



CRESCO

Centro computazionale di RicErca sui Sistemi
COMplessi

Dipartimento di Ingegneria dell'Informazione e Matematica
Applicata

Università di Salerno

Responsabile Unità Operativa

Prof. Ciro D'Apice

LA1-SP I.2 - Sviluppo ed integrazione dell'architettura GRID e di grafica 3 D

Attività I.2.4 Analisi e Realizzazione di Tool innovativi a supporto delle funzionalità
GRID

2d – Integrazione delle componenti software
nell'infrastruttura ENEA-GRID, con il supporto dei
ricercatori ENEA

Integrazione di componenti software nella Windows Farm ENEA GRID

V. 1.0

Indice

1. Varie	3
1.1. Bibliografia e Riferimenti.....	3
1.2. Acronimi.....	3
1.3. Glossario	3
2. Contesto, obiettivi e scopo	5
3. Pubblicazione di codici di simulazione	8
4. Conclusione	13

Indice figure

<i>Figura 1 – Pagina di accesso alla WIN Farm ENEA.....</i>	8
<i>Figura 2 - Accesso alle applicazioni della Server Farm</i>	9
<i>Figura 3 - Operazione di creazione del Profilo.....</i>	10
<i>Figura 4 - Accesso al codice Simulatore Rete</i>	12

1. Varie

1.1. Bibliografia e Riferimenti

- [1] Introduzione alla virtualizzazione
- [2] Virtualizzazione e streaming con SoftGRID
- [3] Virtualizzazione e streaming nella GRID ENEA
- [4] SoftGrid Platform Admin Guide
- [5] <http://www.microsoft.com/systemcenter/softgrid/default.mspix>
- [6] <http://technet.microsoft.com/en-us/softgrid/default.aspx>
- [7] <http://www.microsoft.com/systemcenter/softgrid/default.mspix>

1.2. Acronimi

CRESCO	Centro Computazionale di Ric erca sui Sistemi CO mplexi
SOAP	Simple O bject A ccess P rotocol

1.3. Glossario

Active Directory (AD)

Sistema integrato e distribuito di directory service adottato dai sistemi operativi Microsoft a partire da Windows 2000 Server. In Active Directory sono integrate tutte le applicazioni per la gestione dei servizi di rete e di dominio.

Andrew File System (AFS)

È un file system distribuito sviluppato dalla Carnegie Mellon University.

Downtime

Termine che indica il tempo in cui l'attività di un'azienda è sospesa a causa di guasti.

End User

Termine che indica il consumatore finale di un prodotto o di un servizio software.

Internet Information Services (IIS)

Insieme di servizi Internet-based per sistemi operativi Microsoft. È soprattutto un potente Web Server.

Load Sharing Facility (LSF)

Si tratta di un software di bilanciamento e controllo dei carichi di calcolo in una struttura distribuita, prodotto dalla *Platform Computing*.

Network Attached Storage (NAS)

È un dispositivo collegato ad una rete di computer la cui funzione è quella di condividere tra gli utenti della rete una Area di storage.

Terminal Services (TS)

In un modello centralizzato, un Server TS è in grado di rendere disponibili applicazioni agli utenti, garantendo quindi una più semplice gestione delle applicazioni stesse ed una migliorata sicurezza.

Real Time Streaming Protocol (RSTP)

Il protocollo RSTP ottimizza il flusso di dati. A differenza del protocollo http, RTSP controlla il flusso dati con funzionalità concettualmente simili a quelle di un registratore.

Real Time Transport Protocol (RTP)

Protocollo che permette la distribuzione di servizi che necessitano di trasferimento in tempo reale, come l'interattività audio e video. Ad esempio il "Time stamping" oppure il "monitoring".

Real Time Control Protocol (RTCP)

Protocollo che raccoglie statistiche sulla qualità del servizio del protocollo RTP e trasporta le informazioni riguardo ai partecipanti ad una sessione.

Return Of Investment (ROI)

Indice di redditività del capitale investito.

Storage Area Network (SAN)

Rete ad alta velocità di dispositivi di memorizzazione di massa condivisi.

Total Cost Ownership (TCO)

È un approccio utilizzato per calcolare tutti i costi del ciclo di vita di un'apparecchiatura informatica, per l'acquisto, l'installazione, la gestione, la manutenzione e il suo smantellamento.

Web Services

È un sistema software progettato per supportare l'interoperabilità tra diversi elaboratori su di una medesima rete; caratteristica fondamentale di un Web Service è quella di offrire un'interfaccia software, utilizzando la quale altri sistemi possono interagire con il Web Service stesso attivando le operazioni descritte nell'interfaccia tramite messaggi SOAP.

2. Contesto, obiettivi e scopo

Il Progetto CRESCO: Centro computazionale di **Ric**erca sui Sistemi **CO**mplexi ha come obiettivo la realizzazione, presso il Centro Ricerche ENEA di Portici (NA), di un importante Polo di calcolo multidisciplinare per lo studio dei sistemi complessi di natura biologica e tecnologica, collegato con *facilities* di elaborazione, archiviazione e visualizzazione 3D di dati ubicate presso i Centri ENEA di Brindisi e Trisaia (MT).

Sul piano delle applicazioni e dei contenuti scientifici, il Progetto si focalizza sulle seguenti linee d'attività:

- Linea di Attività 1 (LA1): Realizzazione del Polo di Calcolo e sviluppo di nuove funzionalità di GRID Computing

l'implementazione di soluzioni innovative in tema di architetture di sistemi di calcolo e di GRID computing per le attività di R&S di punta dell'ENEA che richiedano l'utilizzo di risorse computazionali estremamente importanti. Sul versante infrastrutturale il Progetto prevede la realizzazione di una piattaforma di calcolo di assoluta eccellenza in grado di erogare una potenza dell'ordine del Teraflop, corredata di una nuova classe di funzionalità GRID innovative;

- Linea di Attività 2 (LA2): Sviluppo di tecnologie e modelli computazionali per la descrizione di sistemi complessi di origine biologica e di materiali innovativi

lo studio di oggetti biologici dal punto di vista "sistemico" (in silico cell) e lo studio di sistemi naturali (comunità animali e sociali) secondo il paradigma dei sistemi complessi;

- Linea di Attività 3 (LA3): Sviluppo di modelli di simulazione ed analisi delle Reti tecnologiche complesse e delle loro interdipendenze

lo studio di sistemi tecnologici complessi e delle loro mutue interazioni, e la realizzazione di opportuni strumenti per il modelling, la simulazione e il controllo.

Nell'ambito delle attività previste dalla Linea di Attività 1 il sottoprogetto identificato come SP I.2, *Sviluppo ed integrazione dell'architettura GRID e di grafica 3D*, prevede, fra l'altro, l'attività di Analisi e Realizzazione di Tool innovativi a supporto delle funzionalità GRID e in tale contesto, nel quale il Dipartimento di Ingegneria dell'Informazione e Matematica Applicata dell'Università degli Studi di Salerno è coinvolto, è previsto, primariamente, uno studio delle tecnologie di virtualizzazione e di *streaming* delle applicazioni avendo, come prodotto di riferimento, la piattaforma **SoftGrid** realizzato dalla software house **Softtricity**, poi rilevata da Microsoft. SoftGrid è una piattaforma centralizzata di management e di distribuzione di qualunque programma Windows, con la quale è possibile gestire e distribuire applicazioni istantaneamente e su richiesta (*on-demand*). L'introduzione di una tale tecnologia consente, tra l'altro, la possibilità di impiegare applicazioni e soluzioni che girano su piattaforma Microsoft senza la necessità di averle installate sulla propria macchina. Grazie ad un opportuno *client* infatti, anziché installare la versione completa, alla prima richiesta di una o più applicazioni da parte dell'utente, si scarica dal server Softgrid solamente, e quindi anche molto rapidamente, la parte del pacchetto software necessario. Le applicazioni non vengono lanciate sul server come avviene ad esempio con i *Terminal Services*, ma localmente senza comprometterne la rapidità e l'efficienza. La dimensione delle applicazioni influisce solo in parte sulla velocità di esecuzione, sia perché, come detto, il server non distribuisce l'intero software, sia perché la parte di programma ricevuto resta disponibile al client per futuri utilizzi in quanto viene memorizzata nella *cache* locale, permettendone quindi anche il funzionamento "off line" nel caso si verificano problemi alla rete. La più grande novità della piattaforma consiste nel fatto che le applicazioni virtualizzate vengono eseguite sul client in un'area apposita chiamata *System Guard*. Essa consiste in una "gabbia virtuale" che isola il software dagli altri presenti nel sistema e dal sistema operativo stesso. I risultati di tali studi sono riportati in [1] e [2].

Ovviamente la virtualizzazione e lo *streaming* dovranno essere integrate nella GRID ENEA che prevede, quale uno dei suoi componenti cardine, **Citrix Presentation Server**, prodotto mirato alla pubblicazione delle applicazioni. Le ultime versioni del

prodotto, peraltro, sono state integrate con le caratteristiche di *streaming* e virtualizzazione esaminate ponendo l'attenzione su SoftGRID senza peraltro pregiudicare la validità generale degli studi condotti. Quindi la Windows Farm della GRID ENEA potrà impiegare tecnologia Citrix anche per virtualizzare ed effettuare lo *streaming* delle applicazioni pubblicate. I risultati conseguiti sono riportati in [3].

Nel presente documento ci soffermeremo invece sulle attività intraprese per la pubblicazione di alcuni componenti software, sviluppati dal DIIMA dell'Università degli Studi di Salerno nell'ambito delle collaborazioni previste nel progetto CRESCO, che costituiscono un esempio di virtualizzazione, streaming e pubblicazione di applicazioni nell'ambito della GRID ENEA.

3. Pubblicazione di codici di simulazione

Nell'ambito delle attività previste dalla collaborazione fra ENEA ed il Dipartimento di Ingegneria dell'Informazione e Matematica Applicata dell'Università degli Studi di Salerno e relative alla Linea Attività III *Sviluppo di modelli di simulazione ed analisi delle Reti tecnologiche complesse e delle loro interdipendenze*, sottoprogetto S.P.III.3 *Modelli e strumenti di supporto alla ottimizzazione e riconfigurazione delle Reti*, il Dipartimento ha realizzato una serie di codici di calcolo per la modellazione di reti con approccio fluidodinamico. Tali codici di calcolo, realizzati in ambiente Windows, sono accomunati dalla caratteristica di interagire con l'utente lavorando su archivi sia per la fase di input che per quella di output. Si tratta dunque di eseguibili che possono essere lanciati senza bisogno di procedura di installazione e che richiedono all'utente semplicemente la specifica dei percorsi dei file di input e di quelli di output.

Lo scopo che ci si propone è quello di integrare tali codici nella Windows Farm di ENEA-GRID accessibile all'URL <http://winfarm.portici.enea.it>

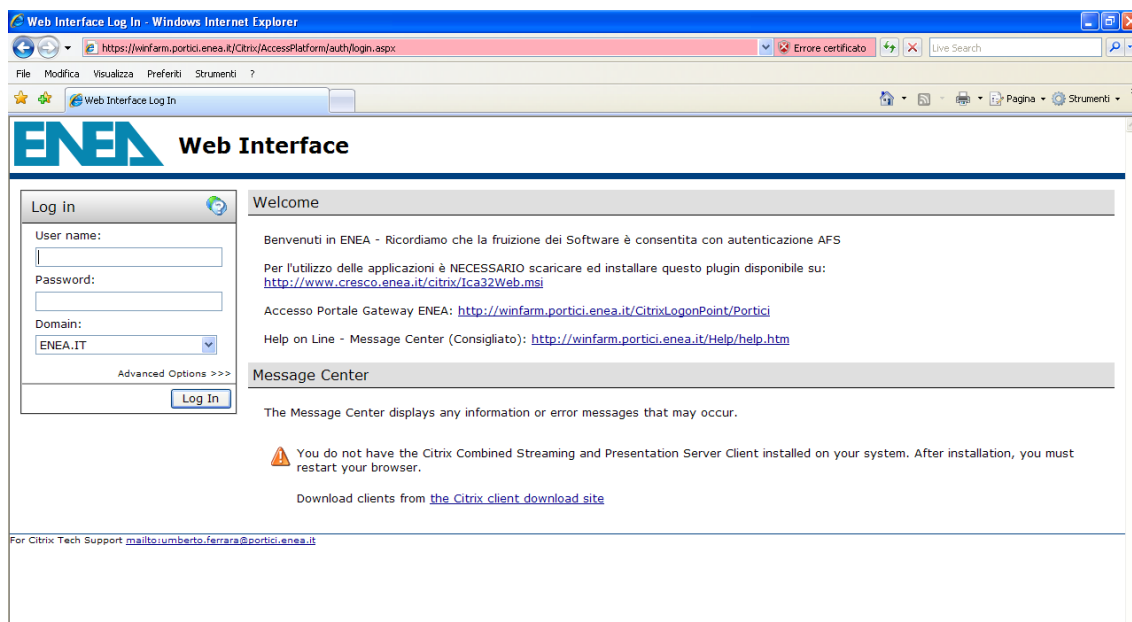


Figura 1 – Pagina di accesso alla WIN Farm ENEA

Il sistema che costituisce il punto di accesso alla Windows Farm è costituito da un *Citrix Web Interface Server* che consente agli utenti di eseguire le proprie applicazioni ICA sui sistemi che agiscono da Citrix Presentation Servers.

La seguente figura mostra la tipica interazione fra la farm di server che eseguono le applicazioni, l'interfaccia Web ed un dispositivo client.

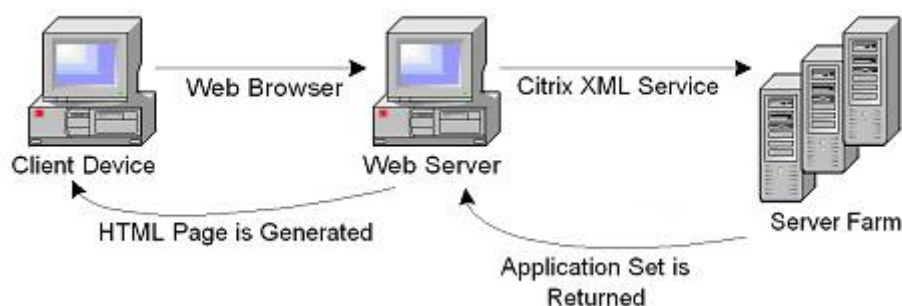


Figura 2 - Accesso alle applicazioni della Server Farm

In generale quindi i passi che sono eseguiti in tale attività sono:

- il dispositivo client impiega un web browser per accedere alla pagina di autenticazione e fornire le proprie credenziali di accesso;
- il server web legge le informazioni utente e le invia ad un opportuno servizio in esecuzione sui server della farm, *Citrix XML Service*, che stabilisce quali siano le applicazioni alle quali l'utente può accedere e restituisce tale set di applicazioni al server Web che le presenta al client;
- l'utente seleziona l'applicazione che desidera eseguire selezionando l'opportuno link sulla pagina HTML che gli è stata presentata ed il server web ricontatta il *Citrix XML Service* per identificare il server nella farm che è meno carico ed ottenerne un ticket corrispondente alle credenziali dell'utente. La coppia (indirizzo server, ticket) viene passata all'interfaccia Web che poi le fornisce al client.

Sui sistemi che costituiscono la Server Farm è in esecuzione il *Citrix Presentation Server* e quindi tali sistemi sono quelli che rendono praticamente disponibili le applicazioni attraverso il meccanismo della pubblicazione.

In linea generale le applicazioni potrebbero essere delle applicazioni fisicamente installate su ciascun server oppure delle applicazioni virtualizzate.

Il chiaro vantaggio di avere delle applicazioni virtualizzate è quello di eliminare la necessità di dover installare l'applicazione su ciascun server ed evitare potenziali conflitti fra le diverse applicazioni installate.

Tenendo ben presente quanto detto nel [3], per poter effettuare lo *streaming* di un'applicazione è necessario averne il **profilo** che viene generato attraverso il prodotto *Citrix Streaming Profiler* che, nello specifico, è installato su una macchina virtuale disponibile su uno dei server della Farm. L'opportunità di installare il *Profiler* su una macchina dedicata scaturisce dalla necessità di eseguire l'operazione su una macchina pulita e riservata allo scopo.

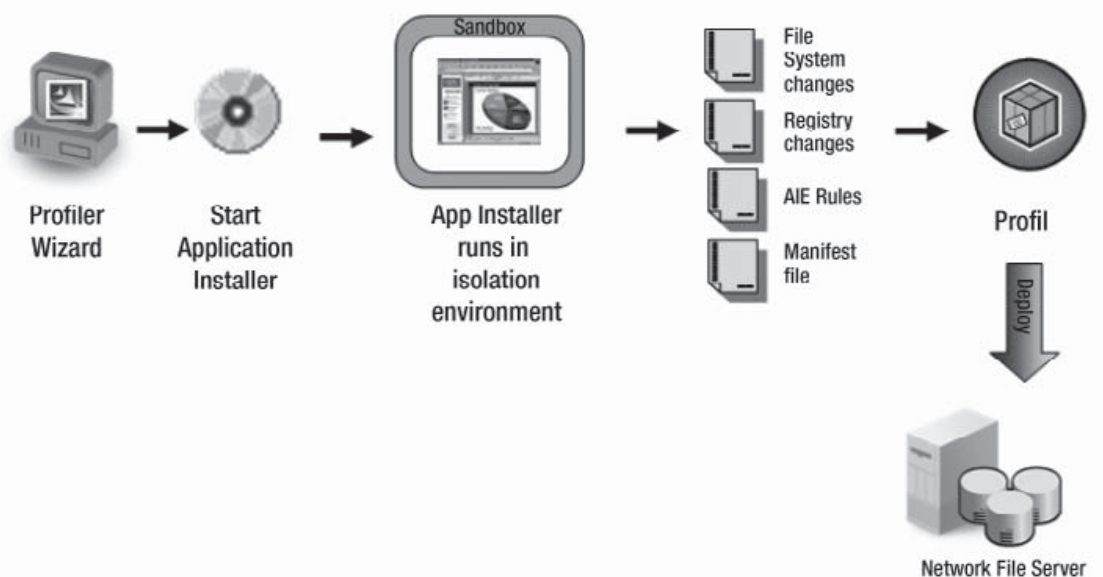


Figura 3 - Operazione di creazione del Profilo

Come si evince dalla figura la costruzione del profilo prevede l'esecuzione dell'installazione del software che si desidera virtualizzare ed il risultato finale sarà costituito da una serie di file costituenti l'applicazione virtualizzata.

Per poter procedere quindi allo *streaming* dell'applicazione sono stati creati, per ciascuno dei quattro codici sviluppati, dei pacchetti di installazione che consentiranno quindi la corretta produzione del profilo.

Il profilo può essere salvato ovunque, ma ovviamente, è opportuno che lo stesso sia salvato su un file system accessibile attraverso la rete. Nel caso specifico della Windows Farm ENEA, il profilo sarà salvato sul file system AFS che è uno dei componenti chiave della GRID ENEA.

Dopo aver creato il profilo è necessario pubblicarlo e tale operazione è eseguita mediante la **Citrix Access Management Console**. In tale fase si preciserà che l'applicazione pubblicata sarà impiegata da uno dei server della Farm, e, si noti, che si potrebbe anche specificare che essa debba essere eseguita sul client finale. In quest'ultimo caso lo *streaming* dell'applicazione, precedentemente virtualizzata, avrebbe come destinazione il sistema client, ossia il sistema che l'utente sta impiegando per accedere all'infrastruttura.

Nello scenario che stiamo esaminando, l'utente che si collega alla Windows Farm ENEA, dopo essersi autenticato potrà attivare il codice di calcolo precedentemente *profilato*. Nella figura di esempio è visualizzato l'accesso ad un codice denominato **Simulatore Rete** e si nota la finestra dei comandi mediante la quale il codice chiede il percorso dei file che costituiscono l'input per la sua corretta esecuzione.

Si consideri che in tale scenario l'utente della GRID ENEA, che si collega alla Windows Farm, ha a disposizione il proprio file system AFS e quindi potrà tranquillamente accedere ai file che ha precedentemente creato ed atti alla descrizione del modello su cui il codice deve lavorare. In maniera analoga i risultati del codice saranno archiviati in file che l'utente potrà salvare su AFS.

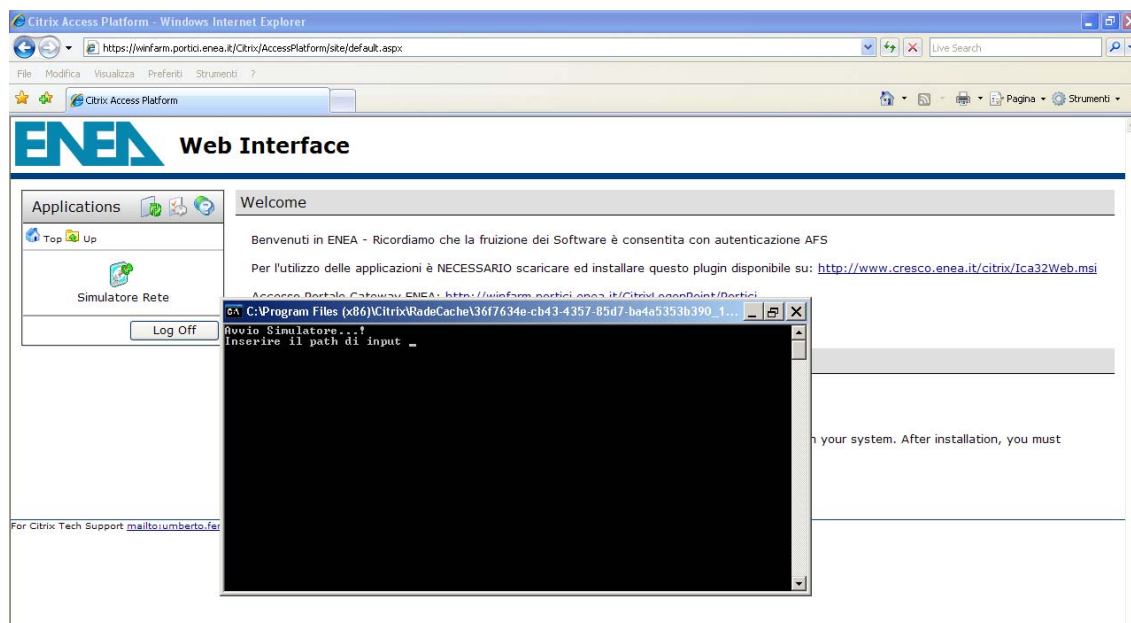


Figura 4 - Accesso al codice Simulatore Rete

Come si evince dalle figure l'accesso alla Windows Farm avviene mediante un browser web (nello specifico si trattava di Internet Explorer) ma è necessario, per poter eseguire un'applicazione, installare un *plugin*, direttamente accessibile tramite un link sulla *home page* della Farm e che, sostanzialmente, implementa all'interno del *browser* il protocollo ICA.

ICA (*Independent Computing Architecture*) è un protocollo proprietario per sistemi di tipo application server, progettato da Citrix Systems. Il protocollo stabilisce una specifica per trasferire i dati fra il client ed il server ma non è legato a nessuna piattaforma specifica. Ciò consente ad esempio che applicazioni Windows ordinarie siano eseguite su un server Windows e che i client supportati possano accedere a tali applicazioni; tali client non dovranno eseguire necessariamente Windows ma potrebbero essere Mac, Unix, Linux. Mediante tale protocollo il carico per il client che attiva l'esecuzione delle applicazioni sul server è minimo in quanto il protocollo prevede solo l'invio di *keystrokes*, *mouse clicks* e *screen updates* tra il client e il server.

4. Conclusione

L'attività ha consentito di mettere a disposizione di tutti gli utenti di ENEA-GRID i codici di modellazione di reti con approccio fluidodinamico sviluppati dal Dipartimento di Ingegneria dell'Informazione e Matematica Applicata dell'Università degli Studi di Salerno.

L'attività ha previsto la creazione di un kit di installazione per ciascun codice al fine di effettuare correttamente la virtualizzazione e quindi la creazione del profilo.

L'utente che si collega alla Windows Farm può lanciare la singola applicazione, che grazie ad un sistema di bilanciamento previsto da Citrix Presentation Server sarà eseguito dal server che in quel momento è meno *carico*.

Su tali server non si dovrà installare nemmeno un byte di codice in quanto essi eseguiranno l'applicazione virtualizzata e disponibile sul file system distribuito AFS.

Gli utenti potranno accedere allo stesso file system sia per fornire i dati di input sia per archiviare i dati di output dei codici.