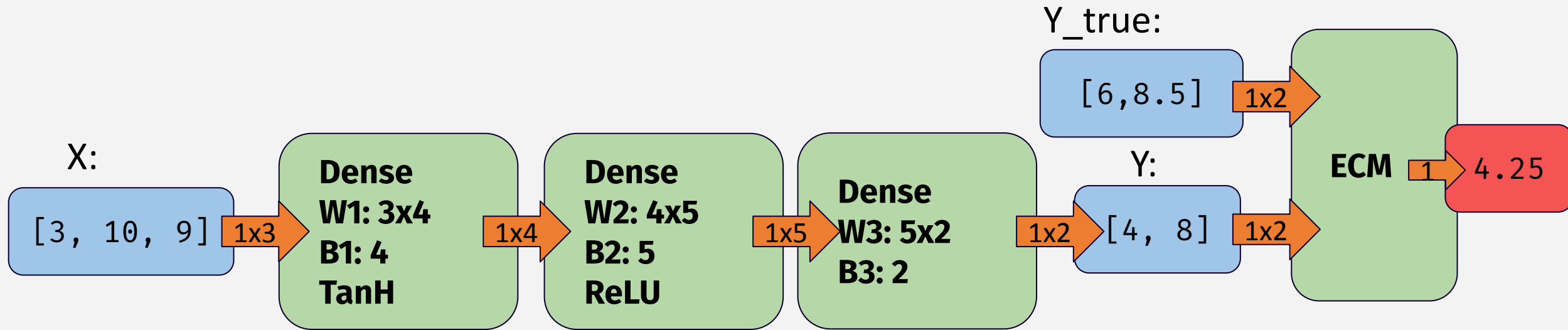


# Backpropagation en Redes Secuenciales

---

# Backpropagation - Capas



- **Forward**

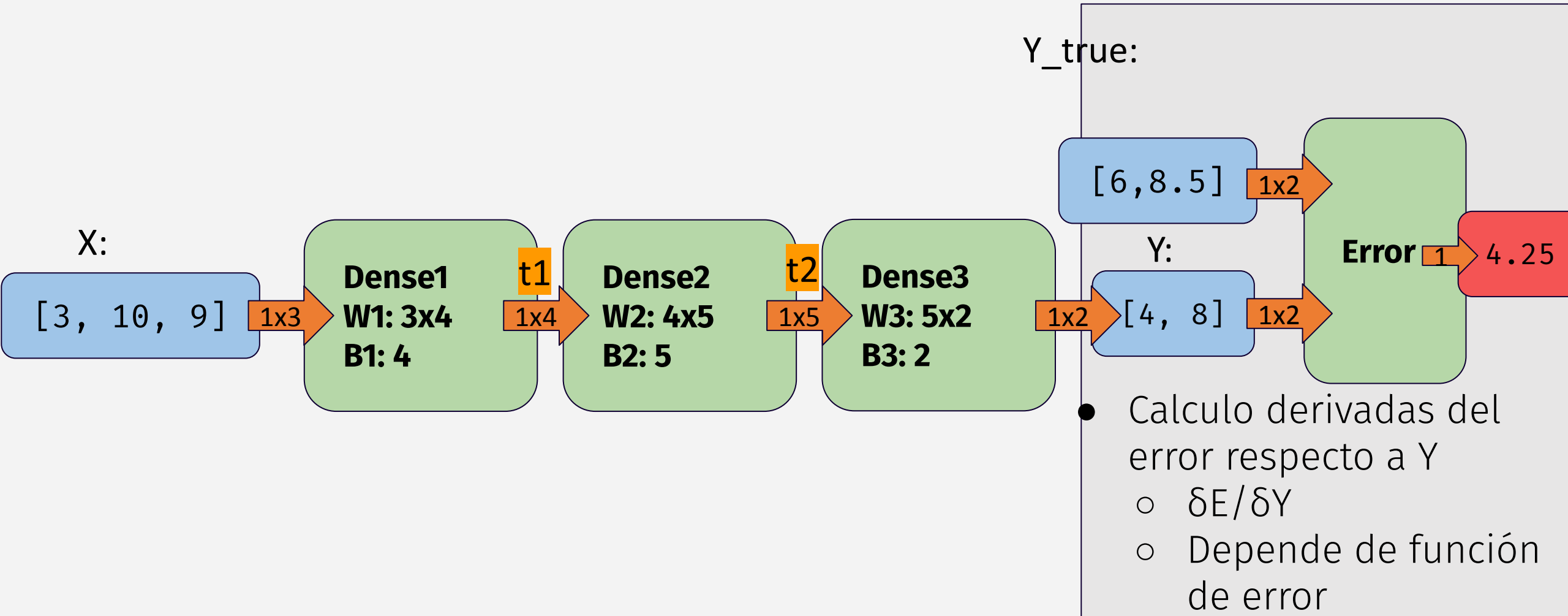
- De entrada a salida
- *Almacenar* valores intermedios

- **Backward**

- Derivadas de salida a entrada
- *Utilizar* valores intermedios

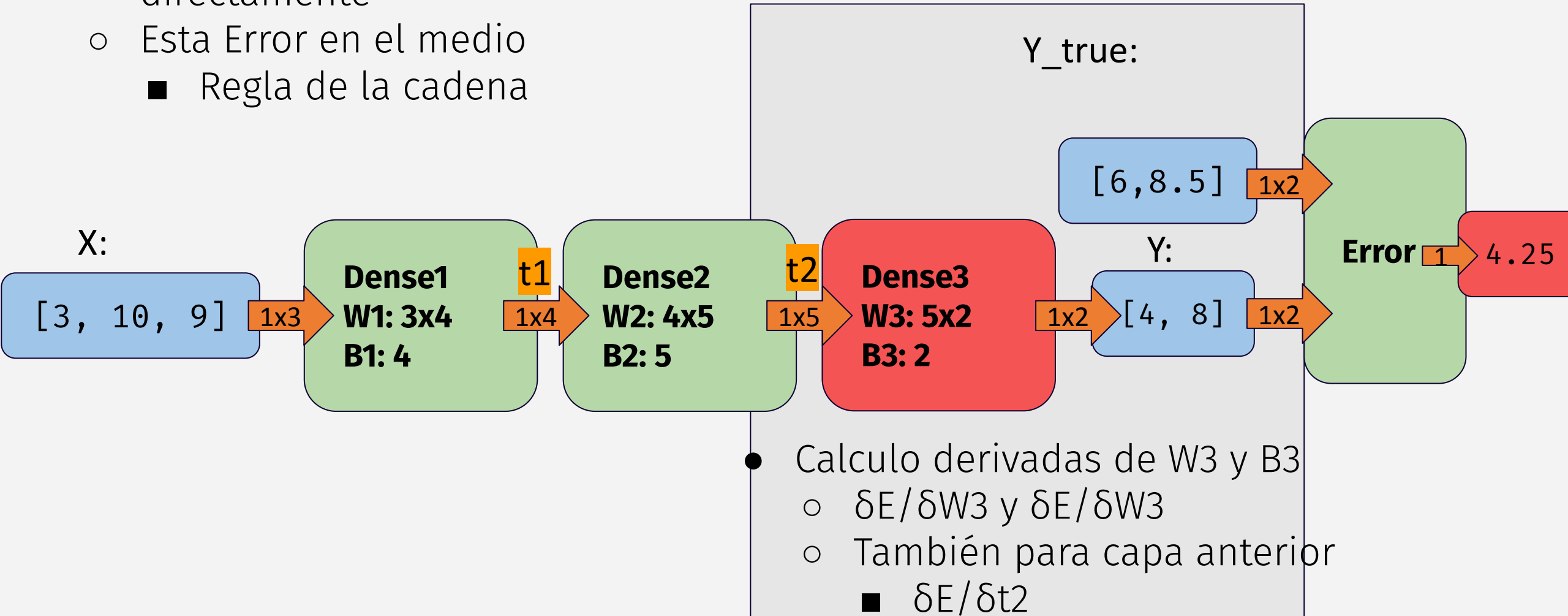
# backward: Cálculo de derivadas por capas

- Derivada del error
  - Se calcula directamente



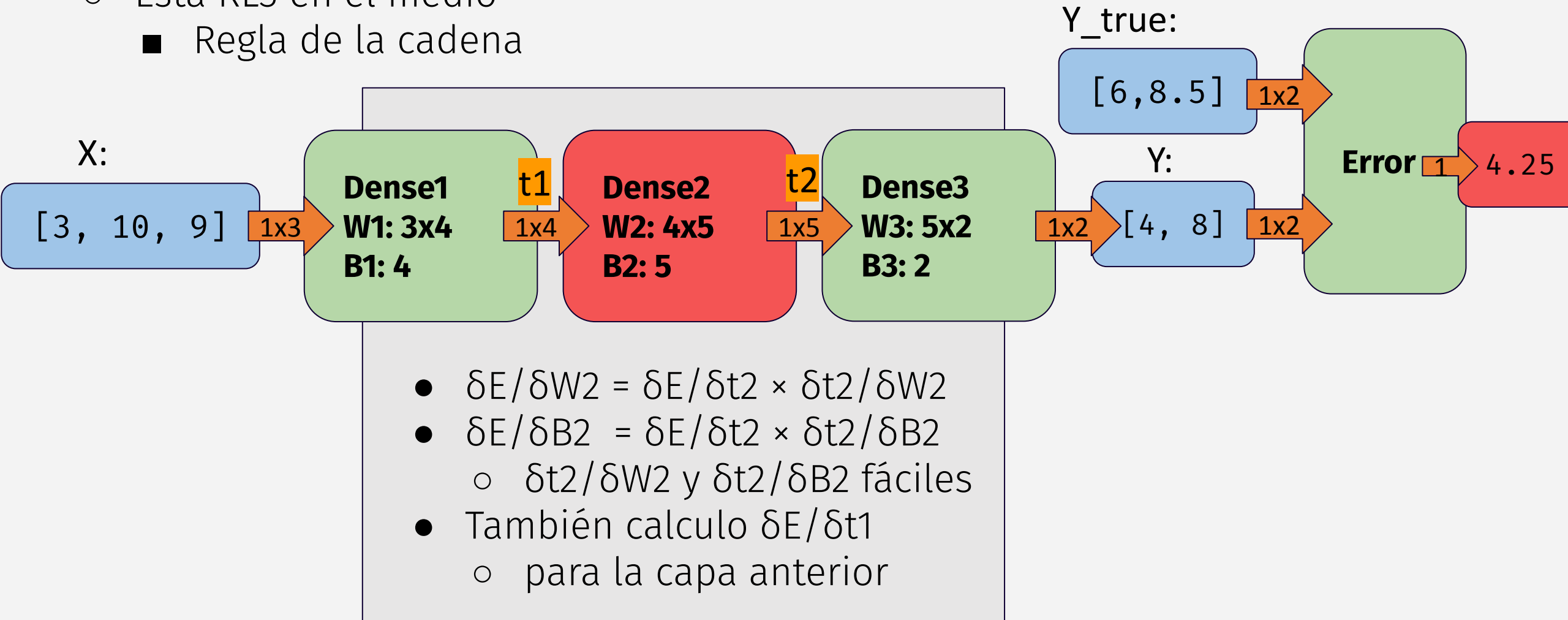
# backward: Cálculo de derivadas por capas

- $\delta E / \delta W_3$  y  $\delta E / \delta B_3$ 
  - No se pueden calcular directamente
  - Esta Error en el medio
    - Regla de la cadena



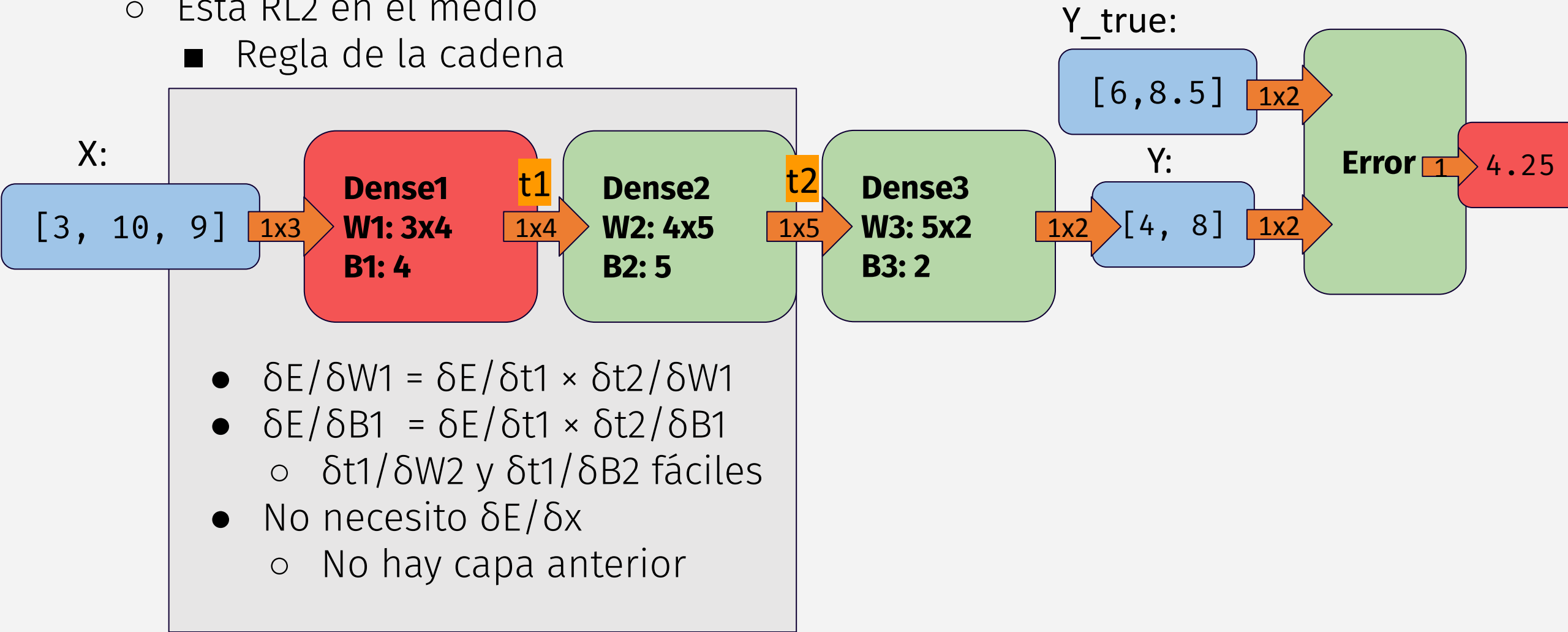
# backward: Cálculo de derivadas por capas

- $\delta E / \delta W2$  y  $\delta E / \delta B2$ 
  - No se pueden calcular directamente
  - Esta RL3 en el medio
    - Regla de la cadena



# backward: Cálculo de derivadas por capas

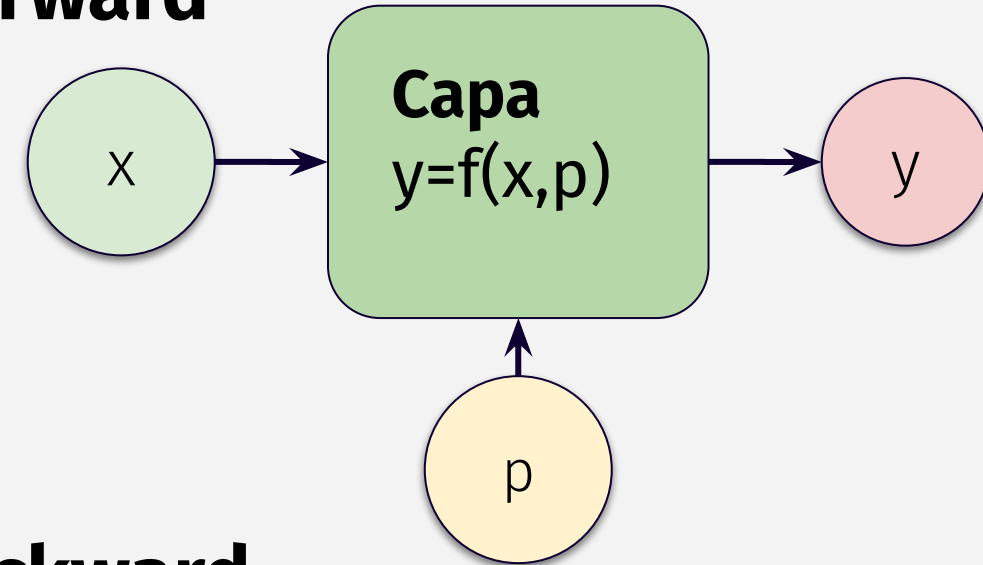
- $\delta E / \delta W1$  y  $\delta E / \delta B1$ 
  - No se pueden calcular directamente
  - Esta RL2 en el medio
    - Regla de la cadena



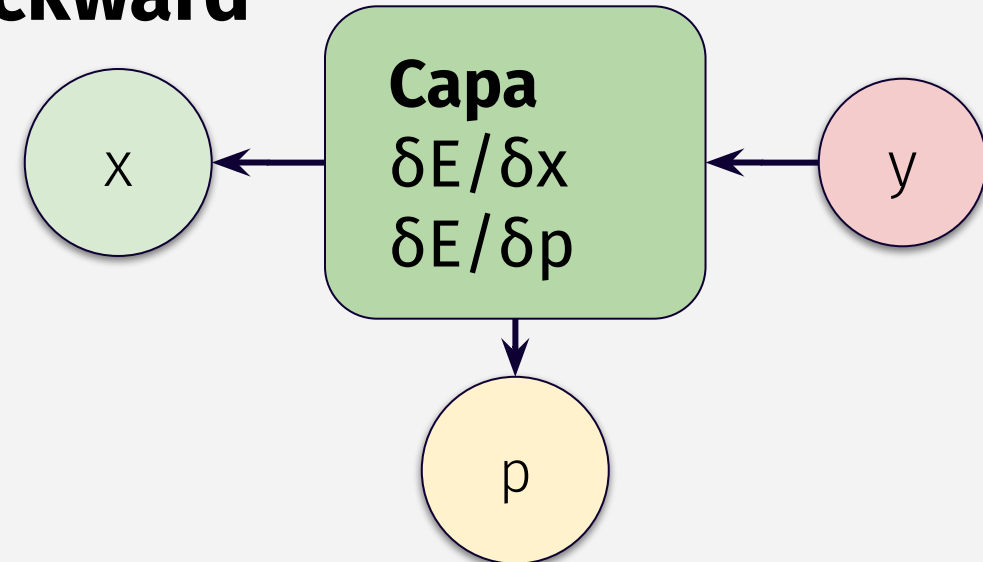
# Cálculo modular por Capas

- Cada capa debe saber calcular
  - Forward:  $y = f(x)$
  - Backward
    - $\delta E / \delta x = \delta E / \delta y * \delta E y / \delta x$
    - $\delta E / \delta p = \delta E / \delta y * \delta E y / \delta p$
- Modularidad
  - Capas independientes
  - Puedo intercambiar

## Forward



## Backward



# Descenso de gradiente en redes

- Descenso de gradiente
  - Igual que antes
  - Más parámetros
  - Misma idea
    - $w = w - \alpha \delta E / \delta w$
- Cálculo de derivadas
  - Backpropagation
  - Frameworks actuales
    - Especificás el **forward**
      - Capas/arquitectura de la red
    - El framework determina el backward
      - Backpropagation automático
    - Funciona por lotes

