Cooperative Checkpointing Theory

Adam J. Oliner, Stanford University
Larry Rudolph, MIT, CSAIL
Ramendra K. Sahoo, IBM, T.J. Watson

April 28th

IPDPS 2006, Rhodes, Greece

The Goal

- We want checkpointing that is practical, efficient, and robust to a wide variety of failure distributions

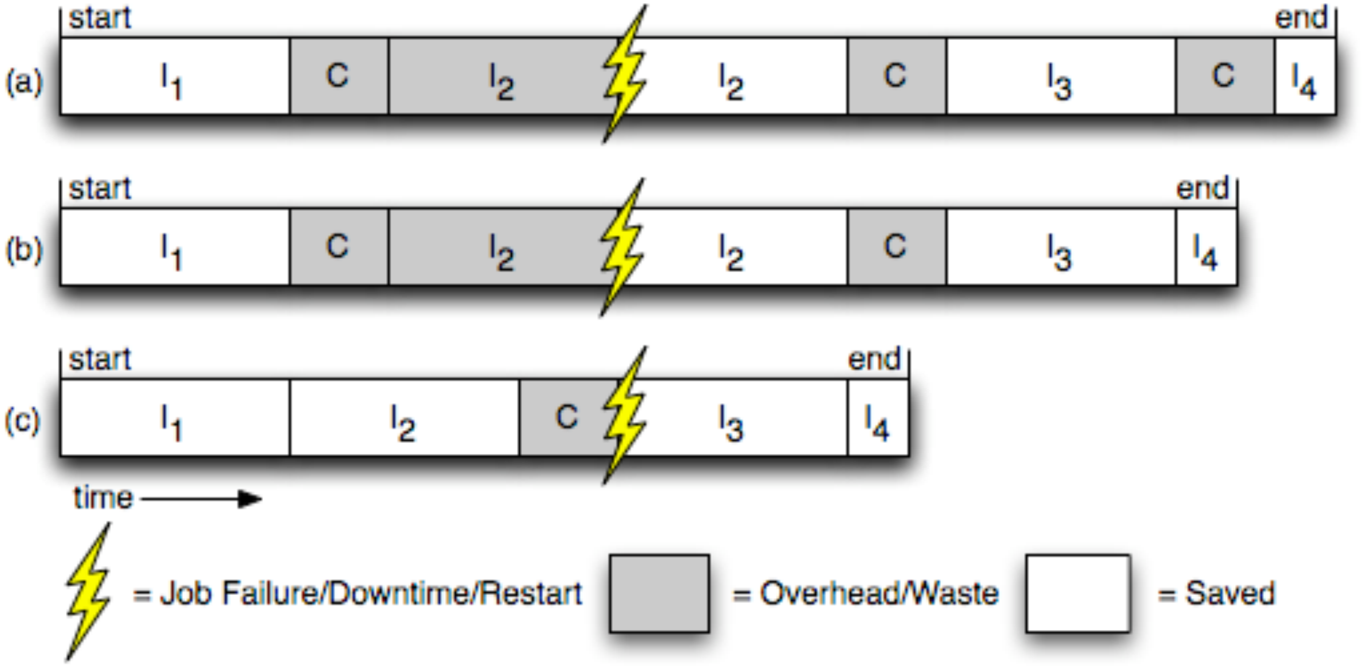
State of the Art

- Periodic Checkpointing
  - When and What
  - Optimal periodic checkpoint interval (I_λ)
- Large systems ➔ application-initiated
  - Smaller checkpoint overhead (C)
- Assumes memoryless failures

Cooperative Checkpointing

1. User inserts checkpoint requests
2. Compiler optimizes
3. Runtime gatekeeper grants/denies
    - Online heuristics = CC algorithm

CC Example

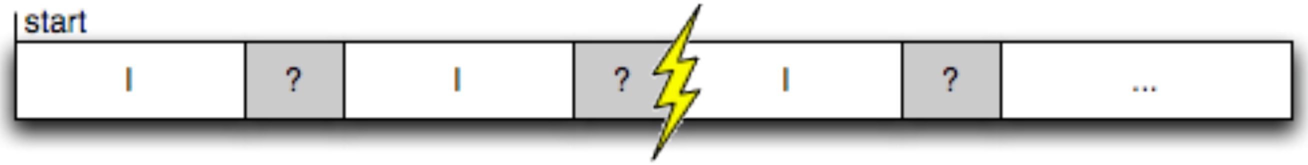


- Intuition: skip checkpoints less likely to be used for rollback

Competitive Analysis

- Given an application with periodic checkpoint *requests*
- Pick which are skipped (A)
- Adversary gets to pick failure times
  - Offline optimal (OPT) knows
- V_A = work saved by A
- Competitive ratio = V_{OPT} / V_A

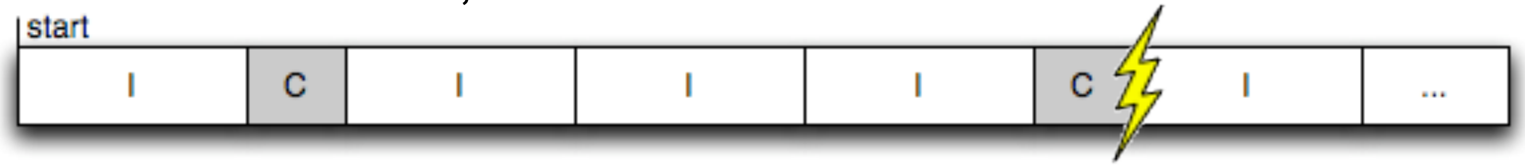
Example Analysis



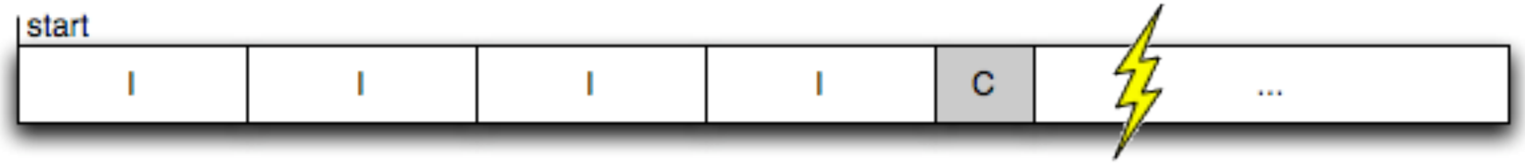
- What is an optimal strategy, given no information?
  - Worst-case analysis

Example Analysis

- We pick $A_{P,3}$



- Zeus zaps us, OPT picks



- Competitive Ratio = $V_{OPT}/V_A = 4/1 = 4$
- **Any algorithm that skips the first checkpoint is non-competitive!**

Expected Competitiveness

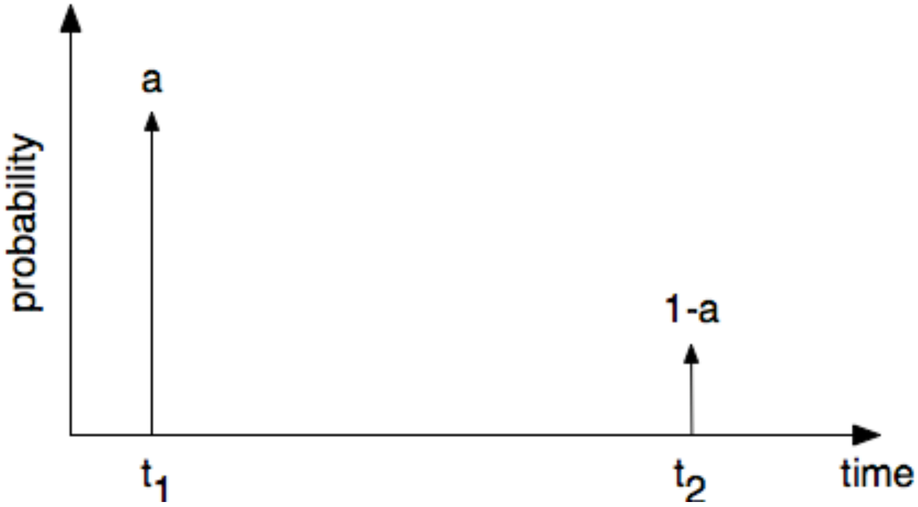
- Given failure distribution: $\mathrm{X}(t)$
- Expected competitive ratio: $E[V_{OPT}]/E[V_A]$
- Pick A to maximize $E[V_A]$

A_λ Is Not Competitive

- A_λ = optimal periodic checkpointing
  - "Sees" $X(t)$ as an exponential distribution
- Proof idea:
  - Pick $X(t)$ so A_λ does worst relative to OPT
  - Try to make $E[V_{OPT}]/E[V_A] \rightarrow \infty$

A_λ Is Not Competitive

- Make A_λ skip first checkpoint
- OPT does not
- Let $X(t) = a\delta(t-t_1)+(1-a)\delta(t-t_2)$



A_λ Is Not Competitive

- $E[F] = a(I+C)+(1-a)t_2$
- Force: $I_\lambda = \sqrt{2CE[F]} \geq 2I$
  - $a(I+C)+(1-a)t_2 \geq 2I^2/C$
- As a➔1:
  - A_λ saves no work arbitrarily often
  - OPT always saves work
  - Competitive ratio ➔ ∞
- A_λ is not competitive. ■

Olympus to Blue Gene/L

- What might X(t) be in practice?
- Blue Gene/L prototype [Sahoo, DSN05]
  - 4,096 nodes
  - Presume linear scaling
- E[F] = 1,459 seconds
- A_λ checkpoints every 17 minutes
  - 26% of machine time spent checkpointing!

Olympus to Blue Gene/L

- But wait…
  - E[F] = 1,459 seconds
  - Max[F] = 504,000 seconds
- Sound familiar?
- **Competitive Ratio = 3.95**

Contributions

- Introduces cooperative checkpointing
- Proposes a model and metrics
- Describes competitive analysis
- Proves A_λ is non-competitive
- Argues practical impact