

Comment observer de l'intérieur un système informatique réparti ?

Claude JARD

Conférence MMI2

9 septembre 2003

Le réseau mondial est en marche...
avec des applications réparties



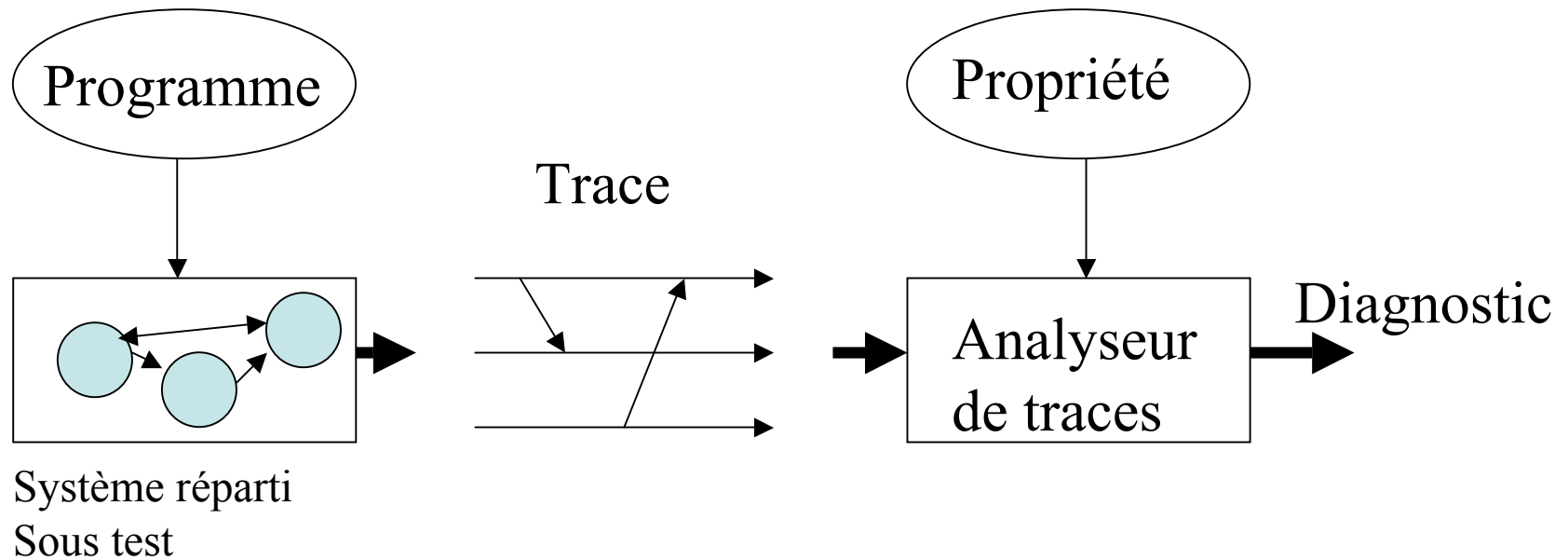
Observation des comportements des systèmes répartis

- Vérification de traces
- Causalité induite par les messages entre les événements observables
- Calcul "au vol" de la causalité
- Notion d'état global et le treillis des idéaux
- Vérification (centralisée)
- Vérification répartie

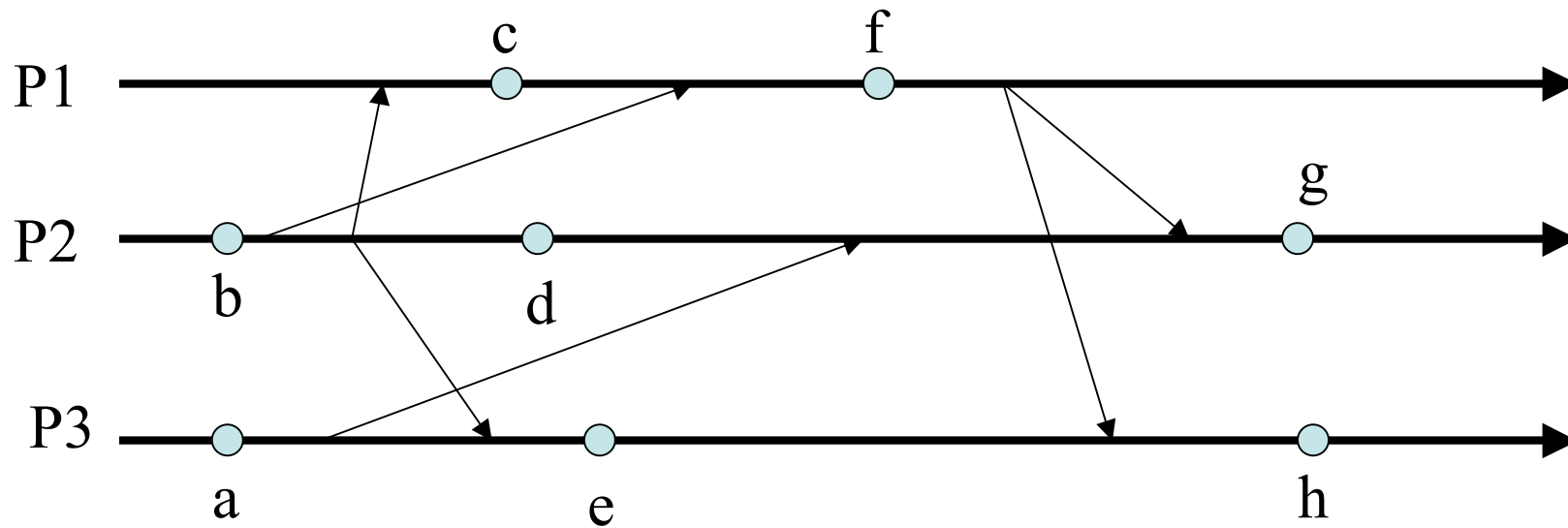
Vérification de traces

T (trace d'exécution répartie) $\models P$ (prédicat)

Applications en diagnostic, débogage, ...



Les comportements vus comme des chronogrammes

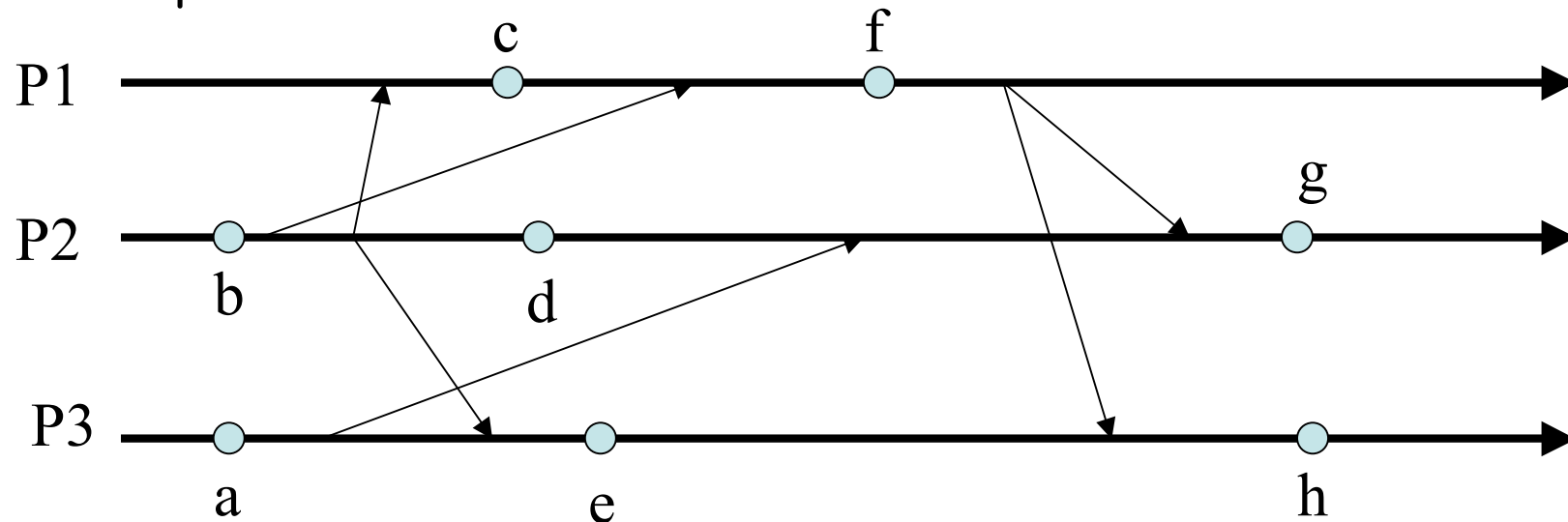


Méthode

- Les programmes sont instrumentés par des sondes qui tracent, pendant l'exécution les événements définis comme observables
- La causalité induite par les échanges de messages est capturée au vol par une technique "d'estampillage" (notion de "temps réparti")
- La détection de propriétés peut être :
 - centralisée (fondée sur le "treillis des idéaux")
 - elle-même répartie

La causalité entre les evts observables, induite par les échanges de messages

- N processus séquentiels (P1 à Pn)
- Les processus effectuent des actions pendant leur exécution. Quelques unes sont définies comme des événements (atomiques)
- La communication synchronise (partiellement) l'activité des processus



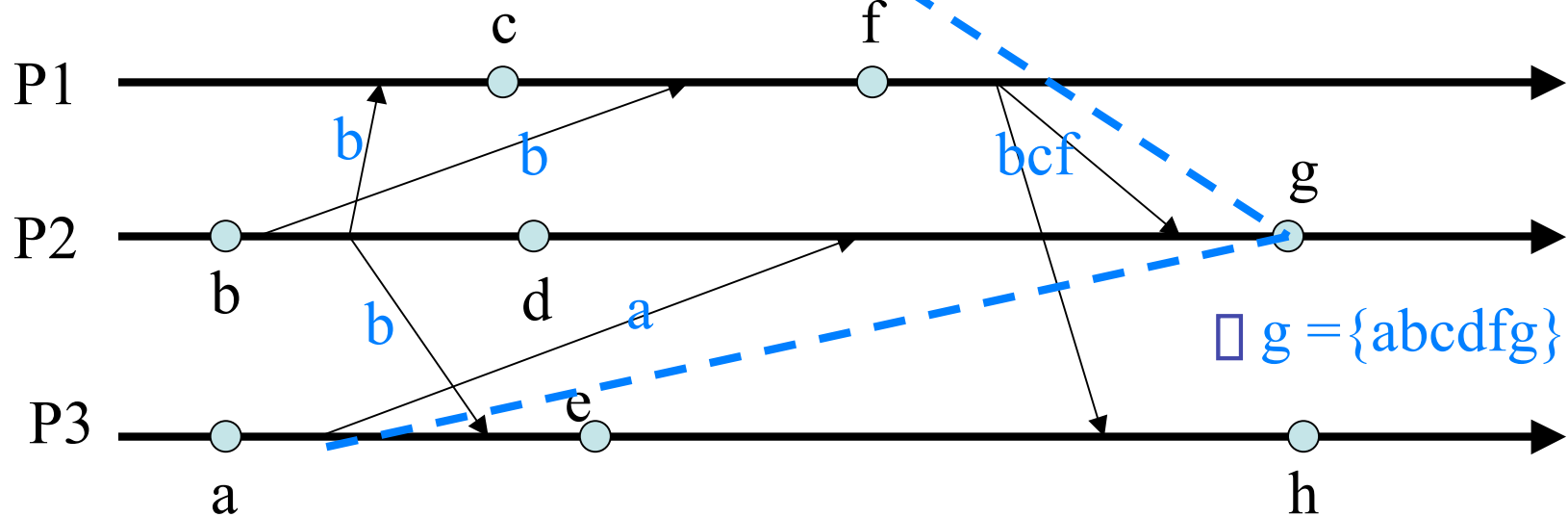
Calcul au vol de la relation de causalité

- Causalité (Lamport 1978) :

$x \leq y$ ssi il existe un chemin de x vers y

- $\square x = \{y \mid y \leq x\}$

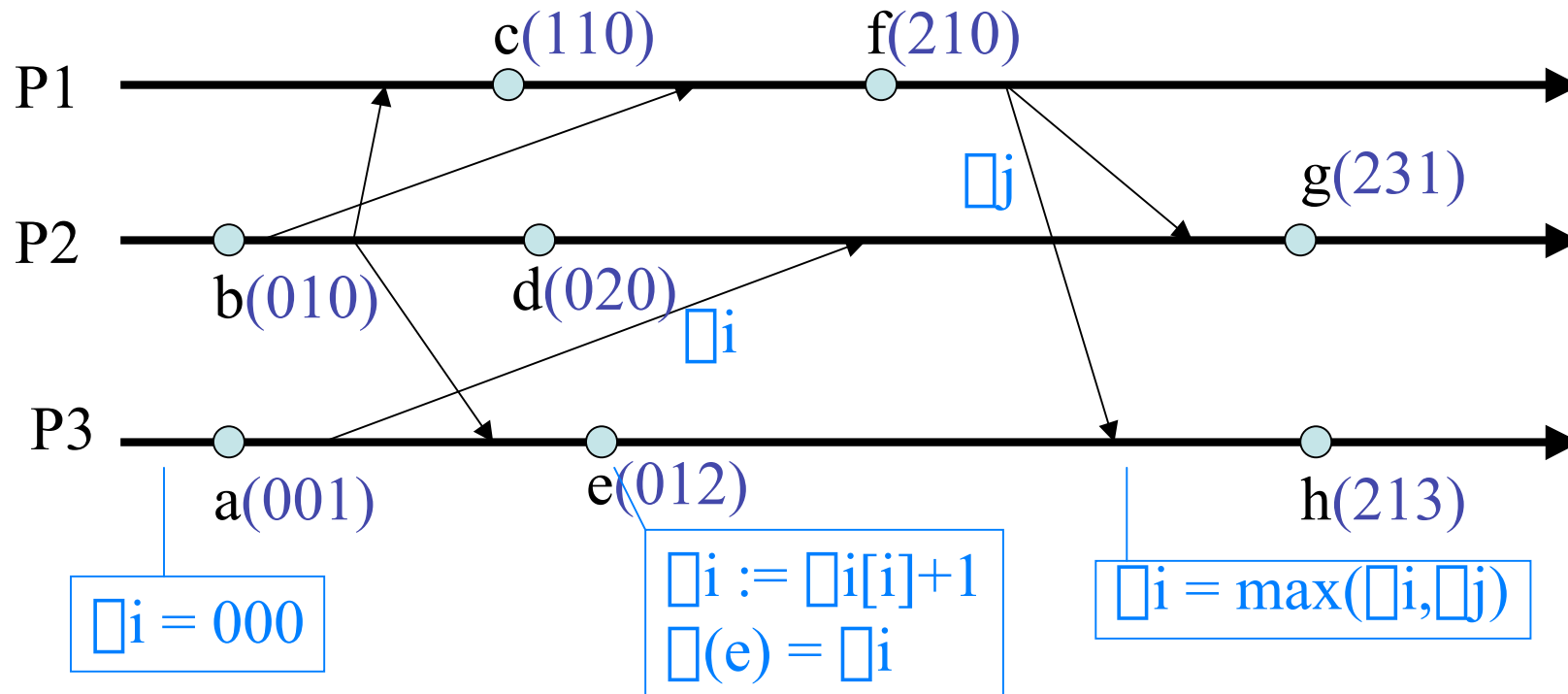
$x \leq y \iff \square x \ni y$



L'estampillage vectoriel (Fidge-Mattern 1988)

$\square(x)[i] =$

nombre d'evts observables de P_i dans le passé causal de x



Modèle mathématique : les ordres partiels

- Forme canonique
- Dirigée de bas en haut
- Réduction transitive

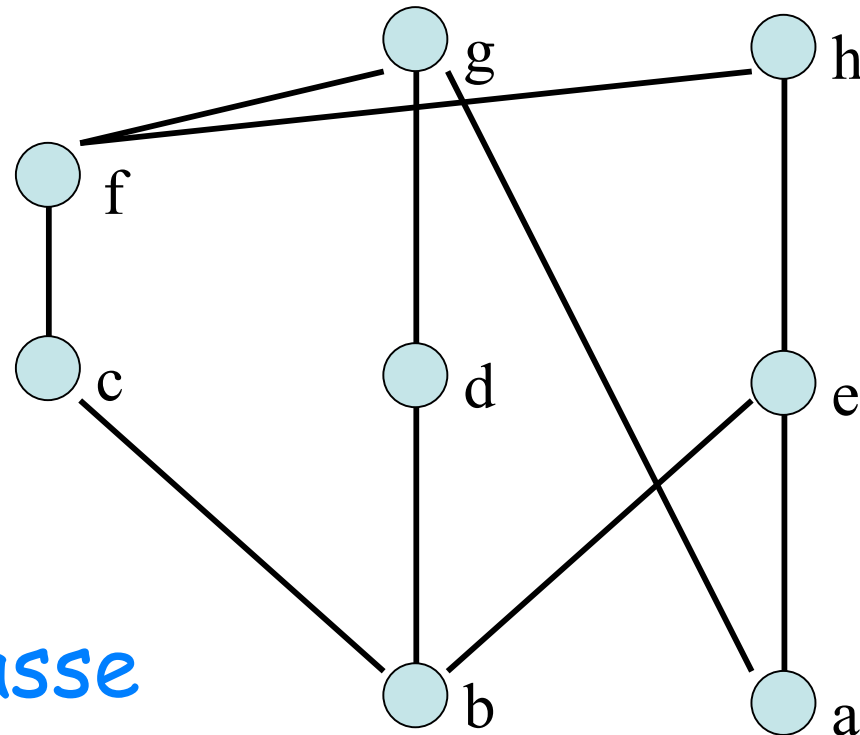
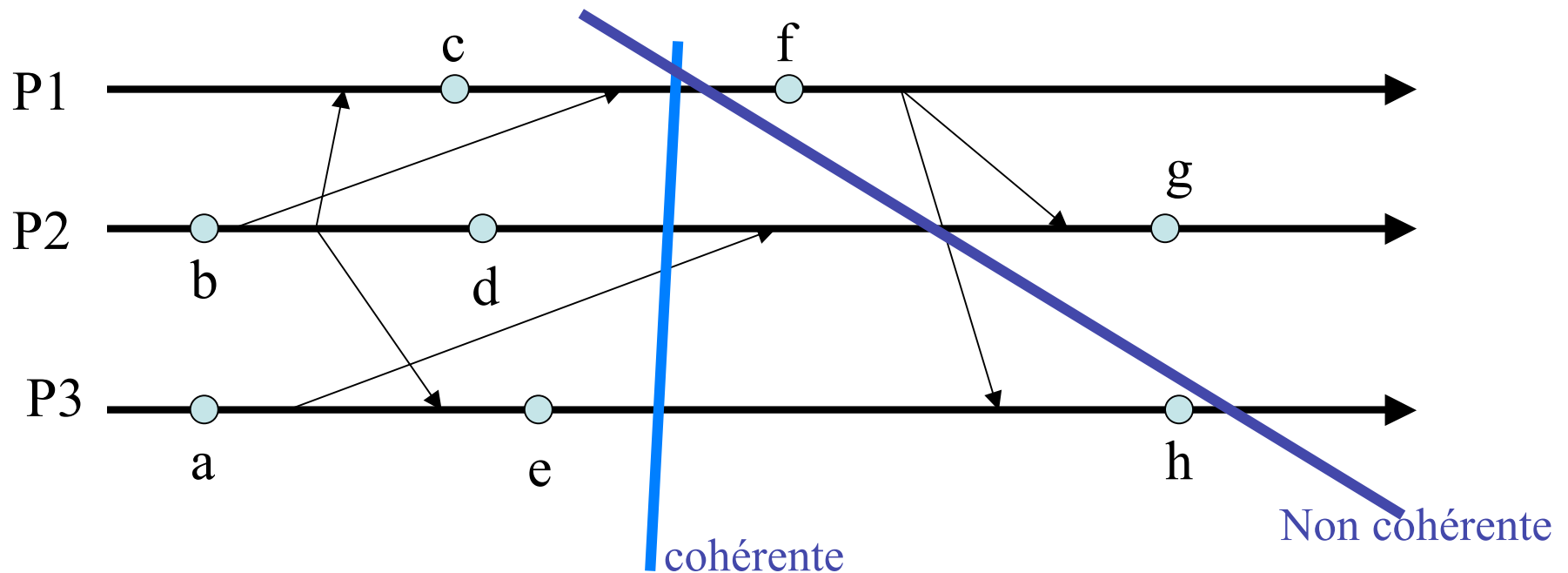


Diagramme de Hasse

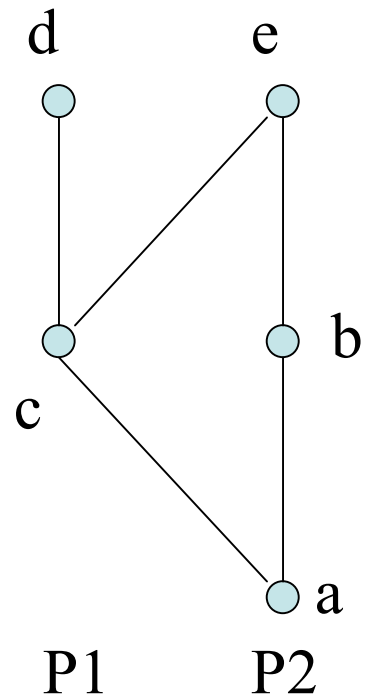
La notion d'état (coupures cohérentes)

- Comme d'habitude, l'état code le passé (coupure)
- Seules les coupures cohérentes sont accessibles
- Les états sont des idéaux : $X \sqcap O / \sqcap X = X$



Le treillis des états

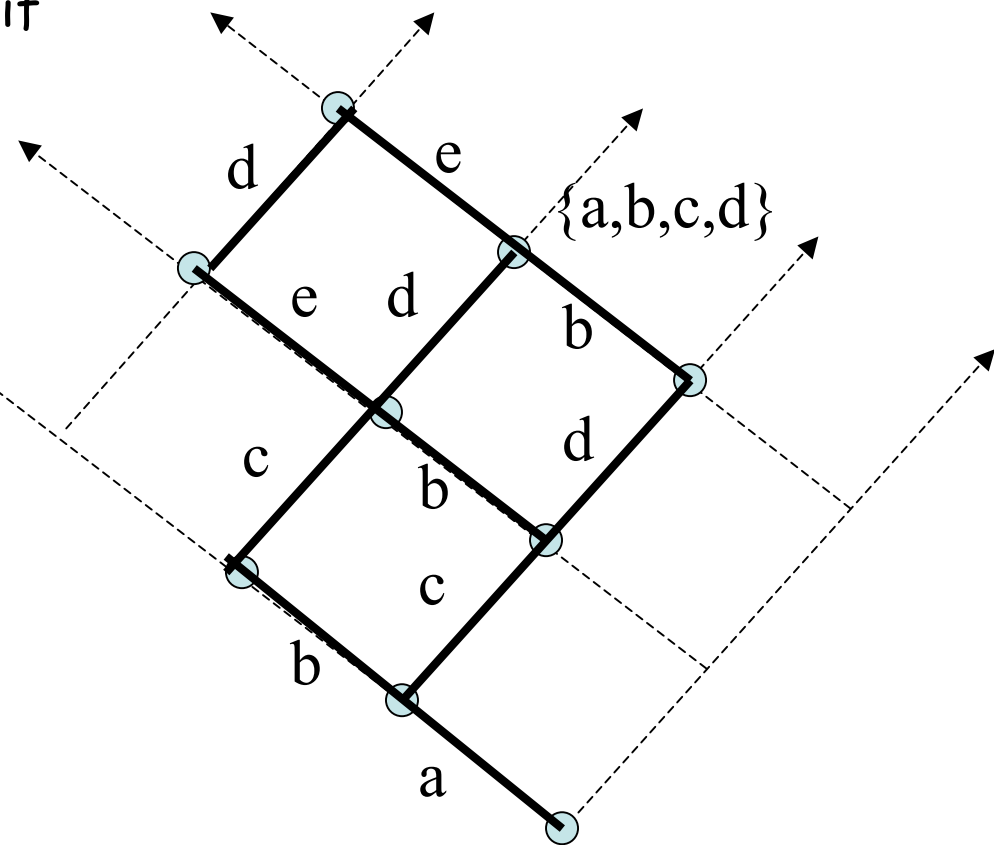
- L'ensemble des idéaux ordonné par l'inclusion ensembliste forme un treillis distributif



Diagramme

$\{\}$
 $\{a\}$
 $\{a,b\}$
 $\{a,c\}$
 $\{a,c,d\}$
 $\{a,b,c\}$
 $\{a,b,c,d\}$
 $\{a,b,c,e\}$
 $\{a,b,c,d,e\}$

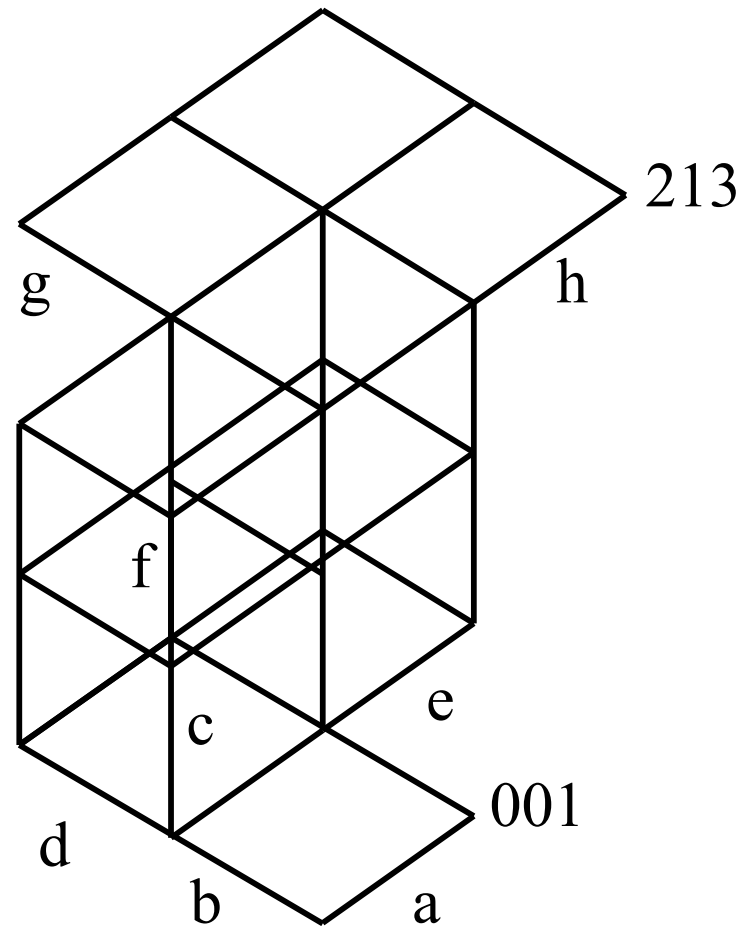
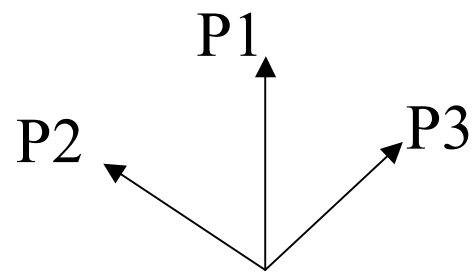
Idéaux



Graphe des états

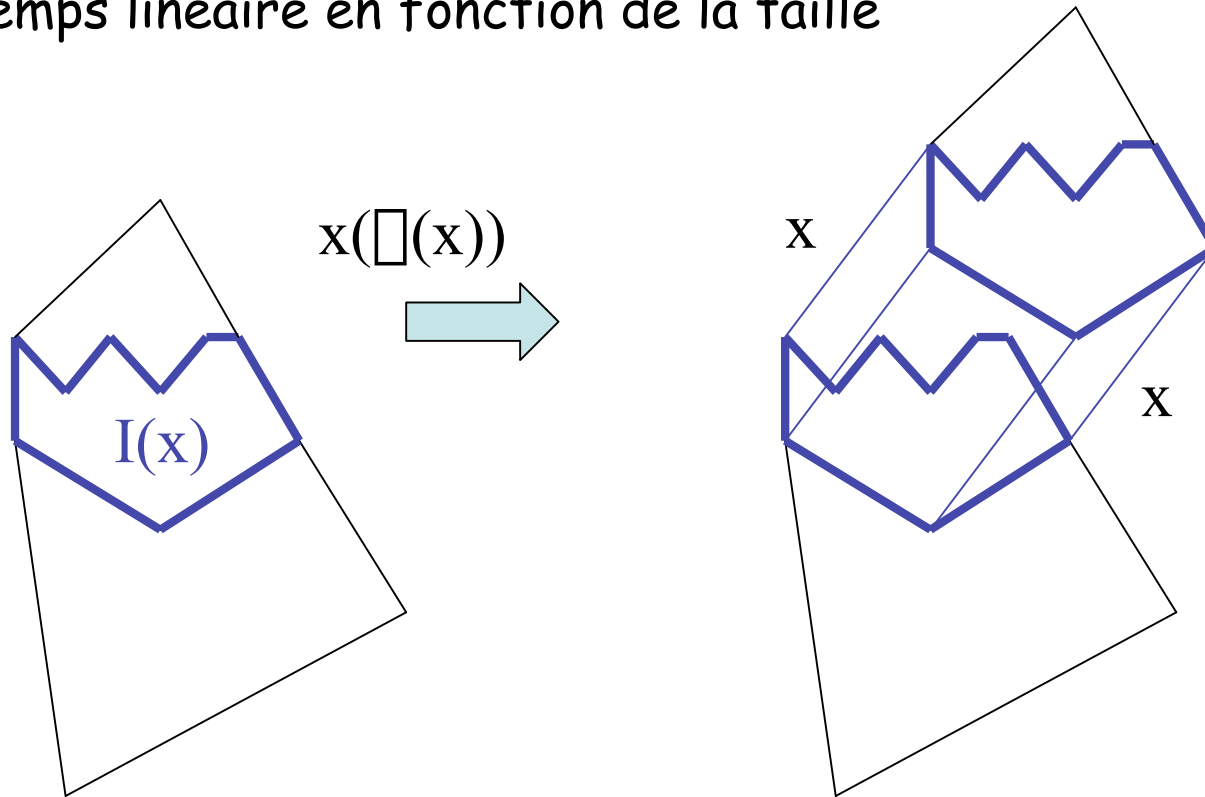
Son plongement dans une grille N-dimensionnelle

- Les estampilles vectorielles forment les coordonnées des états

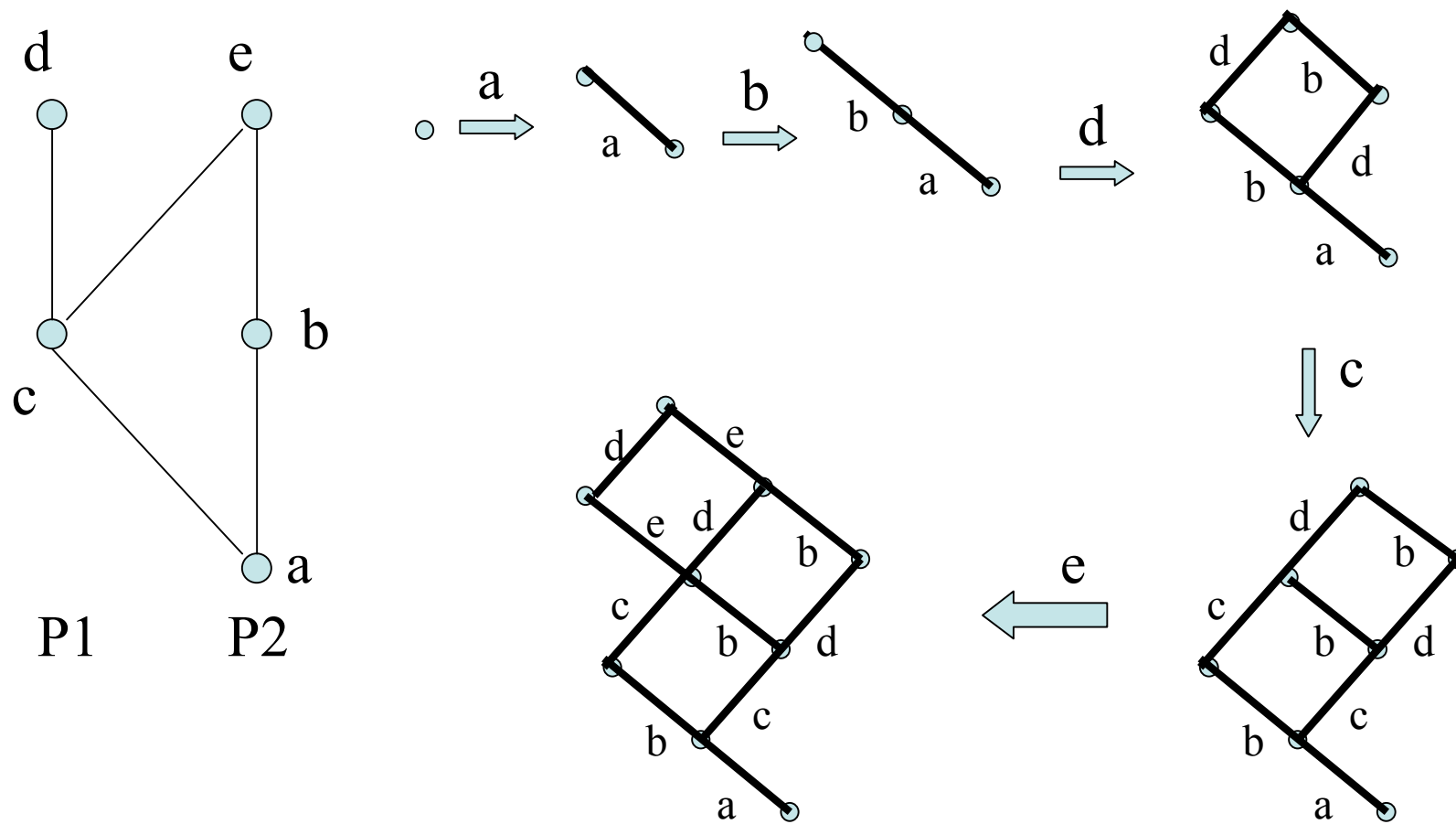


Construction au vol du graphe d'état (Jard-Rampon 1992)

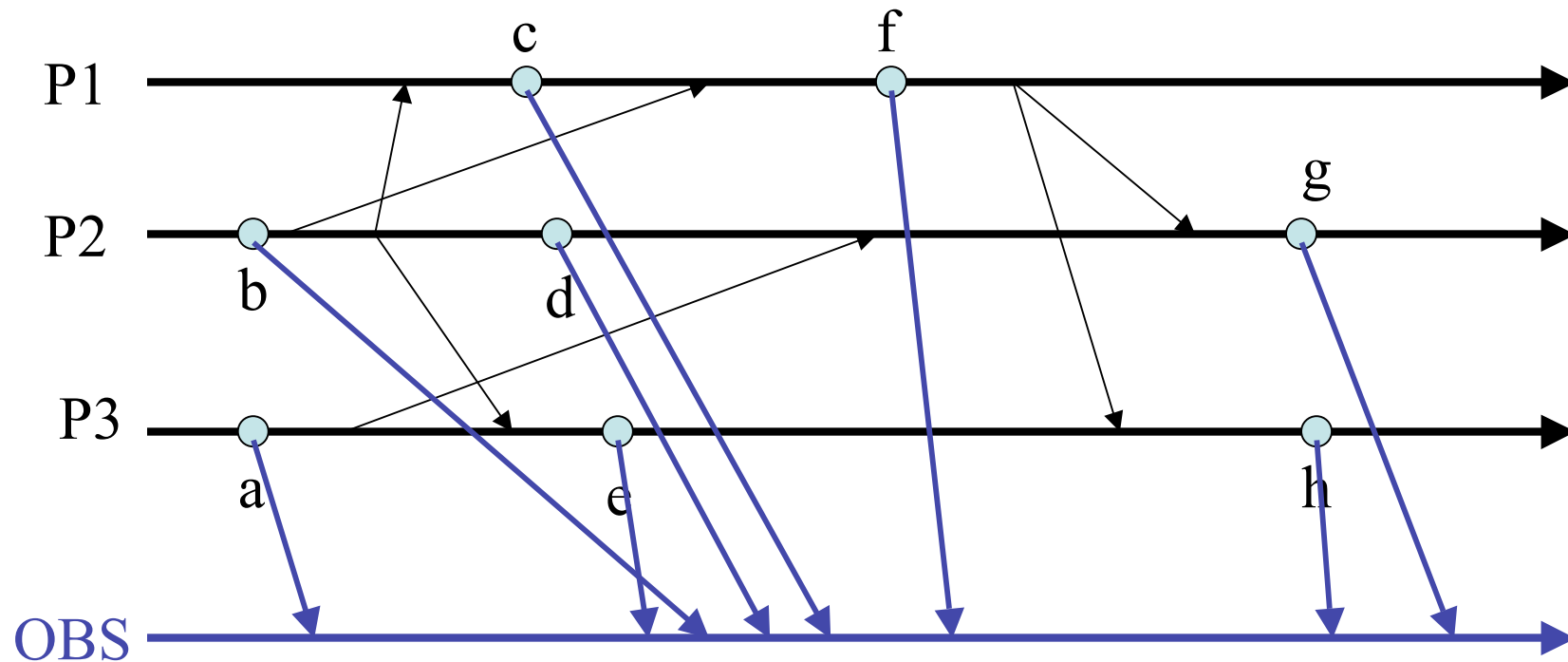
- En temps linéaire en fonction de la taille



Un petit exemple de croissance



La vérification centralisée



Observation causale : retard à la détection

Détection de propriétés régulières

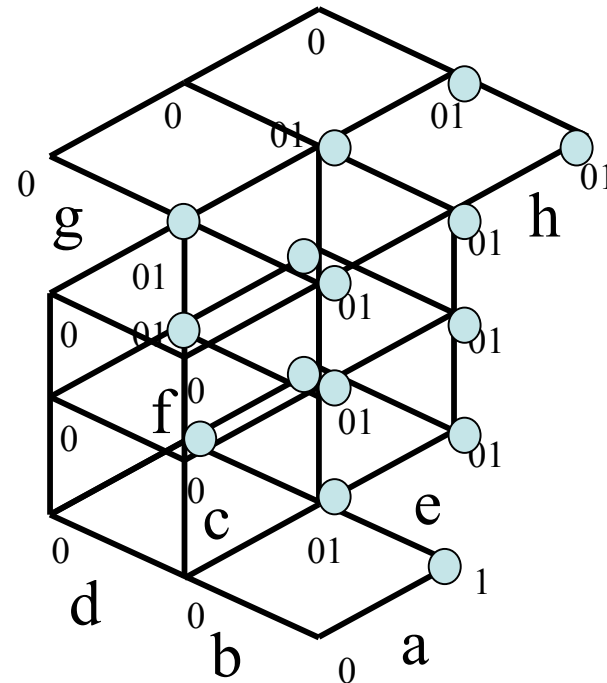
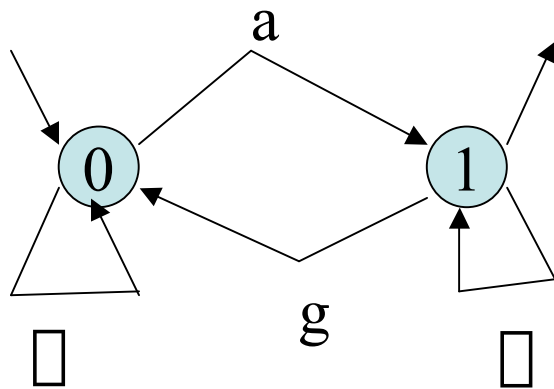
- **Propriété** $\square = \langle \Sigma, Q, q_0, F, \delta \rangle$ (automate),
définit un langage $L(\square) = \{u \in \Sigma^* \mid \delta^*(u, q_0) \in F\}$
- **Graphe d'état** : les transitions sont étiquetées dans Σ ,
pour un état I , $\text{Chemins}(I)$ est l'ensemble des mots menant à I :
 I satisfait \square ssi $\text{Chemins}(I)$ intersecte $L(\square)$
- I satisfait \square ssi $\square(I)$ intersecte F
où $\square(I)$ l'ensemble des états de l'automate \square accédés par les chemins se terminant en I
- **La trace satisfait \square ssi l'état maximal du treillis satisfait \square**

Illustration de l'algorithme de vérification

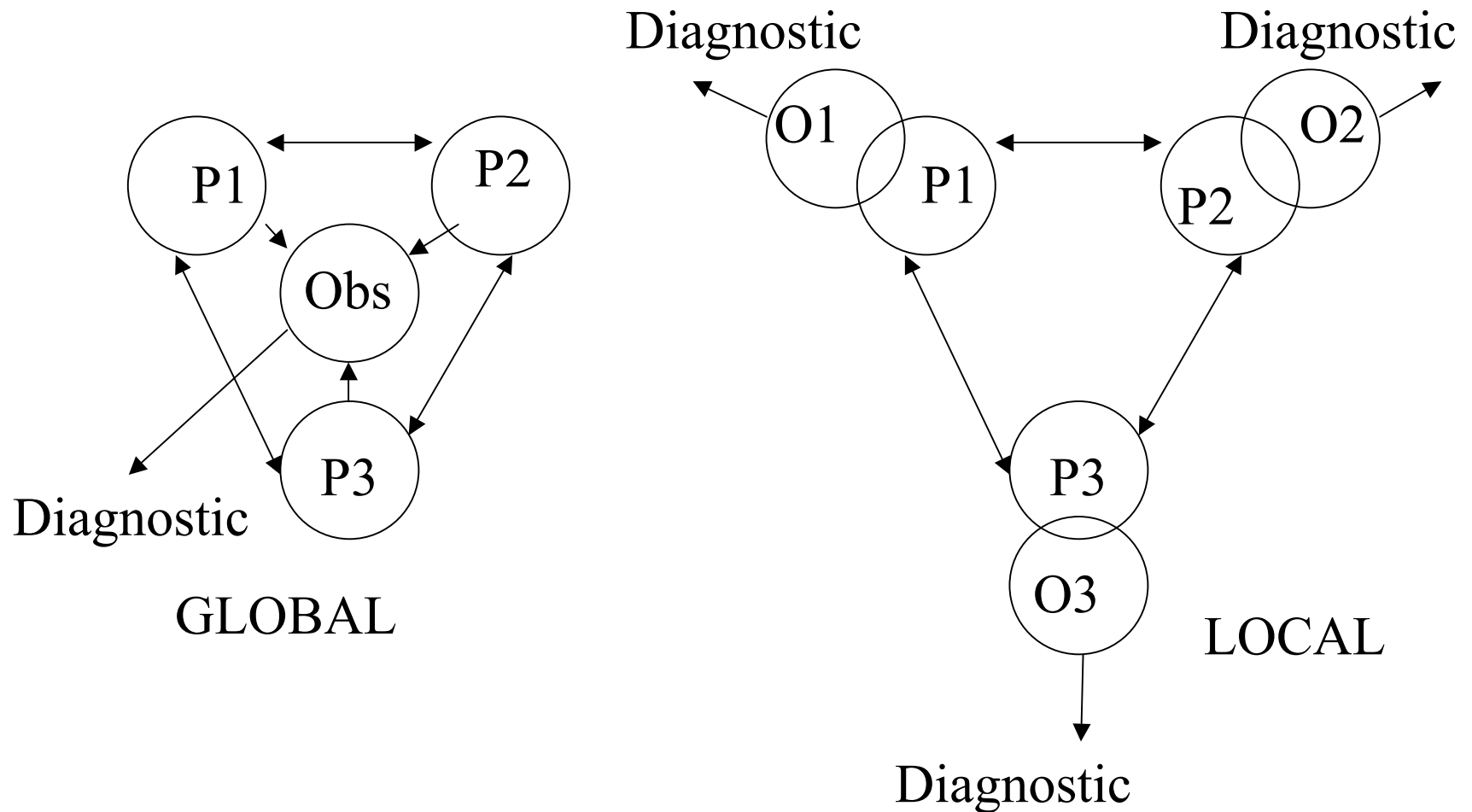
$\square(I)$ peut être calculé incrémentalement

$$\square(I) = \{ \square(\text{label}(J \rightarrow I), q) \text{ pour } q \in \square(J) \text{ et } J \in \square(I) \}$$

La détection est réalisée au vol pendant la croissance du treillis



Pour une détection répartie ("mettre de l'intelligence dans les capteurs")



Principe de la détection locale (Jard-Raynal 1995)

- Les propriétés sont interprétées sur le passé causal des événements observables
- Un evt observable x satisfait une propriété \square ssi il existe un chemin d'extrémité x dans le diagramme de Hasse de l'exécution tel le mot défini par le chemin appartienne au langage de \square
- Par exemple, on peut détecter la violation de propriétés d'ordonnancement causal
- Comme précédemment, la détection peut être effectuée au vol, mais de plus, de façon complètement répartie

Principe de l'algorithme

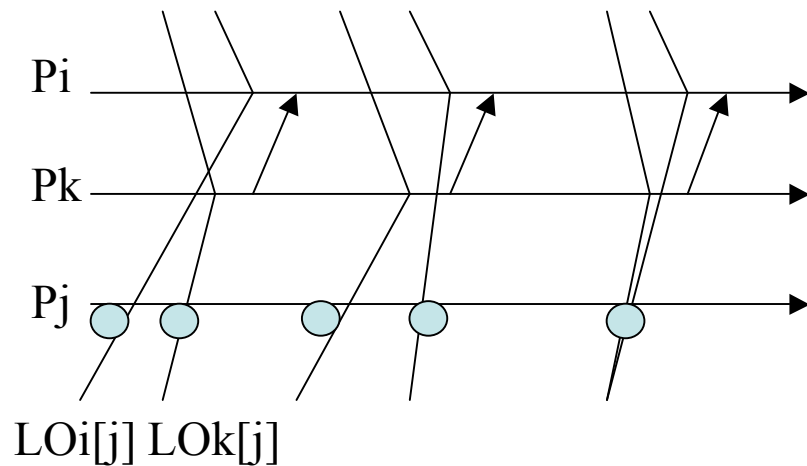
- L'automate \square est connu de chaque processus
- x satisfait \square est calculé localement par le processus qui produit x
- L'information locale est acquise depuis l'extérieur par un mécanisme d'estampillage généralisé
- Chaque processus P_i maintient deux tableaux :
LOi[1..n] et SLOi[1..n]. LOi[j] désigne le dernier evt observable de P_j dans le passé courant de P_i . SLOi[j] est l'information d'état (de \square) correspondante (quand LOi[j] est maximal)

Algorithme du processus P_i (I)

- **Data** : $LO_i[1..n]$ of integer; $SLO_i[1..n]$ of set of states
- **Init** : forall j , $LO_i[j] := 0$; $SLO_i[j] := \{q_0\}$
- **Upon observation of event x** :
 - $LO_i[i] := LO_i[i] + 1$;
 - $SLO_i[i] := \{ \langle q, x \rangle \}$ forall k , forall q in $SLO_i[k]$;
 - forall $j \neq i$, $SLO_i[j] := \{ \}$
- **When sending a message to P_k** :
 - LO_i et SLO_i sont ajoutés au message

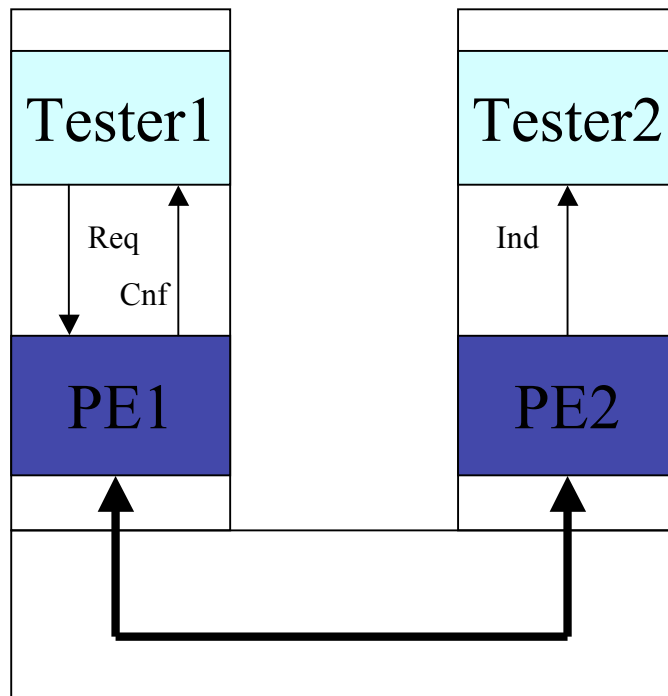
Algorithme du processus P_i (II)

- Upon reception of $\text{msg}(\text{LOk}, \text{SLOk})$ from P_k :
 - forall j , if
 - $\text{LOi}[j] < \text{LOk}[j]$ then $\text{SLOi}[j] := \text{SLOk}[j]$; $\text{LOi}[j] := \text{LOk}[j]$
 - $\text{LOi}[j] > \text{LOk}[j]$ then skip
 - $\text{LOi}[j] = \text{LOk}[j]$ then
 - if $\text{SLOi}[j] \neq \{\}$ and $\text{SLOk}[j] = \{\}$ then $\text{SLOi}[j] := \{\}$

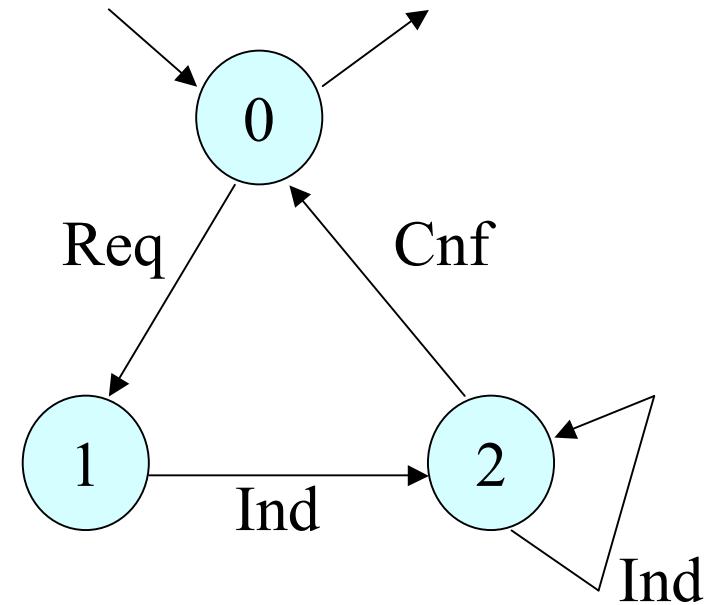


$\text{SLOi}[j]$	$\{\}$	X	$\{\}$	X
$\text{SLOk}[j]$	$\{\}$	X	X	$\{\}$
New $\text{SLOi}[j]$	$\{\}$	X	$\{\}$	$\{\}$

Exemple: a petit protocole de communication

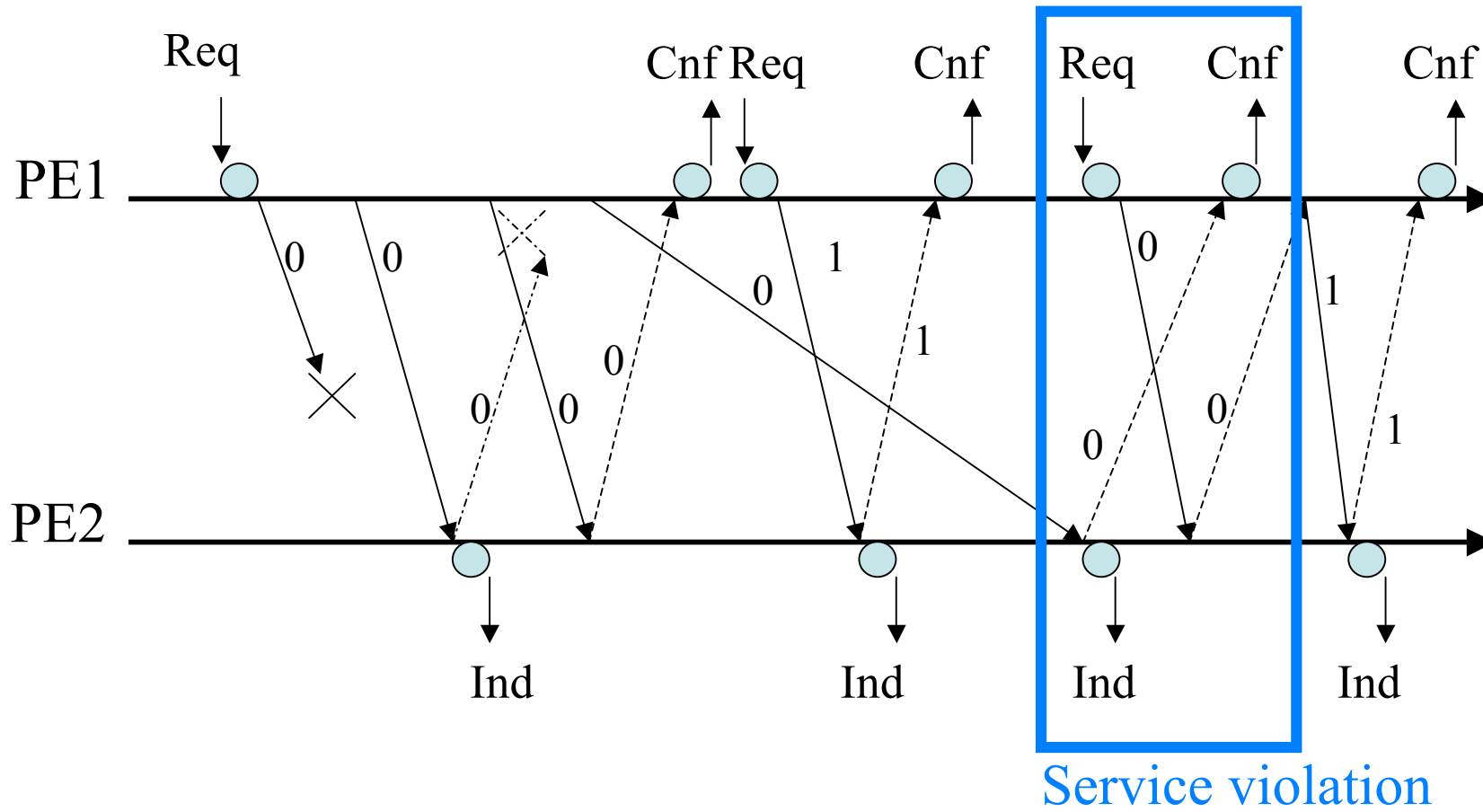


Test architecture

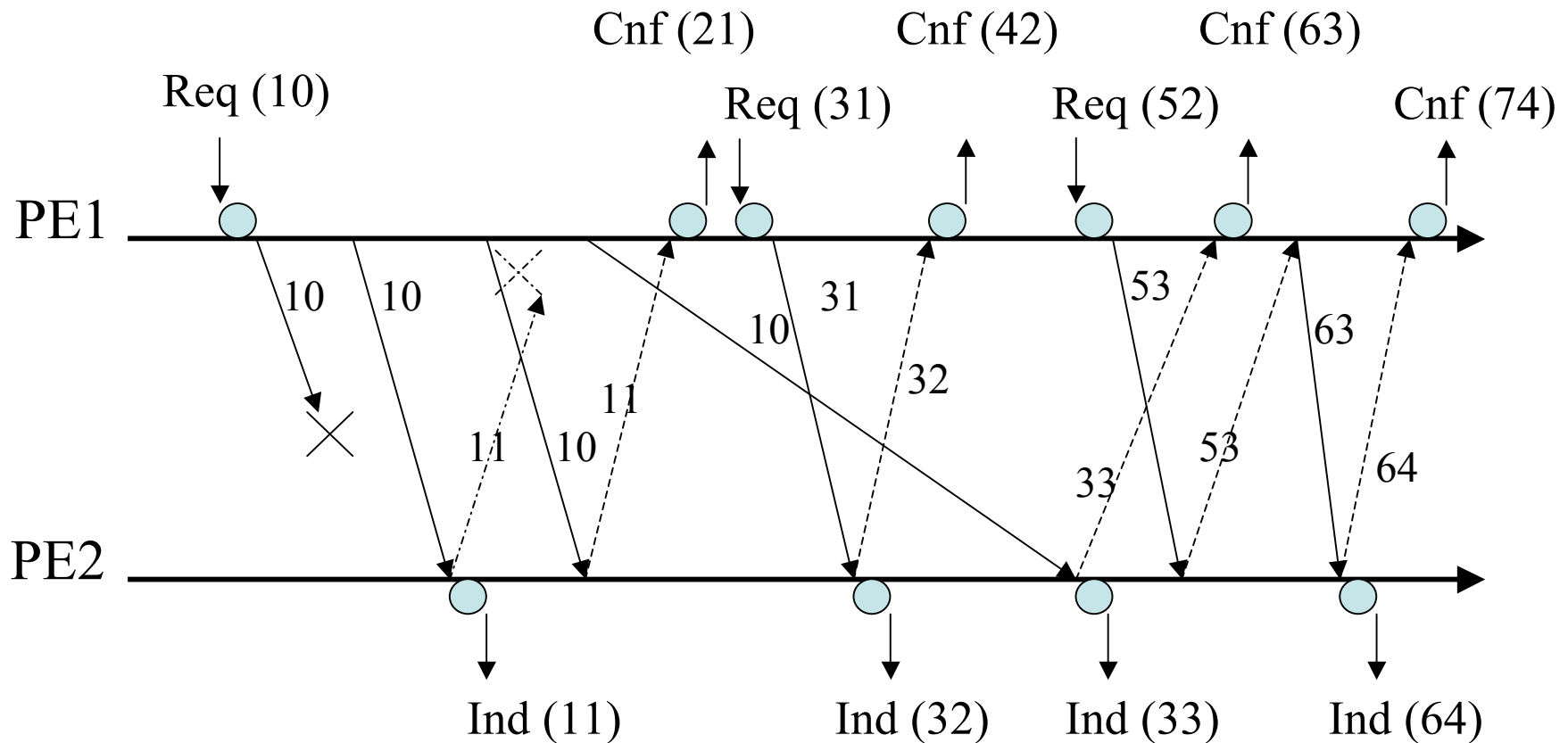


Service property

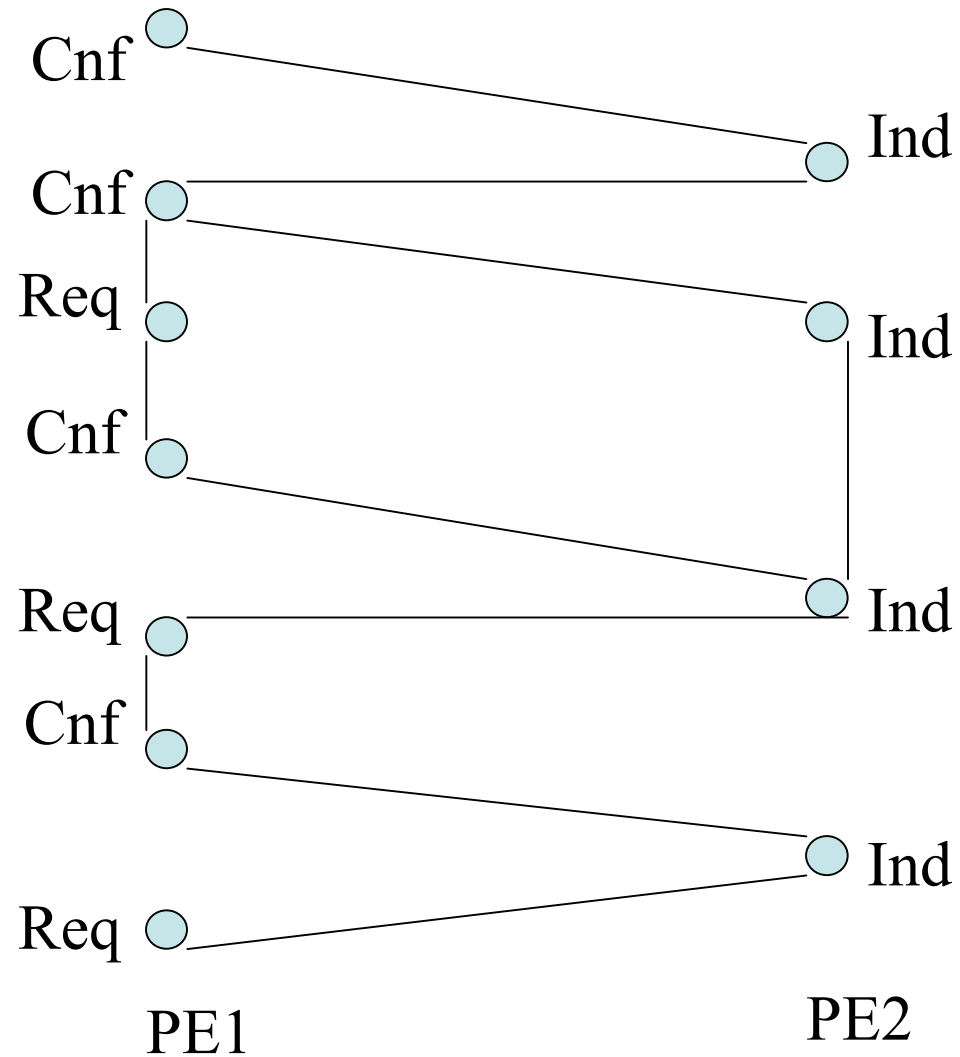
Une exécution erronée (du "Alternating Bit Protocol")



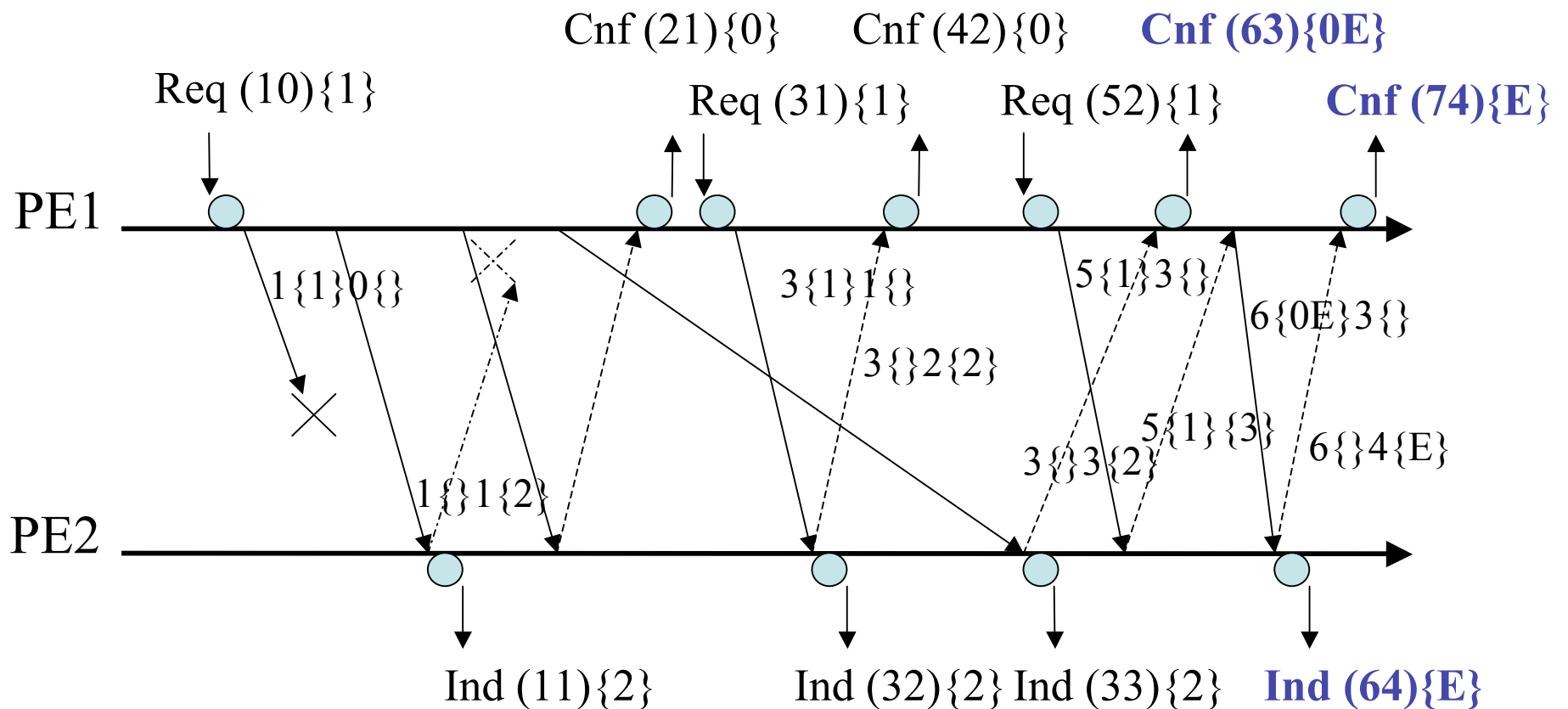
La capture des dépendances causales



Le diagramme de Hasse de l'exécution



Vérification répartie du service



Morale...

- Observer de l'intérieur un système réparti requière une instrumentation pour capturer les dépendences
- La détection de propriétés peut se penser en centralisé ou en réparti
- Centralisé : le treillis des coupures cohérentes
- Réparti (local) : un pas vers le contrôle réparti

MERCI

