HPC

Loi d’Amdahl

Programmation parallèle

Distributed systems

Last update: Fév. 2024

CALCUL DE LA VITESSE

$$\mathrm{speedup}(P) = \frac{t_{\mathrm{seq}}}{t_P}$$

TEMPS PASSÉ PAR UN PROGRAMME SÉQUENTIEL

TEMPS PASSÉ PAR UN PROGRAMME PARALLÈLE ÉQUIVALENT SUR P PROCESSEURS

P = NOMBRE DE PROCESSEURS

CALCUL DE L'EFFICACITÉ

$$e = \frac{\text{speedup}(P)}{P} = \frac{t_{\text{seq}}}{P \times t_P}$$

P = NOMBRE DE PROCESSEURS

TEMPS D'EXÉCUTION DU PROGRAMME PARALLÈLE EXÉCUTÉ SUR P PROCESSEURS.

SCALABILITÉ

$$\text{scalability}(O, P) = \frac{t_O}{t_P} \text{ with } O < P$$

O = NOMBRE DE PROCESSEURS SATISFAISANT LA CONDITION C

OP DOIT ÊTRE INFÉRIEUR AU NOMBRE DE PROCESSEURS DANS LE CALCUL PARALLÈLE

ANALYSE D'ISO-EFFICIENCE ET SCALABILITÉ

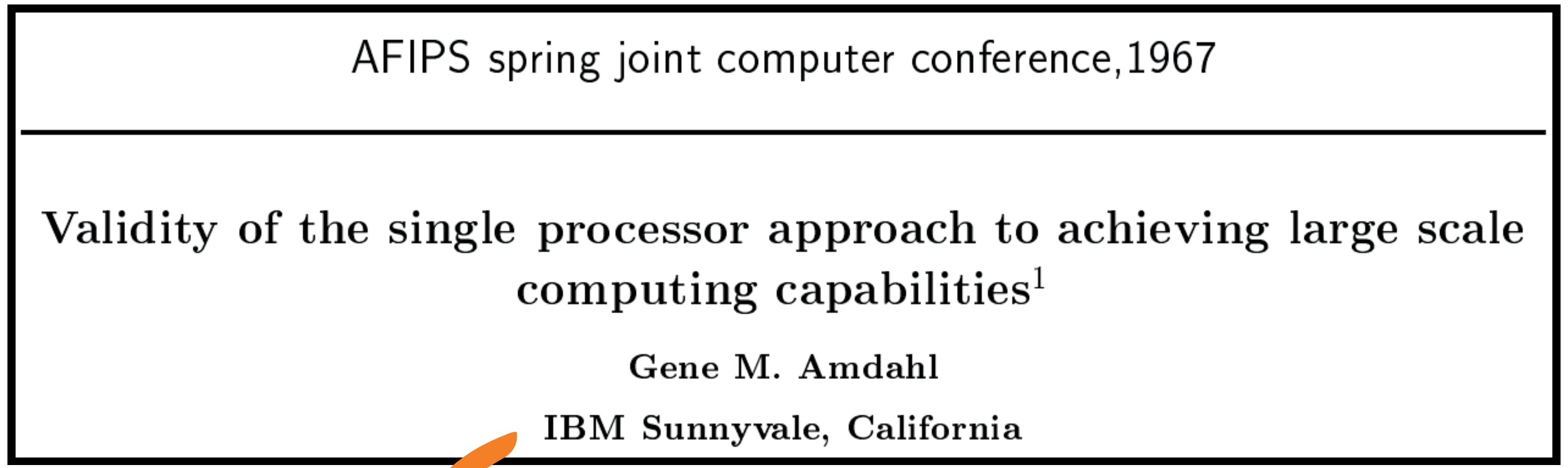
- Pour maintenir une efficacité constante dans les algorithmes parallèles, il est crucial d'ajuster la taille des données en fonction du nombre de processeurs.
- L'analyse d'iso-efficience vise à déterminer le taux de croissance optimal de la taille des données par rapport au nombre de processeurs pour garantir des performances stables lors de l'augmentation du nombre de processeurs.

LOI D'AMDAHL

La loi d'Amdahl est un principe fondamental du calcul parallèle qui décrit le gain de vitesse potentiellement réalisable en parallélisant un calcul.

Elle a été formulée par l'architecte informatique Gene Amdahl en 1967.

La loi d'Amdahl stipule que le gain de vitesse global d'un calcul, lorsqu'il est parallélisé, est limité par la fraction du calcul qui doit être exécutée séquentiellement.

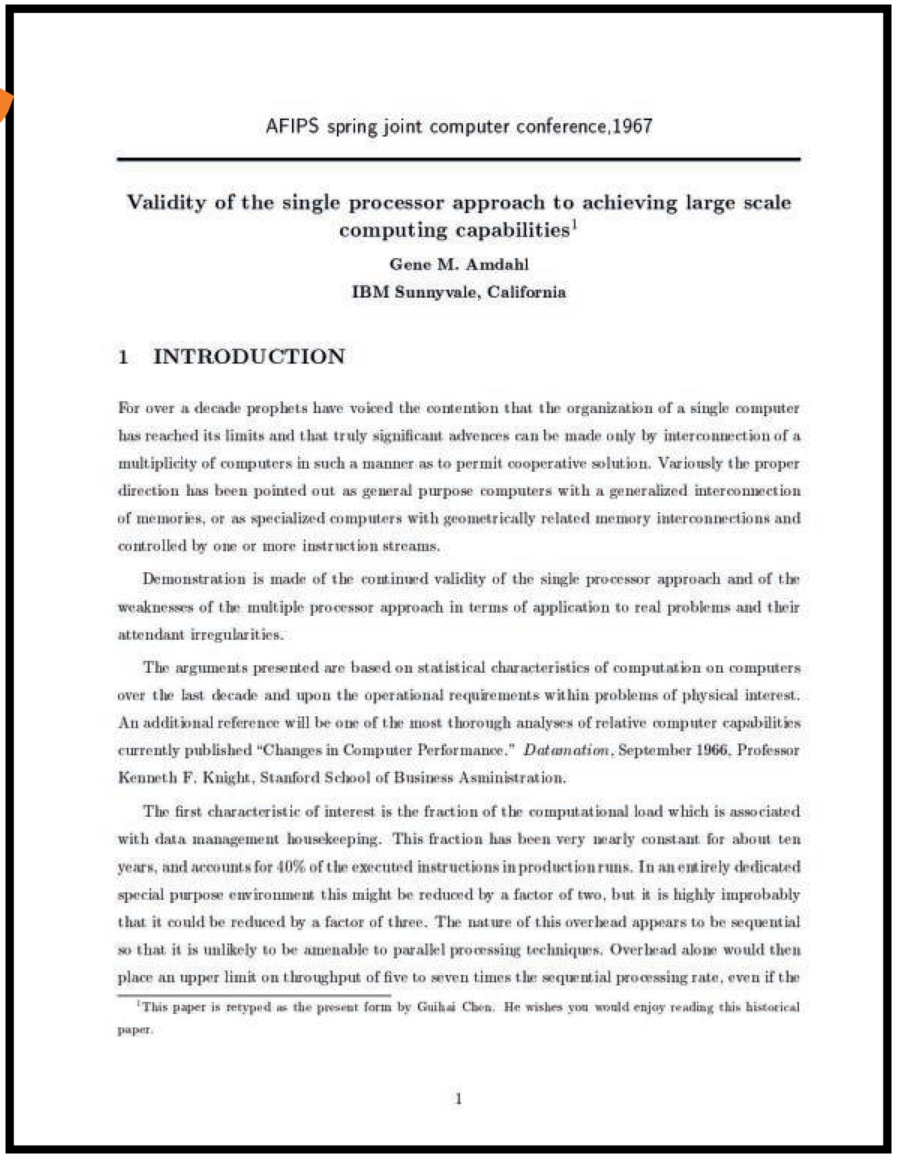


AFIPS spring joint computer conference,1967

Validity of the single processor approach to achieving large scale computing capabilities[1]

Gene M. Amdahl

IBM Sunnyvale, California



AFIPS spring joint computer conference,1967

Validity of the single processor approach to achieving large scale computing capabilities[1]

Gene M. Amdahl

IBM Sunnyvale, California

1 INTRODUCTION

For over a decade prophets have voiced the contention that the organization of a single computer has reached its limits and that truly significant advences can be made only by interconnection of a multiplicity of computers in such a manner as to permit cooperative solution. Variously the proper direction has been pointed out as general purpose computers with a generalized interconnection of memories, or as specialized computers with geometrically related memory interconnections and controlled by one or more instruction streams.

Demonstration is made of the continued validity of the single processor approach and of the weaknesses of the multiple processor approach in terms of application to real problems and their attendant irregularities.

The arguments presented are based on statistical characteristics of computation on computers over the last decade and upon the operational requirements within problems of physical interest. An additional reference will be one of the most thorough analyses of relative computer capabilities currently published "Changes in Computer Performance." *Datamation*, September 1966, Professor Kenneth F. Knight, Stanford School of Business Asministration.

The first characteristic of interest is the fraction of the computational load which is associated with data management housekeeping. This fraction has been very nearly constant for about ten years, and accounts for 40% of the executed instructions in production runs. In an entirely dedicated special purpose environment this might be reduced by a factor of two, but it is highly improbably that it could be reduced by a factor of three. The nature of this overhead appears to be sequential so that it is unlikely to be amenable to parallel processing techniques. Overhead alone would then place an upper limit on throughput of five to seven times the sequential processing rate, even if the

[1]This paper is retyped as the present form by Guihai Chen. He wishes you would enjoy reading this historical paper.

1

LOI D'AMDAHL

$$t_P = \alpha_{\text{seq}} t_1 + (1 - \alpha_{\text{seq}}) \frac{t_1}{P}$$

$$= \alpha_{\text{par}} \frac{t_1}{P} + \alpha_{\text{seq}} t_1 .$$

LOI D'AMDAHL

En supposant que
$t_{\text{seq}} = t_1$
nous déduisons que la vitesse d'une implémentation parallèle utilisant P processeurs peut être exprimée par :

$$\operatorname{speedup}(P) = \frac{t_1}{t_P} = \frac{(\alpha_{\text{par}} + \alpha_{\text{seq}})t_1}{(\alpha_{\text{seq}} + \frac{\alpha_{\text{par}}}{P})t_1} = \frac{1}{\alpha_{\text{seq}} + \frac{\alpha_{\text{par}}}{P}}.$$

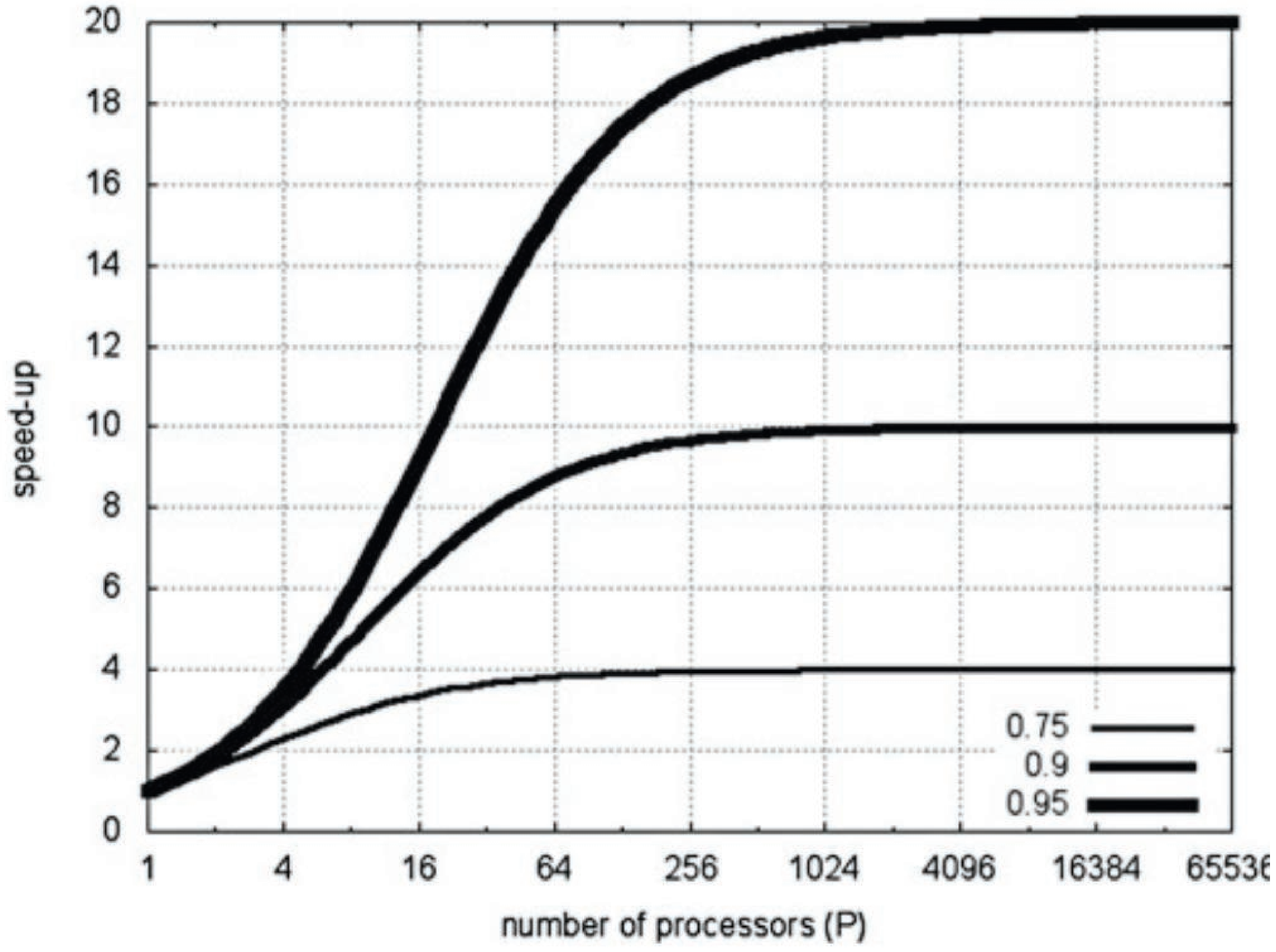
$$\boxed{\lim_{P\to\infty} \operatorname{speedup}(P) = \frac{1}{\alpha_{\text{seq}}} = \frac{1}{1-\alpha_{\text{par}}}.}$$

LOI D'AMDAHL

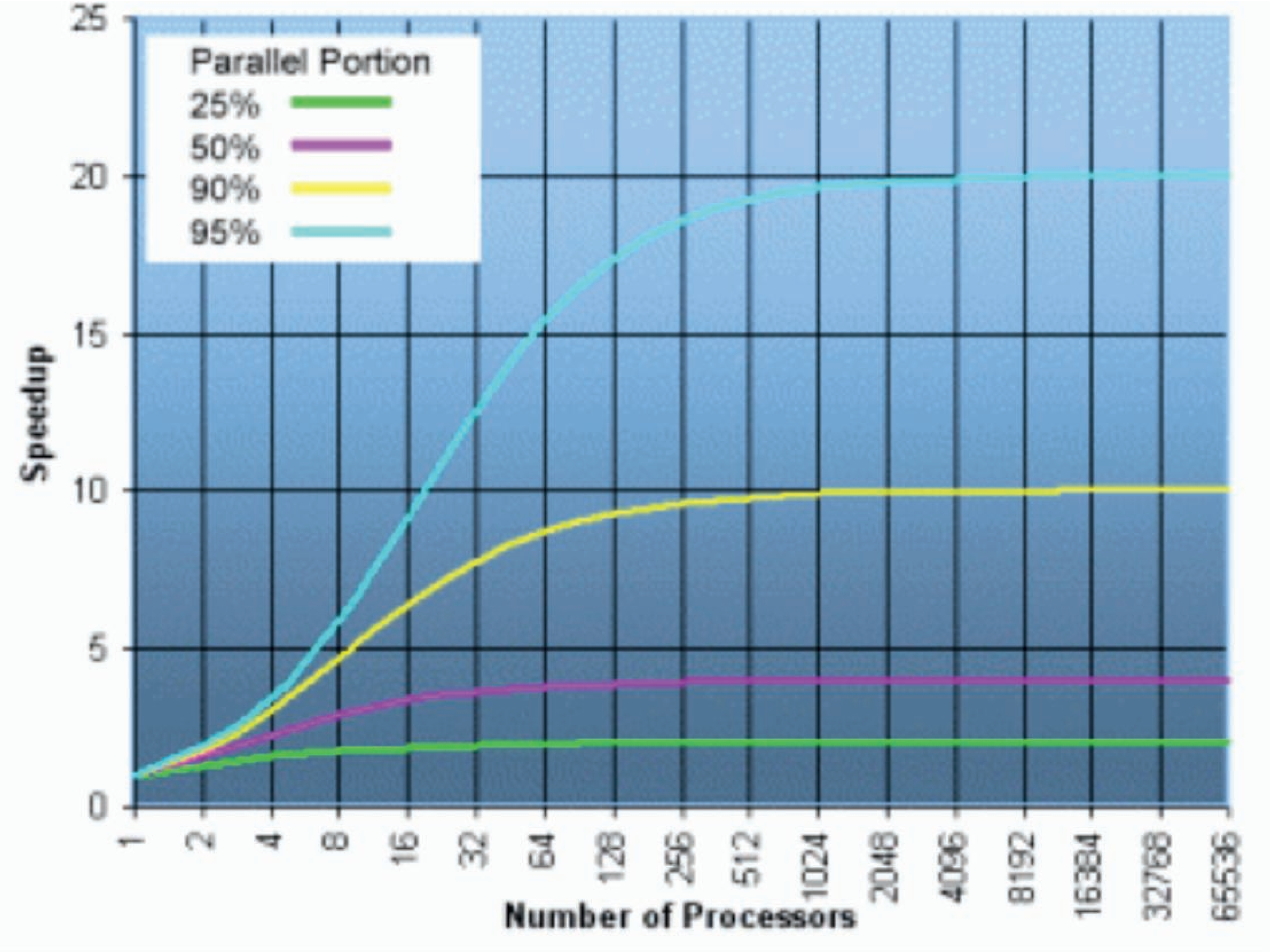
Theorem 1 *Amdahl's law gives the optimal asymptotic speedup of a parallel program as* speedup $= \frac{1}{\alpha_{\text{seq}}} = \frac{1}{1-\alpha_{\text{par}}}$, *where* α_{par} *denotes the proportion of the program that can be parallelized and* $\alpha_{\text{seq}} = 1 - \alpha_{\text{par}}$ *the proportion of code that is intrinsically sequential.*

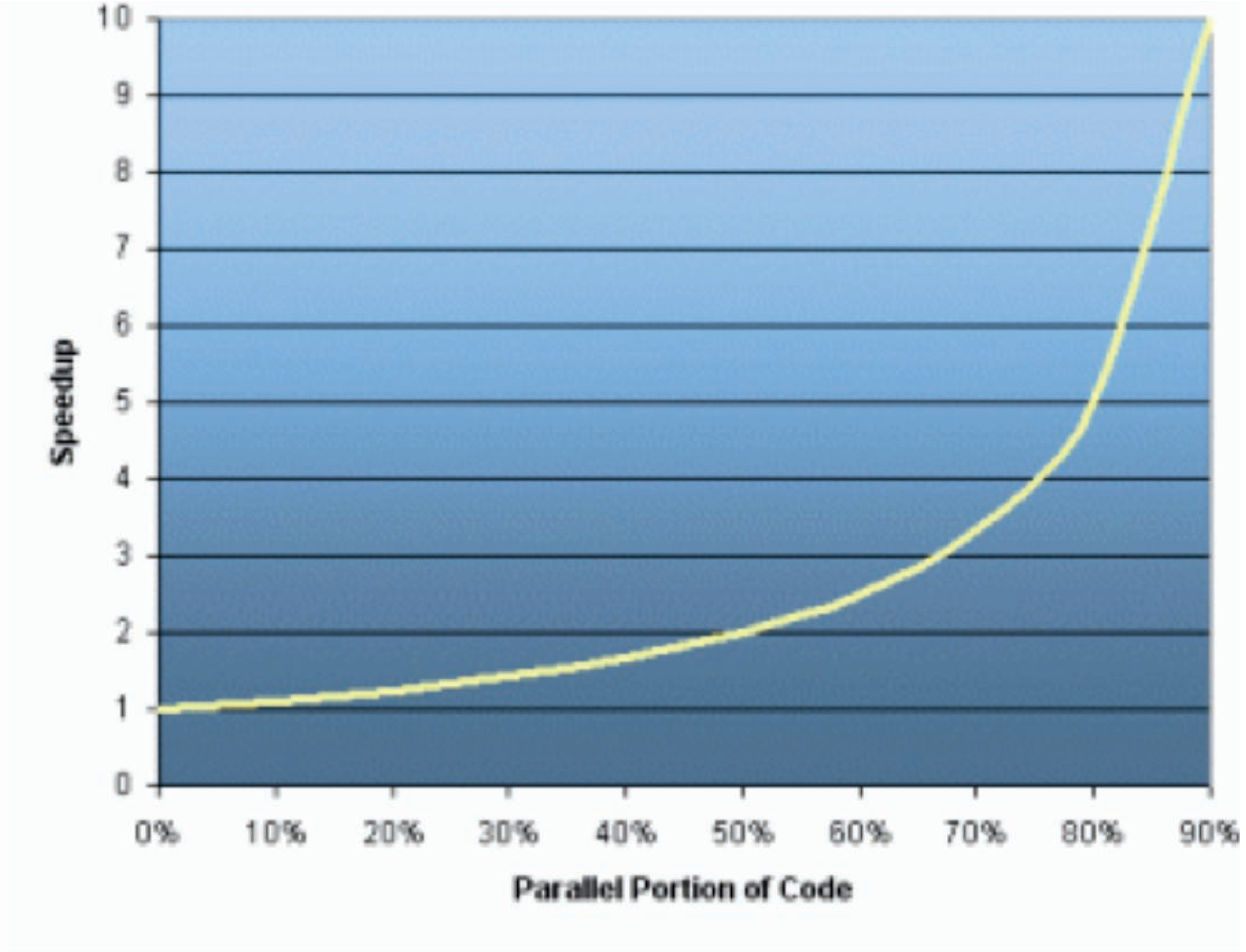
$$\boxed{\lim_{P\to\infty} \text{speedup}(P) = \frac{1}{\alpha_{\text{seq}}} = \frac{1}{1-\alpha_{\text{par}}}.}$$

	speedup			
N	P = .50	P = .90	P = .95	P = .99
10	1.82	5.26	6.89	9.17
100	1.98	9.17	16.80	50.25
1,000	1.99	9.91	19.62	90.99
10,000	1.99	9.91	19.96	99.02
100,000	1.99	9.99	19.99	99.90



LOI D'AMDAHL


Parallel Portion
25%
50%
90%
95%
Speedup
0
5
10
15
20
25
1
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
16384
32768
65536
Number of Processors


Speedup
0
1
2
3
4
5
6
7
8
9
10
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
Parallel Portion of Code

LOI D'AMDAHL : RÉSUMÉ

La loi d'Amdahl met en évidence les points suivants :
Si P=1, alors le Gain de temps est proportionnelle Nombre de processeurs

La loi d'Amdahl met en évidence les points suivants :

1. **Avantages Potentiels** : La loi d'Amdahl montre que la parallélisation d'un programme peut conduire à un gain de vitesse significatif, surtout lorsque qu'une grande partie du calcul peut être parallélisée et lorsque le nombre d'unités de traitement est suffisamment élevé.
2. **Limites** : Cependant, la loi d'Amdahl démontre également qu'il existe des limites inhérentes au gain de vitesse qui peut être obtenu grâce à la parallélisation. Même si une grande partie du calcul peut être parallélisée, il y aura toujours une partie séquentielle qui limitera le gain de vitesse global.
3. **Coûts** : Paralléliser un calcul implique souvent des coûts supplémentaires, tels que le surcoût associé à la parallélisation (par exemple, le surcoût de communication, le surcoût de synchronisation), la complexité de la programmation parallèle et la nécessité de matériel spécialisé dans certains cas.

LOI D'AMDAHL : RÉSUMÉ

La loi d'Amdahl suppose une taille d'entrée de données constante et une charge de travail prescrite une fois pour toutes.

Peu importe le nombre de processeurs disponibles, le speedup maximal théorique sera asymptotiquement borné par l'inverse de la proportion de code séquentiel.

Au delà de cette loi d'Amdahl, jouer sur la proximité de l'emplacement des données de calcul peut réduire drastiquement le temps de calcul observé.
Logique: plus les données sont loin du processeurs, plus long sera le temps de calcul