

Virtual Lab: vantaggi e opportunità per la didattica dell'informatica

Aniello Castiglione, Giuseppe Cattaneo, Luigi Catuogno, Ezio Cerrelli²
Carlo Di Giampaolo, Fulvio Marino, Rodolfo Rotondo³

*Dipartimento di Informatica
Università degli Studi di Salerno
Via Ponte Don Melillo, 84084 - Fisciano (Salerno), Italy
{anicas,cattaneo,luicat,cardig,fulmar}@dia.unisa.it*

²*Dell Italia S.p.A.
Viale Piero e Alberto Pirelli, 6 - 20126 Milano, Italy
ezio_cerrelli@dell.com*

³*VMware Italy
Via Spadolini, 5 / Edificio A – 20141 Milano, Italy
rrotondo@vmware.com*

L'impiego delle infrastrutture di virtualizzazione per l'allestimento di laboratori didattici ha moltiplicato significativamente le risorse e gli strumenti a disposizione dei docenti per corredare i corsi con esercitazioni pratiche effettuate con l'ausilio del computer. In questo articolo, sarà presentata l'esperienza di progettazione ed esercizio di un multi-laboratorio virtuale maturata in un Dipartimento universitario di Informatica e saranno evidenziati i vantaggi offerti dalla nuova infrastruttura. L'articolo si chiude con una panoramica sulle principali tipologie di esercitazioni relative a corsi di informatica che si avvarranno della nuova infrastruttura.

1 Introduzione

Affermatasi negli anni sessanta in ambiente mainframe, la virtualizzazione, nell'informatica, è intesa come la tecnica che sovrappone a una risorsa di sistema, sia essa hardware o software (e.g., un computer, una sua periferica, un servizio del sistema), una sua versione virtuale in tutto simile a quella reale, che fornisce agli utenti la percezione di una risorsa a loro assegnata in modo esclusivo. La virtualizzazione della memoria, la gestione di molteplici device driver virtuali - che nei sistemi multi-tasking rendono disponibili insieme limitati di dispositivi reali ad un numero indefinito di processi concorrenti - non sono altro che consolidate applicazioni di questa tecnica. Recentemente, si è riaffermato, sia a livello consumer, sia a livello *data center*, l'impiego della virtualizzazione del sistema operativo. Pertanto, diverse istanze di sistemi operativi possono

essere rese disponibili su una singola piattaforma hardware, migliorando in maniera significativa l'utilizzo delle risorse hardware ed ottimizzando consumo energetico.

Nel campo educational sono da lungo tempo noti i vantaggi che l'impiego di soluzioni virtualizzate può apportare all'allestimento di piattaforme, laboratori ed esperienze didattiche. La virtualizzazione, infatti, permette di gestire con estrema facilità la realizzazione di reti di *client* e *server ad hoc*, configurati per le specifiche richieste di una particolare esercitazione pratica. Ciò permette di fare esperienze didattiche che non sarebbero possibili altrimenti, quali, ad esempio, sperimentare attività di gestione e manutenzione di applicazioni complesse o di sistemi operativi, così come offrire la possibilità di eseguire applicazioni che richiedono privilegi da superutente. Inoltre la virtualizzazione offre una soluzione economica ad attività di laboratorio che prevedano la gestione di ambienti di rete, permettendo agli studenti di gestire diverse macchine virtuali sul loro PC per simulare un ambiente di calcolo distribuito dove ogni macchina virtuale assolve compiti specifici (DNS, file server, etc.).

Il Dipartimento di Informatica dell'Università di Salerno si è fatto promotore dell'impiego di soluzioni didattiche basate sulla virtualizzazione, realizzando il progetto Virtual Lab, tra i più innovativi del suo genere, ed ormai operativo da tre anni.

In questo articolo sarà presentato il progetto, i suoi obiettivi e saranno evidenziati i risultati in esso ottenuti rispetto alla soluzione preesistente. Seguirà quindi una breve rassegna su alcuni "casi di studio": esempi di attività di laboratorio, abbinate all'insegnamento dell'informatica documentate in letteratura, il cui allestimento trarrà un notevole vantaggio dall'infrastruttura fornita dal progetto Virtual Lab.

2 La soluzione preesistente: problemi e limiti

Tra il 2002 ed il 2004, l'Area Didattica del corso di Laurea in Informatica dell'Ateneo Salernitano aveva già profuso un considerevole sforzo nella ristrutturazione dei laboratori didattici, varando il progetto *e-Class*, finanziato con fondi europei. I due laboratori pre-esistenti da 50 postazioni diskless multiboot, gestiti separatamente, venivano integrati in un'architettura più ambiziosa. La nuova infrastruttura hardware/software prevedeva la gestione centralizzata di oltre 250 postazioni suddivise in 5 aule-laboratorio allestite su due plessi diversi.

L'infrastruttura forniva postazioni multi-sistema Windows/Linux, con *single sign-on*, remote profiles e storage distribuito. Un sistema di management centralizzato permetteva la gestione di complesse policy per il controllo di accesso e per il remote deployment di applicazioni ed aggiornamenti del sistema operativo. In particolare, il nuovo assetto, prometteva un servizio estremamente più efficiente rispetto al passato, soprattutto in termini di tempi di fermo macchina dovuti a manutenzione o a interventi per malfunzionamenti

imprevisti. Infine, la centralizzazione dello storage e dell'accounting avrebbero dovuto rendere più flessibile la gestione degli utenti, dei loro dati e delle loro attività.

Tuttavia, negli anni sono apparsi diversi problemi. In particolare si è evidenziata una notevole difficoltà nella gestione del *deployment* delle applicazioni e/o del sistema operativo locale. Infatti, il deployment delle immagini del sistema operativo locale risultava lento e macchinoso principalmente a causa del sottodimensionamento dell'infrastruttura di rete (100Mb). Inoltre, il deployment di singole applicazioni poteva essere effettuato in maniera efficace soltanto per gli applicativi per cui la casa produttrice forniva il package di installazione via. Nel caso di applicazioni e *toolkit* sperimentali ed *open source* - quali erano la maggioranza degli strumenti richiesti dai docenti - si rendeva necessaria una procedura di installazione manuale. In alcuni casi, per effettuare le nuove installazioni, bisognava chiudere il laboratorio per diversi giorni, o rimandarle ai periodi di pausa delle attività didattiche.

Vi era poi un limite strutturale nell'impostazione del progetto. I laboratori *e-Class* non permettevano, se non al costo di considerevoli interventi *ad hoc*, di impostare tipologie di laboratori differenziati *on-demand*. In altre parole, il sistema operativo locale doveva contenere gli strumenti previsti da tutti gli insegnamenti attivati nel laboratorio. Pertanto, si rese necessario definire un regolamento che prescrivesse le caratteristiche e i tempi dei software da installare. A scadenze precise dell'anno accademico - all'inizio del semestre dei corsi di insegnamento - i docenti dovevano formulare la richiesta di installazione del software richiesto, in modo da consentire allo staff tecnico di preparare le "immagini" ed effettuarne il deployment. Va da se che era sostanzialmente impossibile predisporre *test-bed* per esercitazioni dalle possibili conseguenze distruttive, quali lo sviluppo di componenti del sistema operativo o la sperimentazione di tecniche di intrusione e delle relative contromisure. Tali esercitazioni erano svolte su "immagini" *live* installate su pen-drive USB distribuite agli studenti.

3 Il progetto Virtual Lab

Il progetto Virtual Lab è concepito ponendo al centro lo studente. Esso intende fornire all'utente, indipendentemente dalla postazione di accesso dell'infrastruttura, la possibilità di accedere a una libreria di macchine virtuali. Possono inoltre essere predisposte reti di macchine virtuali dedicate a specifici test-bed previsti per un determinato insegnamento. Le macchine virtuali possono essere sia "usa e getta", sia persistenti, sia personali, sia condivise con altri utenti.

In altre parole, lo studente può accedere a una qualsiasi delle macchine appartenenti al laboratorio virtuale che è stato costruito secondo le esigenze del corso che lo studente segue. L'accesso può aver luogo dalle postazioni situate nelle aule (non più dei PC ma dei semplici thin-client) o da computer remoti su

una sua macchina personale tramite un apposito software denominato *view client*.

Il progetto Virtual Lab poggia su una robusta infrastruttura di virtualizzazione server-side che gestisce su server virtuali tutti i servizi richiesti, fornisce lo storage necessario alla memorizzazione delle librerie di macchine virtuali e, gestisce tutte le politiche di controllo degli accessi.

3.1 Soluzioni Tecnologiche

La virtualizzazione desktop – cioè la sostituzione della postazione di lavoro reale con un'istanza virtuale - sposta la potenza elaborativa e/o lo spazio disco della postazione di lavoro dal desktop dell'utente ai sistemi centrali dei data-center. Negli ambienti desktop tradizionali, l'esigenza di spazio storage per l'utente finale, nonché la protezione dei dati utente, ha sempre di più portato a centralizzare i profili utente e i loro dati su sistemi storage centralizzati, comportando un incremento nell'uso della banda delle trasmissioni di rete (Figura 1).

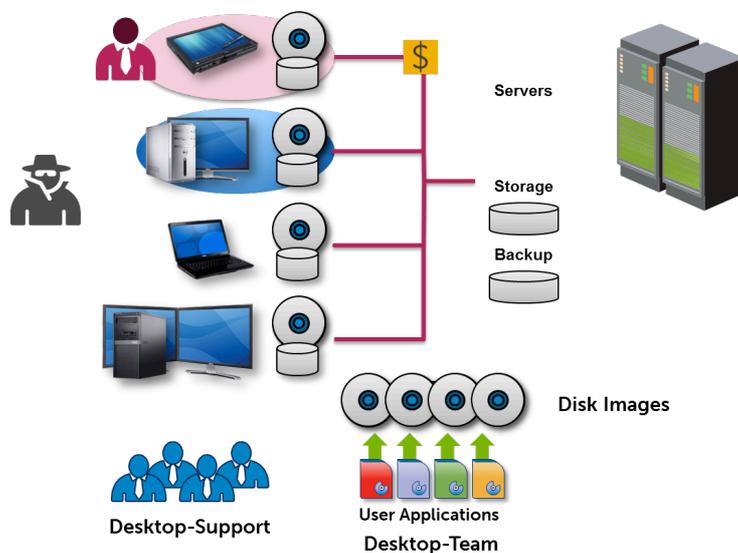


Figura 1 - Deployment di full disk image.

Il passaggio al Virtual Remote Desktop risolve il problema suddetto. La postazione utente infatti, viene generata e gestita nell'infrastruttura server e, attraverso sistemi di *brokering*, distribuita all'utente. Il sistema di visualizzazione dei Desktop Virtuali, il *view client*, invia le sole informazioni necessarie alla visualizzazione del desktop e riceve i soli input per l'operatività della postazione (movimenti del mouse, apertura/chiusura delle finestre, lancio di un applicazione). Il risultato finale consiste in un impegno minimo della banda della

trasmissione di rete ed indipendente dal device con il quale l'utente accede alla postazione Virtual Desktop (Figura 2).

La virtualizzazione dei desktop può essere realizzata mediante la combinazione di tecniche diverse che vanno dall'utilizzo di macchine virtuali su supporto USB, alla remotizzazione dei dischi delle postazioni e la realizzazione del Desktop Streaming, fino alle infrastrutture Virtual Remote Desktop.

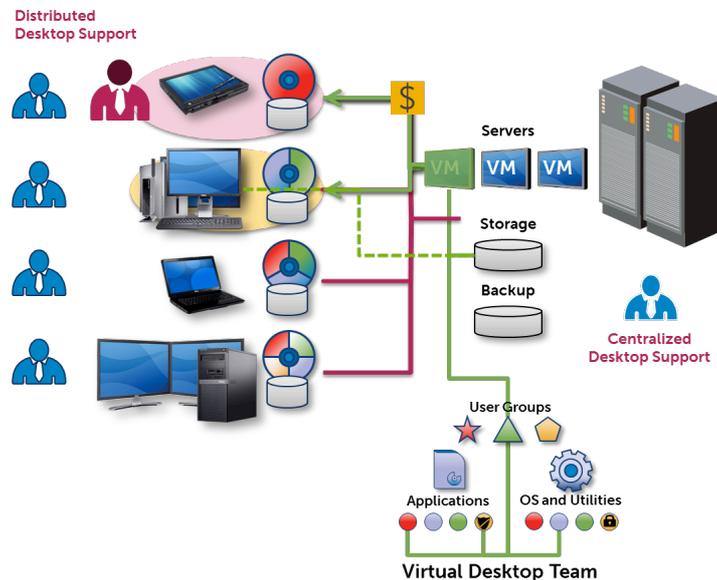


Figura 2 - Virtualizzazione desktop.

Vi sono due tipologie di sistemi virtuali forniti dall'infrastruttura Virtual Lab:

- **Server remoto condiviso.** E' implementato da un server multi-utente, e fornisce all'utenza un sistema standard, erogante servizi generici, e utilizzabile, ad esempio, per la gestione di documenti, accesso ad Internet, etc. L'infrastruttura Virtual Lab fornisce server remoti equipaggiati con sistema operativo Windows/Linux. I dati ed i profili utente sono gestiti in maniera centralizzata ed accessibili indifferentemente dall'uno o dall'altro server. L'accesso al server può aver luogo tramite i consueti protocolli quali RDP, SSH, etc. Tale configurazione si sovrappone funzionalmente alla vecchia infrastruttura e-Class, superandone però i limiti operazionali, come vedremo in seguito.
- **Sistema virtuale dedicato.** In questo caso, l'infrastruttura mette a disposizione dell'utente una macchina virtuale completamente dedicata e che può essere utilizzata a proprio piacimento anche con i privilegi di amministratore. Le macchine virtuali (VM) dedicate sono configurate in modo nativo con Windows secondo dei *template* stabiliti dai docenti dei

corsi ed utilizzano la tecnologia VDI di VMWARE per essere distribuite. Per gli ambienti Linux è fornita una chiave USB allo studente sulla quale è presente una VM unita all'ambiente VMWARE per il boot automatico da USB. Tali macchine possono essere utilizzate per effettuare esperimenti di natura invasiva che renderebbero la macchina non adatta alla gestione di dati persistenti. Un'attività invasiva ad esempio, potrebbe essere quella dell'installazione di un sistema Linux da parte degli utenti, o l'attivazione e sperimentazione di un servizio server (DNS, DHCP, etc.)

Esistono attualmente sul mercato diverse soluzioni che offrono la possibilità di virtualizzare il desktop. La scelta effettuata è stata quella di utilizzare un'infrastruttura per il consolidamento dei servizi lato server mediante il *framework* di virtualizzazione *VMware vSphere*. A ciò è stata aggiunta l'infrastruttura di Virtual Remote Desktop *VMWARE View* per la virtualizzazione del desktop, comunemente identificata con *VMware Virtual Desktop Infrastructure (VDI)*. Gli utenti possono accedere ai propri desktop virtuali in via remota sia da un tradizionale PC sia da un thin-client.

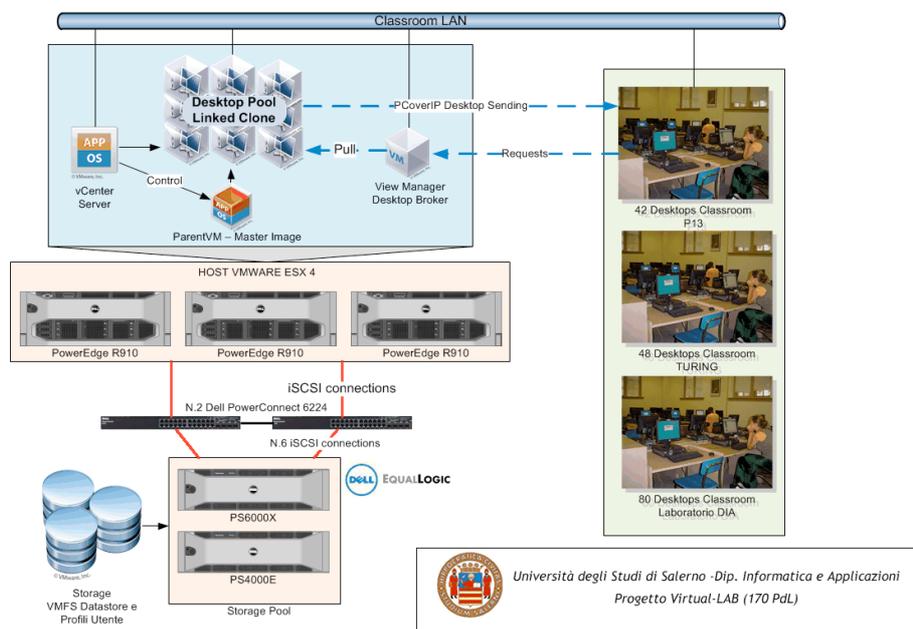


Figura 3 - L'infrastruttura Virtual Lab.

La scelta di VMware per l'implementazione dell'infrastruttura Virtual Lab è dovuta al fatto che VMware rappresenta il fornitore con maggiore esperienza e offre prodotti di una notevole stabilità e varietà. Con *VMware vSphere*, il sistema di VMware pensato per il cloud computing, è possibile amministrare un intero data-center utilizzando tecnologie di virtualizzazione, riconfigurando e riallocando in modo semplice e flessibile le applicazioni e i servizi. Lo strato di virtualizzazione è implementato dall'hypervisor VMware server ESX e gestito

centralmente da *VMware vCenter*. In Figura 3 è rappresentato lo schema dell'infrastruttura VDI realizzata.

4 Vantaggi

I vantaggi del nuovo approccio sono del tutto evidenti. L'infrastruttura Virtual Lab permette la realizzazione e la coesistenza di diverse tipologie di laboratorio. La possibilità di costruire pool di macchine virtuali dedicate alle esigenze di ciascun insegnamento, permette di rimuovere ogni vincolo sulle tipologie di software installabili. Come sarà evidenziato in seguito, il fatto che le onerose procedure di deployment non sono più necessarie, consente una gestione temporalmente più elastica di eventuali modifiche dell'assetto delle macchine virtuali esistenti.

Dal punto di vista delle prestazioni e del miglioramento dell'uso delle risorse, i vantaggi della virtualizzazione sono noti ed ampiamente documentati. Da un punto di vista gestionale, il progetto Virtual Lab rappresenta una considerevole innovazione rispetto alla soluzione preesistente. In questa sezione saranno esposti i vantaggi conseguiti in termini di servizi offerti e di efficienza del management, evidenziando in particolare i miglioramenti ottenuti in merito alle specifiche criticità emerse nell'ambito del precedente progetto *e-Class*.

Come già evidenziato in precedenza, il principale problema sorto durante l'esperienza *e-Class* era legato al deployment del sistema operativo e del software richiesto sulle diverse postazioni client. Tali postazioni ospitavano un sistema operativo locale che a sua volta utilizzava le risorse della server farm alla stregua di servizi di rete.

Nel progetto Virtual Lab, le postazioni client sono sostituite da thin-client o da PC equipaggiati da un sistema operativo minimale su cui è installato soltanto il software necessario per l'accesso all'infrastruttura. Il sistema locale deve essere considerato alla stregua di un firmware che non necessita di aggiornamenti frequenti.

Le postazioni di calcolo sono virtualizzate. La creazione e/o modifica di Desktop virtuali (o Laboratori virtuali) viene effettuata tramite macchine *template* ed il trasferimento fisico delle immagini sulle postazioni non è più necessario. Qualsiasi modifica della configurazione del sistema (ad esempio aggiornamenti o installazione di nuove applicazioni), viene effettuata sul *template*, ed avrà effetto sin dalla successiva richiesta di esecuzione. E' chiaro che questo diverso approccio riduce notevolmente il carico della banda di trasmissione di rete.

Le procedure di backup e di restore dei dati del sistema remoto condiviso sono gestite tramite una applicazione tradizionale. La procedura è estesa anche alle immagini delle macchine virtuali dedicate di cui deve essere garantita la persistenza, senza installare su di esse alcun *agent*. Analogamente, i controlli proattivi dei sistemi antivirus avranno luogo su dei semplici file contenenti le macchine virtuali, senza gravare sulle risorse di calcolo ad esse assegnate.

5 Attività di laboratorio: casi di studio

L'impiego di una infrastruttura di virtualizzazione fornisce nuove e più raffinate soluzioni per la progettazione e la realizzazione di esercitazioni pratiche legate agli insegnamenti in Informatica più tradizionalmente legati all'utilizzo di un laboratorio informatico. Nel seguito, si offre una panoramica sui principali corsi ed argomenti che possono trarre notevole vantaggio da un laboratorio virtuale.

Sistemi Operativi. L'opportunità di accedere ai sorgenti di sistemi operativi quale ad esempio Linux, e la disponibilità di documentazione estremamente dettagliata ed esaustiva (Bovet, Cesati, & Oram, 2002) (Rubini & Corbet, 2001) ha moltiplicato le possibilità di istituire corsi di sistemi operativi che approfondissero gli aspetti implementativi e fornissero agli studenti gli strumenti per la realizzazione di esperienze di altissimo valore professionale (Bovet & Cesati, 2001). Lo svolgimento di un esperimento nel kernel del sistema operativo può compromettere la stabilità del sistema e della sua integrità. In un laboratorio tradizionale, questo costringerebbe ad un periodico refresh delle installazioni delle postazioni di lavoro. In un laboratorio virtualizzato (Nieh & Vaill, 2006) gli studenti possono utilizzare macchine virtuali che sono rese on-demand, accessibili tramite una postazione indipendente dall'ambiente di testing. Effettuare esperimenti su un sistema virtuale non presenta differenze apprezzabili, ed offre notevoli vantaggi riducendo i tempi di recovery e, ad esempio, migliorando le possibilità di diagnostica, debugging e misura delle prestazioni (Dobrilovic & Stojanov, 2006).

Reti di calcolatori. Esercitazioni hand-on nei corsi attinenti al networking (e.g. programmazione su reti, reti di calcolatori) possono richiedere una attrezzatura complessa che include workstation, hub, switch ethernet e router. Inoltre, il setup delle esercitazioni risulta generalmente laborioso, e può richiedere un notevole dispendio di tempo, giocoforza sottratto alle ore di pratica. Il fattore economico non è trascurabile. L'hardware di rete ha un costo elevato, è soggetto a rapida obsolescenza ed inevitabile usura. Generalmente, gli atenei si dotano di tali infrastrutture tramite programmi disponibili in licenza "educational" (Armstrong & Murray, 2005).

Un laboratorio virtuale permette di emulare una rete fisica di computer e di componenti hardware completamente in software (Loddo & Saiu, 2008). Un pool di macchine virtuali può essere istanziato per ospitare postazioni utente ed appliance destinate ad emulare i componenti di rete (Yuan & Zhong, 2007). Una infrastruttura virtualizzata può consentire la creazione e la gestione di laboratori tramite API ed interfacce grafiche.

Sicurezza su reti. L'indispensabilità di predisporre esercitazioni pratiche, in un laboratorio dedicato ai corsi di sicurezza ICT, in cui sia possibile effettuare sperimentazioni estreme, e pregiudicanti l'integrità software dei sistemi utilizzati, è ampiamente riconosciuta (Catuogno & De Santis, 2008). Un approccio alquanto consolidato, per consentire agli studenti di comprendere le problematiche legate alla sicurezza su reti, e familiarizzare con strumenti quali firewall ed *Intrusion*

Detection Systems (IDS), consiste nell'allestire un laboratorio in cui squadre di studenti possano sperimentare tecniche di intrusione e di difesa nell'ambito di vere e proprie competizioni note in letteratura come *cyberwars* (Wagner & Wudi, 2004). L'allestimento di queste vere e proprie arene richiede la configurazione di un ambiente di base comprendente sistemi operativi ed applicazioni, insieme ad un determinato layout di rete ed eventuali componenti hardware aggiuntivi (router, access-point) che aumentino il grado di realismo dell'esercitazione. Il laboratorio deve essere isolato dall'ambiente esterno, in modo che le attività in esso svolte vi rimangano confinate ed inoltre, al termine dell'esercitazione, i computer e le attrezzature utilizzate, devono essere riportate alla loro configurazione iniziale affinché sia possibile effettuare una nuova *cyberwar*. Tali laboratori possono essere molto onerosi, in termini di costi d'impianto e gestione.

L'utilizzo di un laboratorio virtuale, fornisce un ambiente flessibile ed economico per realizzare questo tipo di laboratori (Wang, Hembroff, & Yedica, 2010). Il docente può predisporre gli opportuni *template* di macchine virtuali, configurare una rete virtuale che le connetta secondo la topologia richiesta. La rete così costruita può essere corredata di appliance che simulino il traffico di rete ed azioni malevoli (Aycock, Crawford, & deGraaf, 2008).

Forensic investigation. Le indagini forensi comprendono attività e procedure piuttosto invasive. Potrebbe rendersi necessario, per esempio, smontare il computer oggetto dell'indagine, per asportarne il disco rigido. In un laboratorio didattico destinato ad ospitare più corsi, questo tipo di operazioni non è sempre praticabile. L'uso di macchine virtuali permette di realizzare una parte significativa delle esercitazioni di un laboratorio di informatica forense (Yao, Chu, & Liu, 2010) senza pregiudicare l'integrità delle attrezzature e la loro fruibilità per altri corsi.

6 Conclusioni

I laboratori didattici allestiti grazie a tecnologie basate sulla virtualizzazione offrono numerosi vantaggi tecnici ed economici, migliorando l'utilizzo delle risorse tecniche ed umane impiegate nel loro esercizio. In questo articolo, è stato presentato il progetto di un laboratorio virtuale presso un dipartimento universitario di informatica. Sono stati discussi i vantaggi e le opportunità che esso offre alla didattica e, in particolare, a quegli insegnamenti tradizionalmente più legati all'utilizzo dei laboratori. Occorre sottolineare che, sebbene i casi di studio presentati siano inerentemente legati alle discipline ICT, l'infrastruttura offre una flessibilità d'uso tale da consentirne l'impiego concorrente anche per insegnamenti di altri corsi (e.g., laboratori multimediali di lingue). In particolare, l'infrastruttura Virtual Lab è correntemente utilizzata, negli orari in cui non ospita le attività didattiche, per la conduzione di progetti di ricerca sperimentali con l'ausilio di tecnologie di cloud computing, come, ad esempio, la soluzione di progetti di "molecular dynamics" in collaborazione con un gruppo di ricerca della Facoltà di Farmacia. Va rilevato che tali forme di condivisione delle risorse di

calcolo forniscono un'ottima occasione per la nascita di progetti di ricerca interdisciplinare.

7 Bibliografia

Armstrong, H., & Murray, I. (2005). Incorporating vendor-based training into security courses. *Proceedings of the IEEE Workshop on Information Assurance and Security 2005* (p. 172--175). West Point NY (USA): IEEE.

Aycock, J., Crawford, H., & deGraaf, R. (2008). Spamulator: the Internet on a laptop. *Proceedings of the 13th annual conference on Innovation and technology in computer science education* (p. 142--147). New York NY (USA): IEEE.

Bovet, D. P., & Cesati, M. (2001). A Real Bottom-Up Operating Systems Course. *Operating Systems Review*, 35 (1), 48-60.

Bovet, D., Cesati, M., & Oram, A. (2002). *Understanding the Linux kernel*. Sebastopol CA (USA): O'Reilly Media.

Catugno, L., & De Santis, A. (2008). An internet role-game for the laboratory of network security course. *Proceedings of the 13th annual conference on Innovation and technology in computer science education* (p. 240--244). ACM.

Dobrilovic, D., & Stojanov, Z. (2006). Using virtualization software in operating systems course. *Conference on Information Technology: Research and Education, 2006. ITRE'06* (p. 222--226). New York NY (USA): IEEE.

Loddo, J., & Saiu, L. (2008). Marionnet: A virtual network laboratory and simulation tool. *First International Conference on Simulation Tools and Techniques for Communications, Networks and Systems* (p. 1--8). New York NY (USA): ACM.

Nieh, J., & Vaill, C. (2006). Experiences teaching operating systems using virtual platforms and linux. *ACM SIGOPS Operating Systems Review*, 40 (2), 100-104.

Rubini, A., & Corbet, J. (2001). *Linux Device Drivers*. Sebastopol CA (USA): O'Reilly Media.

Wagner, P., & Wudi, J. (2004). Designing and implementing a cyberwar laboratory exercise for a computer security course. *ACM SIGCSE Bulletin*, 36 (1), 402--406.

Wang, X., Hembroff, G., & Yedica, R. (2010). Using VMware VCenter lab manager in undergraduate education for system administration and network security. *Proceedings of the 2010 ACM conference on Information technology education* (p. 43--52). ACM.

Yao, W., Chu, C., & Liu, B. a. (2010). DESIGNING A VIRTUAL LAB FOR COMPUTER FORENSICS. *Proceedings of the 14th Colloquium for Information Systems Security Education* (p. 1--8). Baltimore MD (USA): CISSE.

Yuan, D., & Zhong, J. (2007). Building a CISCO-feel linux router to enable students learn and configure dynamic routing protocols. *Proceedings of the 8th ACM SIGITE conference on Information technology education* (p. 271--272). New York NY (USA): ACM.