Towards Optimal Multi-Level Checkpointing

Anne Benoit[1], Aurélien Cavelan[1],
Valentin Le Fèvre[1], Yves Robert[1,2], Hongyang Sun[1]

1. ENS Lyon & INRIA, France
2. University of Tennessee Knoxville, USA

hongyang.sun@ens-lyon.fr

JLESC Workshop, Lyon
June 28, 2016

Single-Level Checkpointing

Minimize expected execution overhead $H(W) = \frac{\mathbb{E}(W)}{W} - 1$

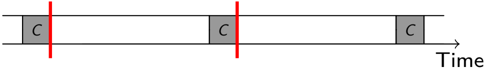


Figure: periodic computing pattern

What is the optimal checkpointing interval?

Single-Level Checkpointing

Minimize expected execution overhead $H(W) = \frac{\mathbb{E}(W)}{W} - 1$

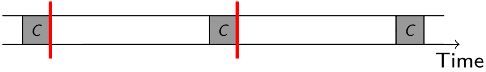


Figure: periodic computing pattern

What is the optimal checkpointing interval?

- Exact solution:

$$H(W) = \frac{e^{\lambda R}(\frac{1}{\lambda} + D)e^{\lambda(W+C)}}{W} - 1, \text{use Lambert function}$$

- First-order approximation [Young/Daly]:

$$W_{\text{opt}} = \sqrt{\frac{2C}{\lambda}}$$

$$H_{\text{opt}} = \sqrt{2\lambda C} + \Theta(\lambda)$$

Single-Level Checkpointing

Minimize expected execution overhead $H(W) = \frac{\mathbb{E}(W)}{W} - 1$

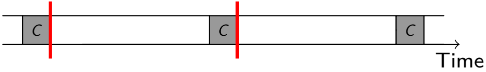


Figure: periodic computing pattern

What is the optimal checkpointing interval?

- Exact solution:

$$H(W) = \frac{e^{\lambda R}(\frac{1}{\lambda} + D)e^{\lambda(W+C)}}{W} - 1, \text{use Lambert function}$$

- First-order approximation [Young/Daly]:

$$W_{\text{opt}} = \sqrt{\frac{2C}{\lambda}}$$
$$H_{\text{opt}} = \sqrt{2\lambda C} + \Theta(\lambda)$$

Scalability problem for large-scale platforms

Multi-Level Checkpointing

E.g., Scalable Checkpoint/Restart (SCR) library, Fault Tolerance Interface (FTI)

- Local memory/SSD, Partner copy/XOR, RS-coding, PFS

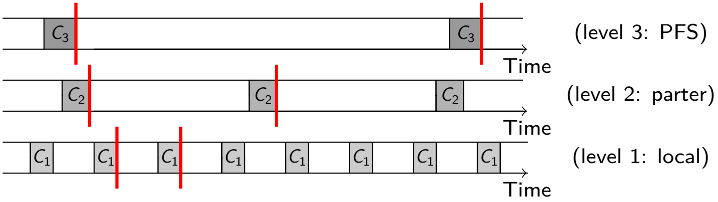
Multi-Level Checkpointing

E.g., Scalable Checkpoint/Restart (SCR) library, Fault Tolerance Interface (FTI)

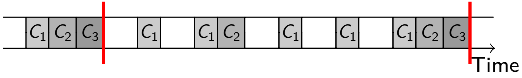
- Local memory/SSD, Partner copy/XOR, RS-coding, PFS

Two approaches:

- Independent checkpointing:



- Synchronized checkpointing:



Two Levels

Easy because pattern repeats (memoryless property)

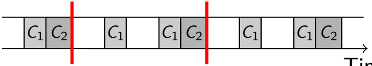


Figure: with N level-1 checkpoints

- Exact solution: very complicated (which error type occurs first?), equal-length chunks, see [1]

[1] S. Di, Y. Robert, F. Vivien, F. Cappello. Toward an optimal online checkpoint solution under a two-level HPC checkpoint model, *IEEE TPDS*, 2016.

Two Levels

Easy because pattern repeats (memoryless property)

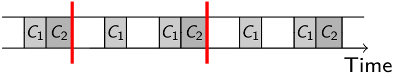


Figure: with N level-1 checkpoints

- Exact solution: very complicated (which error type occurs first?), equal-length chunks, see [1]
- First-order approximation:

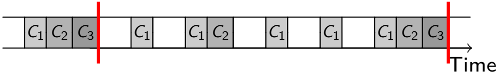
$$H_{\text{opt}} = \sqrt{2\lambda_1 C_1} + \sqrt{2\lambda_2 C_2} + \Theta(\lambda)$$

(obtained for some optimal pattern)

[1] S. Di, Y. Robert, F. Vivien, F. Cappello. Toward an optimal online checkpoint solution under a two-level HPC checkpoint model, *IEEE TPDS*, 2016.

Three Levels

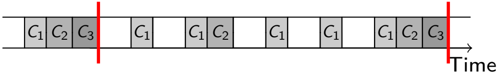
Difficult because sub-patterns may differ



- Exact solution: unknown

Three Levels

Difficult because sub-patterns may differ



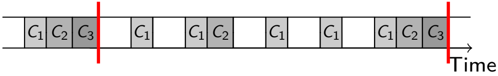
- Exact solution: unknown
- First-order approximation:

$$H_{\text{opt}} = \sqrt{2\lambda_1 C_1} + \sqrt{2\lambda_2 C_2} + \sqrt{2\lambda_3 C_3} + \Theta(\lambda)$$

(obtained for some optimal pattern)

Three Levels

Difficult because sub-patterns may differ



- Exact solution: unknown
- First-order approximation:

$$H_{\text{opt}} = \sqrt{2\lambda_1 C_1} + \sqrt{2\lambda_2 C_2} + \sqrt{2\lambda_3 C_3} + \Theta(\lambda)$$

(obtained for some optimal pattern)

- Choose optimal set of levels:

Level	Overhead
$1, 2, 3$	$\sqrt{2C_1\lambda_1} + \sqrt{2C_2\lambda_2} + \sqrt{2C_3\lambda_3}$
$1, 3$	$\sqrt{2C_1\lambda_1} + \sqrt{2C_3(\lambda_2 + \lambda_3)}$
$2, 3$	$\sqrt{2C_2(\lambda_1 + \lambda_2)} + \sqrt{2C_3\lambda_3}$
3	$\sqrt{2C_3(\lambda_1 + \lambda_2 + \lambda_3)}$

k Levels

Theorem

The optimal k-level pattern, under the first-order approximation, has equal-length chunks at all levels:

Optimal pattern length: $W^{opt} = \sqrt{\dfrac{\sum_{\ell=1}^{k} N_\ell^{opt} C_\ell}{\frac{1}{2}\sum_{\ell=1}^{k} \frac{\lambda_\ell}{N_\ell^{opt}}}}$

Optimal #chkpts at level ℓ: $N_\ell^{opt} = \sqrt{\dfrac{\lambda_\ell}{C_\ell} \cdot \dfrac{C_k}{\lambda_k}}, \quad \forall \ell = 1, \ldots, k$

Optimal pattern overhead: $H_{opt} = \sum_{\ell=1}^{k} \sqrt{2\lambda_\ell C_\ell} + \Theta(\lambda)$

k Levels

Theorem

The optimal k-level pattern, under the first-order approximation, has equal-length chunks at all levels:

Optimal pattern length: $W^{opt} = \sqrt{\dfrac{\sum_{\ell=1}^{k} N_\ell^{opt} C_\ell}{\frac{1}{2}\sum_{\ell=1}^{k} \frac{\lambda_\ell}{N_\ell^{opt}}}}$

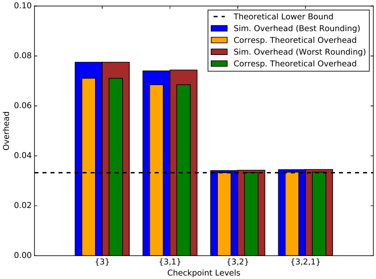
Optimal #chkpts at level ℓ: $N_\ell^{opt} = \sqrt{\dfrac{\lambda_\ell}{C_\ell} \cdot \dfrac{C_k}{\lambda_k}}, \quad \forall \ell = 1, \ldots, k$

Optimal pattern overhead: $H_{opt} = \sum_{\ell=1}^{k} \sqrt{2\lambda_\ell C_\ell} + \Theta(\lambda)$

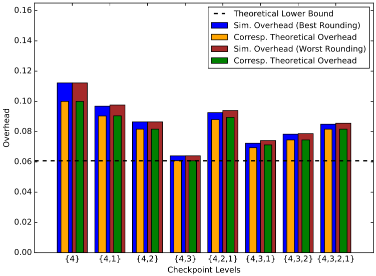
- Dynamic programming algorithm to choose set of levels
- Rounding for integer solution: $n_\ell^{\text{opt}} = \dfrac{N_\ell^{\text{opt}}}{N_{\ell+1}^{\text{opt}}} = \sqrt{\dfrac{\lambda_\ell}{\lambda_{\ell+1}} \cdot \dfrac{C_{\ell+1}}{C_\ell}}$

Simulations

Set	Source	Level	1	2	3	4
(A)	Moody et al. [1]	C (s)	0.5	4.5	1051	-
		MTBF (s)	5.00e6	5.56e5	2.50e6	-
(B)	Balaprakash et al. [2]	C (s)	10	20	20	100
		MTBF (s)	3.60e4	7.20e4	1.44e5	7.20e5



(A)



(B)

[1] A. Moody, G. Bronevetsky, K. Mohror, and B. R. de Supinski. Design, modeling, and evaluation of a scalable multi-level checkpointing system. *Supercomputing*, 2010.

[2] P. Balaprakash, L. A. Bautista-Gomez, M.-S. Bouguerra, S. M. Wild, F. Cappello, and P. D. Hovland. Analysis of the tradeoffs between energy and run time for multilevel checkpointing. *PMBS*, 2014.

Conclusion

Explicit formulas for (almost) optimal multi-level checkpointing

$$H_{\text{opt}} = \sum_{\ell=1}^{k} \sqrt{2\lambda_\ell C_\ell} + \Theta(\lambda)$$

Limitations:

- First-order approximation (accurate for 10,000s of nodes with MTBF in hours; beyond?)
- Independent errors (correlated failure?)