



Università
degli Studi di
Messina

DIPARTIMENTO DI INGEGNERIA

Dependable computing modeling and simulation

Petri nets

Master degree in
Engineering in Computer Science

Introduction

- Markov chains are useful to study a lot of problems
 - Reliability
 - Availability
 - Performance
 - Performability
 - Dependability in general
- They are difficult to be managed

Markov chain based models

- State space
- Representation of a system could be difficult
- A more intuitive representation would be useful

Petri nets

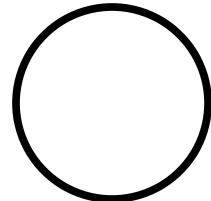
- Petri nets are a tool, also graphical, to represent activity flows
- It is possible to easily represent system interactions
 - Synchronization
 - Sequences
 - Concurrency
 - Conflicts

Definition

- A Petri net is a tuple
- $PN = \{P, T, I, O, M_0\}$
 - P: set of n_P places
 - T: set of n_T transitions
 - I, O: sets of input and output functions
 - $M_0 = \{m_1, m_2, \dots, m_{np}\}$: initial marking (state)

Place and transitions

- A place is drawn as a circle



- A transition is drawn as a bar



Input and output functions

- $I : T \rightarrow Bag(P)$
- $O : T \rightarrow Bag(P)$
- $Bag(A)$ is a set with repeated elements
 - es: $\{p1, p1, p4\} = \{p1^{(2)}, p4\} = \{2, 0, 0, 1\}$
- Input and output arcs are drawn as oriented arcs in graphical representation

Token

- Multiplicity is represented with tokens
- Each place can contain tokens
- A token is drawn as a little black circle
- A token distribution over places identify a Petri net state
- Petri net states are defined through *markings*

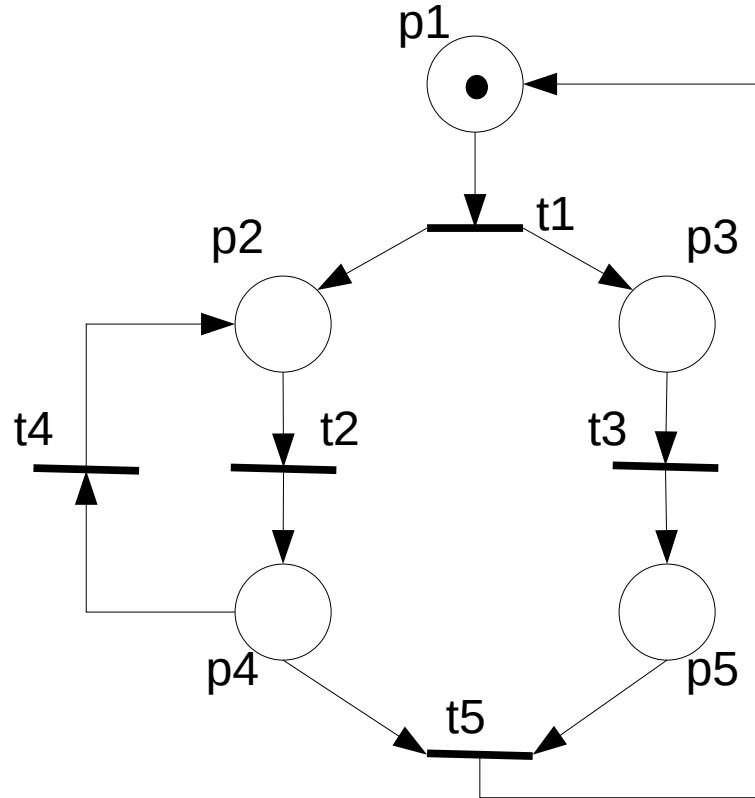
Marking

- $M_0 = \{m_1, m_2, \dots, m_{np}\}$
- $m_i \in \mathcal{N}$; number of *tokens* into place p_i

Notation

- Let us denote with:
 - $I(t)$: the bag of input places of transition t
 - $I(t, p)$: multiplicity of place p in bag $I(t)$
 - $O(t)$: the bag of output places of transition t
 - $O(t, p)$: multiplicity of place p in bag $O(t)$

Example (1)



$$P = \{p_1, p_2, p_3, p_4, p_5\}$$

$$T = \{t_1, t_2, t_3, t_4, t_5\}$$

$$I(t_1) = \{p_1\}, O(t_1) = \{p_2, p_3\}$$

$$I(t_2) = \{p_2\}, O(t_2) = \{p_4\}$$

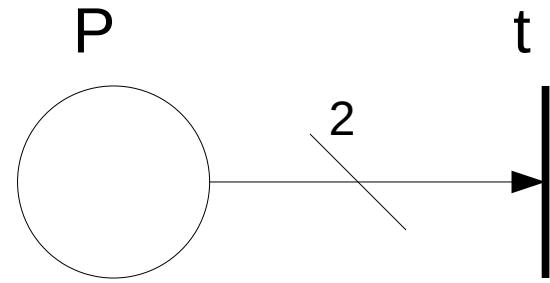
$$I(t_3) = \{p_3\}, O(t_3) = \{p_5\}$$

$$I(t_4) = \{p_4\}, O(t_4) = \{p_2\}$$

$$I(t_5) = \{p_4, p_5\}, O(t_5) = \{p_1\}$$

$$M_0 = \{1, 0, 0, 0, 0\}$$

Example (2)



$$I(t) = \{P\} \quad I(t, P) = 1$$

$$I(t) = \{P^{(2)}\} \quad I(t, P) = 2$$

Enabling rule

$$\forall p \in P_I(t_k) \Rightarrow M_i(p) \geq I(t_k, p)$$

- $P_I(t_k)$: set of input places of transition t_k
- $M_i(p)$: number of tokens in place p in marking M
- Transition t_k is said enabled in marking M iff the number of tokens in its input places is at least equal to the multiplicity of its input arcs

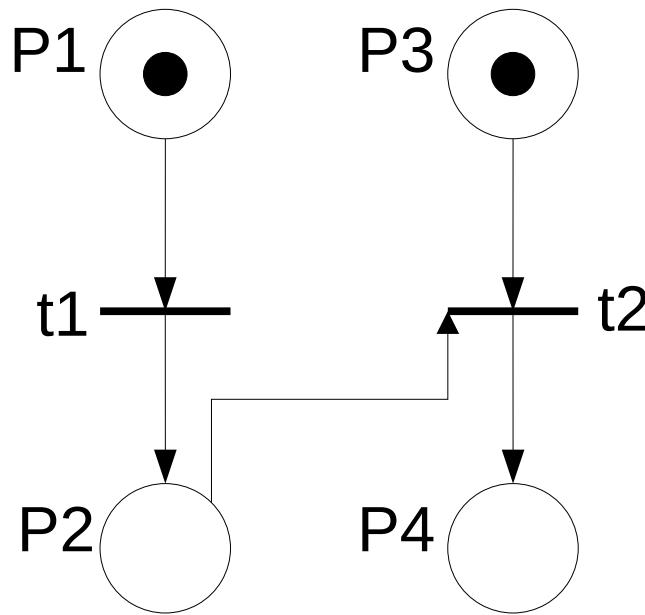
Transition firing

- An enabled transition can fire producing a marking change according to the following rule

$$M_j = M_i - I(t_k) + O(t_k)$$

- Petri net dynamics
- What happen when more transitions are enabled?

Example



$$M_0 = \{1, 0, 1, 0\}$$

$$M_0(P1) = 1 = I(t_1, P1)$$

$$M_0(P3) = 1 = I(t_2, P3)$$

$$0 = M_0(P2) < I(t_2, P2) = 1$$

$$I(t_1) = \{1, 0, 0, 0\}$$

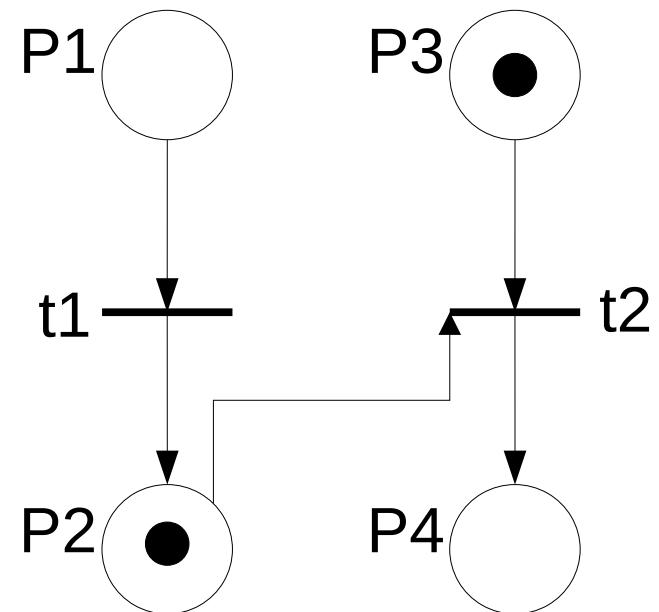
$$O(t_1) = \{0, 1, 0, 0\}$$

$$\{1, 0, 1, 0\} - .$$

$$\{1, 0, 0, 0\} + .$$

$$\{0, 1, 0, 0\} = .$$

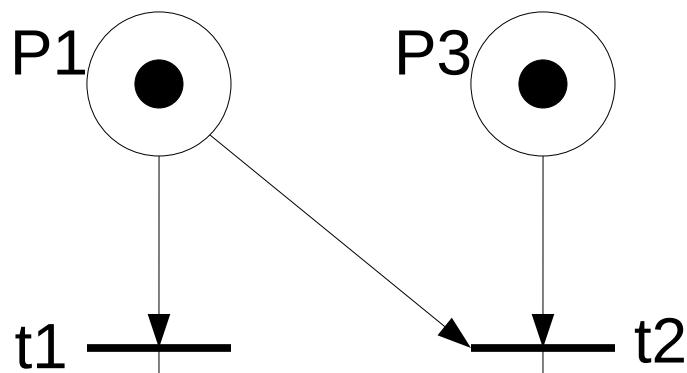
$$\{0, 1, 1, 0\}$$



Definitions

- Two enabled transitions are said
 - **in conflict** when the firing of one of the two **disable** the other
 - **concurrent** when the firing of one of the two **does not disable** the other

Examples



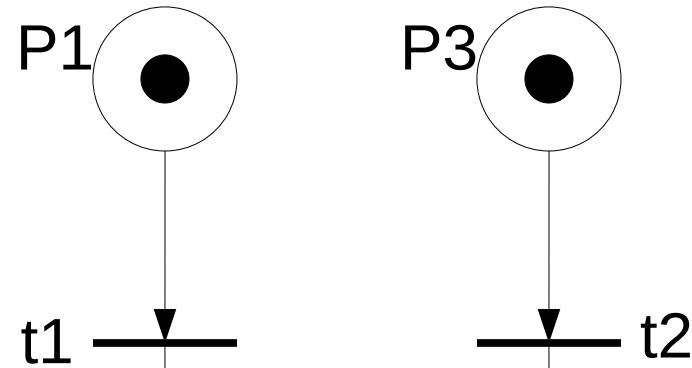
t1

t2

P2

P4

conflict



t1

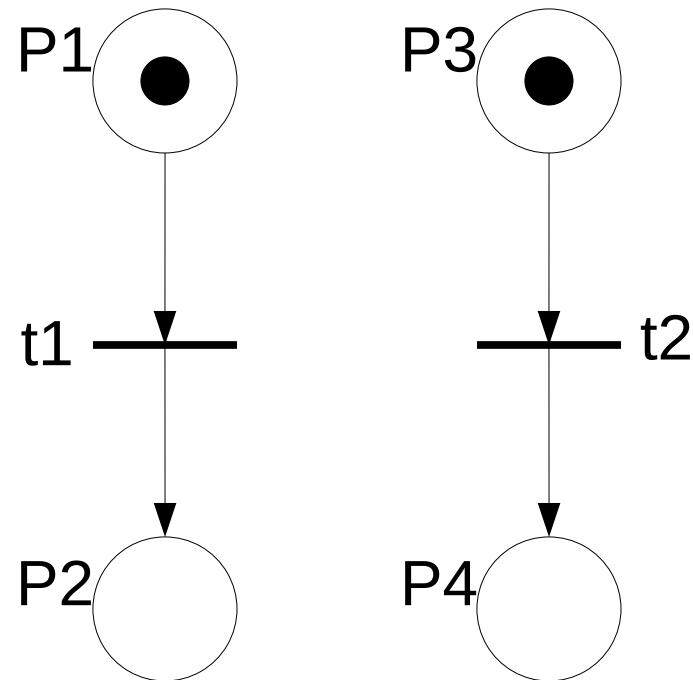
t2

P2

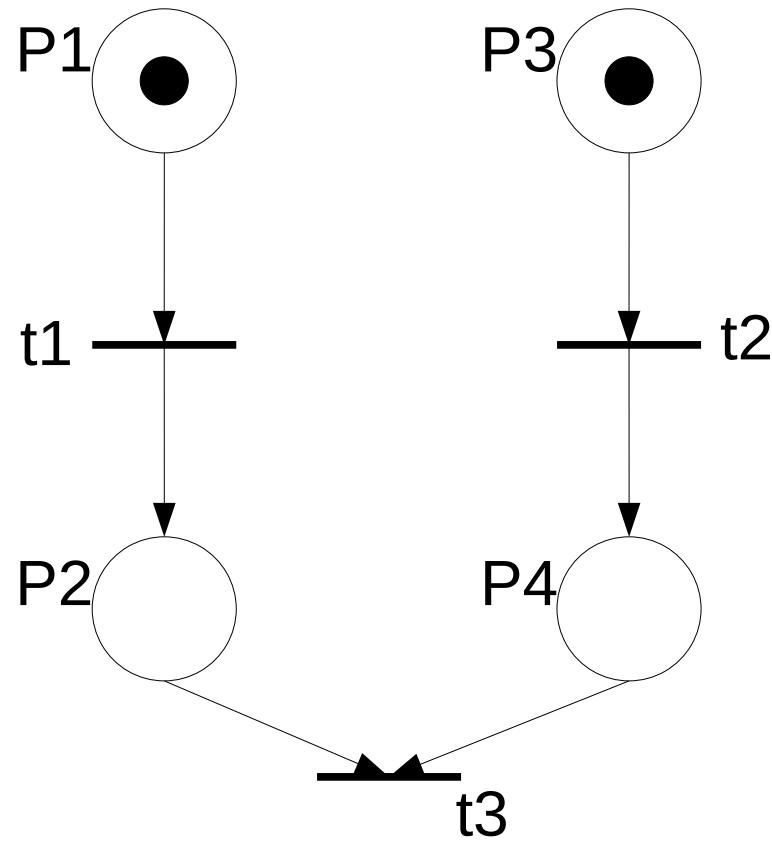
P4

concurrency

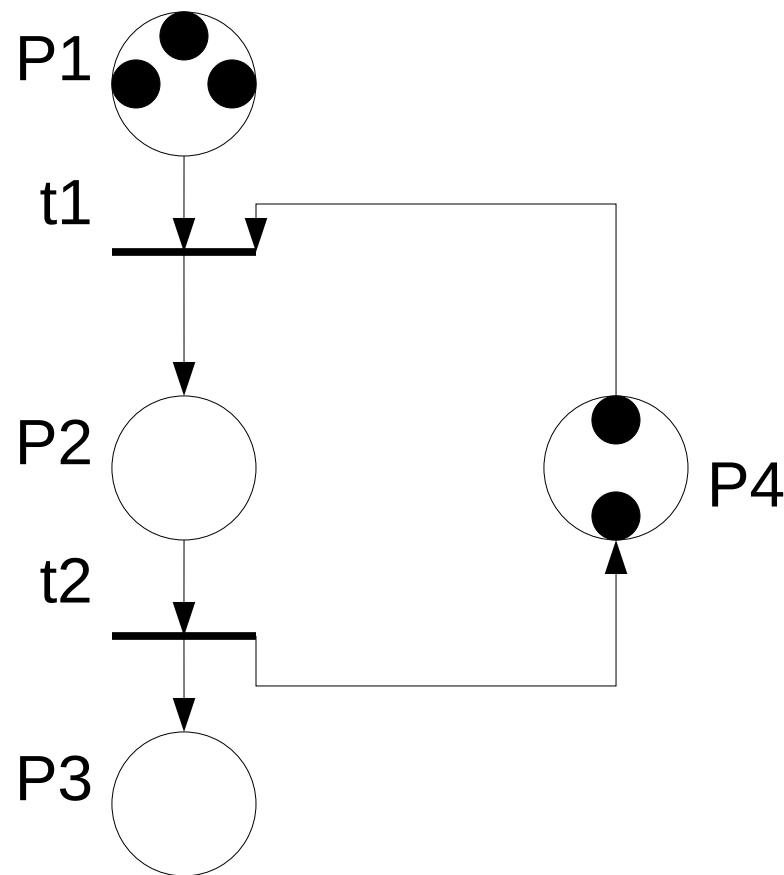
Example - concurrency



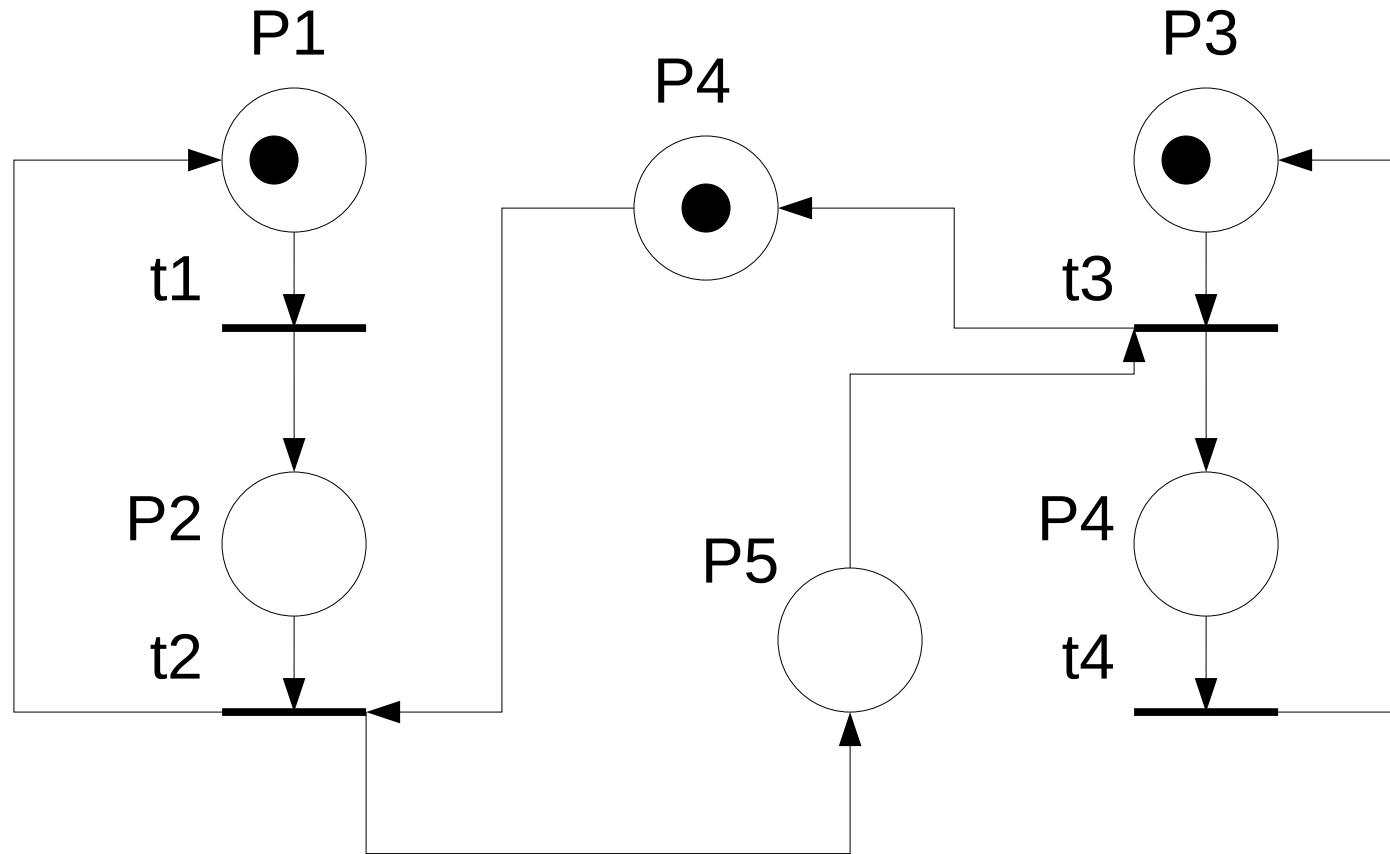
Example - synchronization



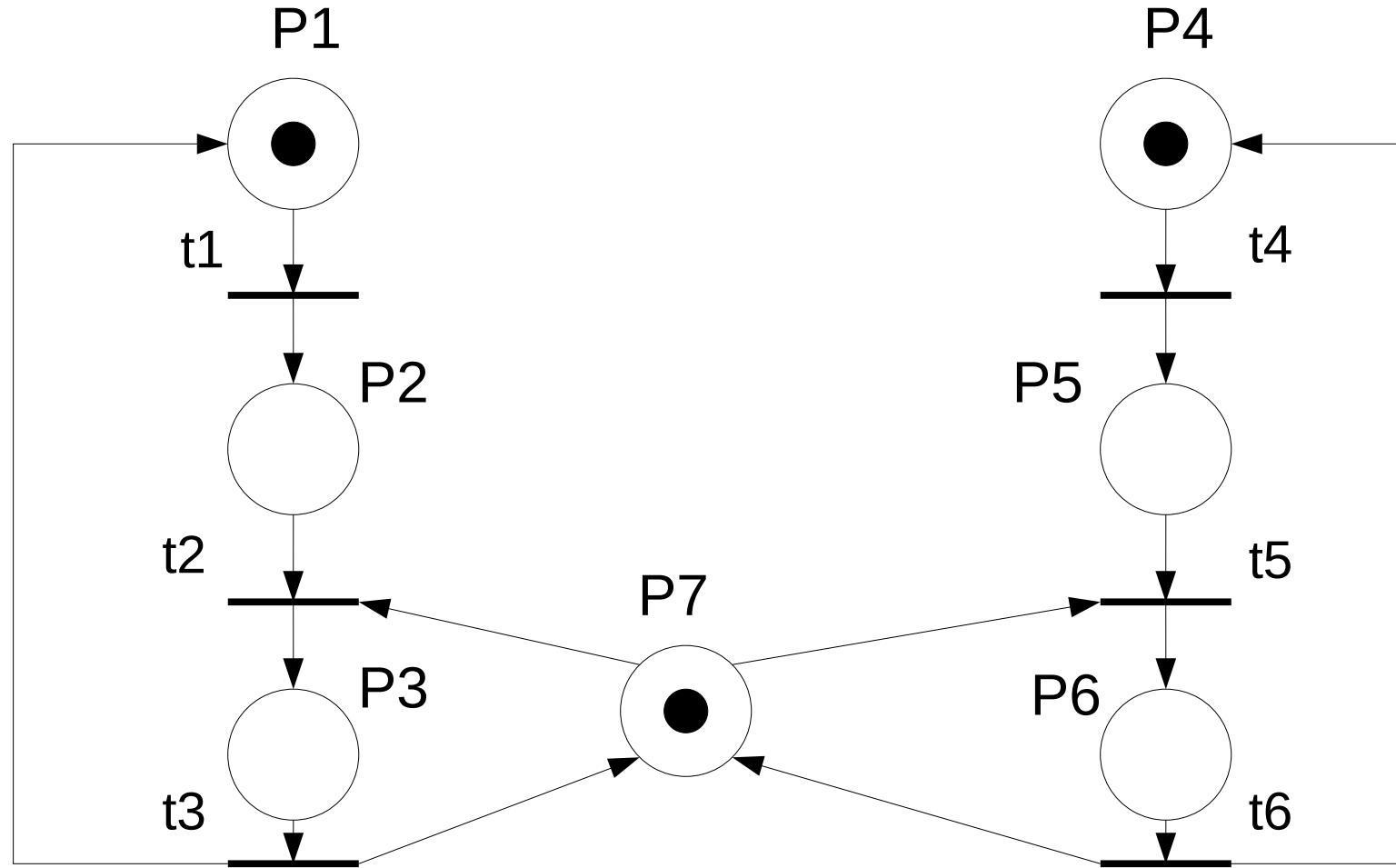
Example – finite resources



Example – producer/consumer

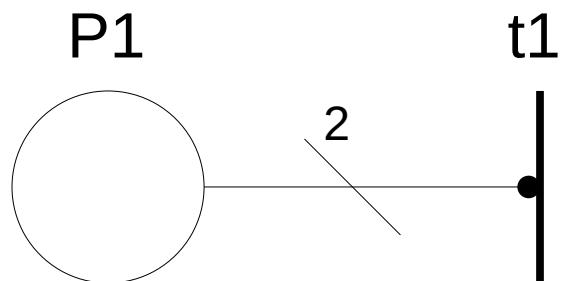


Example – mutual exclusion



Inhibitor arcs

- $PN = \{P, T, I, O, H, M_o\}$
- $H : T \rightarrow Bag(P)$



$$H(t1) = \{P1, P1\}$$

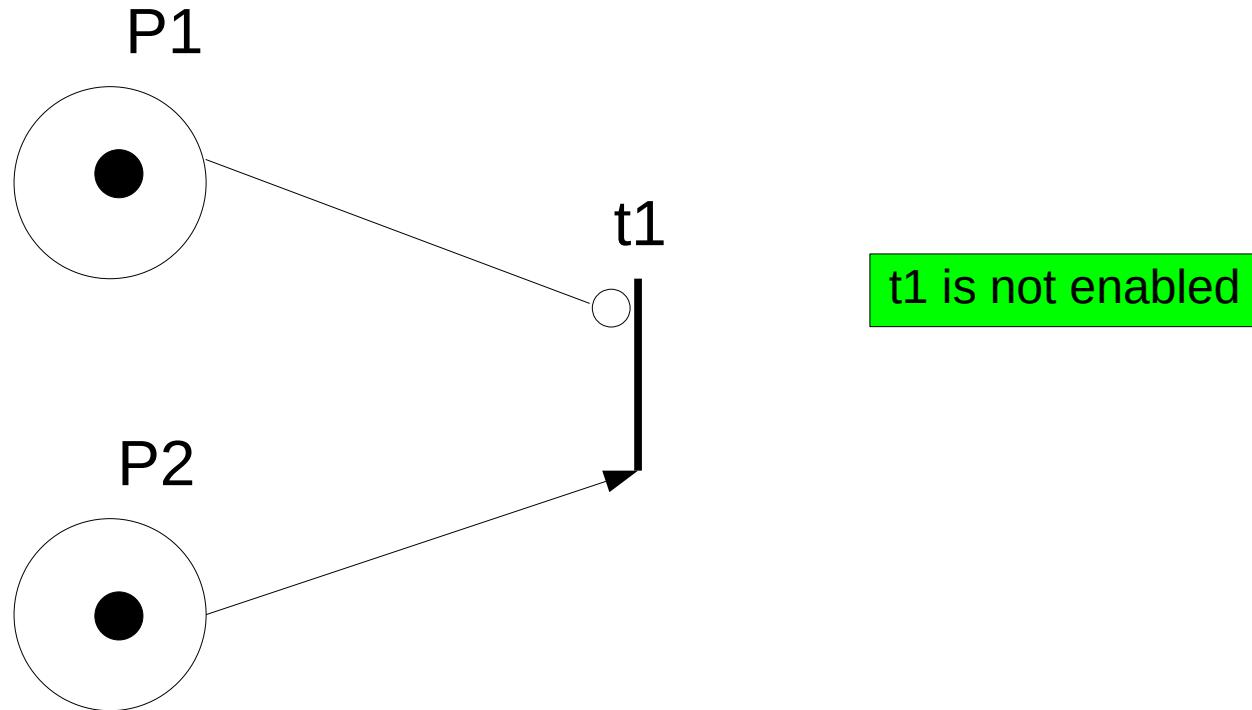
Enabling rule

$$\forall p \in P_I(t_k) \Rightarrow M_i(p) \geq I(t_k, p)$$

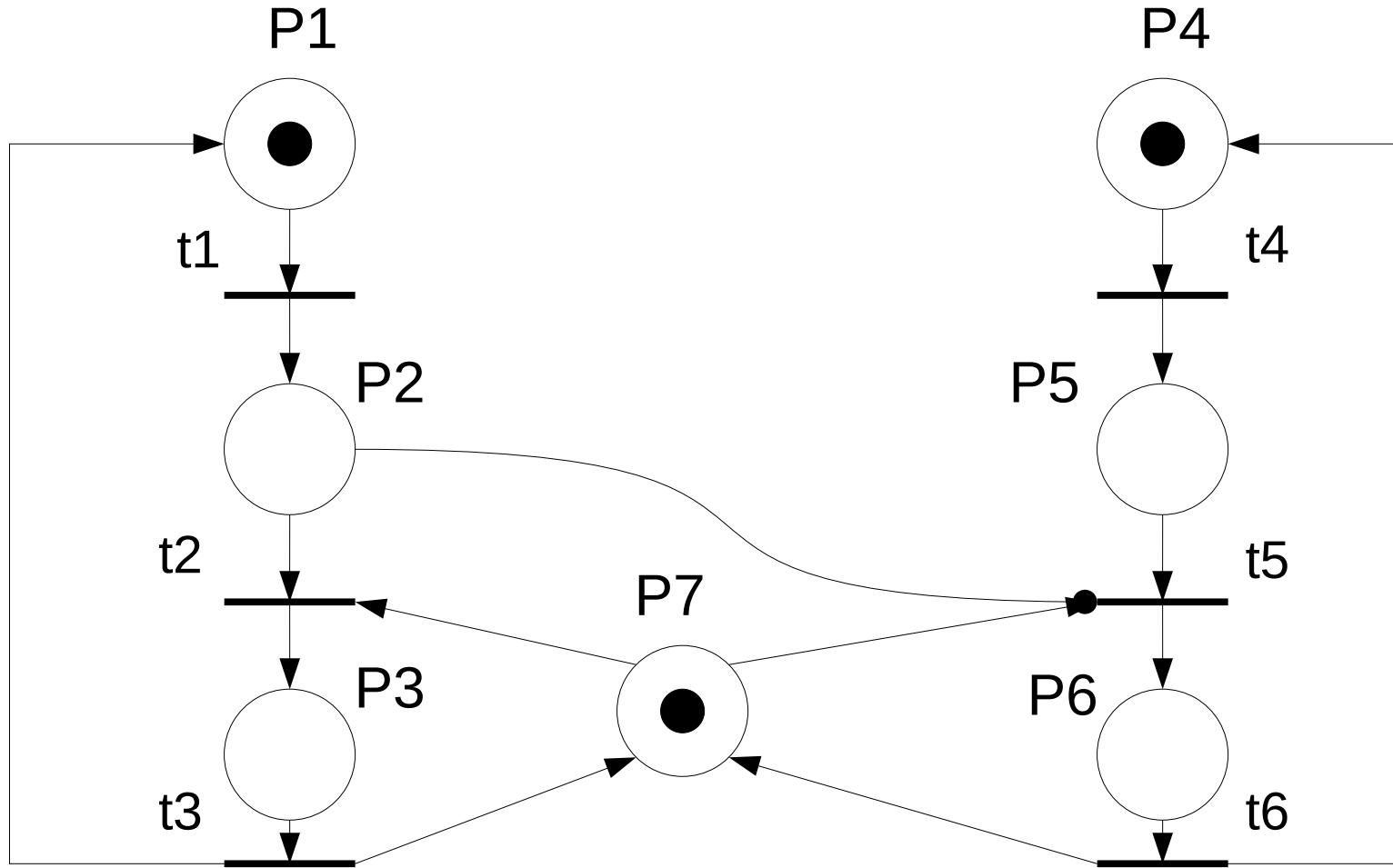
$$\forall p \in P_H(t_k) \Rightarrow M_i(p) < H(t_k, p)$$

1. each input place has at least as many tokens as the multiplicity of the input arcs
2. each inhibitor place has less tokens than the multiplicity of the corresponding arc

Example



Example – priority



Reachability graph

- M_0 is the starting state (initial marking)
- A new marking M_j is computed when an enabled transition $t_k \epsilon T$ in M_i fires
- M_j is said directly reachable from M_i

$$M_i \xrightarrow{t_k} M_j$$

Reachability graph (2)

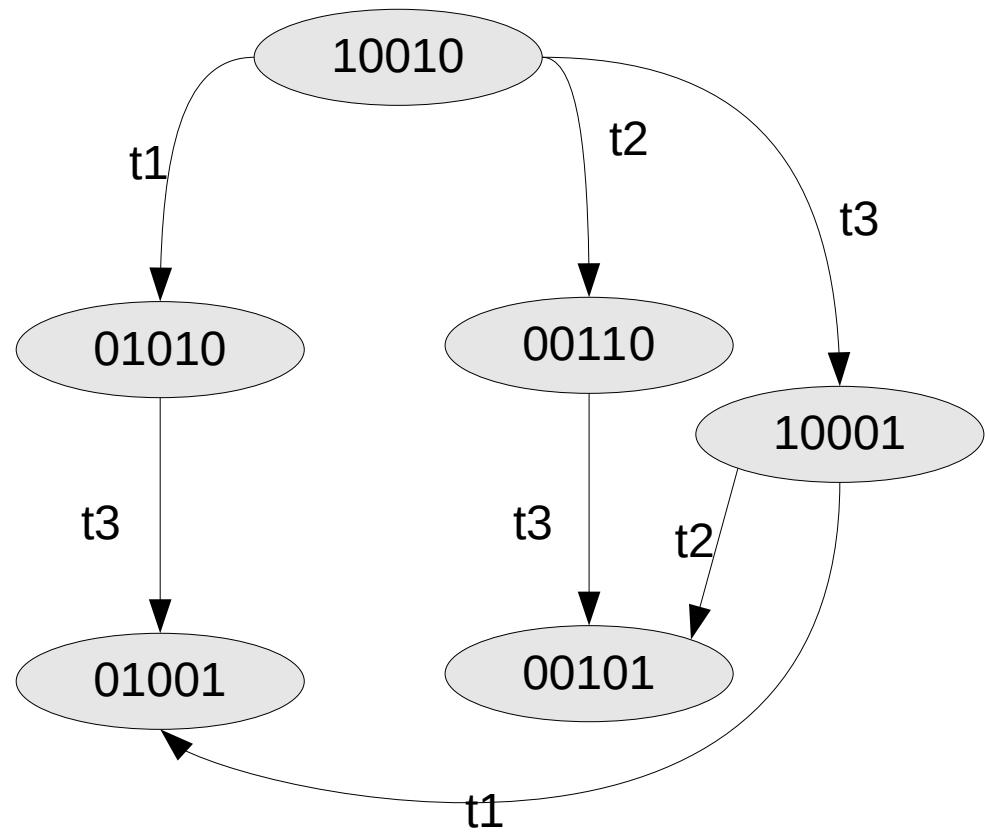
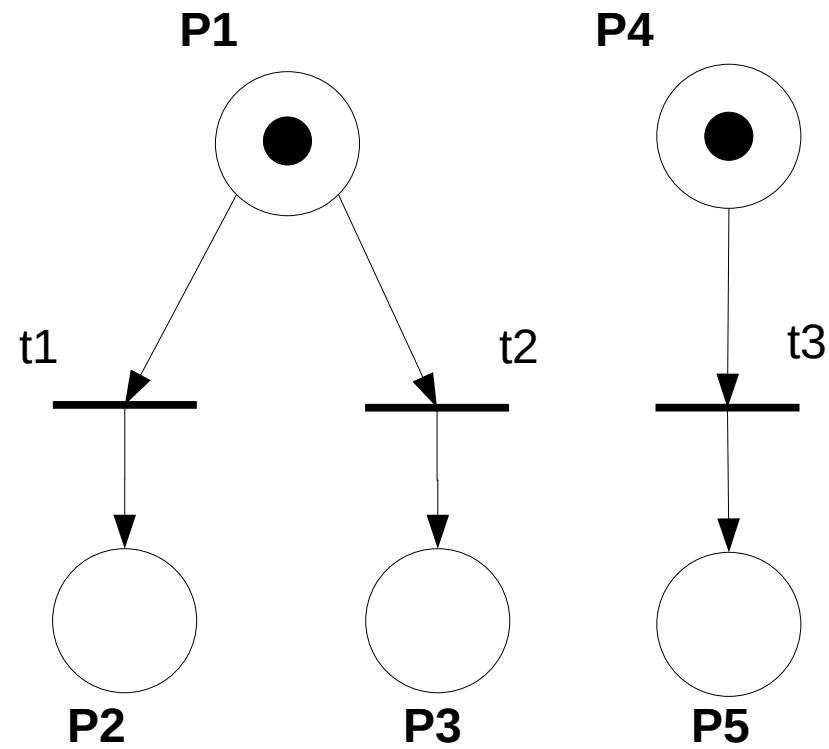
- The repeated application of the firing rule generates the reachability graph
- Petri net dynamics
- A marking M_i is said reachable iff, given M_0 , a firing sequence from M_0 to M_i exists

Reachability graph (3)

- The set of all the reachable markings is said **reachability set**
- It is denoted with $RS(M_0)$
- $RS(M_0)$ + firing events = **reachability graph**
- Firing sequence

$$\varepsilon = \{(M_1, t_1); (M_2, t_2), \dots (M_n, t_n)\}$$

Example



Timed PN

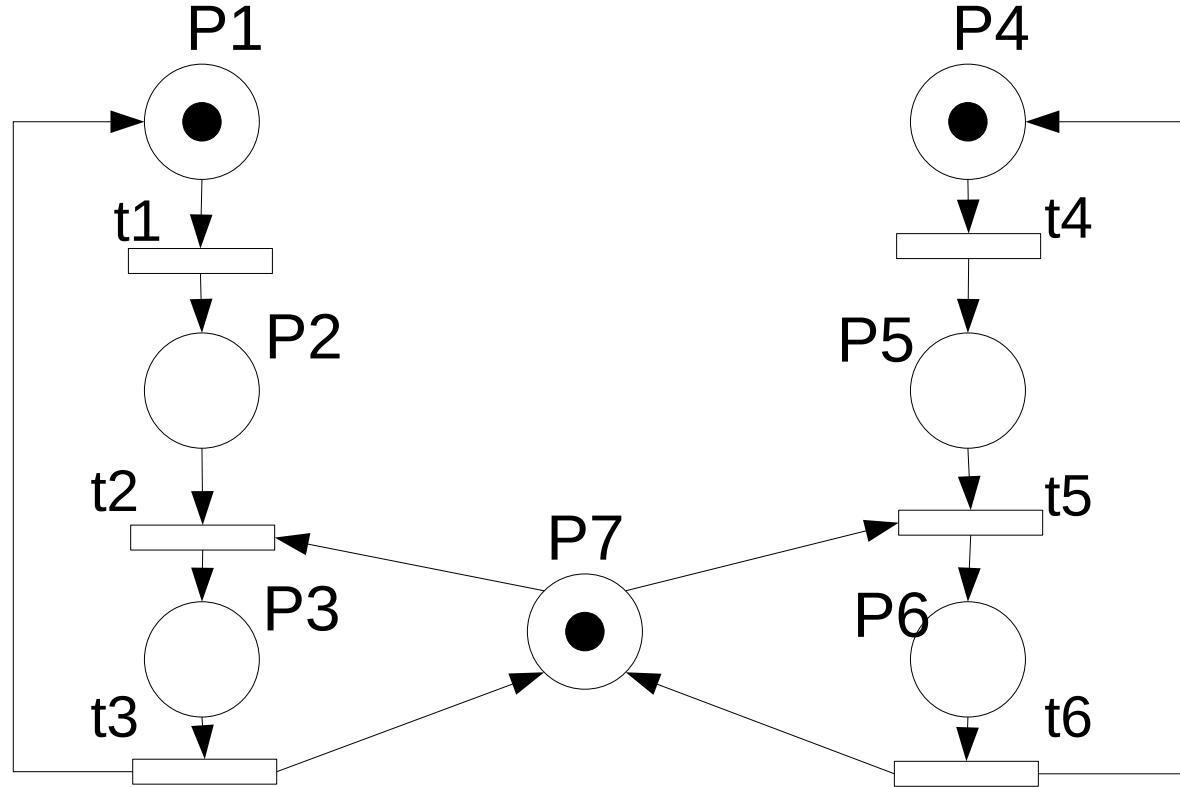
- $TPN = \{P, T, I, O, H, M_0, \Theta\}$
- $\Theta : T \rightarrow R^+$, function assigning a time delay to transitions
- Firing sequence

$$\varepsilon = \{(M_1, t_1, \tau_1), (M_2, t_2, \tau_2); \dots; (M_n, t_n, \tau_n)\}$$

Stochastic Petri nets

- Firing delays are exponentially distributed r.v.
- $SPN = \{P, T, I, O, H, M_0, L\}$
- $L: T \rightarrow R$
 - $L(t_i) = \lambda_i$, cdf's parameter
- Transitions are drawn as empty bar

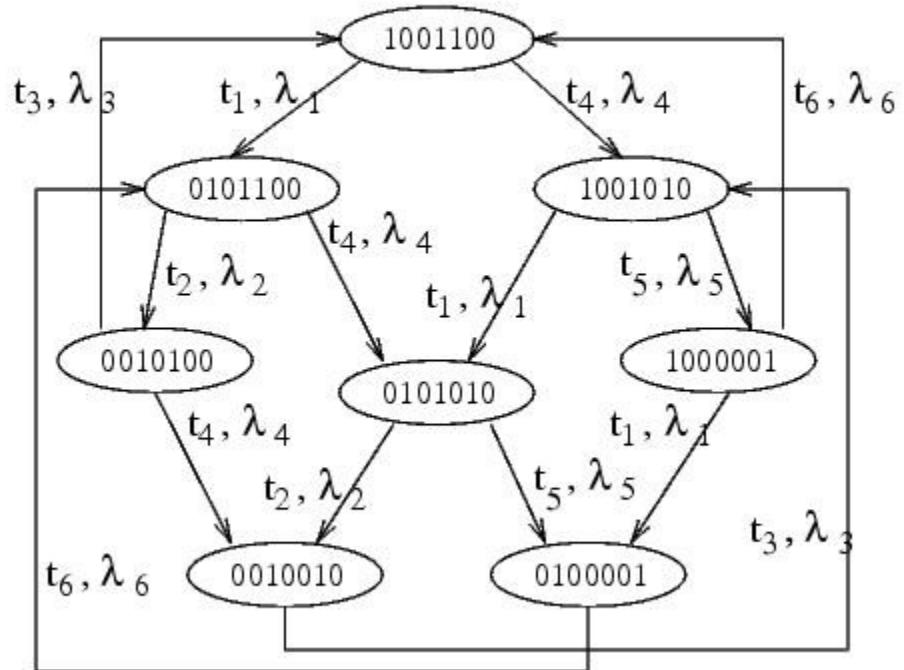
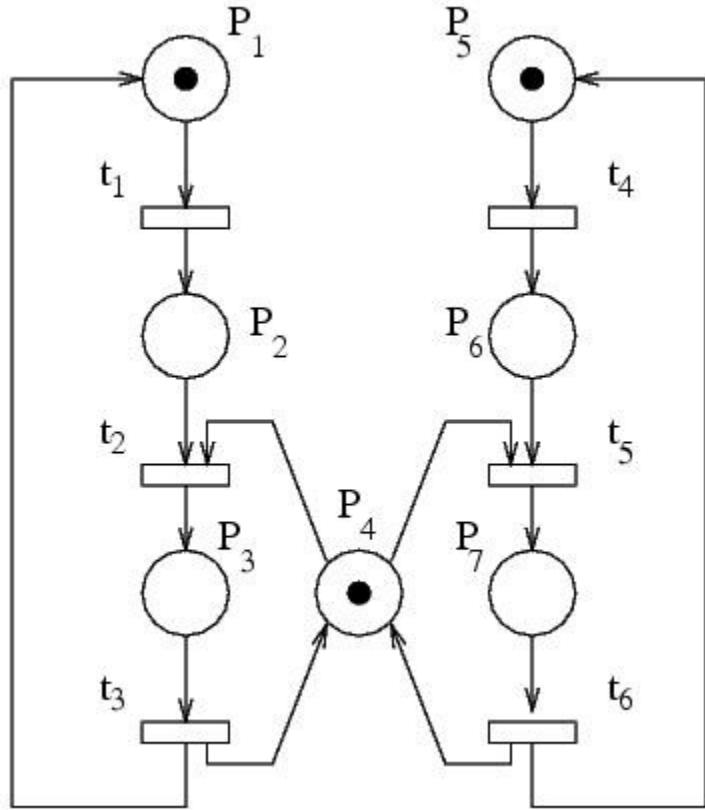
Example



Stochastic Petri nets

- Transition with shortest sampled delay fires
- After the firing, all the enabled transitions start from scratch (memoryless property)

SPN – Reachability graph



Reachability graph labeled with λ_i is a CTMC

SPN - Analysis

$$T_e^{(i)} = \{t \in T : t \text{ is enabled into } M_i\}$$

$$E_j(M_i) = \{t_k \in T_e(i) : M_i \xrightarrow{t_k} M_j\}$$

$$q_{ij} = \begin{cases} \sum_{k : t_k \in E_j(M_i)} \lambda_k & i \neq j \\ - \sum_{k : t_k \in T_e^{(i)}} \lambda_k & i = j \end{cases}$$

- State probabilities = Marking probabilities

Generalized Stochastic Petri Nets

- $GSPN = \{P, T, I, O, H, M_0, L, c\}$
- Immediate and timed transitions
- $L(t_k)$ depends on t_k
 - Timed t_k : $L(t_k) = \lambda_k$
 - Immediate t_k : $L(t_k) = w_k$
- $c: T \rightarrow N$ associates a priority to transitions

GSPN – Enabling rule

$$\forall p \in P_I(t_k) \Rightarrow M_i(p) \geq I(t_k, p)$$

$$\forall p \in P_H(t_k) \Rightarrow M_i(p) < H(t_k, p)$$

$$\forall t \in T_e^{(i)}, t \neq t_k \Rightarrow c(t) \leq c(t_k)$$

- Timed t: $c(t)=0$
- Immediate t: $c(t)>0$
 - Used to model logic events

Definitions

- A marking where some immediate transitions is enabled is said *vanishing*
 - Visit time is null
- A marking where only timed transitions are enabled is said *tangible*

Properties

- Timed and immediate transitions cannot be enabled in the same marking
- If a timed transition is enabled in a marking, all the enabled transitions are timed

Transition firing

- Probabilistic choice

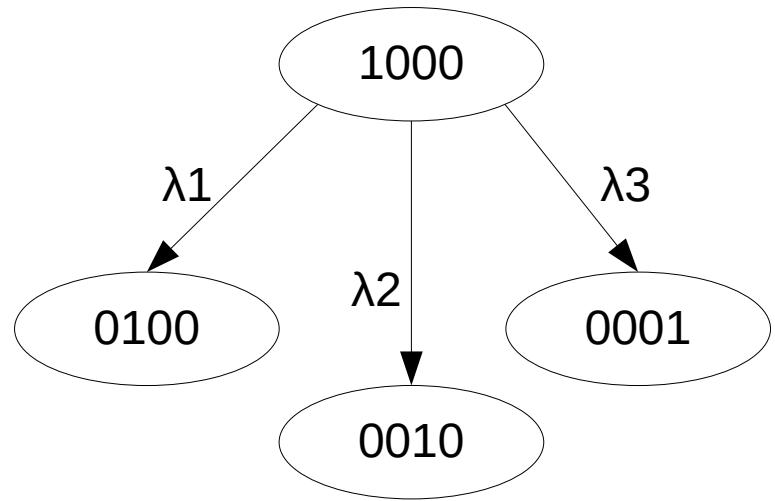
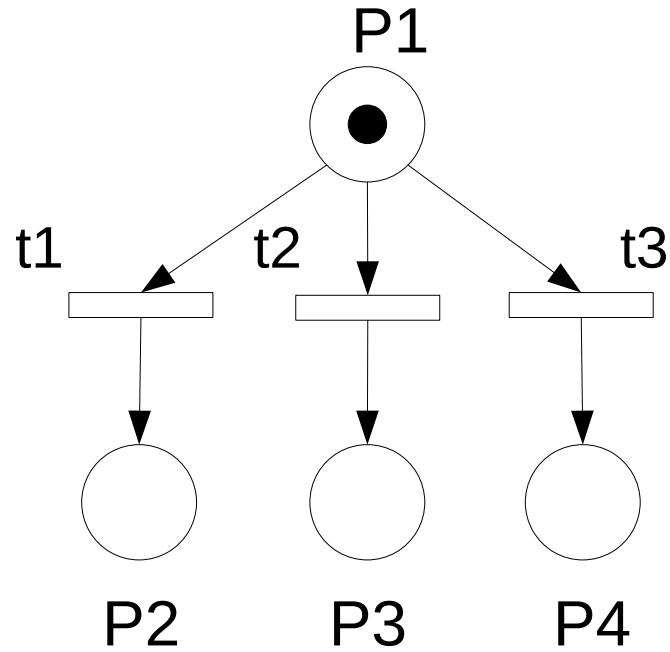
$$s_k^{(i)} = P\{t_k \text{ fires} | M(t) = M_i\} = \frac{w_k}{\sum_{t \in T_e^{(i)}} w_t}$$

Transition enabling

- Three cases can arise in a marking M
 - 1) Timed transitions are enabled
 - 2) One immediate transition is enabled
 - 3) A lot of immediate transitions are enabled

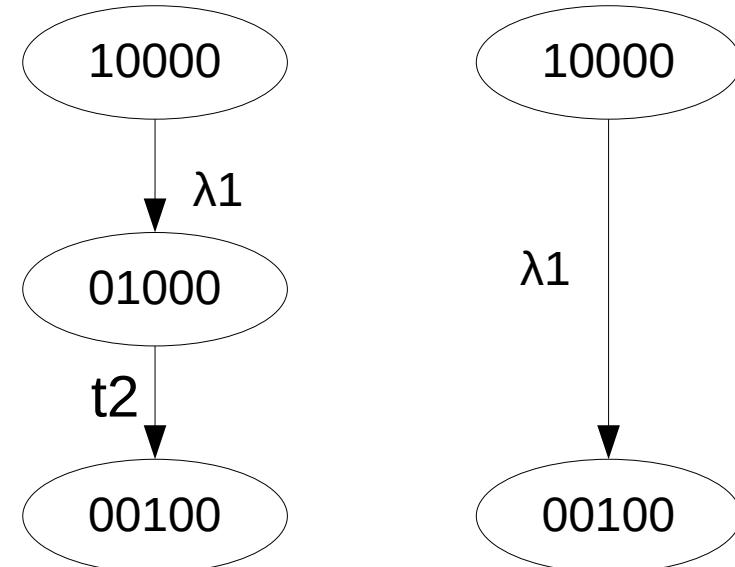
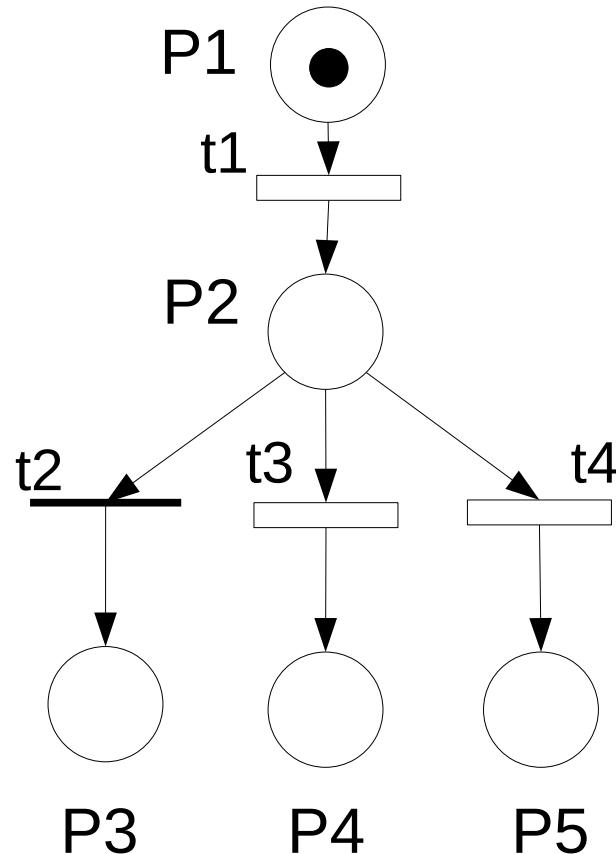
Case 1)

- M is a tangible marking



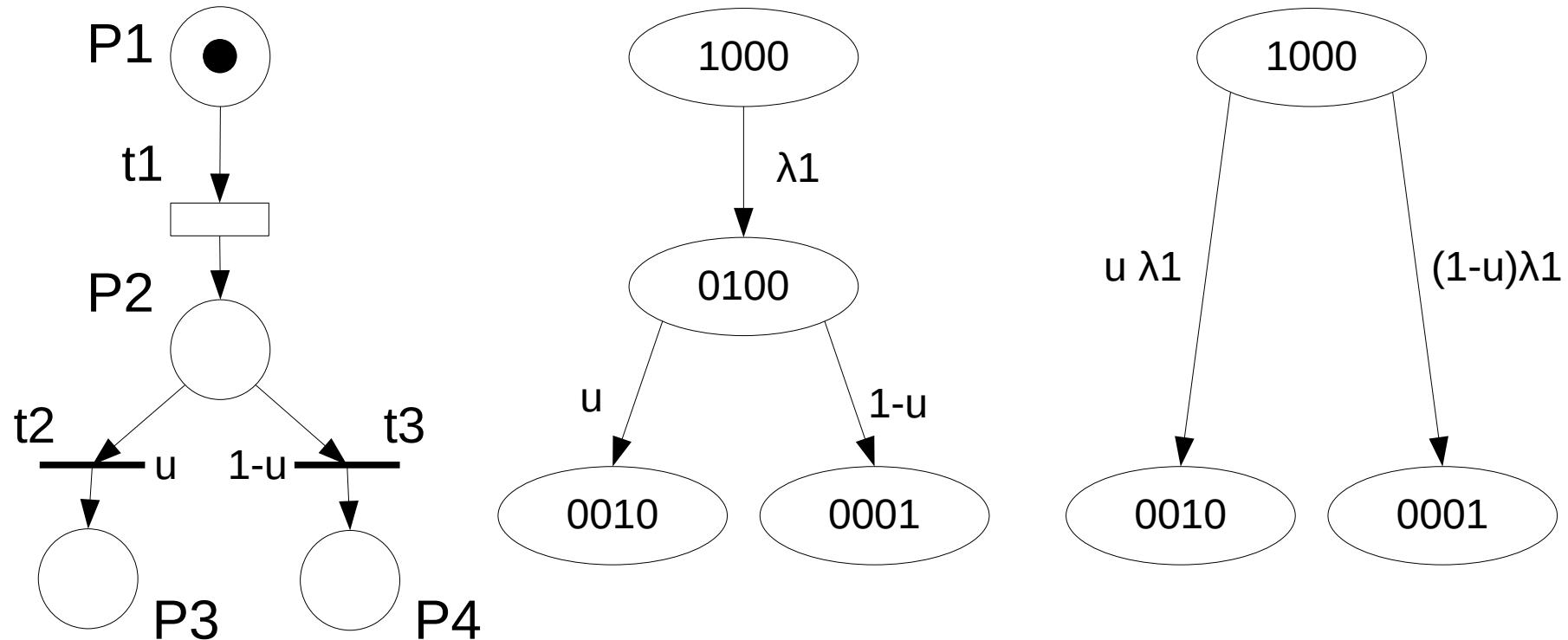
Case 2)

- One immediate transition is enabled



Case 3)

- Many immediate transitions are enabled



- Problem: cycles of vanishing markings

Analysis

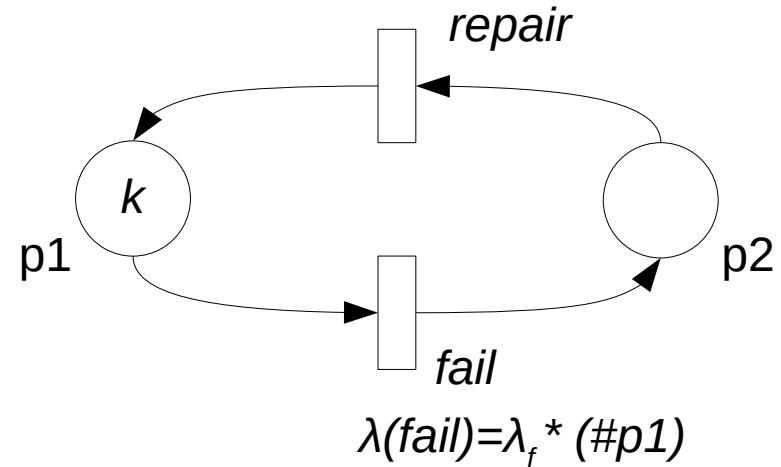
- All the vanishing markings are erased
- Reduced reachability graph
- Some information is lost
- The resulting stochastic process is a CTMC

Marking dependent transitions

- Transition rates may depend on marking
- $L(t_k) = \lambda_k(M)$

example

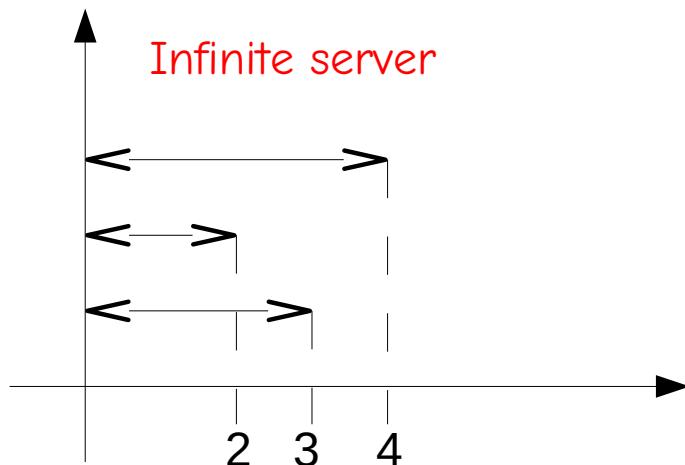
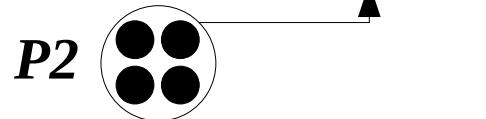
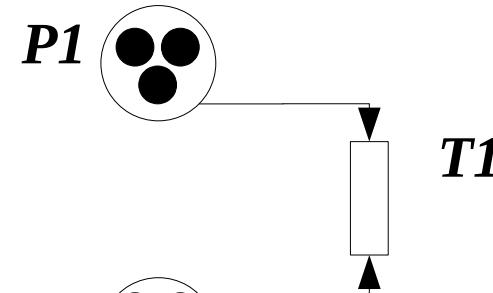
k equal redundant components subject to failures
Only one repair resource



Multiple enabling

- Different semantics can be assumed when more than one token is in a place
 - **single server**: tokens generate serial transitions enabling
 - **infinite-server**: different timers are associated to each token; parallel evolution
 - **multiple server**: tokens are processed in parallel with parallel degree equal to k

Example



Sampled delays:

- $A_1 \rightarrow 3$
- $A_2 \rightarrow 2$
- $A_3 \rightarrow 4$

