A Parallel Implementation for Cellular Potts Model with Software Transactional Memory

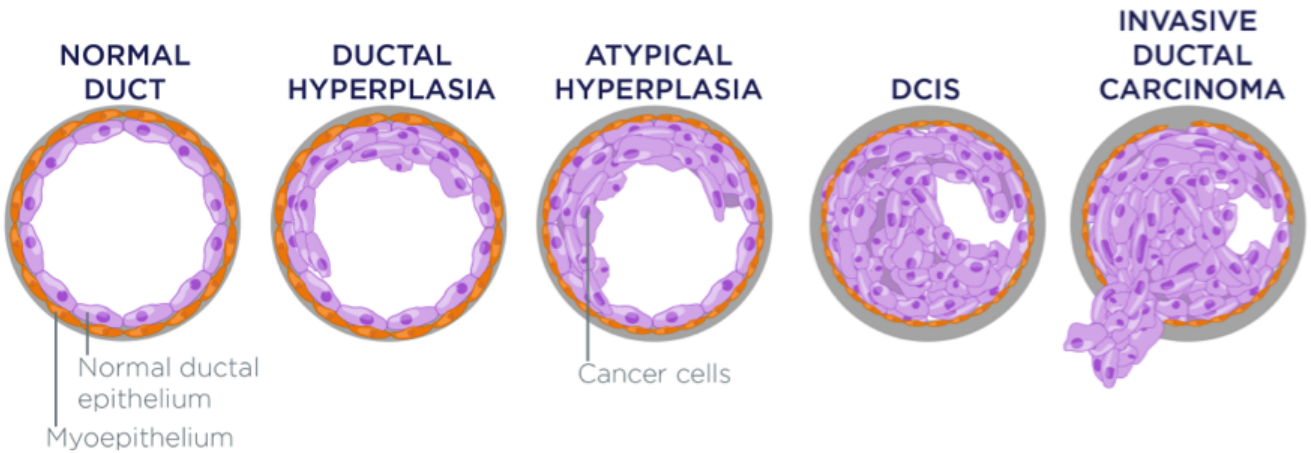
Antonio J. Tomeu[1] Alejandro Gámez[2]
Alberto G. Salguero[1]

[1]Department of Computer Science
University of Cádiz (SPAIN)
[2]IES Mar Mediterráneo
(SPAIN)

Introduction

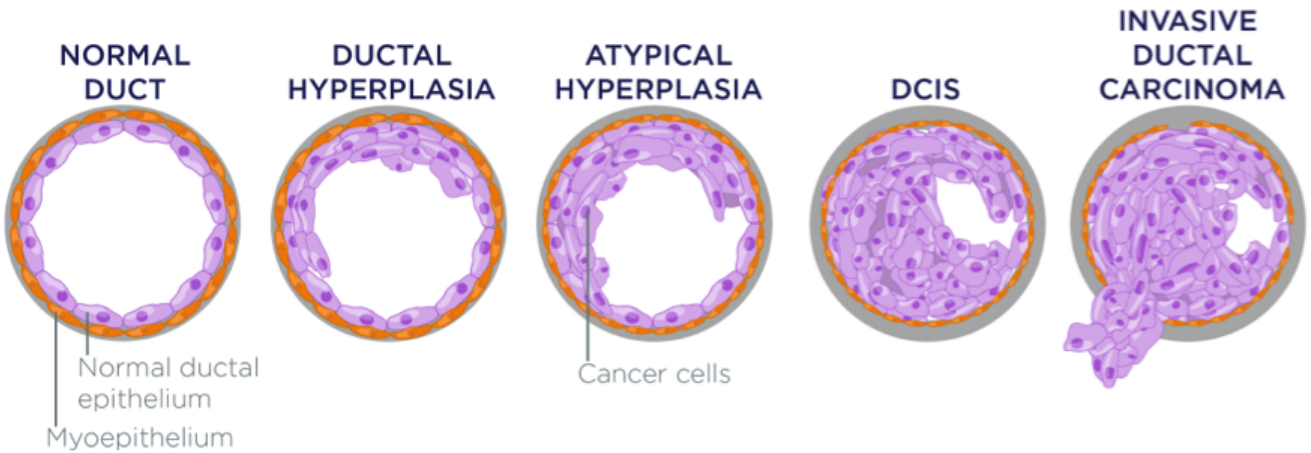
- It is estimated that one in eight women will suffer breast cancer, being approximately 80 % of them ductal carcinomas.
- Computer simulation can be an excellent tool to investigate it.
- Cellular Potts Model (CPM) can be applied to simulate biological systems, in a wide scale range.
- This work proposes a parallel implementation for CPM using Software Transactional Memory.
- Parallel model is applied to model breast cancer, a kind of Ductal Carcinoma In Situ (DCIS).

Biology of Adenocarcinomas in Situ: Natural History



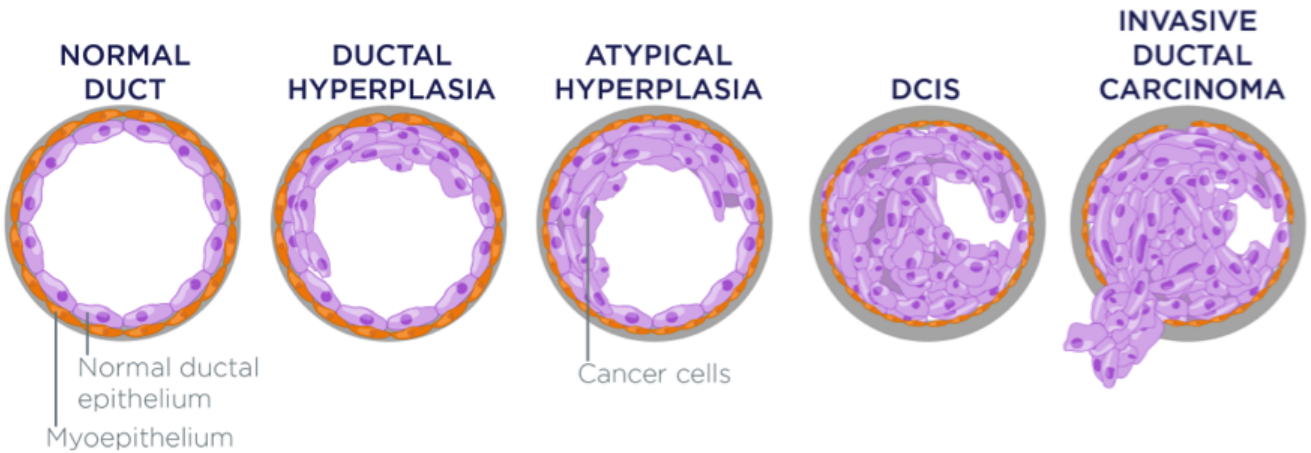
- Ducts are composed of two layers of cells: the innermost layer formed by luminal cells which is envolved by a second layer of myoepithelial cells, wrapped by a basement membrane.

Biology of Adenocarcinomas in Situ: Natural History



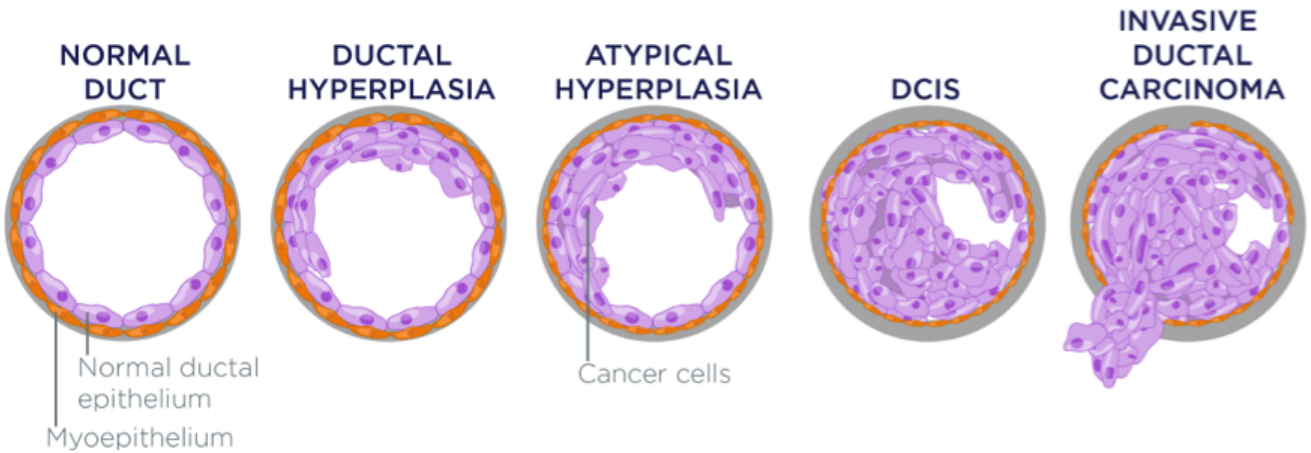
- Breast ductal adenocarcinomas begin with one (or several mutations) in the genomes of these cells.
- The luminal cells invade the light of the duct, breaking the normal double-layer architecture (hyperplasia).

Biology of Adenocarcinomas in Situ: Natural History



- Breast ductal adenocarcinomas begin with one (or several mutations) in the genomes of these cells.
- The luminal cells invade the light of the duct, breaking the normal double-layer architecture (hyperplasia)
- Then go by filling out the light of the duct (DCIS).

Biology of Adenocarcinomas in Situ: Natural History



- And finally breaking the basement membrane and invading the glandular parenchyma.
- At this point the disease acquires an infiltrative characteristic, appears the possibility of metastatic processes and treatment costs rise considerably.

Cellular Potts Model

- This model works with a 2D grid ζ of nodes with null borderline conditions.
- Each node of have coordinates (x, y), and a symbol $k \in \Sigma$.
- A cell is a subset $S = \{(i, j, k) : k \textit{ is constant}\}$ of ζ.

5	5	5	5	5	2	6	6
5	5	5	5	2	2	2	6
7	5	1	1	2	2	2	6
7	1	1	1	2	2	2	3
7	7	1	1	1	8	3	3
7	7	1	1	8	8	3	3
7	7	7	9	8	8	8	3
9	9	9	9	8	8	8	3

Cellular Potts Model

- One node (belonging to a cell or to an empty space) is randomly selected in each time step.
- This cell tries to change its location, size, etc.

Cellular Potts Model

- ~~One node (belonging to a cell or to an empty space) is randomly selected in each time step.~~
- A target node is randomly selected in each time step.
- ~~This cell tries to change its location, size, etc.~~
- Surrounding cells (neighbors) tries to occupy that node.

Cellular Potts Model: Probability For Change

$$P\left(\alpha\left(i,j,k\right)\rightarrow\alpha\left(i',j',k'\right)\right)=\begin{cases} e^{-\frac{\Delta H}{T_m}} & if\ \Delta H>0 \\ 1 & if\ \Delta H\leq 0 \end{cases}$$

where

- the triplet (i,j,k) specifies a node of the grid..
- and T_m is the temperature parameter.
- Change is accepted or rejected based on an energy function called the hamiltonian H.

Cellular Potts Model: The Hamiltonian H

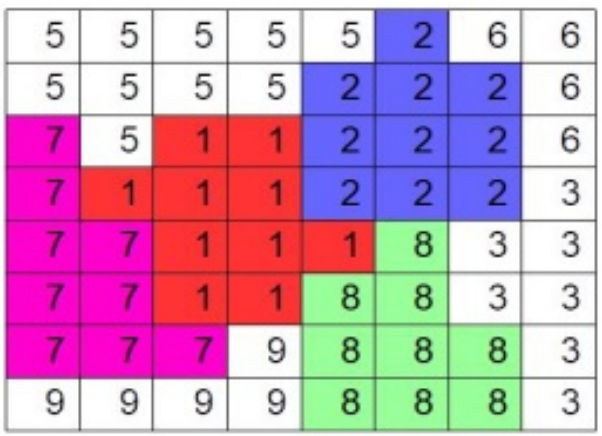
$$H = \sum_{(i,j,k),(i',j',k')} J_{\tau(\alpha(i,j,k)),\tau(\alpha(i',j',k'))} \left(1 - \delta_{\alpha(i,j,k),\alpha(i',j',k')}\right)$$
$$+ \sum_{\alpha} \lambda_V(\tau)(V(\alpha) - V_t(\alpha))^2$$
$$+ \sum_{\alpha} \lambda_S(\tau)(S(\alpha) - S_t(\alpha))^2$$

where

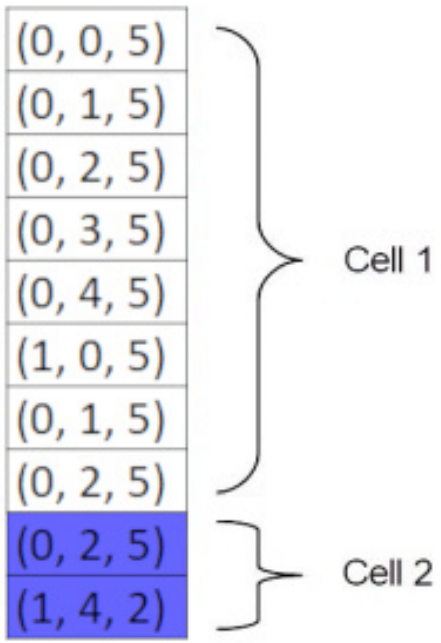
- τ represents the type of agent (luminal, extracellular or myoepithelial cells).
- The first term describes the energy of adhesion between a cell and its neighbors.
- The second term defines the volume and the degree of compressibility of the cell.
- The third term models the elasticity of the cell.

Cellular Potts Model: Data Structures

- A 2D array to simulate the grid ζ.
- A list Ξ of nodes thas have been processed.
- The list Ξ wil be processed under transactions.
- Parallel task choose a node, evaluate H and apply δ.
- Data structures are accessed within transactions.



2D-array ζ (AtomicInteger)

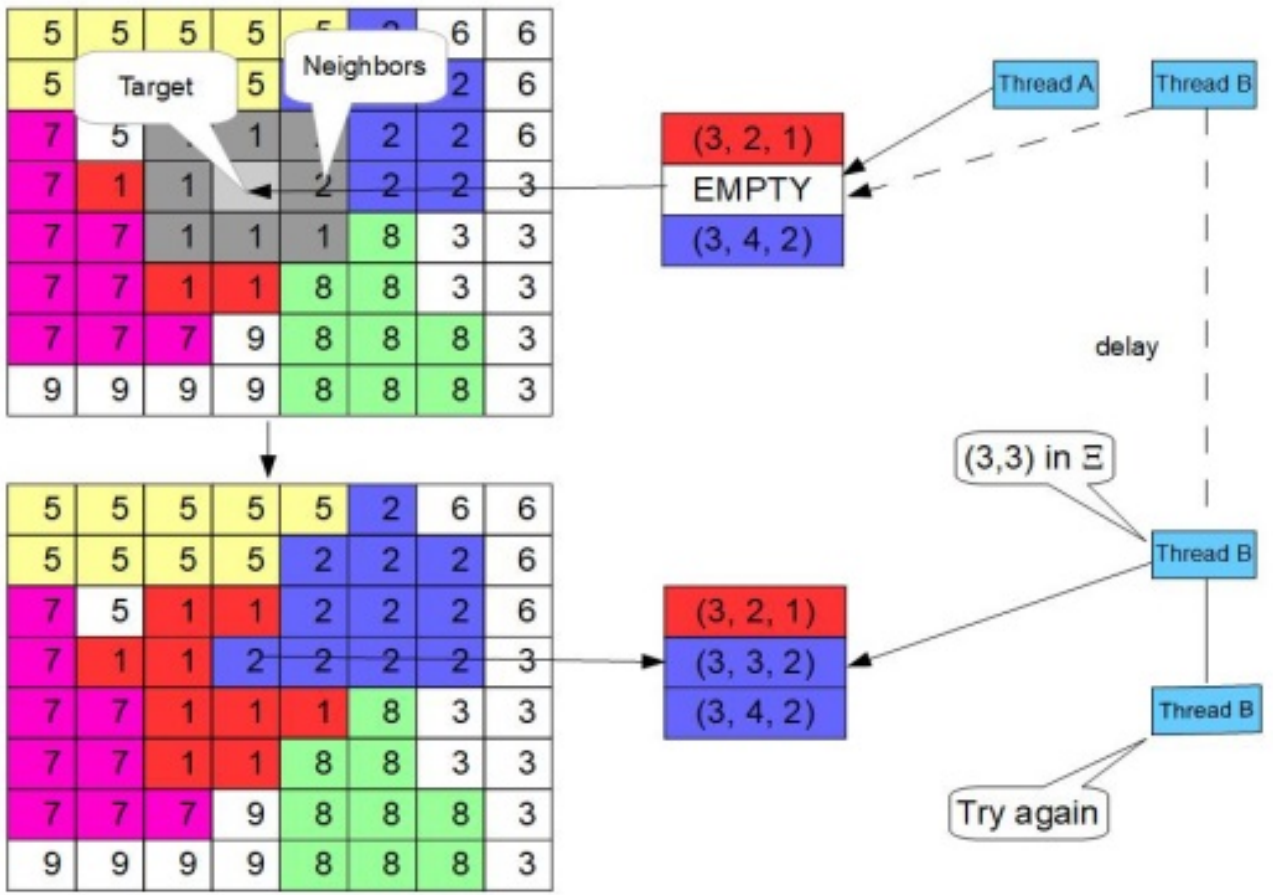


Ξ Autosynchronized List

Cellular Potts Model: Algorithm I

```
1  Algorithm Evolve(zeta, ji){
2  for(i=1, i<niterations/ntasks, i++){
3        x=random(xmax);
4        y=random(ymax):
5        runInTransaction(){if !((x,y) in ji){
6         cell= zeta[x][y];
7         ji.add((x,y,k));}
8         else goto 1.1
9        }
10   neighbourX = getRandom(rangXMin,rangXMax);
11   neighbourY = getRandom(rangYMin,rangYMax);
12   cellNeighbour = zeta [neighbourX][neighbourY];
13   J=getEnergyAdhesionForNeighbors(cell);
14   V=lambdaV*getVolumeForCell(cell);
15   S=lambdaS*getSurfaceCell(cell);
16   deltaH=J+V+S;
17   if(deltaH>0)
18         P=Math.exp(-deltaH/params.getTemperature());
19    else if(deltaH<=0)P=1;
20   if(p>=random()) zeta[x][y]=cellNeighbour();
21 }}
```

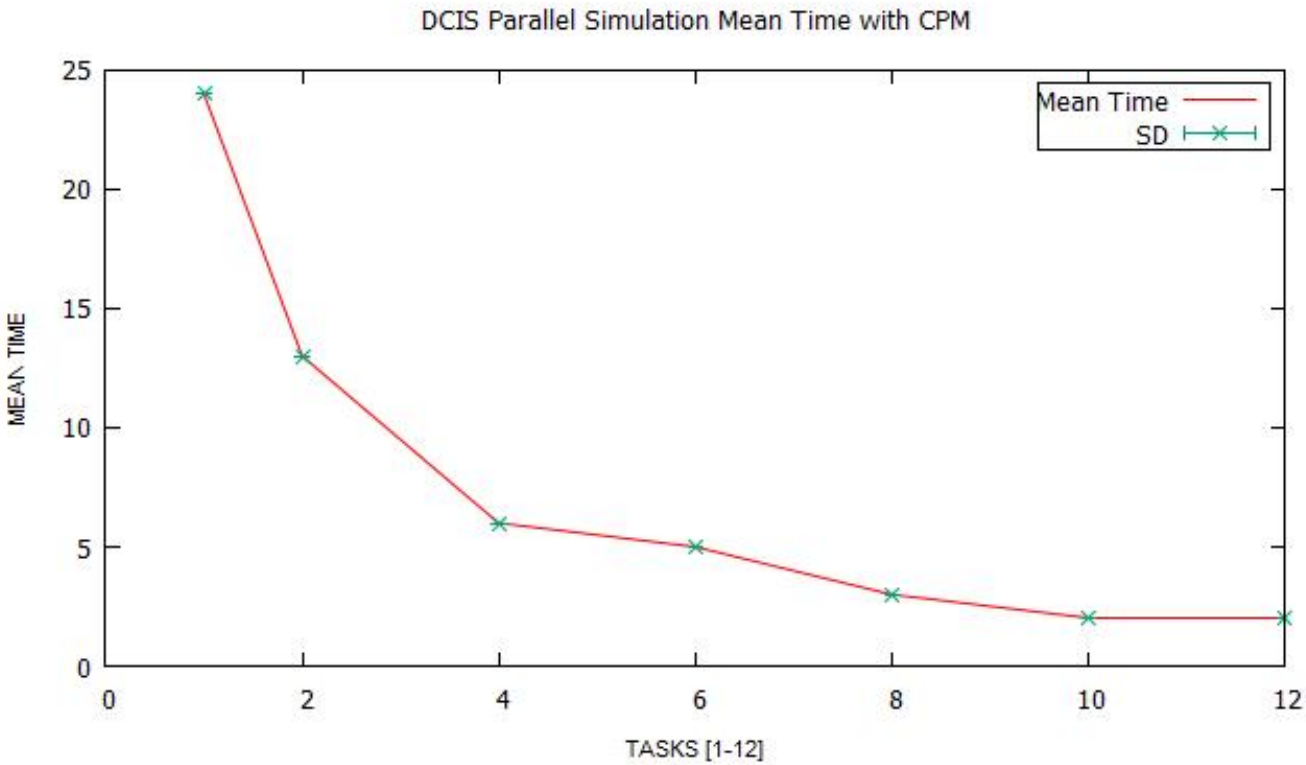
Cellular Potts Model: How It Works...



The Experiment

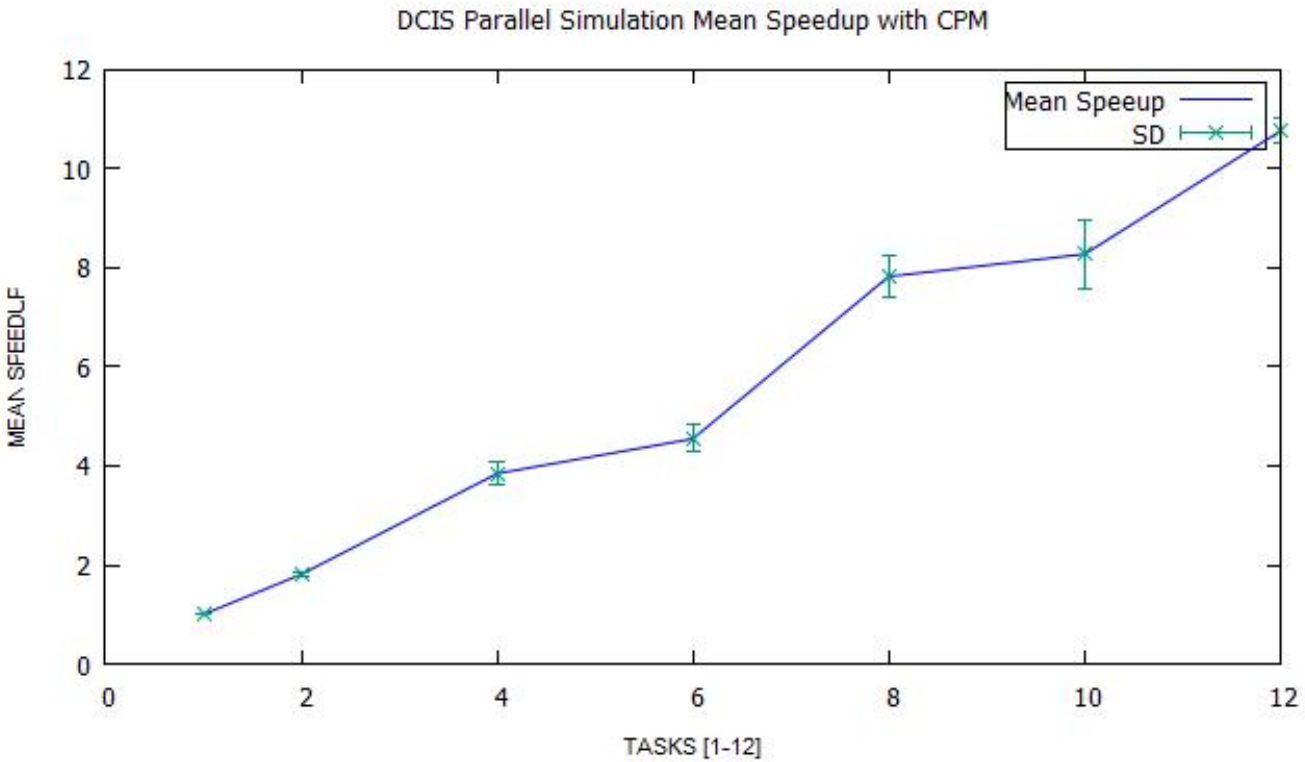
- Implementation with Java and Clojure (for STM).
- From a normal duct to a DCIS duct.
- Simulation carried out on 8 nodes of our cluster of processors
- For each node: two Intel Xeon E5 processors (16 cores), 2.6 GHZ and 128GB RAM yielding 20.8 GFLOPS.
- Grid of 900×900 and 10^9 iterations.
- Number of Parallel Task: 2 to 16.
- Mean execution times and speedups are calculated.

The Experiment: Execution Times



- Sequential: 24.13 seconds.
- Locks-based technique: 6.32 seconds (12 tasks).
- STM technique: 2.28 seconds (12 tasks).

The Experiment: Speedups



- Maximum speedup: 10.76 (12 threads).
- No improvement for 16 threads.

Conclussions and Future Work

- The work proposes a general procedure to do parallel simulations for CPM model.
- We apply the procedure to Breast Adenocarcinoma in situ (DCIS).
- We protect data structures within transactions, and divide the work between parallel task.
- We obtain a maximun speedup for 12 parallel tasks.
- We check that subsequent increases in the number of tasks do not offer performance improvements.
- Good scalability for parallel implementations.
- We will focus our future work in:
  - the application of the model to other glandular neoplams in situ.
  - the development of a data partition scheme for GPU architectures.

Thank you for your attention...
Questions?