When Amdahl Meets Young/Daly

Aurélien Cavelan[1] Jiafan Li[3] Yves Robert[1,2]
Hongyang Sun[1]

[1]ENS Lyon & INRIA, France.

[2]University of Tennessee Knoxville, USA.

[3]East China Normal University, China.

IEEE Cluster'16@Taipei, Taiwan
September 14, 2016

What is the optimal number of processors to execute a parallel job obeying Amdahl's law on a failure-prone platform?

Amdahl's Law

Speedup with P processors and α sequential fraction:

$$S(P) = \frac{1}{\alpha + \frac{1-\alpha}{P}}$$

- Bounded above by $1/\alpha$
- Strictly increasing function of P

Amdahl's Law

Speedup with P processors and α sequential fraction:

$$S(P) = \frac{1}{\alpha + \frac{1-\alpha}{P}}$$

- Bounded above by $1/\alpha$
- Strictly increasing function of P

Allocating processors on a failure-prone platform?

- Same speedup ☺
- More errors/failures ☹

$$\text{MTBF } \mu_P = \frac{\mu_{\text{ind}}}{P}$$

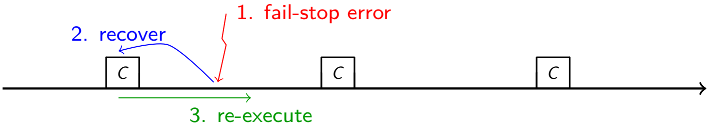
- Increased resilience overhead ☹

Resilience for HPC

Fail-stop errors: e.g., resource crash, node failure

- Instantaneous error detection

Standard approach: periodic checkpointing, rollback and recovery

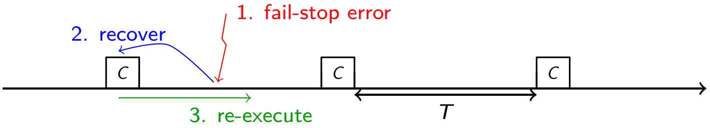


Resilience for HPC

Fail-stop errors: e.g., resource crash, node failure

- Instantaneous error detection

Standard approach: periodic checkpointing, rollback and recovery



Optimal checkpointing interval à la Young/Daly:

$$T^* = \sqrt{2\mu C}$$

where μ is MTBF and C is checkpointing time

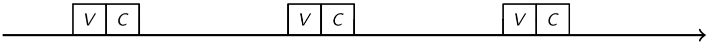
- First-order approximation formula
- With fixed processor allocation

Coping with Silent Errors

Silent errors (or Silent Data Corruptions or SDCs): e.g., soft faults in L1 cache, ALU, double bit flip, due to cosmic radiation, packaging pollution, etc.

- Arbitrary detection latency

Promising approach: combine checkpointing with verification (for error detection)

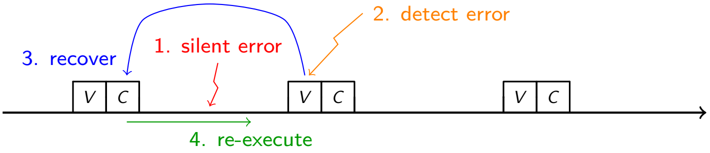


Coping with Silent Errors

Silent errors (or Silent Data Corruptions or SDCs): e.g., soft faults in L1 cache, ALU, double bit flip, due to cosmic radiation, packaging pollution, etc.

- Arbitrary detection latency

Promising approach: combine checkpointing with verification (for error detection)

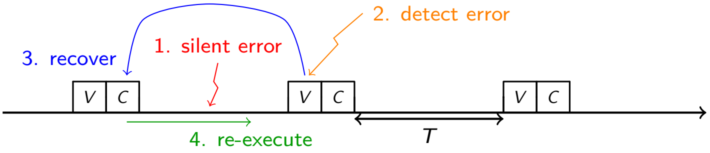


Coping with Silent Errors

Silent errors (or Silent Data Corruptions or SDCs): e.g., soft faults in L1 cache, ALU, double bit flip, due to cosmic radiation, packaging pollution, etc.

- Arbitrary detection latency

Promising approach: combine checkpointing with verification (for error detection)



- Extension of Young/Daly: $T^* = \sqrt{\mu(V + C)}$
- Many methods to detect silent errors

Methods for Detecting Silent Errors

General-purpose approaches

- Replication [*Fiala et al. 2012*] or triple modular redundancy and voting [*Lyons and Vanderkulk 1962*]

Application-specific approaches

- Algorithm-based fault tolerance (ABFT): checksums in dense matrices
  Limited to one error detection and/or correction in practice [*Huang and Abraham 1984*]
- Partial differential equations (PDE): use lower-order scheme as verification mechanism [*Benson, Schmit and Schreiber 2014*]
- Generalized minimal residual method (GMRES): inner-outer iterations [*Hoemmen and Heroux 2011*]
- Preconditioned conjugate gradients (PCG): orthogonalization check every k iterations, re-orthogonalization if problem detected [*Sao and Vuduc 2013, Chen 2013*]

Data-analytics approaches

- Dynamic monitoring of HPC datasets based on physical laws (e.g., temperature limit, speed limit) and space or temporal proximity [*Bautista-Gomez and Cappello 2014*]
- Time-series prediction, spatial multivariate interpolation [*Di et al. 2014*]

When Amdahl Meets Young/Daly

Optimizing performance (overhead $H = 1/S$):

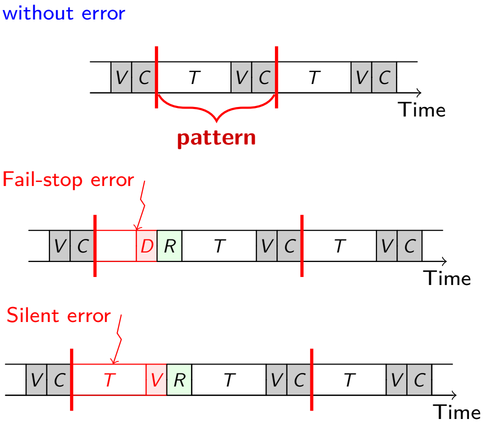
- Optimal number of processors P^*
- Optimal checkpointing interval T^*

When Amdahl Meets Young/Daly

Optimizing performance (overhead $H = 1/S$):

- Optimal number of processors P^*
- Optimal checkpointing interval T^*

Coping with both fail-stop and silent errors:



Models

Error model: exponential distribution, $\lambda_{\text{ind}} = 1/\mu_{\text{ind}}$
(memoryless and independent)

	error rate	error probability
Fail-stop errors	$\lambda_P^f = f\lambda_{\text{ind}}P$	$q_P^f = 1 - e^{-\lambda_P^f T}$
Silent errors	$\lambda_P^s = s\lambda_{\text{ind}}P$	$q_P^s = 1 - e^{-\lambda_P^s T}$

Resilience model:

Checkpointing time	$C_P = a + b/P + cP$
Verification time	$V_P = v + u/P$
Down time (fail-stop)	D

All coefficients (a, b, c, v, u, f, s, D) are assumed to be constants

Main Results

Exact execution time of a pattern in expectation (see paper)

First-order approximation of optimal P^*, T^* and H^*

Main Results

Exact execution time of a pattern in expectation (see paper)

First-order approximation of optimal P^*, T^* and H^*

- **Case 1**: checkpoint cost increases with P ($C_P = cP + o(P)$)

$$P^* = \left(\frac{1}{c\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}}\right)^{1/4}\left(\frac{1-\alpha}{2\alpha}\right)^{1/2} = \Theta(\lambda_{\text{ind}}^{-1/4})$$

$$T^* = \left(\frac{c}{\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}}\right)^{1/2} = \Theta(\lambda_{\text{ind}}^{-1/2})$$

$$H^* = \alpha + 2\left(4\alpha^2(1-\alpha)^2 c\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}\right)^{1/4} = \Theta(\lambda_{\text{ind}}^{1/4})$$

- **Case 2**: checkpoint/verif. cost constant ($C_P + V_P = d + o(1)$)

$$P^* = \left(\frac{1}{d\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}}\right)^{1/3}\left(\frac{1-\alpha}{\alpha}\right)^{2/3} = \Theta(\lambda_{\text{ind}}^{-1/3})$$

$$T^* = \left(\frac{d^2}{\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}}\right)^{1/3}\left(\frac{\alpha}{1-\alpha}\right)^{1/3} = \Theta(\lambda_{\text{ind}}^{-1/3})$$

$$H^* = \alpha + 3\left(\alpha^2(1-\alpha)d\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}\right)^{1/3} = \Theta(\lambda_{\text{ind}}^{1/3})$$

Main Results

Exact execution time of a pattern in expectation (see paper)

First-order approximation of optimal P^*, T^* and H^*

- **Case 1**: checkpoint cost increases with P ($C_P = cP + o(P)$)

$$P^* = \left(\frac{1}{c\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}}\right)^{1/4}\left(\frac{1-\alpha}{2\alpha}\right)^{1/2} \quad = \Theta(\lambda_{\text{ind}}^{-1/4})$$

$$T^* = \left(\frac{c}{\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}}\right)^{1/2} \quad = \Theta(\lambda_{\text{ind}}^{-1/2})$$

$$H^* = \alpha + 2\left(4\alpha^2(1-\alpha)^2 c\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}\right)^{1/4} \quad = \Theta(\lambda_{\text{ind}}^{1/4})$$

- **Case 2**: checkpoint/verif. cost constant ($C_P + V_P = d + o(1)$)

Processors ↑ $$P^* = \left(\frac{1}{d\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}}\right)^{1/3}\left(\frac{1-\alpha}{\alpha}\right)^{2/3} \quad = \Theta(\lambda_{\text{ind}}^{-1/3})$$

Interval ↓ $$T^* = \left(\frac{d^2}{\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}}\right)^{1/3}\left(\frac{\alpha}{1-\alpha}\right)^{1/3} \quad = \Theta(\lambda_{\text{ind}}^{-1/3})$$

Overhead ↓ $$H^* = \alpha + 3\left(\alpha^2(1-\alpha)d\left(\frac{f}{2}+s\right)\lambda_{\text{ind}}\right)^{1/3} \quad = \Theta(\lambda_{\text{ind}}^{1/3})$$

Limitation of First-Order Approximation

Difficulty with other (less practical) cases:
e.g., $C_P + V_P = h/P$ or $\alpha = 0$

Observation: Suppose $P = \Theta(\lambda_{\text{ind}}^{-x})$ and $T = \Theta(\lambda_{\text{ind}}^{-y})$. Then, for first-order approx. to accurately estimate error probabilities (e.g., $e^{-\lambda_P C_P}$, $e^{-\lambda_P V_P}$ and $e^{\lambda_P T}$), we need:

$$x < \delta, \text{ where } \delta = \begin{cases} 1/2 & \text{if } c \neq 0 \\ 1 & \text{if } c = 0 \end{cases}$$

$$x + y < 1$$

$$\Rightarrow \; P \cdot T < 1/\lambda_{\text{ind}} = \mu_{\text{ind}} \text{ (MTBF)}$$

Limitation of First-Order Approximation

Difficulty with other (less practical) cases:
e.g., $C_P + V_P = h/P$ or $\alpha = 0$

Observation: Suppose $P = \Theta(\lambda_{\text{ind}}^{-x})$ and $T = \Theta(\lambda_{\text{ind}}^{-y})$. Then, for first-order approx. to accurately estimate error probabilities (e.g., $e^{-\lambda_P C_P}$, $e^{-\lambda_P V_P}$ and $e^{\lambda_P T}$), we need:

$$x < \delta, \text{ where } \delta = \begin{cases} 1/2 & \text{if } c \neq 0 \\ 1 & \text{if } c = 0 \end{cases}$$

$$x + y < 1$$

$$\Rightarrow \quad P \cdot T < 1/\lambda_{\text{ind}} = \mu_{\text{ind}} \text{ (MTBF)}$$

Possible solution: second or high-order approximations with numerical methods

Simulation Settings

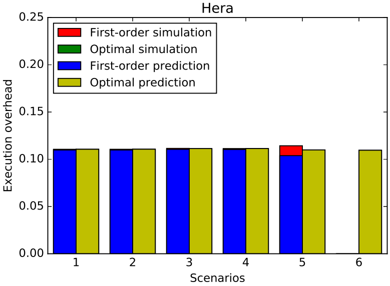
Table: Model parameters from SCR library [Moody et al. 2010]

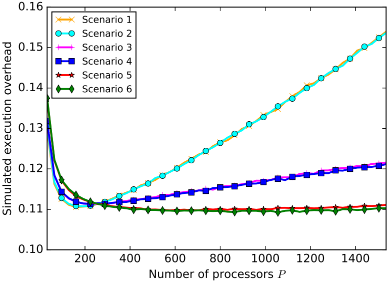
Platform	Hera	Atlas	Coastal	Coastal SSD
λ_{ind}	1.69e-8	1.62e-8	2.34e-9	2.34e-9
f	0.2188	0.0625	0.1667	0.1667
s	0.7812	0.9375	0.8333	0.8333
P	512	1024	2048	2048
C_P	$300s$	$439s$	$1051s$	$2500s$
V_P	$15.4s$	$9.1s$	$4.5s$	$180s$

Table: Different resilience scenarios

Scenario	1	2	3	4	5	6
C_P	cP	cP	a	a	b/P	b/P
V_P	v	u/P	v	u/P	v	u/P

Simulation Results

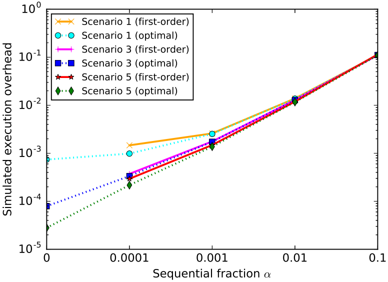


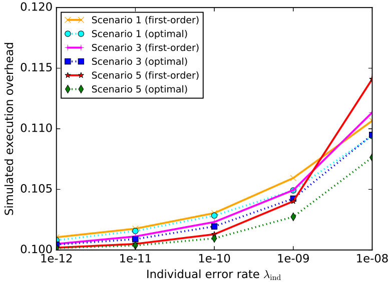


$$\alpha = 0.1$$

Simulation Results

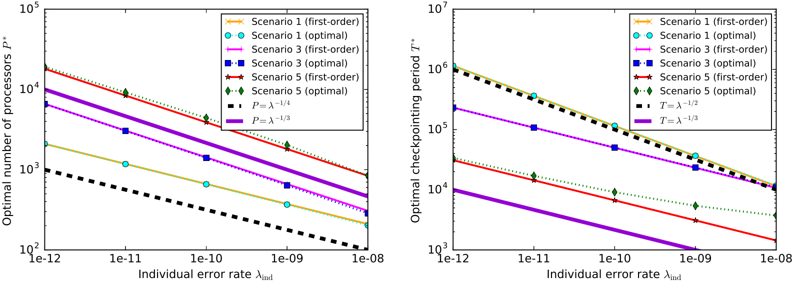
- **Impact of sequential fraction α and error rate λ_{ind}**





Simulation Results

- **Order of optimal P^* and T^***



Conclusion

What to remember

- Optimal P^* and T^* as function of MTBF $\mu_{\text{ind}} = 1/\lambda_{\text{ind}}$
  1. Checkpointing cost increases with P
     $\Rightarrow P^* = \Theta(\lambda_{\text{ind}}^{-1/4}), T^* = \Theta(\lambda_{\text{ind}}^{-1/2})$
  2. Checkpointing/verification cost remains constant
     $\Rightarrow P^* = \Theta(\lambda_{\text{ind}}^{-1/3}), T^* = \Theta(\lambda_{\text{ind}}^{-1/3})$

Future work

- Explore different speedup profiles, weak scaling, higher-order approximations