

Genome Simulator

Umang Patel
ujp2001

MS (Machine Learning) , Columbia University

Personal Website : <http://umangpatel17.wordpress.com/>

Binary File, Manual, Code : <https://github.com/17patelumang/GenomeSimulator.git>

What is Genome Simulator ?

- A program to simulate sequences drawn from population .

Old Simulator

- Start with samples and move back ward in time to simulate recombination and migration event .

Syntax/Semantic of new Simulator

- Key words :-
- gen
- node
- size
- edge
- rate
- migration

Sample Text File

```
admixtureGeneration = gen(15,2); |  
  
nodeEuToAA = node(admixtureGeneration);  
nodeAfToAA = node(admixtureGeneration);  
nodeAA = node(admixtureGeneration);  
  
sizeEuAtAdmixture = size(100000,1);  
sizeAfAtAdmixture = size(70000,1);  
ancestralSizeAA = size(10000,2);  
  
EuropeansPostAdmixture = edge(node(G0), nodeEuToAA, size(1000000), size(98000,1));  
AfricansPostAdmixture = edge(node(G0), nodeAfToAA, size(700000,1), size(62000,1));  
AfricanAmericans = edge(node(G0), nodeAA, size(1000000,2), ancestralSizeAA);  
migrationBranchEU = edge(nodeAA, nodeEuToAA, size(2000,2));  
migrationBranchAf = edge(nodeAA, nodeAfToAA, size(8000,2));  
  
nodeSplitAfEU = node(gen(400,1));  
ancientBranchEU = edge(nodeEuToAA, nodeSplitAfEU, sizeEuAtAdmixture, size(50000,1));  
ancientBranchAf = edge(nodeAfToAA, nodeSplitAfEU, sizeAfAtAdmixture, size(50000,1));  
  
nodeAncestralAll = node(gen(600));  
ancestralSize = size(10000);  
edge(nodeSplitAfEU, nodeAncestralAll, size(100000,1), ancestralSize);  
edge(nodeAncestralAll, node(Ginf), ancestralSize);  
  
rateAf-AA=rate(0.001,1);  
  
migration(EuropeansPostAdmixture, AfricansPostAdmixture, rateAf-AA );
```

gen

- **generation_variable=**
gen(generation_value,offset_parameter)
- Generation_value – value of generation
- Offset_parameter- 1 or 2
- Eg:- admixtureGeneration = gen(15,2) ;
- admixtureGeneration is the name of generation variable
- 15 is the value of generation , offset 2

node

- **Node_name= node(generation_variable)**
- Each node is population in new model
- Eg:- AfricanPopulation =
node(admixtureGeneration);
- AfricanPopulation – name of node/population
- admixtureGeneration – name of generation varialbe

size

- **size_variable= size(size_value);**
- Used to specify size of population .
- Eg:- size_of_Arficanpopulation= size(100000,1);

edge

- Edge_name= edge(start_node, end_node, start_node_size ,end_node,size);
- Used to specify how one population interacts with other .
- Eg:- africanpopulation_american_population= edge(AfricanPopulation, AmericanPopulation, size(1000000), size(98000,1));

rate

- Rate_name=rate(rate_number)
- Do not give rate name as R1 or R2 (pre defined internally)
- Rate_number – 0.1·1
- Eg:- rateAfrican-American=rate(0.001,1);

migration

- migration(starting edge,ending edge,rate variable);
- Migration always occurs between edges.
- Eg:- migration(EuropeansPostAdmixture, AfricansPostAdmixture, rateAf-AA);

Need to remember syntax ?

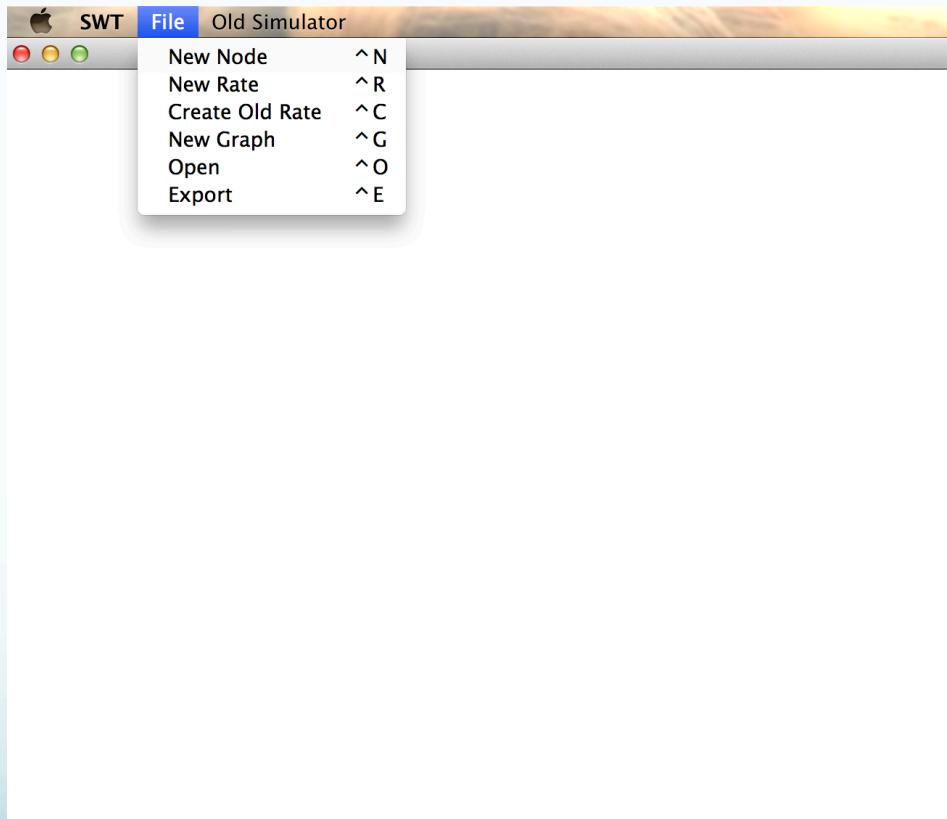
- No , need . GUI !

Feature

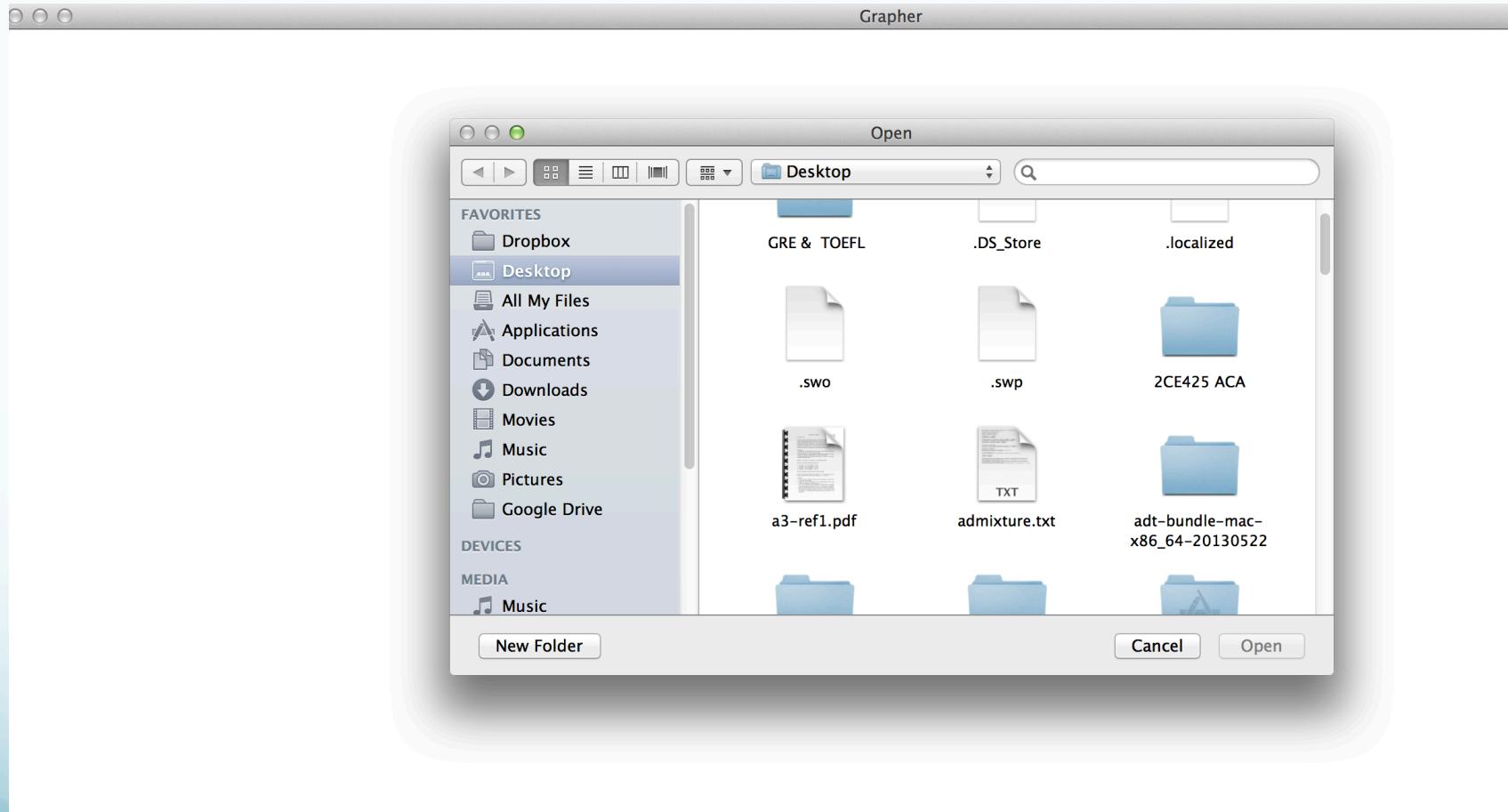
- Read file and produce graph
- Create and graph and export it to syntax of simulator
- Add node
- Add edge between any node to any other node
- Migration add/show from GUI
- Output in old Genome Simulator format made by University of Michigan

Read file and produce graph

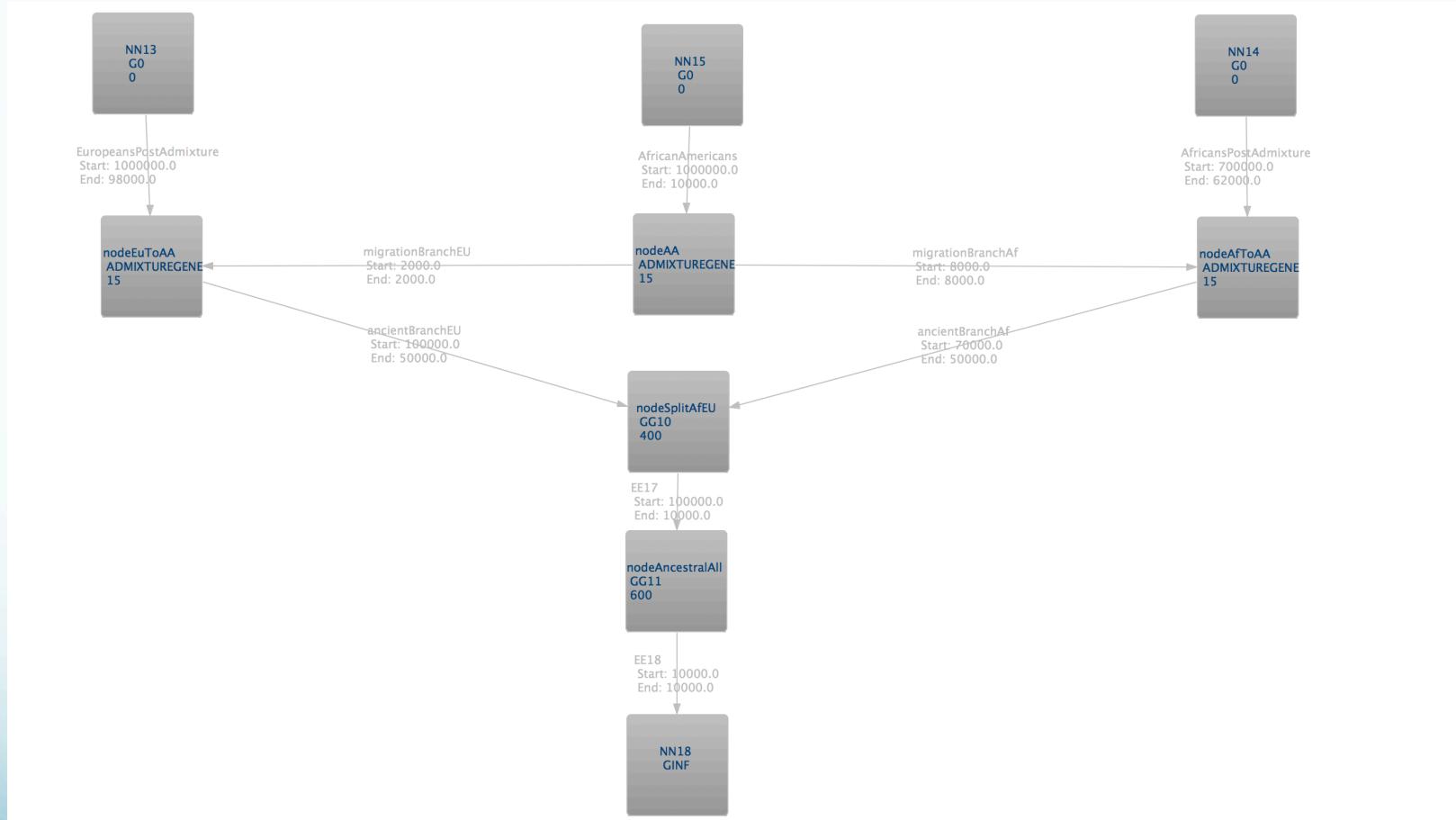
- File -> Open



Read file and produce graph

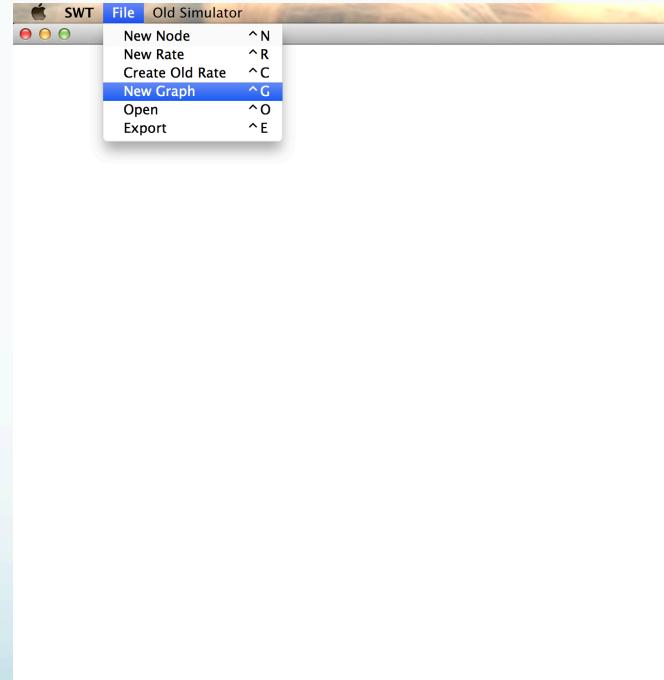


Read file and produce graph

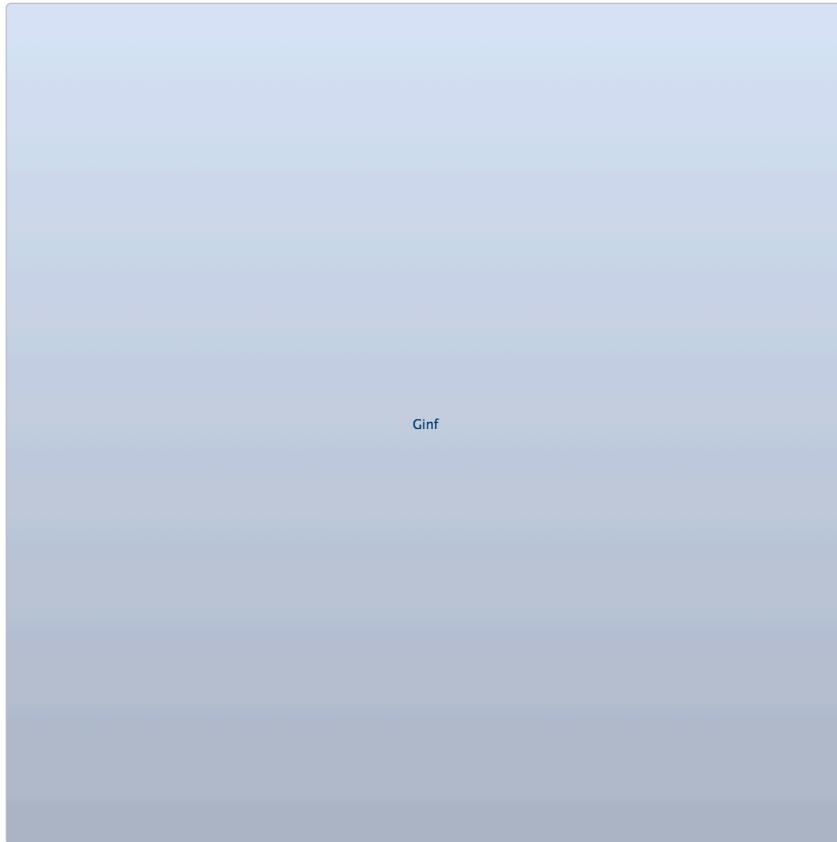


Create and graph and export it to syntax of simulator

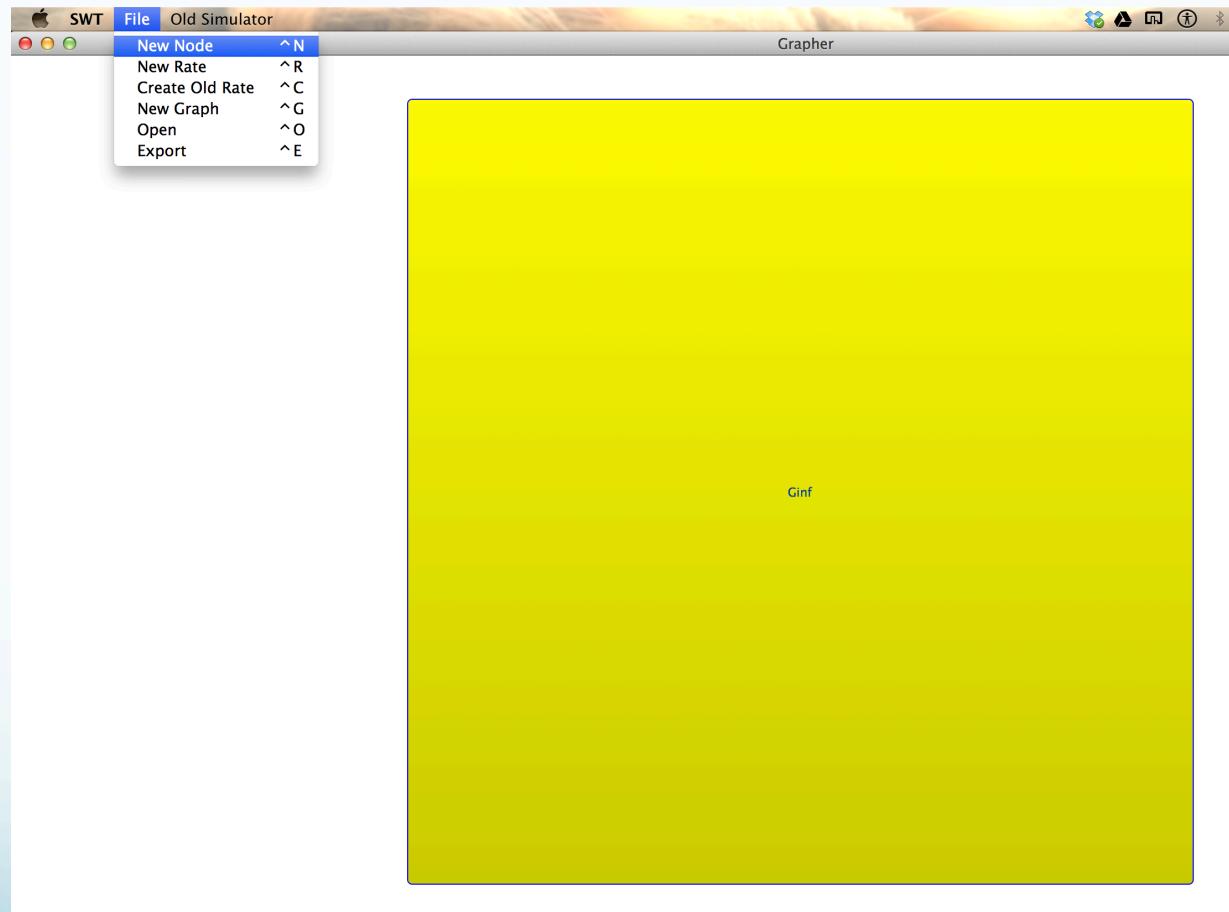
- File-> New Graph -> (Draw as you want) -> File -> Export



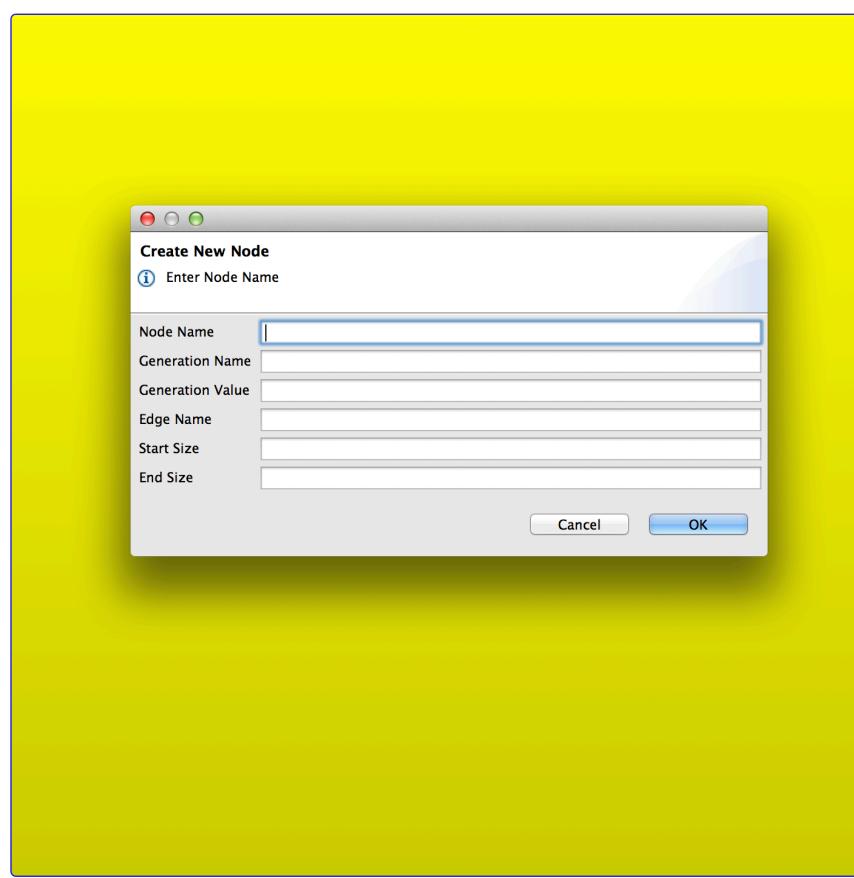
Create and graph and export it to syntax of simulator



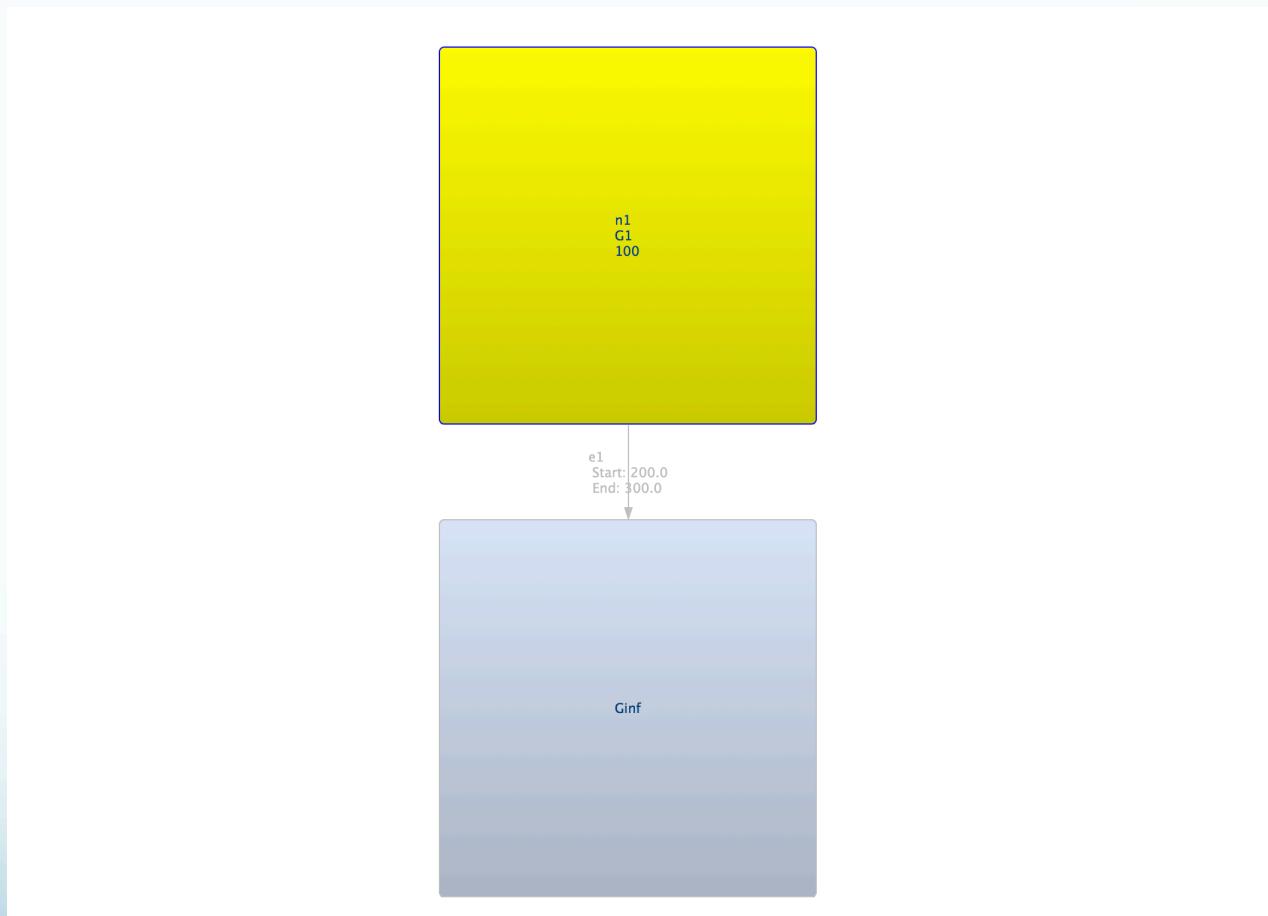
Create and graph and export it to syntax of simulator



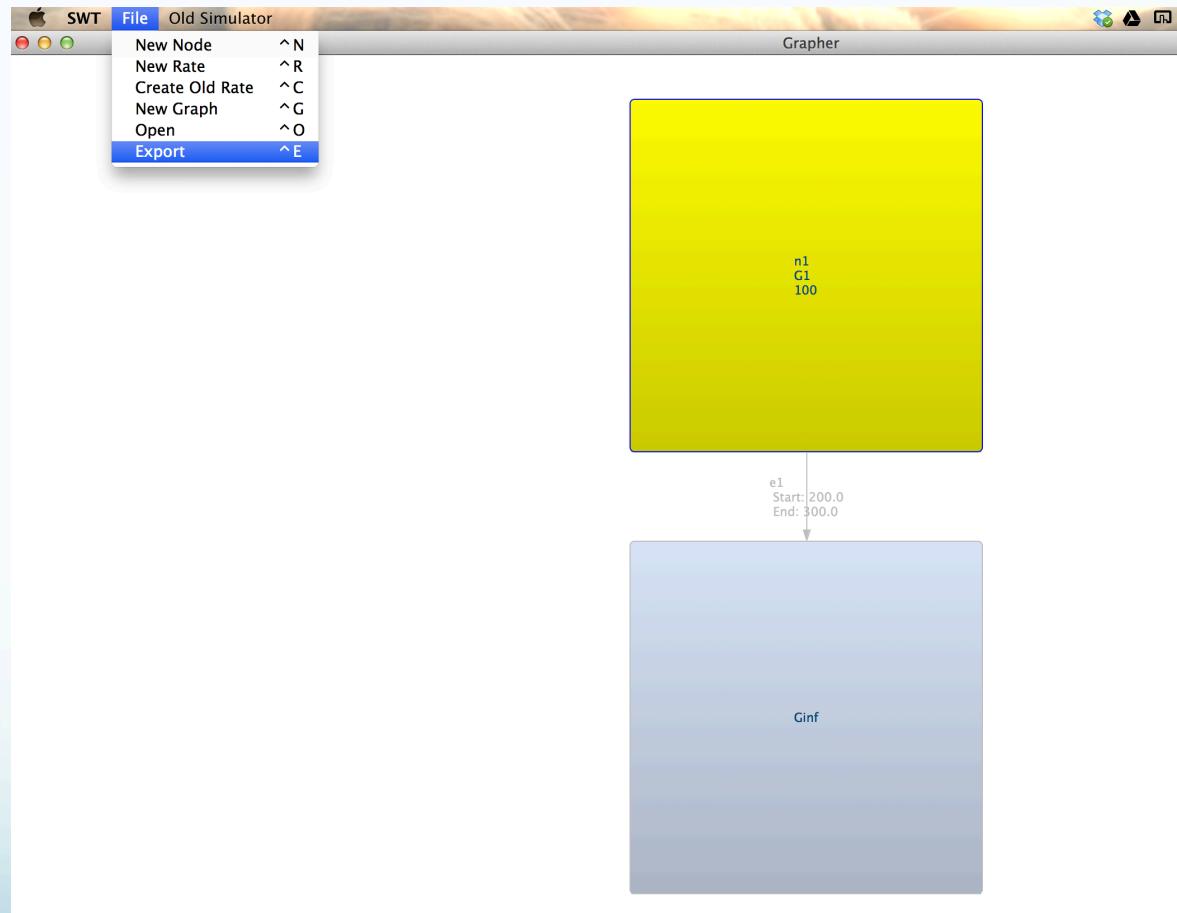
Create and graph and export it to syntax of simulator



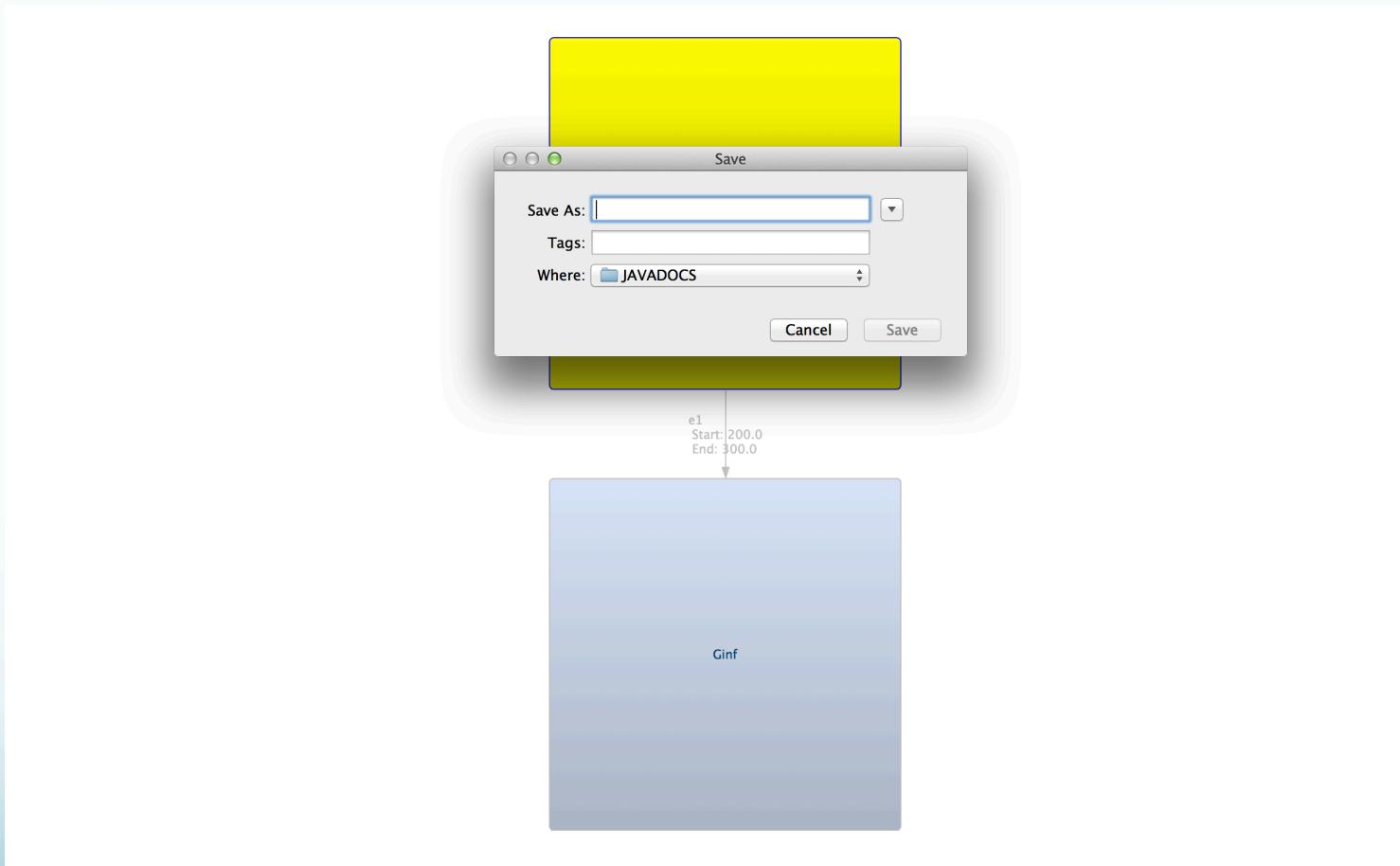
Create and graph and export it to syntax of simulator



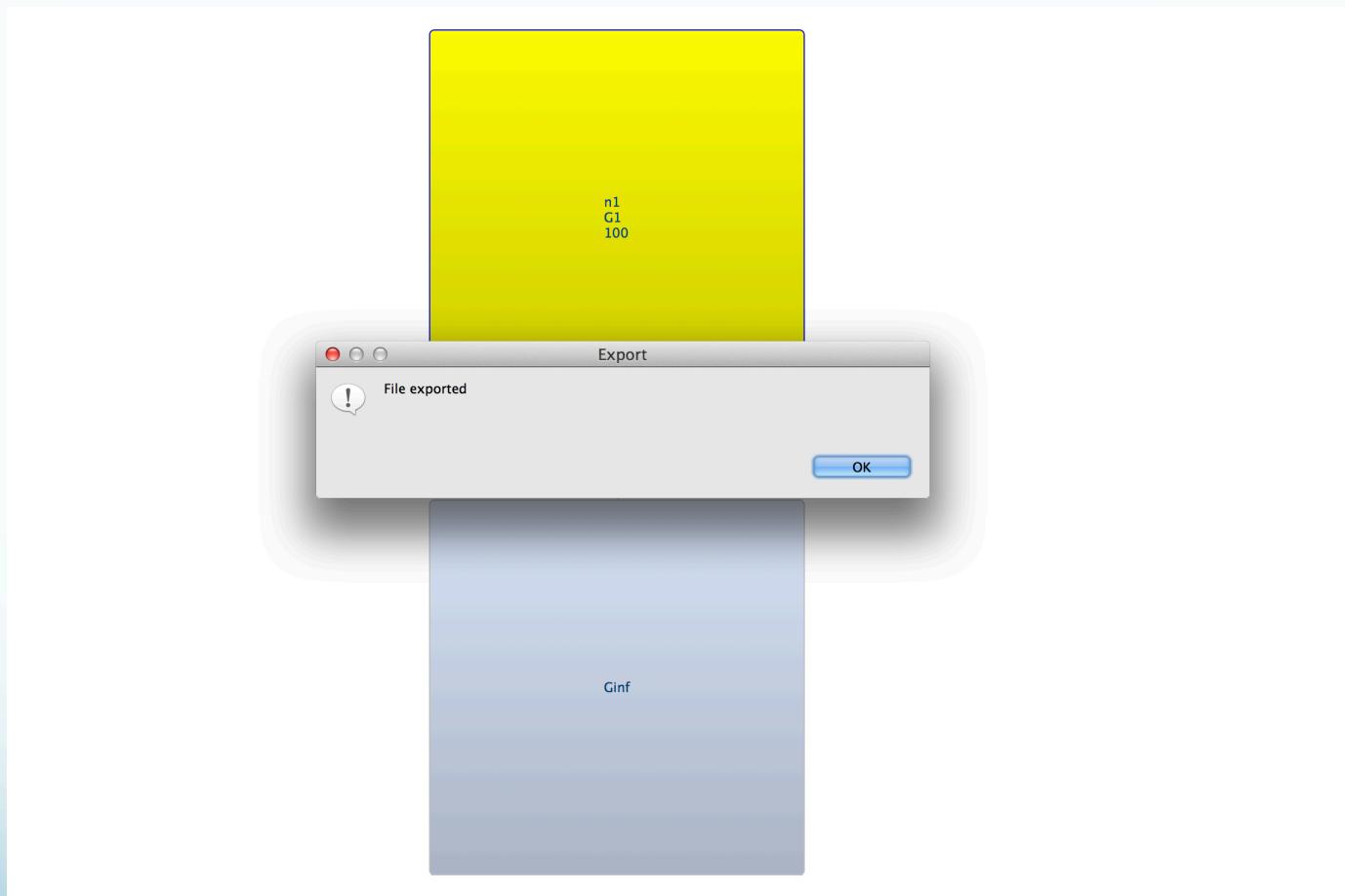
Create and graph and export it to syntax of simulator



Create and graph and export it to syntax of simulator



Create and graph and export it to syntax of simulator



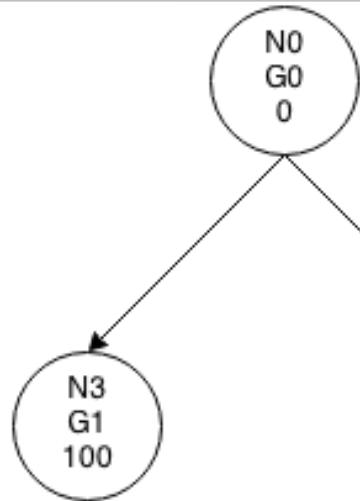
Create and graph and export it to syntax
of simulator

```
G1=gen(100,2);  
  
n1=node(G1);  
gnode=node(Ginf);  
  
e1=edge(n1,gnode,size(200));
```

Add node

- Click on node above which you want to add node and select ‘New Node’ from menu .
- To add final node give Generation name as ‘G0’ and Generation Value as ‘0’
- Each node in generation just below G0 generation , should have individual nodes in G0 generation having generation name as G0 and generation value as 0.
- **To add node above given node : File -> New Node**

Add node



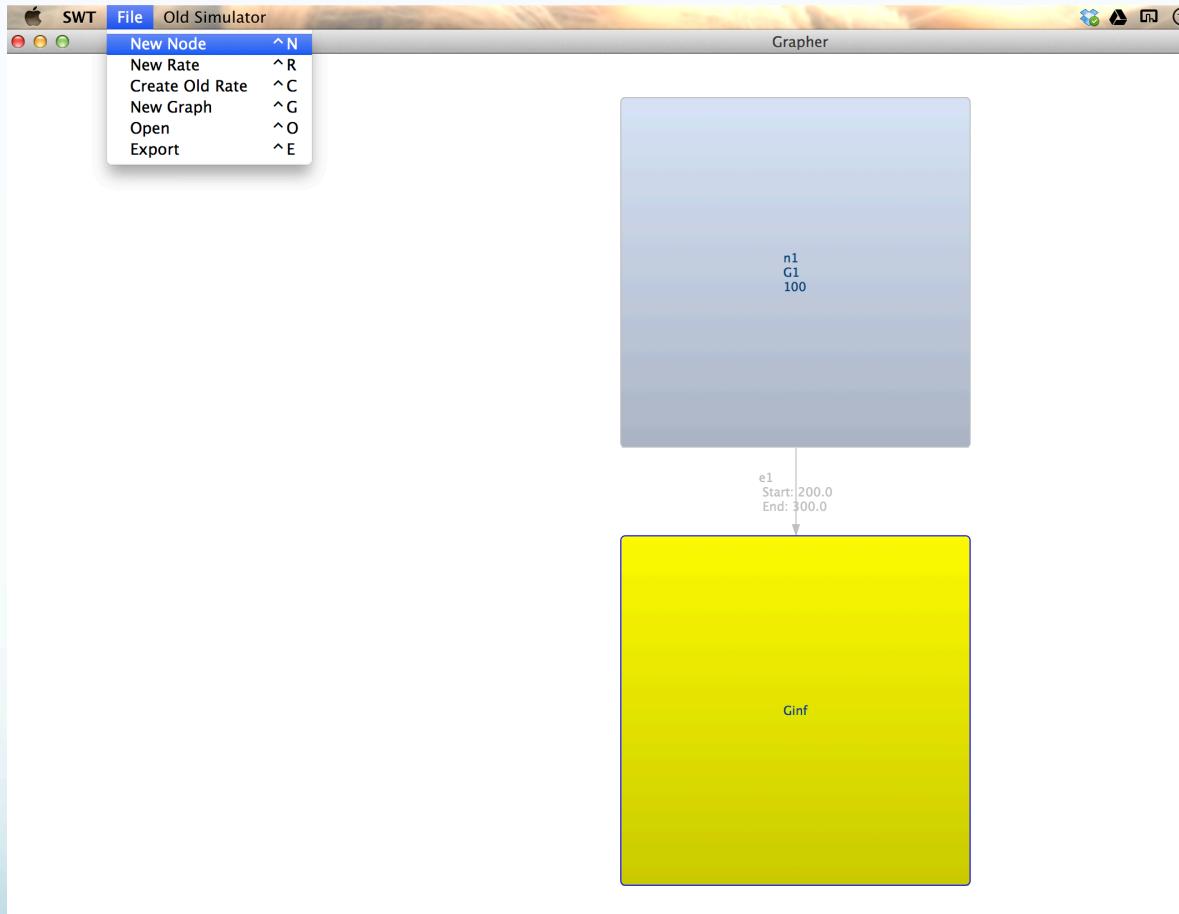
Wrong



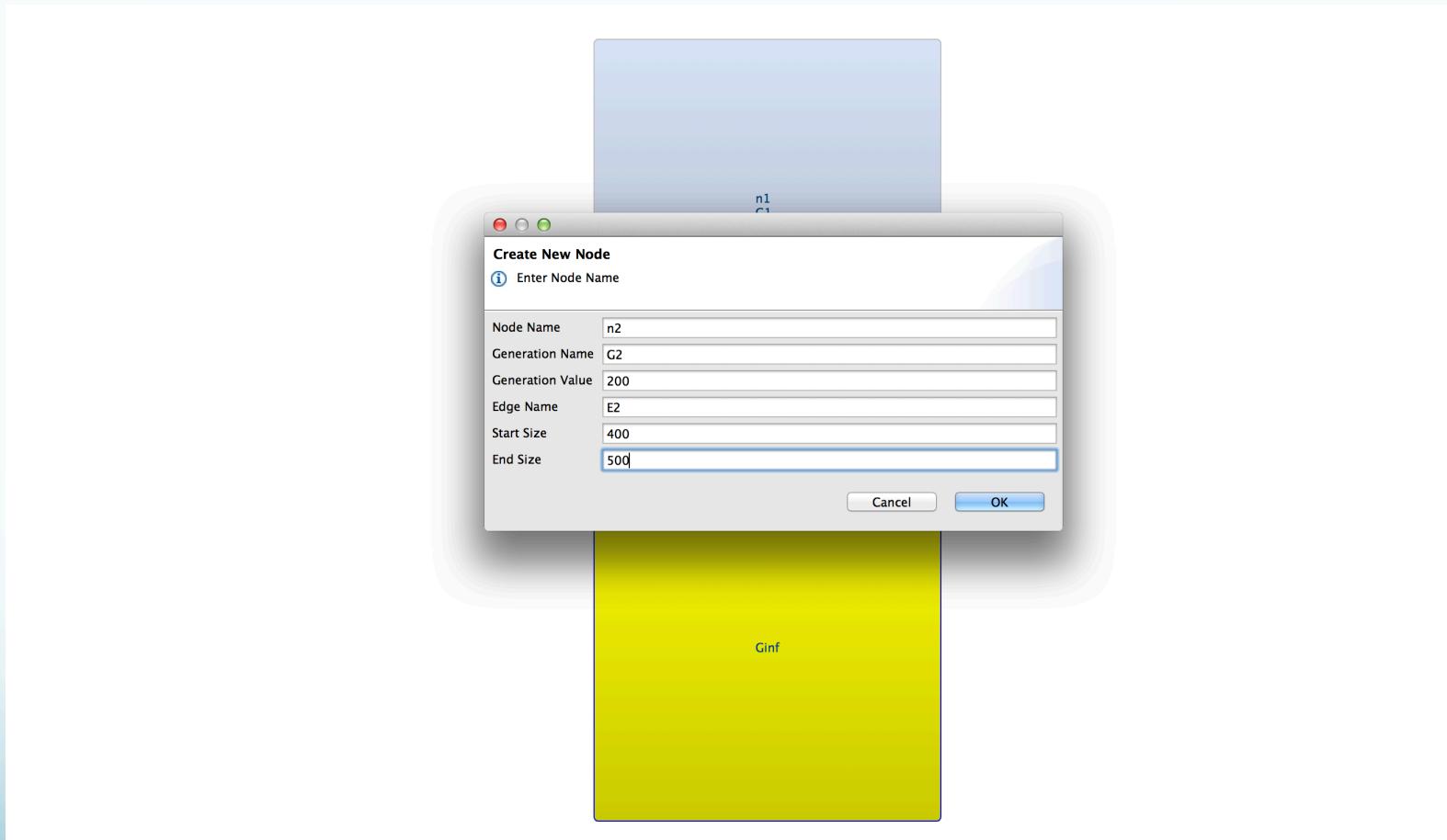
Right

ONLY APPLICABLE TO G0 AND JUST ABOVE GENERATION (HERE G1-100)

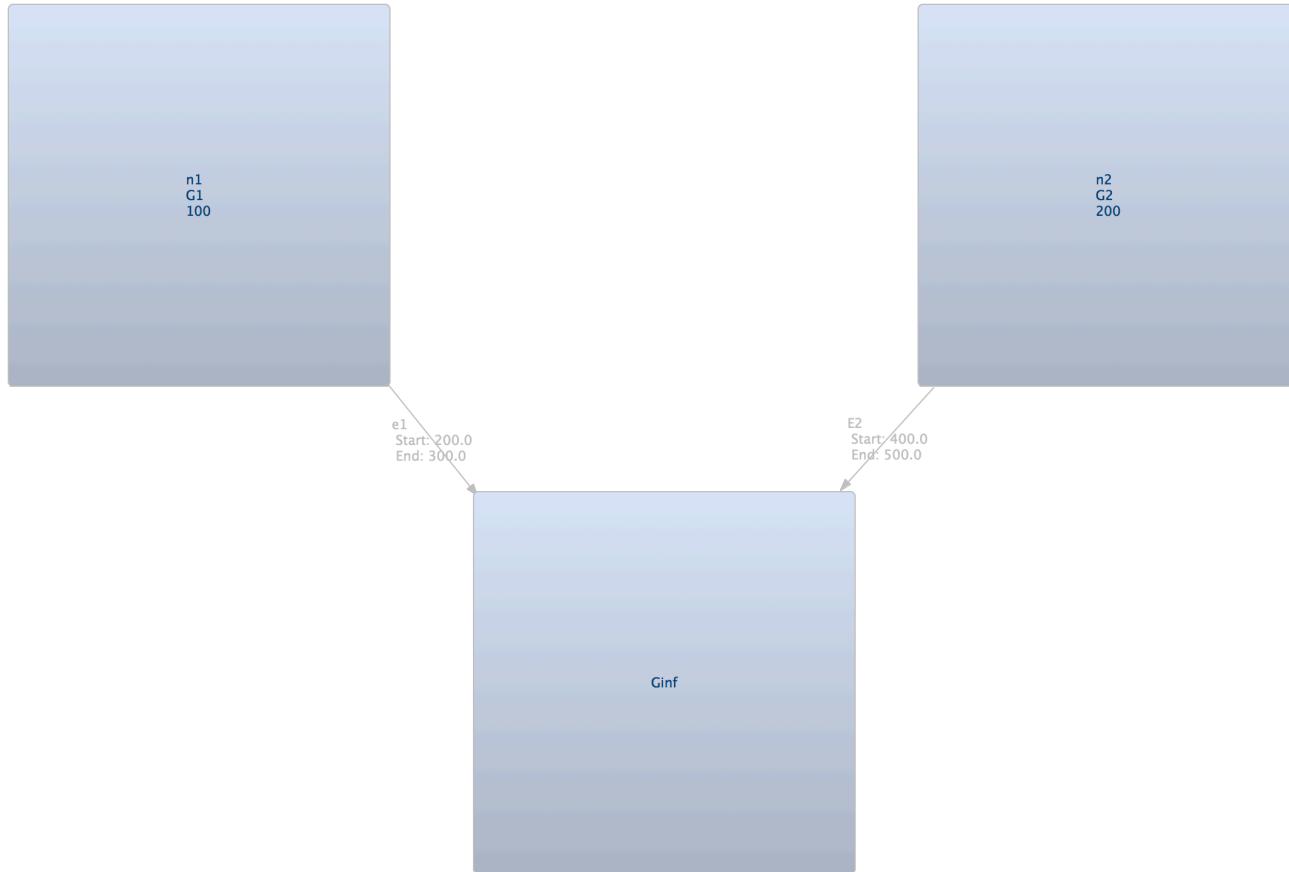
Add node



Add node



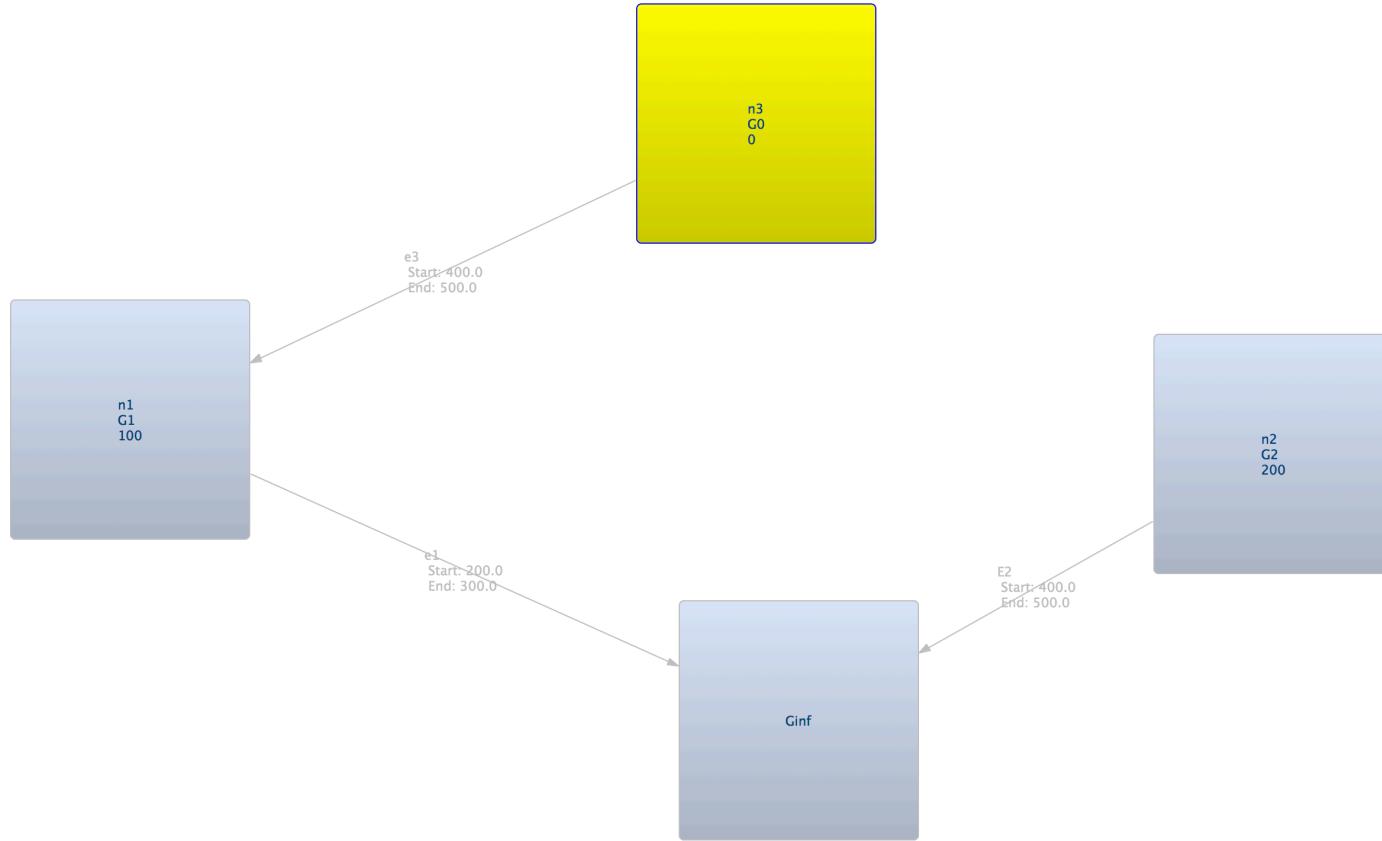
Add node



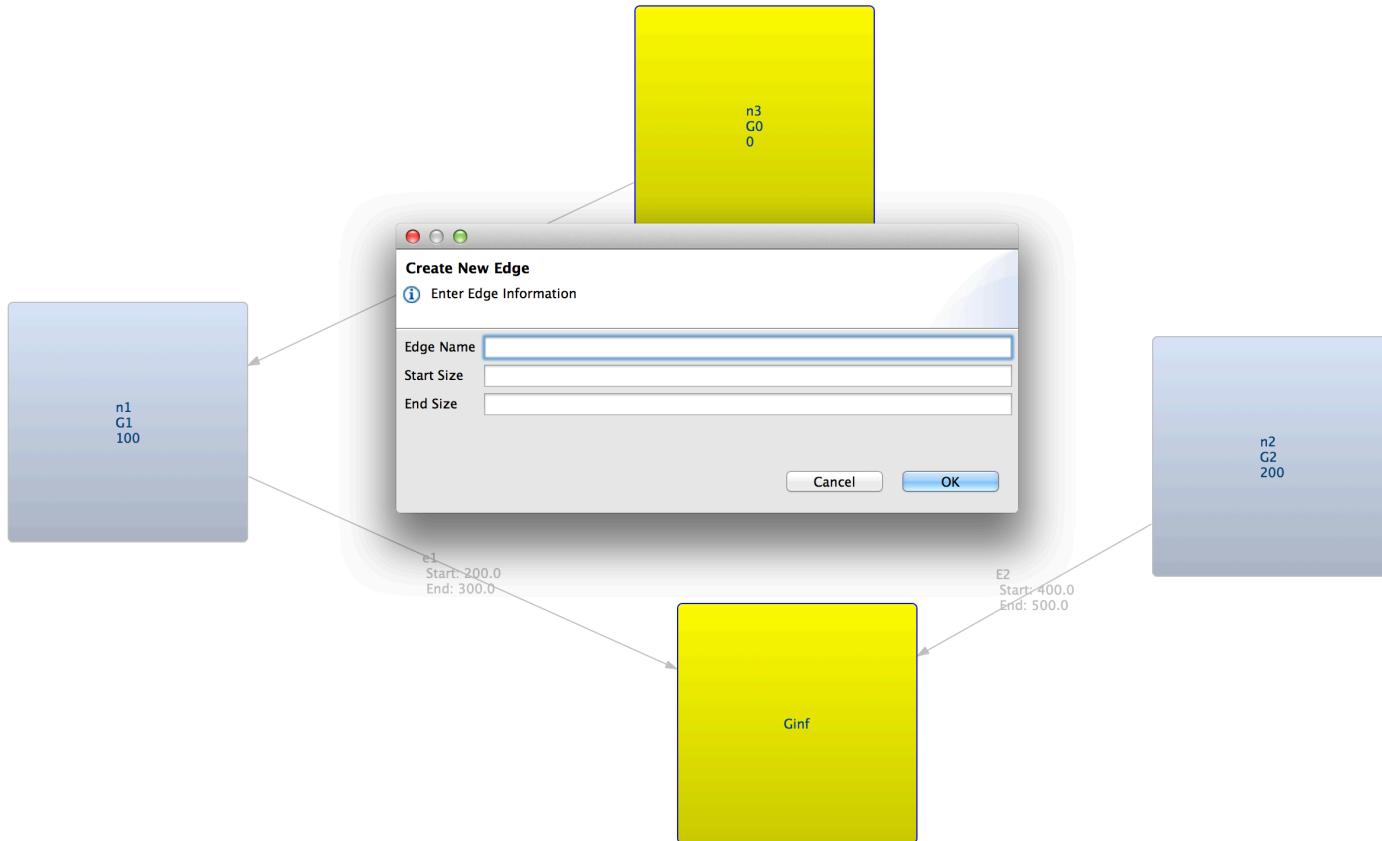
Add edge between any node to any other node

- Click on first node from which you want the edge to start .
- Press CTRL
- Click on second node at which you want to end edge .

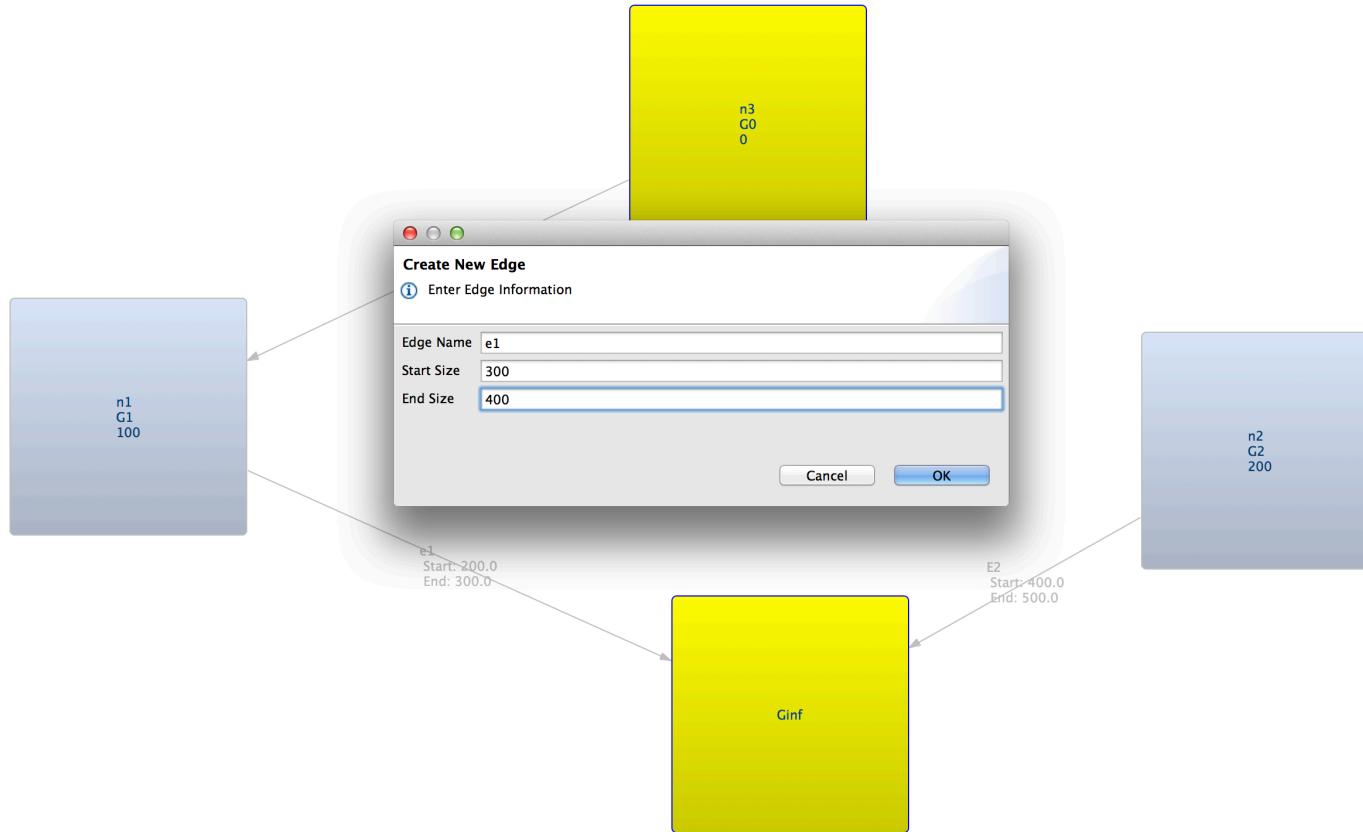
Add edge between any node to any other node



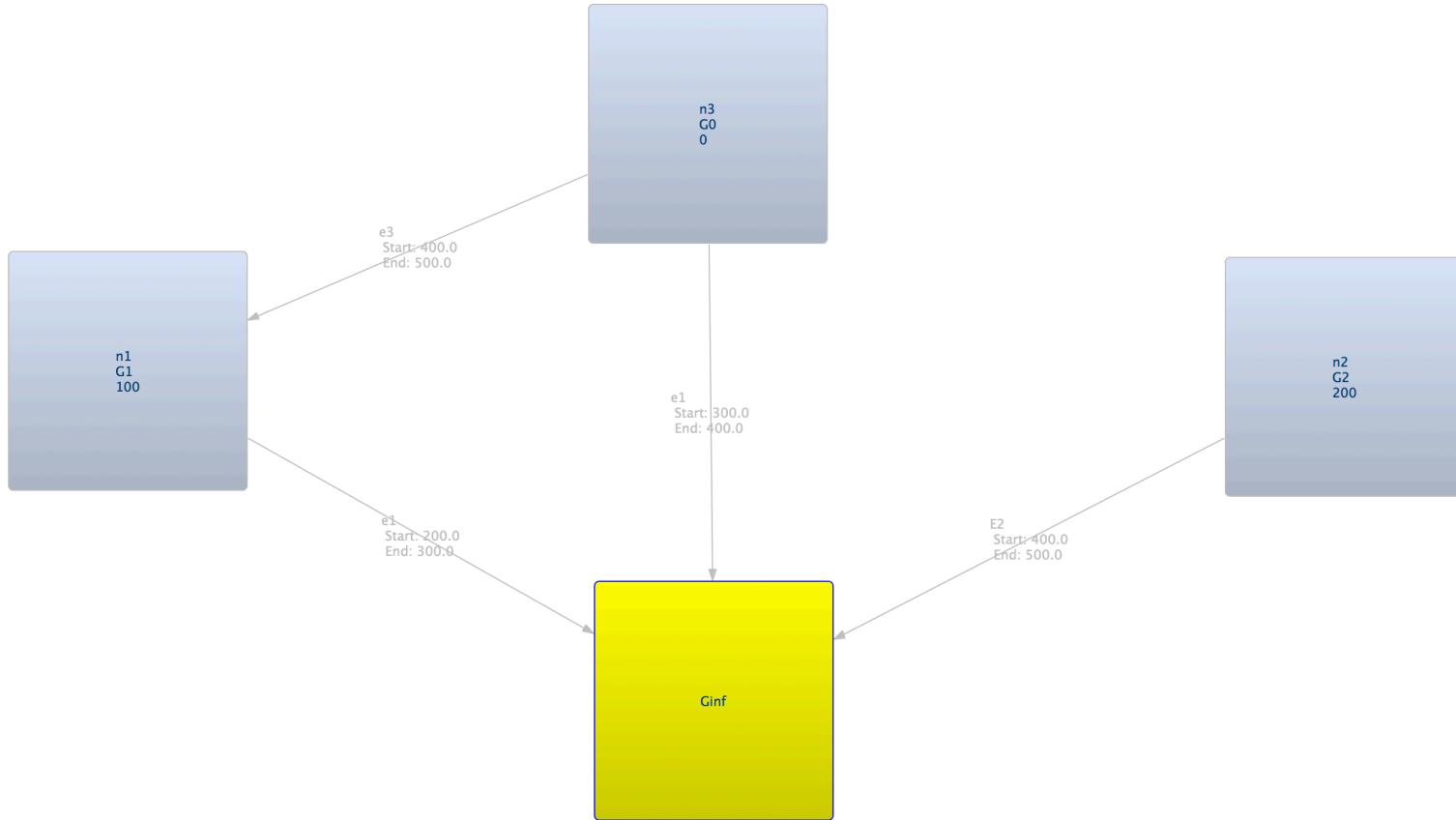
Add edge between any node to any other node



Add edge between any node to any other node



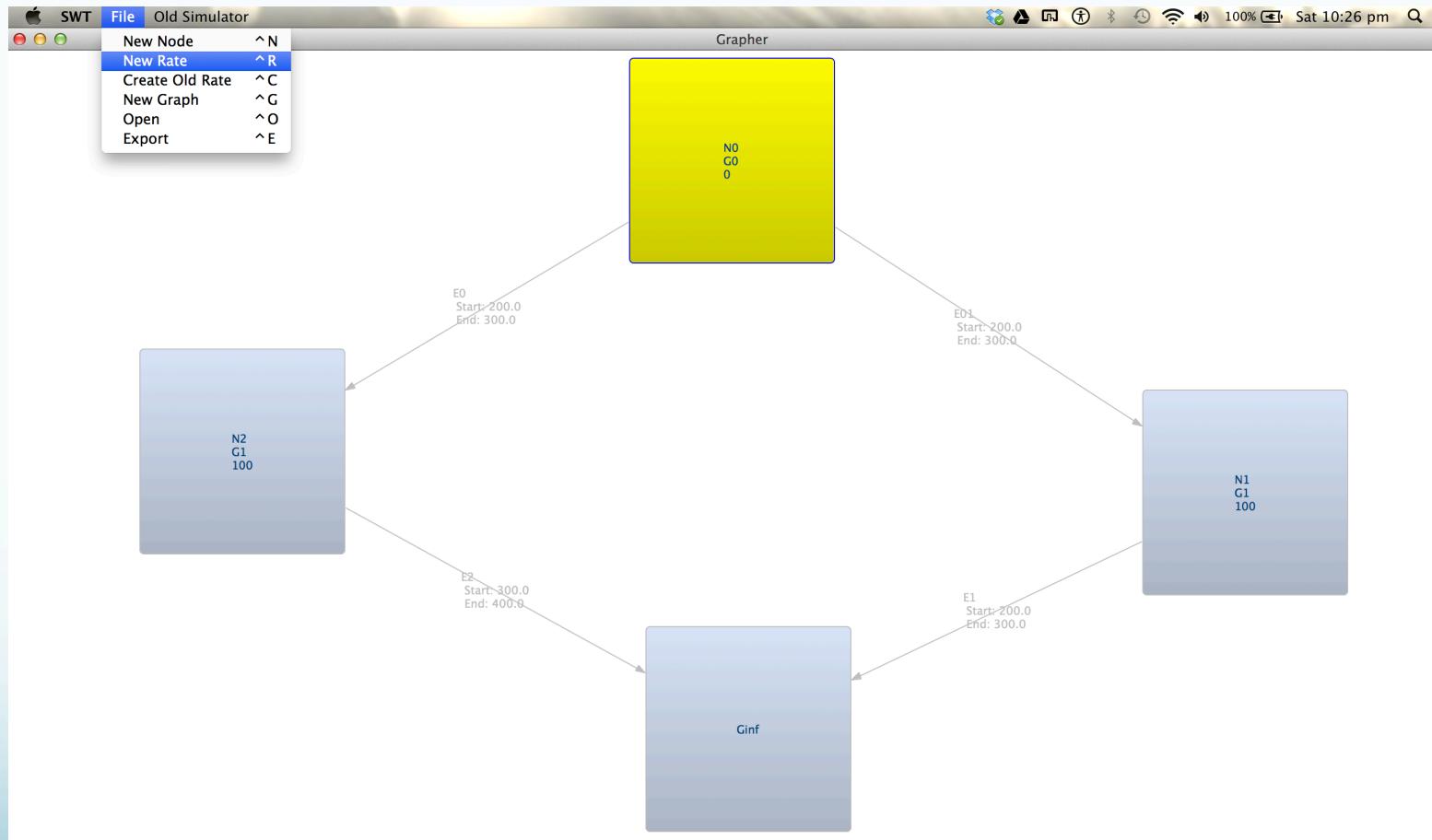
Add edge between any node to any other node



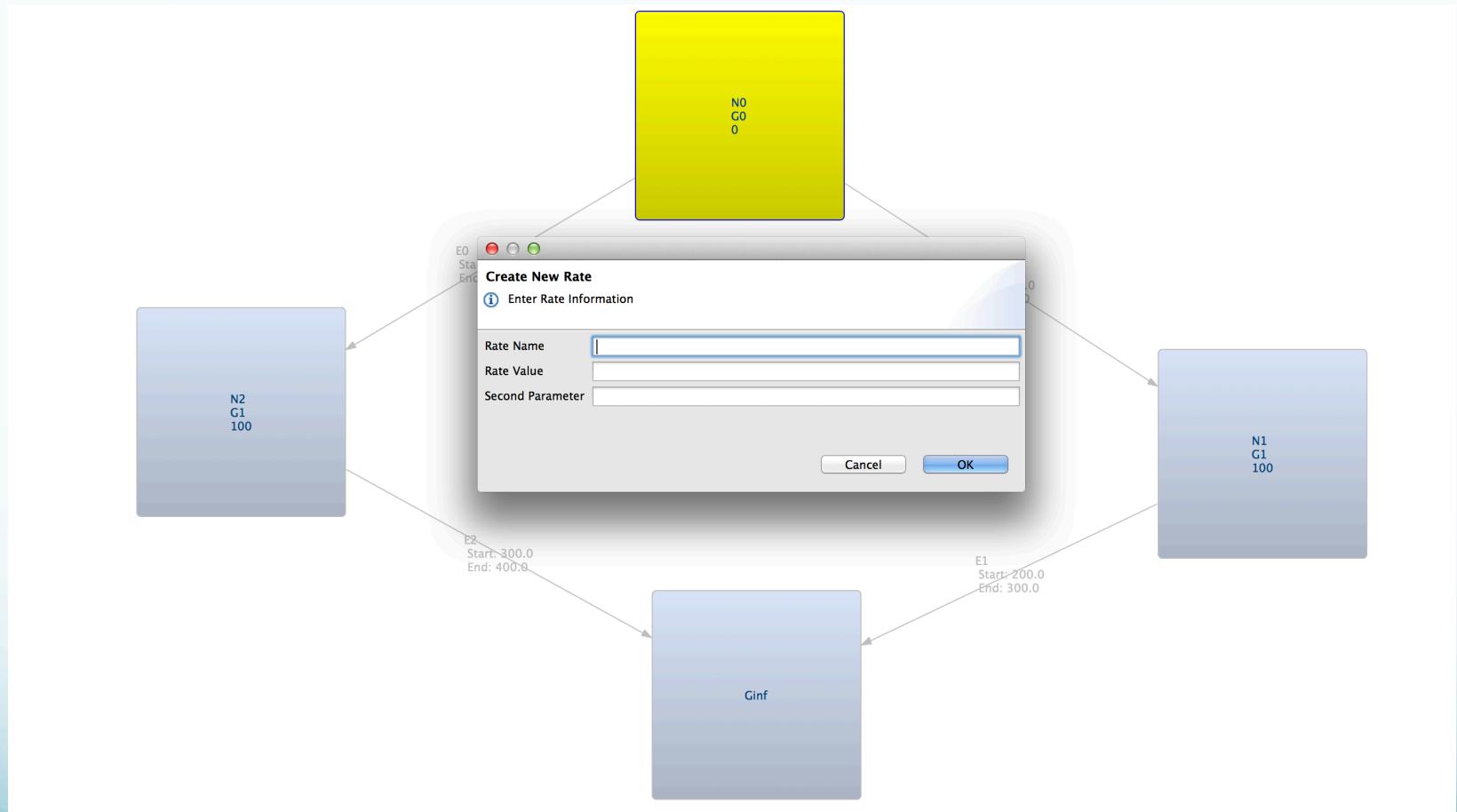
Migration add/show from GUI

- MIGRATION ADDED BETWEEN NODES OF SAME GENERATION
- First enter Rate.
- Click on the first edge between which you want to add migration .
- Hold CTRL
- Click on second edge on which you want to add migration .

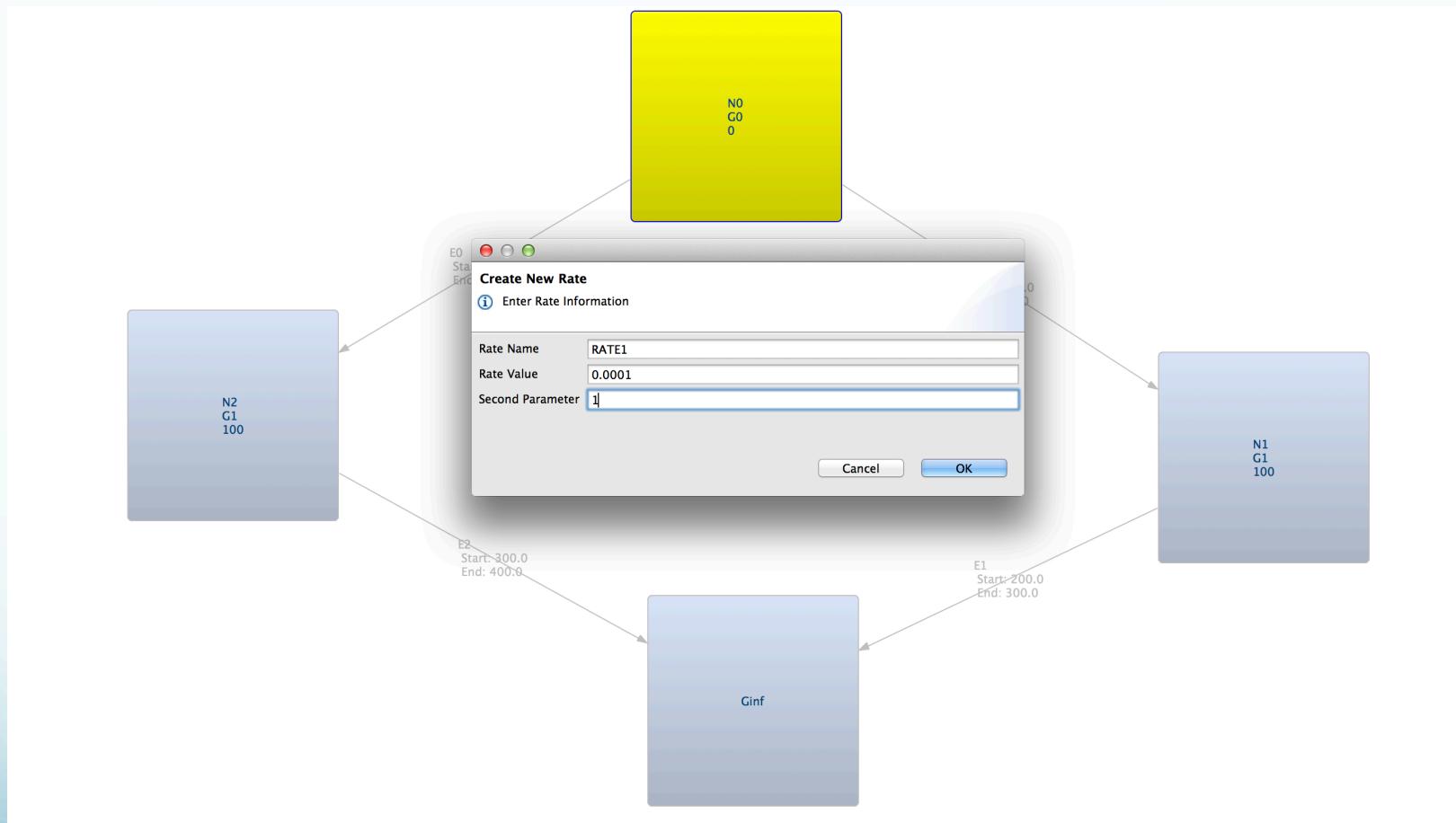
Migration add/show from GUI



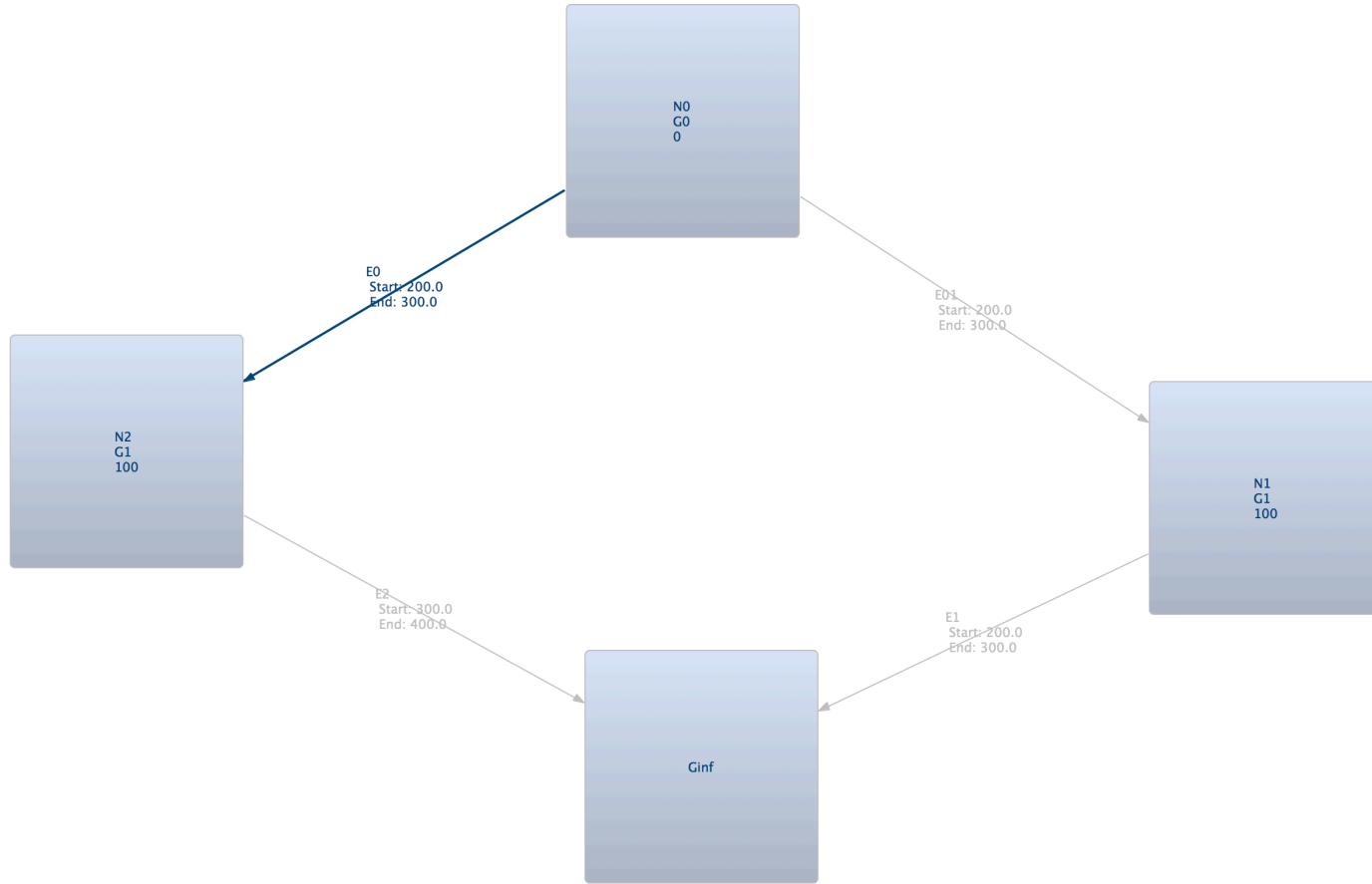
Migration add/show from GUI



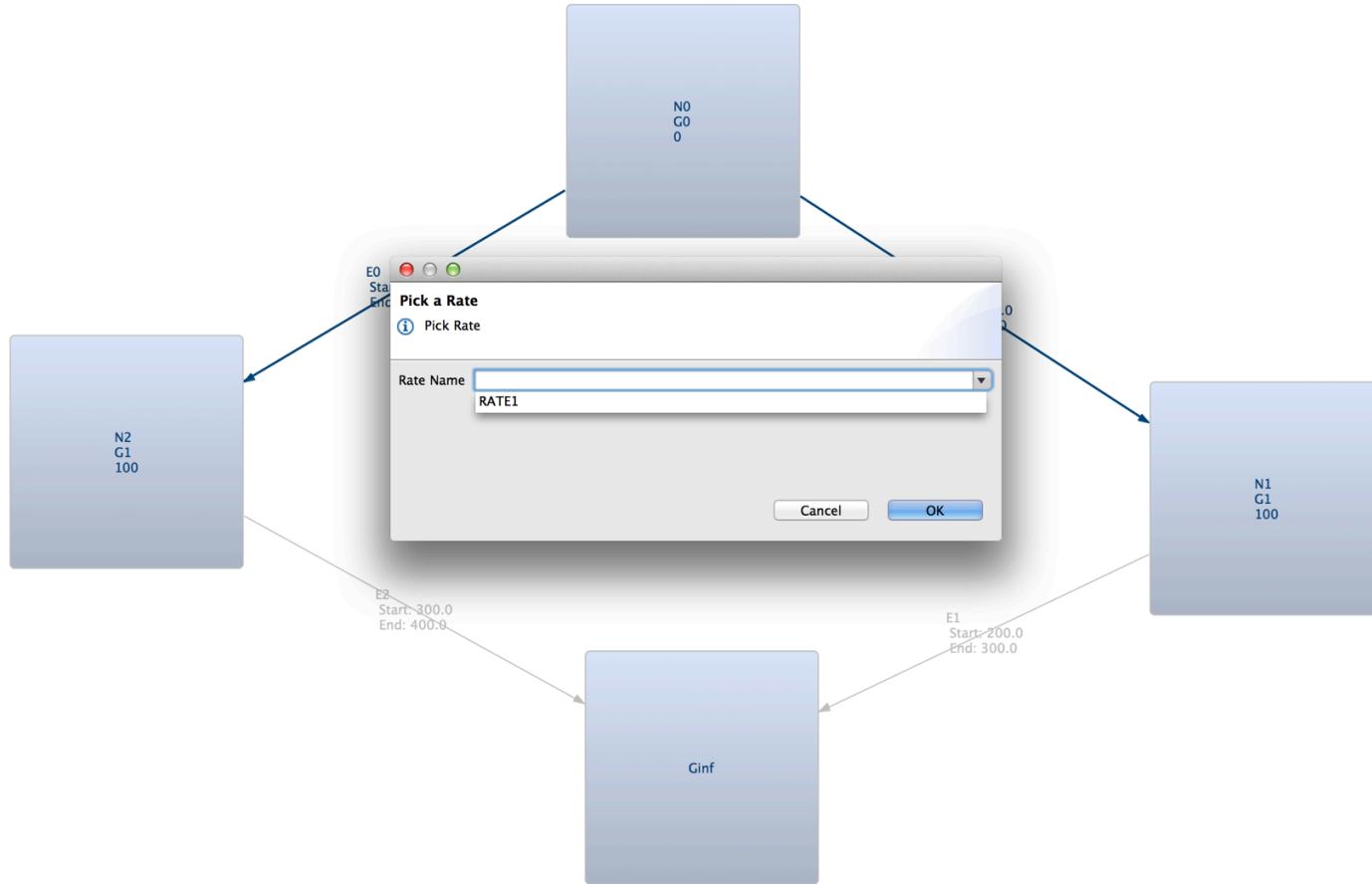
Migration add/show from GUI



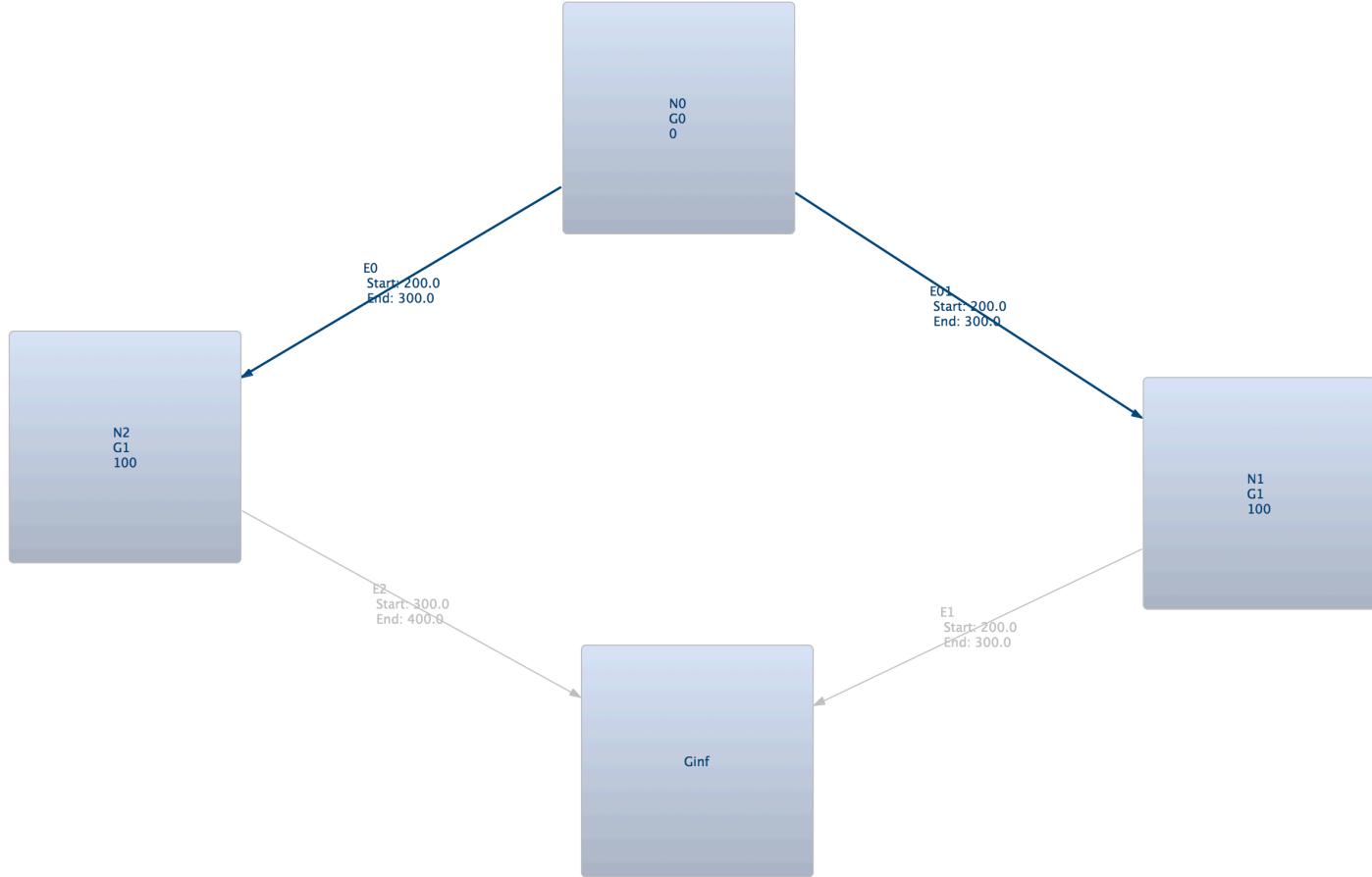
Migration add/show from GUI



Migration add/show from GUI

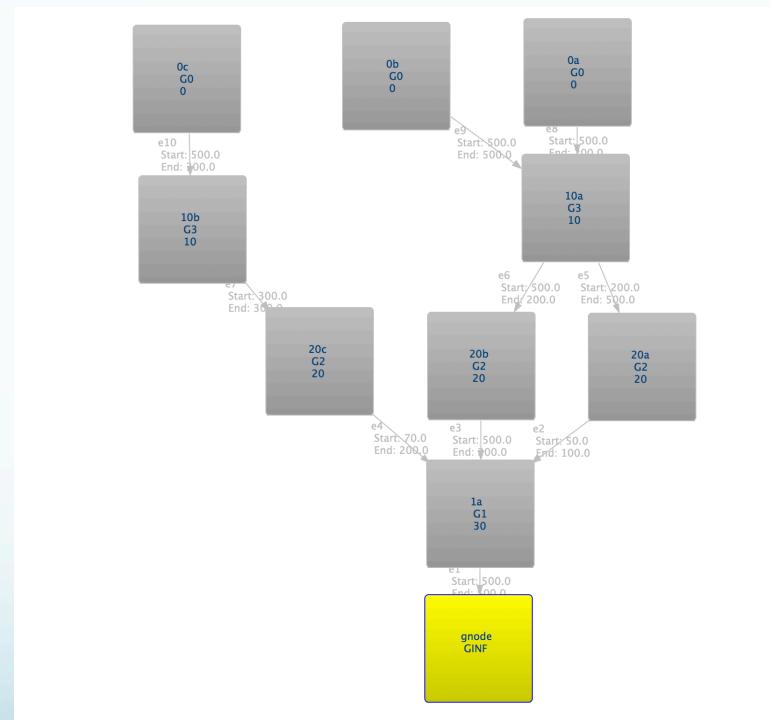


Migration add/show from GUI

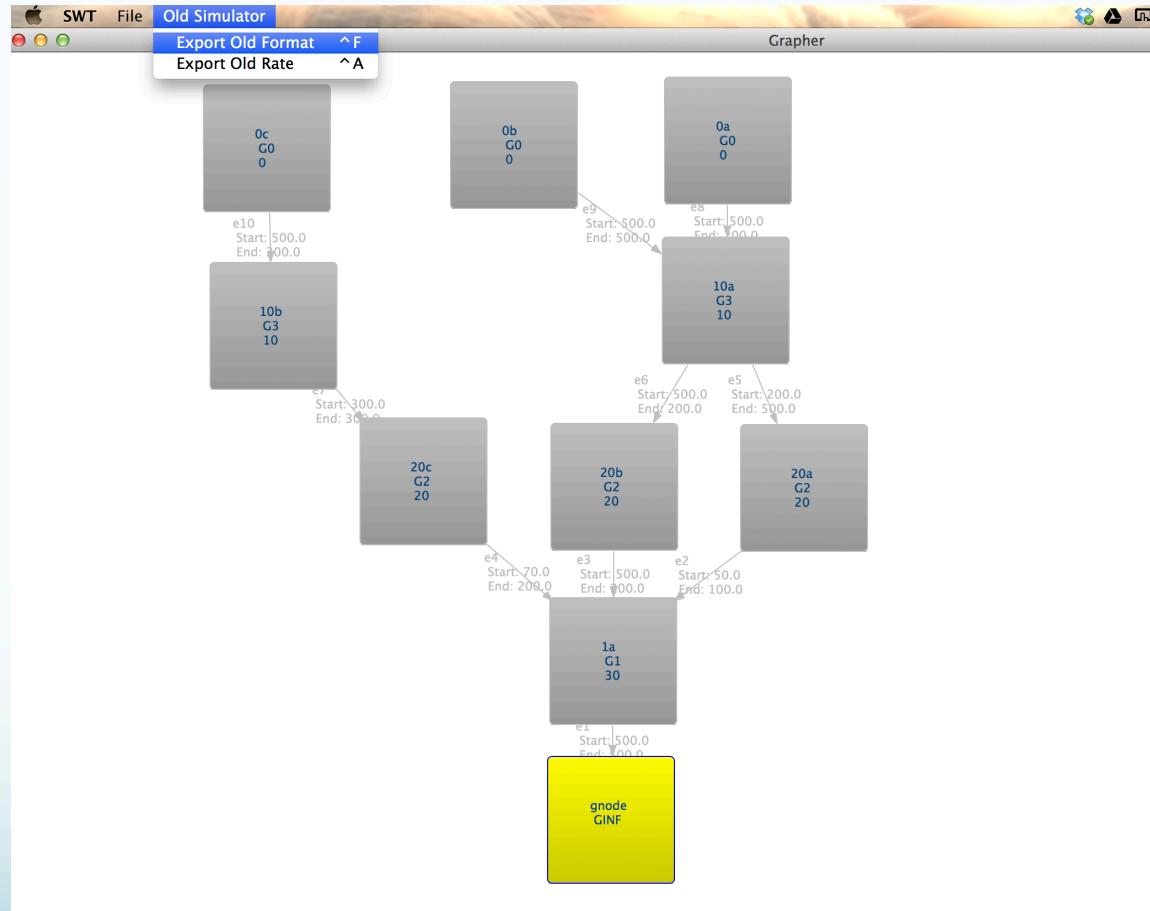


Output in Old Genome Simulator format

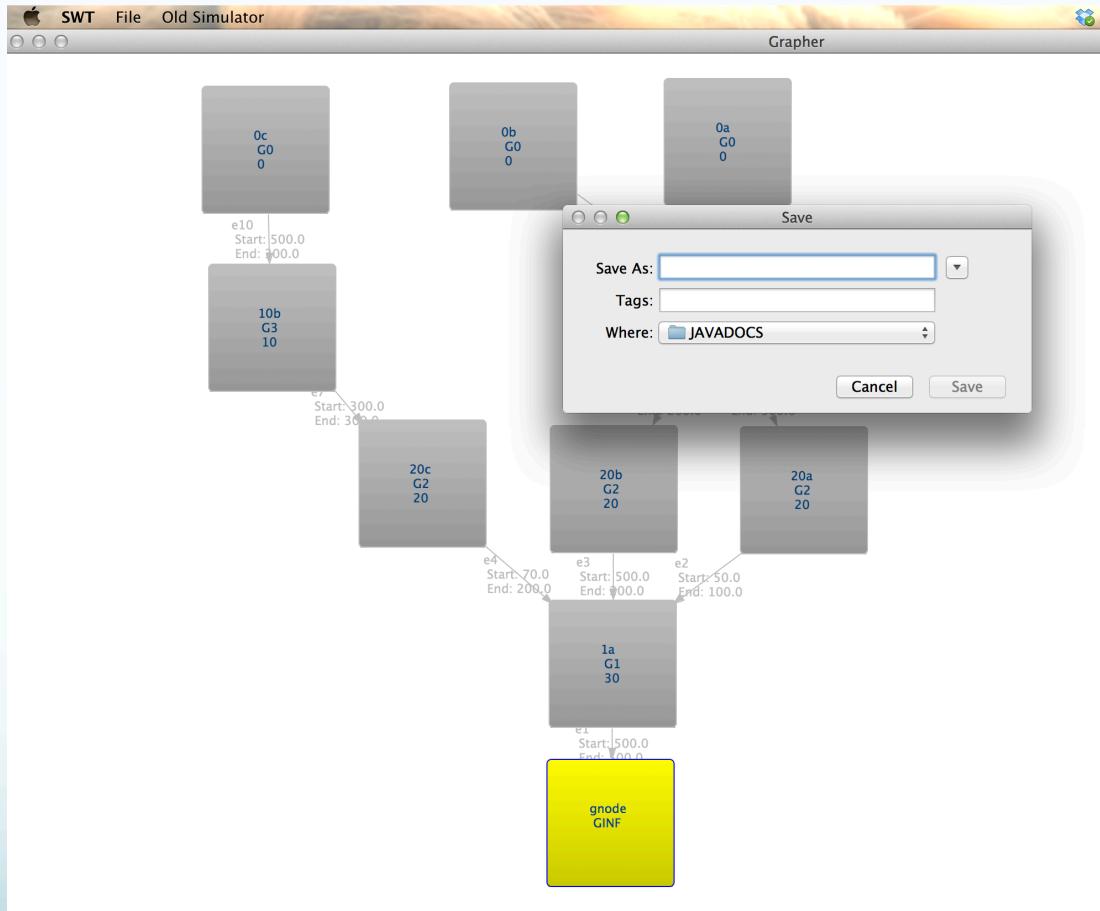
Source: <http://www.sph.umich.edu/csg/liang/genome/>



Output in Old Genome Simulator format



Output in Old Genome Simulator format



Output in Old Genome Simulator format

```
0 500 500 500
1-1 2-2 3-2
10 300 700
1-1 2-2 2-3
20 300 500 200
1-1 2-1 3-1
30 500
```

Thank You

For constant guidance given by

- Dr. Itsik Peer .
<http://www.cs.columbia.edu/~itsik/home.html>
- Dr. Pier Palamara.
<http://www.cs.columbia.edu/~pier/>