Sicurezza nei modelli peer-to-peer

F.Baiardi
Dipartimento di Informatica, Centro Serra
Università di Pisa
f.baiardi@unipi.it

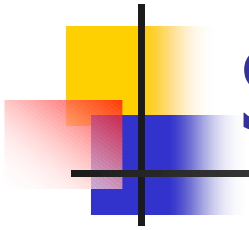
Credits

- Stefano Suin (unipi serra)
- Claudio Telmon
- Paolo Mori (iit cnr)
- Gli studenti del corso Sicurezza delle reti

Peer-to-peer

- Definizione di un servizio (calcolo, memorizzazione) mediante condivisione di risorse
- condivisione
  - A livello di risorse logiche
  - A livello di risorse fisiche

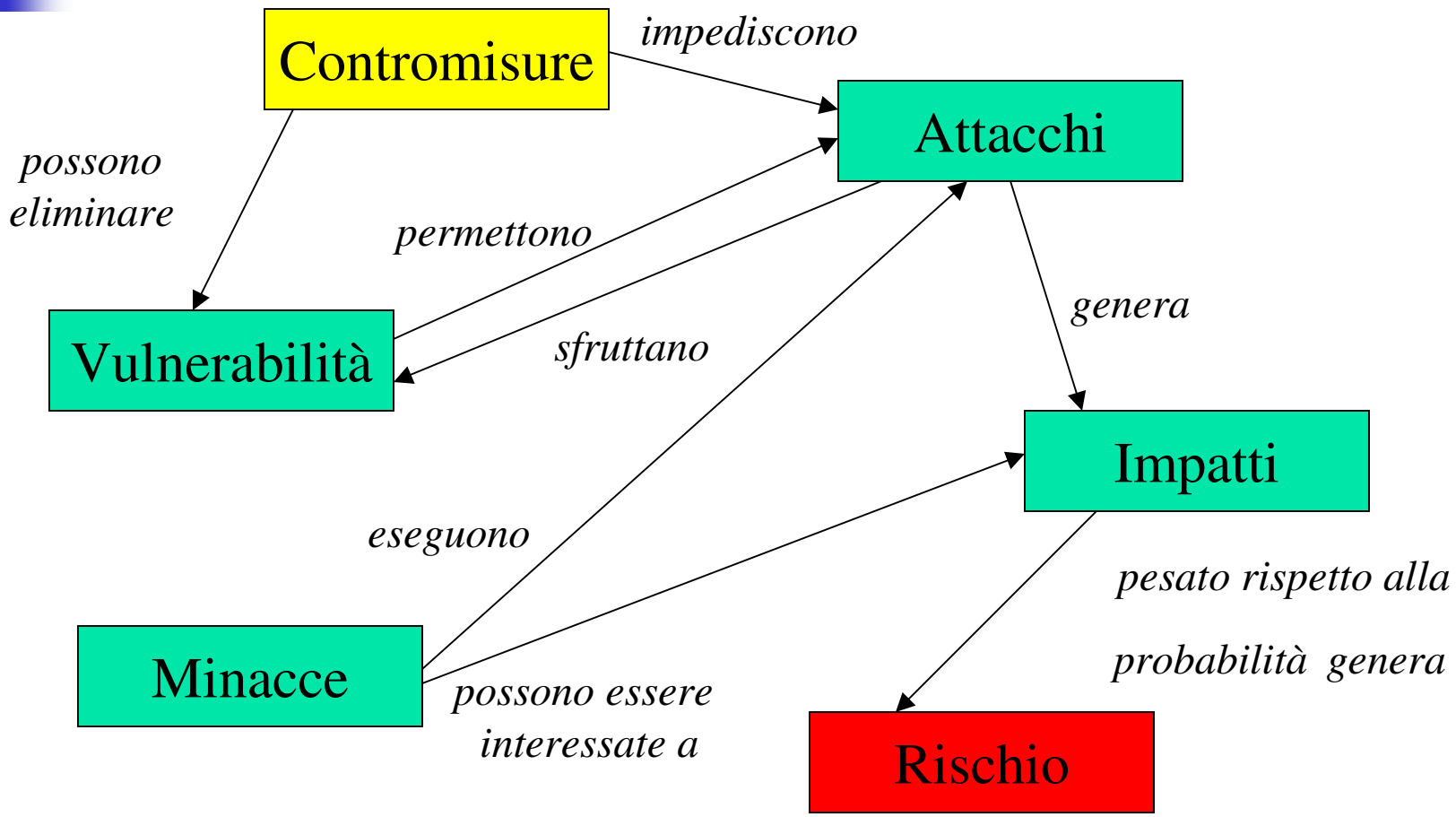
  nel range da file exchange al grid computing, la condivisione è comunque presente
- autonomia
  - dei singoli nodi
  - sull'unirsi o lasciare una rete p2p



Sicurezza

- Tre attributi fondamentali per parlare di sistema sicuro
  - Confidenzialità delle informazioni (dati)
  - Integrita
    - Delle risorse
    - Dei dati
  - Disponibilità
    - Delle risorse
    - Dei dati

Terminologia



Meccanismi fondamentali

- Controllo e gestione delle risorse
- Ridondanza
  - Nei controlli
  - Nelle risorse
- Crittografia
  - Chiave privata
  - Chiave pubblica
- Firma elettronica
- Protocolli sicuri per scambio informazioni

Sicurezza peer-to-peer - I

- Tre prospettive
  - Chi fornisce la risorsa condivisa F
  - Chi accede la risorsa condivisa U
  - Attacco esterno (situazione classica)
- Le prime due
  - importanti per definire un modello di business
  - attualmente risolte mediante autenticazione ed in base alla fiducia tra gli utenti
  - é realistico se la fiducia dipende non dalla persona ma da come la persona (o chi per lui) gestisce un sistema????

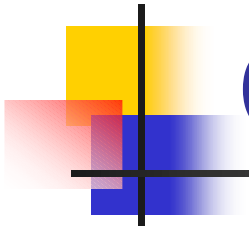
Sicurezza peer-to-peer - II

Asimmetria fondamentale

- Sono disponibili metodi e strumenti per proteggere F = chi fornisce risorse condivise in modo da garantire una certo livello di controllo sulla condivisione
- Non disponibili metodi e tecnologie generali per salvaguardare U, chi accede ed usa una risorsa condivisa, da F, chi fornisce la risorsa

Terzo punto di vista

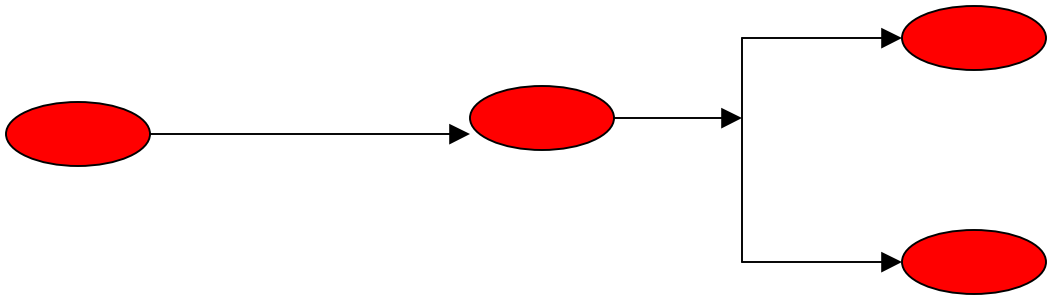
- Attaccante esterno
- Ha come obiettivo
  - F
  - U
  - Le risorse logiche/fisiche del sistema p2p
  - Il servizio
- Ha a disposizione tutte le soluzioni classiche per rete informatica più quelle permesse dalla tecnologia p2p



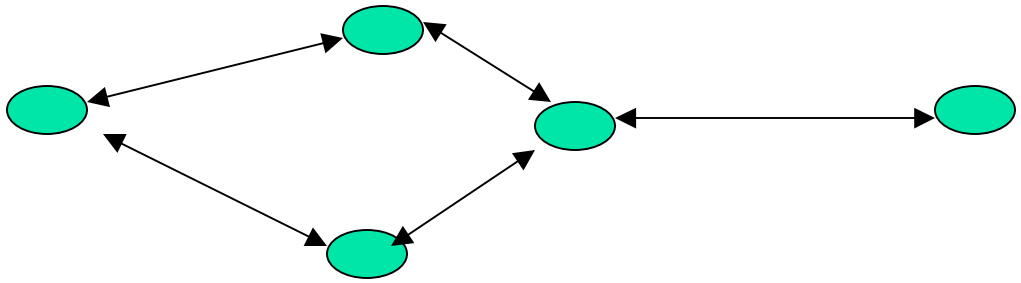
Overlay

- Nozione fondamentale nel p2p è quella di overlay
- Definizione di un nuovo servizio distribuito a partire dai servizi di comunicazione esistenti
- I nuovi servizi utilizzano sostanzialmente i servizi di comunicazione e di risoluzione dei nomi esistenti
- Questa nozione è fondamentale per la sicurezza

Overlay della topologia



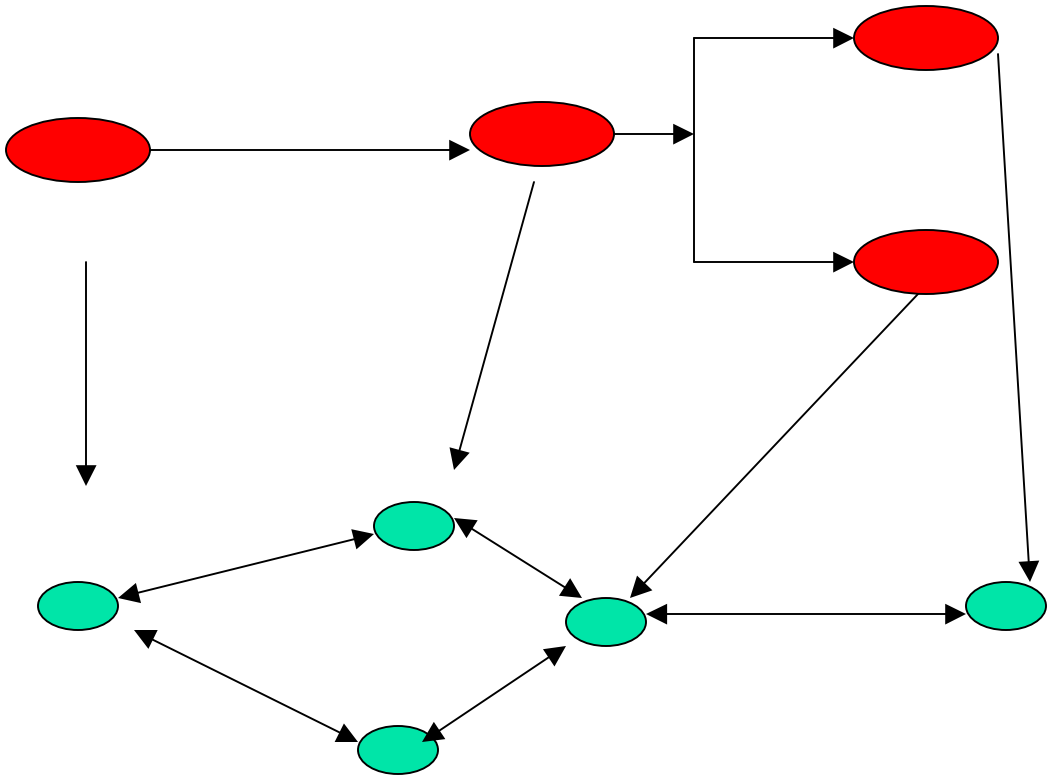
da mappare su



Nuova topologia

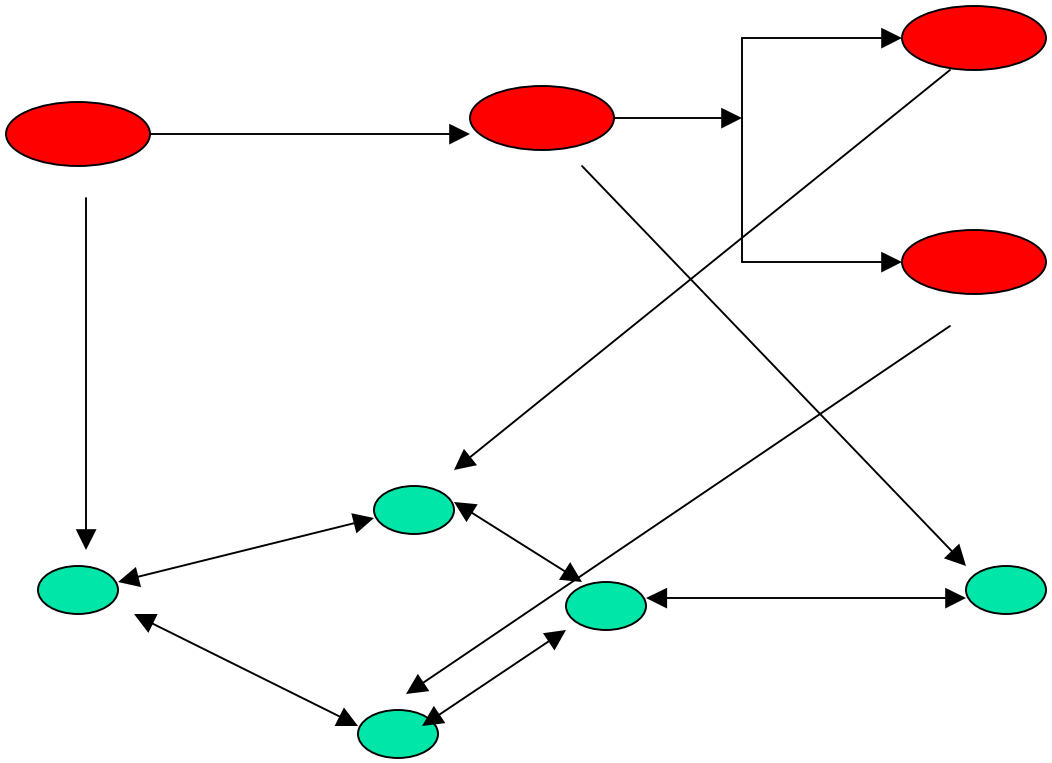
- Nuovo spazio dei nomi
  - Strutturato
  - Piatto
- Nuove regole di indirizzamento
- Nuove regole di instradamento
- Nuove bande di comunicazione
- Nuove regole di filtraggio

Overlay della topologia



Overlay efficiente

Overlay della topologia



Overlay inefficiente

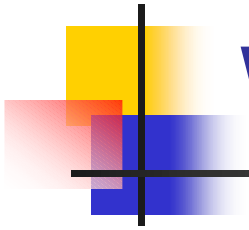
Overlay ed efficienza

- È comunque possibile
  - individuare i punti di inefficienza mediante i servizi dello stack originale
  - riorganizzare overlay
- Vantaggi
  - è spesso più semplice misurare l'utilizzo della nuova topologia che della rete fisica
  - si possono definire attributi di sicurezza della topologia (overlay= vpn)

Attaccante esterno

- Ha un insieme di informazioni non banali
  - Un overlay della rete fisica
  - Tutti i nodi dell'overlay utilizzano lo stesso software
  - Un attacco che funziona contro un nodo funziona contro tutti gli altri

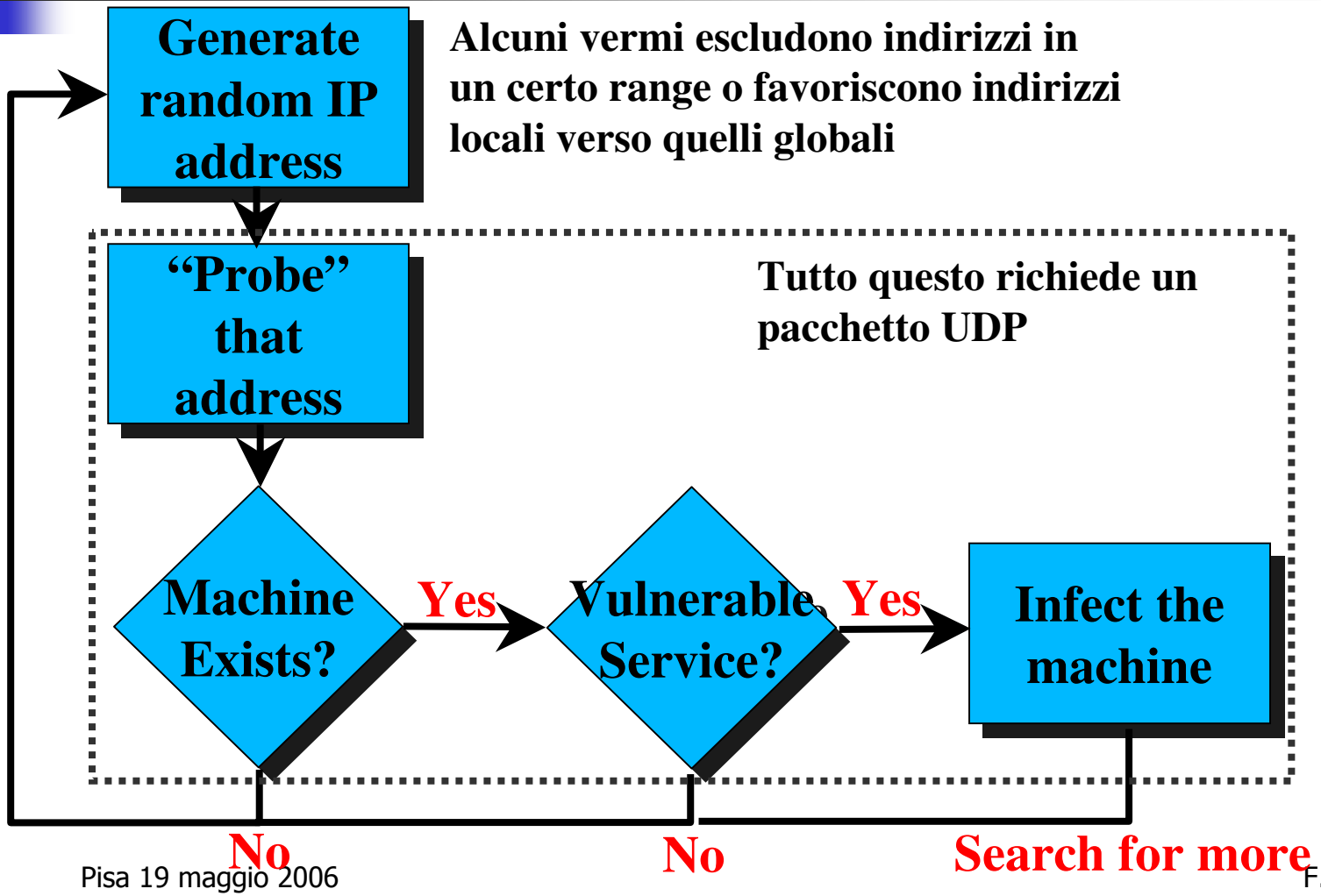
  $\Rightarrow$ si semplifica la scrittura di worm



Worm

- Comportamento tipico
  - Generazione di un indirizzo casuale
  - Verifica se esiste un nodo corrispondente
  - Verifica di vulnerabilità
  - Attacco e copia autonoma sul nuovo nodo
- Tutti i passi sono semplificati nel caso di p2p

Come funziona un verme in generale



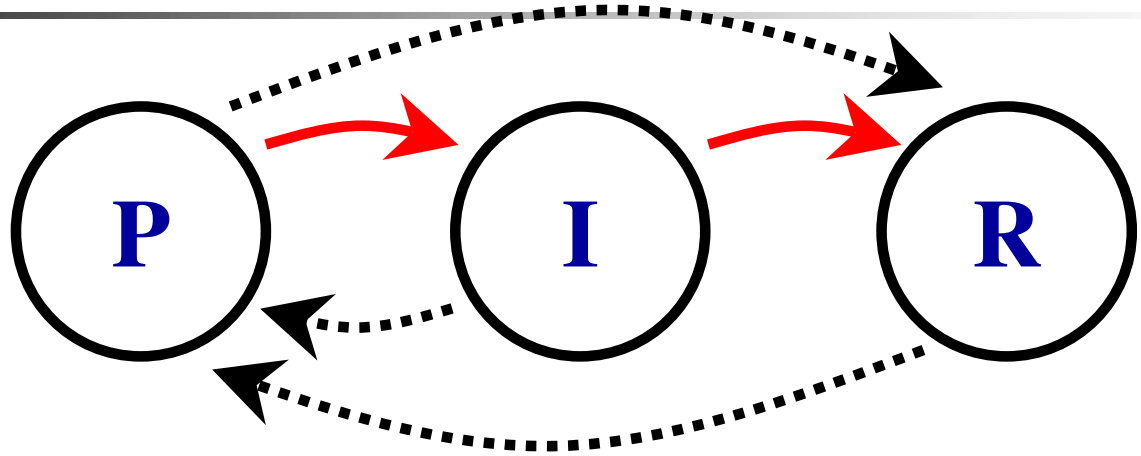
Funzione di generazione indirizzo

- Due insiemi possibili
  - Locale (denso)
  - Globale (poco denso)
- Se percentuale di locale è troppo bassa, può essere eliminato localmente prima di trovare un host da infettare
- Se percentuale di locale è troppo alta, una volta infettati tutti gli host si sprecano risorse perché ci sono intersezioni tra worm diversi
- Esistono rapporti tra percentuali che facilitano la diffusione del worm altri che la rallentano

Cosa cambia con p2p

- Uno spazio uniforme che poi viene mappato in nodi locali o globali
- Lo spazio è denso
- E' più semplice individuare nodi
- Aumenta l'efficacia

Modello per la propagazione



Stati del modello

- **Potenziale = Host vulnerabile all'infezione**
- **Infettato = Host contagiato dal verme**
- **Recovered = Host non vulnerabile**

Sequenza tipica (frecce rosse)

- **Host esegue una versione di XXX con vulnerabilità (potenziale).**
- **Verme sfrutta la vulnerabilità ed attacca l'host (infettato).**
- **Amministratore nota problemi, elimina verme ed installa patch (recovered).**

Un pò di matematica

Epidemiologia classica

- **[Kermack and McKendrick, 1927]**
- **Il cammino rosso del lucido precedente (P a I, I a R)**
- **Le equazioni**

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -\beta si$$

$$\frac{\mathrm{d}i}{\mathrm{d}t} = \beta si - \gamma i$$

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \gamma i$$

s = popolazione potenziale

i = popolazione infetta

r = popolazione recovered

***Beta* = tasso di infezione**

***Gamma* = tasso di recovery**

Gamma ignorato perchè trascurabile nei tempi considerati

Curva soluzione del sistema

Qualche dato…

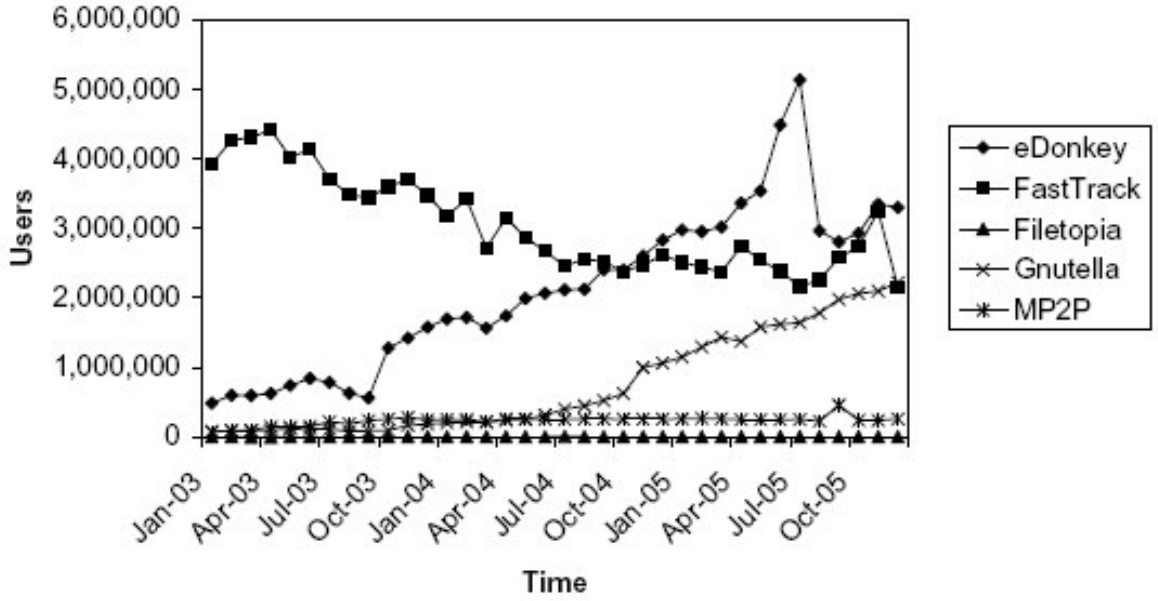


Figure 3: User populations for each of the file sharing services

Qualche dato…

Network	Worms	Variants
Multiple Networks	78	154
FastTrack only	62	122
eDonkey only	3	3
Gnutella only	1	3
FileTopia only	0	4
MP2P only	0	0

Numero di worm individuati
da gen 2003 a dic 2005

Attacchi contro il p2p

- Spazio degli attacchi è bidimensionale
- Attacchi contro i dati
  - File poisoning = inserimento di fake nella rete
- Attacchi contro il controllo
  - Sybil attack
  - Eclipse attack

Sybil attack

- Basato sulla generazione di più identità per l'overlay da parte di un nodo fisico
- Numero di identità possibili è limitato solo dalle risorse del nodo
- Difese
  - Dipendenza del nome da indirizzo IP
  - Punto di centralizzazione per l'inserimento nella rete
  - Protocolli basati su crittografia asimmetrica

Eclipse Attack

- Non è un attacco diverso ma semplicemente la prosecuzione di quello precedente
- Chi controlla un grande numero di nodi dell'overlay può
  - avvelenare i file
  - interrompere il routing dell'overlay

Sicurezza del fornitore

- Può essere migliorata mediante una generalizzazione il concetto di overlay che è alla base del p2p
- Passare dal concetto di overlay della topologia di interconnessione a quello di overlay completo tra reti
- Si sfrutta la tecnologia di virtualizzazione che è una delle più antiche e conosciute dell'informatica

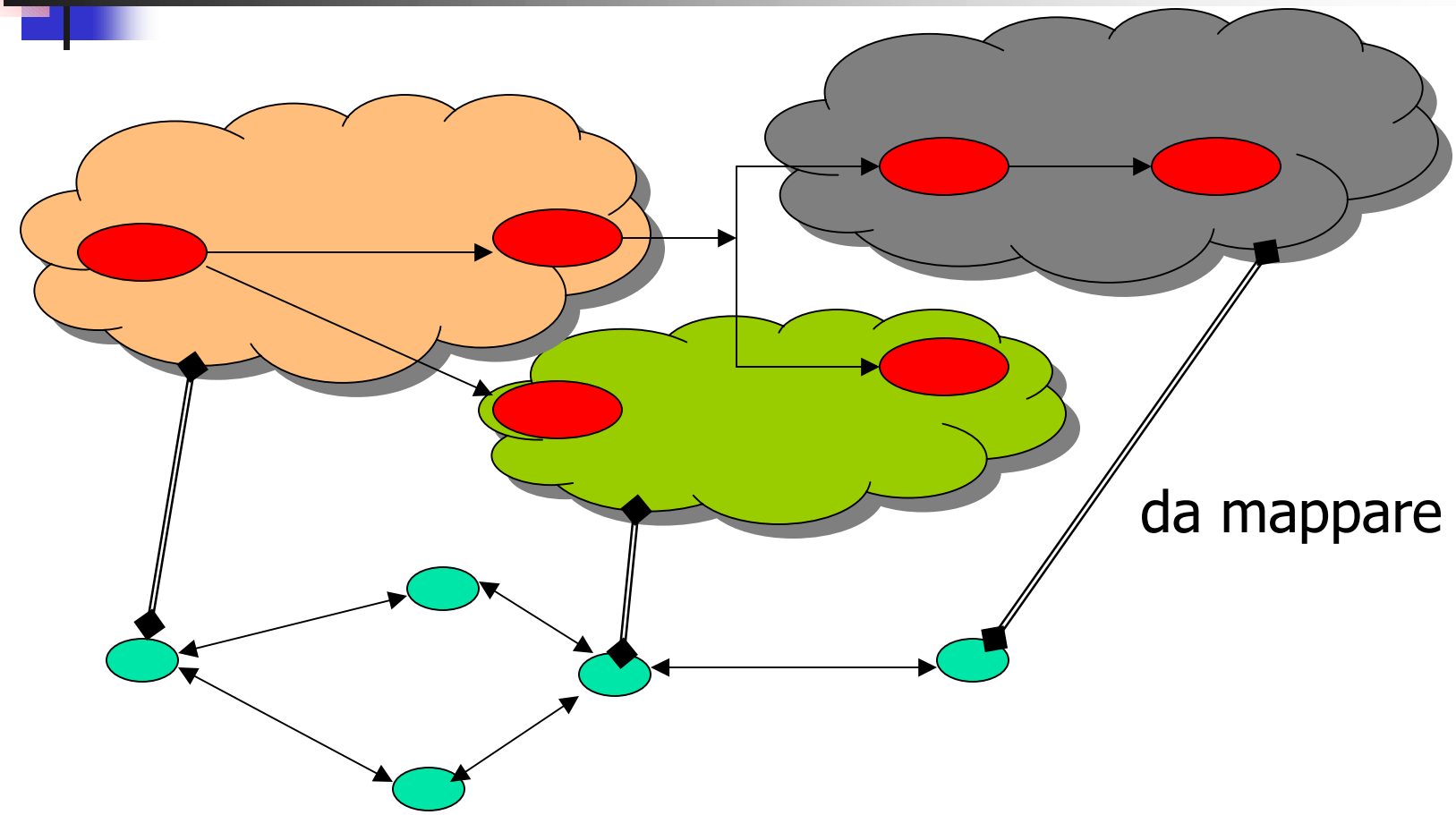
Overlay Completo

- Le tecnologie di virtualizzazione permettono definire reti virtuali complete date da
  - nodi virtuali
  - connessioni virtuali = overlay della topologia
- Il fornitore mette a disposizione non un solo nodo o le risorse di un nodo ma una rete completa che viene poi allocata sui componenti di un sistema reale

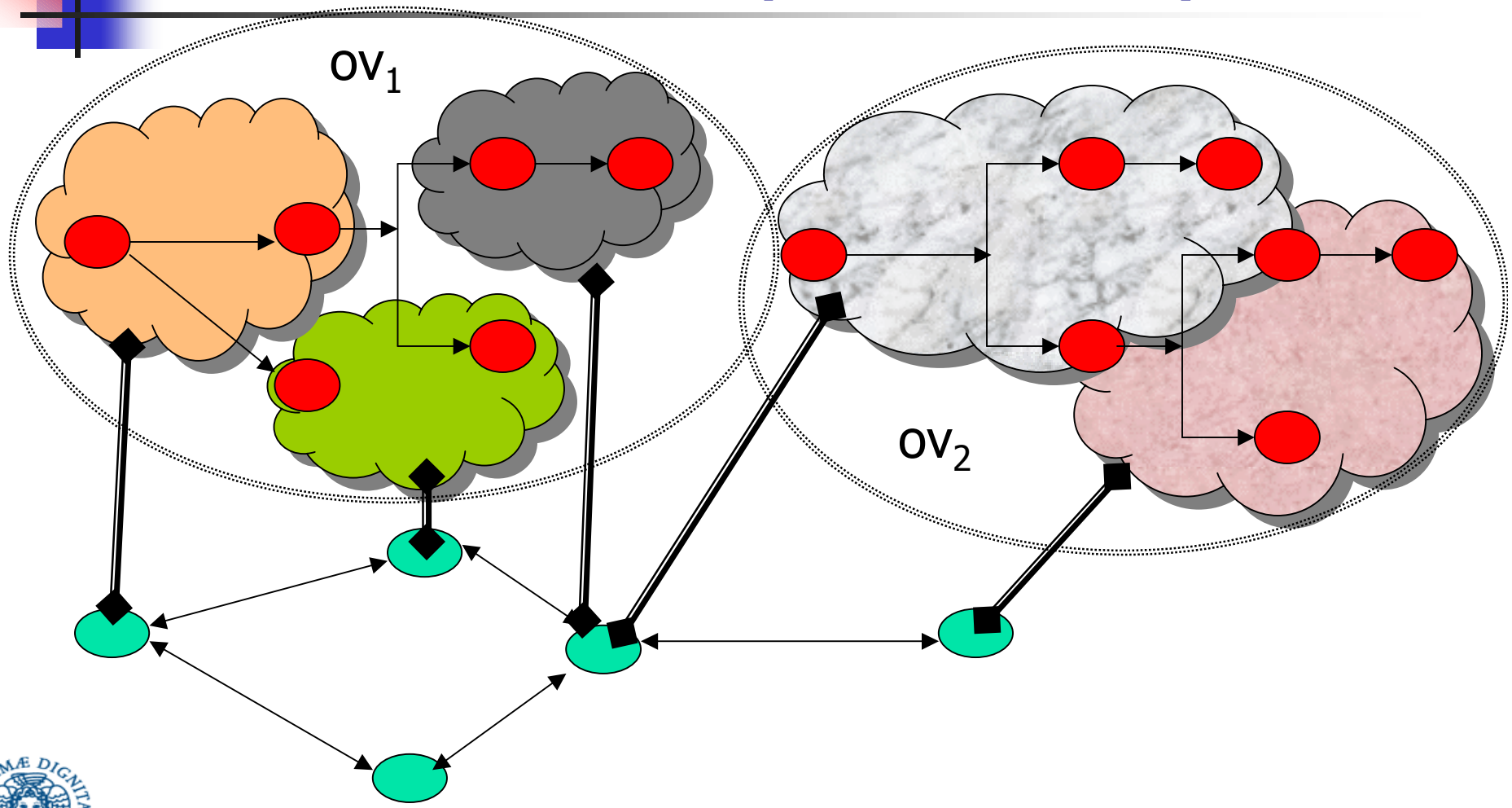
Allocazione di overlay

- In modo statico o dinamico sfruttando adeguatamente i meccanismi di migrazione delle macchine virtuali
- Permette di superare eventuali disomogeneità delle risorse
- Aumenta sicuramente l'efficienza nell'uso delle risorse

Allocazione di overlay



Allocazione di più overlay



Allocazione di overlay

- In generale l'allocazione di overlay mappa le risorse virtuali in quelle fisiche mediante uno schema molti a uno e mai uno a molti =

  un nodo virtuale su un unico nodo fisico

- Lo schema uno a molti (un nodo virtuale su più nodi fisici) è
  - interessante teoricamente
  - poco importante

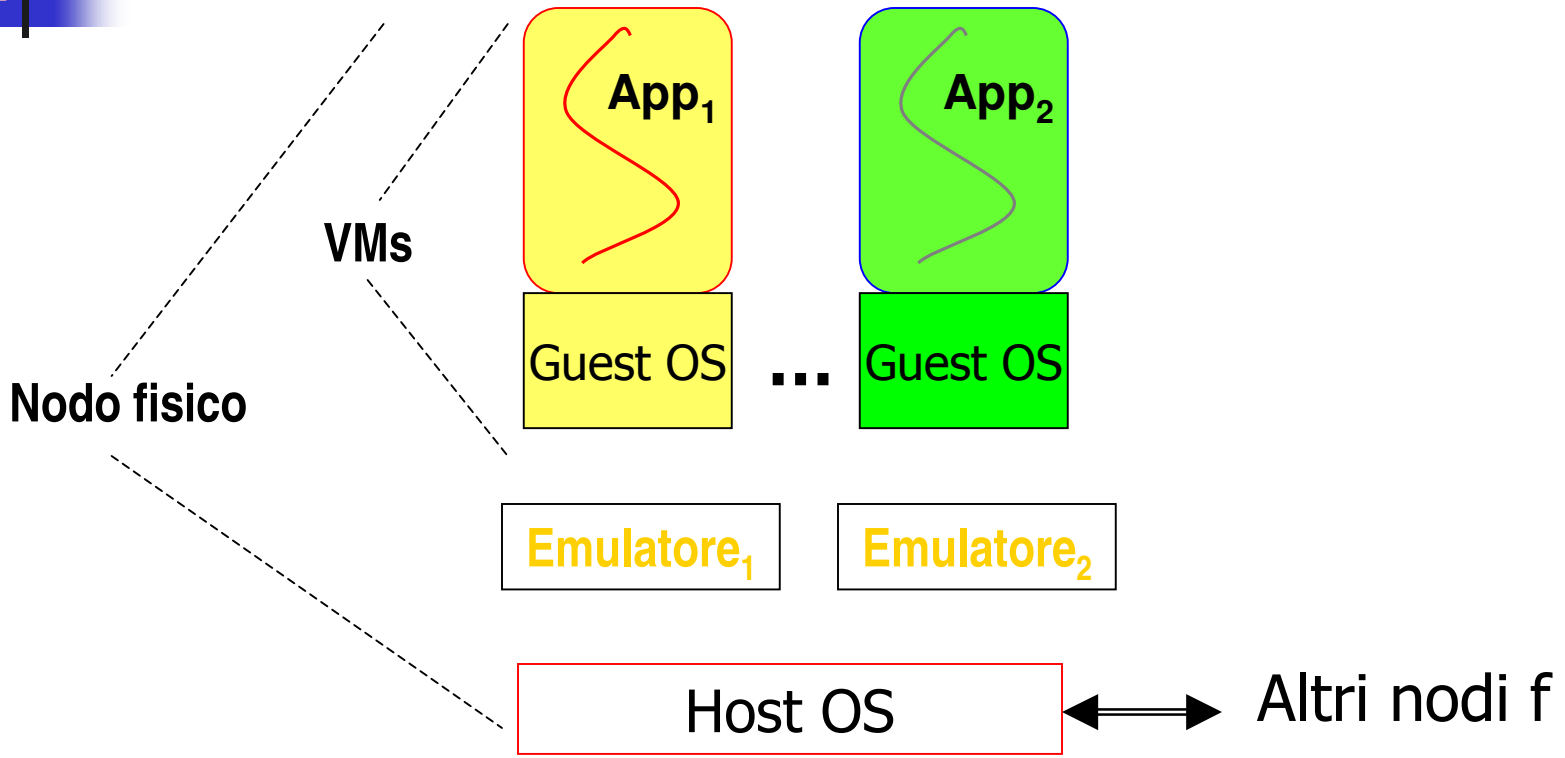
  con l'avvento di chip multi core il problema è come sfruttare la potenza del singolo nodo

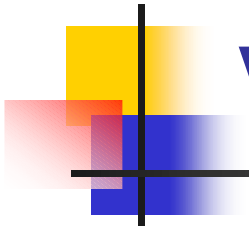
Overlay completo: un punto di vista

In un primo punto di vista

- nodo di elaborazione
  - è un sistema tradizionale
  - esegue mediante host os un insieme di applicazioni ognuna delle quali è una macchina virtuale
  - Ogni mv esegue guest os+applicazioni
- overlay della topologia avviene in una delle strategie standard di p2p

Un primo punto di vista





Vantaggi

- Portabilità del software
- Confinamento dei guasti e degli errori
- Gestione più semplice dei diritti in host os
- Scelta del guest os anche in base al livello di fiducia nell'applicazione da eseguire
- Garanzia di quote di risorse condivise sfruttando, se esistono, i meccanismi host os

Vantaggio unico

- È possibile definire condizioni globali sullo stato complessivo del so e della applicazione per scoprire
  - Malfunzionamenti
  - Attacchi
- Proprietà spesso sfruttata per definire honeypot o honeynet

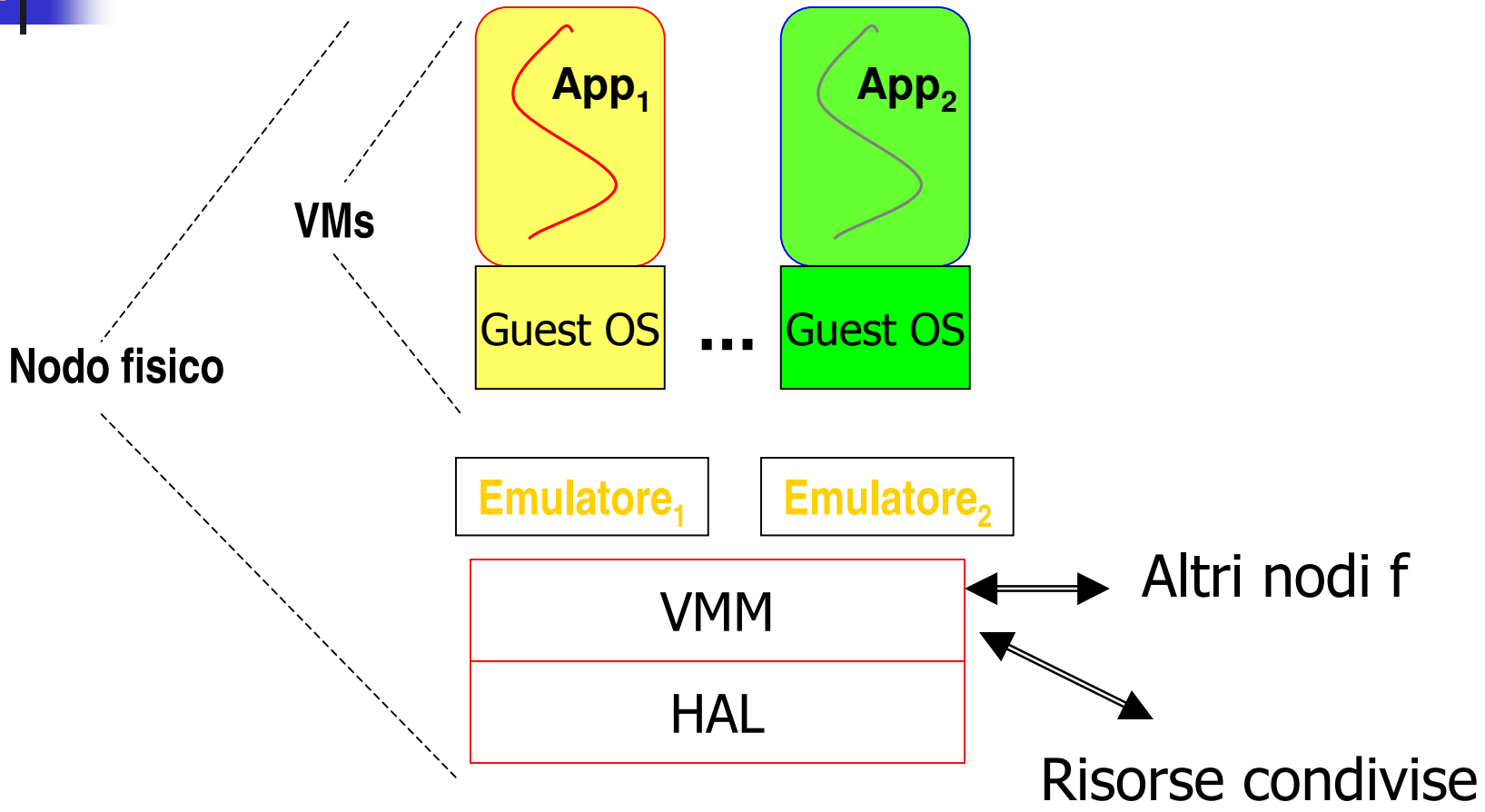
Svantaggio

- Difficile garantire una cooperazione controllata tra overlay diversi
- Il forte confinamento tra overlay diversi è fortemente diminuito se si ammette una cooperazione non controllata tra overlay

Una possibile soluzione - 1

- Il nodo non esegue un full os ma un virtual machine monitor
- Astrazione di componenti per i/o
- Il fornitore definisce una politica di sicurezza per il vmm che definisca
  - Operazioni permesse ai singoli overlay
  - Interazioni permesse tra overlay
  - Gestione delle risorse condivise

Una alternativa



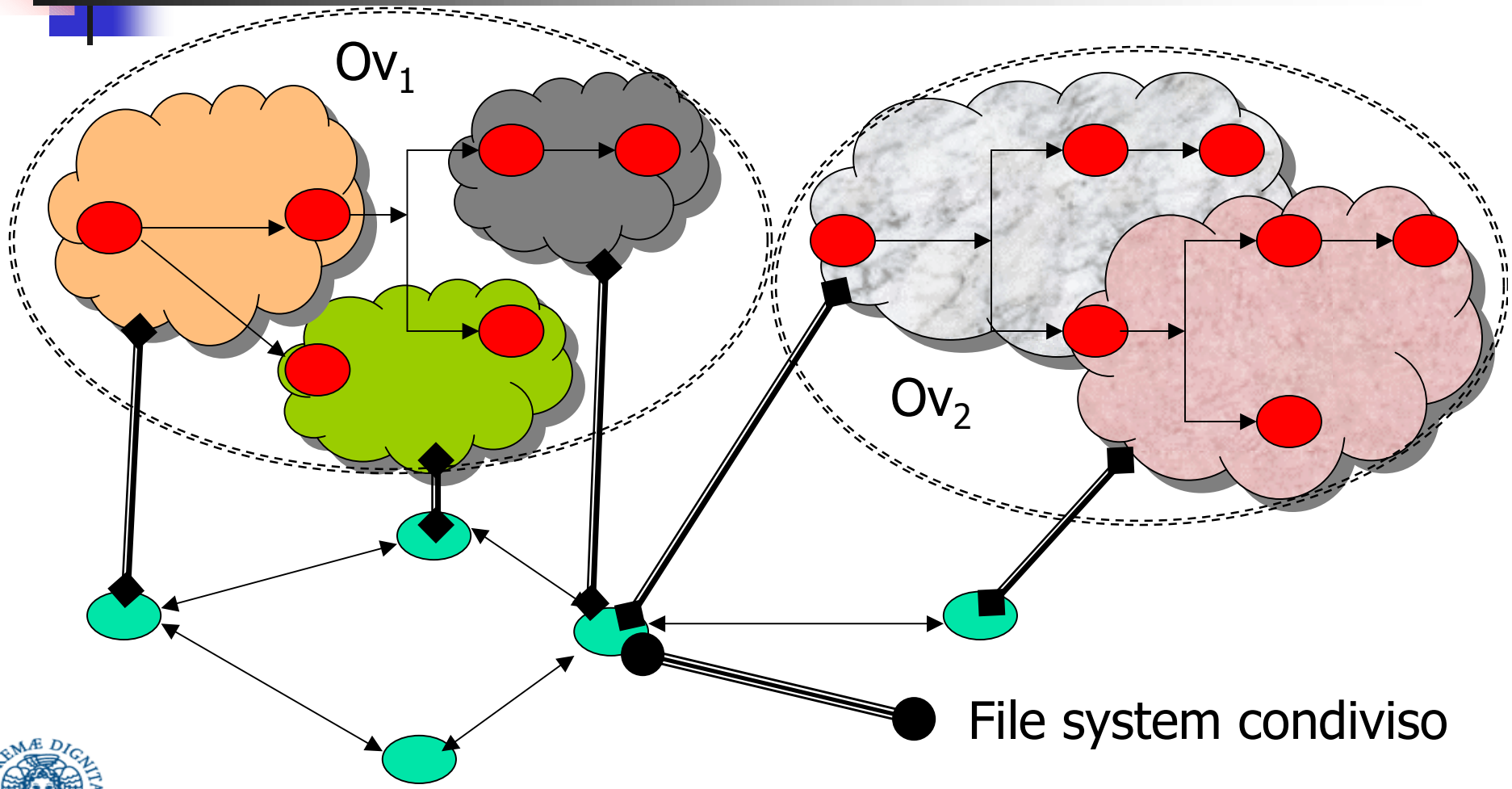
Una possibile alternativa - 2

- Il passaggio da full os a vmm permette
  - L'aumento delle prestazioni
  - La certificazione più semplice delle proprietà vmm
  - L'integrazione dei controlli di guest os e vmm
- Per semplificare la politica di sicurezza i controlli del vmm possono essere associati all'overlay e non alle applicazioni
- Si conservano i vantaggi di poter misurare semplicemente diversi valori interessanti
- Resta comunque elevata la complessita VMM per le risorse condivise (file etc)

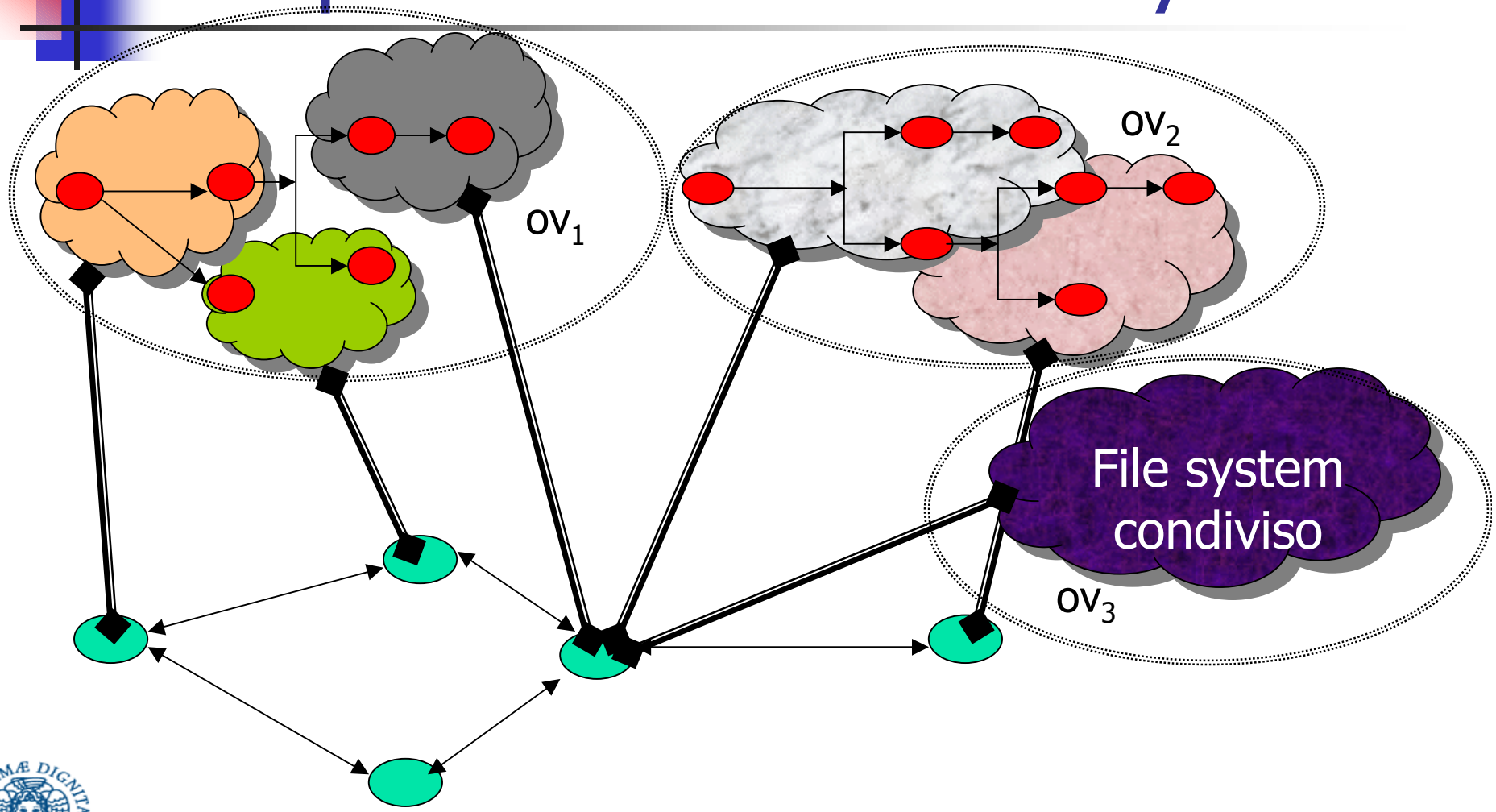
Una possibile alternativa - 3

- La cooperazione tra overlay sfrutta un ulteriore overlay che implementi la gestione delle risorse condivise
- Controlli del VMM
  - Legalità delle comunicazioni
    - tra overlay
    - Con overlay per risorse condivise
  - Confinamento sulle risorse condivise e non gestite dall'ulteriore overlay

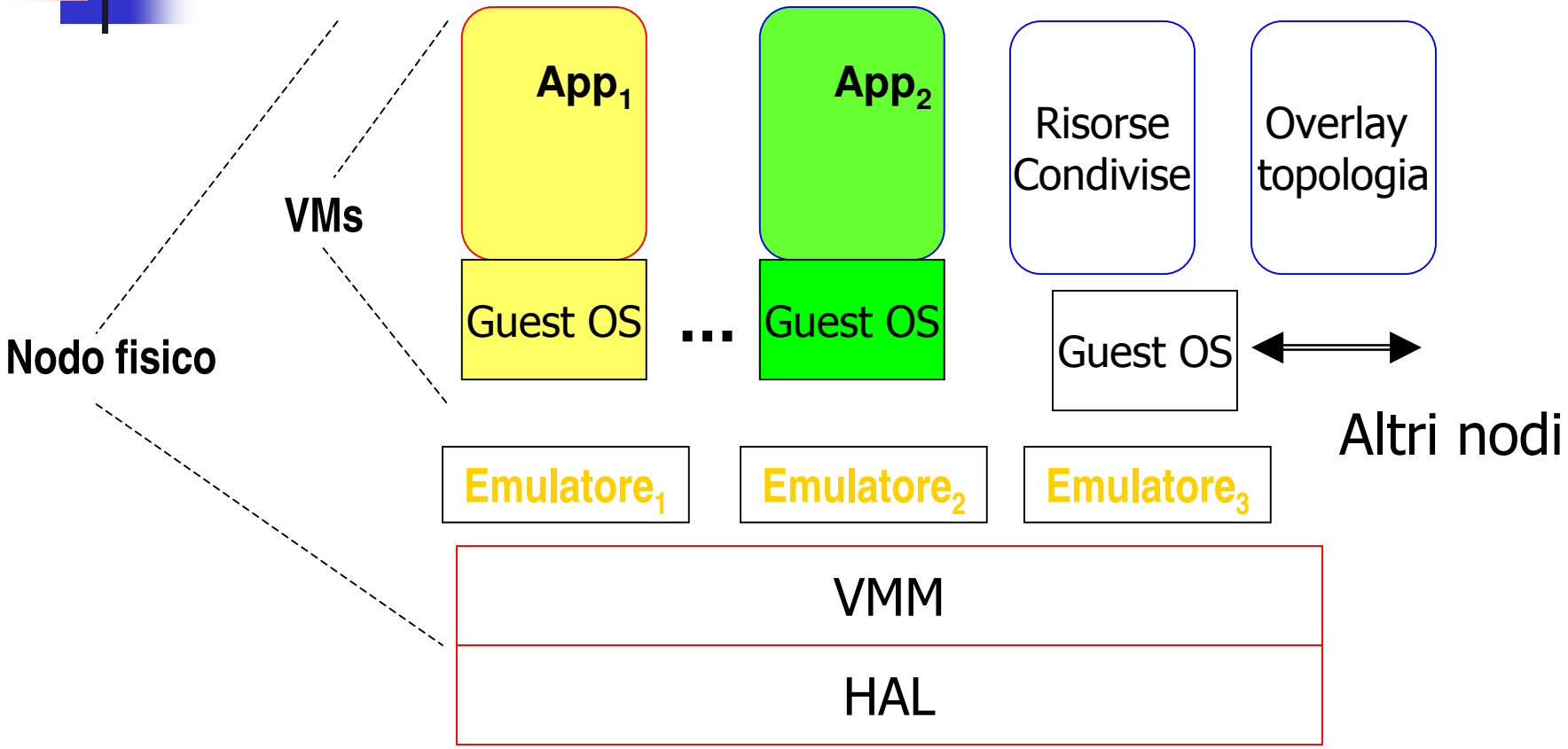
Cooperazione tra overlay



Cooperazione tra overlay



Una alternativa



Cooperazione tra overlay

- Trasmissione delle richieste al nuovo overlay può utilizzare le stesse strategie utilizzate per introdurre un proxy
- Deve comunque rimanere il controllo del VMM sulla interazione tra due overlay qualunque

Riassumendo - 1

- Possiamo avere in un overlay più nodi virtuali partizionando tra loro i compiti di
  - Esecuzione applicazione
  - Gestione overlay topologia
  - Gestione risorse condivise
  - Protezione overlay da worm etc.
- Gli overlay possono essere
  - Creati allo start up ed associati a classi di utenti (grid)
  - Creati dinamicamente in base al carico
  - Riallocati (migrare)

Riassumendo - 2

- Possiamo recuperare a livello di nodi virtuali quanto perso in sicurezza nelle interazioni all'interno del nodo
- Stiamo ripercorrendo la strada dei firewall ma all'interno dei nodi
- Forse l'unico modo di sperare in un aumento di sicurezza è la legge di Moore …

E' una visione realistica???

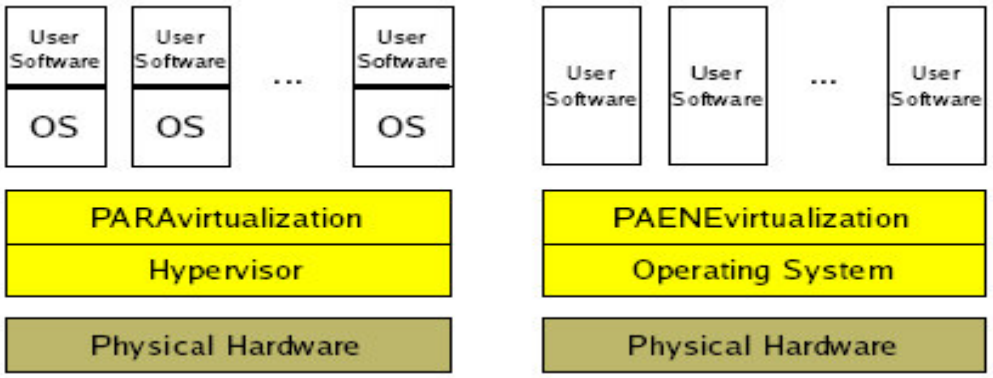
- Numerose tecnologie o prodotti commerciali basati su vm/vmm che hanno come obiettivo il confinamento
  - Vmware
  - Xen
  - User Mode Linux
  - Virtual server
  - Netop
  - …

E' una visione realistica???

- Progetti di ricerca su condivisione controllata
  - Planet Lab Proper
  - Virtuoso
  - Ostia
  - Sharp
- Evoluzione di architetture
  - Intel VT-I, VT-X (ex Vanderpool)
  - Hyperthreading + multicore (adatte per vm)
- Implementazioni alternative di VM
  - Interpretazione pura (con o senza caching)
  - Paravirtualization (Xen)
  - Paenevirtualization

Approcci

- Interpretazione + modifiche dinamiche per problemi x86= VMWare
- Paravirtualization = si espone esistenza di vmm al sistema operativo (Xen)
- Paenevirtualization = no untrusted os = ogni utente ha l'illusione di essere l'unico utente



Hyperthreading + multicore

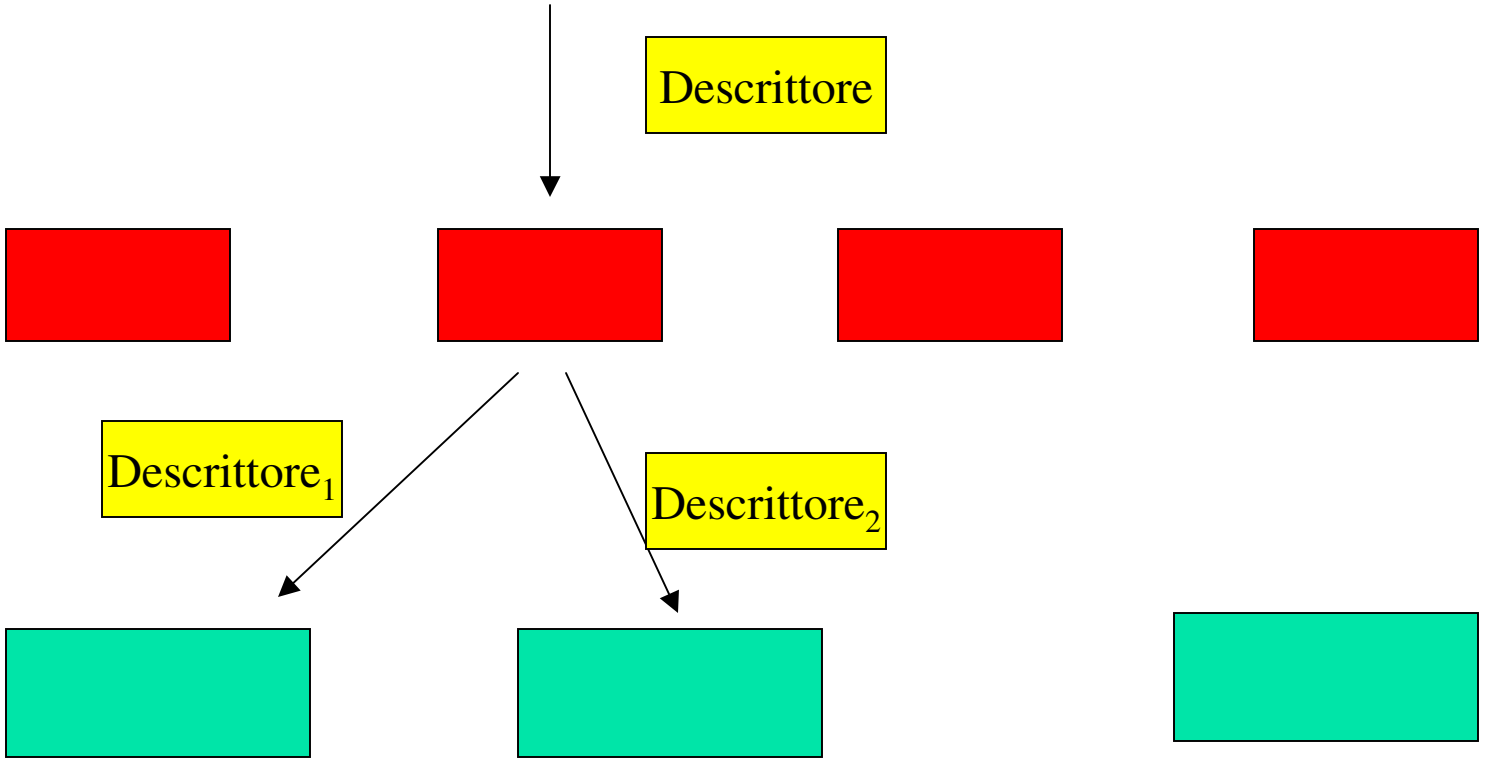
È interessante studiare cosa succede se alcuni thread (e quindi alcuni core) sono dedicati a

- personal firewall
- personal honeypot
- personal overlay topology

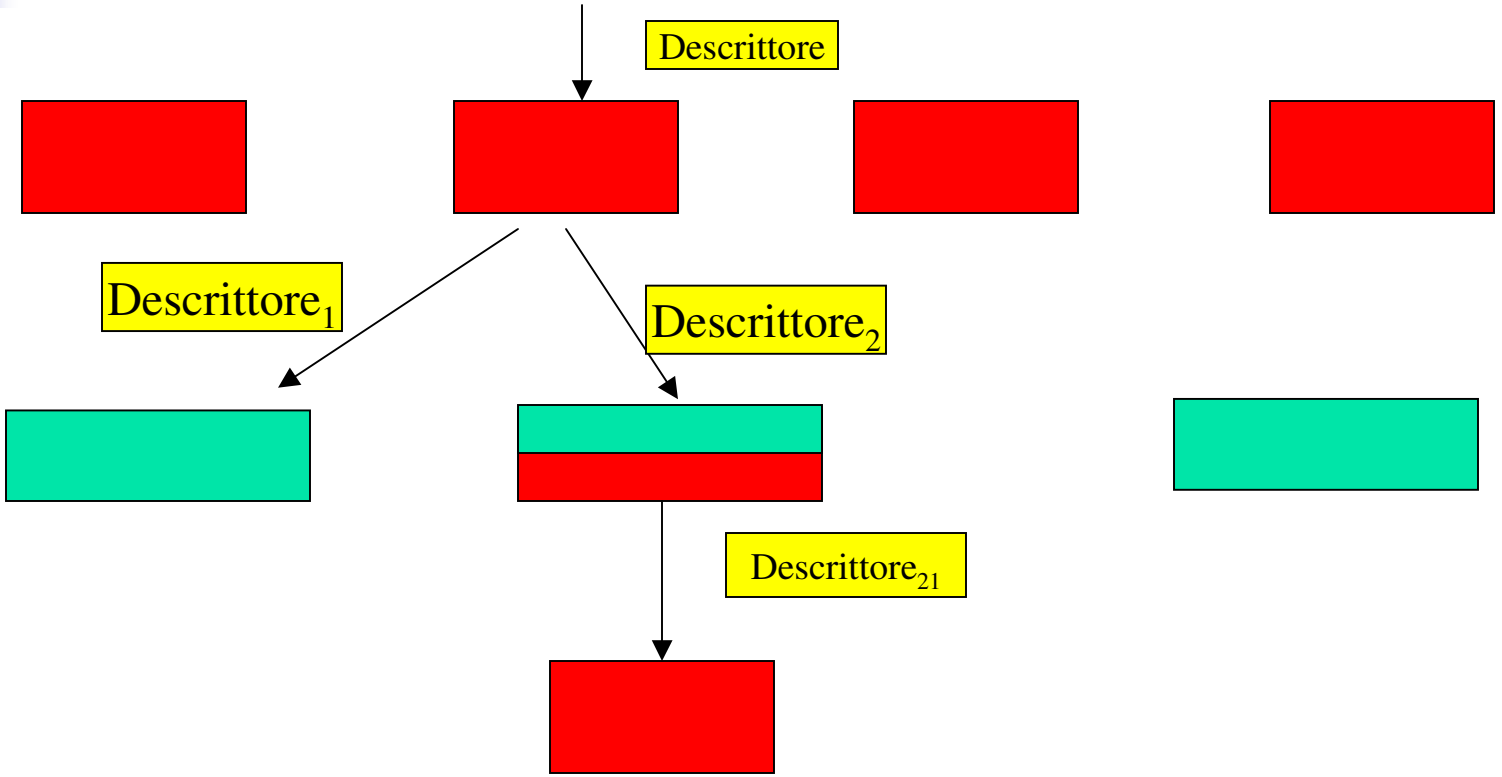
Punti deboli

- Gestione dei dispositivi
  - Auspicabile una visione "a processi" dei dispositivi dove interazione avviene mediante scambio di descrittori delle richieste di I/O
  - Più adatta ad un sistema a "nucleo minimo" (per qualcuno esokernel) in cui il nucleo implementi multiprogrammazione e livello HAL

Gestione dispositivi



Gestione dispositivi virt^2



Asimmetria del p2p

- U (utilizzatore) esegue
  - un programma P
  - con dati D

  sulle risorse fornite da F (fornitore)
- F può difendersi da U per evitare side effects inattesi monitorando D e P
- U ha pochi mezzi per difendersi da F

Asimmetria - Confidenzialità

- Quale protezione per chi mappa il proprio overlay = utilizzatore
  - Possibilità di verificare la consistenza dei risultati ottenuti
  - Estremamente difficile garantire la confidenzialità di
    - Input
    - Output
    - Parametri interni

Asimmetria – Confidenzialità

- Ci sono diversi casi in cui U vuole garantire la confidenzialità ed integrità dei programmi e dei dati che usa
- In generale i meccanismi per garantire confidenzialità ed integrità sono legate ad encryption
- Può funzionare nel caso di un servizio p2p di tipo supporto di memoria distribuito ma non in quello di calcolo

Asimmetria – Confidenzialità

- Alcuni risultati preliminare su
  - adozione di codifiche one-time pad per proteggere
    - Input
    - Output
  - Uso di frammentazione, encryption ed anonimity per dati memorizzati su supporti di più fornitori
- Più difficile proteggere i parametri dei modelli di simulazione utilizzati
- Modello di business commerciale???