A different re-execution speed can help

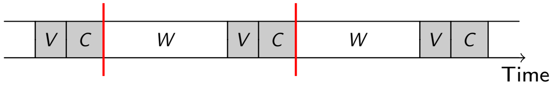
Anne Benoit, Aurélien Cavelan, Valentin Le Fèvre,
Yves Robert, Hongyang Sun

LIP, ENS de Lyon, France

PASA Workshop, in conjunction with ICPP'16
August 16, 2016

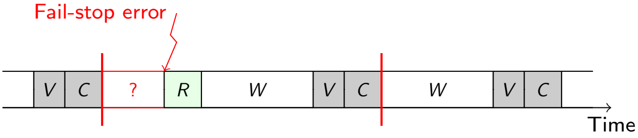
Motivation: Resilience

- Large-scale platforms: increasingly subject to errors
- Major challenge for Exascale: frequent striking of silent errors
- How to deal with these errors? Verification + Checkpoint/Restart
- Verification mechanism: general-purpose (replication, triplication) or application-specific
- *Verified checkpoints*: a verification is performed just before each checkpoint

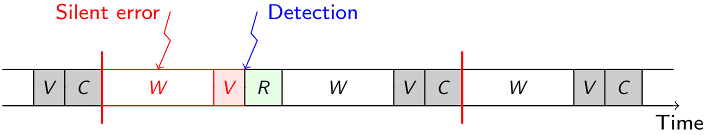


Silent vs Fail-stop errors

- C: time to checkpoint; λ: error rate (platform MTBF $\mu = 1/\lambda$); V: time to verify; R: time to recover
- Optimal checkpointing period W for fail-stop errors (Young/Daly): $W = \sqrt{2C/\lambda}$ $(V = 0)$



- Silent errors: $W = \sqrt{(V + C)/\lambda}$ ($C \to V + C$; missing factor 2)



Motivation: Energy consumption

- Power requirement of current petascale platforms = small town
- Need to reduce energy consumption of future platforms
- Popular technique: dynamic voltage and frequency scaling (DVFS)
- Lower speed $\rightarrow$ energy savings: when computing at speed σ, power proportional to σ^3 and execution time proportional to $1/\sigma$
  $\rightarrow$ (dynamic) energy proportional to σ^2
- Also account for static energy: trade-offs to be found
- Realistic approach: minimize energy while guaranteeing a performance bound

Motivation: Energy consumption

- Power requirement of current petascale platforms = small town
- Need to reduce energy consumption of future platforms
- Popular technique: dynamic voltage and frequency scaling (DVFS)
- Lower speed $\rightarrow$ energy savings: when computing at speed σ, power proportional to σ^3 and execution time proportional to $1/\sigma$
  $\rightarrow$ (dynamic) energy proportional to σ^2
- Also account for static energy: trade-offs to be found
- Realistic approach: minimize energy while guaranteeing a performance bound
- $\Rightarrow$ At which speed should we execute the workload?

Outline of the talk

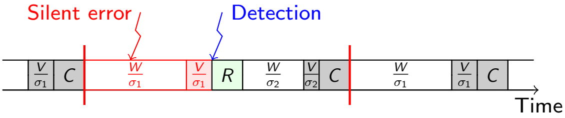
- Model and optimization problem
- Optimal pattern size and speeds
- Simulations
- Extensions: both fail-stop and silent errors
- Conclusion

Framework

- Divisible-load applications
- Subject to silent data corruption
- Checkpoint/restart strategy: periodic patterns that repeat over time
- Verified checkpoints
- Is it better to use two different speeds rather than only one? What are the optimal checkpointing period and optimal execution speeds?

Model

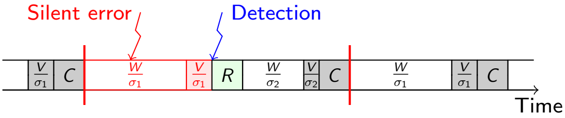
- Set of speeds $S = \{s_1, \ldots, s_K\}$:
  $\sigma_1 \in S$ speed for first execution, $\sigma_2 \in S$ speed for re-executions



With a silent error

Model

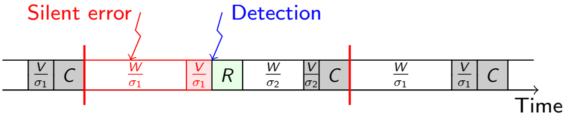
- Set of speeds $S = \{s_1, \ldots, s_K\}$:
  $\sigma_1 \in S$ speed for first execution, $\sigma_2 \in S$ speed for re-executions
- Silent errors: exponential distribution of rate λ



With a silent error

Model

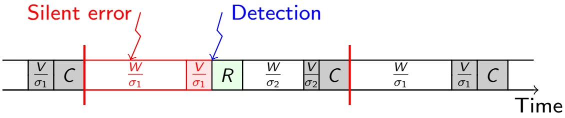
- Set of speeds $S = \{s_1, \ldots, s_K\}$:
  $\sigma_1 \in S$ speed for first execution, $\sigma_2 \in S$ speed for re-executions
- Silent errors: exponential distribution of rate λ
- Verif.: V units of work; checkpointing: time C; recovery: time R



With a silent error

Model

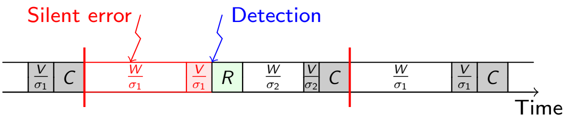
- Set of speeds $S = \{s_1, \ldots, s_K\}$:
  $\sigma_1 \in S$ speed for first execution, $\sigma_2 \in S$ speed for re-executions
- Silent errors: exponential distribution of rate λ
- Verif.: V units of work; checkpointing: time C; recovery: time R
- P_{idle} and P_{io} constant; and $P_{\text{cpu}}(\sigma) = \kappa\sigma^3$



With a silent error

Model

- Set of speeds $S = \{s_1, \ldots, s_K\}$:
  $\sigma_1 \in S$ speed for first execution, $\sigma_2 \in S$ speed for re-executions
- Silent errors: exponential distribution of rate λ
- Verif.: V units of work; checkpointing: time C; recovery: time R
- P_{idle} and P_{io} constant; and $P_{\text{cpu}}(\sigma) = \kappa\sigma^3$
- Energy for W units of work at speed σ: $\frac{W}{\sigma}(P_{\text{idle}} + \kappa\sigma^3)$
  Energy of a verification at speed σ: $\frac{V}{\sigma}(P_{\text{idle}} + \kappa\sigma^3)$
  Energy of a checkpoint: $C(P_{\text{idle}} + P_{\text{io}})$
  Energy of a recovery: $R(P_{\text{idle}} + P_{\text{io}})$



With a silent error

Problem

Optimization problem BiCrit:

$$\text{Minimize } \frac{\mathcal{E}(W, \sigma_1, \sigma_2)}{W} \text{ s.t. } \frac{\mathcal{T}(W, \sigma_1, \sigma_2)}{W} \leq \rho,$$

- $\mathcal{E}(W, \sigma_1, \sigma_2)$ is the expected energy consumed to execute W units of work at speed σ_1, with eventual re-executions at speed σ_2
- $\mathcal{T}(W, \sigma_1, \sigma_2)$ is the expected execution time to execute W units of work at speed σ_1, with eventual re-executions at speed σ_2
- ρ is a performance bound, or admissible degradation factor

Computing expected execution time

Proposition 1

For the BiCrit *problem with a single speed,*

$$\mathcal{T}(W,\sigma,\sigma) = C + e^{\frac{\lambda W}{\sigma}}\left(\frac{W+V}{\sigma}\right) + \left(e^{\frac{\lambda W}{\sigma}} - 1\right) R$$

Proposition 2

For the BiCrit *problem,*

$$\mathcal{T}(W,\sigma_1,\sigma_2) = C + \frac{W+V}{\sigma_1} + \left(1 - e^{-\frac{\lambda W}{\sigma_1}}\right) e^{\frac{\lambda W}{\sigma_2}} \left(R + \frac{W+V}{\sigma_2}\right)$$

Proof of Proposition 1

Proof.

The recursive equation to compute $\mathcal{T}(W, \sigma, \sigma)$ writes:

$$\mathcal{T}(W, \sigma, \sigma) = \frac{W + V}{\sigma} + p(W/\sigma)\left(R + \mathcal{T}(W, \sigma, \sigma)\right) + (1 - p(W/\sigma))C,$$

where $p(W/\sigma) = 1 - e^{-\frac{\lambda W}{\sigma}}$. The reasoning is as follows:

- We always execute W units of work followed by the verification, in time $\frac{W+V}{\sigma}$;
- With probability $p(W/\sigma)$, a silent error occurred and is detected, in which case we recover and start anew;
- Otherwise, with probability $1 - p(W/\sigma)$, we simply checkpoint after a successful execution.

Solving this equation leads to the expected execution time. □

Proof of Proposition 2

Proof.

The recursive equation to compute $\mathcal{T}(W, \sigma_1, \sigma_2)$ writes:

$$\mathcal{T}(W, \sigma_1, \sigma_2) = \frac{W + V}{\sigma_1} + p(W/\sigma_1)\left(R + \mathcal{T}(W, \sigma_2, \sigma_2)\right) + \left(1 - p(W/\sigma_1)\right)C,$$

where $p(W/\sigma_1) = 1 - e^{-\frac{\lambda W}{\sigma_1}}$. The reasoning is as follows:

- We always execute W units of work followed by the verification, in time $\frac{W+V}{\sigma_1}$;
- With probability $p(W/\sigma_1)$, a silent error occurred and is detected, in which case we recover and start anew at speed σ_2;
- Otherwise, with probability $1 - p(W/\sigma_1)$, we simply checkpoint after a successful execution.

Solving this equation leads to the expected execution time. □

Computing expected energy consumption

Proposition 3

For the BICRIT *problem,*

$$\mathcal{E}(W,\sigma_1,\sigma_2) = \left(C + \left(1 - e^{-\frac{\lambda W}{\sigma_1}}\right) e^{\frac{\lambda W}{\sigma_2}} R\right)(P_{io} + P_{idle}) + \frac{W+V}{\sigma_1}(\kappa\sigma_1^3 + P_{idle}) + \frac{W+V}{\sigma_2}(1 - e^{-\frac{\lambda W}{\sigma_1}})e^{\frac{\lambda W}{\sigma_2}}(\kappa\sigma_2^3 + P_{idle})$$

Power spent during checkpoint or recovery: $P_{io} + P_{idle}$; power spent during computation and verification at speed σ: $P_{cpu}(\sigma) + P_{idle} = \kappa\sigma^3 + P_{idle}$. From Proposition 2, we get the expression of $\mathcal{E}(W,\sigma_1,\sigma_2)$.

Finding optimal pattern length (1)

To get closed-form expression for optimal value of W, use of first-order approximations, using Taylor expansion $e^{\lambda W} = 1 + \lambda W + O(\lambda^2 W^2)$:

$$\frac{\mathcal{T}(W,\sigma_1,\sigma_2)}{W} = \frac{1}{\sigma_1} + \frac{\lambda W}{\sigma_1\sigma_2} + \frac{\lambda R}{\sigma_1} + \frac{\lambda V}{\sigma_1\sigma_2} + \frac{C + V/\sigma_1}{W} + O(\lambda^2 W) \quad (1)$$

$$\begin{aligned}\frac{\mathcal{E}(W,\sigma_1,\sigma_2)}{W} &= \frac{\kappa\sigma_1^3 + P_{\text{idle}}}{\sigma_1} + \frac{\lambda W}{\sigma_1\sigma_2}(\kappa\sigma_2^3 + P_{\text{idle}}) \\ &+ \frac{\lambda R}{\sigma_1}(P_{\text{io}} + P_{\text{idle}}) + \frac{\lambda V}{\sigma_1\sigma_2}(\kappa\sigma_1^3 + P_{\text{idle}}) \\ &+ \frac{C(P_{\text{io}} + P_{\text{idle}}) + V(\kappa\sigma_1^3 + P_{\text{idle}})/\sigma_1}{W} + O(\lambda^2 W) \quad (2)\end{aligned}$$

Finding optimal pattern length (2)

Theorem 1

Given σ_1, σ_2 and ρ, consider the equation $aW^2 + bW + c = 0$, where $a = \frac{\lambda}{\sigma_1\sigma_2}$, $b = \frac{1}{\sigma_1} + \lambda\left(\frac{R}{\sigma_1} + \frac{V}{\sigma_1\sigma_2}\right) - \rho$ and $c = C + \frac{V}{\sigma_1}$.

- *If there is no positive solution to the equation, i.e., $b > -2\sqrt{ac}$, then* BiCrit *has no solution.*
- *Otherwise, let W_1 and W_2 be the two solutions of the equation with $W_1 \leq W_2$ (at least W_2 is positive and possibly $W_1 = W_2$). Then, the optimal pattern size is*

$$W_{\text{opt}} = \min(\max(W_1, W_e), W_2), \tag{3}$$

$$\text{where } W_e = \sqrt{\frac{C(P_{\text{io}} + P_{\text{idle}}) + \frac{V}{\sigma_1}(\kappa\sigma_1^3 + P_{\text{idle}})}{\frac{\lambda}{\sigma_1\sigma_2}(\kappa\sigma_2^3 + P_{\text{idle}})}}. \tag{4}$$

Finding optimal pattern length (3)

Proof.

Neglecting lower-order terms, Equation (2) is minimized when $W = W_e$ given by Equation (4).

Two cases:

- ρ is too small $\Rightarrow$ no solution
- $W_2 > 0$:
  - $W_e < W_1$
  - $W_1 \leq W_e \leq W_2$
  - $W_e > W_2$

Using that the energy overhead is a convex function, we get the result (W_{opt} is in the interval $[W_1, W_2]$) □

Finding optimal speed pair

- Speed pair (s_i, s_j), with $1 \leq i, j \leq K$: $\rho_{i,j}$ is the minimum performance bound for which the BiCrit problem with $\sigma_1 = s_i$ and $\sigma_2 = s_j$ admits a solution
- For each speed pair, compute W_1, W_2 the roots of $aW^2 + bW + c$; discard pairs with $\rho < \rho_{i,j}$
- For each remaining speed pair (σ_1, σ_2), compute W_{opt} and associated energy overhead
- Select speed pair (σ_1^*, σ_2^*) that minimizes energy overhead
- Time $O(K^2)$, where K is the number of available speeds, usually a small constant

Simulation setup

- Platform parameters, based on real platforms

Platform	λ	$C = R$	V
Hera	3.38e-6	$300s$	15.4
Atlas	7.78e-6	$439s$	9.1
Coastal	2.01e-6	$1051s$	4.5
Coastal SSD	2.01e-6	$2500s$	180.0

- Power parameters, determined by the processor used

Processor	Normalized speeds	$P(\sigma)$ (mW)
Intel Xscale	$0.15, 0.4, 0.6, 0.8, 1$	$1550\sigma^3 + 60$
Transmeta Crusoe	$0.45, 0.6, 0.8, 0.9, 1$	$5756\sigma^3 + 4.4$

- Default values: P_{io} equivalent to power used when running at lowest speed; $\rho = 3$

Simulation results, using Hera/XScale configuration

A different re-execution speed does help!
And all speed pairs can be optimal solutions (depending on ρ)!

σ_1	**Best** σ_2	W_{opt}	$\frac{\mathcal{E}(W_{opt},\sigma_1,\sigma_2)}{W_{opt}}$
0.15	0.4	1711	466
0.4	**0.4**	2764	416
0.6	0.4	3639	674
0.8	0.4	4627	1082
1	0.4	5742	1625

$\rho = 8$

σ_1	**Best** σ_2	W_{opt}	$\frac{\mathcal{E}(W_{opt},\sigma_1,\sigma_2)}{W_{opt}}$
0.15	-	-	-
0.4	**0.4**	2764	416
0.6	0.4	3639	674
0.8	0.4	4627	1082
1	0.4	5742	1625

$\rho = 3$

σ_1	**Best** σ_2	W_{opt}	$\frac{\mathcal{E}(W_{opt},\sigma_1,\sigma_2)}{W_{opt}}$
0.15	-	-	-
0.4	-	-	-
0.6	**0.8**	4251	690
0.8	0.4	4627	1082
1	0.4	5742	1625

$\rho = 1.775$

σ_1	**Best** σ_2	W_{opt}	$\frac{\mathcal{E}(W_{opt},\sigma_1,\sigma_2)}{W_{opt}}$
0.15	-	-	-
0.4	-	-	-
0.6	-	-	-
0.8	**0.4**	4627	1082
1	0.4	5742	1625

$\rho = 1.4$

Simulations - Impact of the parameters (1)

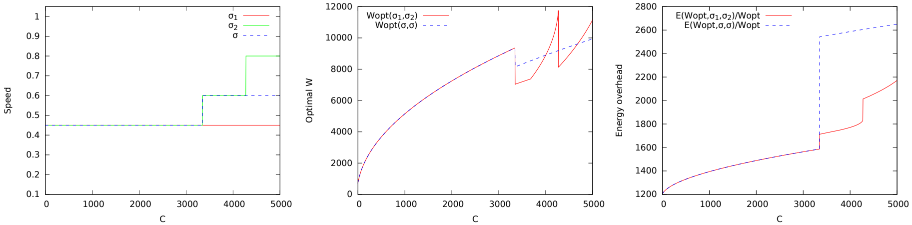


Figure: The optimal solution (speed pair, pattern size, and energy overhead) as a function of the checkpointing time c in Atlas/Crusoe configuration.

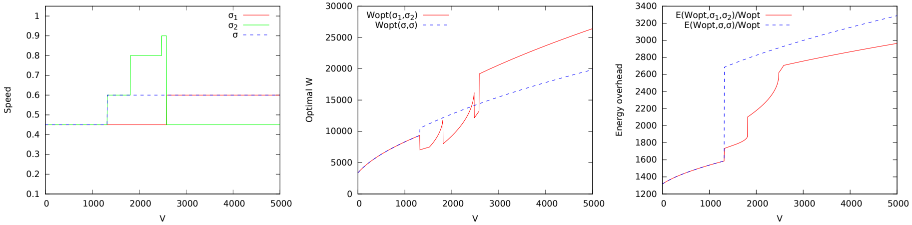


Figure: The optimal solution (speed pair, pattern size, and energy overhead) as a function of the verification time v in Atlas/Crusoe configuration.

Dotted line: one single speed; up to 35% improvement with two speeds

Simulations - Impact of the parameters (2)

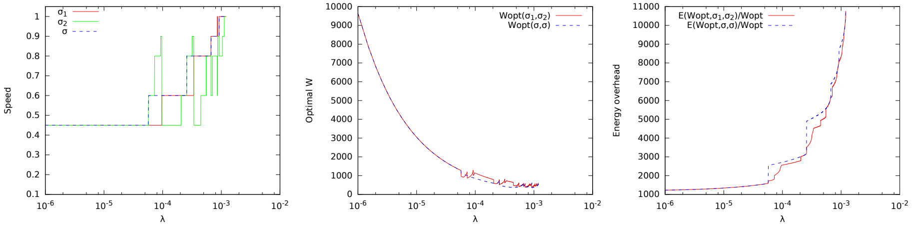


Figure: The optimal solution (speed pair, pattern size, and energy overhead) as a function of the error rate λ in Atlas/Crusoe configuration.

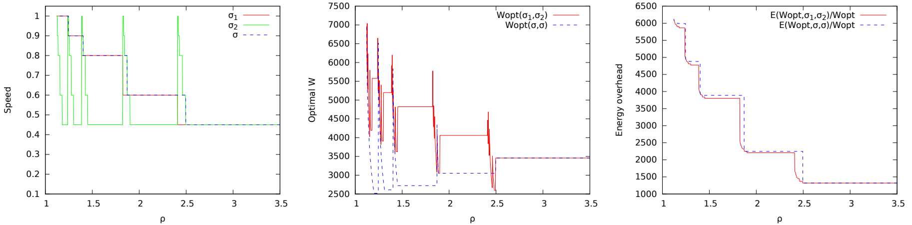


Figure: The optimal solution (speed pair, pattern size, and energy overhead) as a function of the performance bound ρ in Atlas/Crusoe configuration.

Two speeds: checkpoint less frequently and provide energy savings

Simulations - Impact of the parameters (3)

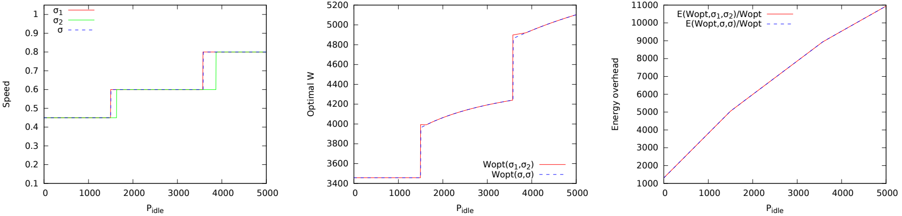


Figure: The optimal solution (speed pair, pattern size, and energy overhead) as a function of the idle power P_{idle} in Atlas/Crusoe configuration.

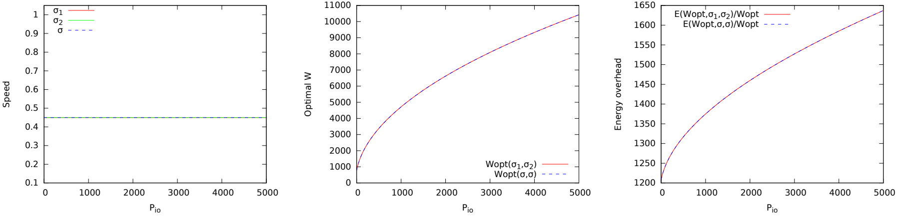


Figure: The optimal solution (speed pair, pattern size, and energy overhead) as a function of the I/O power P_{io} in Atlas/Crusoe configuration.

Increase of W and E with P_{idle} and Pio; P_{io} has no impact on speeds

Extensions: With fail-stop errors

- f: proportion of fail-stop errors
- s: proportion of silent errors

Proposition 4

With fail-stop and silent errors,

$$\frac{\mathcal{T}(W,\sigma_1,\sigma_2)}{W} = \cdots + \left(\frac{(f+s)}{\sigma_1\sigma_2} - \frac{f}{2\sigma_1^2}\right)\lambda W + O(\lambda^2 W). \quad (5)$$

$$\frac{\mathcal{E}(W,\sigma_1,\sigma_2)}{W} = \cdots + \left(\frac{(f+s)(\kappa\sigma_2^3 + P_{\text{idle}})}{\sigma_1\sigma_2} - \frac{f(\kappa\sigma_1^3 + P_{\text{idle}})}{2\sigma_1^2}\right)\lambda W + O(\lambda^2 W) \quad (6)$$

Limit of the first-order approximation

For BiCrit, the first-order approximation leads to a solution iff

$$\left(2\left(1+\frac{s}{f}\right)\right)^{-1/2} < \frac{\sigma_2}{\sigma_1} < 2\left(1+\frac{s}{f}\right)$$

Use second-order approximation? Open problem in the general case!

Interesting case

Theorem 2

When considering only fail-stop errors with rate λ, the optimal pattern size W to minimize the time overhead $\frac{\mathcal{T}(W,\sigma,2\sigma)}{W}$ is

$$W_{\text{opt}} = \sqrt[3]{\frac{12C}{\lambda^2}}\sigma$$

- Young/Daly's formula: $W_{\text{opt}} = \sqrt{2C/\lambda}\sigma = O(\lambda^{-1/2})$
- Here: $W_{\text{opt}} = O(\lambda^{-2/3})$

Conclusion

- A different re-execution speed indeed helps saving energy while satisfying a performance constraint
- Silent errors: extension of Young/Daly formula $\rightarrow$ general closed-form solution to get optimal speed pair and optimal checkpointing period (first-order)
- Extensive simulations: up to 35% energy savings, any speed pair can be optimal

- BiCrit still open for general case with both silent and fail-stop errors
- Interesting case with fail-stop errors and double re-execution speed: $O(\lambda^{-2/3})$ vs $O(\lambda^{-1/2})$
- New methods needed to capture the general case