

# I chip che renderanno la cloud più 'verde'

Giorgio Fusari

**I servizi It veicolati attraverso la nuvola informatica stanno esplodendo sul mercato, ma per rendere i data center più efficienti e sostenibili anche nei prossimi anni è necessario innovare le architetture di microprocessore per le macchine server**

Il paradigma di It rappresentato dal cloud computing permette oggi di utilizzare risorse informatiche hardware e software, distribuite e virtualizzate nel data center di un provider, fruendole come servizio attraverso la nuvola (Internet) e pagandole secondo la formula del canone a consumo. In funzione dei modelli di provisioning, si possono distinguere alcune tipologie fondamentali di servizi cloud: soprattutto, si parla di servizi SaaS (Software as-a-service), PaaS (Platform as-a-service) e IaaS (Infrastructure as-a-service), ormai disponibili sul mercato attraverso numerosi provider. Si parla anche di cloud pubbliche e private. Nel primo caso si intendono servizi disponibili per i normali utenti o le aziende, tramite fornitori come Amazon, Microsoft, IBM o Fujitsu. Nel secondo, ci si riferisce di norma a infrastrutture cloud gestite internamente da una singola organizzazione.

**Strada segnata**

L'utilizzo di applicazioni e servizi It attraverso il cloud

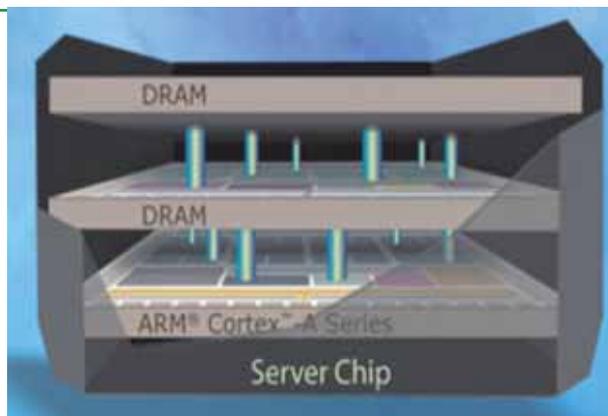
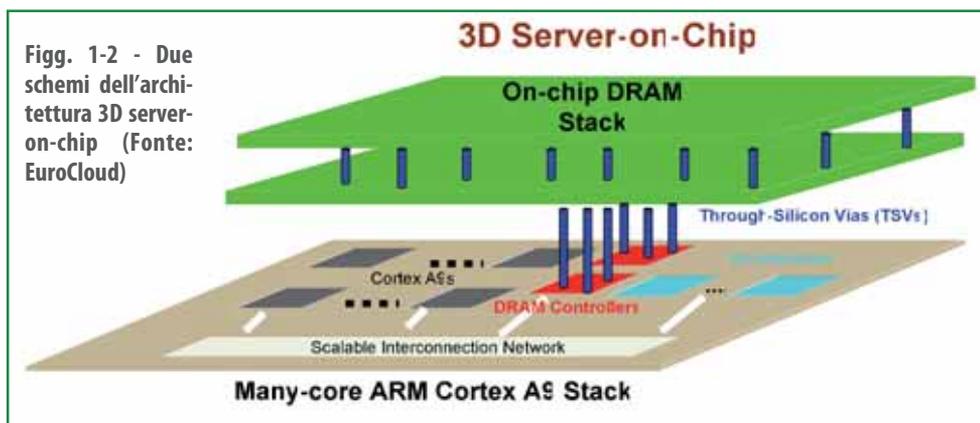
computing sta emergendo in maniera netta a livello globale e diventando sempre più diffuso. Sia fra i normali utenti, che amano la 'nuvola' per elaborare e memorizzare file, ad esempio attraverso servizi come le Google Apps o Dropbox, sia per le aziende di varie categorie e

dimensioni che, dopo la fase degli interventi di consolidamento e virtualizzazione delle risorse It, sulla cloud cominciano a trasferire servizi e applicazioni, anche se di solito di natura non business-critical.

È vero che in Italia i timori delle aziende sulle architetture di cloud computing,

soprattutto legati ai problemi di sicurezza e privacy dei dati sensibili memorizzati nei server remoti che forniscono i servizi, hanno incominciato in parte a dissolversi, grazie a un'aumentata conoscenza della 'nuvola'. È vero anche che le continue rassicurazioni dei fornitori sull'affidabilità e sui livelli di protezione garantiti dalle più moderne tecnologie di cifratura delle informazioni hanno contribuito a far crescere la fiducia. Tuttavia, soppesando i rischi di natura tecnica e legale, da chiarire al momento di definire e stipulare l'accordo Sla (Service level agreement) con il cloud provider, anco-

ra molte imprese restano convinte che sia più sicuro realizzare una cloud privata, comprando i server e gestendoli in-house, a cura dello staff It. Una delle paure più grandi resta legata al fatto che nella cloud le informazioni sensibili dell'utente possono essere memorizzate in data center ubicati in diverse zone geografiche del mondo, dove l'esistenza di giurisdizioni differenti da quelle del paese d'origine (si pensi alle differenze fra Europa e Stati Uniti in materia di privacy) non rende certo chiare le attribuzioni di responsabilità, e le relative garanzie, in caso di perdita, danneggiamento o fuga



## DUE APPROCCI PER OTTIMIZZARE L'ATTIVITÀ DEI CORE

L'approccio per ridurre i consumi dei data center può avvenire su due fronti, spiega Giorgio Richelli, systems architect di Ibm Italia. "Una strada è quella di usare processori tradizionali, lavorando poi sul software e sull'efficienza dei sistemi di raffreddamento". Nei sistemi convenzionali, quindi nei processori non pensati per ottimizzare i consumi, occorre agire sulle funzionalità di controllo della cpu. "Ad esempio, si possono usare i sistemi di scheduling intelligenti: quando si ha un data center con molti processori non utilizzati in maniera costante al 100 per cento, la cosa migliore da fare è metterli in uno stato in cui consumano poco". A seconda delle tipologie di core, si può intervenire sui diversi meccanismi di regolazione. Oppure, per ridurre i consumi del data center si può osservare, a fronte delle misurazioni eseguite all'interno del chip, come variano le performance del job in esecuzione, in funzione della frequenza di clock della cpu. Ciò permette di arrivare a stabilire quali sono i migliori rapporti fra prestazioni e consumi, per tutti i differenti tipi di workload che lo scheduler deve gestire.

Sull'altro versante, la via alternativa per ottimizzare l'energia spesa nelle sale dati è usare processori di tipo diverso, specializzati, e in grado di fornire determinati livelli di performance in rapporto ai consumi. Ad esempio, si può talvolta accettare il compromesso di ridurre un po' le prestazioni, per risparmiare una certa quantità di energia. In Ibm, Richelli cita il progetto Blue Gene, in cui i processori sono studiati per ottimizzare l'efficienza energetica e quindi il rapporto prestazioni (gigaflops) per watt. In generale, le performance rispetto ai consumi non hanno un andamento lineare, e dipendono dal tipo di lavoro svolto dal processore.



Il supercomputer Blue Gene

di informazioni preziose per il business dell'azienda. Con la cloud privata però, l'impresa utente non ottiene i vantaggi chiave della nuvola nella sua più pura accezione, ossia la cloud pubblica. La sola attraverso cui un'organizzazione può permettersi di evitare nuovi investimenti in server fisici e altre risorse It (personale di supporto), acquisendo direttamente via rete le applicazioni, la capacità di storage o la potenza computazionale necessaria per la propria attività.

### Energia, un grosso problema da risolvere

La grande eco mediatica e i vantaggi in termini di costi

e comodità prospettati agli utenti non significano che il cloud computing sia tutto rose e fiori, come spesso viene dipinto. Dal punto di vista di chi deve fornire i servizi cloud, e soprattutto pensando ai grandi cloud provider, a prescindere da altre varie tipologie di problemi, resta da risolvere la delicata questione centrale della gestione dei pesanti consumi di energia dei server e delle infrastrutture tecnologiche ospitati all'interno dei data center, da dove i servizi cloud vengono erogati. Già nel 2006, l'Epa (Environmental protection agency), l'agenzia per la protezione ambienta-

le americana, aveva stimato che negli Stati Uniti l'elettricità consumata dai server nei centri dati, comprese le infrastrutture ausiliarie e di raffreddamento, rappresentava circa l'1,5 per cento dell'utilizzo della corrente elettrica a livello nazionale. Un altro rapporto, elaborato nel 2008 dalla società di consulenza McKinsey & Company, prevedeva che, a livello mondiale, le emissioni di gas serra prodotte dai data center quadruplicheranno per il 2020. Ma, già oggi, la tendenza relativa ai fabbisogni di energia nei centri dati appare chiara: l'adozione crescente di servizi cloud anche in

mobilità, attraverso i dispositivi tablet e smartphone, e i trend derivanti dalla 'consumerizzazione' dell'It nel mondo aziendale, come il Byod (Bring your own device), stanno facendo sviluppare enormemente il traffico di informazioni e i dispositivi in gioco. Di conseguenza, nei prossimi anni, i server presenti nei data center, oltre a soddisfare le più classiche necessità di tipo aziendale o applicazioni come quelle di Hpc (High performance computing), attraverso la cloud dovranno indirizzare anche le esigenze di comunicazione, storage e gestione dati di miliardi di thin client e dispositivi mobili, sempre più usati da normali consumatori, professionisti e addetti aziendali.

In questo scenario, abbassare i consumi in un'ottica di sostenibilità ambientale e di business diventa strategico: se tradizionalmente i computer e server delle sale dati sono stati sviluppati sulla base di normali architetture x86, divenute 'commodity', oggi i maggiori produttori di schede e server stanno lavorando per la fornitura ai data center di macchine e attrezzature a corredo sempre più energy-efficient e specializzate nell'ottimizzazione dei consumi. In particolare, sul mercato si sviluppano le offerte di microprocessori multicore dotati di funzionalità di power management, cioè della capacità di scalare in modo dinamico la

frequenza e la tensione del dispositivo in funzione dei carichi di lavoro. Aumenta anche l'utilizzo dei sistemi di alimentazione ad alta efficienza, e delle ventole di raffreddamento in grado di variare la velocità in funzione della necessità di raffreddamento del server.

### EuroCloud, microchip 3D per server a basso consumo

I cloud data center tradizionali, 'affamati' di energia, non sono sostenibili nel lungo periodo. Lo dichiara anche un comunicato della Commissione europea, che per il 2020 prefigura un universo digitale delle dimensioni di 35 zettabyte, rispetto agli 1,8 zettabyte del 2011. A quel punto le attività di alimentazione e raffreddamento dei centri dati diventeranno diseconomiche per le società europee, se non verranno introdotte innovazioni a livello dei chip e dei server.

Ecco perché, per trovare soluzioni più efficienti rispetto alle piattaforme server convenzionali, basate su microprocessori e sistemi di memoria che generano notevoli sfide per i responsabili dei sistemi informativi a livello di alimentazione, raffreddamento e spazio occupato nella sala dati, è stato avviato nel gennaio 2010 il progetto EuroCloud. Un progetto europeo che ha ricevuto un fondo di finanziamento di 3,3 milioni di euro per tre anni, con completamento a fine 2012, e

che ha visto coinvolte le società Arm, Nokia, l'istituto di ricerca belga Imec, il politecnico di Losanna e l'università di Cipro (Ucy).

Al centro del progetto congiunto la realizzazione di speciali microchip 3D, chiamati '3D server-on-chip' e ideati sulla base delle tecnologie di microprocessore low-power usate nel settore dei telefoni cellulari. Questi chip sono dieci volte meno costosi, e in grado di funzionare consumando dieci volte meno energia rispetto alle architetture server convenzionali. A livello strutturale, tali chip 3D integrano memorie Dram impilate fra loro e interconnesse con la piattaforma Arm Cortex-A9 tramite tecnologia Tsv (Through silicon via). Ciò permetterà di costruire singoli server dotati di centinaia di core embedded con l'obiettivo, realizzabile in futuro, di arrivare a costruire data center con un milione di core.

I 3D server-on-chip puntano a consentire la costruzione di data center super efficienti, compatti e puliti dal punto di vista ecologico, per la fornitura di servizi cloud 'green', in grado anche di rafforzare la posizione europea nell'area del computing sostenibile, e di far risparmiare a molti provider miliardi di euro. Il vantaggio è che, integrando un core di elaborazione low-power in stretta prossimità con la memoria, e in tre dimensioni, è possibile accedere alla memoria on-chip usando

molto meno energia, perché le distanze fisiche fra i componenti diventano più brevi. Fra gli altri obiettivi del progetto EuroCloud, c'è anche lo studio del funzionamento del software nei data center,

di Melbourne. In sintesi, in determinate circostanze d'uso, e soprattutto in condizioni di utilizzo intenso, come nel caso dello storage as-a-service (es. Dropbox), del software as-



Fig. 3 - Un data center Ibm

e la misurazione e analisi del comportamento delle applicazioni cloud, per arrivare alla progettazione di chip e piattaforme server ottimizzati per questo tipo di utilizzo.

In aggiunta, altri studi sono stati focalizzati sui sistemi di interconnessione e su nuove metodologie per organizzare meglio l'accesso alle memorie on-chip. Tuttavia, va anche detto che i vantaggi ottenibili con i chip per server ideati nel progetto EuroCloud, dieci volte più efficienti di quelli convenzionali, possono essere in parte ridotti dai livelli di consumo energetico che spesso si raggiungono utilizzando servizi e applicazioni nella nuvola: il riferimento è a una ricerca australiana, condotta sul tema del green cloud computing dall'università

a-service (es. GoogleDocs) o della potenza computazionale fruita sotto forma di servizio (es. Amazon Elastic Cloud), la nuvola fa consumare anche molta più energia del computing svolto in maniera convenzionale, in cui ciascun utente esegue gli stessi compiti ma localmente, dal proprio pc sulla scrivania. Il fabbisogno di elettricità è maggiore, perché nel calcolo dell'energia necessaria per l'uso dei servizi cloud non basta semplicemente stimare l'entità dei consumi di tutti i microprocessori del data center, ma occorre considerare anche l'energia consumata dagli apparati e attrezzature di storage e da tutti i dispositivi di networking (switch, router, reti di comunicazione) necessari per il trasporto dei dati in remoto. ■