

# Live Migration of Virtual Machines

October 2014

Soramichi Akiyama

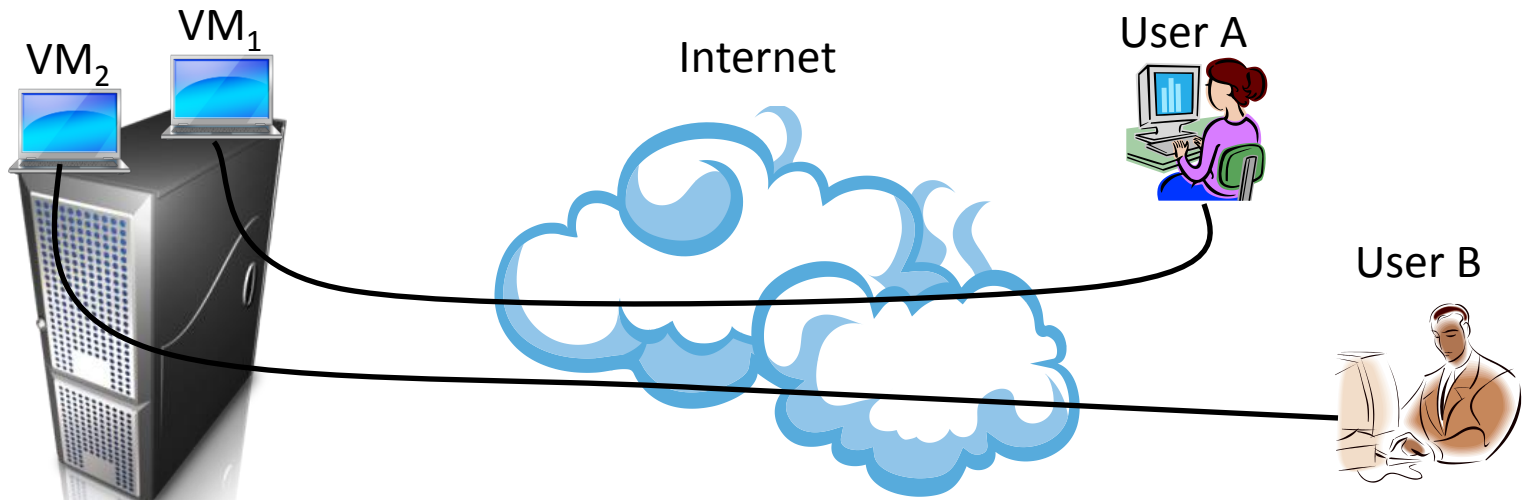
<http://www.soramichi.jp/>

# 仮想マシン(Virtual Machine, VM)

IaaSクラウドではユーザは仮想マシンを利用

例: Amazon EWSで仮想マシンを借りwebサイトを構築

リソースの有効利用	1物理マシンに複数VM
ユーザの分離	自分がVMを支配しているように見える
構成変更が容易	VMのメモリが足りないので追加等



物理マシン(Physical Machine, PM)

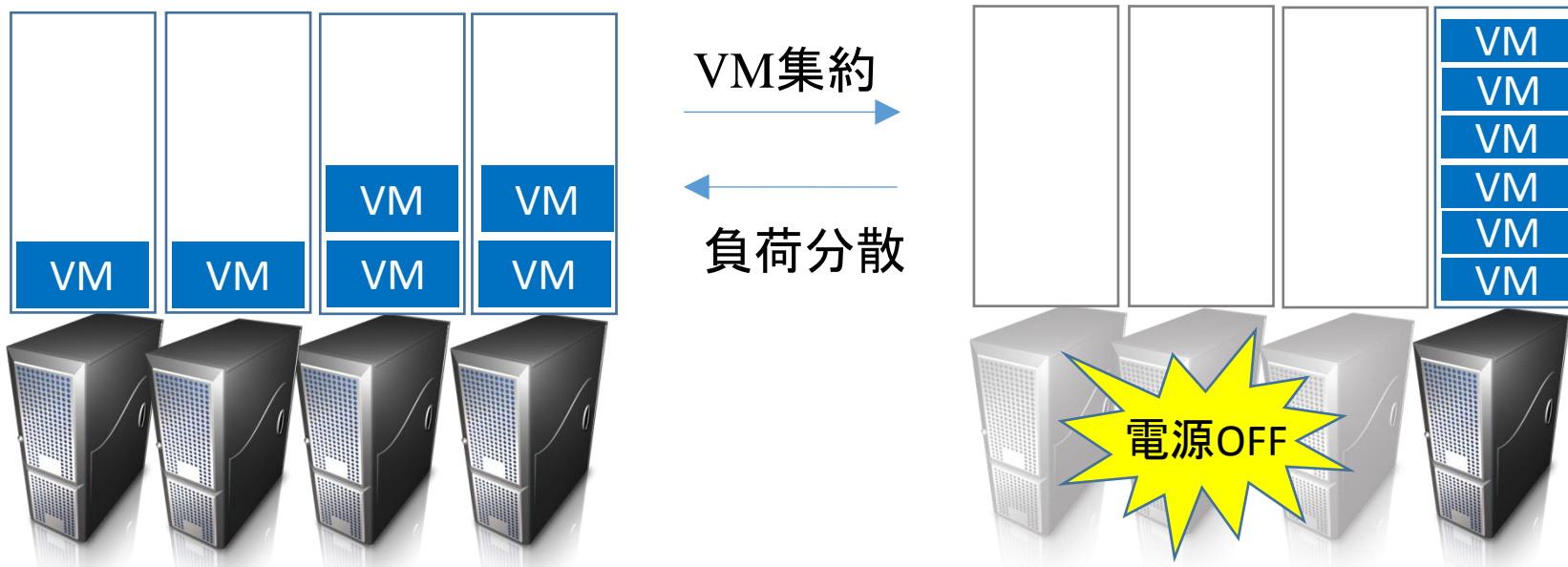
# VMの移動 (Migration)

## VMをPM間で移動

- “Virtual” Machine だからこそ可能

消費電力削減  
PMのメンテナンス  
処理効率化

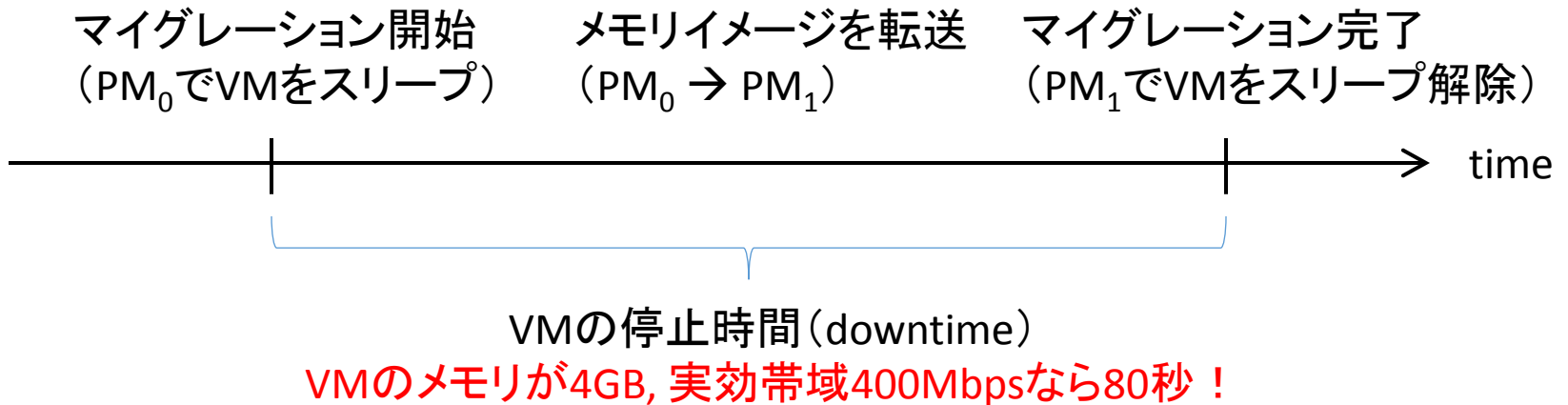
負荷の低いVMを集約  
PM<sub>0</sub>の全VMをPM<sub>1</sub>へ移動  
必要なデータの近くへVMを移動



# “Cold” Migration

## 最も原始的なMigration

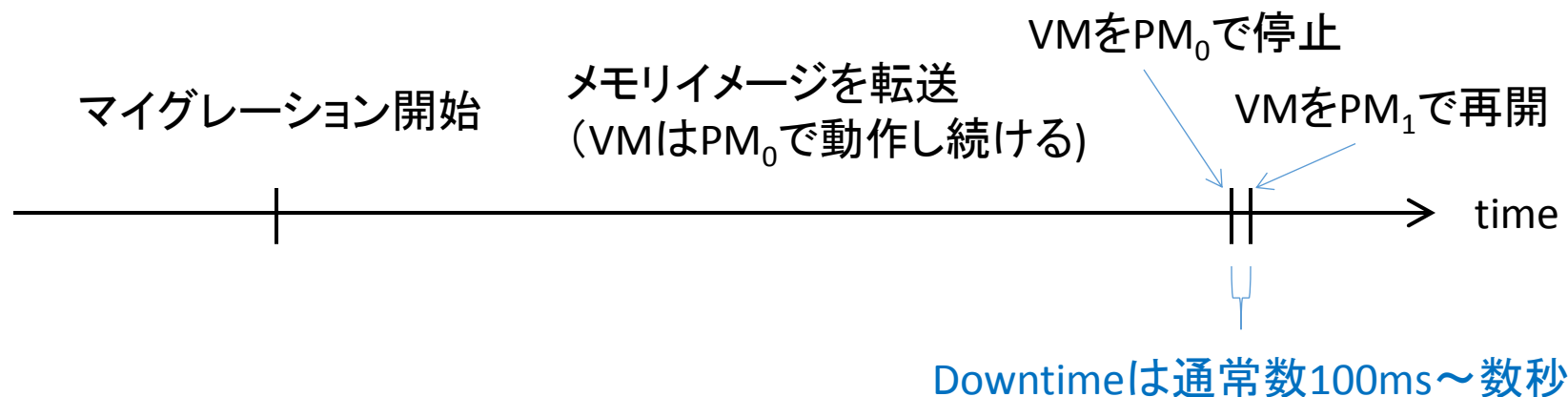
1. VMをスリープ
  2. 別のPMにメモリイメージをコピー
  3. コピー先のPMでVMをスリープから解除
- 簡単だが、**マイグレーション中にVMが停止**



# “Live” Migration

VMのユーザにはMigrationを意識させたくない

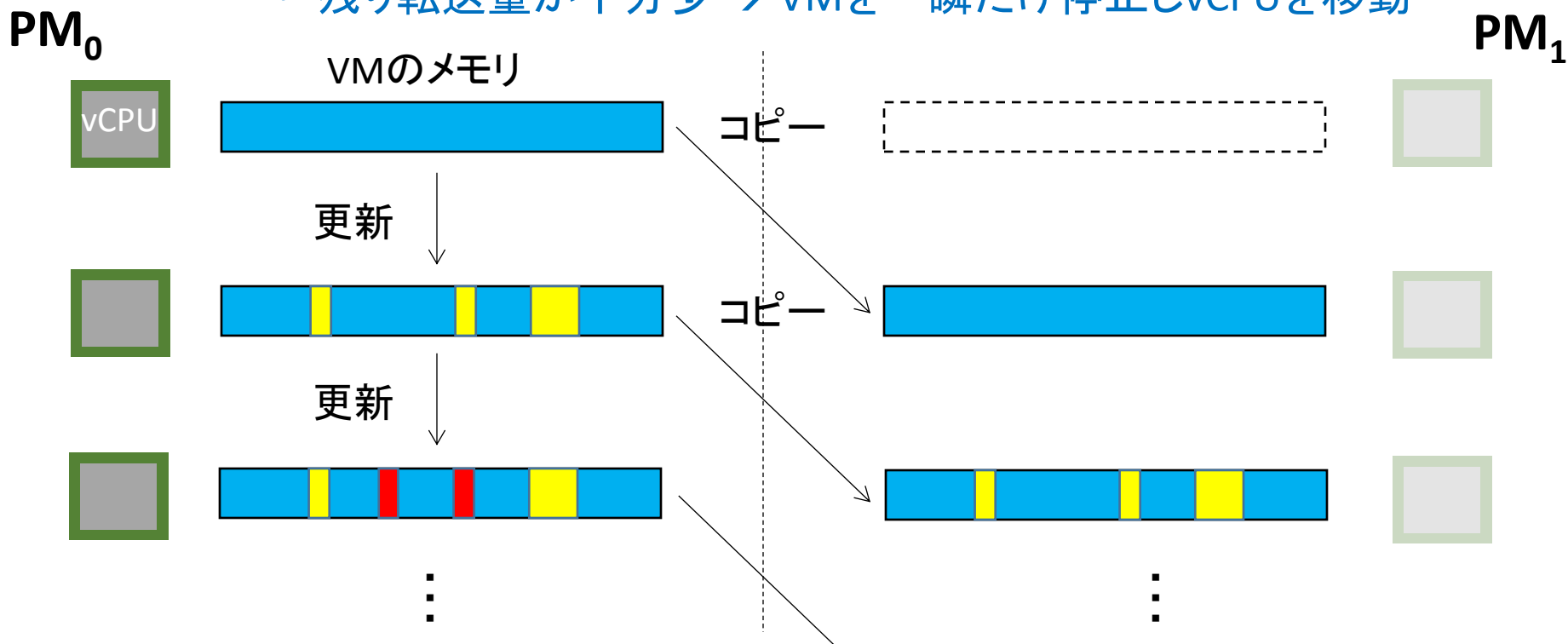
- クラウドでは 99.99... % のavailabilityが要求される
- VMを動かしたまま移動 (Live Migration)



# Pre-copy Live Migration

## 技術的課題: 動作中のVMをどうやって移動?

- **メモリがコピー中に更新される**
- **更新された部分を繰り返しコピーして対応**
  - 残り転送量が十分少 → VMを一瞬だけ停止しvCPUを移動



# 発展的研究(★は重要論文)

**課題1:** メモリ更新が速すぎ、残り転送量が”十分少なく”ならない事がある

- ★Hines *et al.*, “Post-copy based live virtual machine migration using adaptive pre-paging and dynamic self-ballooning”, *VEE'09*
- ★Liu *et al.*, “Live migration of virtual machines based on full system trace and replay”, *HPDC'09*
- ★Svärd *et al.*, “Evaluation of Delta Compression Techniques for Efficient Live Migration of Large Virtual Machines”, *VEE'11*

**課題2:** ストレージの転送はどうする？(NFS等で共有しない場合)

- Knauth *et al.*, “dsync: Efficient Block-wise Synchronization of Multi-Gigabyte Binary Data”, *LISA'13*

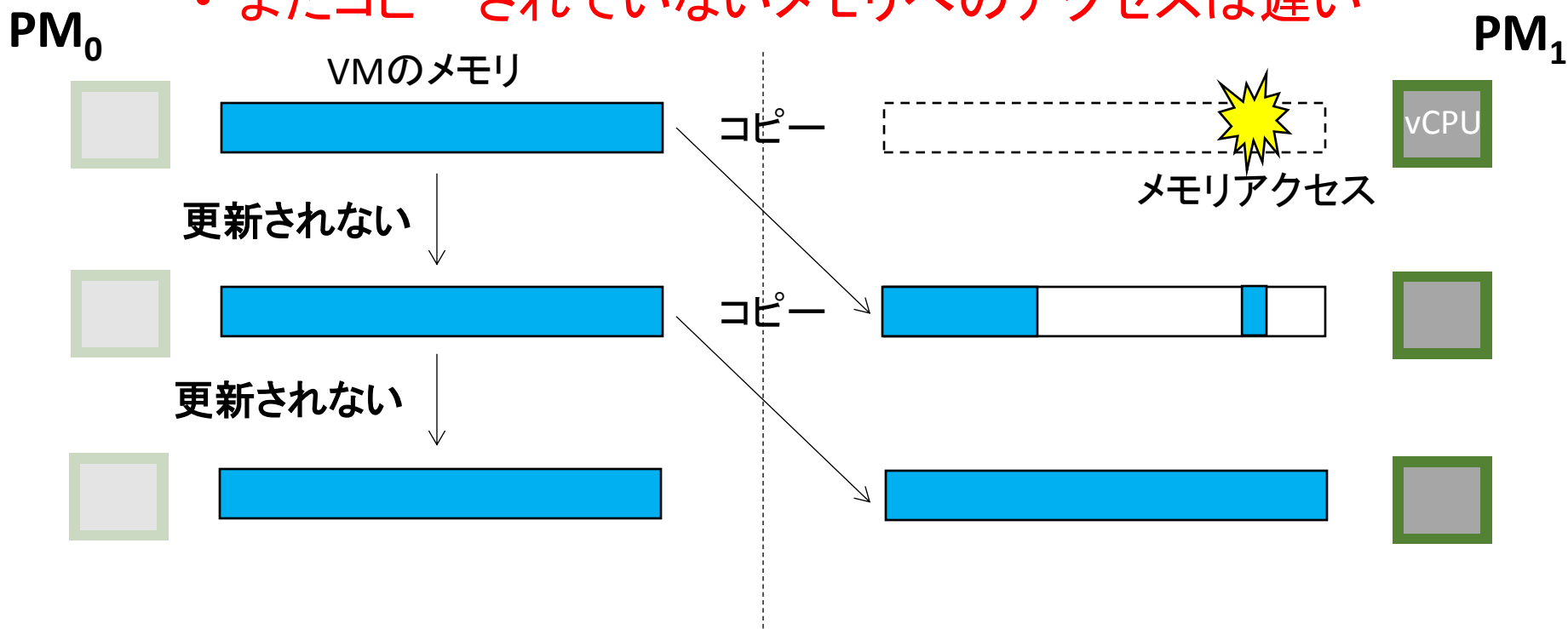
**課題3:** より高度なマイグレーション

- ★Zhai *et al.*, “Live migration with pass-through device for Linux VM”, *Linux Symposium 2008* (パススルーデバイスを持つVMをマイグレーション)
- Tullsen *et al.*, “Execution Migration in a heterogeneous-ISA chip multiprocessor”, *ASPLOS'13* (x86からARMへprocess migrationする。VMへの応用が期待される。)

# Post-copy based Live Migration (Hines *et al.*)

- 課題1: メモリの更新が速すぎ残り転送量が十分小さくならない

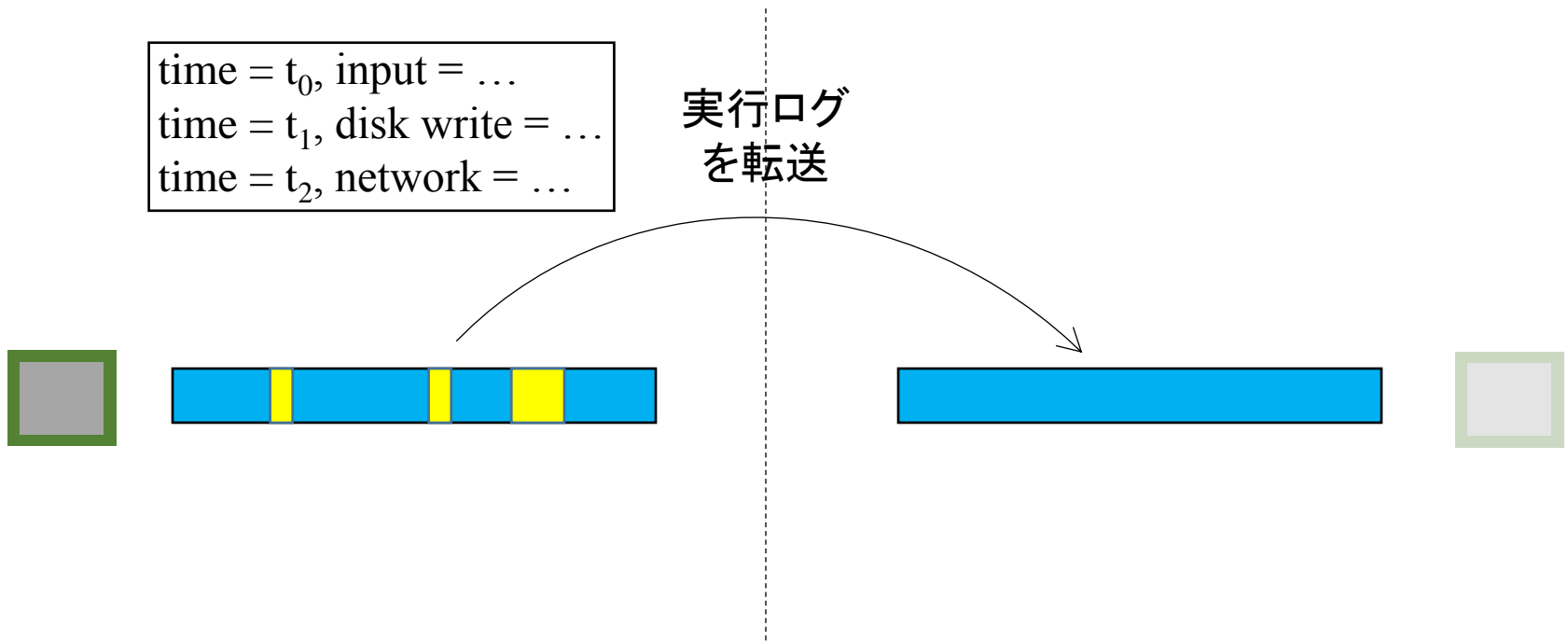
- vCPUを先に移動し、メモリを後から(post-)コピー
- まだコピーされていないメモリへのアクセスは遅い





# Live Migration based on System Trace & Replay (Liu *et al.*)

- 課題1: メモリの更新が速すぎ残り転送量が十分小さくない
  - 更新されたメモリ自体ではなく、実行ログを転送
  - 一般に実行ログはそのログが生成するメモリより小さい



# Delta Compression (Svärd *et al.*)

- 課題1: メモリの更新が速すぎ残り転送量が十分小さくない

- 1メモリページ(4096bytes)のうち、実際に更新されるのは多くの場合数% → 差分のみを転送
- 通常のmigrationでは更新されたページは丸ごと転送

