubuntu	2.5	476	173
FACE	17.66	Windows XP	2.1	279	109
Null	0	Ubuntu	2.5	476	0.32

* 8GB Disk with qcow2

** Memory snapshot of 2GB Main Memory

VM Synthesis Time

Application	Overlay Transfer (s)	Decompression (s)	Apply Delta (s)	Run KVM (s)	Total (s)
Gimp	38	17.2	18.7	2.1	76.0
MOPED	44.9	23.3	17.9	4.0	90.1
FACE	30.3	14.5	9.7	4.0	58.5
Null*	0.1	0.04	10.7	4.0	15.6

* Null case is for comparison and does not install anything

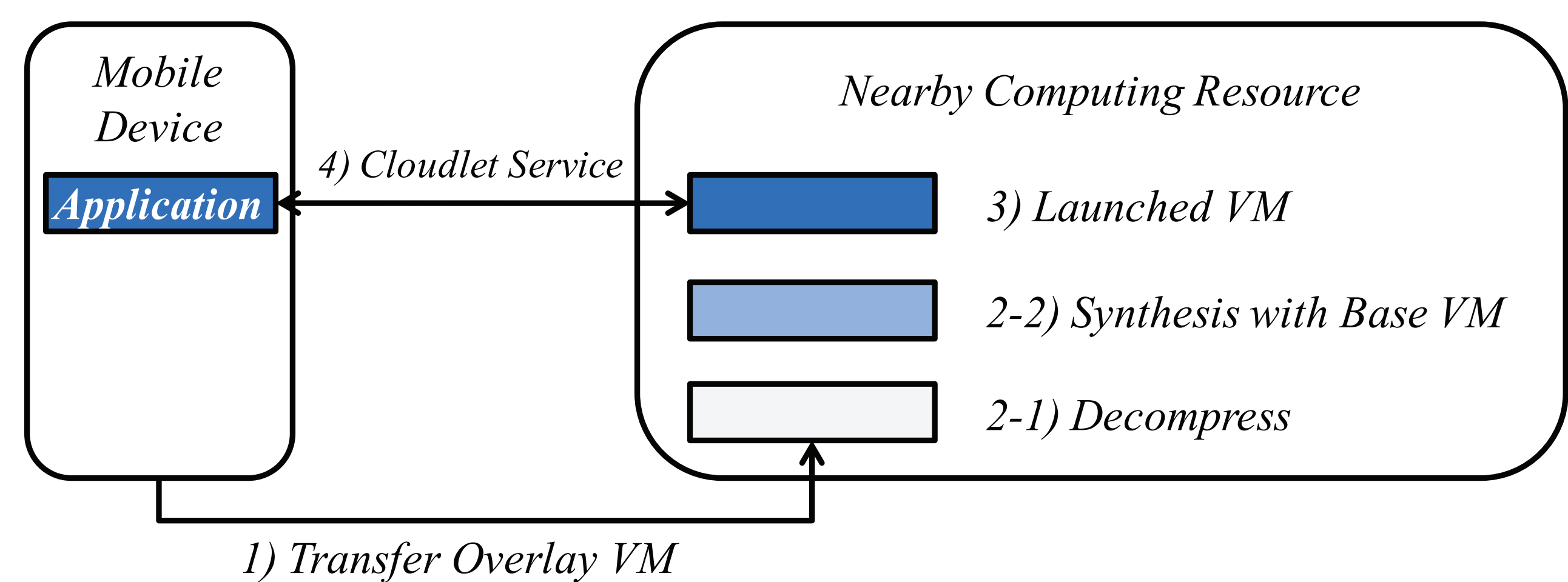


DYNAMIC VM SYNTHESIS

Problems in VM: too big to carry, transfer, and launch

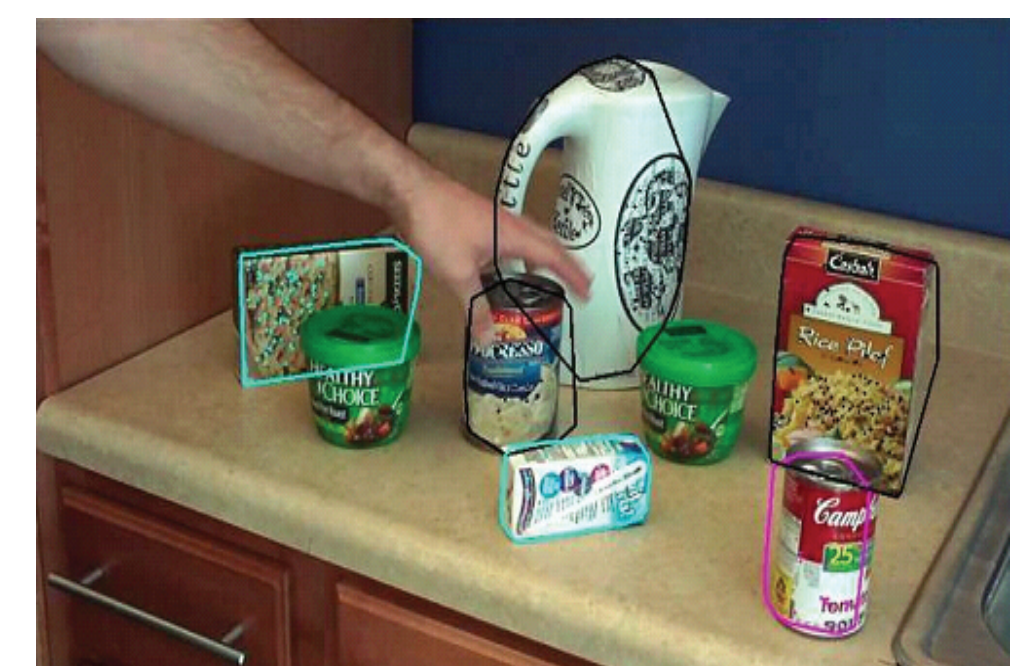
Solution: Dynamic VM Synthesis

- **Base VM:** Cloudlet pre-fetches large, widely-used VM
- **VM overlay:** Mobile device delivers small patch just before use
- Cloudlet discards VM after use/Or caches for future reuse



PROTOTYPE APPLICATIONS

- **Object-Recognition (MOPEd)**
 - Object recognition and pose estimation program (CMU RI)
- **Face Detection and Recognition**
 - OpenCV
- **Application characteristics**
 - Requires high computation and interactivity
 - Future Candidates
 - Indoor navigation
 - Language translation ...



WORK IN PROGRESS

Goal: Reducing VM synthesis time

- Discovery, transfer, and launch to happen within 5 seconds

Work in progress

- Exploit parallelism for VM synthesis
- Multi-layer overlays
 - Potential to reuse, improve caching of overlays
 - Reduce size of final client-provided patches
- Launch before completing synthesis
 - Overlap synthesis and execution, hide latency

