

## Esercizio 1

Si consideri il cluster di Server Web in figura 1. Il sistema è composto da un Web Switch che riceve le richieste dagli utenti e le distribuisce ad un set di N Server Web. Il Web Switch è connesso alla rete internet mediante il link L1. I Web Server e il Web Switch sono tra di loro connessi mediante la LAN L2.

Ogni server è caratterizzato da una CPU, da uno o più dischi ed una scheda di rete (lw2). La CPU dei server ha un tasso di servizio pari a 150 richieste al secondo (indipendentemente dal tipo di richieste). Il disco ha le seguenti caratteristiche: Seek Time 8.5ms, Controller Time 0.2ms, Rotational Speed 7200 RPM, transfer rate 100MB/sec, block size 2048 byte.

Il Web Switch è caratterizzato da una CPU e da due scheda di rete (ls1 e ls2), una che lo connette alla rete L1 ed una che lo connette alla rete L2.

La rete L2 è una rete Ethernet 100Mbit/sec e sia i Web Server che il Web Switch sono ad essa connessi mediante delle schede di rete con un transfer rate di 100Mbit/sec.

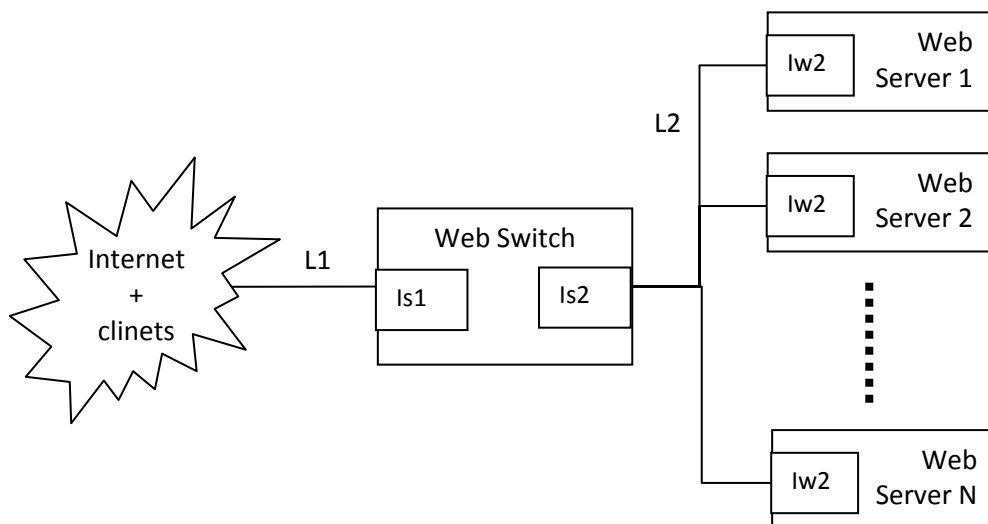


Figura 1 Schema del sistema

Si suppone che il carico del sistema è caratterizzato come segue

<b>Richieste per sessione</b>	Gaussiana inversa	$x > 0, \mu = 3.86, \lambda = 9.46$
<b>User Think Time</b>	Pareto	$x \geq k, \alpha = 1.4, k = 1$
<b>Numero di oggetti per richiesta</b>	Pareto	$x \geq k, \alpha = 1.33, k = 2$
<b>Dimensione pagina HTML</b>	Lognormale	$x > 0, \mu = 7.63, \sigma = 9.46$
	Pareto	$x \geq k, \alpha = 1, k = 10240 \text{ (byte)}$
<b>Embedded Objects</b>	Lognormale	$x > 0, \mu = 8.215, \sigma = 1.46$

Si chiede agli studenti di studiare il comportamento del sistema utilizzando un modello di simulazione e, fatta l'opportuna caratterizzazione del carico, mediante la risoluzione di un modello analitico.

Nello specifico si chiede di:

- Dimensionare opportunamente il sistema per far si che sia in grado di sopportare un carico di 300 utenti al secondo con un'utilizzazione massima della risorsa collo di bottiglia compresa tra il 65 ed il 70%.
- Studiare le prestazioni del sistema al variare delle politiche di selezione dei server (ROUND ROBIN, RANDOM, LEAST LOADED)
- Valutare che impatto ha sulle prestazioni del sistema l'introduzione di
  - o una scheda di rete aggiuntiva per i web server (per la gestione separata delle richieste in ingresso ed in uscita)
  - o un link in uscita, dimensionato opportunamente, per lo smistamento delle risposte (Vedi figura 2)
- Valutare che impatto ha, sulle prestazioni del sistema, l'introduzione di un proxy server che offra un cache hit rate del 40%.

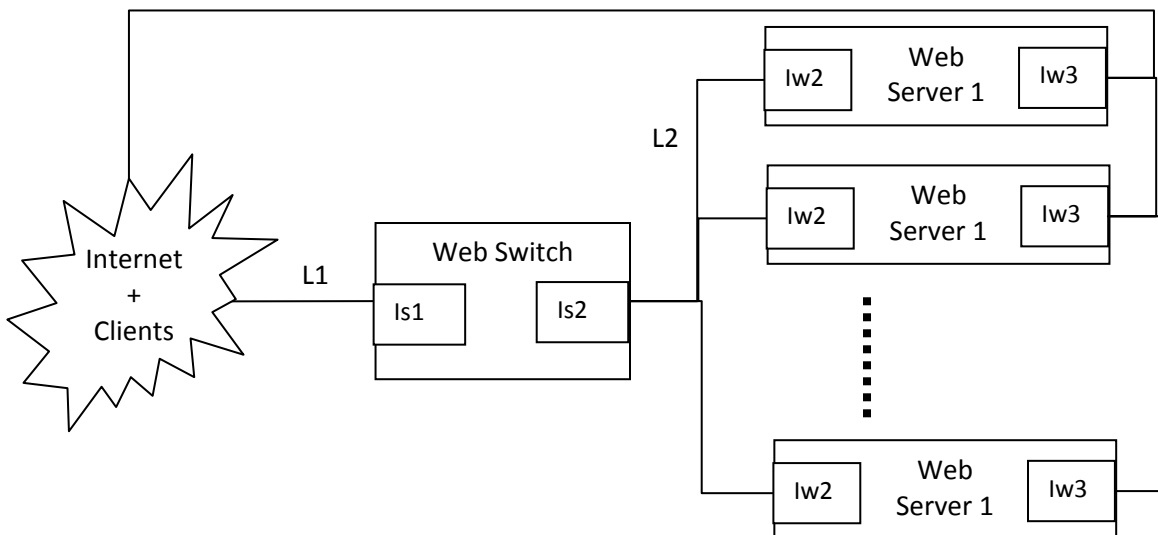


Figura 2: Schema del sistema dopo aver introdotto un link di uscita aggiuntivo.