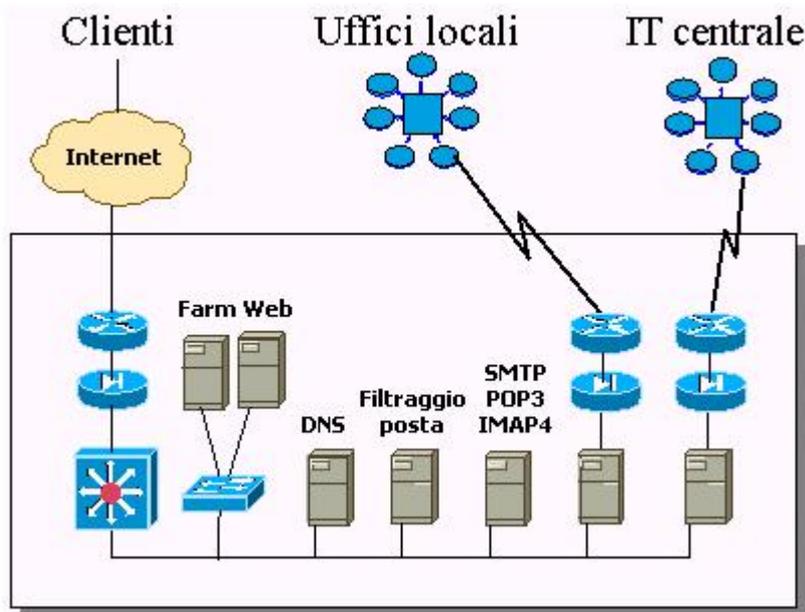


Importanza del dimensionamento delle risorse

Il dimensionamento delle risorse ha lo scopo di garantire che le capacità IT (ad esempio il supporto applicativo, le infrastrutture fisiche e così via) sulle quali si basano le soluzioni degli ASP (*application services provider*) siano in grado di soddisfare le mutevoli richieste dei clienti, che le risorse esistenti vengano utilizzate al meglio e che i necessari aggiornamenti vengano completati nel modo più rapido ed economico possibile.

Importanza del dimensionamento delle risorse

Per un *provider* di soluzioni applicative è importante riuscire a soddisfare le richieste dei clienti. Se non si dispone delle capacità sufficienti per offrire i livelli di servizio richiesti è possibile incorrere in tempi di risposta prolungati, *time-out*, errori e interruzione delle applicazioni.



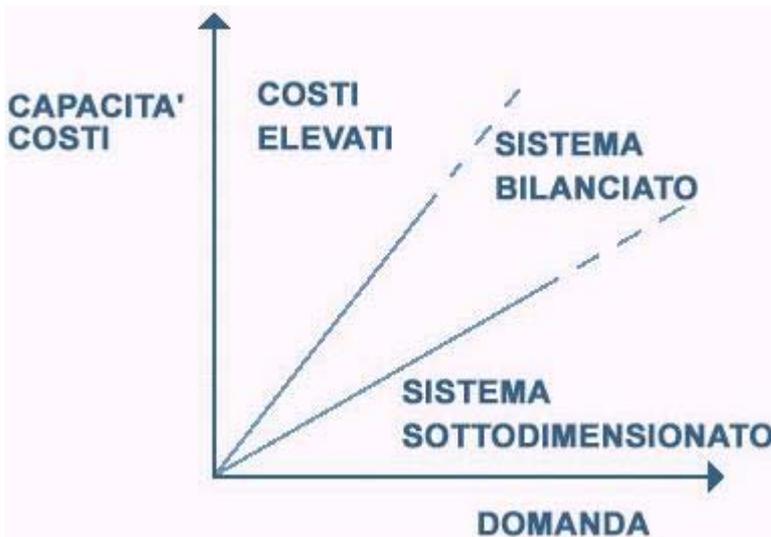
Per evitare negazioni di servizio è fondamentale che gli ASP creino infrastrutture capaci di gestire non solo i livelli *standard* di richieste, ma anche volumi di richieste molto più elevati. Nell'ambito della gestione delle capacità vengono utilizzati sistemi OSS (*Operations Support Systems*) per verificare che l'utilizzo delle risorse esistenti nelle soluzioni ASP sia adeguato ai livelli di servizio stabiliti.

Se eseguito correttamente, il processo di gestione delle risorse consente di valutare la quantità di *hardware* necessaria per gestire il volume di richieste degli utenti relative a una particolare soluzione ASP. Questo tipo di valutazione è utile per individuare i punti deboli delle infrastrutture su cui si basano le soluzioni ASP e i colli di bottiglia che possono causare inefficienza nelle prestazioni. Mediante l'aggiunta di componenti *hardware* o la ridefinizione dei requisiti di una soluzione ASP è possibile ottimizzarne le prestazioni eliminando i punti deboli.

Il dimensionamento è un processo mediante il quale il personale tecnico di un ASP predispone la struttura della soluzione applicativa per futuri ampliamenti che consentano di far fronte a un numero maggiore di richieste. Se per la gestione delle risorse si utilizzano strumenti automatici e adatti a diversi tipi di applicazioni, è possibile ridurre la necessità di intervento umano, aumentare il grado di soddisfazione dei clienti e la redditività del modello ASP.

Concetti fondamentali del dimensionamento delle risorse per gli ASP

Il dimensionamento delle risorse ha la funzione di valutare se l'infrastruttura è in grado di offrire un livello di prestazioni accettabile e adeguato alle esigenze.



Il processo include:

- il monitoraggio delle richieste attuali dei clienti relativamente alle risorse IT e la creazione di previsioni in merito ai futuri requisiti dei clienti.
- La canalizzazione delle richieste dei clienti, ad esempio in combinazione con i *Service Level Agreement*, allo scopo di distribuire gli effetti dei picchi di utilizzo.
- L'acquisizione delle risorse necessarie in base ai requisiti dei clienti e la creazione di previsioni in relazione alle esigenze future (come regola generale è preferibile disporre di capacità superiori rispetto alle effettive necessità in modo da essere in grado di fare fronte ad eventuali picchi imprevisti. I costi derivanti dalla mancata capacità di soddisfare le richieste dei clienti sono infatti notevolmente maggiori rispetto ai costi necessari per l'acquisizione delle risorse necessarie).
- Il monitoraggio delle prestazioni e del *throughput* dei servizi ASP.
- Il monitoraggio delle prestazioni dei componenti che costituiscono l'infrastruttura di supporto e l'esecuzione di attività di ottimizzazione che consentano di trarre il massimo vantaggio dalle risorse esistenti.
- La definizione di un piano relativo alla disponibilità di risorse che consenta all'ASP di offrire servizi con un livello di qualità conforme a quanto stabilito nel *Service Level Agreement*.

Il dimensionamento delle risorse ha lo scopo di bilanciare i seguenti fattori:

- **Offerta e domanda.** Il processo consente di verificare che le risorse disponibili (ad esempio la potenza del processore e lo spazio di archiviazione) corrispondano alle richieste attuali e future dei clienti delle soluzioni ASP.
- **Costi e capacità.** È necessario verificare attentamente che i costi sostenuti per l'acquisizione delle capacità siano giustificabili non solo dal punto di vista delle esigenze aziendali, ma anche che le capacità acquisite consentano di utilizzare le risorse nel modo più efficiente possibile.

Un metodo estremamente semplicistico per la valutazione delle capacità consiste nel calcolare il numero di utenti che accedono a una determinata soluzione ASP in relazione al carico di richieste che ogni utente invia ai diversi componenti della soluzione. Questo semplice calcolo può essere utilizzato per potenziare le risorse che possono ridurre le capacità del sistema (CPU, RAM, spazio

su disco e larghezza di banda) in modo che siano in grado di supportare i livelli di utilizzo presenti e futuri.

Dimensionamento di un sito

Il processo di dimensionamento delle risorse è costituito da diversi sottoprocessi, che includono a loro volta varie attività.



Tali sottoprocessi sono descritti di seguito:

- **gestione delle richieste.** Questo sottoprocesso ha l'obiettivo di garantire che i requisiti aziendali futuri dell'ASP in relazione ai servizi IT vengano considerati, pianificati e implementati in modo tempestivo. A questo scopo, il *team* preposto alla gestione delle risorse ne analizza l'utilizzo attuale all'interno delle varie soluzioni ASP per valutare le tendenze predominanti e sviluppare le necessarie previsioni. I requisiti futuri vengono definiti dalla direzione amministrativa dell'ASP, che ha il compito di analizzare costantemente le esigenze presenti e future della clientela.
- **Gestione del carico di lavoro.** Questo sottoprocesso ha l'obiettivo di interpretare le richieste degli utenti in termini di carico di lavoro sui vari componenti della soluzione ASP (le diverse applicazioni utilizzate per creare la soluzione) e utilizzare questa analisi per determinare le risorse necessarie. Il processo traduce in carico di lavoro sia le richieste presenti sia le richieste previste per il futuro.
- **Gestione delle prestazioni.** Questo sottoprocesso ha l'obiettivo di verificare le prestazioni delle soluzioni ASP utilizzate dai clienti e le prestazioni delle singole risorse sulle quali tali soluzioni si basano. Mediante tale processo è possibile monitorare le prestazioni di tutte le soluzioni ASP (in base ai parametri definiti nei contratti SLA) e delle relative risorse sottostanti. Tutti i dati raccolti vengono registrati, analizzati e documentati. Se necessario, il *team* preposto al dimensionamento delle risorse verificherà che le prestazioni delle soluzioni siano conformi ai requisiti aziendali.

I dati raccolti vengono analizzati e utilizzati per eseguire prove di ottimizzazione e definire profili mediante i quali identificare e impostare valori soglia e avvisi. Quando vengono generati report delle eccezioni o avvisi, è necessario analizzarli e documentarli e quindi intraprendere le necessarie azioni correttive. In genere è preferibile impostare le soglie e gli avvisi su valori inferiori rispetto a quelli stabiliti nel **Service Level Agreement**. In questo modo sarà possibile intraprendere le necessarie azioni correttive prima che si verifichino effettive inadempienze ai contratti SLA.

I servizi e le modifiche

I servizi

La gestione dei servizi rappresenta uno dei punti chiave e ha lo scopo di fornire e supportare i servizi IT che corrispondono ai requisiti aziendali dei clienti dell'ASP. La gestione delle relazioni con i clienti e la gestione dei servizi sono processi strettamente correlati. Tuttavia, l'obiettivo della

gestione dei servizi consiste soprattutto nel fornire i servizi concordati con i clienti e, pertanto, privilegia gli aspetti di tipo operativo e tattico, mentre la gestione delle relazioni con i clienti ha una funzione più che altro strategica. La gestione delle relazioni con i clienti consente inoltre di fare previsioni sui costi che sarà necessario sostenere per soddisfare i futuri requisiti dei clienti relativamente alle prestazioni.

Le modifiche

La gestione delle modifiche rappresenta un aspetto molto importante nell'ottica dell'integrità e del corretto funzionamento di un sistema. Il processo di controllo delle modifiche offre l'opportunità di sottoporre le modifiche a una procedura di approvazione e di analizzare attentamente le modifiche richieste dai clienti. Questa analisi consente di valutare le richieste, ordinare le risorse necessarie in tempo utile e adeguare i piani relativi alle capacità in base alle modifiche implementate. Il controllo delle modifiche consente di definire uno schema dettagliato che specifica come e quando le modifiche verranno implementate. In questo modo è possibile ridurre il rischio che vengano apportate modifiche all'ambiente ASP senza che prima ne vengano attentamente valutate le possibili conseguenze sulle richieste dei clienti e sulle risorse.

Durante il processo di gestione delle modifiche è opportuno creare un documento che definisca in modo dettagliato il flusso di lavoro e le modifiche da implementare, incluso lo scopo della modifica, il risultato che questa comporterà e il momento in cui dovrà essere implementata. Questa definizione viene utilizzata dai processi di gestione della disponibilità e delle risorse per determinare gli effetti delle modifiche sulle capacità dei sistemi.

Valutazione dei rischi

Il processo di gestione delle modifiche dovrebbe includere una valutazione dei rischi che le modifiche da implementare potrebbero comportare sulle capacità e le richieste esistenti. Per ogni modifica viene valutato il livello di rischio potenziale (che può essere assente, di *media* entità o molto elevato), che rappresenta un fattore fondamentale di cui è necessario tenere conto durante il processo di approvazione della modifica.

- **Processo di Approvazione.** Il processo di approvazione si basa sulle informazioni finora descritte. Per semplificare lo svolgimento del processo, è consigliabile che le modifiche vengano sottoposte agli organi di approvazione utilizzando un formato *standard*. Ad esempio, è possibile creare un modulo di esempio precompilato che chi propone le modifiche potrà modificare in base alle esigenze. È inoltre importante definire indicazioni di massima in relazione ai tempi che dovrebbero intercorrere tra la proposta e l'approvazione delle modifiche. Per le modifiche più urgenti o che è necessario implementare prima dei tempi stabiliti dovrà essere utilizzato lo stesso formato, sollecitandone tuttavia l'approvazione.
- **Procedure di verifica delle modifiche.** In seguito all'implementazione e all'esecuzione della modifica, è necessario fornire agli organi di approvazione le informazioni necessarie per verificare che la modifica abbia prodotto gli effetti desiderati. Quindi, in base ai risultati della valutazione, è possibile intraprendere eventuali azioni correttive (ad esempio, se la modifica non ha avuto gli effetti voluti potrebbe essere necessario eseguirne il *roll-back*).

Equazione per il calcolo delle capacità

Per stabilire la quantità di risorse necessaria per un determinato servizio occorre analizzare i **picchi** di utilizzo. Se si è in grado di soddisfare le richieste dei clienti durante i periodi di massimo utilizzo del servizio, le risorse utilizzate saranno sicuramente sufficienti nei periodi in cui il servizio viene utilizzato in misura minore.

L'equazione su cui si basa il dimensionamento delle risorse include tre **fasi**. Nella prima fase

vengono valutate le richieste, nella seconda vengono calcolati i carichi di lavoro e nell'ultima fase viene calcolata la quantità di risorse necessarie. Di seguito viene riportato un esempio semplificato dell'equazione appena descritta.

Calcolo delle richieste

Richieste totali = Numero di utenti simultanei x Richieste dei singoli utenti

Calcolo dei carichi di lavoro

Carico di lavoro A = Richieste totali x Carico di lavoro per le singole richieste dell'applicazione A

Carico di lavoro B = Richieste totali x Carico di lavoro per le singole richieste dell'applicazione B

Calcolo delle risorse

Spazio di archiviazione (MB) =

Carico di lavoro A x MB necessari per ogni unità del carico di lavoro A

+ Carico di lavoro B x MB necessari per ogni unità del carico di lavoro B

Potenza CPU = Carico di lavoro A x Potenza CPU necessaria per ogni unità del carico di lavoro A

+ Carico di lavoro B x Potenza CPU necessaria per ogni unità del carico di lavoro B

Larghezza di banda =

Carico di lavoro A x Larghezza di banda per ogni unità del carico di lavoro A

+ Carico di lavoro B x Larghezza di banda per ogni unità del carico di lavoro B

Nell'esempio si presuppone che entrambe le applicazioni risiedano nello stesso *computer*.

Risolvendo l'equazione sarà possibile stabilire la quantità di utenti che l'applicazione può gestire nei periodi di massimo utilizzo. Questo metodo consente di calcolare facilmente il numero di utenti che l'applicazione è in grado di supportare. Facciamo inoltre alcune considerazioni:

- diminuendo il carico di lavoro sull'applicazione correlato alle richieste dei singoli utenti è possibile incrementare il numero di utenti supportati. A questo scopo è necessario pianificare, programmare e configurare adeguatamente l'applicazione per ottimizzare l'utilizzo delle risorse.
- incrementando le capacità delle risorse che possono limitare le prestazioni del sistema (ad esempio RAM, larghezza di banda, licenze e *server*) è possibile incrementare il numero di utenti supportati.

La scalabilità **verticale** consente di aggiungere capacità a una risorsa esistente. Un esempio di scalabilità verticale è rappresentato dall'aggiunta di più processori, RAM o spazio su disco a un *server* esistente. La scalabilità **orizzontale** implica invece l'aggiunta di risorse, quali *server*, connessioni di rete o personale.

Risorse che possono ridurre le capacità del sistema

Il primo passaggio nell'identificazione delle risorse che possono ridurre le capacità del sistema consiste nell'analizzare i termini del cosiddetto *Service Level Agreement* (ciò che si è offerto al cliente). Generalmente, gli ASP sono responsabili della gestione dei seguenti aspetti delle soluzioni:

- infrastruttura della soluzione applicativa;
- connettività di rete;

- interfaccia del cliente con l'applicazione.

Indipendentemente dal metodo di distribuzione utilizzato per l'applicazione, ai fini della pianificazione delle capacità è necessario monitorare quattro risorse fondamentali, ovvero processore, memoria (RAM), array/sottosistema di dischi e rete. Di seguito vengono descritti gli aspetti di ognuna di queste risorse che è necessario monitorare nel contesto di una soluzione ASP:

- **Processore.** L'utilizzo del processore da parte di un determinato componente *hardware*. Un processore che viene utilizzato per più dell'**80%** può indicare l'esistenza di un problema di insufficienza delle capacità. Non è raro che un'applicazione includa processi a *thread* singolo, ovvero processi che possono accedere a un solo processore per eseguire le transazioni. È possibile quindi che un processo poco efficiente utilizzi un processore al 100 per cento mentre tutti gli altri processi nella stessa partizione lo utilizzano allo 0 per cento. Una situazione di questo tipo indica dei problemi nel codice dell'applicazione, che dovrebbe pertanto essere riscritta.
- **Memoria.** L'utilizzo della memoria dovrebbe essere monitorato per tutte le partizioni di memoria allocate in modo indipendente. La memoria insufficiente rappresenta uno dei problemi più comuni per gli ASP. In genere, i problemi di utilizzo della memoria si manifestano inizialmente attraverso incrementi di utilizzo del processore e dell'*array* di dischi, in quanto queste risorse vengono utilizzate maggiormente per gestire la memoria virtuale ripartendone i dati tra l'*array* di dischi e la RAM disponibile.
- **Dischi.** I segnali che indicano un problema di utilizzo dell'*array* di dischi consistono generalmente nella mancanza di spazio su disco o nell'utilizzo **eccessivo** di una parte dell'*array*.
- **Rete.** Per rete si intendono tutti gli aspetti che riguardano le risorse che risiedono sulla LAN aziendale, inclusi l'utilizzo delle singole schede di rete, i tempi di latenza, la larghezza di banda disponibile e le prestazioni di *switch* e *router*.

Per ognuna di queste risorse è possibile eseguire un monitoraggio più o meno dettagliato. In genere, il sistema operativo o i sottosistemi di dischi dispongono di meccanismi in grado di monitorare queste risorse in modo molto preciso.

Siti ridimensionabili

Cambia la risoluzione del monitor e il sito cambia. E' un errore frequente di chi si trova a sviluppare per la prima volta pagine *Web*. Spesso si ragiona basandosi su parametri cartacei, quando tutto rimane fermo, fisso e immutabile. Così si cerca di ottimizzare il sito in modo che non compaiano barre orizzontali o verticali, ma poi ci si dimentica di domandarsi come si vedrà il sito a una risoluzione del monitor diversa da quella in cui si lavora?.

Uno dei pregi del *Web* è infatti la relativa indipendenza dei **contenuti** dai **contenitori**: le pagine *Web* possono essere viste infatti in moltissimi modi. Il che vuol dire anche che posso girare nel *Web* a partire da qualsiasi piattaforma (*Windows*, *Linux* o *Macintosh*), con qualsiasi *browser* (*Internet Explorer*, *Netscape Navigator*, *Mozilla*, *Opera*...), e a qualsiasi risoluzione dello schermo (640x480, 800x600, 1024x768, eccetera). Ma potrei vedere pagine *Web* anche su PALM!

Bisogna sapere adattare il proprio sito a molteplici situazioni. Oggi la maggior parte degli utenti *Internet* imposta il proprio monitor in questo modo (dati tratti dal sito **Netmechanic**):

- il 50% degli utenti utilizza una risoluzione di 800x600;
- il 40% naviga a 1024x768;
- circa il 10% utilizza ancora una risoluzione di 640x480.

I dati sono solo indicativi e non tengono conto del *target*: infatti se dovessimo sviluppare un sito di

ICT, dovremmo anche tenere conto del fatto che chi opera nel ramo informatico di solito utilizza monitor piuttosto grandi che gli consentono una risoluzione di 1024 x 768 (con cui è più semplice lavorare...).

In ogni caso, visto che oramai gli utenti che utilizzano risoluzioni del monitor di 800x600 e quelli che invece preferiscono 1024x768 si equivalgono, è opportuno sviluppare il sito in modo che si veda in maniera corretta a entrambe le risoluzioni.

Come impostare un sito indipendente dalla risoluzione

Iniziamo dunque a ottimizzare il nostro sito a 800 x 600: si tratta di evitare che nel sito compaiano barre orizzontali e verticali (in quest'ultimo caso solo se lo si ritiene opportuno). Impostiamo dunque il sito a una grandezza di 780x450 (con *Windows* XP, che utilizza delle barre più grosse, queste grandezze diventano: 779x430).

Dopodiché abbiamo diverse possibilità. A risoluzioni superiori di 800x600:

1. il sito rimane allineato a sinistra;
2. il sito rimane centrato a qualsiasi risoluzione;
3. il sito si ridimensiona, occupando la totalità dello schermo;

A seconda del tipo di *layout* potremo preferire una soluzione oppure l'altra. Ma analizziamo le varie situazioni nel dettaglio.

1. **Sito allineato a sinistra.** Si tratta per lo più di una svista o di un errore di progettazione. Tuttavia in particolari contesti il sito potrebbe rimanere tranquillamente sulla sinistra del monitor: in questo caso il *webmaster* giocherà abilmente con il colore di sfondo in modo da mimetizzare il tutto. Così la parte che a 1024 o a risoluzioni superiori viene lasciata vuota, apparirà invece come logica continuazione del *layout*.
2. **Sito centrato.** È la soluzione che prediligono molti portali o testate giornalistiche (**HTML.it** stesso è impostato in questo modo). Per lo più il sito viene sviluppato con una testata (*header*), una colonna centrale, una o più colonne laterali, e una chiusura (*footer*). In effetti ricorda molto la disposizione tradizionale dei quotidiani. La centratura del sito viene per lo più ottenuta con un semplice:

```
<DIV ALIGN="center">
```

aperto subito dopo il `<BODY>` e chiuso prima della chiusura dello stesso `</BODY>`.

3. **Sito ridimensionabile.** Il sito occupa la totalità del monitor a qualsiasi risoluzione (partendo però da 800x600). In questo caso utilizzeremo le tabelle in percentuale in modo che ci siano una o più celle mobili, in grado di occupare tutto lo spazio necessario.

Utilizziamo le tabelle in percentuale

Abbiamo visto che per far sì che un sito si ridimensioni a tutta pagina occorre mantenere varie celle fisse e lasciare una o più celle libere di ridimensionarsi a proprio piacimento: saranno queste celle libere a colmare lo spazio che altrimenti risulterebbe vuoto. Potrebbe sembrare particolarmente laborioso l'utilizzo delle tabelle in percentuale. A prima vista saremmo infatti portati a credere di dover utilizzare le proporzioni matematiche per ri-calcolare la percentuale delle celle di tutto il sito.

Nulla di più sbagliato. Esiste infatti un piccolo escamotage (che funziona correttamente con tutti i *browser*), che ci permette di risparmiare un sacco di lavoro.

Incominciamo a procurarci una gif trasparente di 1 *pixel* x 1 *pixel*. Si chiama *shim* ed è un'immagine vuota che nell'ordinaria costruzione di un sito ci è utile in una molteplicità di situazioni:

- si sa che, se non c'è nulla all'interno di una cella, *Netscape 4* non la vede. È allora indispensabile riempire la cella con qualsiasi cosa: un *break* (
), un non-breaking-space (), oppure la nostra immagine vuota. (*Netscape* in origine aveva creato un tag apposito - lo <SPACER> - per risolvere il problema delle celle).
- Un altro problema che si verifica con *Netscape 4* è che, quando ci sono delle tabelle annidate, da una cella interna non è possibile vedere lo sfondo che sta sopra. Anche in questo caso la situazione si risolve mettendo un'immagine vuota di sfondo. Così:

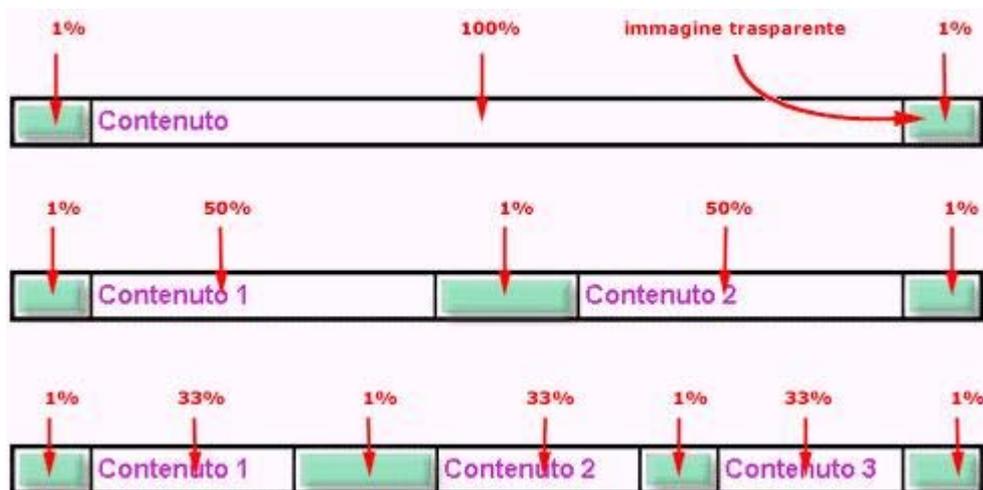
```
<table width="500" border="1" cellspacing="0" cellpadding="0" height="400">
  <tr>
    <td background="imgs/sfondoOmino.gif" align="center" valign="middle">
      <table width="100%" border="0" cellspacing="0" cellpadding="0">
        <tr> <td background="imgs/shim.gif"> </td>
          <td background="imgs/shim.gif"> </td>
          <td background="imgs/shim.gif"> </td>
        </tr>
      </table>
    </td>
  </tr>
</table>
```

- Se non ci fosse l'immagine trasparente, con *Netscape 4* la tabella interna avrebbe uno sfondo bianco.

La nostra immagine vuota, lo shim, ci viene in aiuto anche nel caso delle tabelle in percentuale. Per quanto possa sembrare strano, infatti, non occorre utilizzare delle percentuali realistiche, ma basta dare a una cella il valore dell'1% perché questa si schiacci sul proprio contenuto, basta darle invece un valore del 100% perché si allarghi al massimo. È evidente, a questo punto, che le dimensioni non vanno più *date* alle celle (che saranno dell'1% o del 100%), ma va attribuito al contenuto.

Si tratta dunque di prendere il nostro shim, dargli le dimensioni che preferiamo, e quindi schiacciargli contro la cella: la dimensione della cella viene così tenuta dallo *shim*. Stesso procedimento nel caso in cui avessimo del contenuto vero contro cui schiacciare la cella (ricordarsi di usare il NOWRAP per le scritte).

La pagina non risulta appesantita dall'utilizzo di tutte queste immagini vuote: l'immagine viene infatti caricata una sola volta. Ovviamente se le celle ridimensionabili sono più di una, il 100% deve essere diviso fra tutte le celle che devono essere ridimensionate. Nella figura sottostante si trovano tre esempi.



Ecco il codice corrispondente al terzo esempio:

```
<table width="100%" border="2" cellspacing="0" cellpadding="3"
bordercolor="#000000">
  <tr>
    <td width="1%"></td>
    <td width="33%">Contenuto 1</td>
    <td width="1%"></td>
    <td width="33%">Contenuto 2</td>
    <td width="1%"></td>
    <td width="33%">Contenuto 3</td>
    <td width="1%"></td>
  </tr>
</table>
```

Come si vede la somma totale delle celle è di 103% e le proporzioni non rispettano la realtà, ma il *layout* viene visualizzato correttamente, perché la larghezza del 100% è espressa nel tag `<TABLE>`, che controlla le celle sottostanti. Proprio per questo stesso motivo (che cioè le percentuali non rispecchiano più la realtà), non ci sarà da stupirsi se per avere tre celle della stessa misura e che si ridimensionano allo stesso modo, dovremo dare, ad esempio, alla prima il 60% alla seconda il 40% e alla terza il 30%.