INFN Istituto Nazionale di Fisica Nucleare

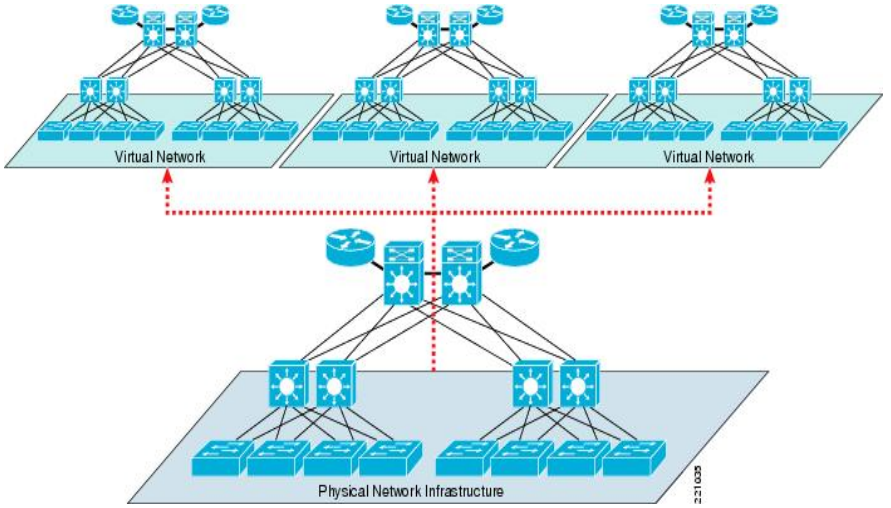
WNoDeS
Dynamic Virtual Networks

Sommario

- Contesto e requisiti
- Reti virtuali dinamiche
- Scelte implementative
- Architettura: descrizione e proprietà
- Test e risultati
- Integrazione in WNoDeS
- Sviluppi futuri

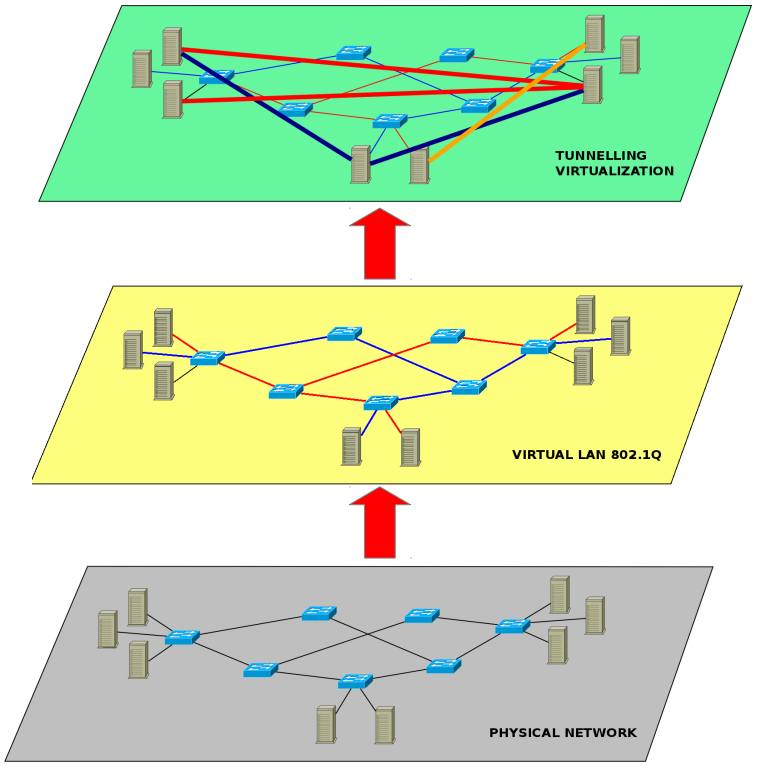
Contesto di utilizzo

- Presso il CNAF viene utilizzato WNoDeS anche per fornire servizi Cloud.
- In un cloud IaaS, i consumatori spesso richiedono un accesso con privilegi di root alle loro VM.
- Si vuole realizzare un sistema di **virtualizzazione della rete** per la separazione del traffico degli utenti.
- Proprietà:
  - **Scalabilità**
  - **Resilienza**
  - **Sicurezza**
  - **Disponibilità**



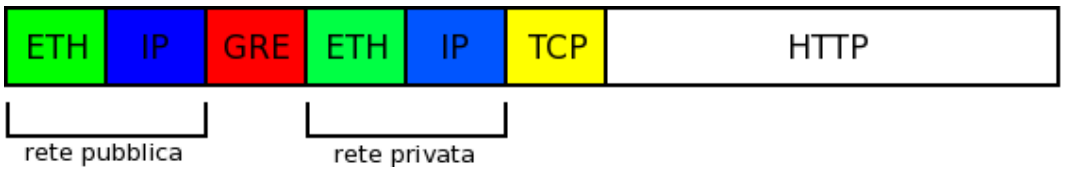
Dynamic Virtual Networks

- La definizione dinamica di VLAN 802.1Q, se pur teoricamente possibile, spesso è non praticabile.
- Requisiti fondamentali del sistema:
  - → **mobilità** delle entità
  - → **scalabilità**
  - → **isolamento** del traffico
  - → non alterare la rete fisica
- Nuova virtualizzazione a livello L3 → **tunneling** mediante protocollo GRE.



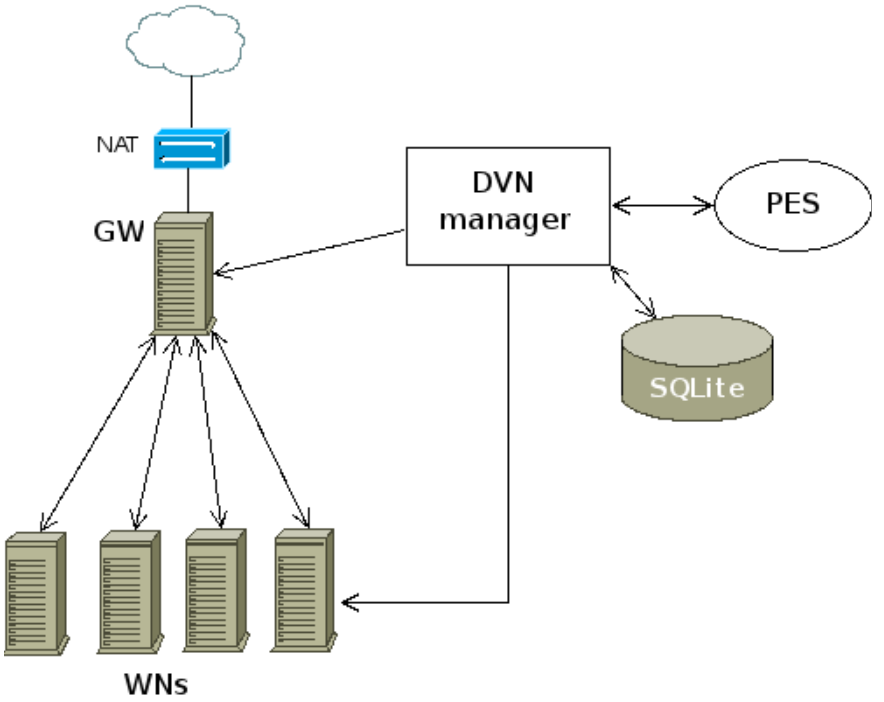
Generic Routing Encapsulation

- **Standard** definito nella RFC 2890.
- Compatibile con la maggior parte dei sistemi operativi (escluso FreeBSD) e con la quasi totalità dei router.
- Tunnel che incapsulano un frame Ethernet all'interno di un datagram IP ("**Ethernet over IP**").
- Supportato dagli switch virtuali, come Open vSwitch.
- Interfacce **gretap** definite mediante iproute2, si possono connettere direttamente ai bridge.



Architettura

- Tutto il traffico delle macchine cloud circola all'interno dei tunnel.
- Topologia "**Hub and Spoke**".
- Ogni rete virtuale usa una subnet differente.
- Componenti:
  - Nodo GW e Compute Resource
  - DVN manager
  - Policy Enforcement Service



Policy Enforcement Service

- Componente **centralizzato** con lo scopo di contenere le definizioni delle policy delle reti virtuali.
- Le policy sono definite in un **meta-linguaggio** di alto livello.
- Il PES si occupa di tradurre le policy in file di configurazione da applicare sui nodi.
- Soluzione pensata per gestire policy per **nodi eterogenei**.
- Si occupa anche di generare **regole anti-spoofing** di MAC e IP.

Realizzazione del PES

- **Esempio**

XML policy file

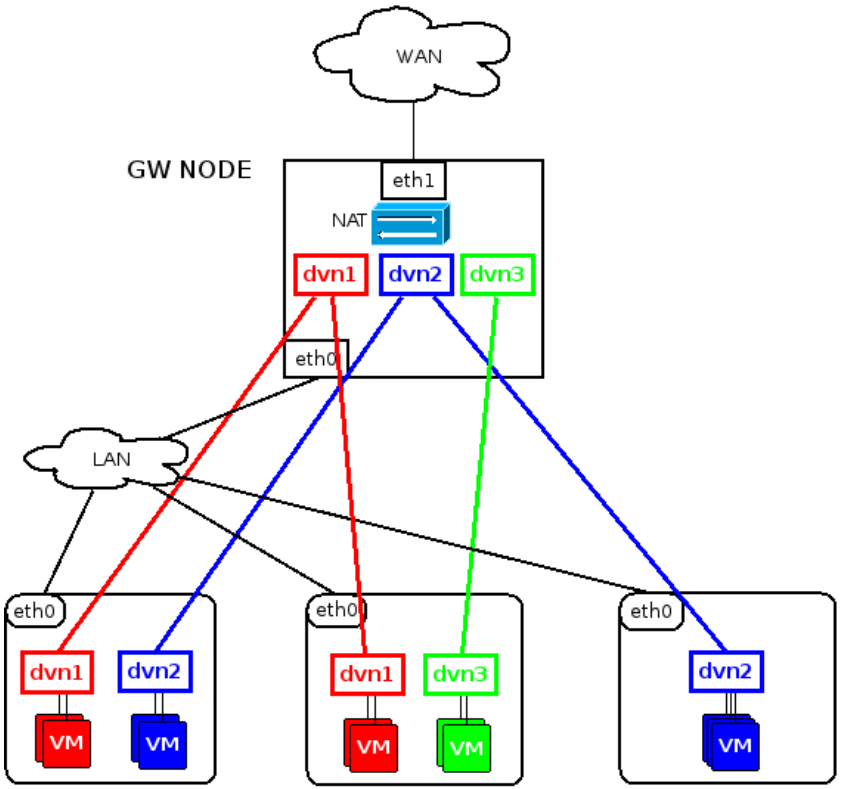
```
<network type='virtual'>
     <name>wlan0001</name>
     ...
     <bridge>wlan0001</bridge>
     ...
     <dns>enabled</dns>
     <dhcp>enabled</dhcp>
     <subnet>
          <address>10.0.1.0</address>
          <mask>24</mask>
     </subnet>
     <firewall>
          <tcp>
               <whitePort>http</whitePort>
               <whitePort>https</whitePort>
               <whitePort>ssh</whitePort>
          </tcp>
          <default>drop</default>
      </firewall>
</network>
```

IPTABLES script

```
iptables -N wnod-wlan0001

iptables -A INPUT -i wlan0001 -j wnod-wlan0001
iptables -A FORWARD -i wlan0001 -j wnod-wlan0001
iptables -A OUTPUT -o wlan0001 -j wnod-wlan0001

iptables -A wnod-wlan0001 -s 10.0.1.0/24 -p udp \
     -m multiport --dports domain -j ACCEPT

iptables -A wnod-wlan0001 -s 10.0.1.0/24 -p udp \
     -m multiport --dports bootps:bootpc -j ACCEPT

iptables -A wnod-wlan0001 -s 10.0.1.0/24 -p tcp \
     -m multiport --dports http,https,ssh -j ACCEPT

iptables -A wnod-wlan0001 -j DROP
```

Dettagli della soluzione

- Usando gli strumenti di networking Linux, viene creato un diverso bridge per ogni rete virtuale.
- Sui Compute Resource:
  - un bridge e un'interfaccia GRE per ogni rete virtuale.
- Sul nodo GW:
  - Ad ogni bridge sono connessi tanti tunnel quanti sono gli host fisici coinvolti nella rete virtuale.



Caratteristiche 1/2

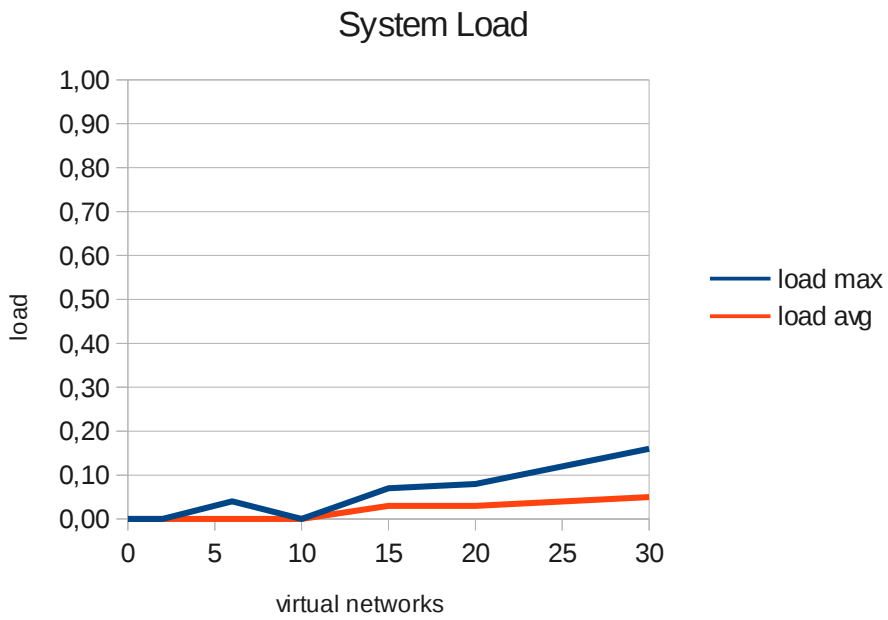
- **Scalabilità**
  - Il numero di reti virtuali dipende da come si partiziona il pool delle sotto-reti;
  - Per aggiungere nuovi host basta configurare un bridge con un tunnel GRE verso il nodo GW.
- **Resilienza**
  - Il nodo GW è il *single point of failure* dell'architettura.

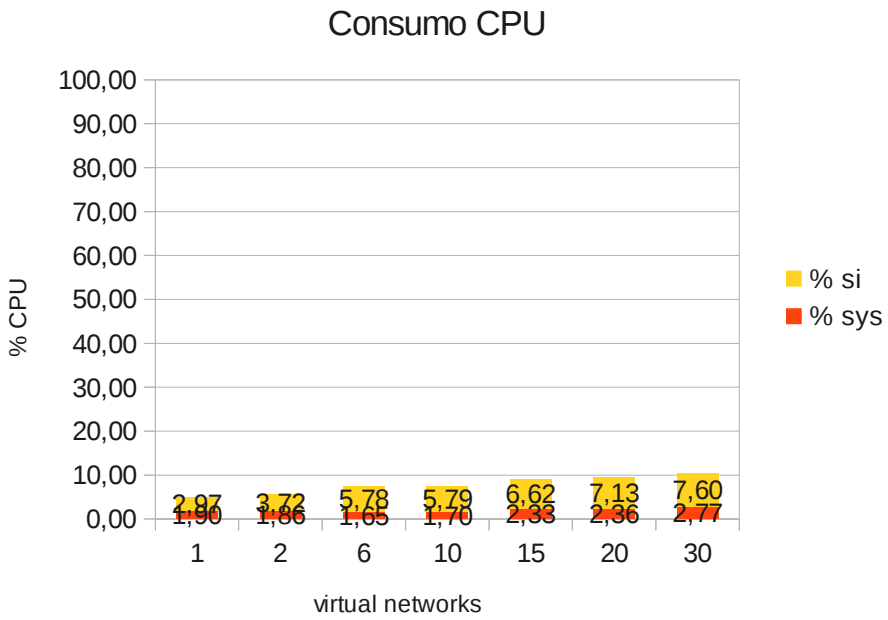
Caratteristiche 2/2

- **Sicurezza**
  - L'isolamento a L2 viene soddisfatto mediante separazione dei bridge.
  - Il forwarding a livello del kernel è disabilitato.
  - A L3 ogni utente ha una sotto-rete differente e viene impedito il routing fra le diverse sotto-reti.
- **Disponibilità**
  - L'accesso alle risorse di rete è strettamente legato al funzionamento del nodo GW.

Performance

- **Test di carico** effettuati per valutare il comportamento del nodo GW.
- In ciascuna rete virtuale il traffico viene generato mediante *iperf* per 10 minuti.





Limitazioni e soluzioni

- Il nodo GW è il collo di bottiglia

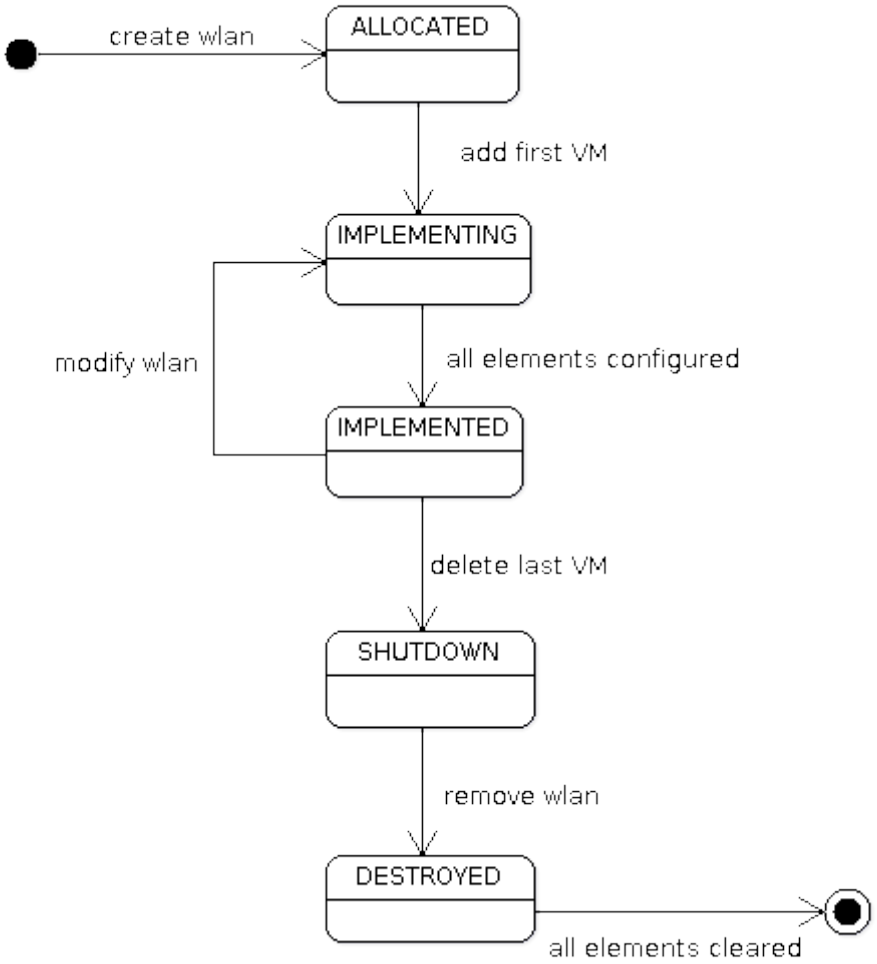
  ⟶ può essere replicato
- Il nodo GW è interamente realizzato in software

  ⟶ può essere sostituito da router hardware
- Le limitazioni sono dovute alla topologia “hub and spoke”

  ⟶ si può passare all'uso di gruppi multicast (filosofia VxLAN).

Integrazione in WNoDeS

- Ogni rete virtuale è chiamata *WLAN* (WNoDeS LAN)
- Ad ogni WLAN è associato un numero intero, chiamato *WID* (WNoDeS ID) e una subnet di classe C (10.*a*.*b*.0/24).
- Tutti i nodi hanno un "demone" che gestisce la creazione/rimozione di bridge e tunnel.
- *DVN manager* invia i comandi di configurazione ai demoni mediante messaggi TCP/IP.



Sviluppi futuri

- Risoluzione delle limitazioni trovate: abbandono della topologia “hub and spoke” in favore dei gruppi multicast.
- Ampliare i servizi offerti ai consumatori: load balancing, QoS, VPN.
- Supporto per l'allocazione di reti virtuali in configurazione multi-sito.
- Cluster on-demand.

Conclusioni

- Soluzione di virtualizzazione della rete mediante protocollo di **tunneling**.
- Scelto GRE in una topologia "**Hub and Spoke**"
- **Replicazione** di bridge e tunnel per isolamento
- Il lavoro svolto verrà presentato a **CHEP 2012** come poster dal titolo:

  "Creating Dynamic Virtual Networks for network isolation to support Cloud computing and virtualization in large computing centers"