

# Security-aware Program Transformations

Massimo Bartoletti, Pierpaolo Degano, Gian Luigi Ferrari

Dipartimento di Informatica, Università di Pisa



# Stack Inspection (1)

- access control mechanism based on the analysis of the execution stack (stack of method frames)
- a security policy maps each class to a *protection domain* (a named set of permissions)
- to check if a permission  $P$  is granted:

```
for each frame in the call stack (starting from top)
    if  $P$  is not granted to the frame
        throw an AccessControlException
    if the frame is privileged
        return
```

# Stack Inspection (2)

- *lazy* evaluation strategy: the one shown above
  - slow security checks
  - prevents from interprocedural optimizations
  - + no update of the security context

# Stack Inspection (2)

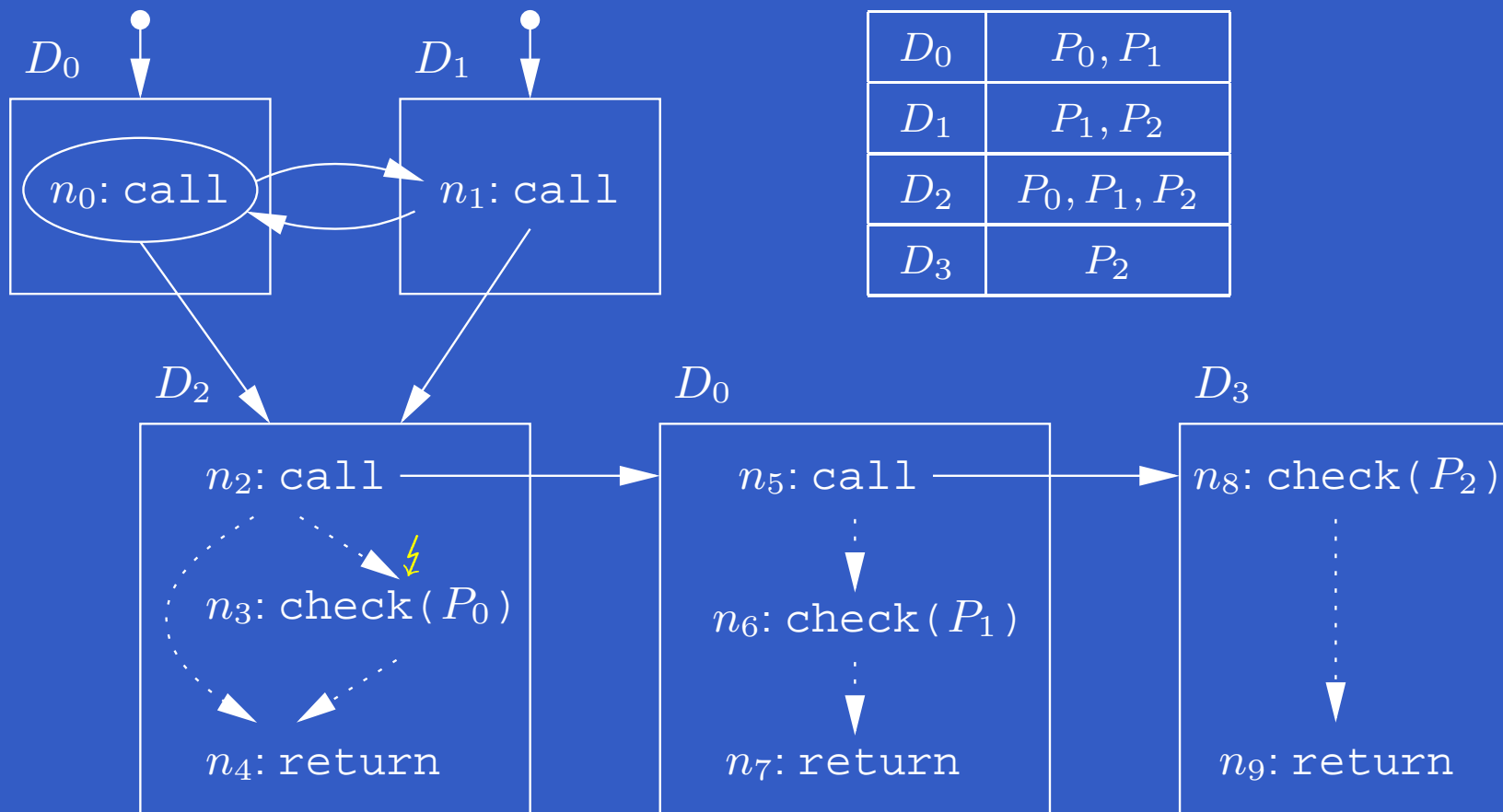
- *lazy* evaluation strategy: the one shown above
  - slow security checks
  - prevents from interprocedural optimizations
  - + no update of the security context
- *eager* evaluation strategy: the set of granted permissions is updated at each method call
  - + fast security checks
  - + allows for interprocedural optimizations  
(in combination with *security passing style*)
  - update of the security context

# Program Model

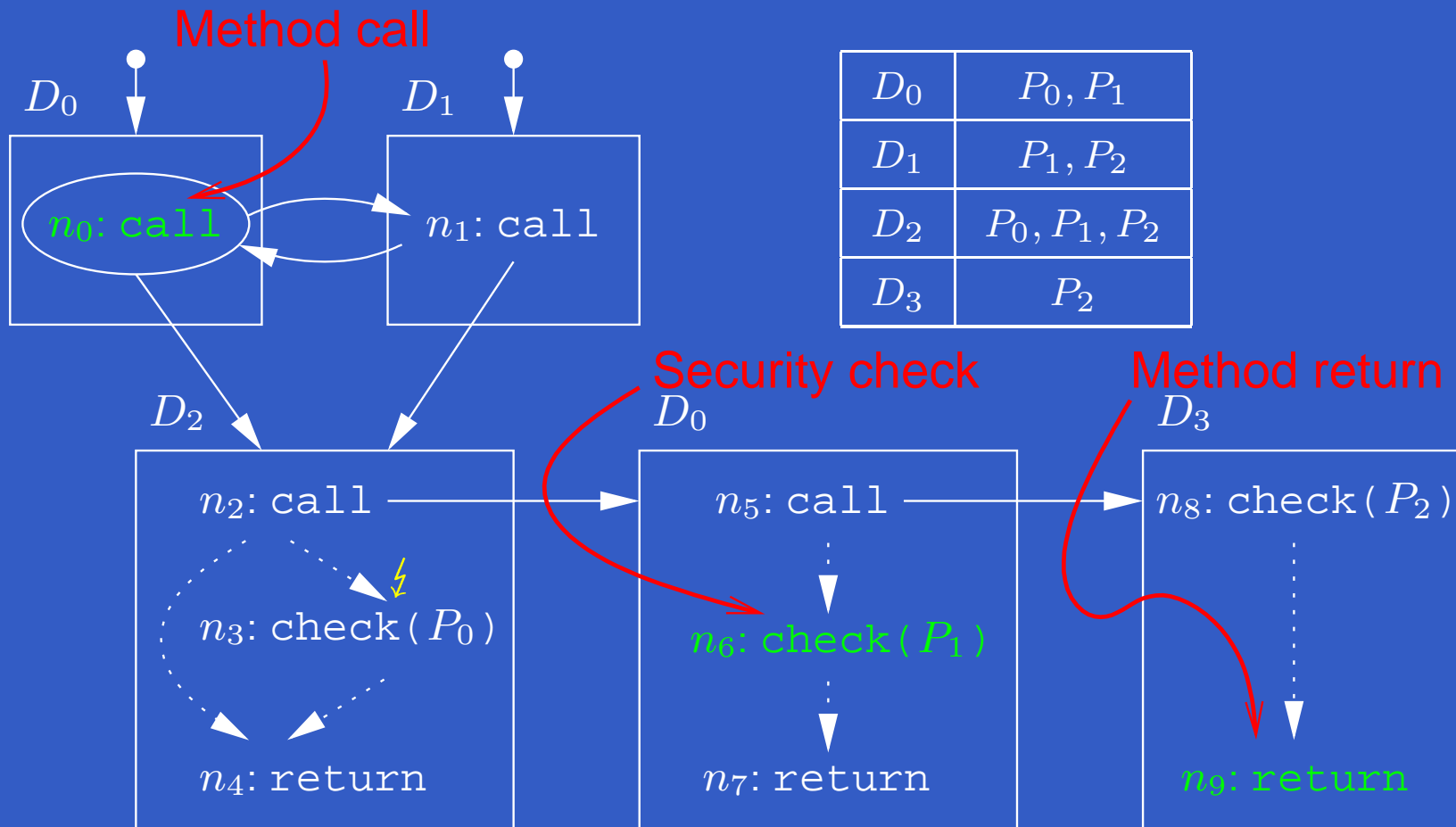
Java Bytecode  $\xrightarrow{CFA}$  Control Flow Graph

- control flow + security checks
- no data flow
- conditional construct  $\longrightarrow$  nondeterminism
- dynamic dispatching  $\longrightarrow$  nondeterminism

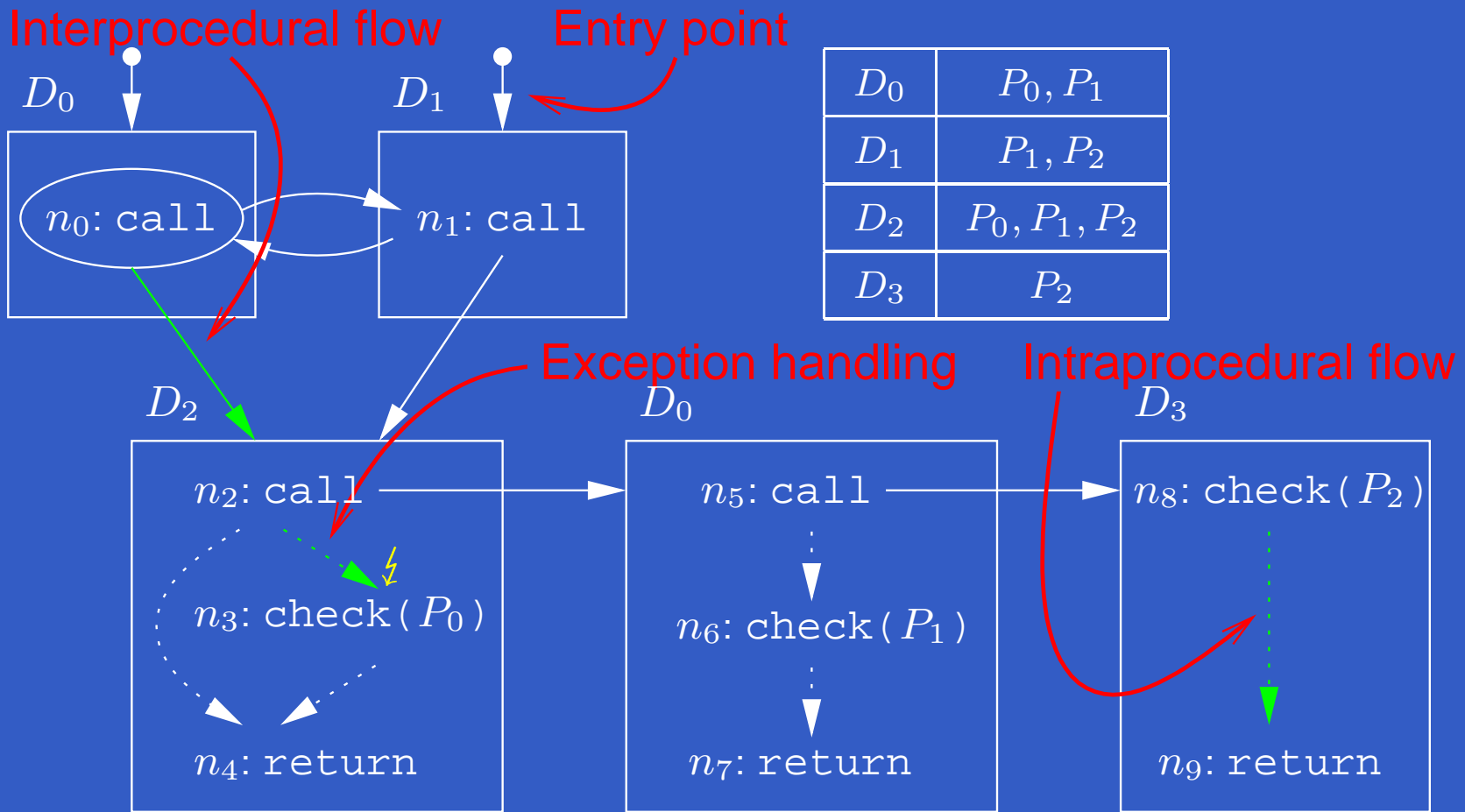
# Program Model - syntax (1)



# Program Model - syntax (2)

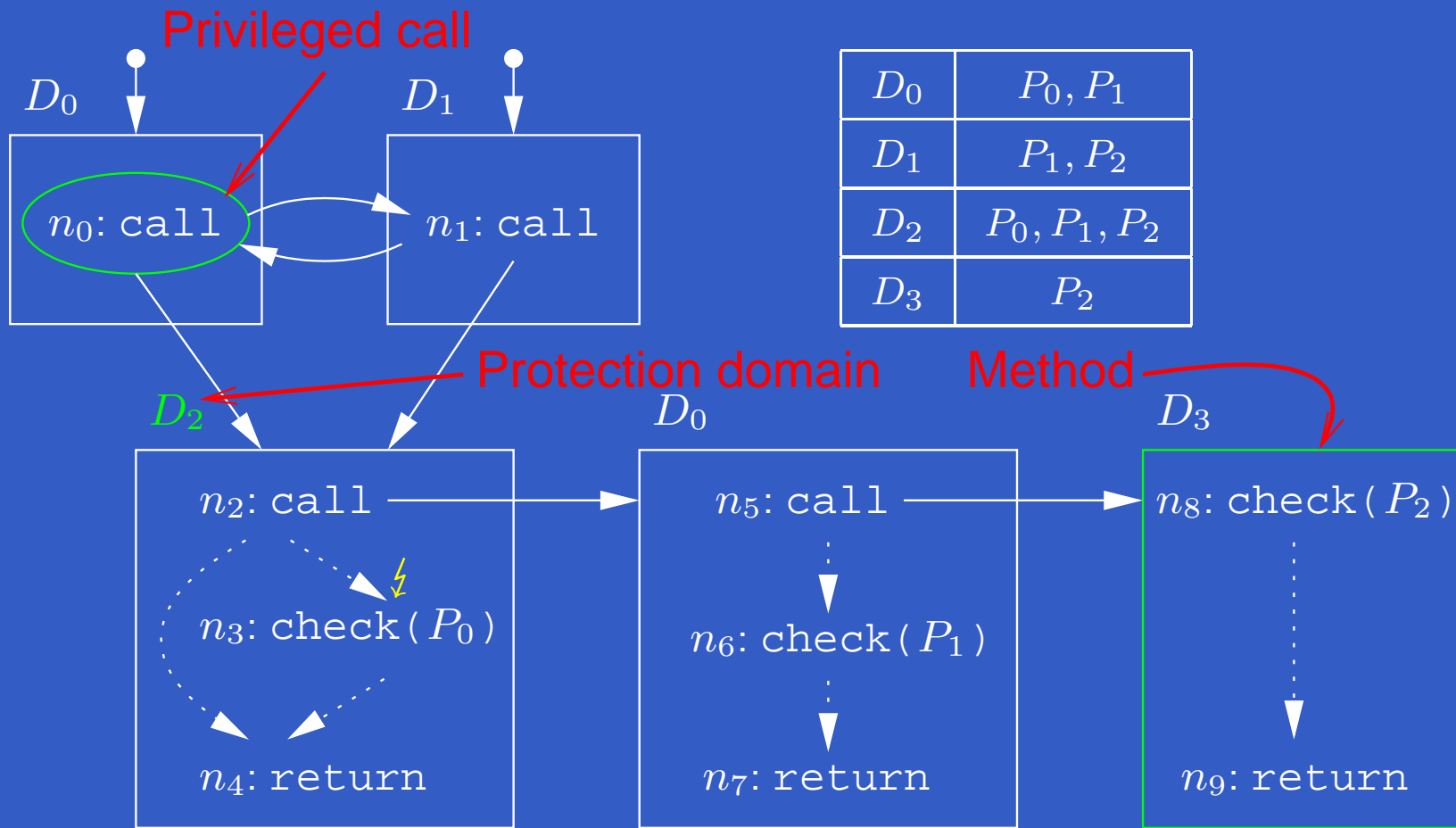


# Program Model - syntax (3)

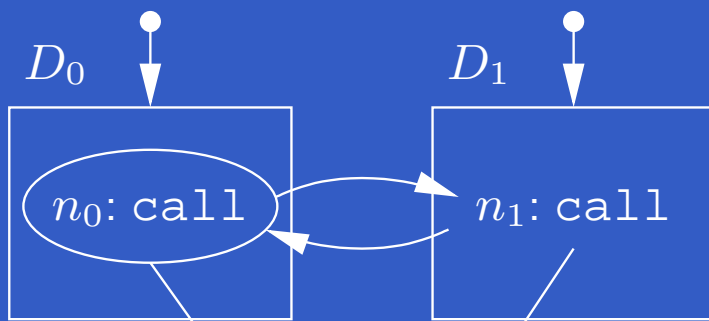




# Program Model - syntax (4)

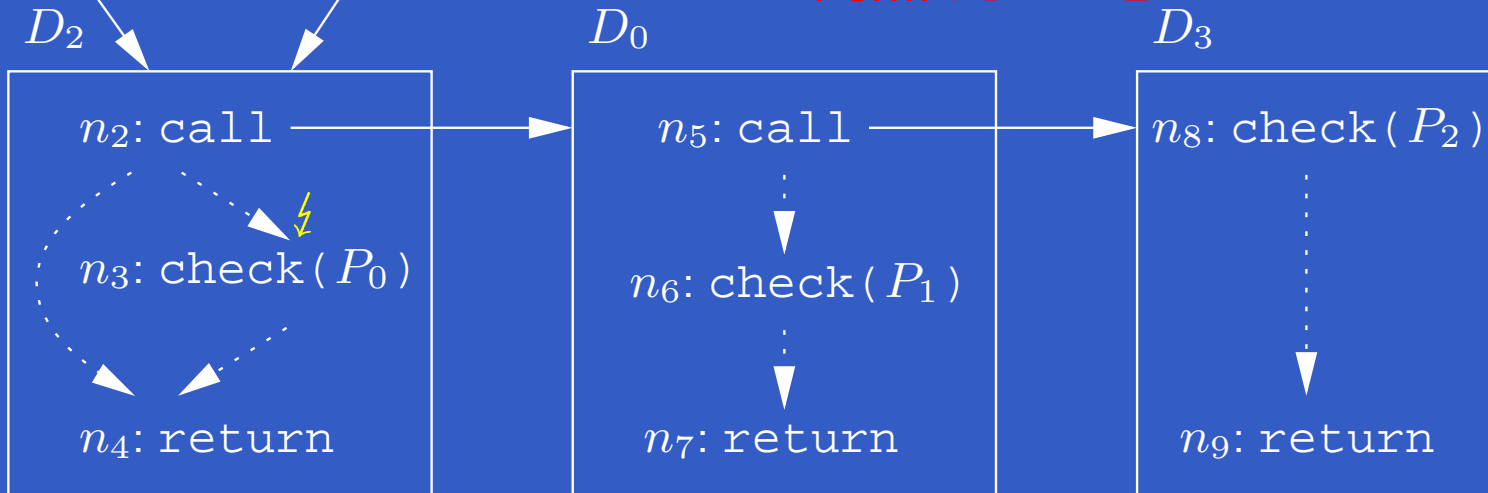


# Program Model - syntax (5)



| $\mathcal{D}$ | $\mathcal{P}$   |
|---------------|-----------------|
| $D_0$         | $P_0, P_1$      |
| $D_1$         | $P_1, P_2$      |
| $D_2$         | $P_0, P_1, P_2$ |
| $D_3$         | $P_2$           |

$\text{Perm} : \mathcal{D} \rightarrow 2^{\mathcal{P}}$



# Program Model - semantics (1)

- state = call stack + exception flag

Example:

$$\langle [n_0, \dots, n_k], true \rangle = [n_0, \dots, n_k] \downarrow$$


# Program Model - semantics (1)

- state = call stack + exception flag

Example:

$$\langle [n_0, \dots, n_k], true \rangle = [n_0, \dots, n_k] \downarrow$$

top node



# Program Model - semantics (1)

- state = call stack + exception flag

Example:

$$\langle [n_0, \dots, n_k], true \rangle = [n_0, \dots, n_k] \downarrow$$

an exception is active!

# Program Model - semantics (1)

- state = call stack + exception flag

Example:

$$\langle [n_0, \dots, n_k], true \rangle = [n_0, \dots, n_k] \downarrow$$

- stack inspection  $\sigma \vdash P$
- transition relation  $\langle \sigma, x \rangle \triangleright \langle \sigma', x' \rangle$
- reachability relation  $G \triangleright \langle \sigma, x \rangle$  when there is a trace from  $\langle [], false \rangle$  to  $\langle \sigma, x \rangle$

# Program Model - semantics (2)

## Stack inspection

$$\frac{}{[] \vdash P} \quad [\vdash_1]$$

$$\frac{P \in \text{Perm}(n) \quad \sigma \vdash P}{\sigma : n \vdash P} \quad [\vdash_2]$$

$$\frac{P \in \text{Perm}(n) \quad \text{Priv}(n)}{\sigma : n \vdash P} \quad [\vdash_3]$$

# Program Model - semantics (3)

## Method call/return

$$\frac{\bullet \longrightarrow n}{[] \triangleright [n]}$$

$[ \triangleright_{entry} ]$

$$\frac{\ell(n) = \text{call} \quad n \longrightarrow n'}{\sigma : n \triangleright \sigma : n : n'}$$

$[ \triangleright_{call} ]$

$$\frac{\ell(m) = \text{return} \quad n \dashrightarrow n'}{\sigma : n : m \triangleright \sigma : n'}$$

$[ \triangleright_{ret} ]$



# Program Model - semantics (4)

## Security checks

$$\frac{\ell(n) = \text{check}(P) \quad \sigma : n \vdash P \quad n \dashrightarrow n'}{\sigma : n \triangleright \sigma : n'}$$

$[\triangleright_{\text{pass}}]$

$$\frac{\ell(n) = \text{check}(P) \quad \sigma : n \not\vdash P}{\sigma : n \triangleright \sigma : n \downarrow}$$

$[\triangleright_{\text{fail}}]$

# Program Model - semantics (5)

## Exception handling

$$\frac{n \dashrightarrow_{\downarrow} n'}{\sigma : n_{\downarrow} \triangleright \sigma : n'} \quad [\triangleright_{catch}]$$

$$\frac{n \not\rightarrow_{\downarrow}}{\sigma : n_{\downarrow} \triangleright \sigma_{\downarrow}} \quad [\triangleright_{propagate}]$$

# The Trace Permissions Analysis (1)

- for each node  $n$ , it computes the *security contexts*  $\tau(n)$  of all  $\sigma$  such that  $G \triangleright \sigma$
- Security context of a state  $\sigma$ :

$$\Gamma([\ ] ) = \emptyset \quad \Gamma(\sigma : n) = \begin{cases} \{\mathbf{Dom}(n)\} & \text{if } \mathbf{Priv}(n) \\ \Gamma(\sigma) \cup \{\mathbf{Dom}(n)\} & \text{otherwise} \end{cases}$$

- Set of permissions granted to a security context  $\gamma$ :

$$\Pi(\gamma) = \bigcap_{D \in \gamma} \mathbf{Perm}(D)$$

# The Trace Permissions Analysis (2)

- For each state  $\sigma$  and permission  $P$ :

$$\sigma \vdash P \iff P \in \Pi(\Gamma(\sigma))$$

- For each solution  $\tau$  and state  $\sigma : n$

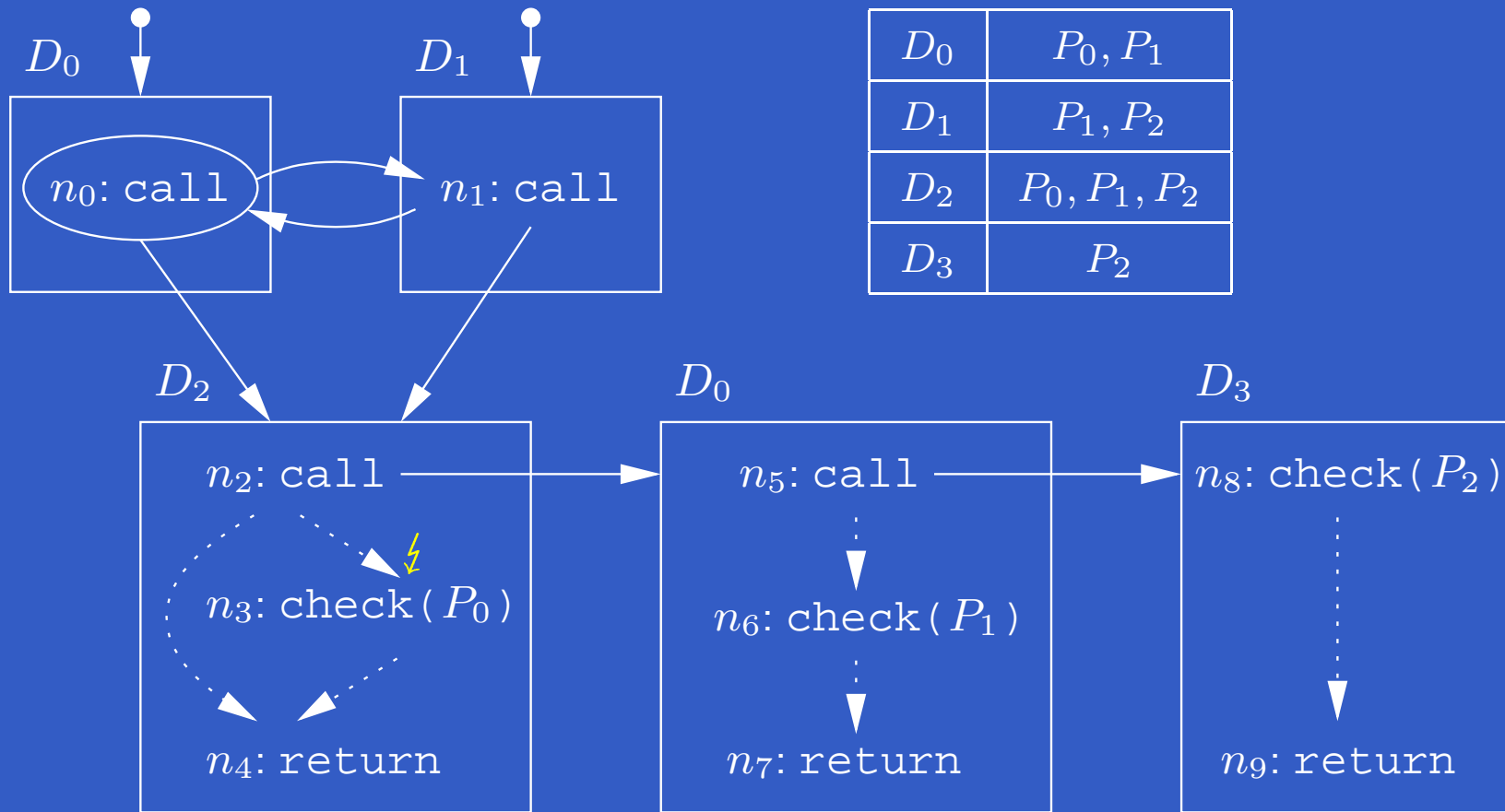
$$G \triangleright \sigma : n \implies \exists \gamma \in \tau(n). \gamma = \Gamma(\sigma : n)$$

- For the *minimal* solution  $\tau$  and each state  $\sigma : n$

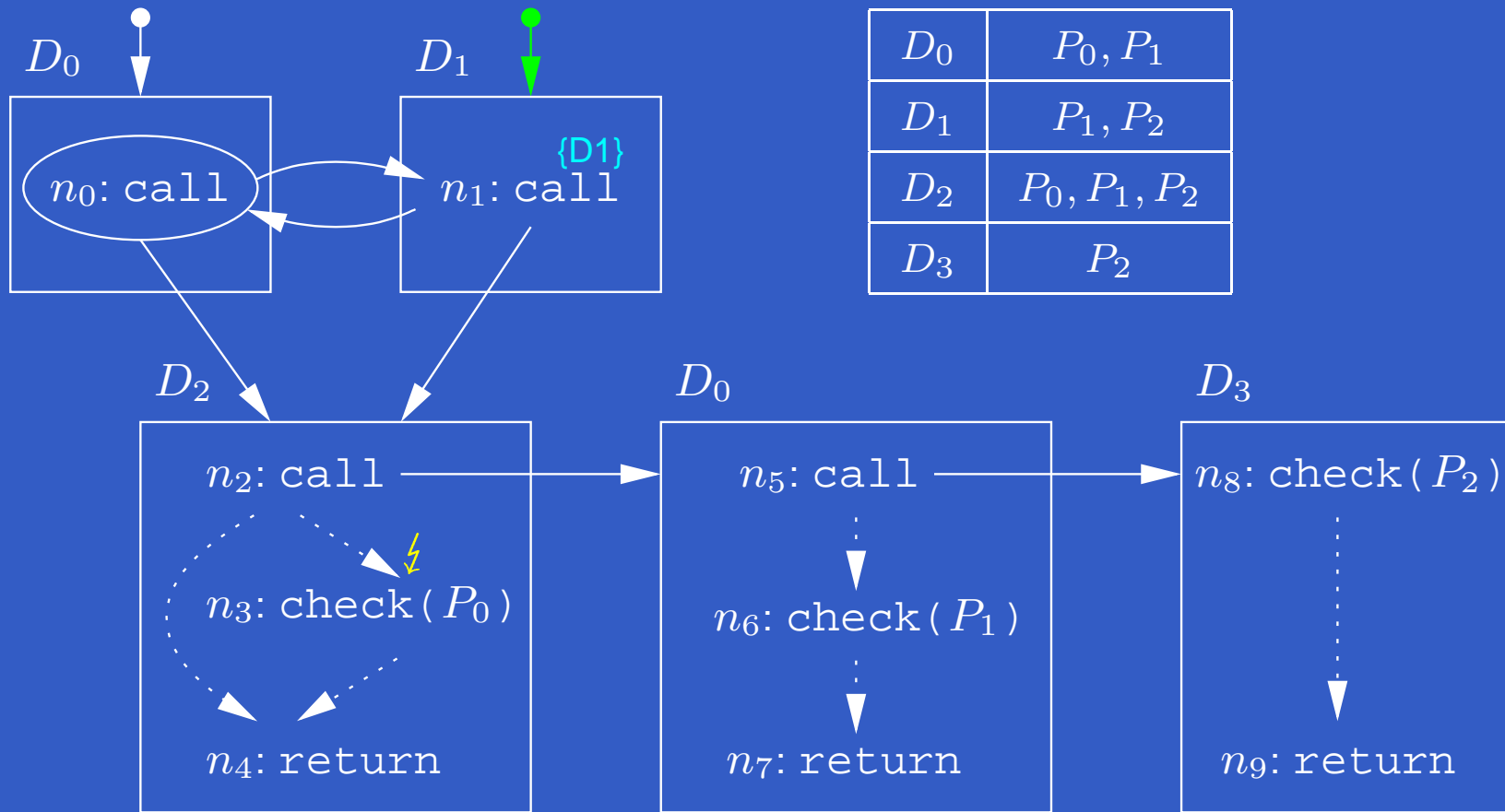
$$\gamma \in \tau(n) \implies \exists \sigma. G \triangleright \sigma : n \wedge \gamma = \Gamma(\sigma : n)$$

- the minimal solution is computed in  $\mathcal{O}(N)$  by our worklist algorithm ( $N$  is the number of nodes)

# The Trace Permissions Analysis (3)

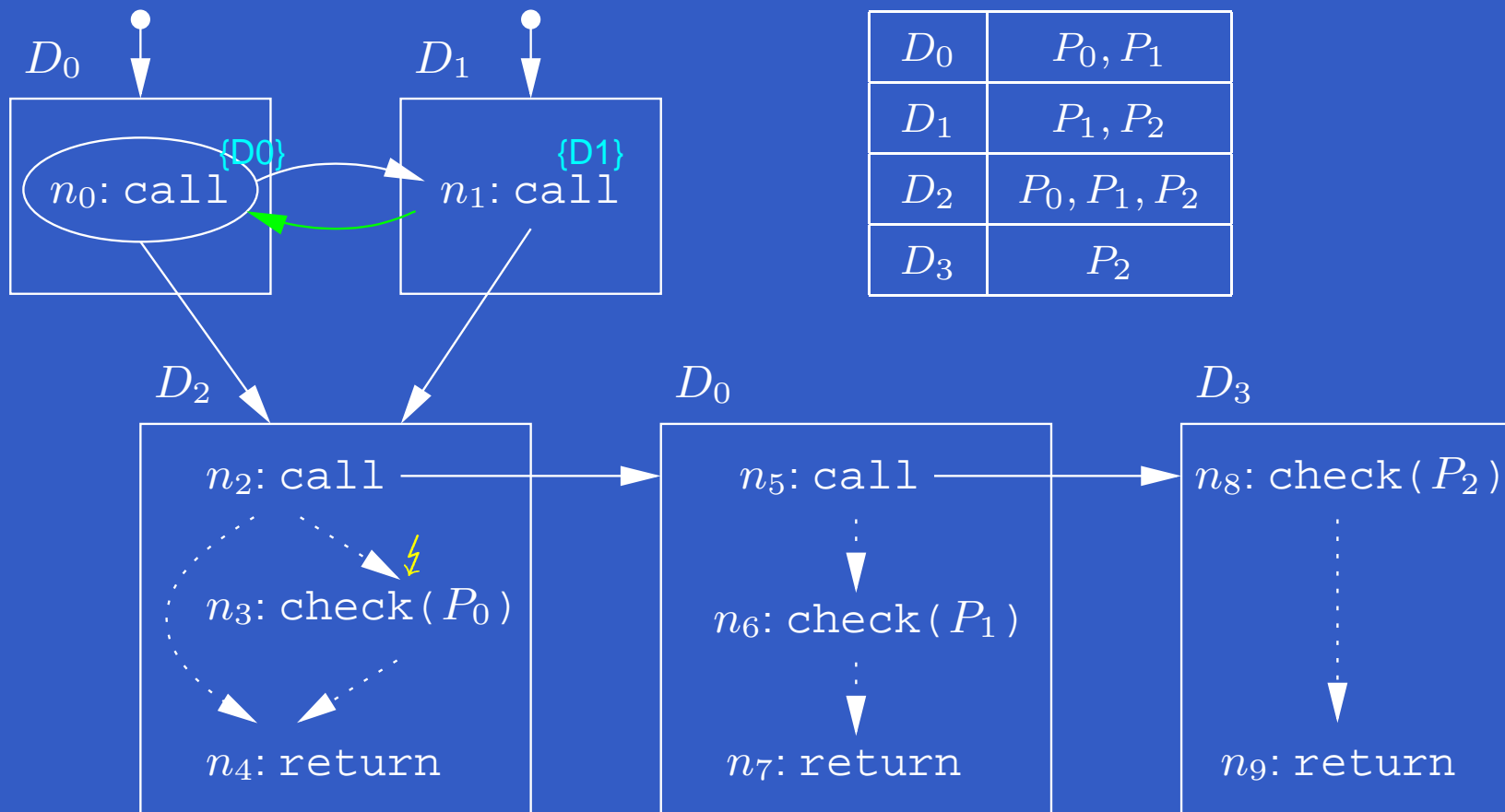


# The Trace Permissions Analysis (3.1)



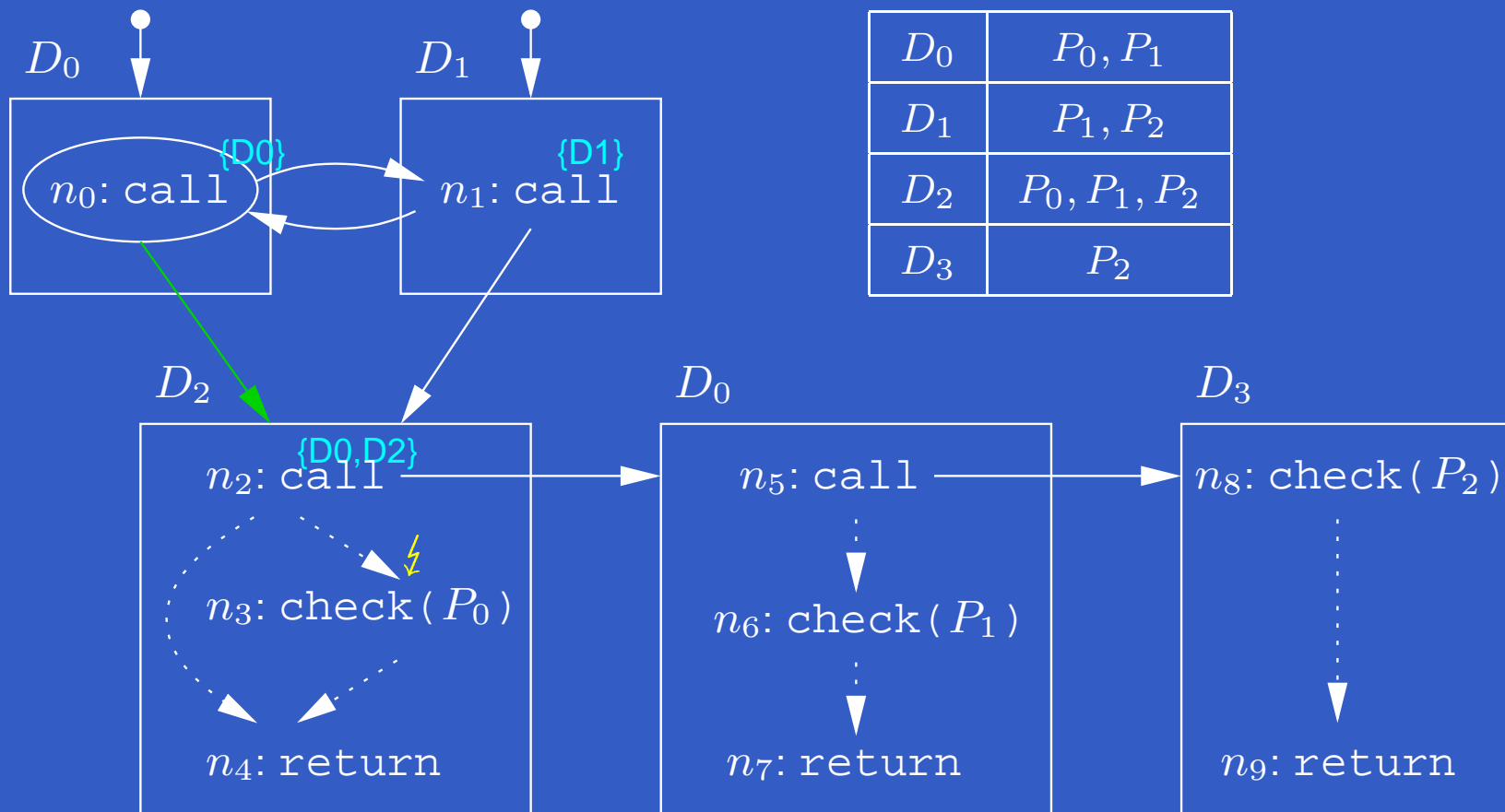
$[] \triangleright [n_1]$

# The Trace Permissions Analysis (3.2)



$[] \triangleright [n_1] \triangleright [n_1, n_0]$

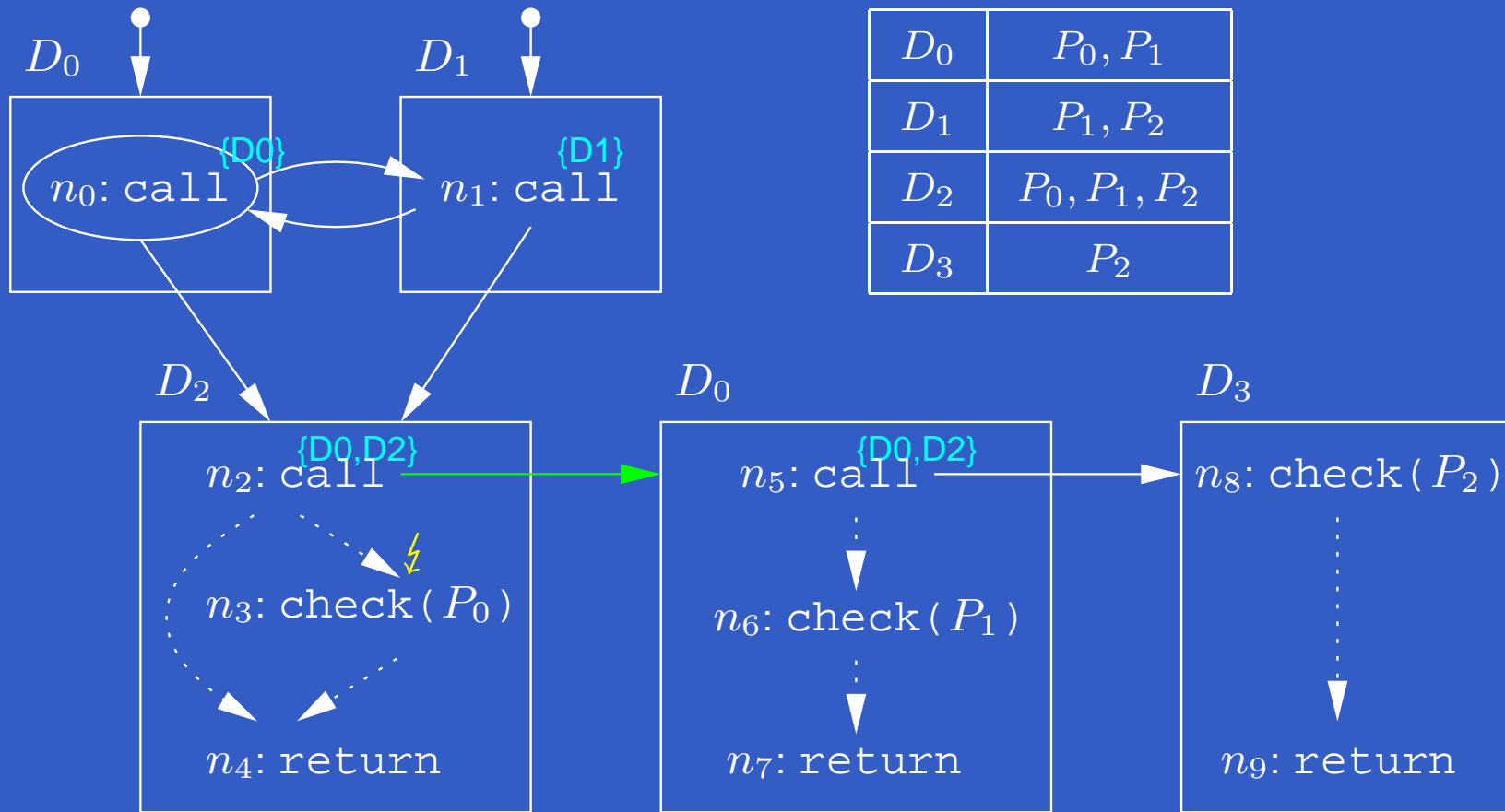
# The Trace Permissions Analysis (3.3)



$[] \triangleright [n_1] \triangleright [n_1, n_0] \triangleright [n_1, n_0, n_2]$

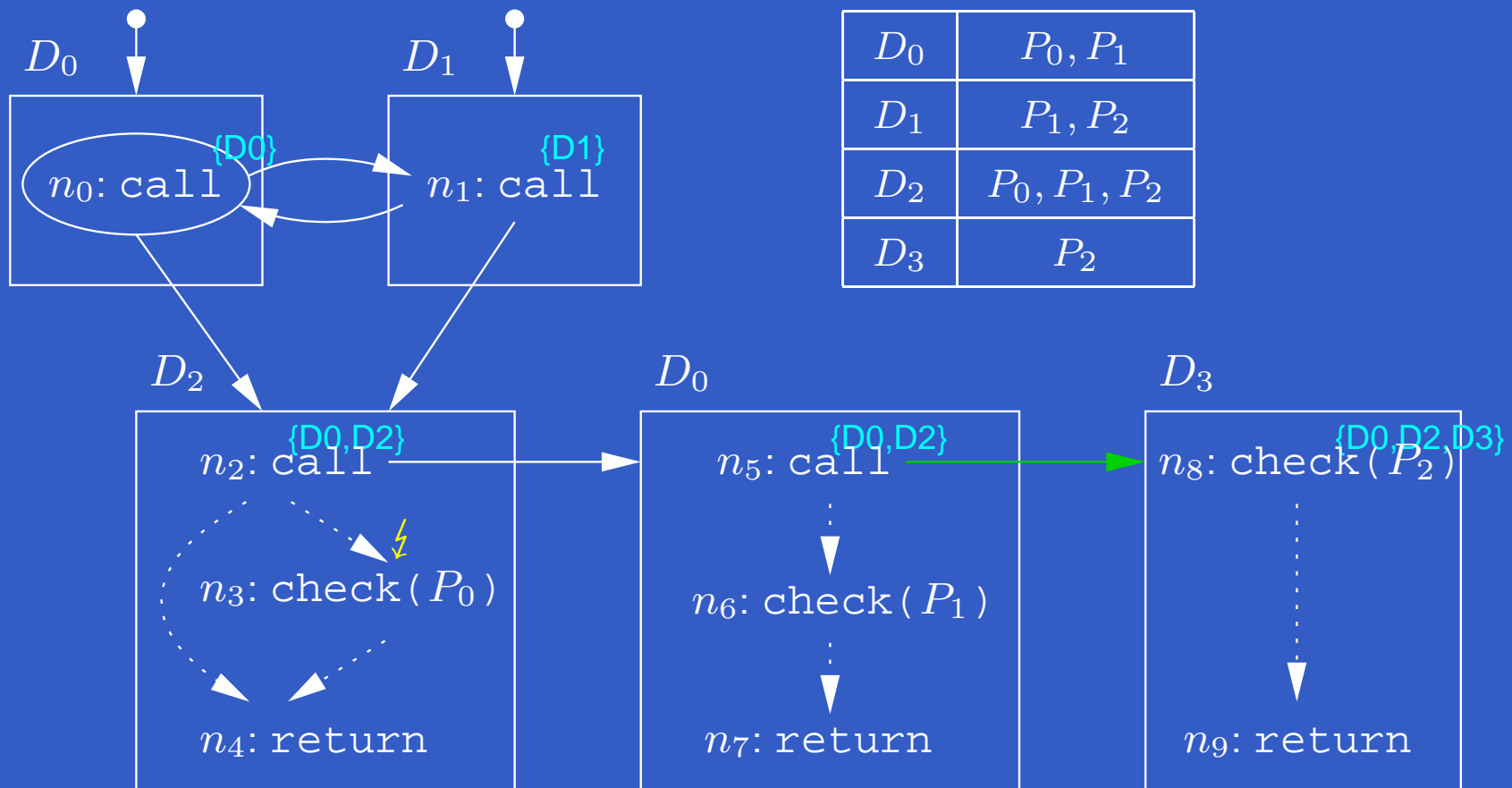


# The Trace Permissions Analysis (3.4)



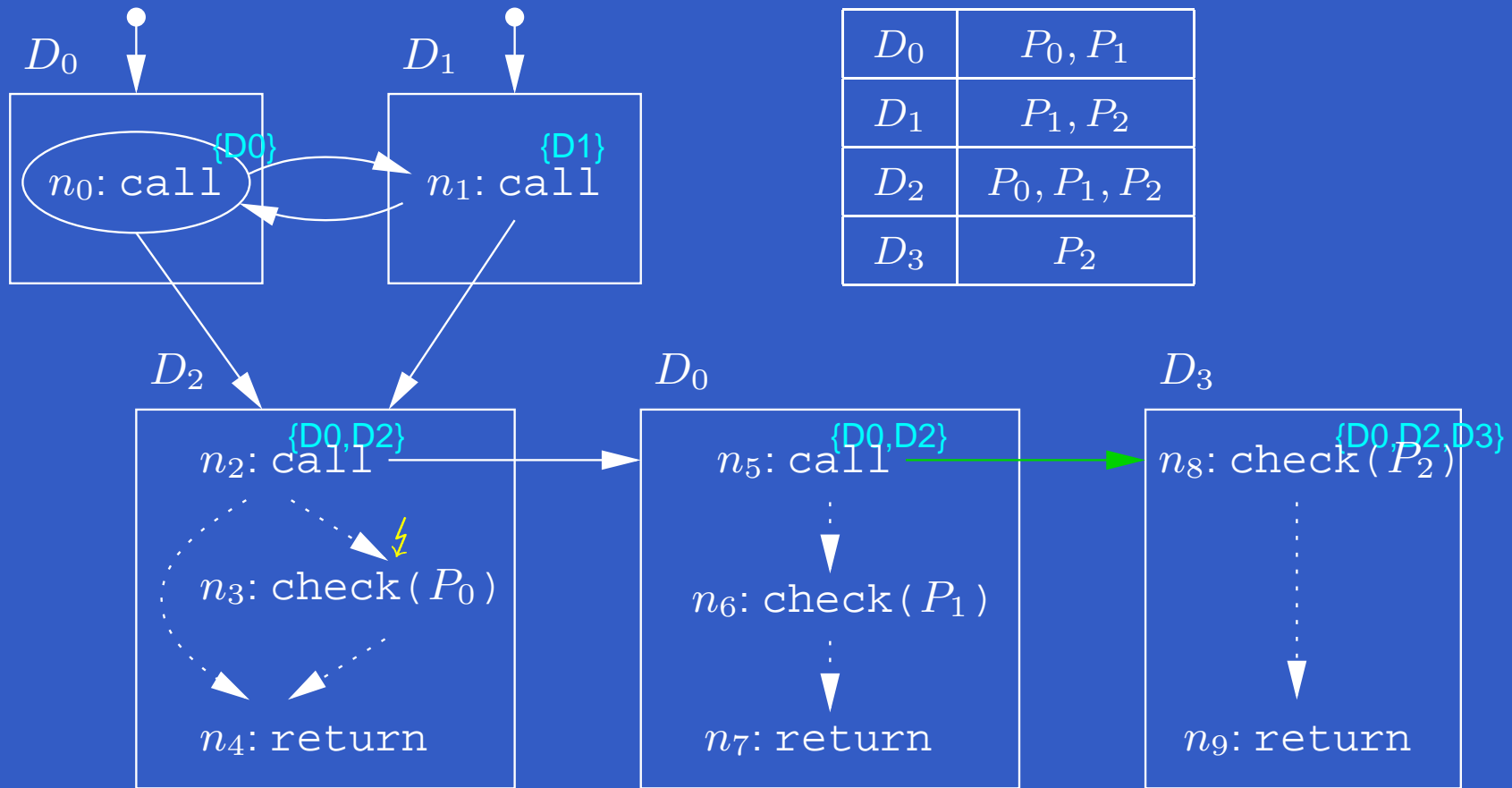
$[] \triangleright [n_1] \triangleright [n_1, n_0] \triangleright [n_1, n_0, n_2] \triangleright [n_1, n_0, n_2, n_5]$

# The Trace Permissions Analysis (3.5)



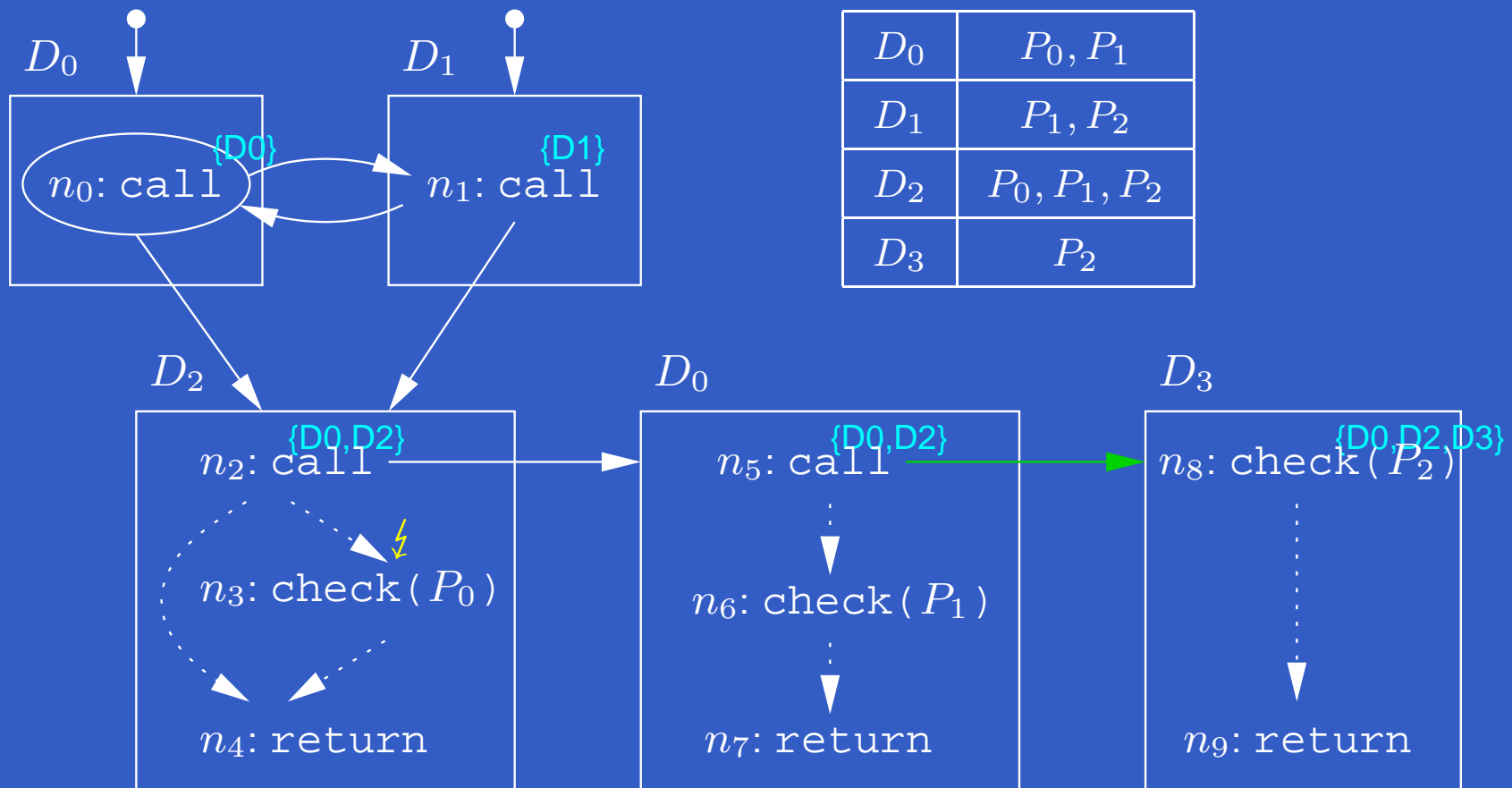
$\dots \triangleright [n_1, n_0, n_2, n_5] \triangleright [n_1, n_0, n_2, n_5, n_8]$

# The Trace Permissions Analysis (3.5)



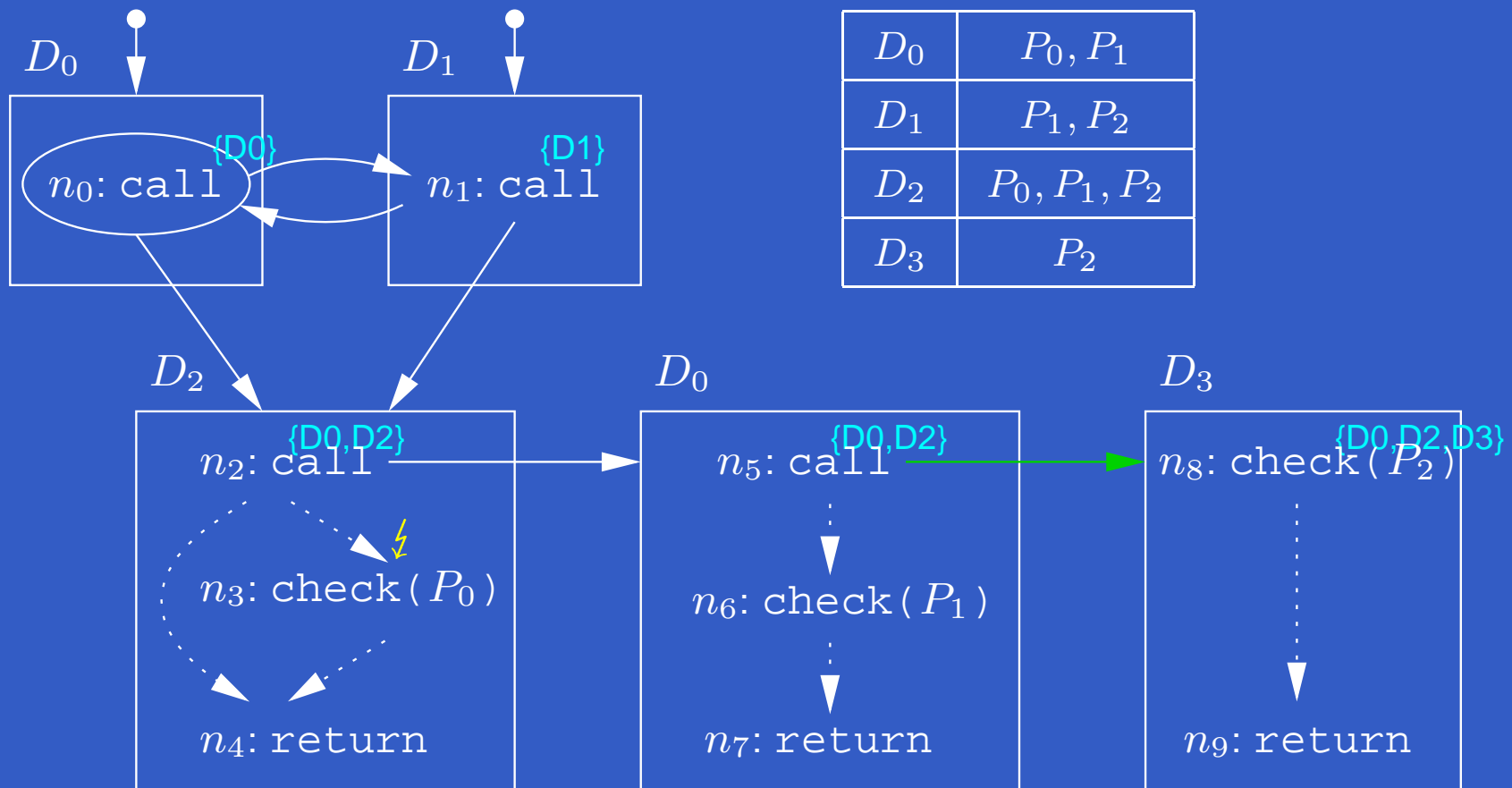
$\dots \triangleright [n_1, n_0, n_2, n_5] \triangleright [n_1, n_0, n_2, n_5, n_8] \not\vdash P_2$

# The Trace Permissions Analysis (3.5)



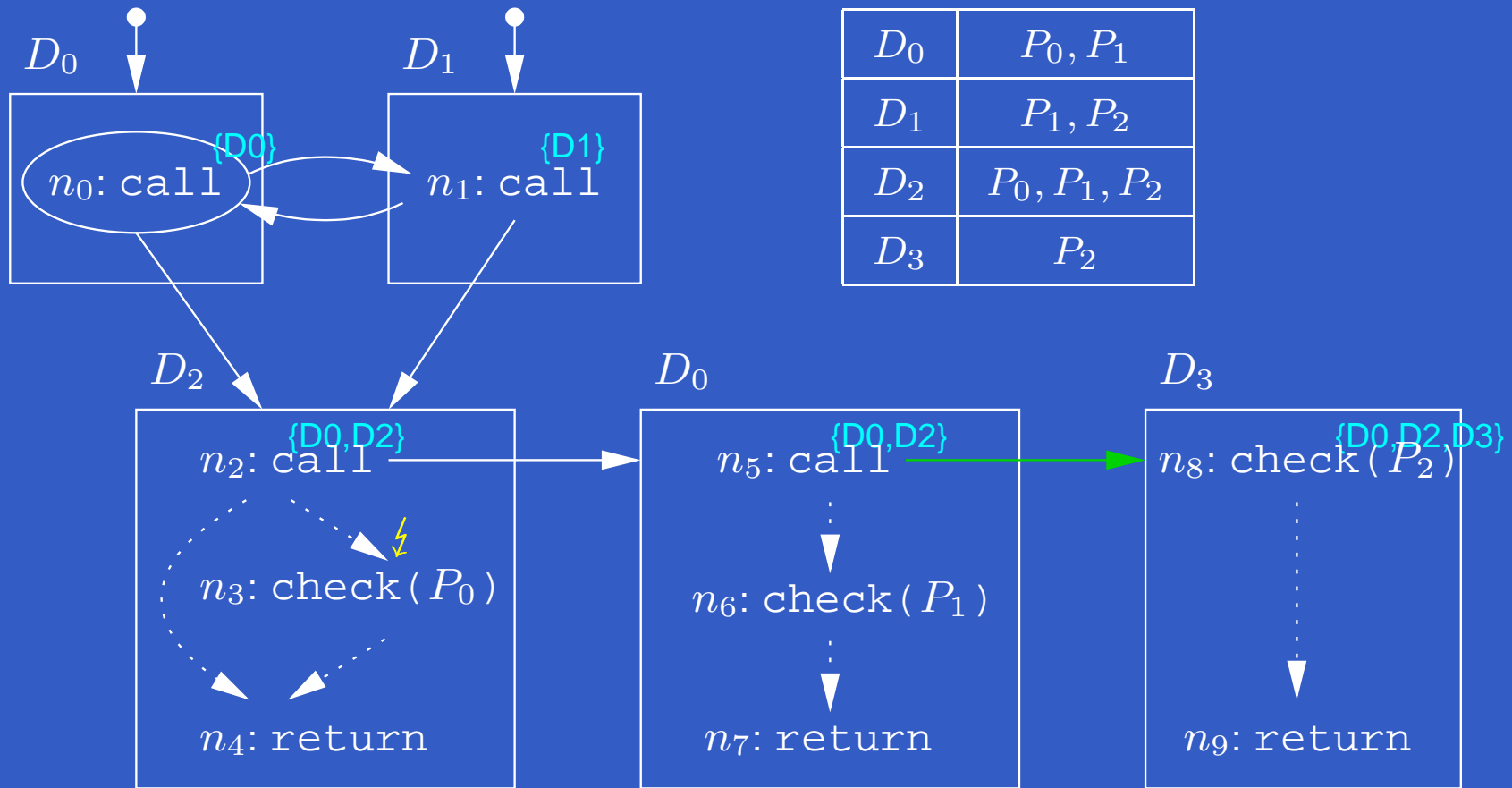
... ▷  $[n_1, n_0, n_2, n_5, n_8]$  ⚡

# The Trace Permissions Analysis (3.5)



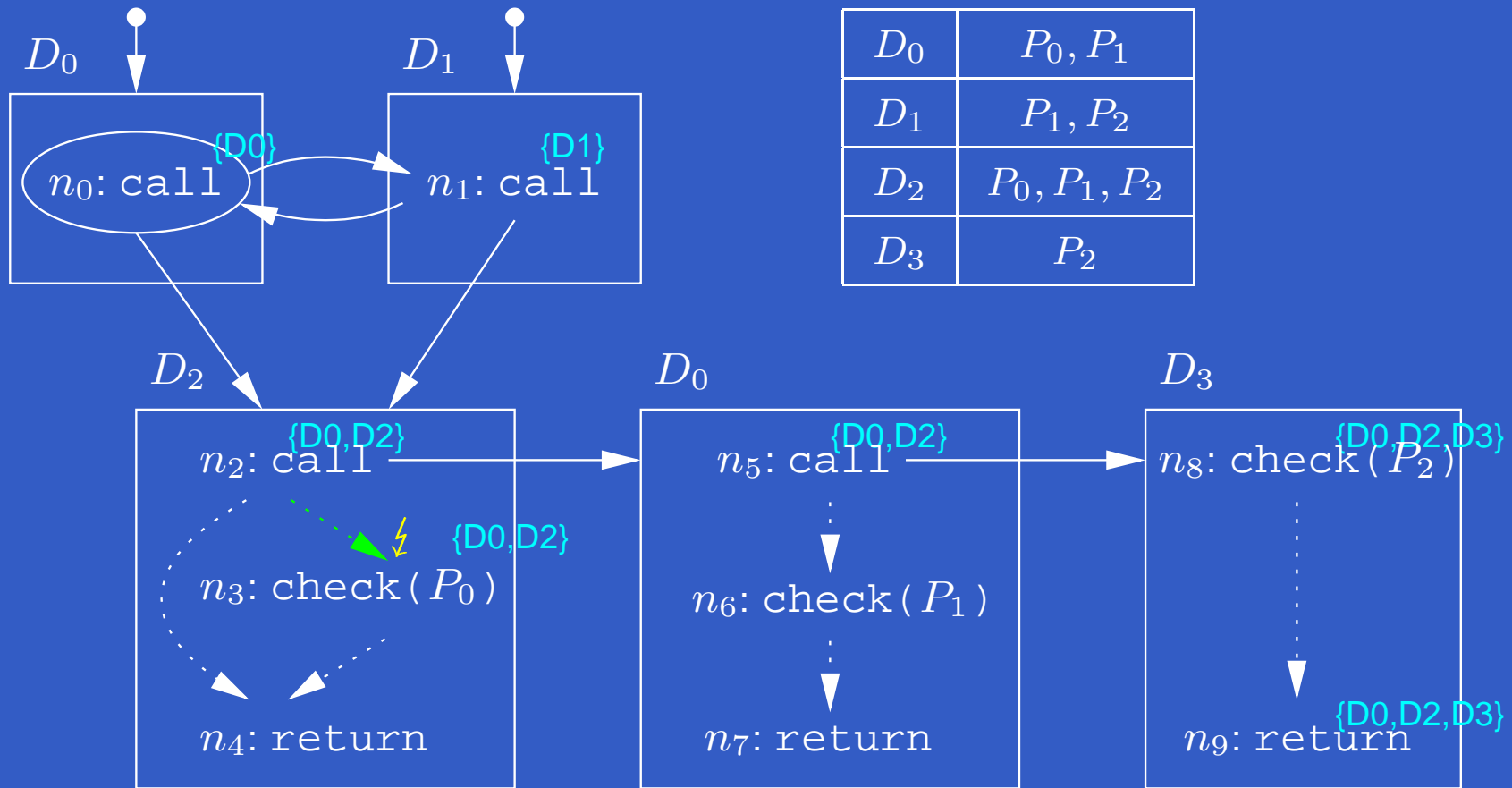
$\dots \triangleright [n_1, n_0, n_2, n_5, n_8] \text{⚡} \triangleright [n_1, n_0, n_2, n_5] \text{⚡}$

# The Trace Permissions Analysis (3.5)



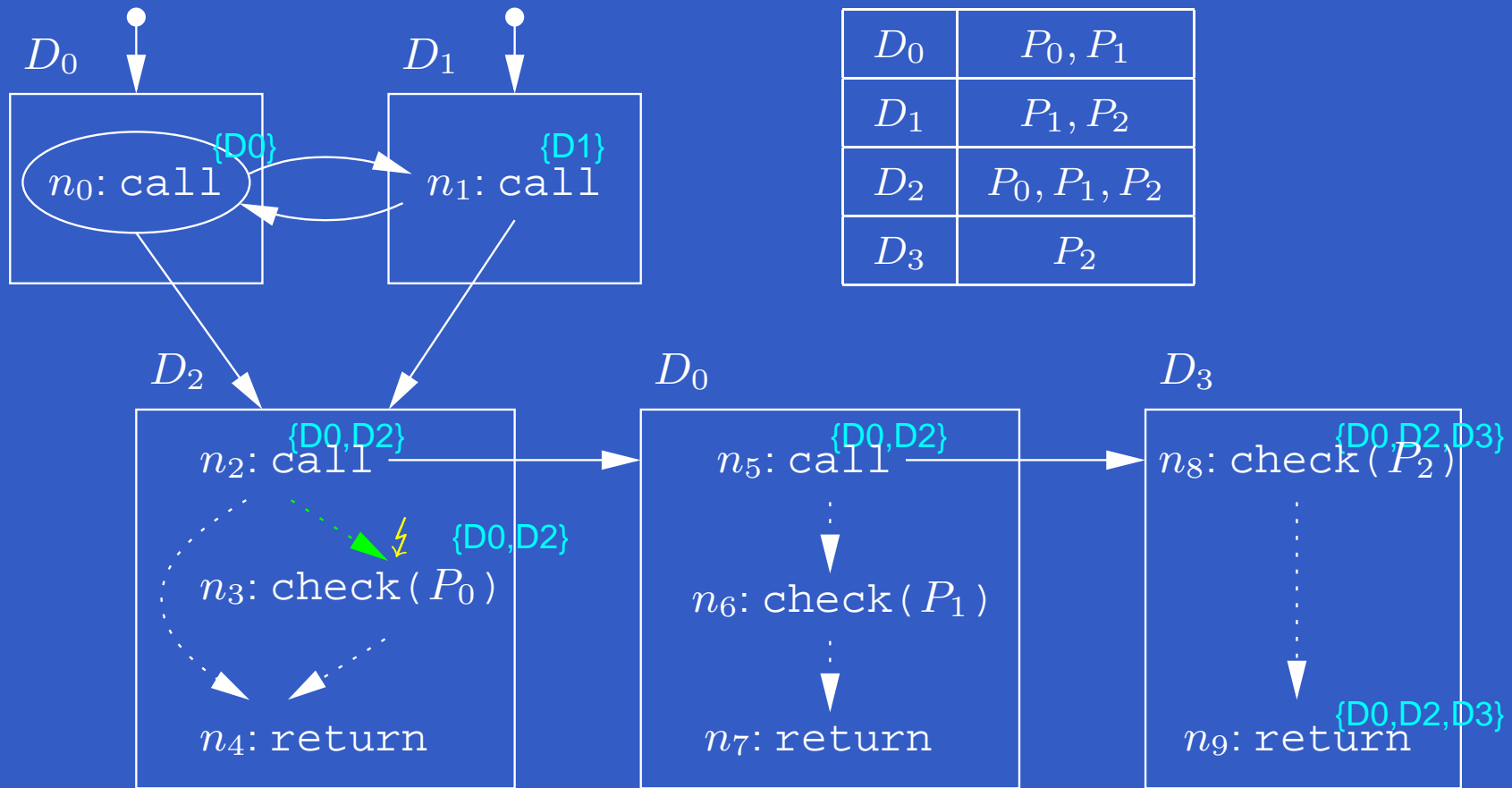
$\dots \triangleright [n_1, n_0, n_2, n_5, n_8] \downarrow \triangleright [n_1, n_0, n_2, n_5] \downarrow \triangleright [n_1, n_0, n_2] \downarrow$

# The Trace Permissions Analysis (3.6)



... ▷  $[n_1, n_0, n_3]$

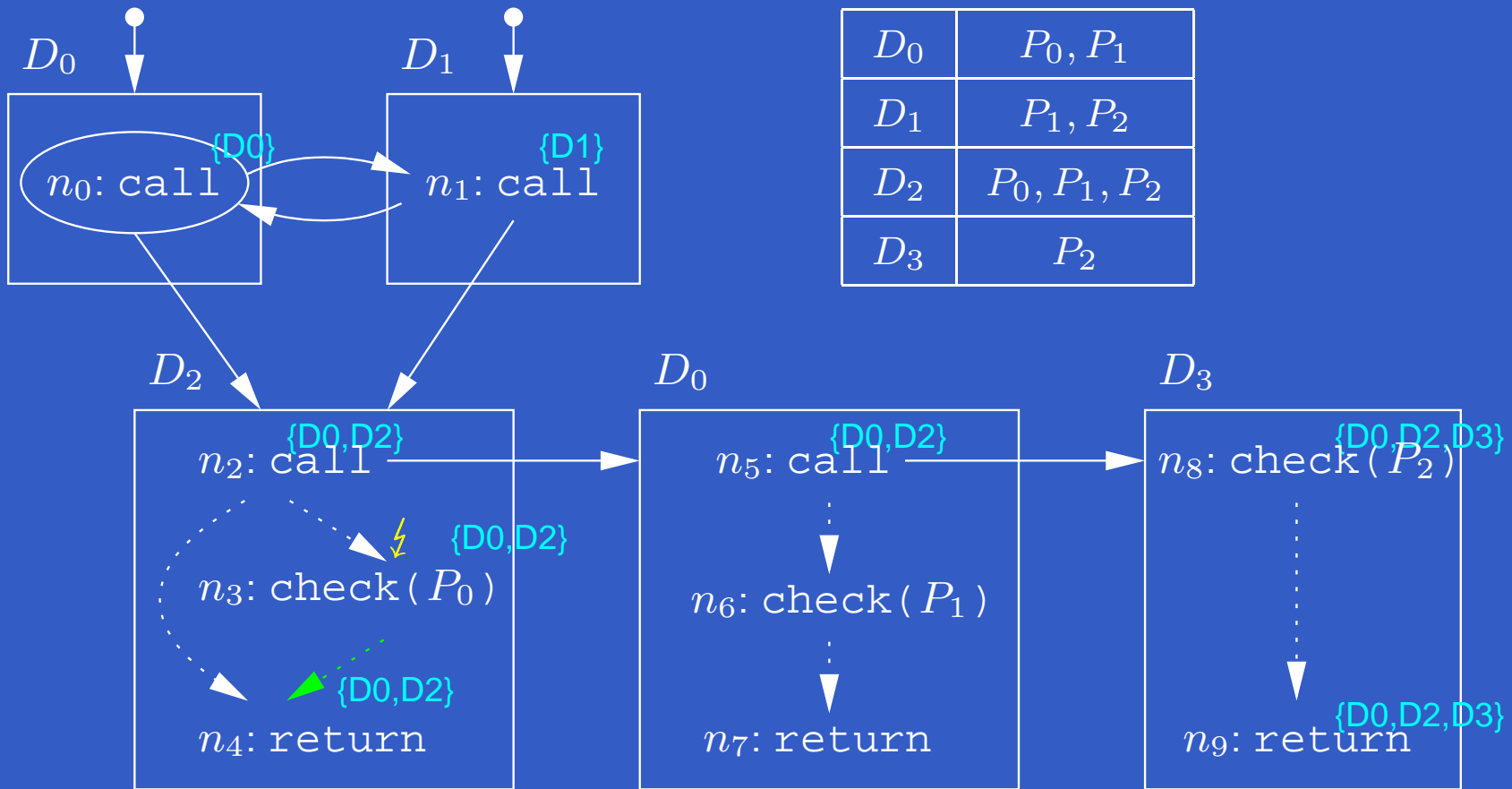
# The Trace Permissions Analysis (3.6)



$\dots \triangleright [n_1, n_0, n_3] \vdash P_0$



# The Trace Permissions Analysis (3.7)



$\dots \triangleright [n_1, n_0, n_3] \triangleright [n_1, n_0, n_4]$

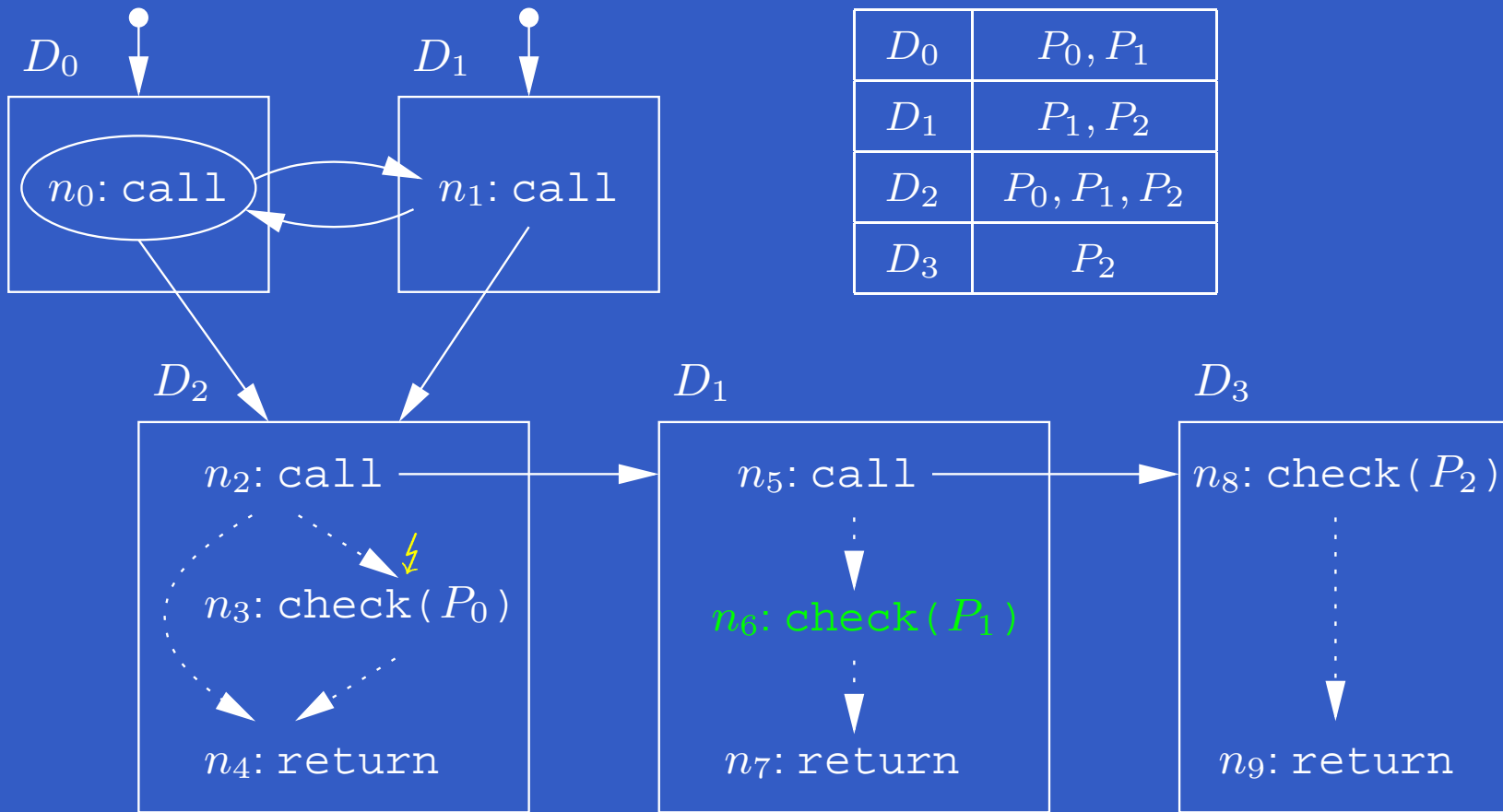
# Optimizations

- Elimination of the redundant checks
- Dead code elimination
- Method inlining
- Tail call elimination
- Fast implementation of eager stack inspection

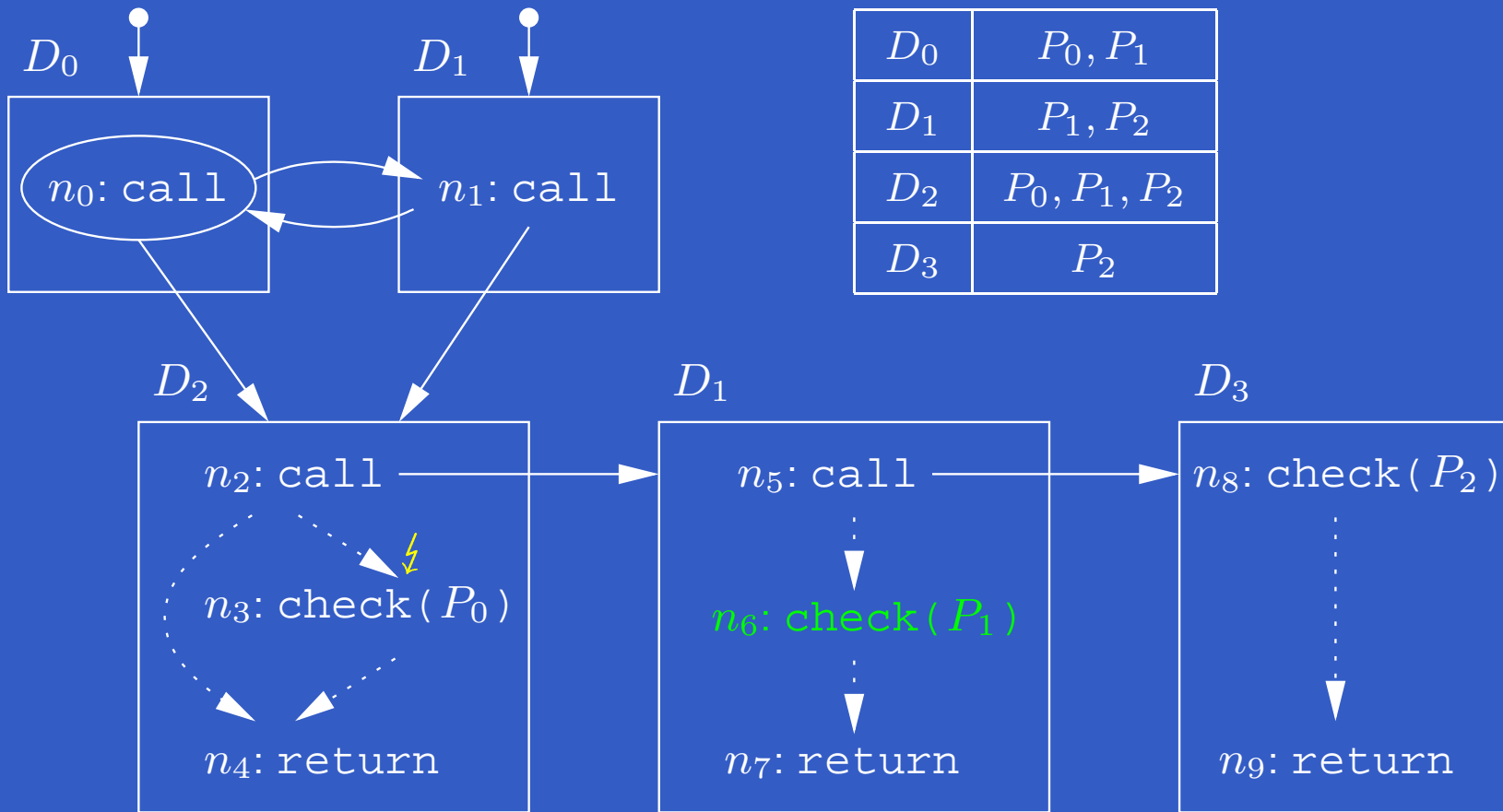
# Optimizations

- Elimination of the redundant checks
- Dead code elimination
- Method inlining
- Tail call elimination
- Fast implementation of eager stack inspection

# Redundant Checks Elimination (1)

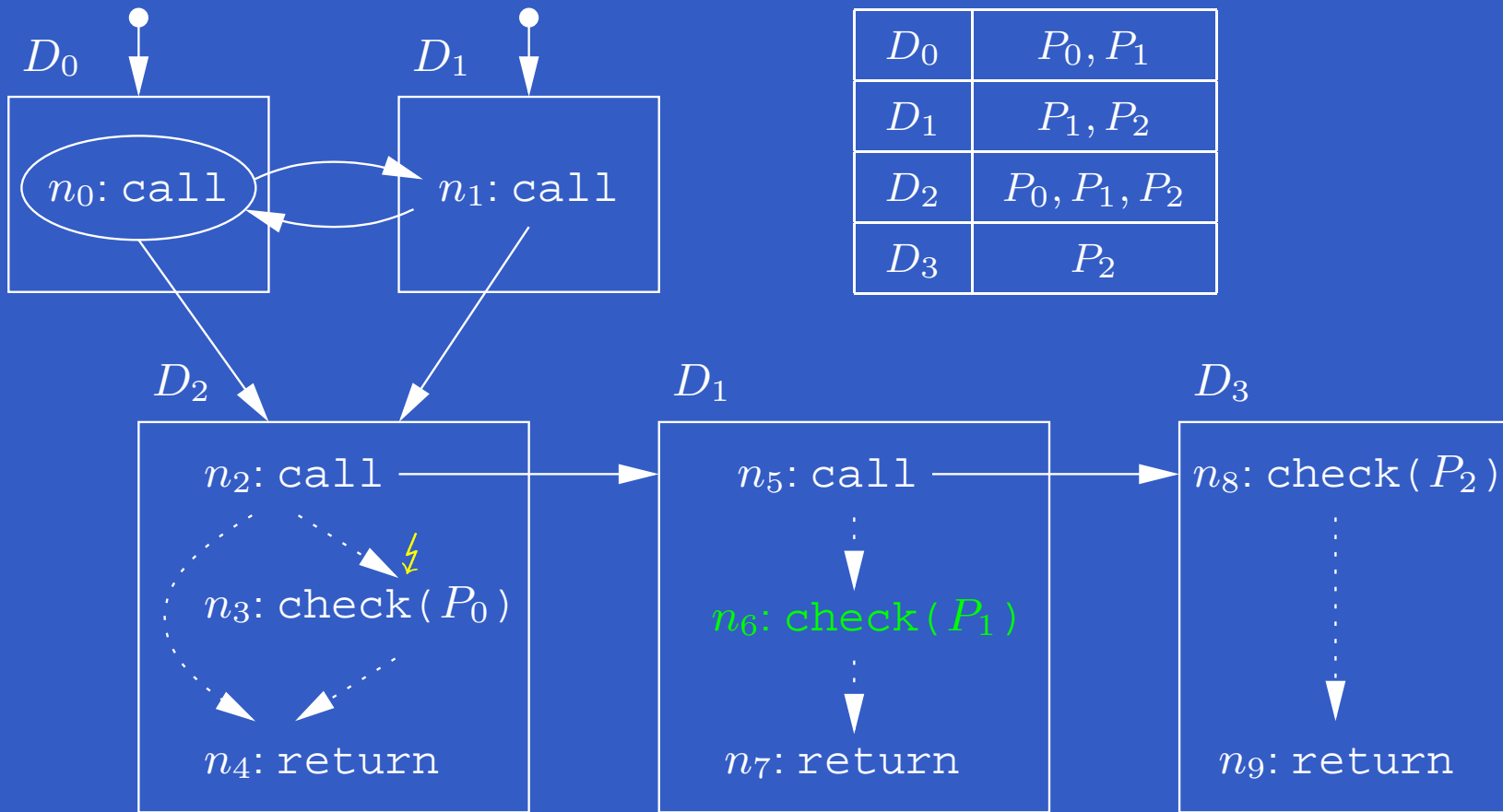


# Redundant Checks Elimination (1)



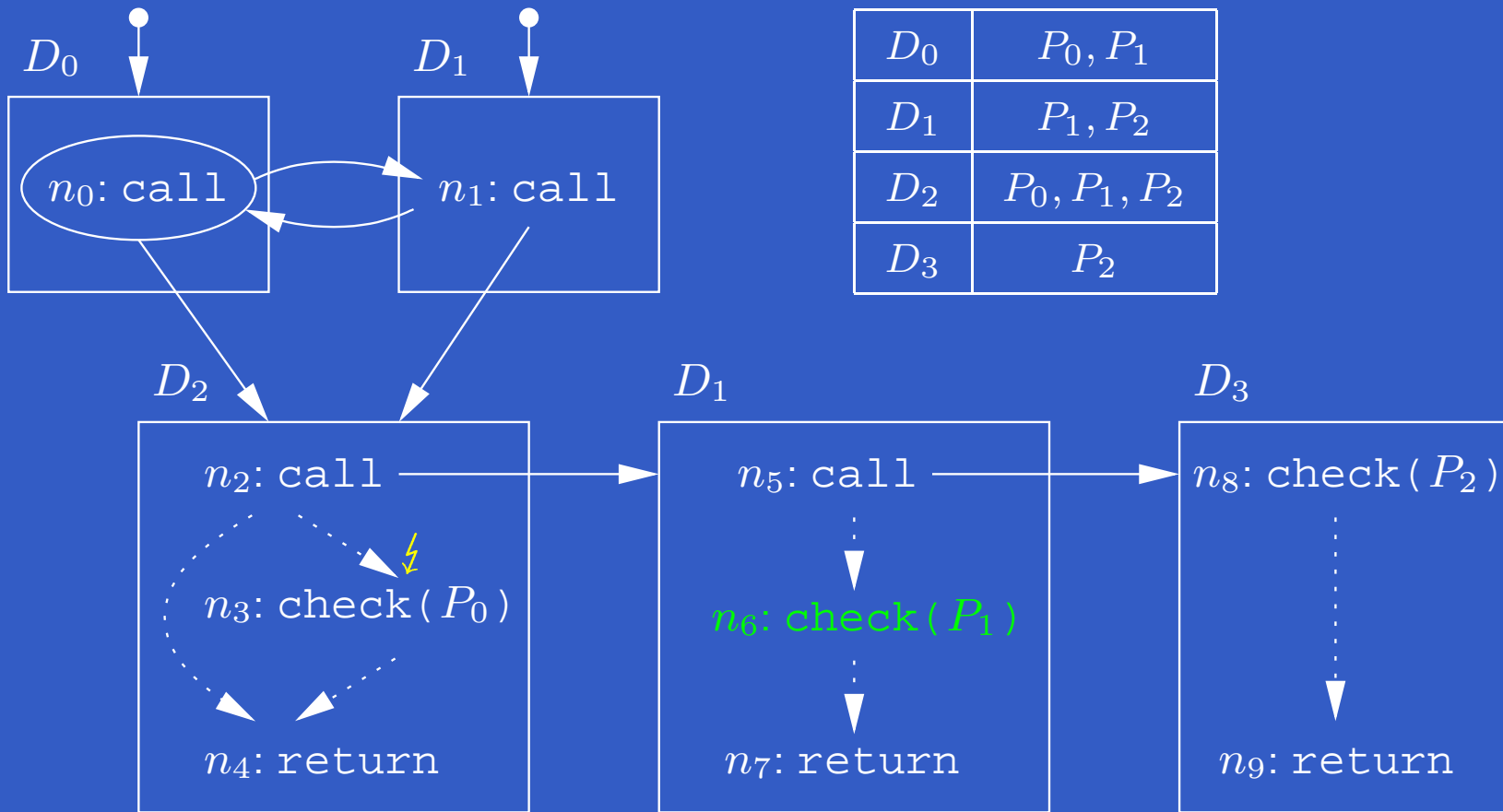
$$\tau(n_6) = \{\{D_1, D_2\}\}$$

# Redundant Checks Elimination (1)



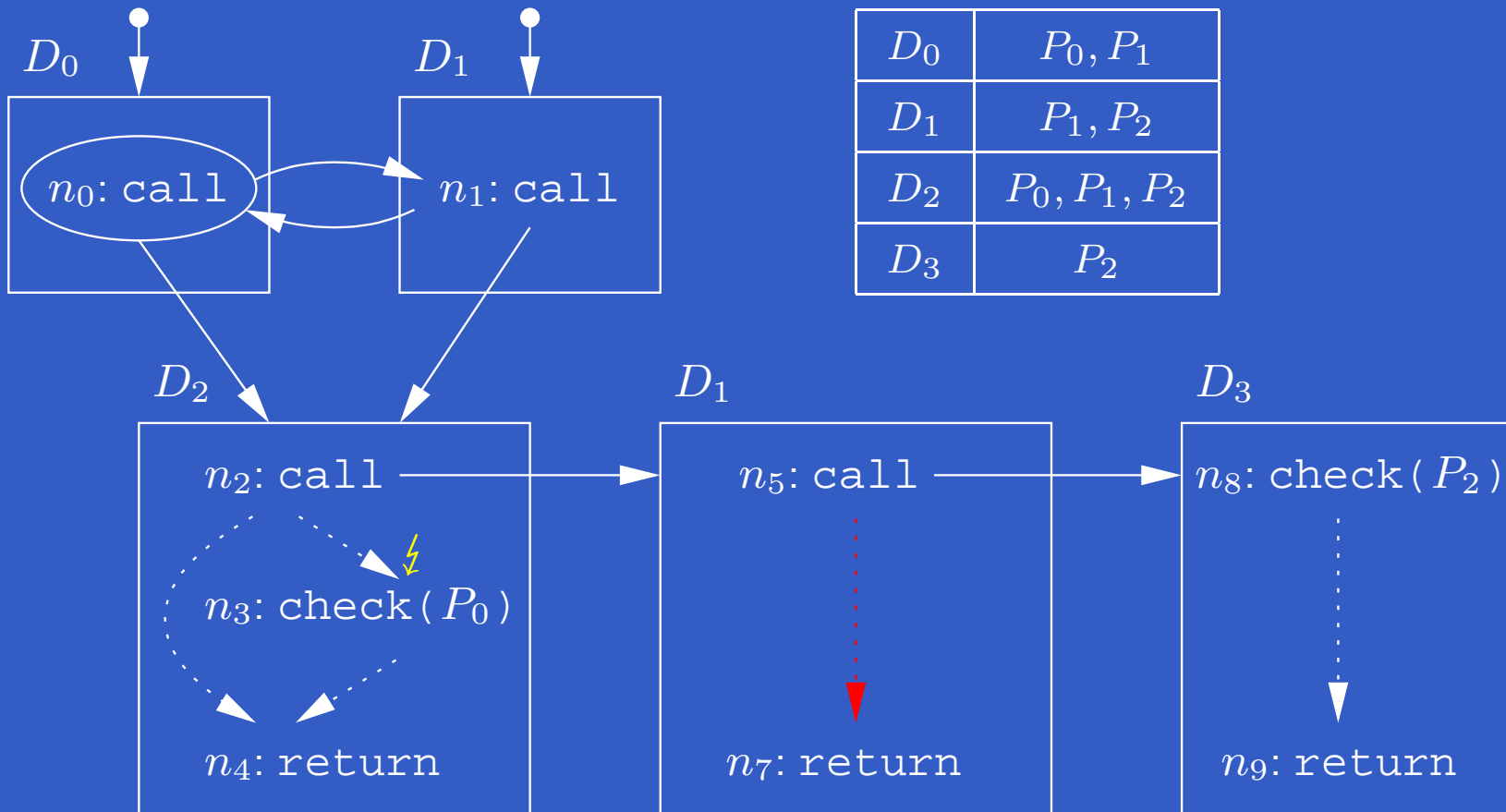
$$\Pi(\{D_1, D_2\}) = \text{Perm}(D_1) \cap \text{Perm}(D_2) = \{P_1, P_2\}$$

# Redundant Checks Elimination (1)



$P_1 \in \Pi(\{D_1, D_2\}) \implies n_6$  is redundant

# Redundant Checks Elimination (1)





# Redundant Checks Elimination (2)

- a check node  $n$  for permission  $P$  is *redundant* when:

$$\forall \sigma \in N^*. \quad G \triangleright \sigma : n \implies \sigma : n \vdash P$$

# Redundant Checks Elimination (2)

- a check node  $n$  for permission  $P$  is *redundant* when:

$$\forall \sigma \in N^*. \quad G \triangleright \sigma : n \implies \sigma : n \vdash P$$

- Let  $\tau \models TP^=(G, Perm)$ . For each check node  $n$ , define:

$$\Pi(n) = \bigcap \{ \Pi(\gamma) \mid \gamma \in \tau(n) \}$$

$\Pi(n)$  is the set of permissions (statically) granted to  $n$ .

# Redundant Checks Elimination (2)

- a check node  $n$  for permission  $P$  is *redundant* when:

$$\forall \sigma \in N^*. \quad G \triangleright \sigma : n \implies \sigma : n \vdash P$$

- Let  $\tau \models TP^=(G, Perm)$ . For each check node  $n$ , define:

$$\Pi(n) = \bigcap \{ \Pi(\gamma) \mid \gamma \in \tau(n) \}$$

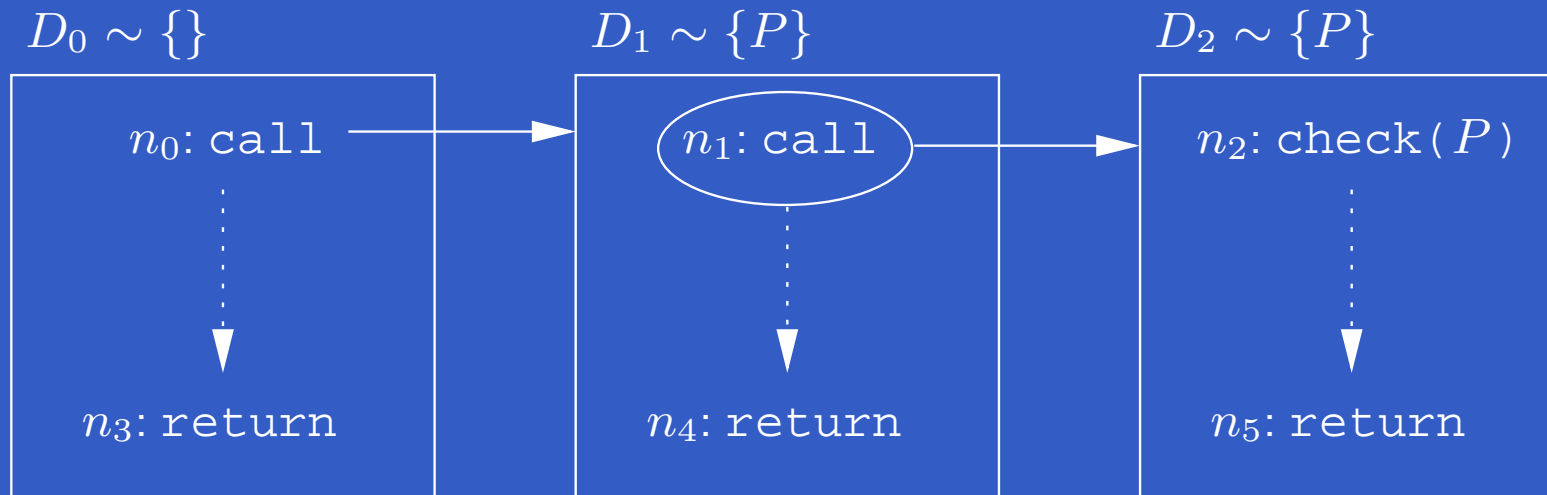
$\Pi(n)$  is the set of permissions (statically) granted to  $n$ .

- Correctness of the optimization:

$$n \text{ is redundant} \iff P \in \Pi(n)$$

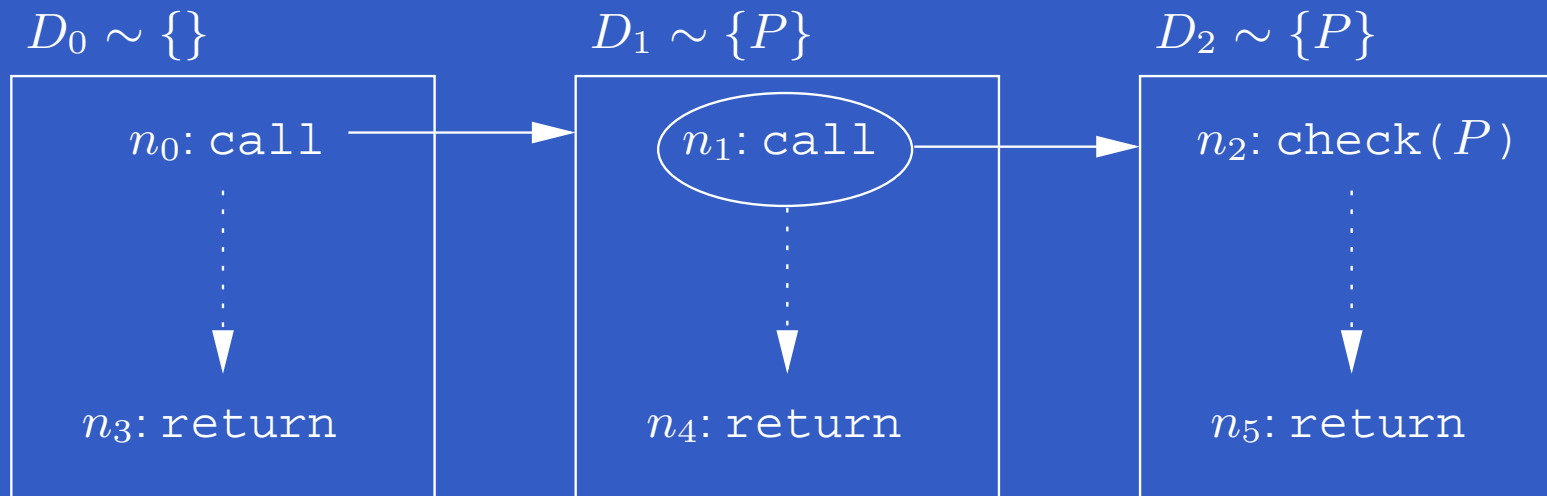
# Method inlining (1)

## Example 1 (before inlining)



# Method inlining (1)

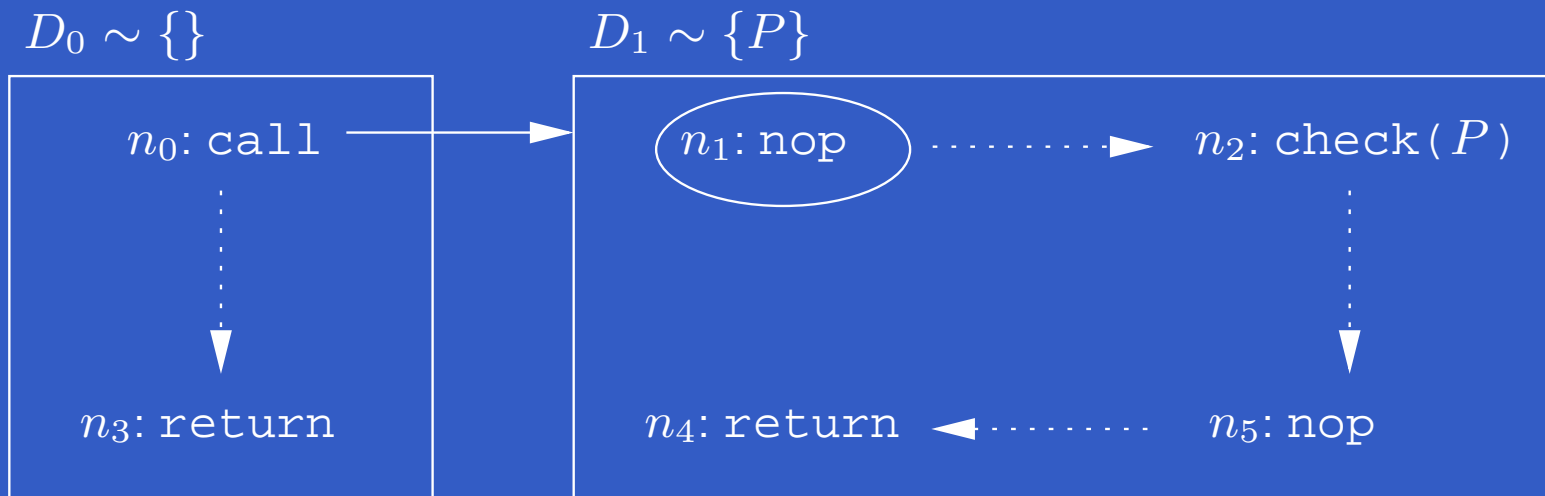
## Example 1 (before inlining)



$$G \triangleright [n_0] \triangleright [n_0, n_1] \triangleright [n_0, n_1, n_2] \vdash P$$

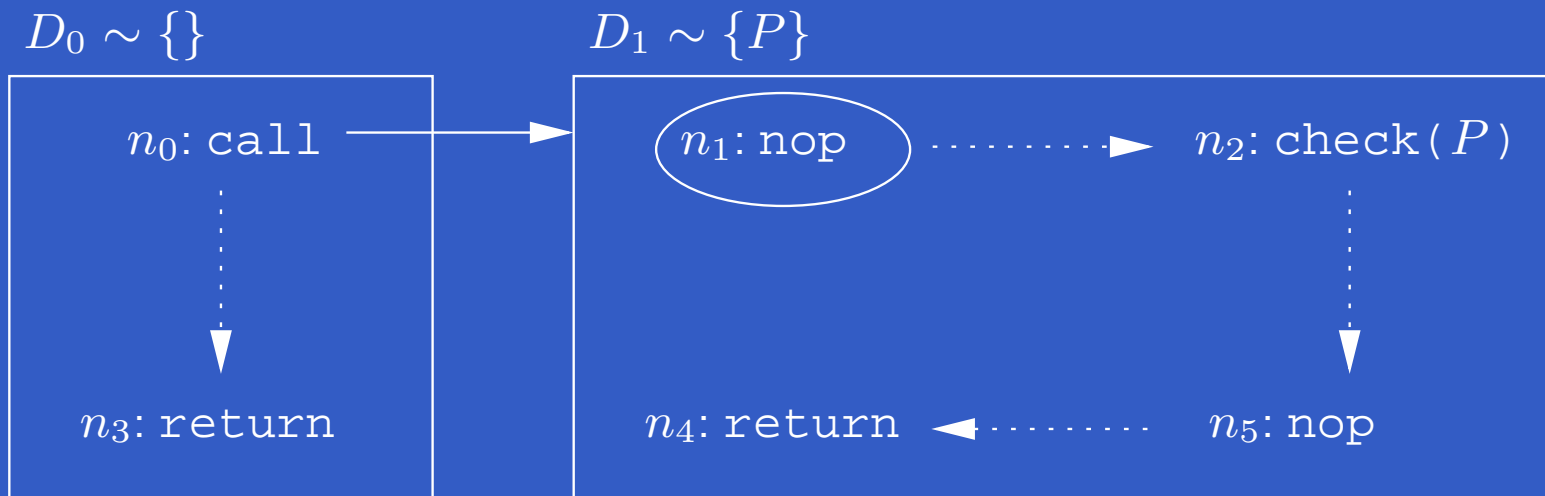
# Method inlining (2)

## Example 1 (after inlining of $n_1$ )



# Method inlining (2)

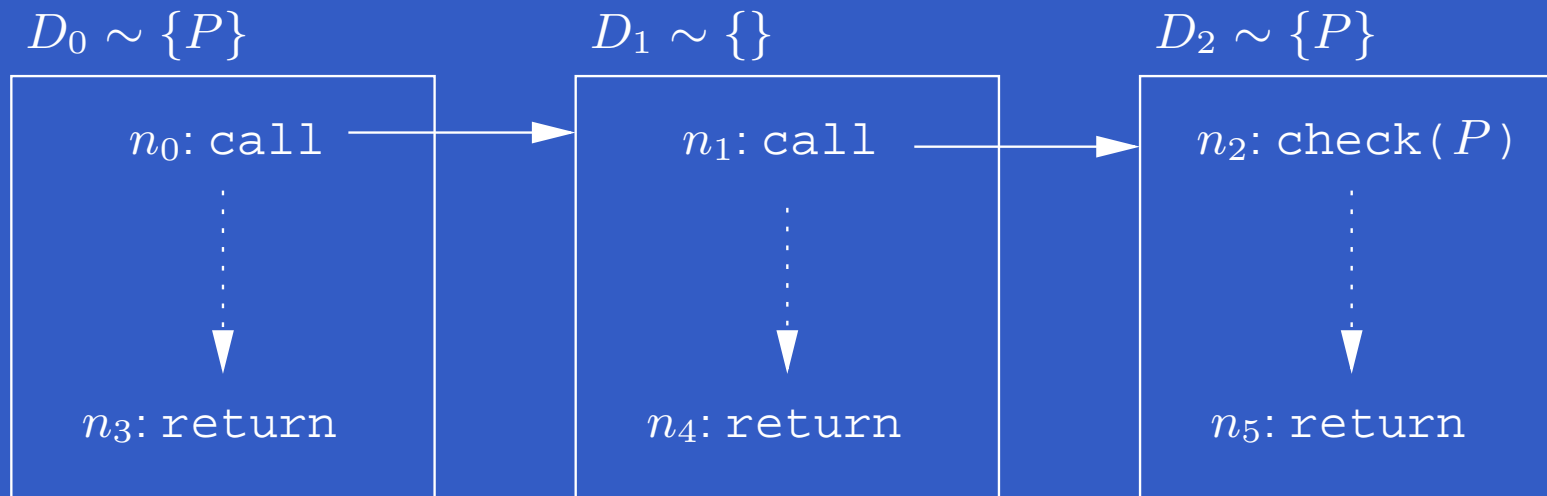
Example 1 (after inlining of  $n_1$ )



$$\dot{G} \triangleright [n_0] \triangleright [n_0, n_1] \triangleright [n_0, n_2] \not\vdash P$$

# Method inlining (3)

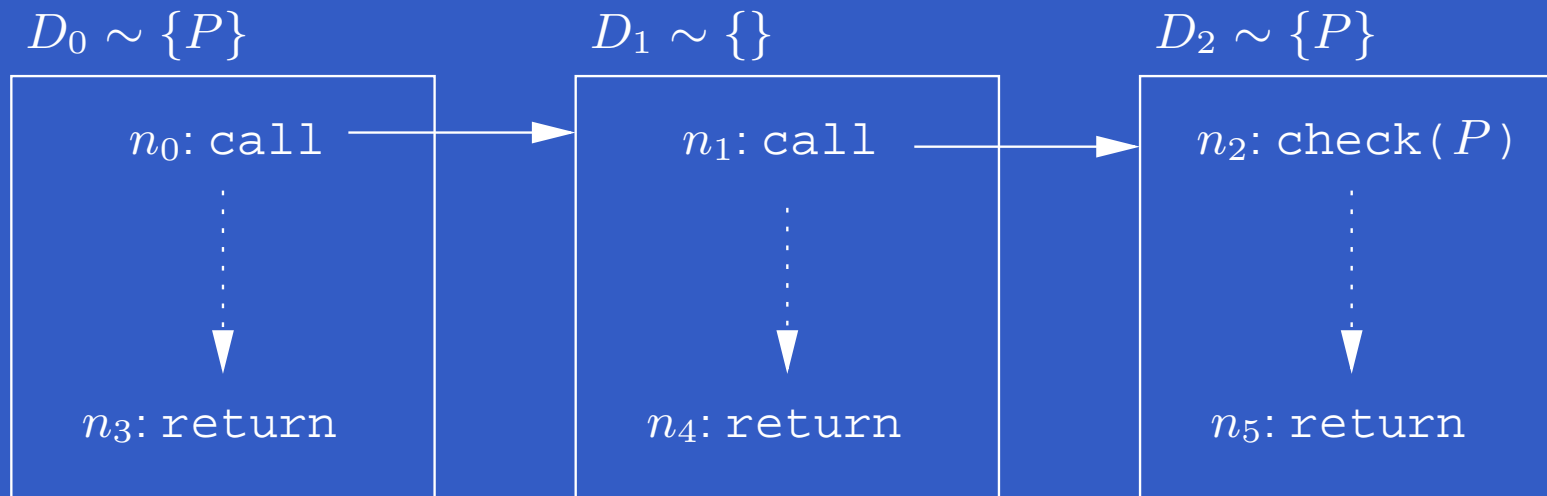
## Example 2 (before inlining)





# Method inlining (3)

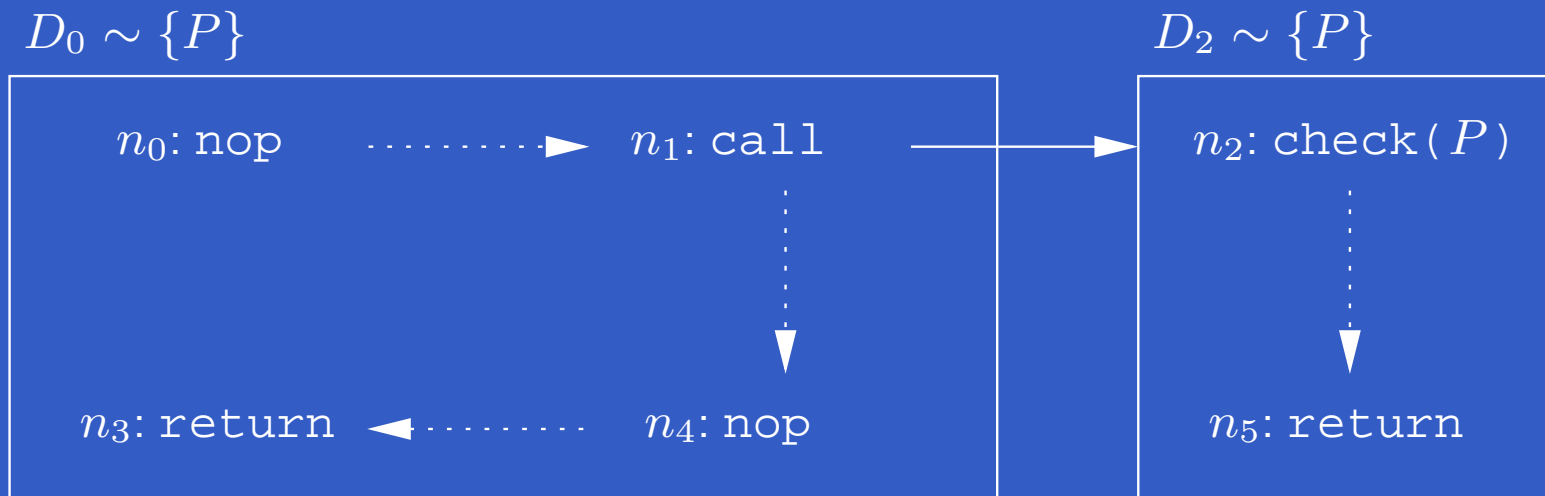
## Example 2 (before inlining)



$G \triangleright [n_0] \triangleright [n_0, n_1] \triangleright [n_0, n_1, n_2] \not\vdash P$

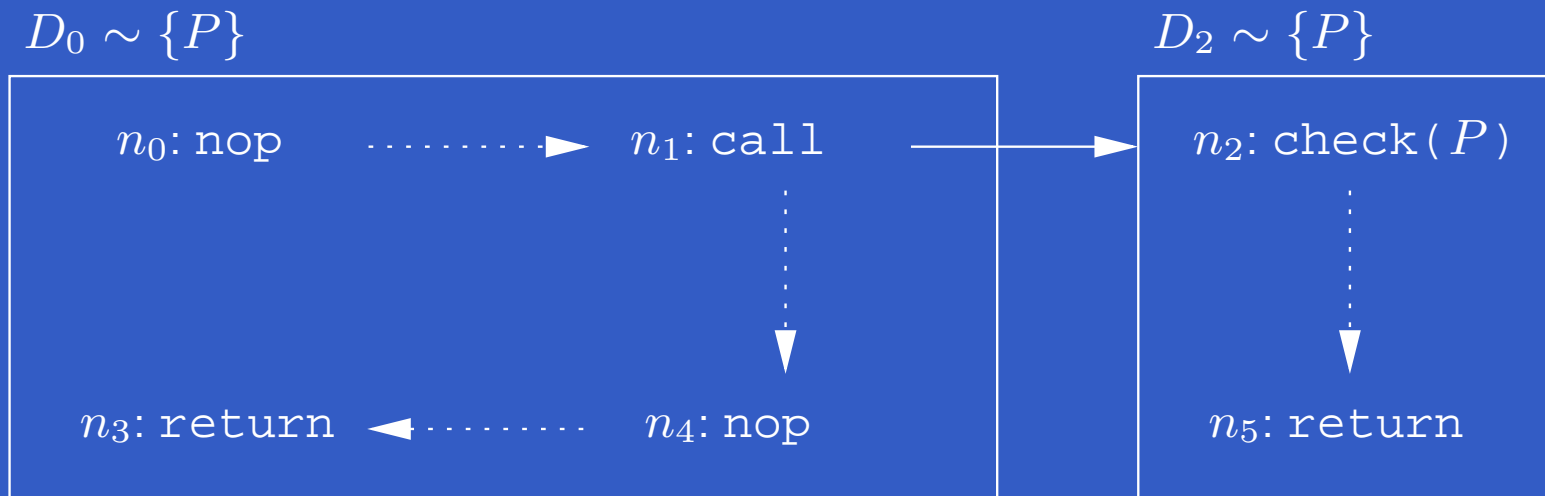
# Method inlining (4)

## Example 2 (after inlining of $n_0$ )



# Method inlining (4)

## Example 2 (after inlining of $n_0$ )



$$\dot{G} \triangleright [n_0] \triangleright [n_1] \triangleright [n_1, n_2] \vdash P$$

# Method inlining (5)

Let  $\dot{n} \longrightarrow n'$  be the call candidate for inlining. We require:

- static dispatching, non-recursiveness:

$$\forall m' \in N. \dot{n} \longrightarrow m' \implies m' = n' \wedge m' \notin \mu(\dot{n})$$

# Method inlining (5)

Let  $\dot{n} \longrightarrow n'$  be the call candidate for inlining. We require:

- static dispatching, non-recursiveness:

$$\forall m' \in N. \dot{n} \longrightarrow m' \implies m' = n' \wedge m' \notin \mu(\dot{n})$$

- original version inlining:

$$\forall m \in N. m \longrightarrow n' \implies m = \dot{n}$$

# Method inlining (5)

Let  $\dot{n} \longrightarrow n'$  be the call candidate for inlining. We require:

- static dispatching, non-recursiveness:

$$\forall m' \in N. \dot{n} \longrightarrow m' \implies m' = n' \wedge m' \notin \mu(\dot{n})$$

- original version inlining:

$$\forall m \in N. m \longrightarrow n' \implies m = \dot{n}$$

- isolation of the protection domain of  $n'$ :

$$\forall m \notin \mu(n'). \text{Dom}(m) \neq \text{Dom}(n')$$

# Method inlining (6)

The key idea is the following:

- method inlining is safe iff the outcome of the security checks is preserved.

# Method inlining (6)

The key idea is the following:

- method inlining is safe iff the outcome of the security checks is preserved.
- let  $\text{Dom}(n) = D$ ,  $\text{Dom}(n') = D'$ . We define:

$$\text{Inl}_n(\gamma) = \begin{cases} \gamma & \text{if } D' \notin \gamma \\ (\gamma \setminus \{D'\}) \cup \{D\} & \text{otherwise} \end{cases}$$



# Method inlining (6)

The key idea is the following:

- method inlining is safe iff the outcome of the security checks is preserved.
- let  $\text{Dom}(\dot{n}) = D$ ,  $\text{Dom}(n') = D'$ . We define:

$$\text{Inl}_{\dot{n}}(\gamma) = \begin{cases} \gamma & \text{if } D' \notin \gamma \\ (\gamma \setminus \{D'\}) \cup \{D\} & \text{otherwise} \end{cases}$$

- the three conditions above guarantee that  $\text{Inl}_{\dot{n}}(\gamma)$  is the context *after* the inlining of  $\dot{n}$ .

# Method inlining (7)

The correctness of method inlining is decided as follows:

- assume that a solution  $\tau$  to the TP analysis is available. We assign a fresh name to  $\text{Dom}(n')$ , then we restart the worklist algorithm from  $\dot{n}$ .

# Method inlining (7)

The correctness of method inlining is decided as follows:

- assume that a solution  $\tau$  to the TP analysis is available. We assign a fresh name to  $\text{Dom}(n')$ , then we restart the worklist algorithm from  $\dot{n}$ .
- each time we reach a node  $\ell(n) = \text{check}(P)$ , we check that, for each context  $\gamma \in \tau(n)$ ,

$$P \in \Pi(\gamma) \iff P \in \Pi(\text{Inl}_{\dot{n}}(\gamma))$$

# Method inlining (7)

The correctness of method inlining is decided as follows:

- assume that a solution  $\tau$  to the TP analysis is available. We assign a fresh name to  $\text{Dom}(n')$ , then we restart the worklist algorithm from  $\dot{n}$ .
- each time we reach a node  $\ell(n) = \text{check}(P)$ , we check that, for each context  $\gamma \in \tau(n)$ ,

$$P \in \Pi(\gamma) \iff P \in \Pi(\text{Inl}_{\dot{n}}(\gamma))$$

- if this is true for each node reached after the call  $\dot{n}$ , then  $\dot{n}$  is inlineable in  $G$ .

# Method inlining (8)

We define the  *$\dot{n}$ -inlined version* of a CFG

$G = \langle N \cup \{n_\varepsilon\}, E, \text{Priv}_G, \text{Dom}_G \rangle$  as

$\dot{G} = \langle N \cup \{n_\varepsilon\}, E, \text{Priv}_{\dot{G}}, \text{Dom}_{\dot{G}} \rangle$ , where:

$$\text{Priv}_{\dot{G}}(n) = \begin{cases} \text{true} & \text{if } \text{Priv}_G(\dot{n}) \text{ and } \dot{n} \longrightarrow \mu(n) \\ \text{Priv}_G(n) & \text{otherwise} \end{cases}$$

$$\text{Dom}_{\dot{G}}(n) = \begin{cases} \text{Dom}_G(\dot{n}) & \text{if } \dot{n} \longrightarrow \mu(n) \\ \text{Dom}_G(n) & \text{otherwise} \end{cases}$$

# Method inlining (9)

## Method call

$$\frac{\ell(n) = \text{call} \quad n \longrightarrow n' \quad n \neq \dot{n}}{\sigma : n \triangleright_{inl}^{\dot{n}} \sigma : n : n'}$$

$$\frac{\ell(\dot{n}) = \text{call} \quad \dot{n} \longrightarrow n'}{\sigma : \dot{n} \triangleright_{inl}^{\dot{n}} \sigma : n'}$$

# Method inlining (10)

## Method return

$$\frac{\ell(n') = \text{return} \quad \dot{n} \dashrightarrow m \quad \dot{n} \not\rightarrow \mu(n')}{\sigma : n : n' \triangleright_{inl}^{\dot{n}} \sigma : m}$$

$$\frac{\ell(n') = \text{return} \quad \dot{n} \dashrightarrow m \quad \dot{n} \longrightarrow \mu(n')}{\sigma : n' \triangleright_{inl}^{\dot{n}} \sigma : m}$$

# Method inlining (11)

Theorem (*Correctness of method inlining.*)

If  $\dot{n}$  is inlineable in  $G$  and  $\dot{G}$  is the  $\dot{n}$ -inlined version of  $G$ :

$$\langle \sigma_0, x_0 \rangle \triangleright \cdots \triangleright \langle \sigma_k, x_k \rangle$$

$$\iff$$

$$\langle \dot{\sigma}_0, x_0 \rangle \triangleright_{\dot{n}_{inl}} \cdots \triangleright_{\dot{n}_{inl}} \langle \dot{\sigma}_k, x_k \rangle$$

where  $\sigma_0 = []$ ,  $x_0 = false$ , and  $\dot{\sigma}_i = inl_{\dot{n}}(\sigma_i)$  for  $i \in 0..k$ .

$$\frac{inl_{\dot{n}}(\sigma) = \dot{\sigma} \quad top(\sigma) \neq \dot{n}}{inl_{\dot{n}}(\sigma : n') = \dot{\sigma} : n'}$$

$$\frac{inl_{\dot{n}}(\sigma) = \dot{\sigma}}{inl_{\dot{n}}(\sigma : \dot{n} : n') = \dot{\sigma} : n'}$$



# Conclusions

- interprocedural optimizations in presence of stack inspection
  - + based on solid static techniques (CFA)
  - + no update of the security context
  - + dynamic linking is possible
    - overhead at linking time / deoptimization
- TO DO:
  - parametric permissions (ongoing work)
  - dynamic policies (ongoing work)
  - implementation & performance evaluation

# Appendix - Def. of the TP Analysis (1)

$$TP_{in}(n) = \bigcup_{(m,n) \in E} TP_{out}(m, n)$$

$$TP_{out}(m, n) = \begin{cases} \{\{\mathbf{Dom}(n)\}\} & \text{if } \bullet \longrightarrow n \\ \{\gamma \cup \{\mathbf{Dom}(n)\} \mid \gamma \in TP_{call}(m)\} & \text{if } m \longrightarrow n \\ TP_{trans}(m) & \text{if } m \dashrightarrow n \\ TP_{catch}(m) & \text{if } m \dashrightarrow \text{⚡} n \end{cases}$$

$$TP_{call}(n) = \begin{cases} \{\{\mathbf{Dom}(n)\}\} & \text{if } \mathbf{Priv}(n) \text{ and } TP_{in}(n) \neq \emptyset \\ TP_{in}(n) & \text{otherwise} \end{cases}$$

# Appendix - Def. of the TP Analysis (2)

$$TP_{trans}(n) = \begin{cases} \{\gamma \in TP_{in}(n) \mid P \in \Pi(\gamma)\} & \text{if } \ell(n) = \text{check}(P) \\ \{\gamma \in TP_{in}(n) \mid \text{Trans}(n, \{\text{Dom}(n)\})\} & \text{if } \ell(n) = \text{call}, \text{Priv}(n) \\ \{\gamma \in TP_{in}(n) \mid \text{Trans}(n, \gamma)\} & \text{otherwise} \end{cases}$$

$$\text{Trans}(n, \gamma) \stackrel{\text{def}}{=} \exists m \in \rho(n). \gamma \cup \{\text{Dom}(m)\} \in TP_{in}(m)$$

$$TP_{catch}(n) = \begin{cases} \{\gamma \in TP_{in}(n) \mid P \notin \Pi(\gamma)\} & \text{if } \ell(n) = \text{check}(P) \\ \{\gamma \in TP_{in}(n) \mid \text{Catch}(n, \{\text{Dom}(n)\})\} & \text{if } \ell(n) = \text{call}, \text{Priv}(n) \\ \{\gamma \in TP_{in}(n) \mid \text{Catch}(n, \gamma)\} & \text{otherwise} \end{cases}$$

$$\text{Catch}(n, \gamma) \stackrel{\text{def}}{=} \exists m \in \xi_1(n). \gamma \cup \{\text{Dom}(m)\} \in TP_{catch}(m)$$