COMPUTERSCIENCE

RRP Edge Computing System

Guangli Dai, Pavan Kumar Paluri, Albert M. K. Cheng
University of Houston, Texas, USA

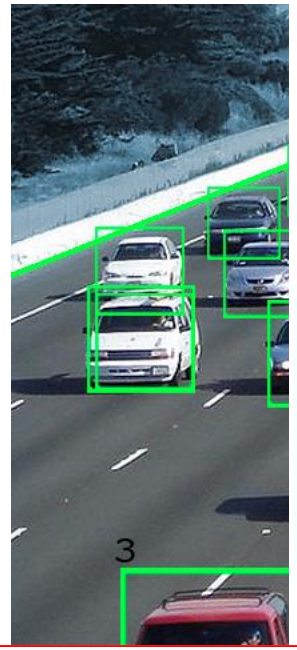
- Outline
  - Motivation
    - Edge Computing System (ECS)
    - Regularity-based Resource Partition (RRP)
  - Architecture
    - RRP-ECS
    - Centralized Task-Partition Architecture
  - Theoretical obstacles and our answer
    - Regular-Partition Periodic Task Mapping (RPM)
    - Regular-Partition Compositional Task Mapping (RCM)

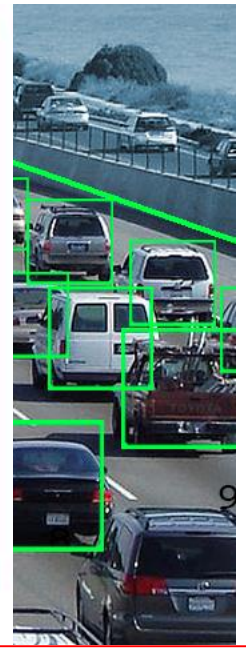
Speaker: Guangli Dai

Motivation

Background:

- Internet of Things (IoT)
- Applications that need more computational power and lower latency.
- e.g.: Smart Surveillance, Real-time City transportation analysis.





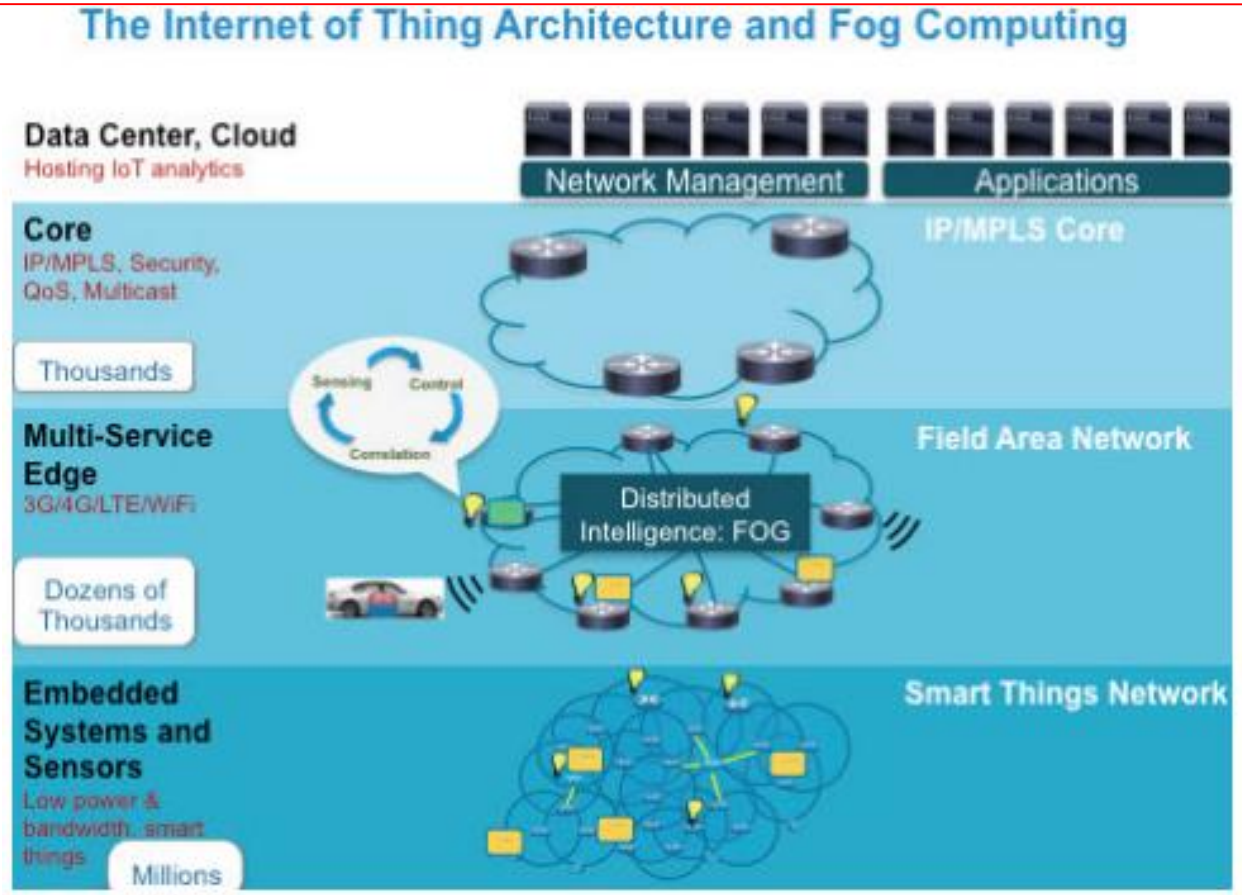
http://www.trafficvision.com

https://www.consumerreports.org/privacy/facial-recognition-who-is-tracking-you-in-public1/

Motivation

Why Edge Computing System?

- Local connection
- Low latency
- Large bandwidth



[Bonomi et al. 2012]

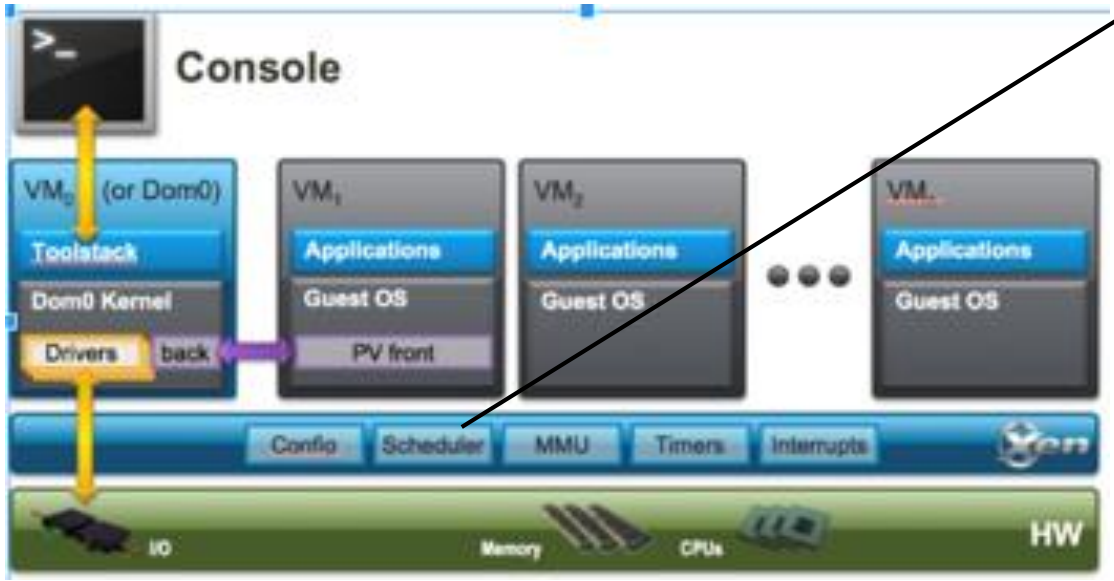
Motivation

What can Regularity-based Resource Partition (RRP) model offer?

- Efficiency:
  - Physical computing resource -> multiple partitions (Magic7, MulZ).
  - Scheduler: Partition-time table.
- Security:
  - Independent partitions with fixed shares of CPU.
- Uniform measurement:
  - Availability factor α and regularity k

Motivation

Saves time in scheduler by using
PT(Partition - Time slice) table.
The table is initialized during the booting.

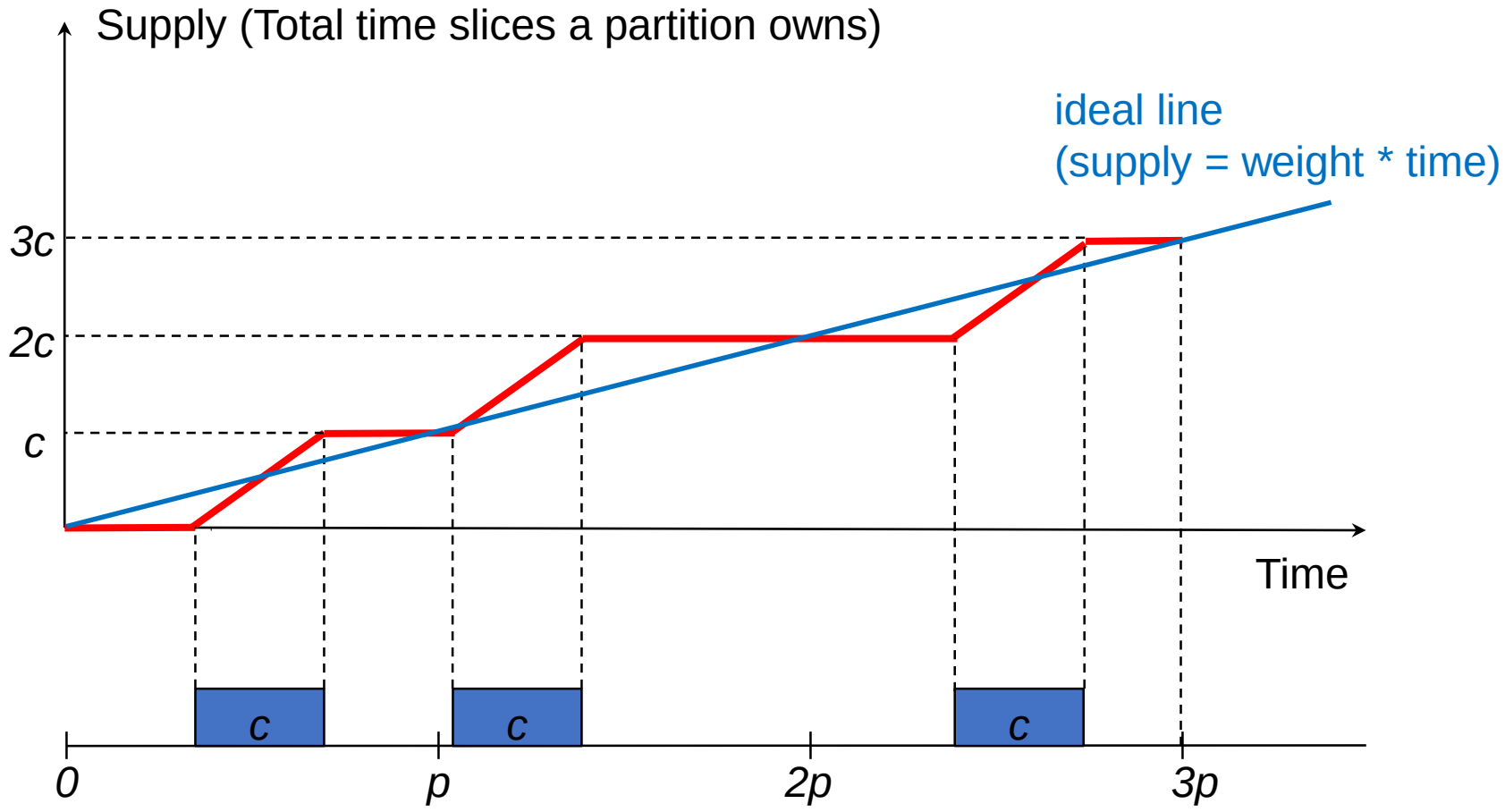


Partition	Time Slice
P1	0
P2	1
P1	2
so on..	⋮

Quantum=10Millisecs

https://wiki.xen.org/wiki/Xen_Project_Software_Overview

Motivation



Availability factor: $\frac{c}{p}$

[Li&Cheng 2017]

Architecture

How to construct an RRP-ECS?

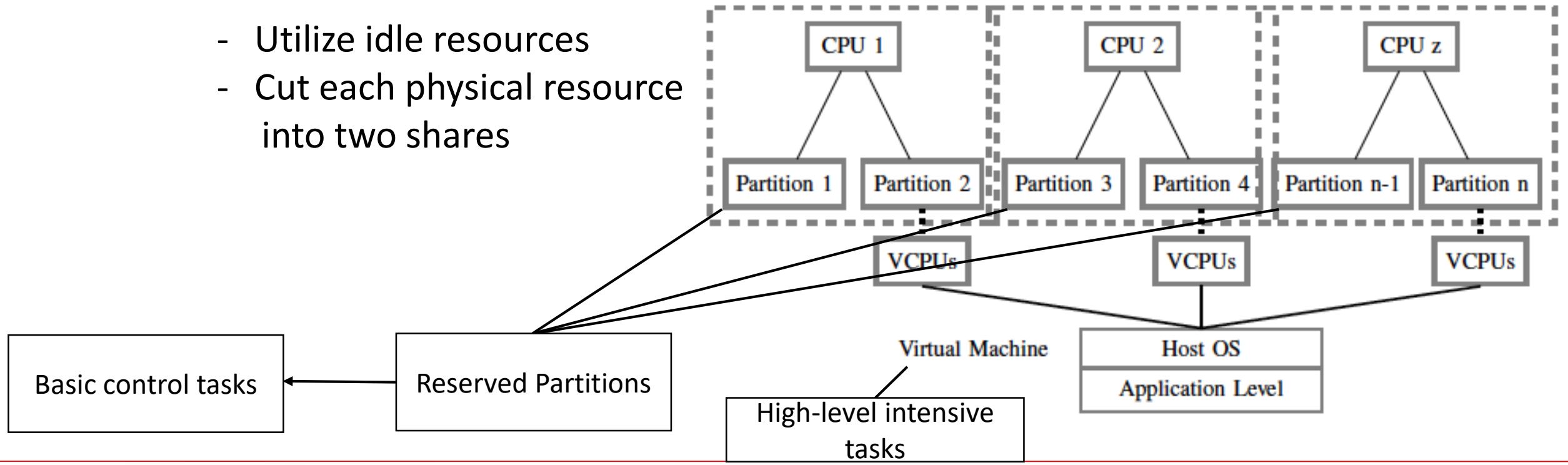
What are tasks for smart objects?

1. Basic control tasks: fundamental and low-cost.
2. High-level intensive tasks: e.g., build neural network based on the data collected.

Architecture

How to construct an RRP-ECS?

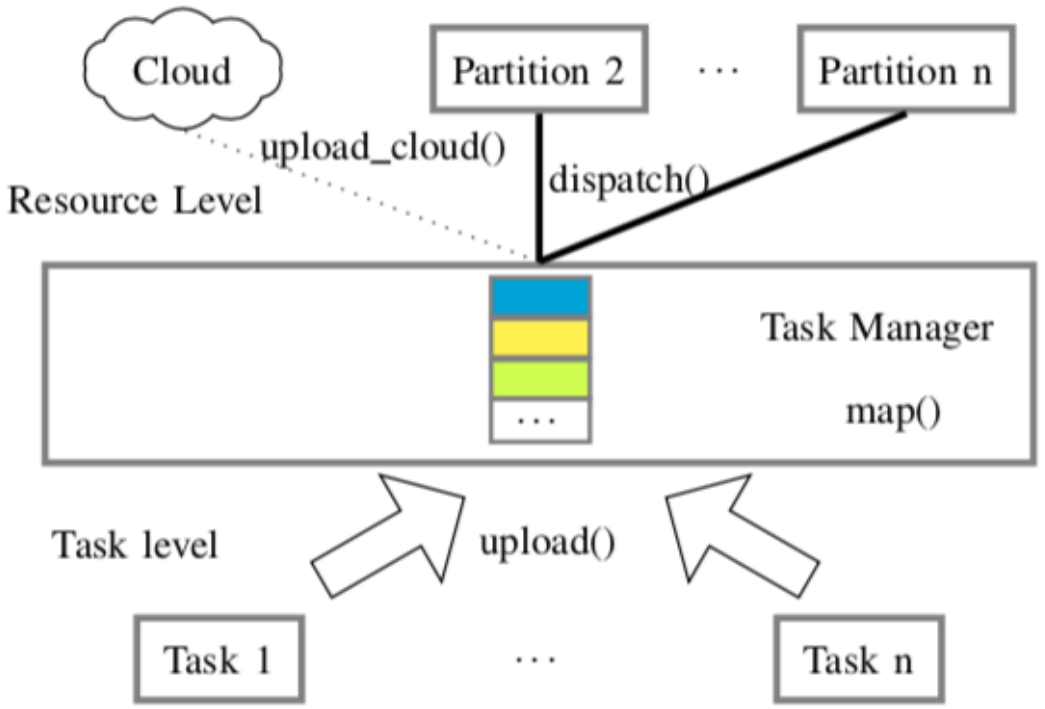
- Utilize idle resources
- Cut each physical resource into two shares



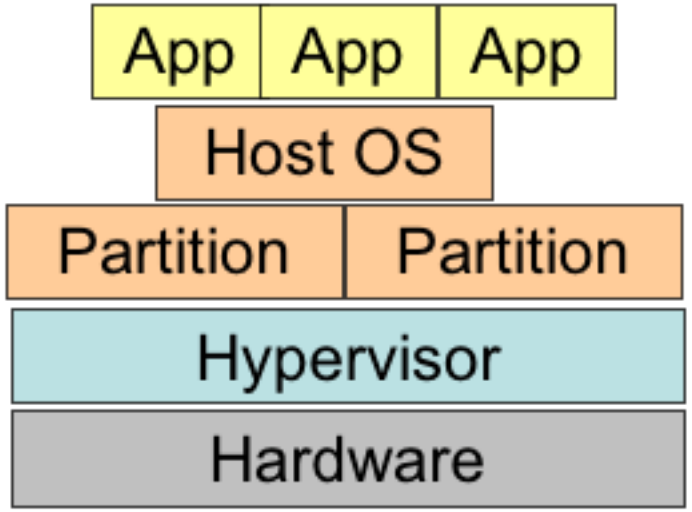
Architecture

Centralized Task-Partition Architecture

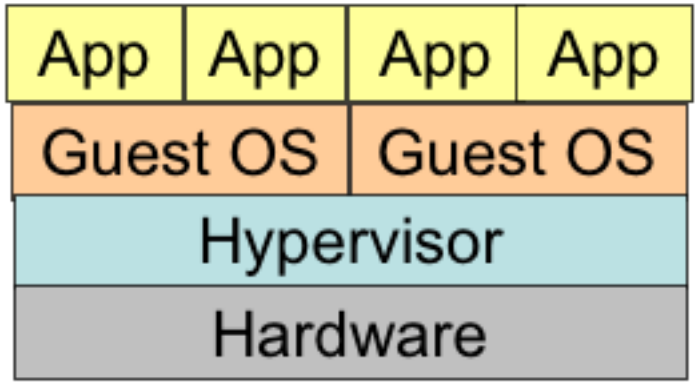
- Transparency
- Easy to implement with Xen
- Portability



Architecture



Proposed Architecture



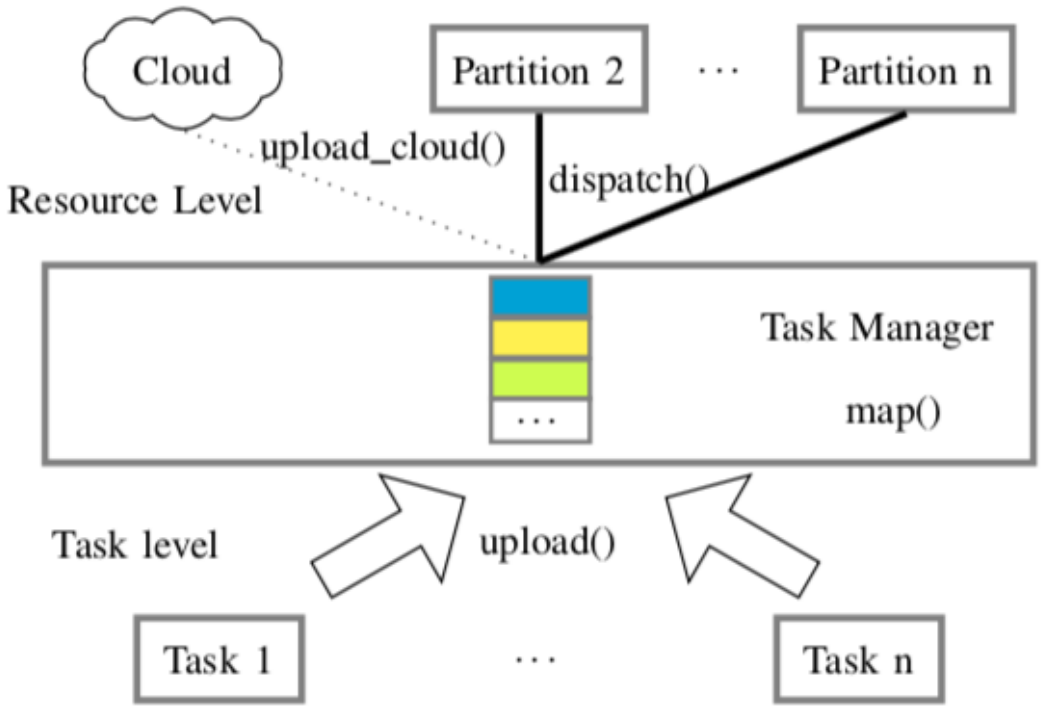
Existing Xen Architecture

X

Architecture

Centralized Task-Partition Architecture

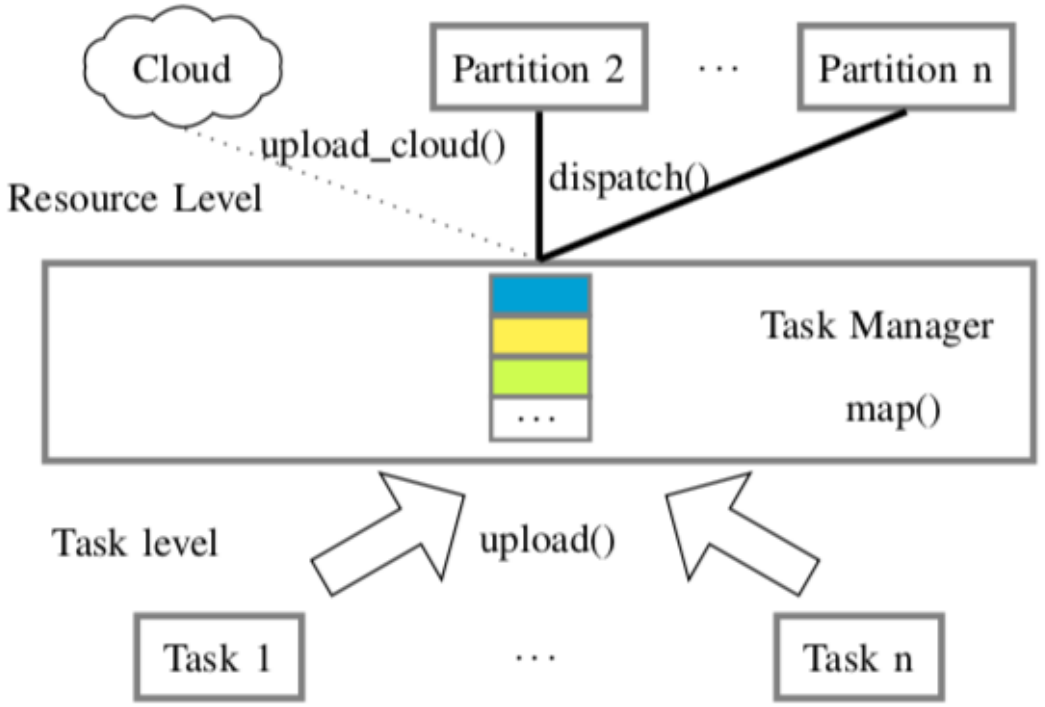
- upload()
- map()
- dispatch()
- upload_cloud()



Theoretical obstacles

map():

- Scheduling tasks on a partition is the same as that on a core.
- Find a partition a task should be assigned to.
- Similar to an online Multiple Knapsack Problem (MKP).



Theoretical obstacles

Schedulability test:

For each task T_j,

c_j: worst case execution time (WCET) of T_j.

p_j: relative deadline of T_j.

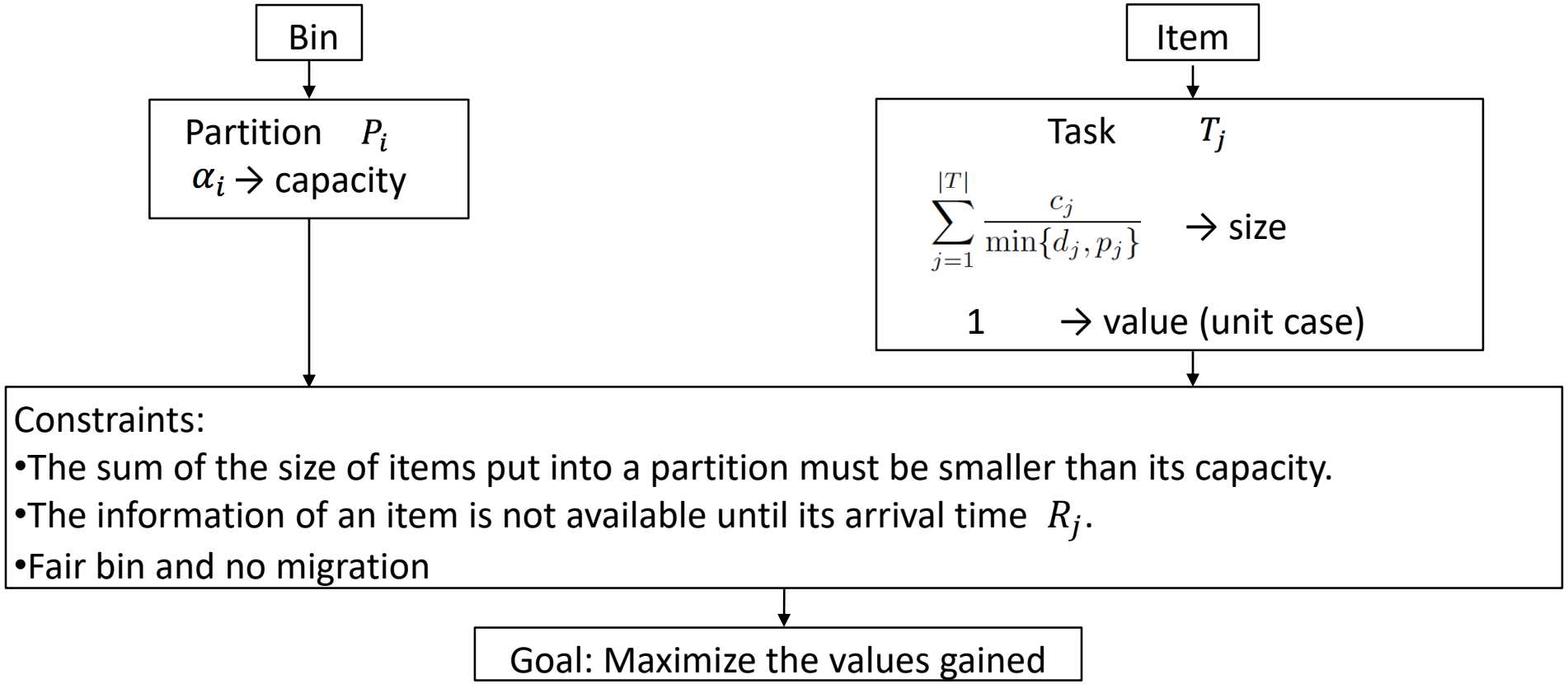
d_j: period of task T_j.

Theorem 2: A task set T is schedulable on a regular partition $\mathbb{P}_j$ if

$$\sum_{i=1}^{|T|} \frac{c_i}{\min\{d_i, p_i\}} \leq \alpha_j \qquad (1)$$

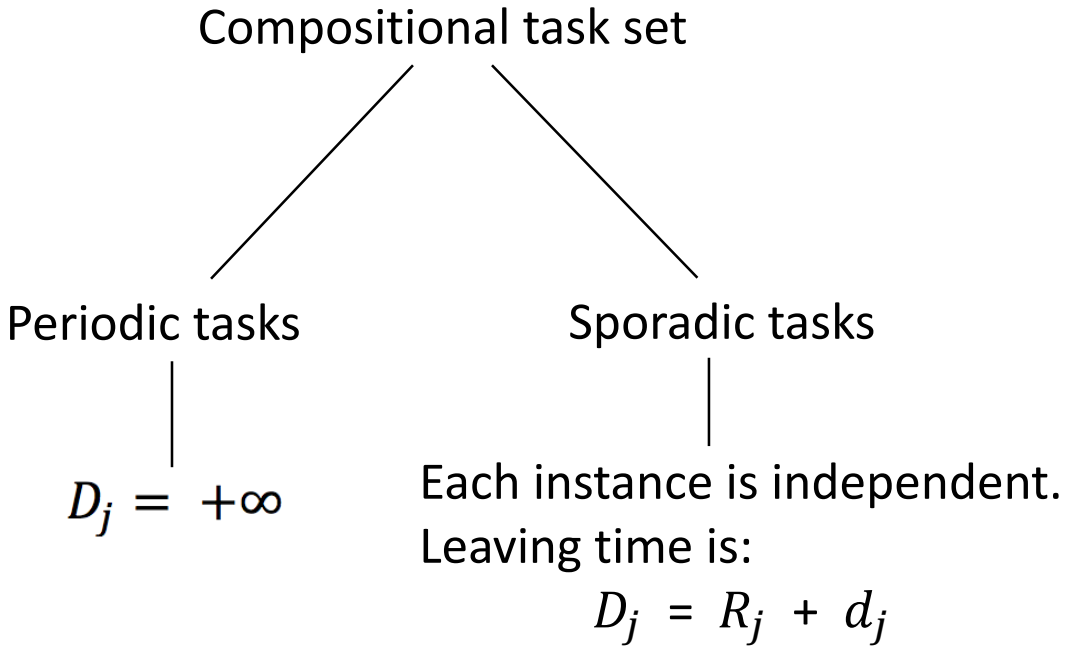
Theoretical obstacles

Regular-Partition Periodic Task Mapping (RPM)



Theoretical obstacles

Regular-Partition Compositional Task Mapping (RCM)



- Include sporadic tasks into the system
- After D_j , task T_j leaves the system forever.

Theoretical obstacles

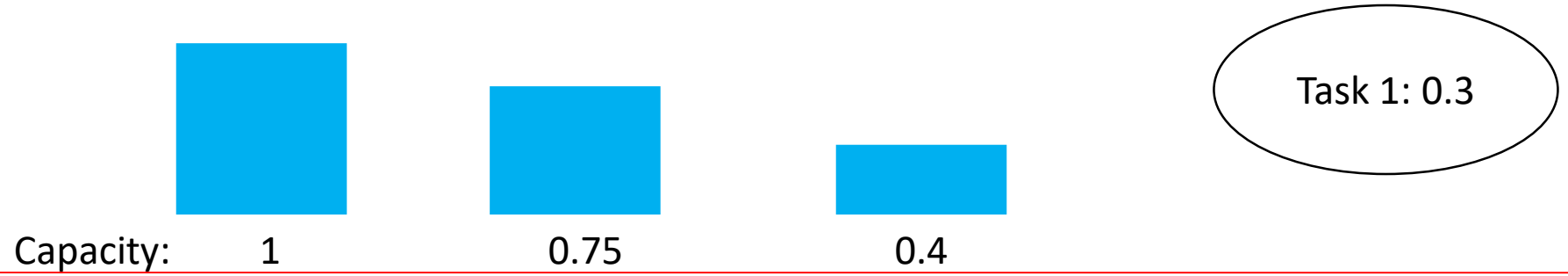
What makes RPM and RCM different?

1. Online model.
2. Both the bin capacity and item size here do not have a basic unit (no integer assumption).
3. Irremovable entities and fair constraints.
4. In RCM, items leave the bin.

Theoretical obstacles

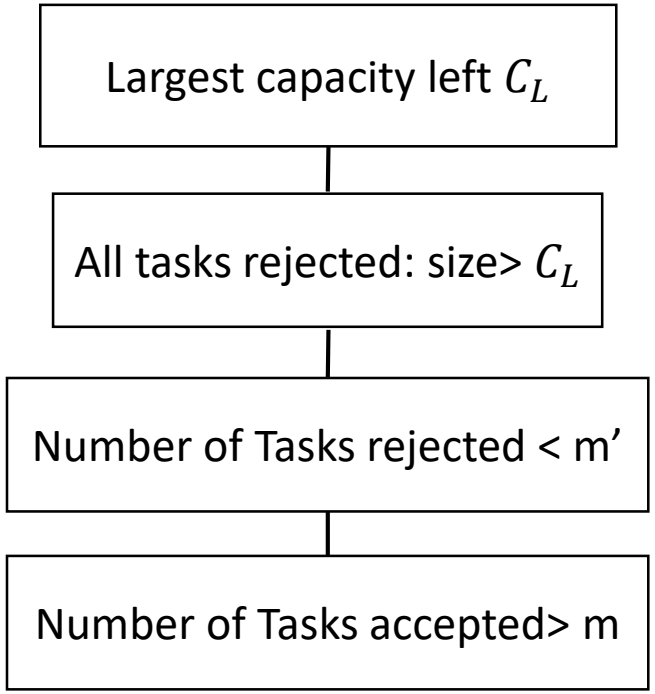
Algorithm Best Fit (BF):
Always find the partition that can just accommodate the task.

Lower bound of Best Fit:
With the same partition set, if the optimal algorithm can accommodate the whole task set, BF can at least accommodate half of it.

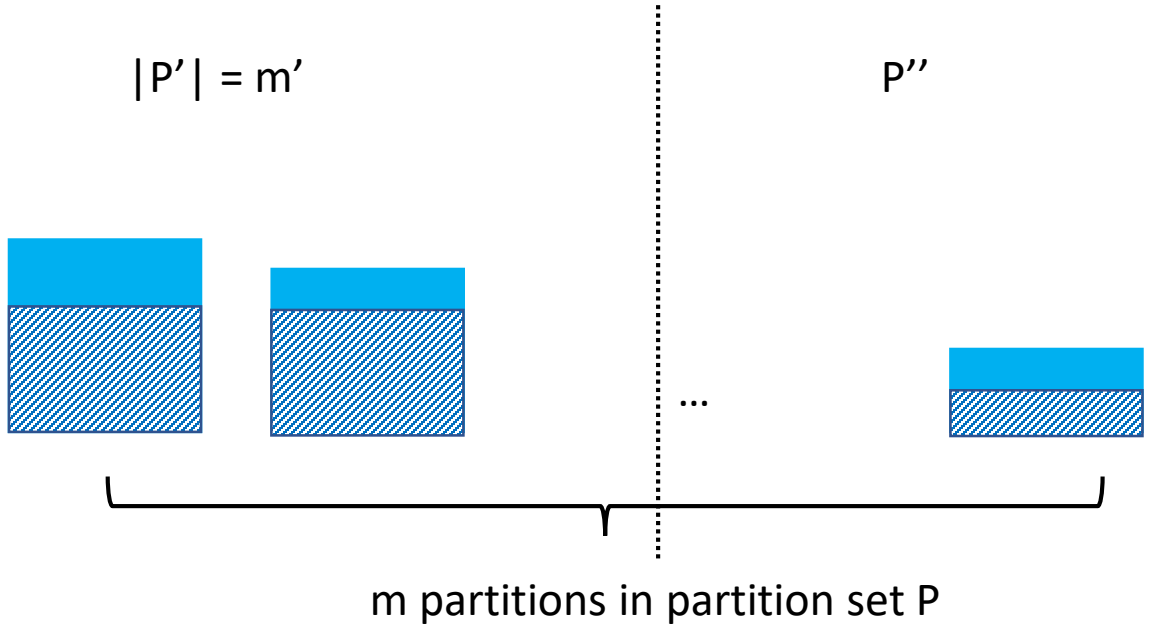


Theoretical obstacles

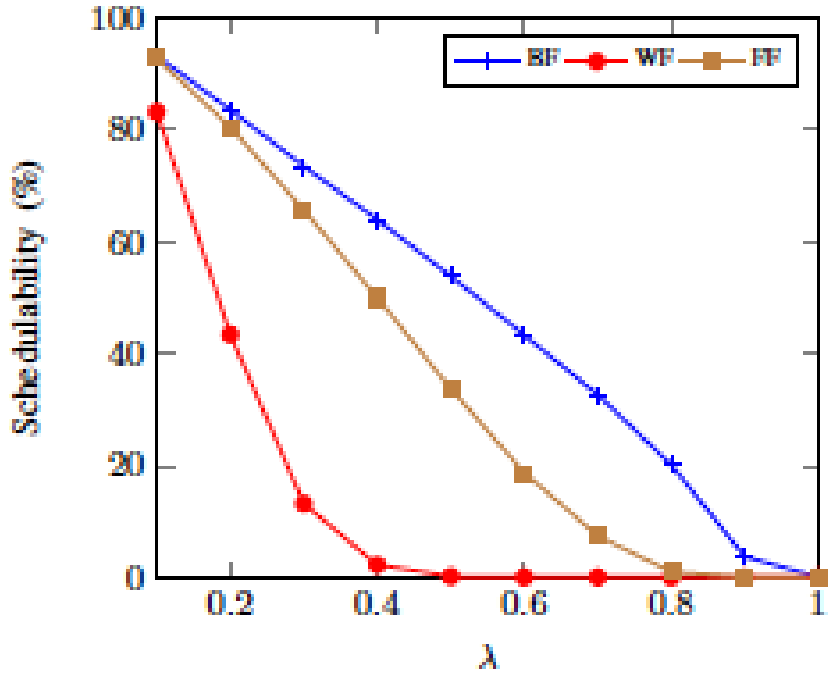
Algorithm Best Fit (BF) competitive ratio $\frac{1}{2}$ on accommodating sequence.



$$\frac{m}{m+m'} \geq \frac{1}{2}$$

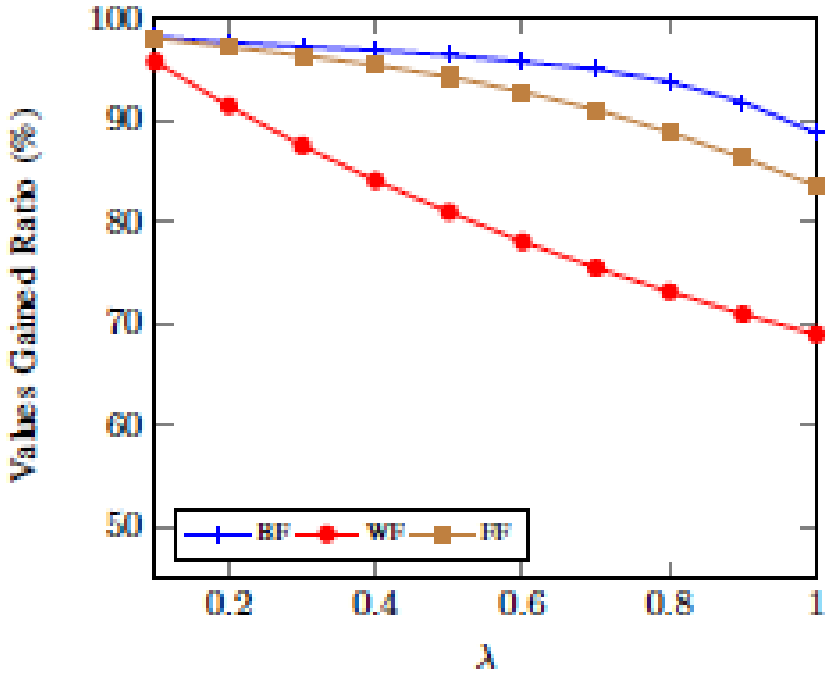


Theoretical obstacles



(a) Schedulability

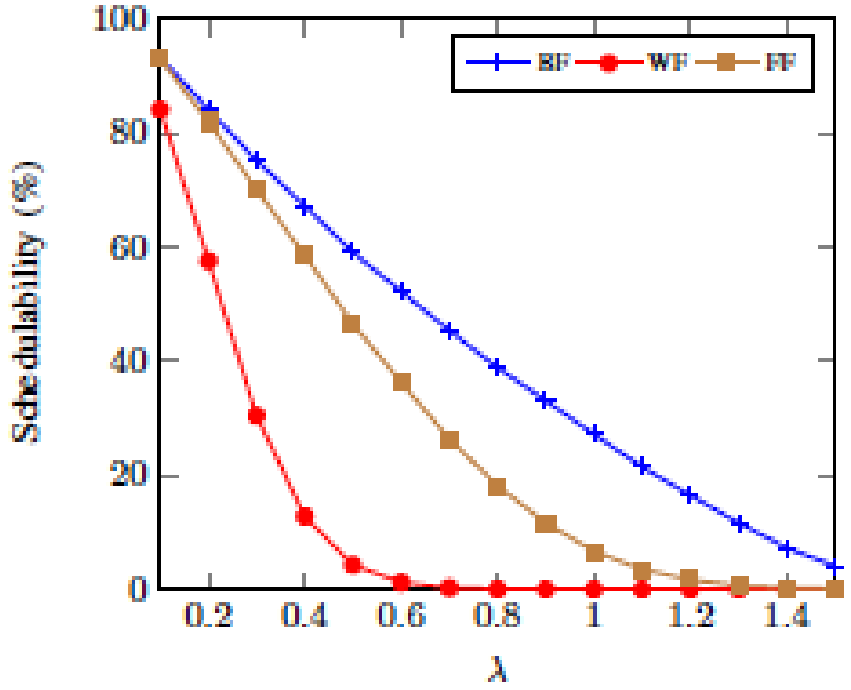
λ:
Ratio between total size of items and total capacities of partitions.



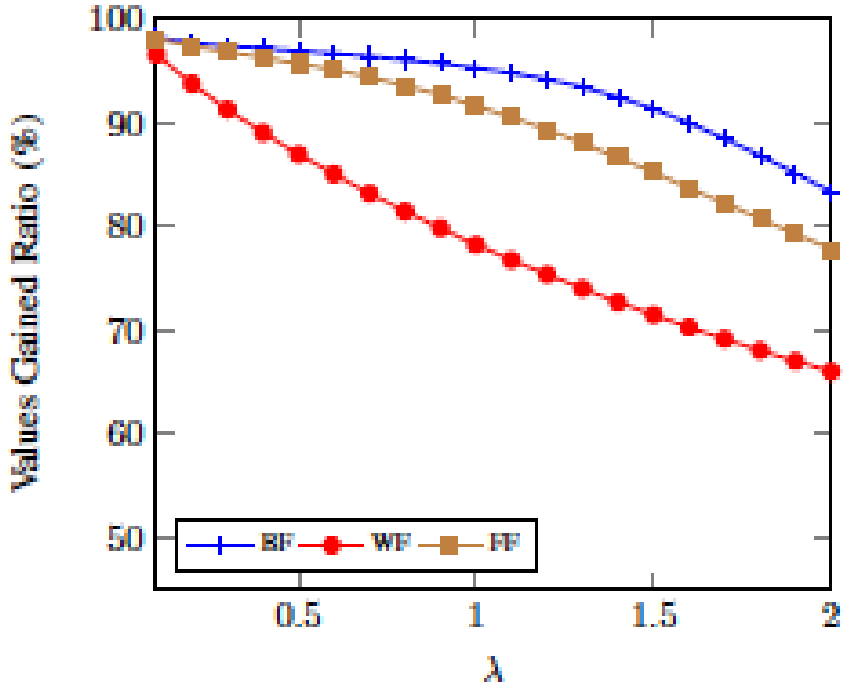
(b) Values Gained Ratio in Unit Case

Performance in RPM

Theoretical obstacles



(a) Schedulability



(b) Values Gained Ratio in Unit Case

Performance in RCM

Q&A

Thank you!

The New Architecture model

- Based on existing Xen interface.
- Introduces a new virtualized layer between the domains and vcpus, what we call **Partitions.**
- Partitions contain VCPUs that contain all the task parameters (relative deadline, absolute deadline and period of the task).
- In Xen, each VCPU is regarded as a single task and we are going to abide by this concept.