22 СЕНТЯБРЯ 2017, АСПИРАНТСКИЙ СЕМИНАР

Анализ глобального времени в алгоритмах параллельного моделирования дискретных событий

Зиганурова Лилия

National Research University Higher School of Economics, Moscow
Scientific Center in Chernogolovka, Chernogolovka, Moscow area

IN COLLABORATION WITH:

L.N. Shchur

1. Scientific Center in Chernogolovka, Moscow area, Russia
2. National Research University Higher School of Economics, Moscow, Russia
3. Landau Institute for Theoretical Physics, Moscow, Russia

M. Novotny

1. Mississippi State University, USA

OVERVIEW

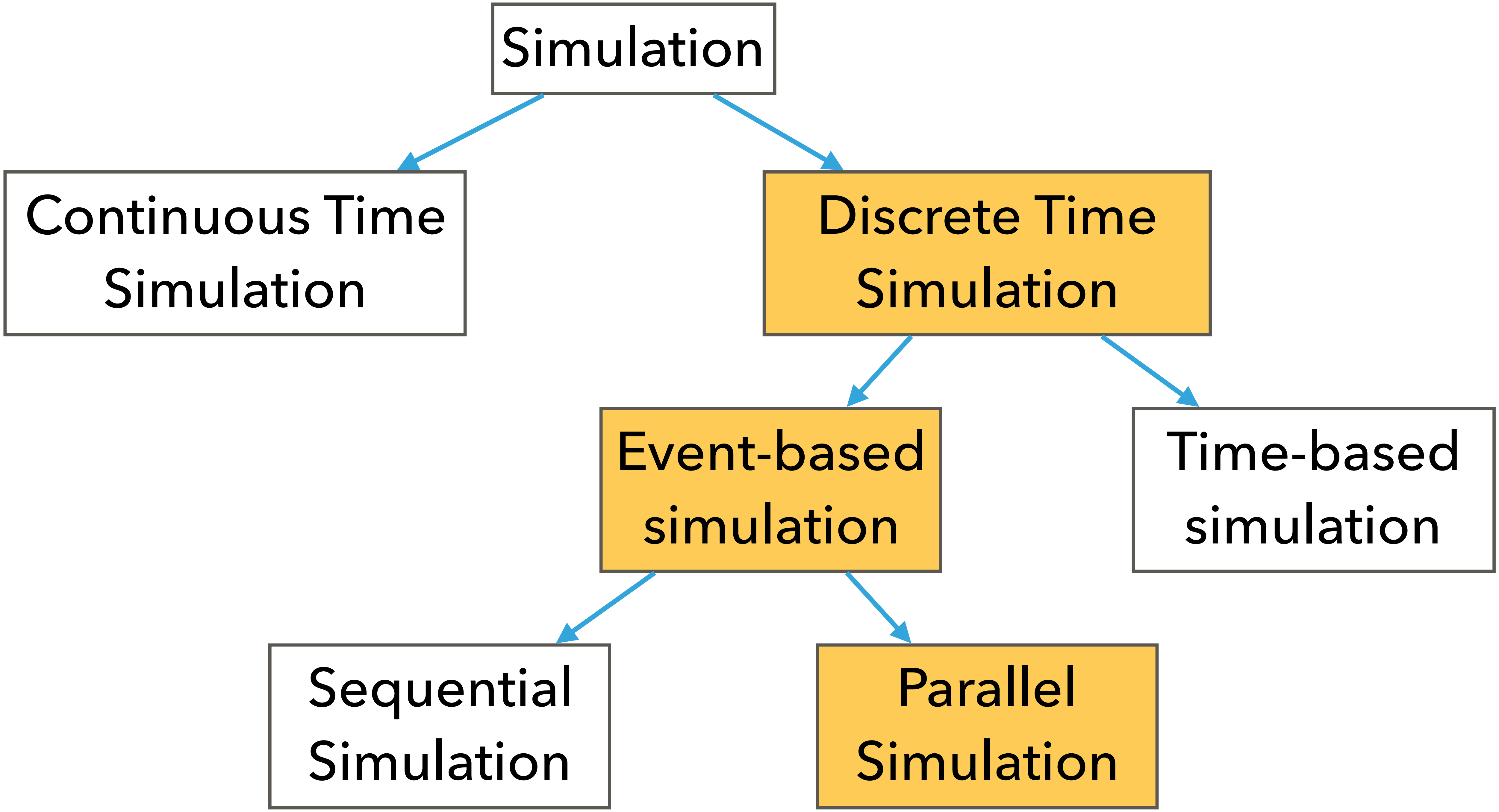
- What is the Parallel Discrete Event Simulation?
- Conservative synchronisation algorithm
- Underlying topologies
- The simulation of the algorithm on Small-World networks
- Results: How the underlying topology affects the synchronisation in PDES?

MOTIVATION

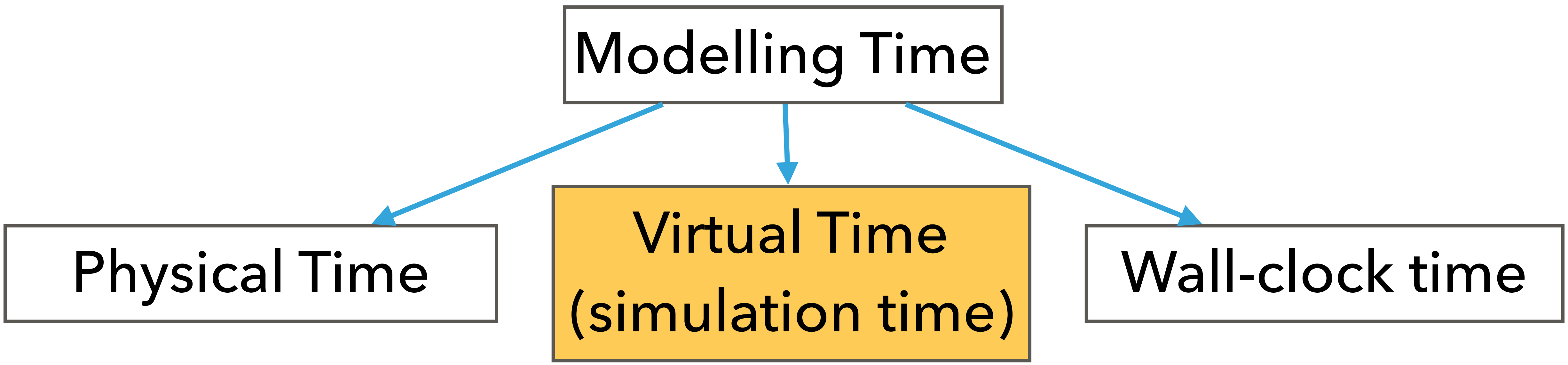
- Modern computer systems: **10^4 of nodes**
- Each node may have many CPUs, cores, and numerical accelerator
- CPU's (cores, threads, …) must be **synchronised** to efficiently execute one parallel program
- High performance computing requires new approaches to programming models

Parallel Discrete Event Simulation is a method of large-scale simulation which allows to execute a single program on a parallel computer.

CLASSIFICATION OF COMPUTER MODELLING



CLASSIFICATION OF MODELLING TIME



Время в физической моделируемой системе

Например, моделируем систему из N одинаковых молекул в объеме V и при температуре T в течение **10 минут**

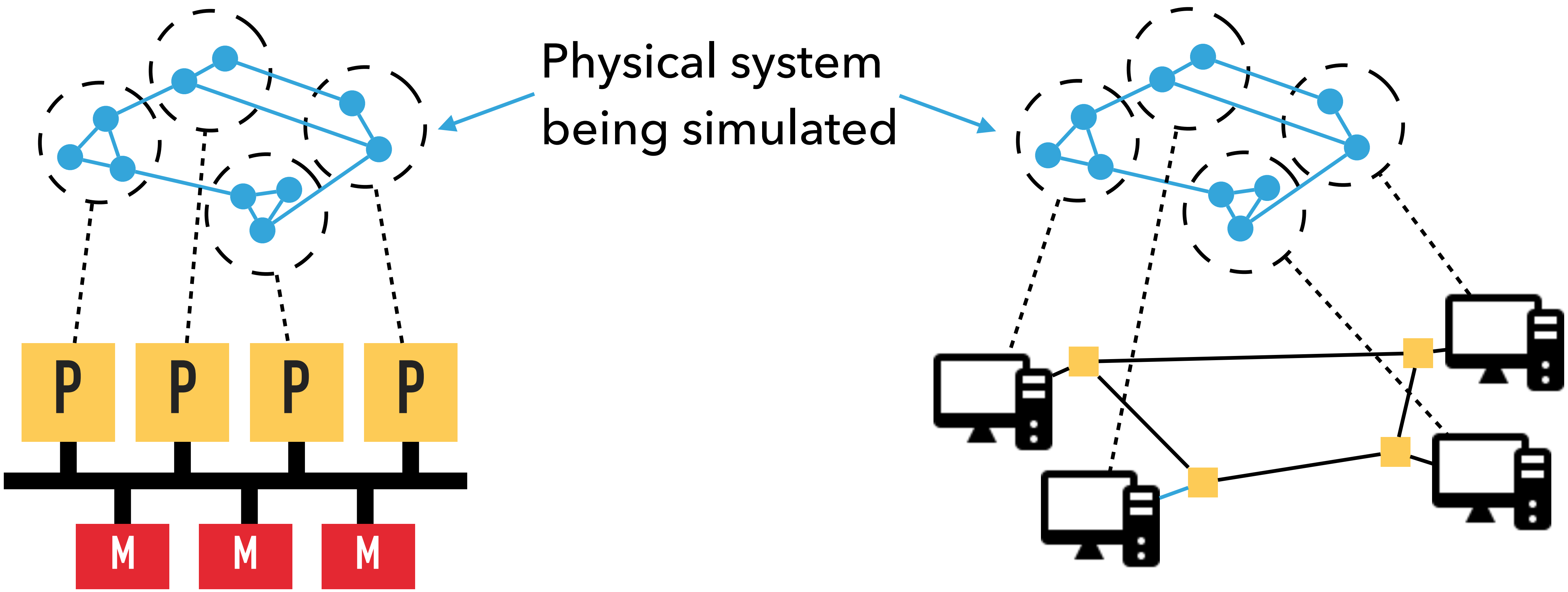
Виртуальное или модельное время

Пусть шаг моделирования 1сек. Тогда дискретная переменная t в программе будет меняться **от 0 до 600**

Время работы программы

Мы запустили программу и она завершила моделирование за **30 секунд**

PARALLEL AND DISTRIBUTED SIMULATION*



Parallel simulation involves the execution of a *single* simulation on a collection of **tightly** coupled processors (e.g. a shared memory multiprocessor)

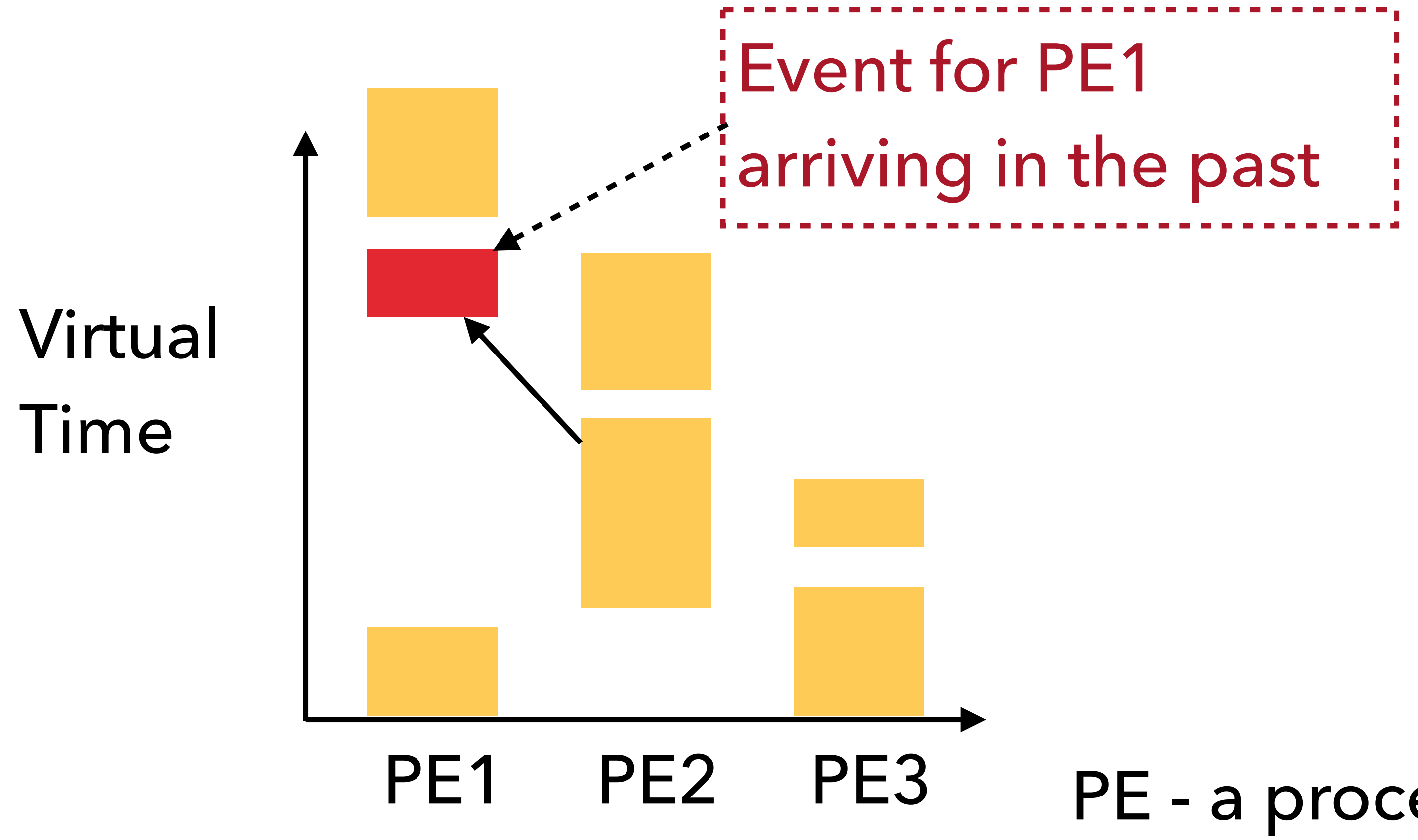
Distributed simulation involves the execution of a *single* simulation on a collection of **loosely** coupled processors (e.g. PCs interconnected by a LAN or WAN)

**from R.Fujimoto[1]*

ESSENTIAL PROPERTIES OF PDES:

- Changes in subsystems occur at some instant of time and are called **discrete events.**
- To preserve causality between dependent objects some **synchronisation protocol** is used.
- Using the **virtual time concept**.
- Communication between parallel processes goes via timestamped messages.
- No information exchange and no shared memory between subsystems.

SYNCHRONISATION IN PDES



PE - a processing element (node/core/thread/etc.)

processed event

“straggler” event

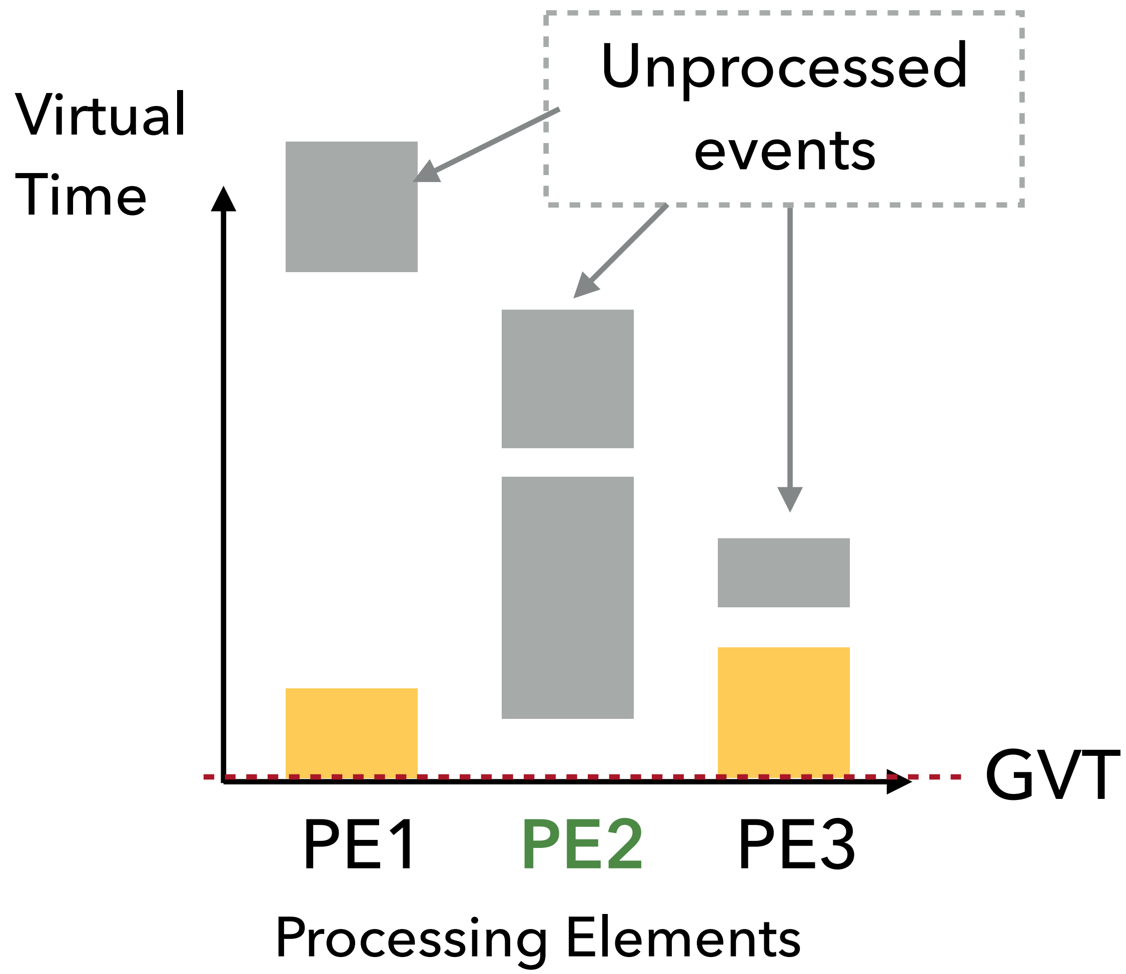
Local causality constrain: events must be processed in time-stamped order

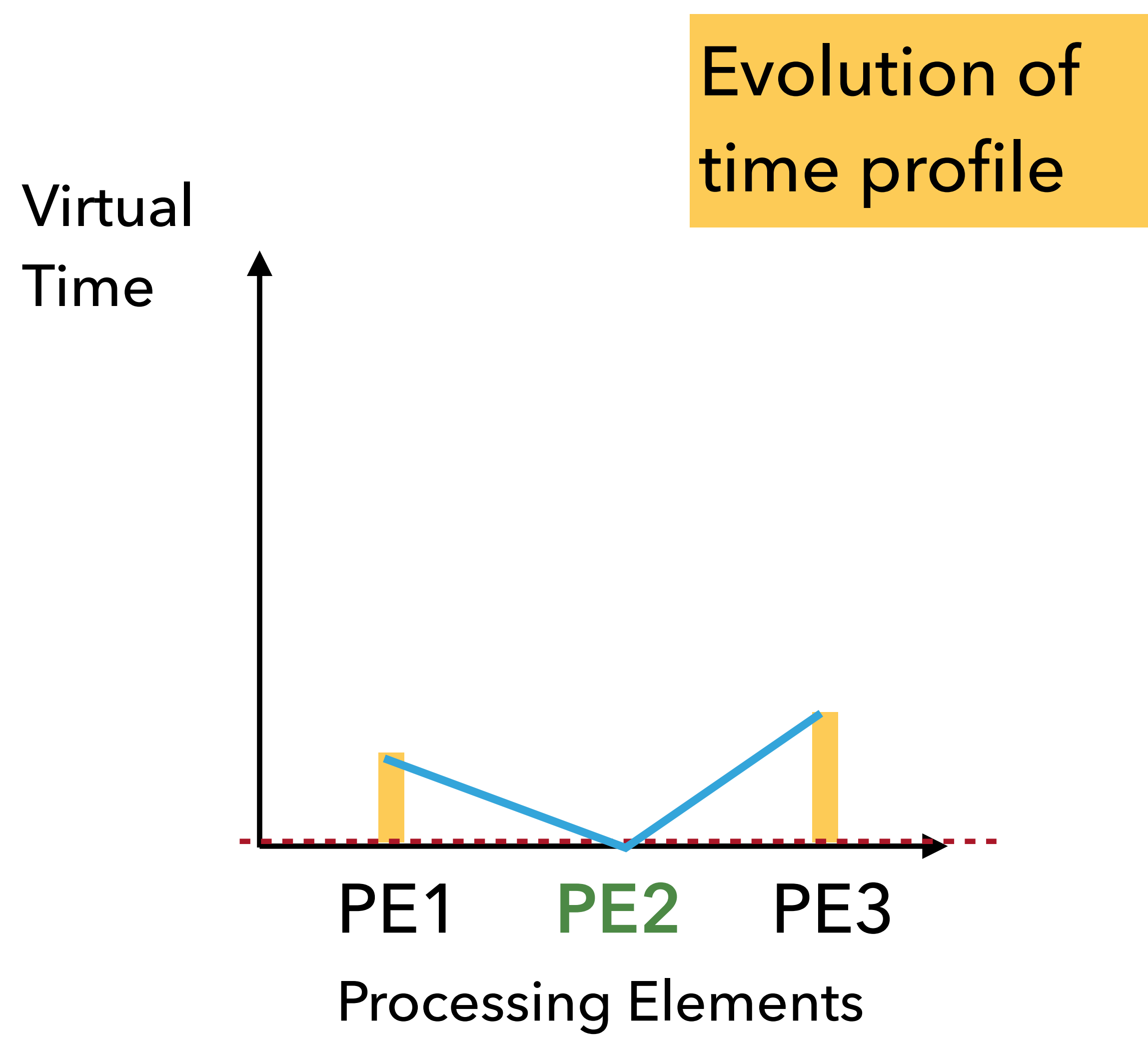
SYNCHRONISATION ALGORITHMS

- **Conservative[1]** - avoids all possible causality violations by checking the causality relations between dependent events at each discrete step of simulation
- **Optimistic[2]** - allows some causality errors, has a roll-back mechanism
- **Freeze-and-Shift[3]**

CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

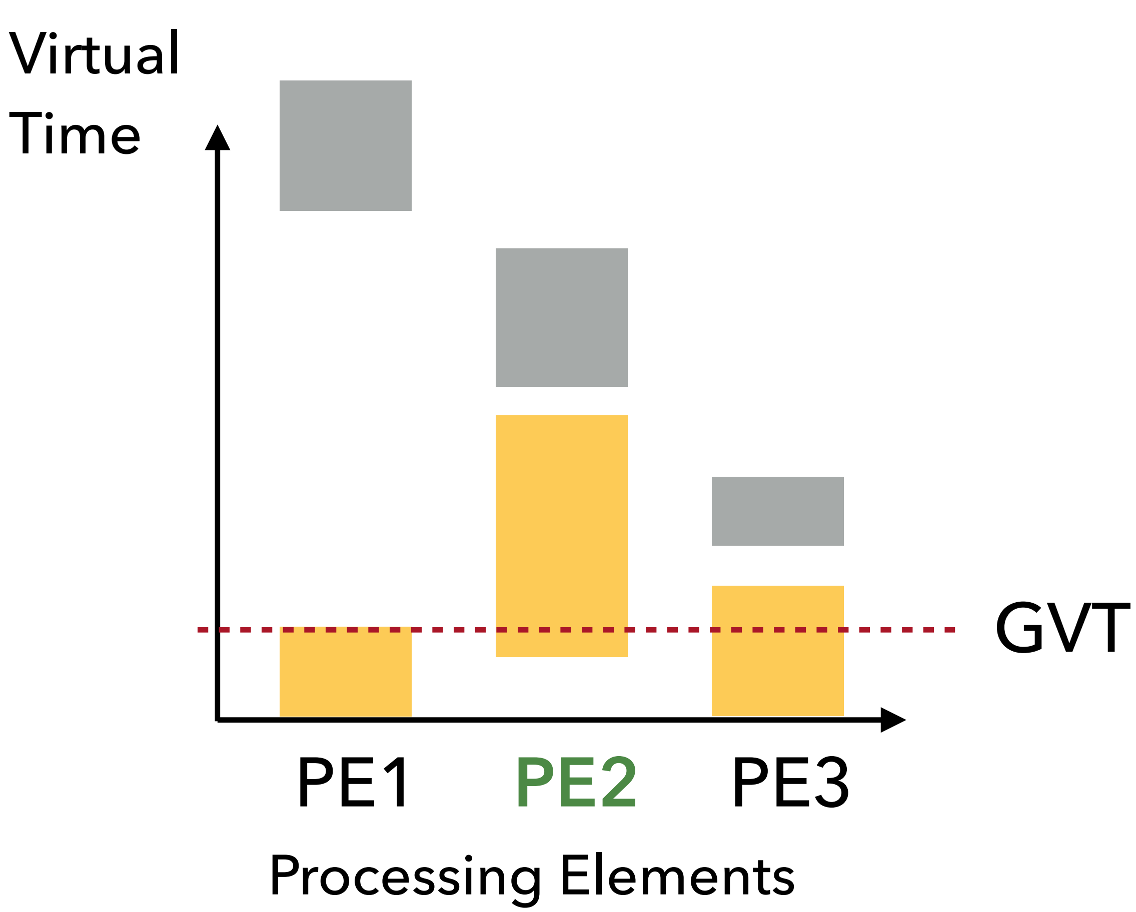


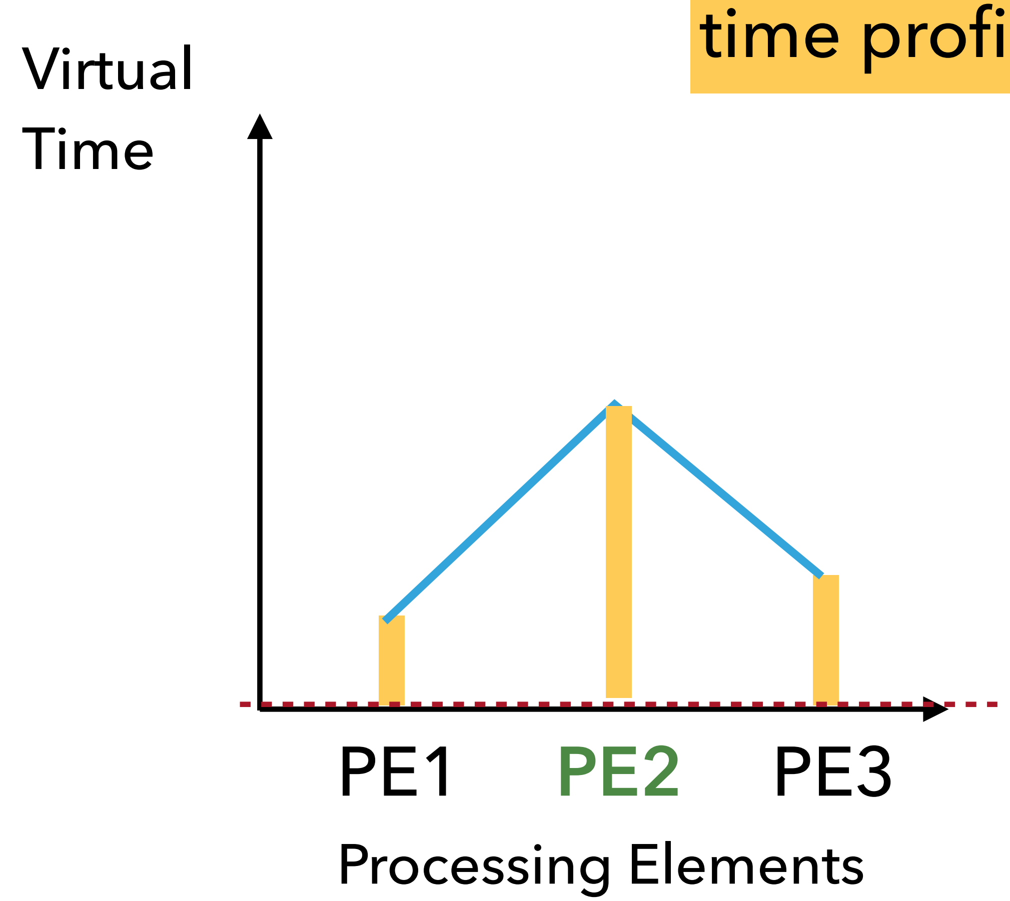


CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

Evolution of time profile

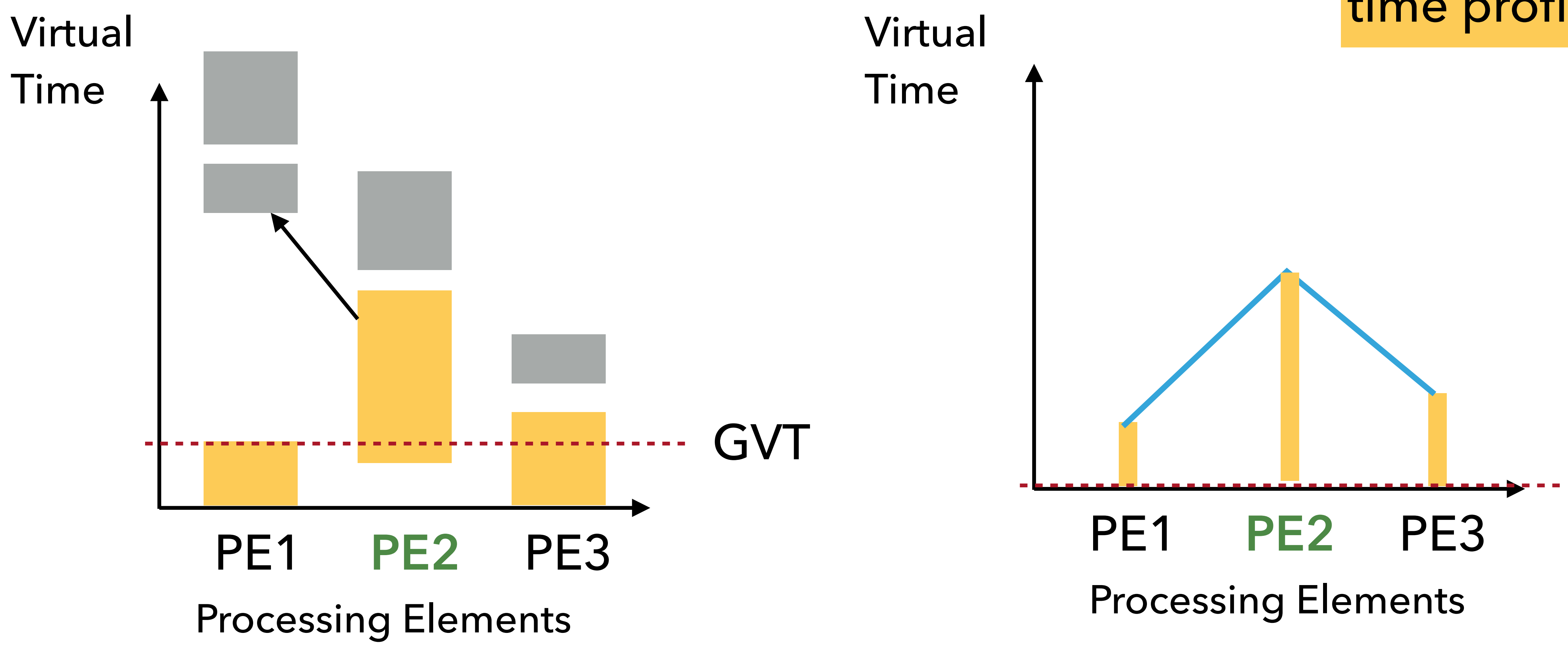




CONSERVATIVE SYNCHRONISATION ALGORITHM

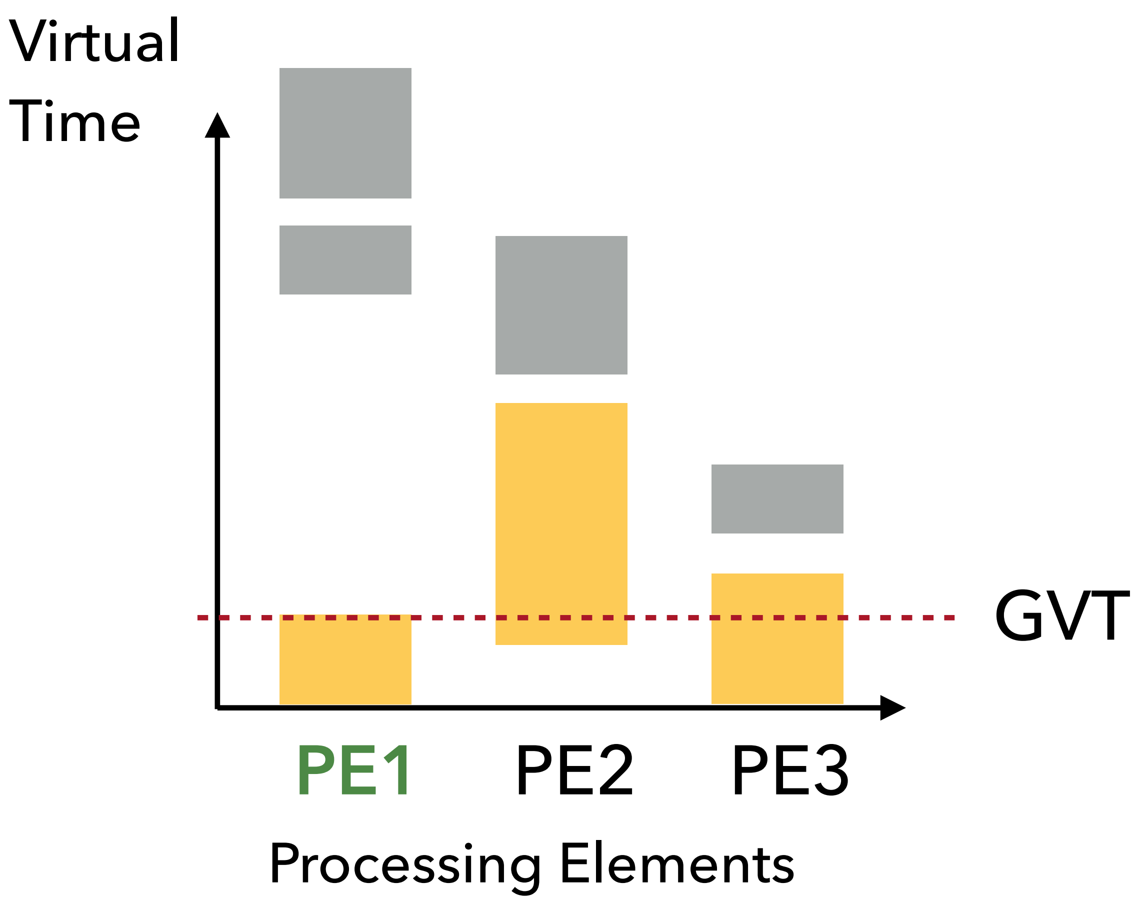
Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

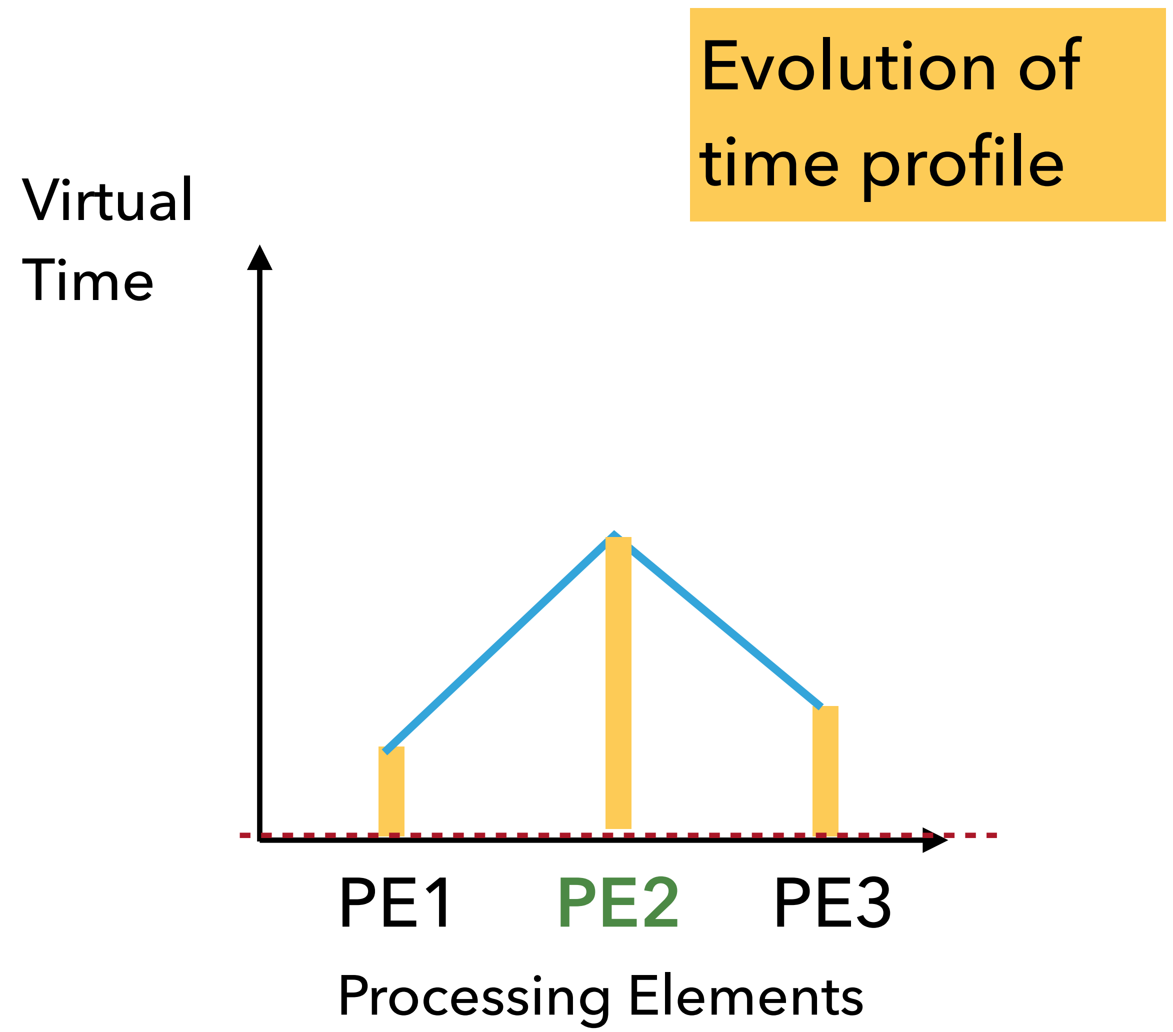
Evolution of time profile



CONSERVATIVE SYNCHRONISATION ALGORITHM

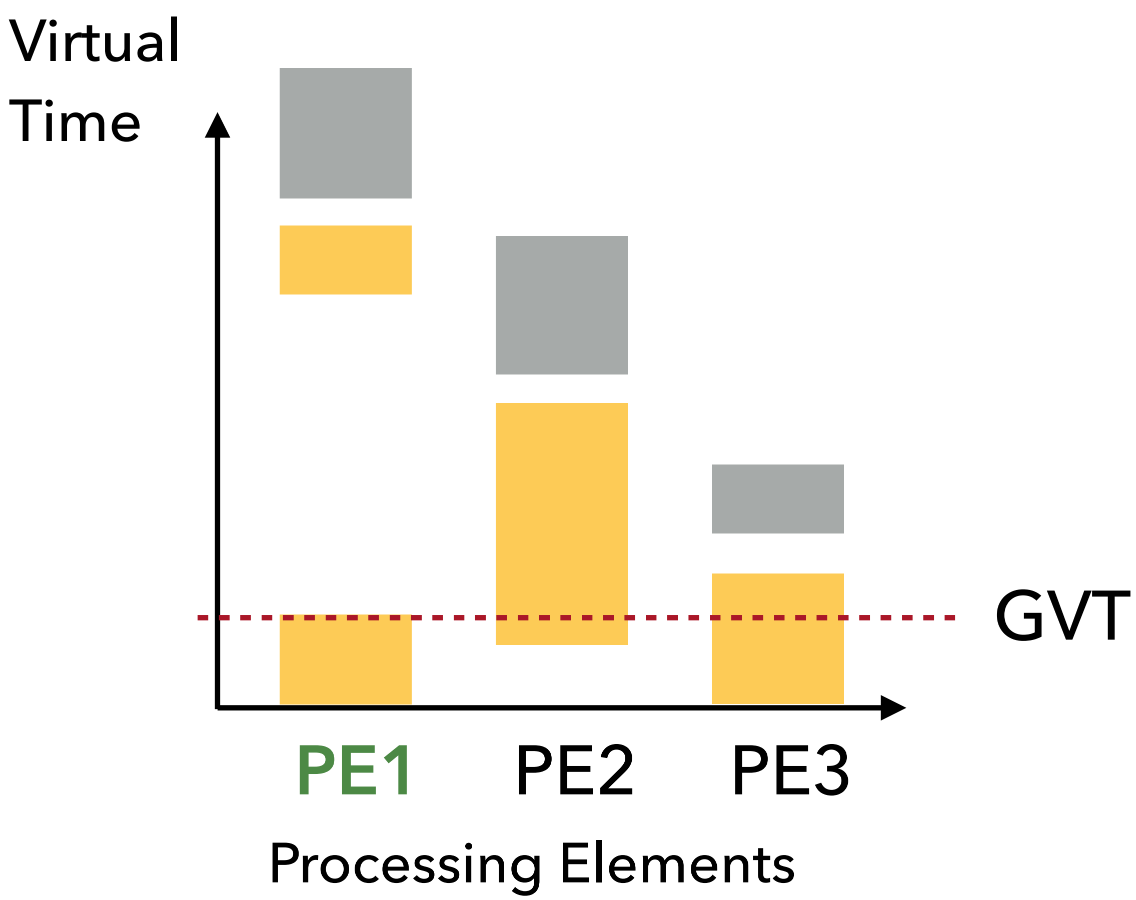
Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.



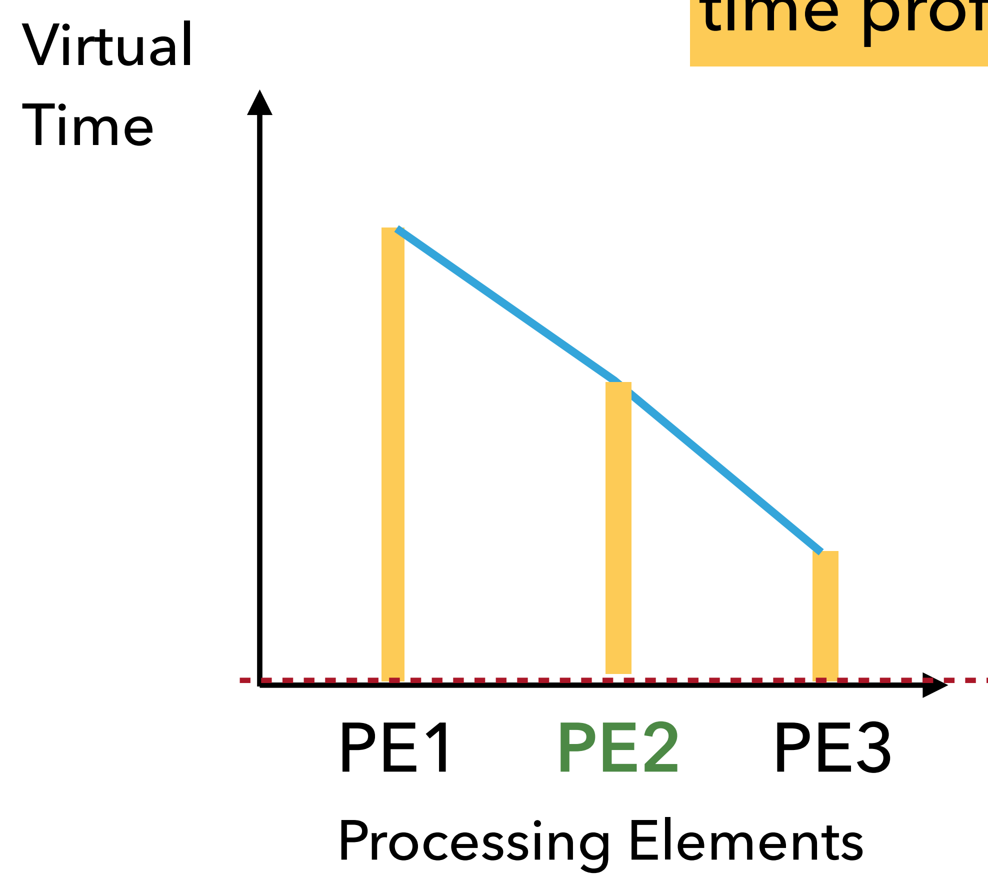


CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

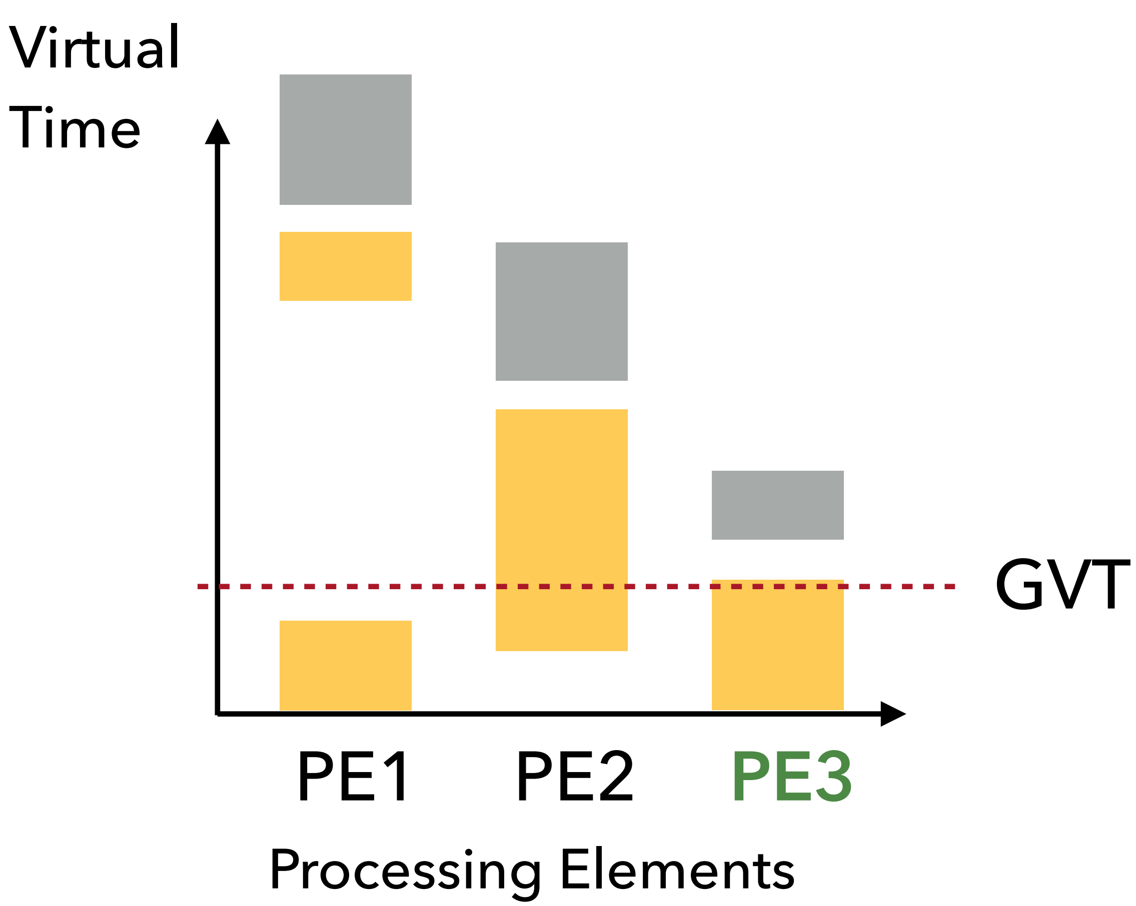


Evolution of time profile

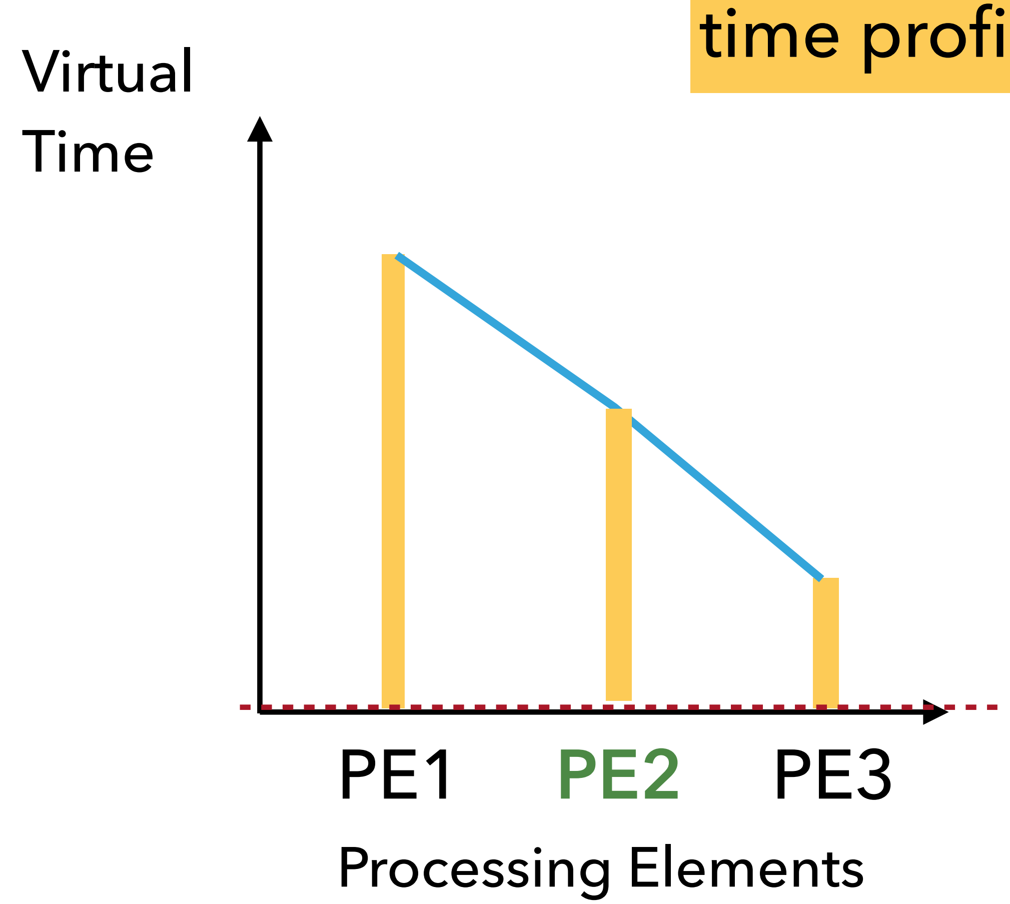


CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

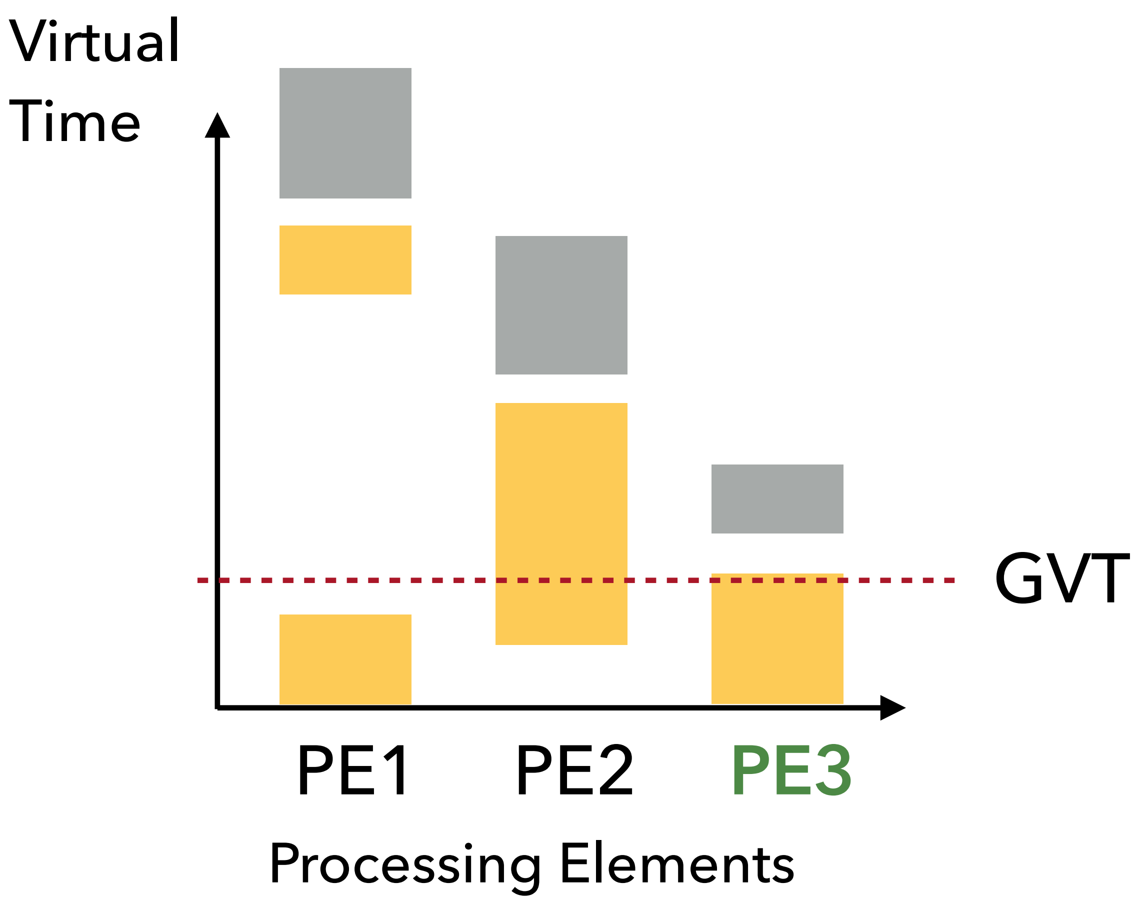


Evolution of time profile

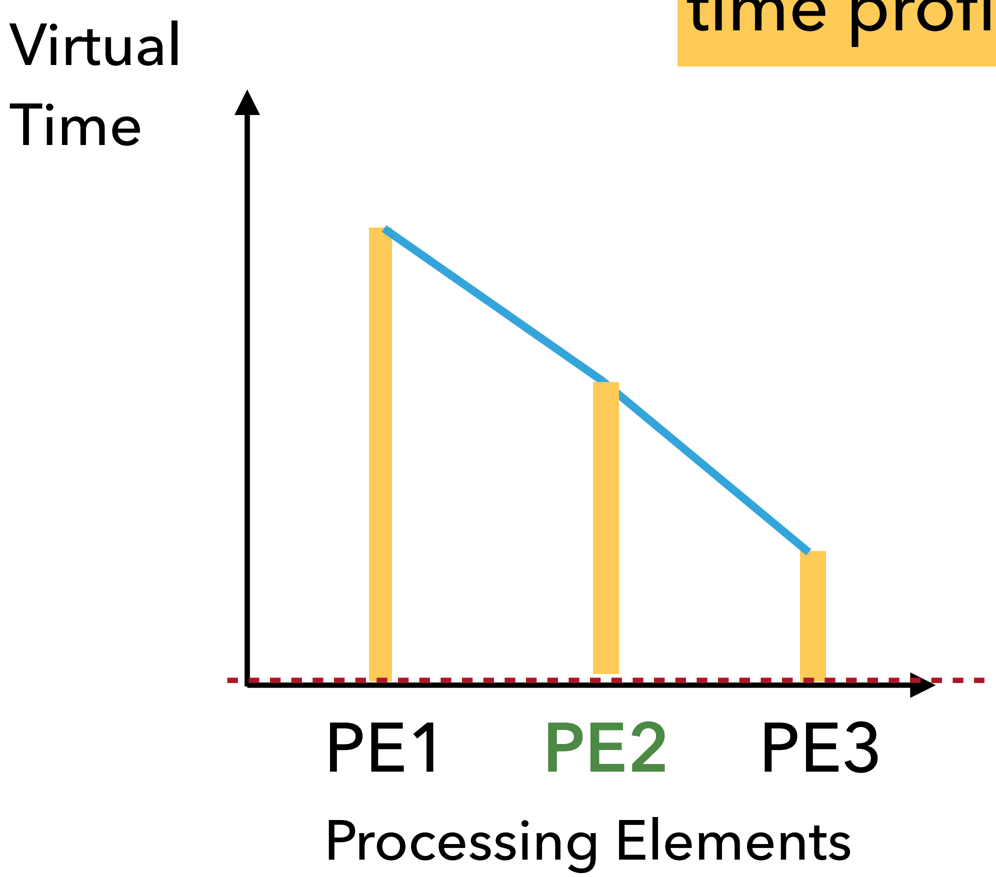


CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

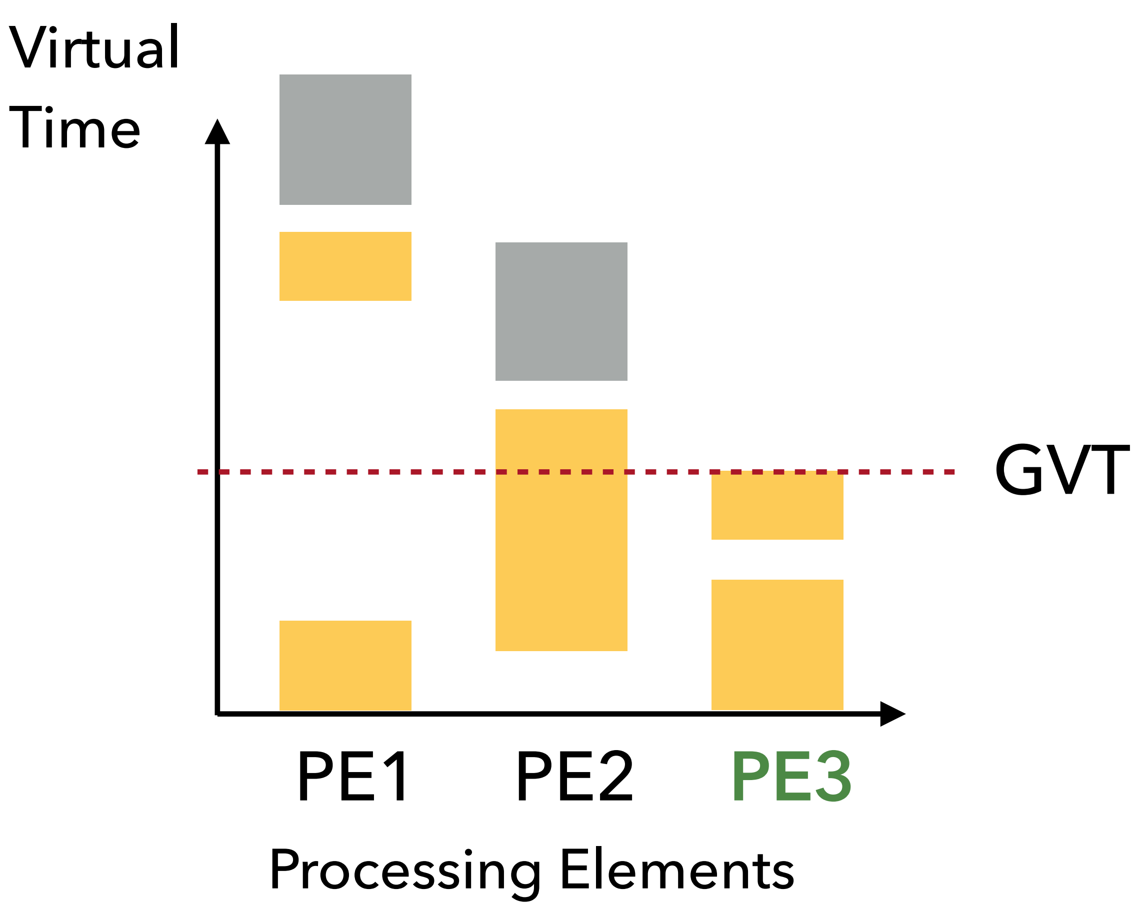


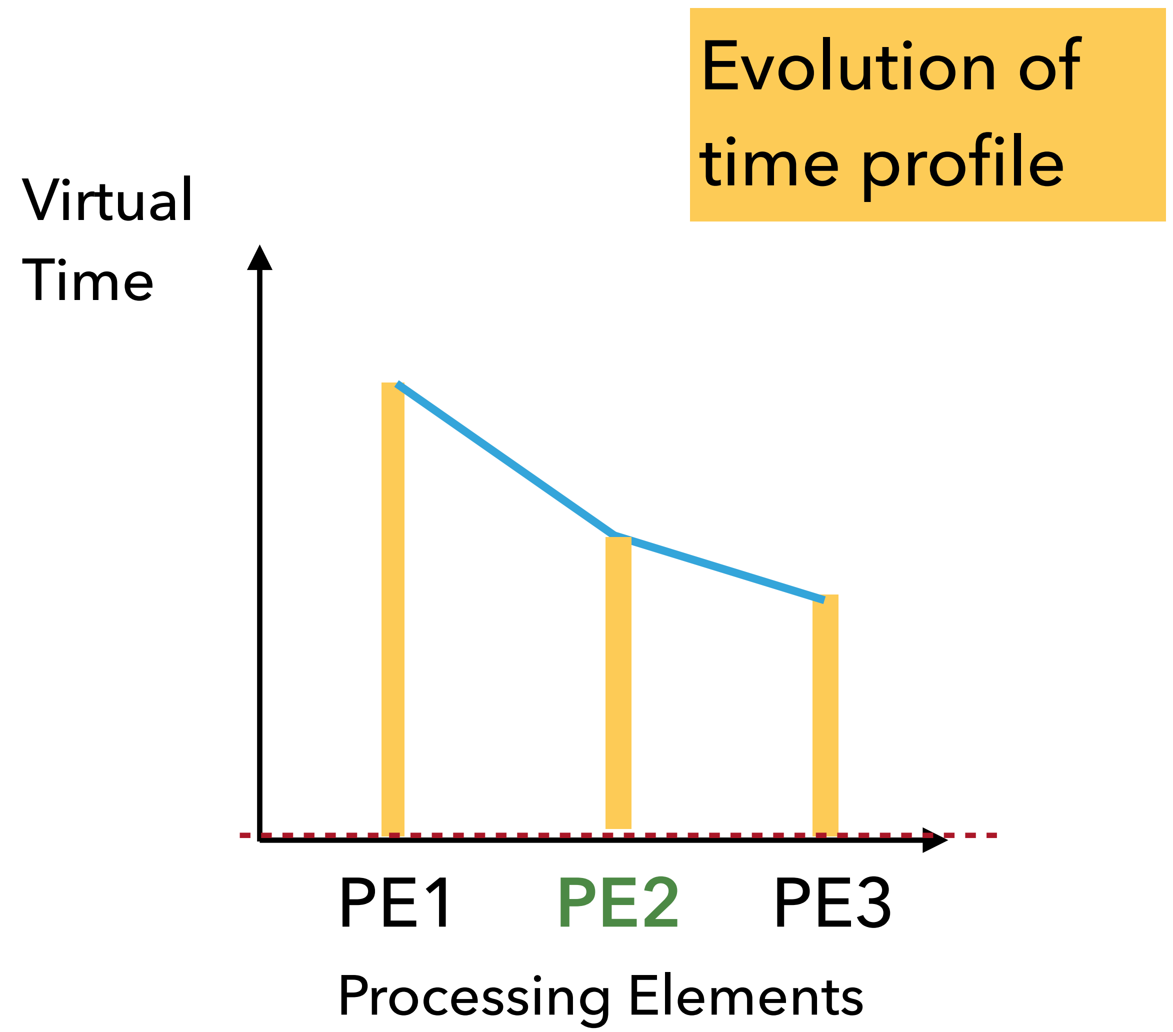
Evolution of time profile



CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

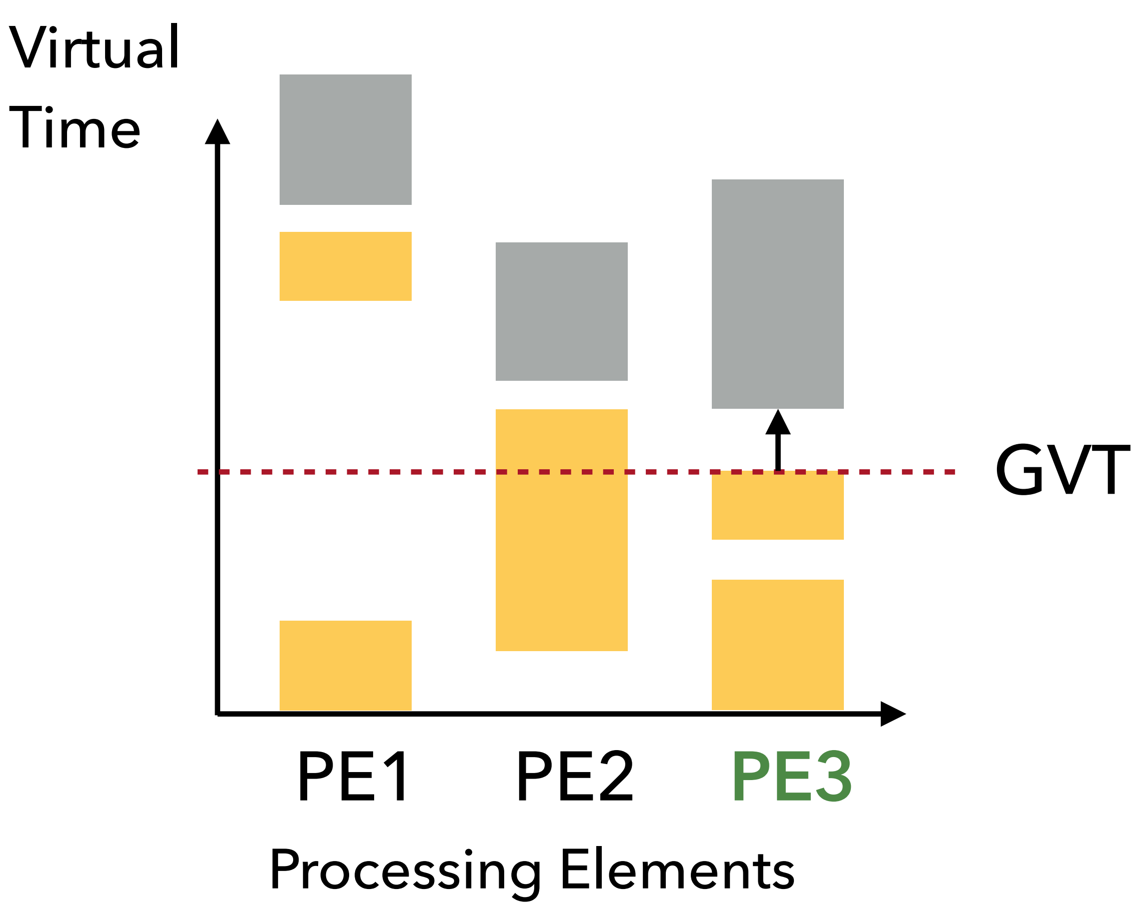


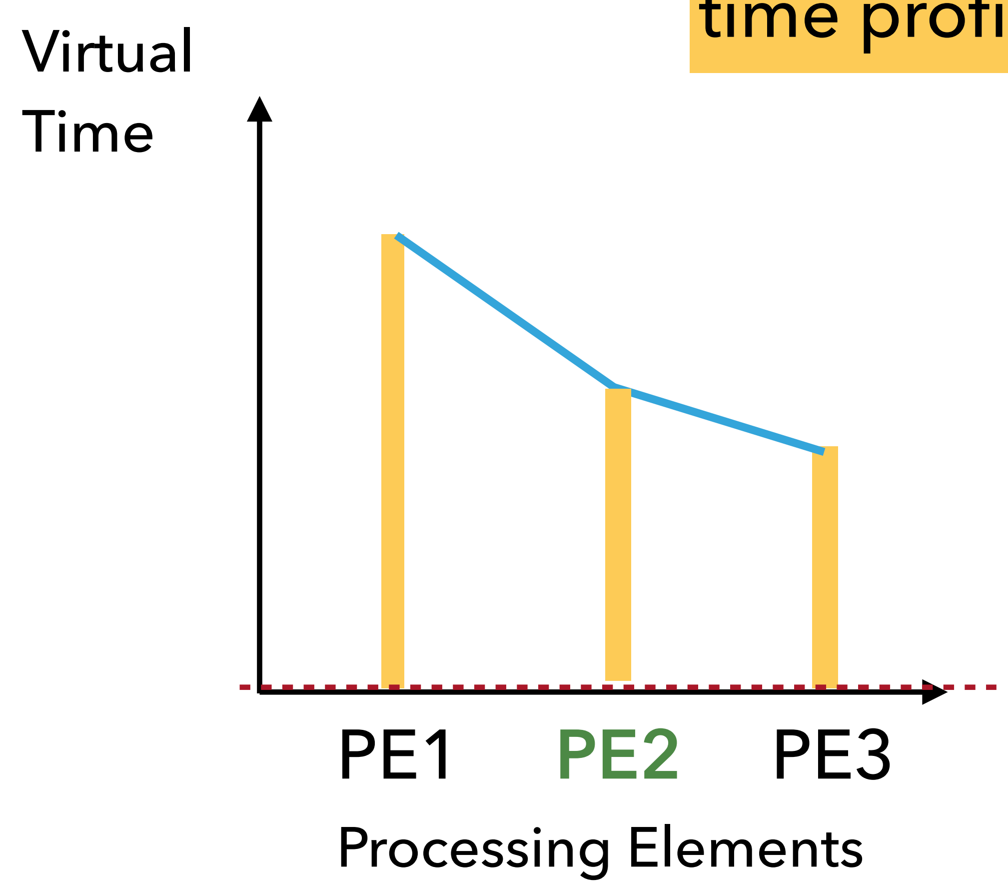


CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

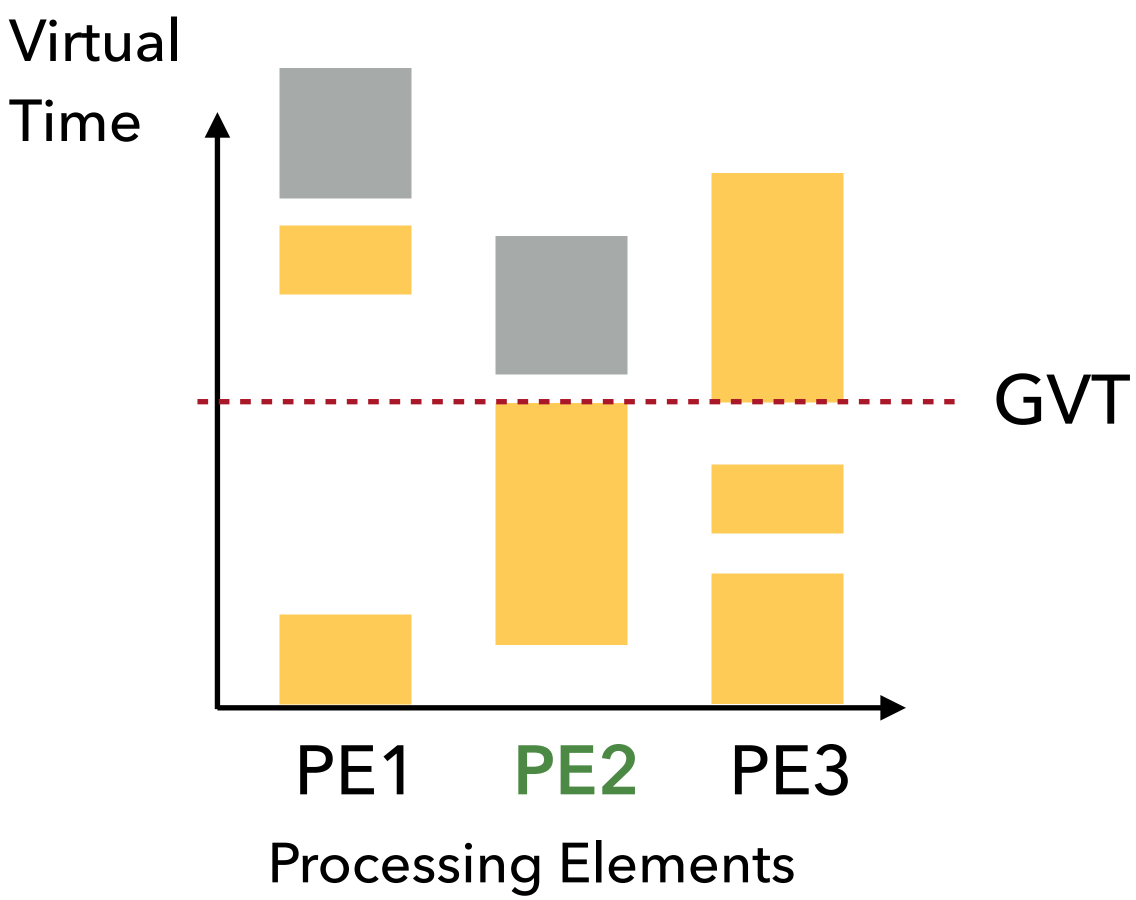
Evolution of time profile



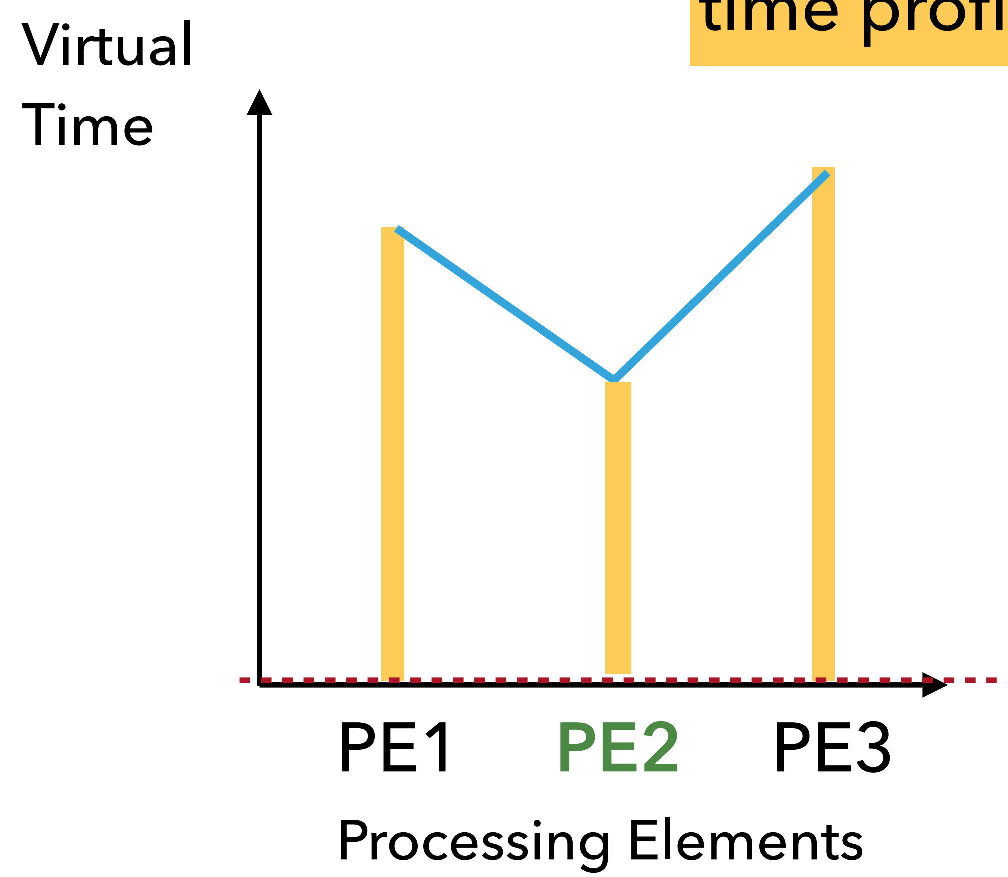


CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

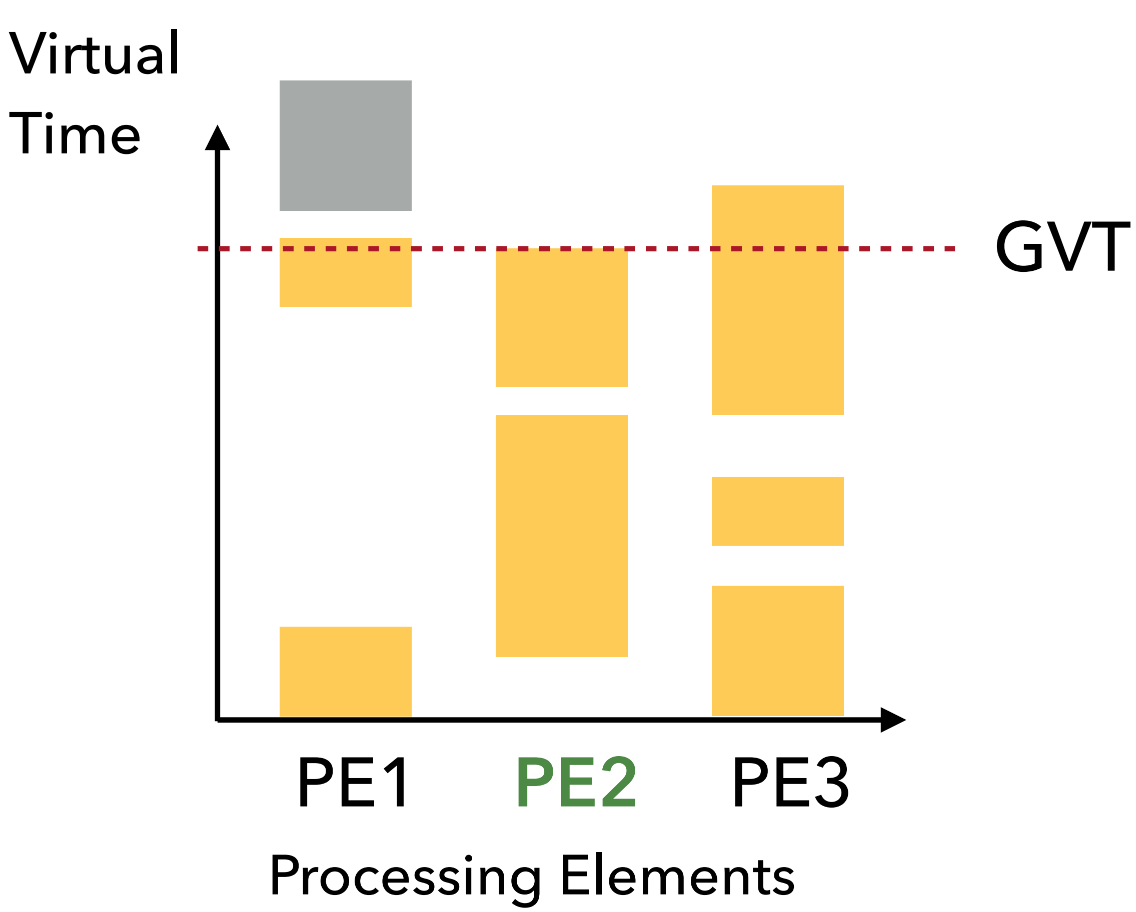


Evolution of time profile

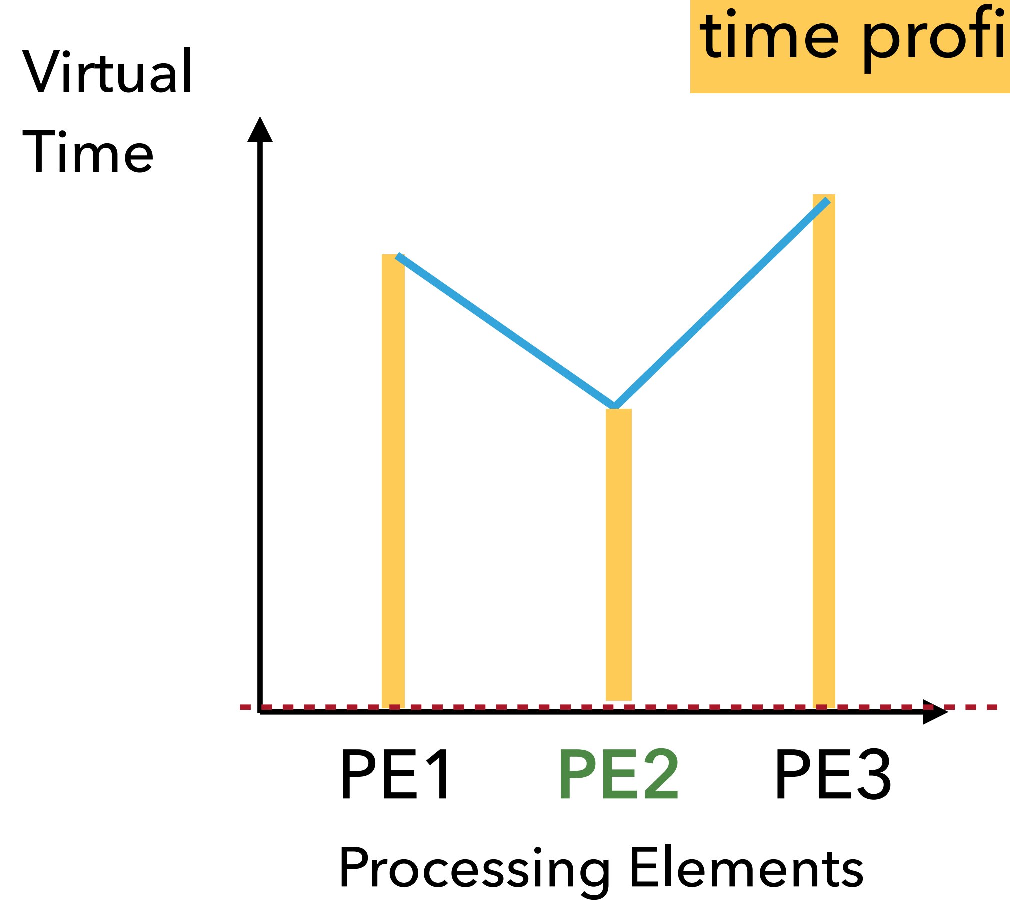


CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

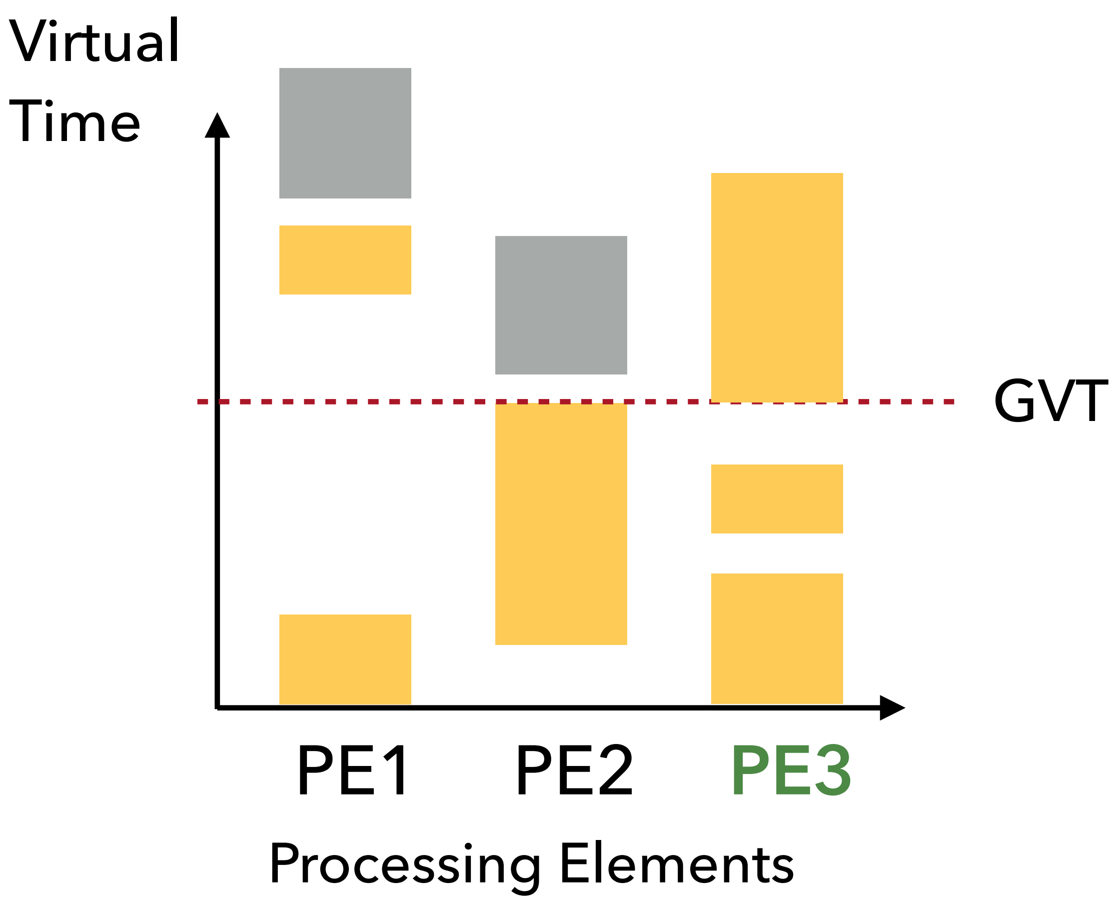


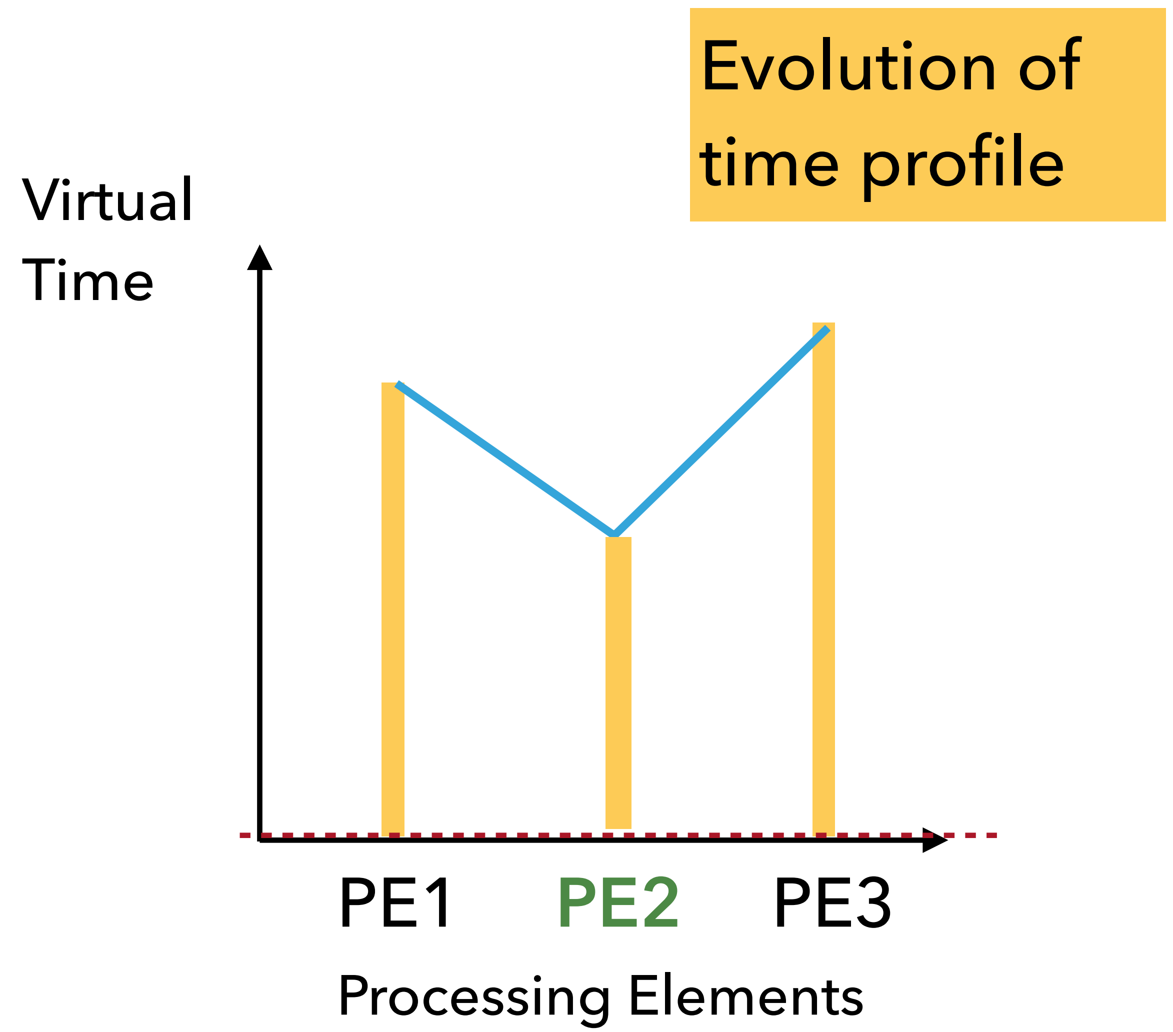
Evolution of time profile



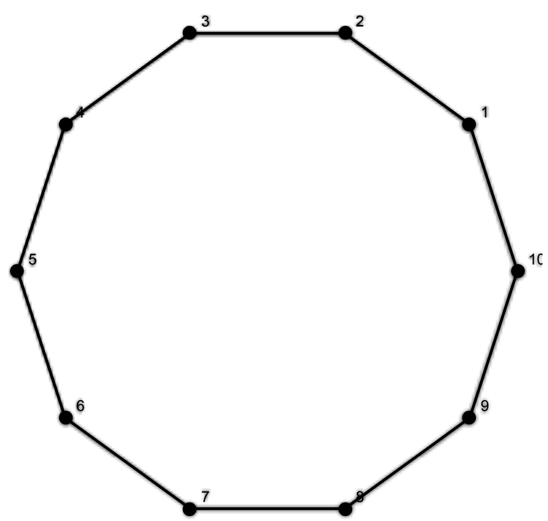
CONSERVATIVE SYNCHRONISATION ALGORITHM

Only those PEs, whose current time is lower than the time of their neighbours (i.e. the PEs which it is connected with), may proceed with computations. These PEs are called active. Such scheme guarantees that causality will be preserved.

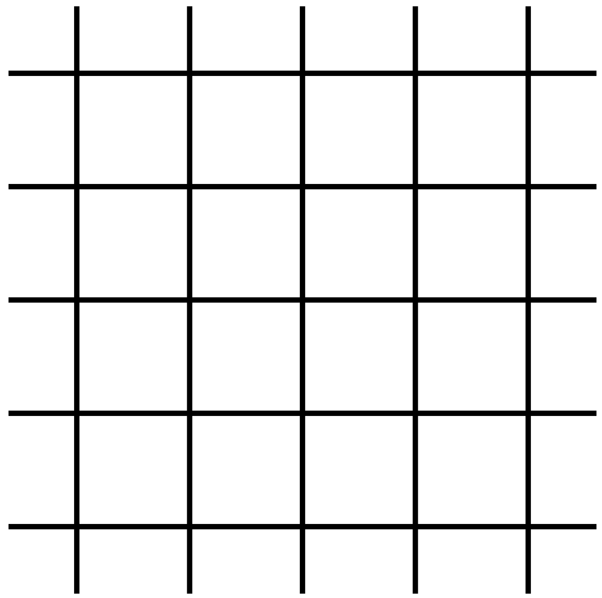




HOW LONG-RANGE LINKS AFFECTS SYNCHRONISATION?



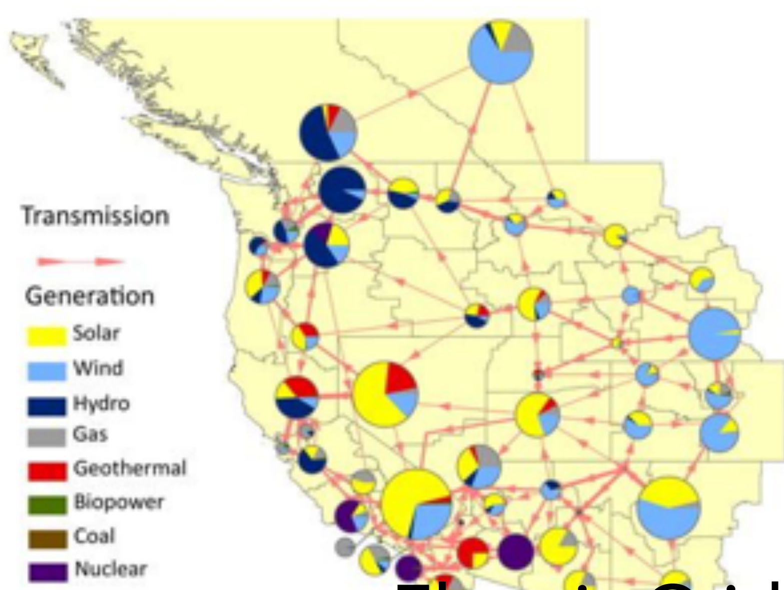
Regular ring lattice



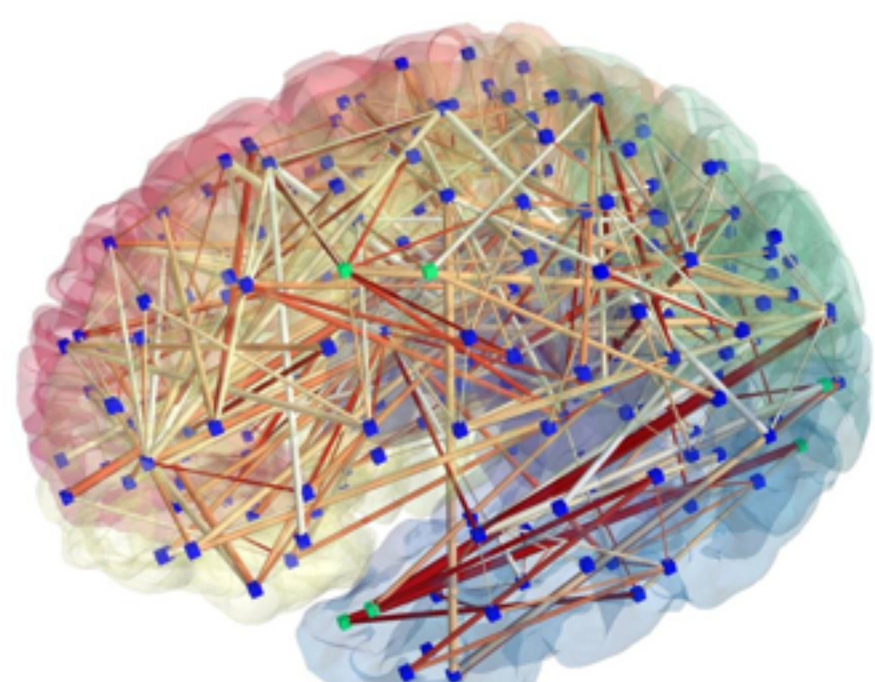
2d regular lattice

Social Networks

Vertices - PEs
Edges - communications



Electric Grid

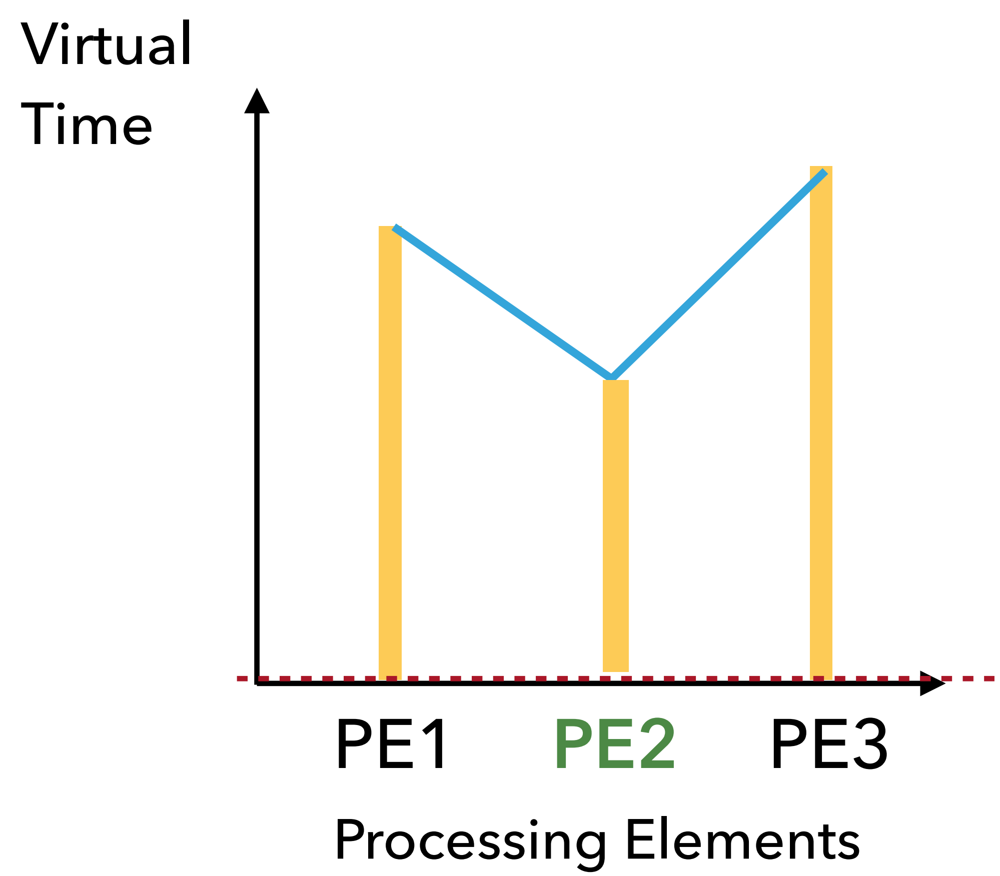


Brain neurons

THE OBJECT OF THE RESEARCH

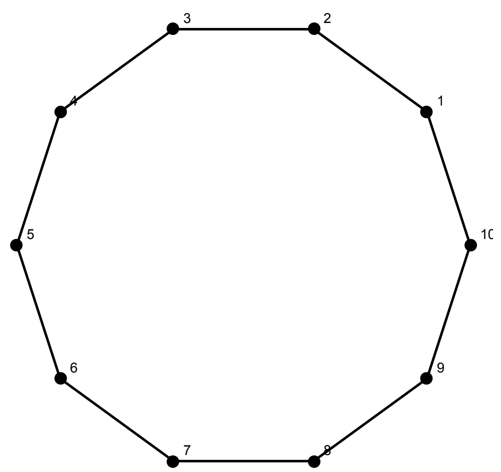
We study the scalability properties of the conservative synchronisation algorithm on small-world communicational network.

Evolution of LVT time profile

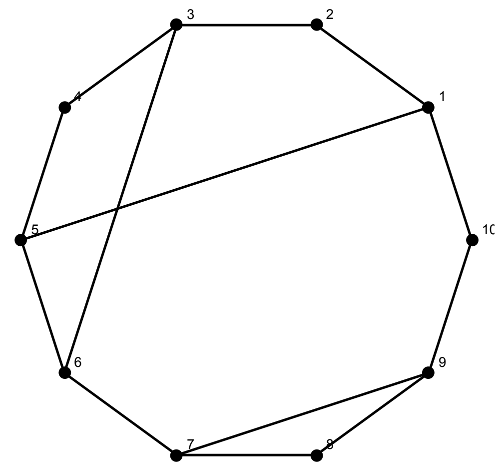


THREE TYPES OF NETWORKS

regular



Small-World-like[4]

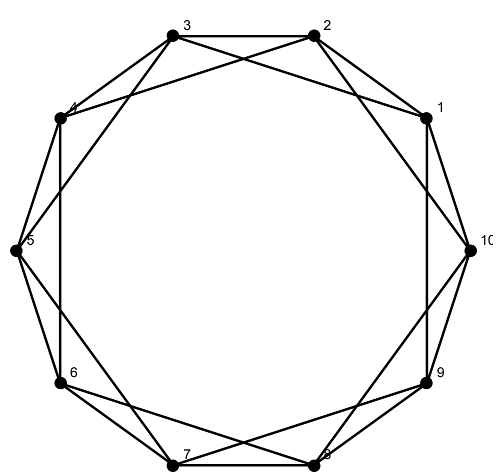


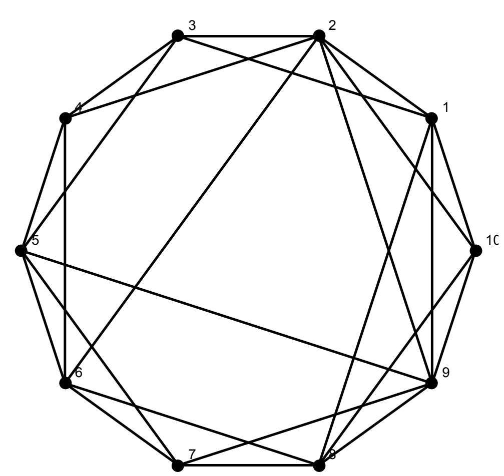
1. Average shortest path ~ log(N)
2. Clustering coefficient ≈ 0

p - concentration of added links

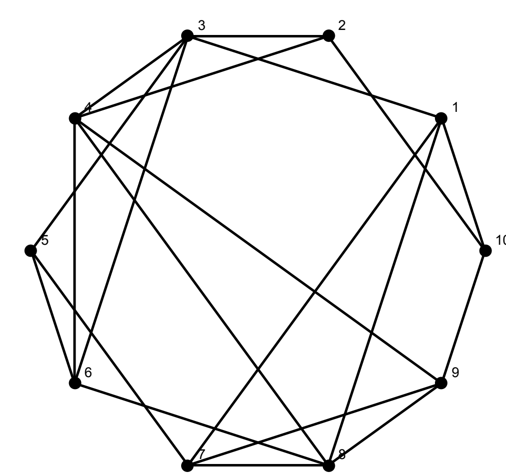
p=0

0 < p << 1





“hard” - links are added



“soft” - links are rewritten

1. Average shortest path ~ log(N)
2. High clustering coefficient

THE MODEL OF EVOLUTION OF THE LOCAL VIRTUAL TIME PROFILE

Initiate an array **t** of length ***N***, t[i]=0, i=0..N

For *k*=0..*M*, where *M* - number of time steps, do:

If (t[i]<=t[k]), where k is index of connected PE:

{t[i]+= a random value from Poisson distribution}

N_{active} +=1;

Calculate the average time

$$\overline{\tau}(t) = 1/N \sum_{i=1}^{N} \tau_i$$

Calculate the average width

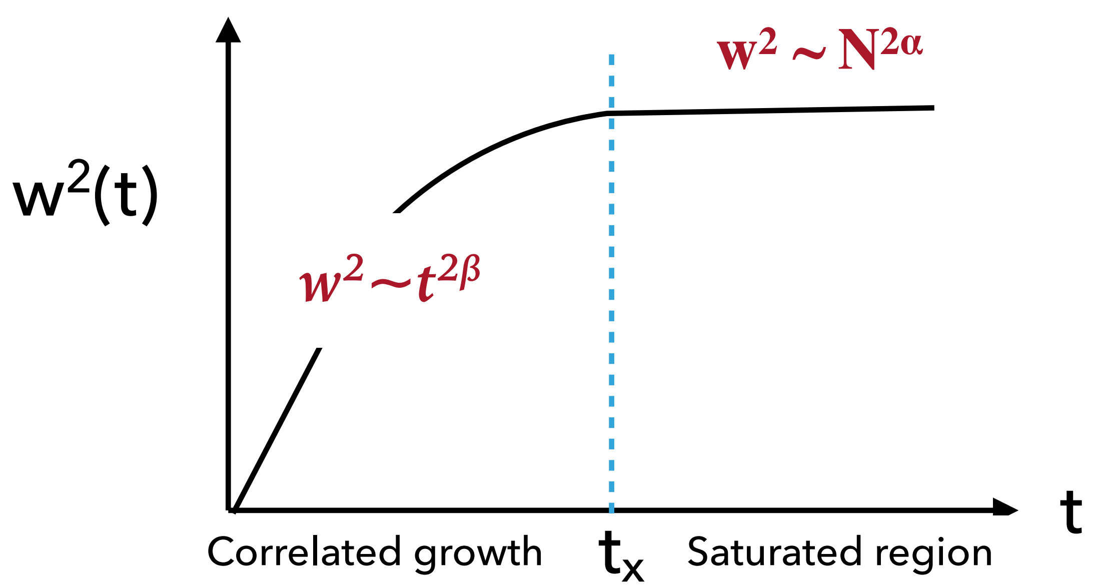
$$\langle w^2(t) \rangle = \langle \frac{1}{N} \sum_{i=1}^{N} [\tau_i(t) - \overline{\tau}(t)]^2 \rangle$$

Calculate the average utilisation

$$\langle u(t) \rangle = \frac{N_{active}}{N}$$

RESULT FOR 1-DIMENSIONAL CASE [5]

Kardar-Parisi-Zhang Universality class for growth models



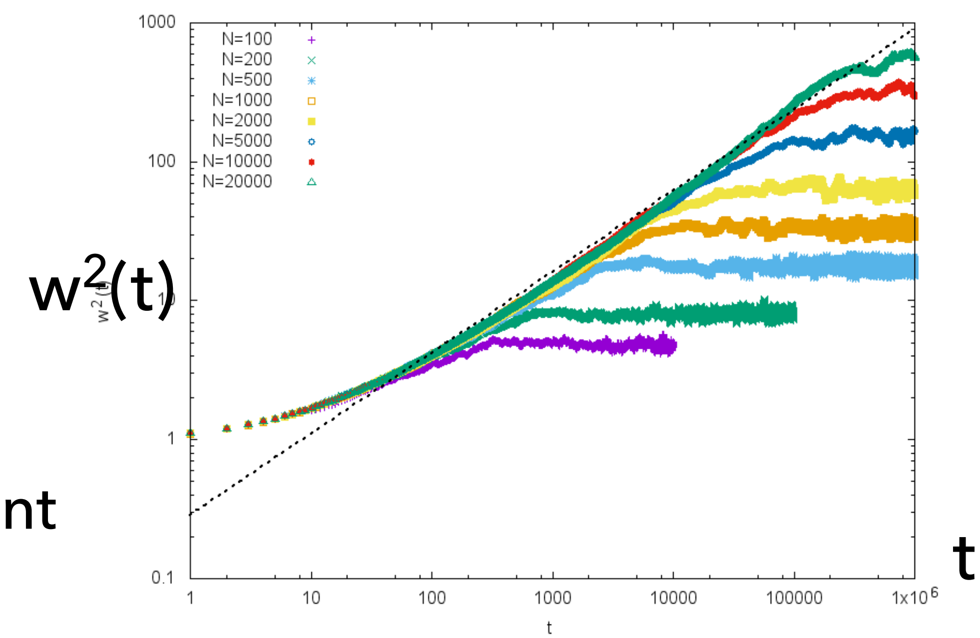
$$\alpha = 1/2$$

Growth exponent

$$\beta = 1/3$$

Roughness exponent

Model of the conservative algorithm on 1D topology



high of surface

$$\frac{\partial h}{\partial t} = \nu \nabla^2 h + \lambda (\nabla h)^2 + \eta,$$

surface tension

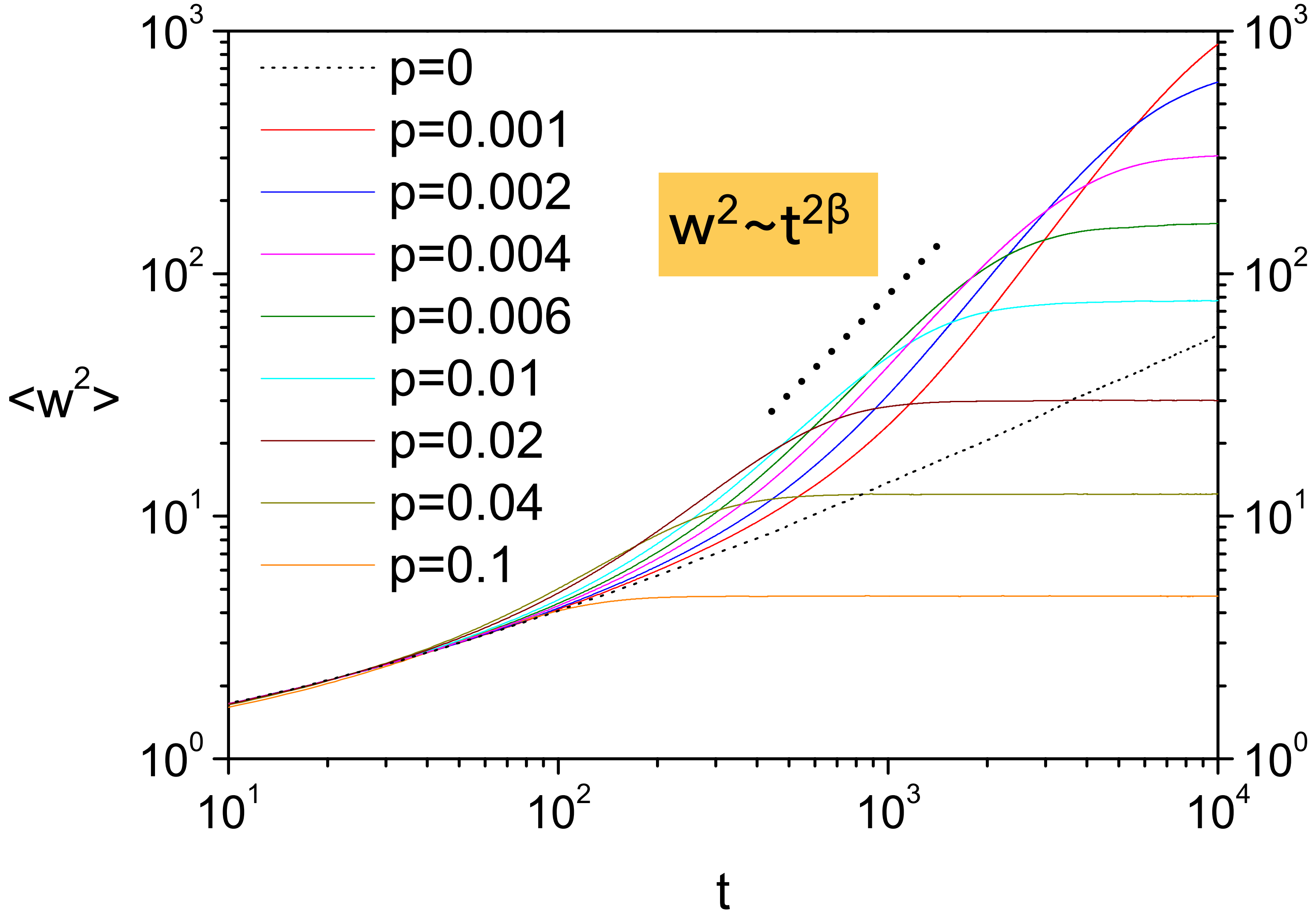
Gaussian noise

CONCLUSION FOR 1D:

Desynchronisation increases as the number of PEs goes to infinity

RESULT #1 GROWTH EXPONENT

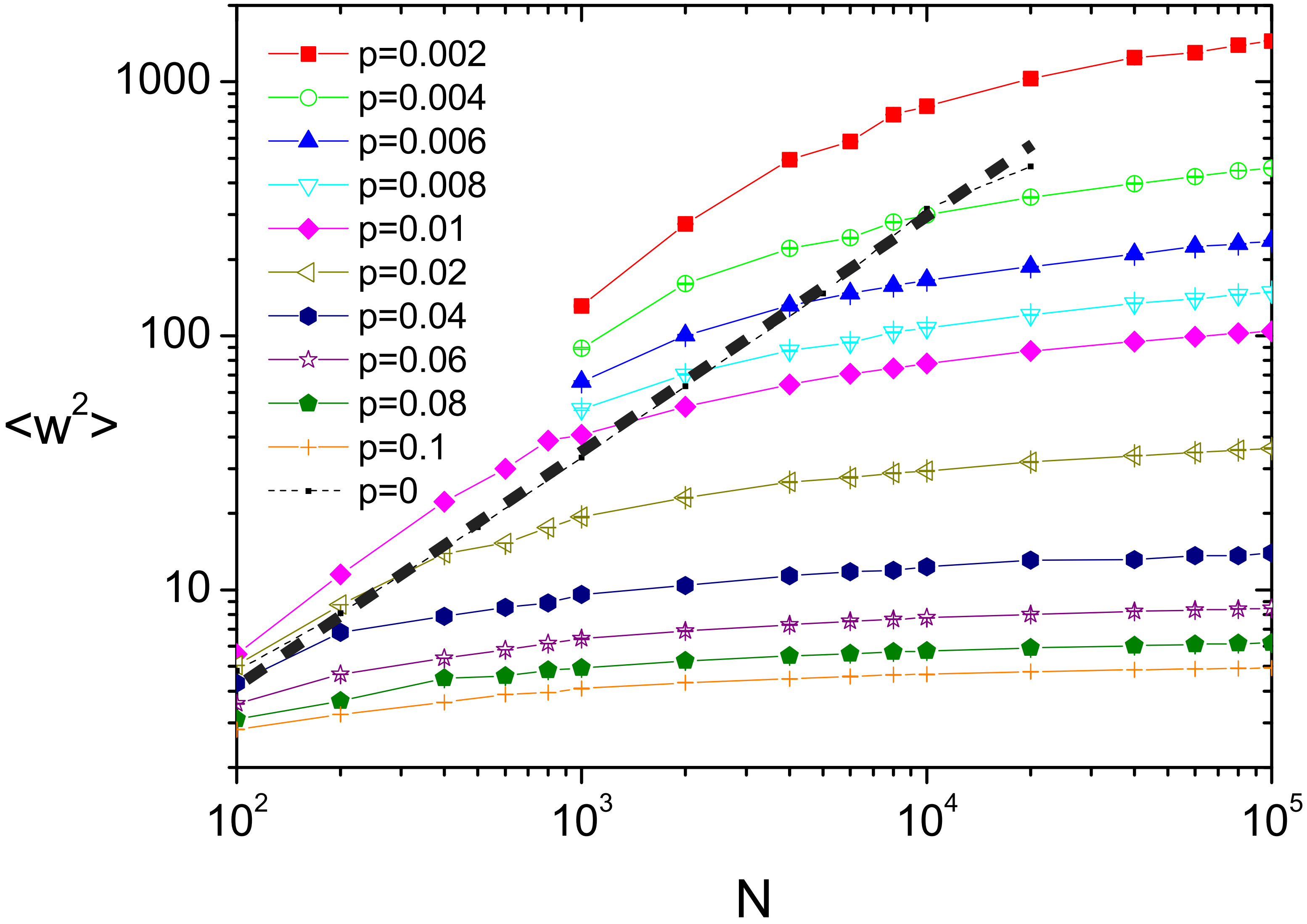
Growth exponent **β** logarithmically depends on the parameter ***p***



$$\beta \sim -0.162(2)\ln(p)$$

RESULT #2 ROUGHNESS EXPONENT

The width (i.e. desynchronisation) remains constant as the number of processes goes to infinity.



RESULT #3 AVERAGE SPEED

Average utilisation reduces slowly with ***p***

$$\langle u \rangle = \langle u_0 \rangle - 0.078(3) p^{0.092(3)}$$

$$\langle u_0 \rangle \approx 0.246$$

Analytical result from [1]

$$\langle u \rangle \simeq \frac{1}{4} + \frac{\sqrt{p}}{4\pi} + \mathrm{O}(p)$$

[1] H.Guclu, G.Korniss, Z.Toroczkay, and M.A.Novotny, Small-World Synchronized Computing Networks for Scalable Parallel Discrete-Event Simulations, Lext. Notes Phys. ***650****, 255-275 (2004)*

CONCLUSION

The underlying topology affects the scalability properties of the PDES algorithms. Adding a small fraction of long synchronisation links makes the conservative algorithm **fully scalable**, namely:

1) the average progress rate remains positive - **no deadlocks**,

2) the desynchronisation degree of the LVT profile becomes **finite**, when the number of PEs goes to infinity.

REMARKS

- Synchronisation in conservative PDES algorithm depends only on average shortest path and does not depend on clustering degree of a network
- Our model occupies intermediate position between the regular case and the mean-field case.

Regular lattice	Small-World	Fully connected (mean-field)
$\langle l \rangle \sim N$ $p = 0$	$\langle l \rangle \sim \log(N)$ $0 < p << 1$	$\langle l \rangle = 1$ $p \to \infty$

$\langle l \rangle$ - average shortest path

N - system size

p - concentration of long-range connections

"SCIENTIFIC DELTA"

OLD RESULTS

- 1D model of the conservative algorithm[5]
- Small-World model of the conservative algorithm[6]
- Case-studies on ROSS and WRAPED2 simulators with analysis of efficiency in terms of number of processed events per time[7]

NEW RESULTS

- 1D model of the optimistic algorithm[8]
- 3 different types of Small-World model of the conservative and optimistic algorithm[9]

FUTURE WORK

- Run test models on ROSS simulator and analyse the efficiency in terms of local virtual time evolution
- Probably, implement an efficient molecular dynamic algorithm using PDES method

REFERENCIES

1. *Fujimoto R.M.* Parallel Discrete Event Simulation // Commun. ACM. 1990. V. 33. P. 31–53.

2. *Jefferson D.R.* Virtual Time // ACM Trans. on Programming Languages and Systems (TOPLAS). 1985. V. 7. P. 404–425.

3. Shchur, Lev N., and M. A. Novotny. "Evolution of time horizons in parallel and grid simulations." Physical Review E 70.2 (2004): 026703.

4. Watts, Duncan J., and Steven H. Strogatz. "Collective dynamics of'small-world'networks." nature 393.6684 (1998): 440.

5. G. Korniss, Z. Toroczkai, M. Novotny, and P. A. Rikvold, Physical review letters 84, 1351 (2000).

6. Korniss G. et al. Suppressing roughness of virtual times in parallel discrete-event simulations //Science. – 2003. – Т. 299. – №. 5607. – С. 677-679.

7. Wilsey, P.A.: Some Properties of Events Executed in Discrete-Event Simulation Models. In: Proceedings of the 2016 annual ACM Conference on SIGSIM Principles of Advanced Discrete Simulation, pp. 165-176 (2016). ACM, New York (2016).

8. *Ziganurova L.F., Shchur L.N., Novotny M.A.* Model for the evolution of the time profile in optimistic parallel discrete event simulations // Journal of Physics: Conference Series **681** (2016) 012047

9. Ziganurova, Liliia & Shchur, Lev. (2017). Properties of the Conservative Parallel Discrete Event Simulation Algorithm. in Lecture Notes in Computer Science series, Vol. 10421, p. 246-253. doi: 10.1007/978-3-319-62932-2_23.