

Contents

1	Introduzione	1
1.1	Importanza del performance modelling . . .	2
1.2	Valutazione delle performance	3
1.2.1	Obiettivi della valutazione delle performance	4
1.3	Modellazione	5
1.3.1	Tecniche di performance	6

1 Introduzione

Terminologia:

- Sistema: insieme di risorse hardware e software
- Metriche: criteri per confrontare le prestazioni del sistema, qualsiasi esso sia. Ad esempio:
 - il tempo di risposta
 - throughput: "produttività" del sistema per unità di tempo, in base a cosa il sistema produce (es richieste per unità di tempo)
- Workload: richieste che gli utenti sottomettono per un servizio che il sistema fornisce. Esempi:
 - istruzioni CPU che un certo server può ricevere
 - query al DB
- Tecnica: misure, simulazioni e modelli analitici

1.1 Importanza del performance modelling

Nonostante viviamo e sperimentiamo la sua importanza quotidianamente, ha una sensibilità poco importante. Ultimi anni:

1. Google down: 14/12/2020, problema per cui per circa 2h e 15 min tutti i servizi sono andati giù su scala mondiale. Didattica a distanza e smart working bloccati.
Problema è stato il blocco di qualsiasi servizio per l'accesso tramite autenticazione e quindi tutte le applicazioni coinvolte. La capacità ridotta del sistema centrale di gestione delle identità e di autenticazione di Google
2. Cashbak IO Pagopa: 7-10/12/2020. Milioni di download e di accessi, fino a 14000/s ed un autenticazione molto lenta, le troppe richieste hanno saturato le porte disponibili per l'accodamento delle richieste.
Blocco nell'inserimento dei metodi di pagamento, con annesso crollo dei servizi di push (che mette in coda le richieste che arrivano), dovuto alla lentezza dell'autenticazione.
Bottleneck nell'autenticazione e problema nella gestione non appropriata delle richieste
3. Signal: 16/01/2021, aumento improvviso dei download del circa 4200% in una settimana. Primo rallentamento del servizio, seguito da una parziale interruzione. La soluzione è stata di creare una replica del back-end su altri server.

Moltissimi altri casi di questo tipo, oggi i sistemi hanno

un livello di complessità molto alta come anche la loro composizione che è più complessa, la loro evoluzione è sempre più rapida. Inoltre, sono sempre più essenziali per il business e questo richiede la necessità di strumenti e tecniche che assistano progettisti, service provider etc... che permettano di capire a pieno il comportamento dei sistemi, in tutte le fasi:

- Progetto ed implementazione
- Dimensionamento
- Vita ed evoluzione del sistema

Tutte le tecniche per la valutazione delle prestazioni consentono anche una visione ed una conoscenza del sistema che il sistema stesso non offre: studiare il comportamento mediante un modello consente di vedere aspetti che non avremo potuto vedere in altro modo.

Tutte le figure che hanno a che fare con sistemi di questo tipo devono avere un background di tecniche di valutazione delle prestazioni.

1.2 Valutazione delle performance

Non è importante solo a livello industriale, ma anche nella ricerca accademica, nel momento in cui si vuole valutare una nuova proposta. Anche in questo caso la modellazione può essere molto utile. Nell'industria è essenziale per mantenere dei livelli di qualità alti, espressi negli SLA: il service provider deve poter identificare in maniera rapida dove c'è un problema che potrebbe portare al crollo delle performance garantite (per non andare in

contro a penali) e di fare un tuning del sistema per ritornare al livello di qualità che doveva essere garantito. Un buon modello di valutazione delle prestazioni ci dà una conoscenza profonda del comportamento del sistema: perché il sistema si comporta in un certo modo e perché lo fa, quali sono i limiti di quel comportamento ed i punti critici nel caso ci fossero problemi e fosse necessario migliorare le performance.

1.2.1 Obiettivi della valutazione delle performance

Alcuni esempi e obiettivi per un sistema:

- Capacity planning: determinare il numero e la taglia dei componenti del sistema
- Tuning del sistema: messa a punto del sistema, quando c'è qualcosa che si evidenzia nel comportamento del sistema che porta al degrado delle performance, bisogna determinare l'ottimo per il valore del parametro
- Bottleneck identification: determinare le performance di un bottleneck, identificarlo per poter capire qual'è la risorsa che saturerà per prima.
- Caratterizzazione del carico: solitamente ben modellate con distribuzioni esponenziali (o al più a fase), quindi che hanno la maggior parte della loro probabilità distribuite su valori piccoli, ovvero con tempi piccoli. Le probabilità di valori grandi sono piuttosto basse. Heavy-tail, la coda della distribuzione che modella i valori grandi non è così trascurabile.
- Previsione del carico (forecasting): tende ad oscillare, può essere critica

Quale può essere l'approccio: cominciare con una visione completa del sistema, degli obiettivi dello studio e dell'applicazione. Il punto di partenza è questo, anche se poi le tecniche non devono essere condizionate da tale visione.

Una volta fatto ciò, ci sono molti approcci ed hanno due casi limite:

- Intuizione ed estrapolazione delle tendenze: richiede alto grado di esperienza e di capacità intuitiva. Rapido e flessibile, ma l'accuratezza dei risultati a cui si arriva va guardato con un certo sospetto, in quanto sono derivanti dall'esperienza
- Valutazione sperimentale delle alternative: sempre possibile, ma costosa in quanto non è detto che sia generalizzata. Un esperimento porta alla conoscenza accurata del sistema sotto determinate assunzioni. Accuratezza eccellente, ma la valutazione sperimentale è molto più complessa e poco flessibile.

Entrambe le strade hanno vantaggi e svantaggi, tra questi due approcci estremi si colloca la modellazione, che prende lati positivi di uno e dell'altro ed anche i limiti.

1.3 Modellazione

Un modello è un'astrazione del sistema, la modellazione è il tentativo di distillare dalla quantità enorme di dettagli di quel sistema esattamente quegli aspetti e non di più che sono essenziali al comportamento del sistema rispetto agli obiettivi posti.

Il modello va definito e per fare questo occorrono:

- capacità di astrazione
- parametrizzazione del modello
- valutazione

Rispetto ai pro e contro dei due approcci, la modellazione è:

- più affidabile dell'approccio intuitivo
- meno costoso dell'approccio sperimentale

1.3.1 Tecniche di performance

Le tecniche di performance sono tecniche matematiche e computazionali per analizzare le performance del sistema stocastico (non ha un comportamento puramente deterministico)

Modellare: racchiude tutto il framework concettuale che descrive il sistema, la soluzione di divide in:

- tecniche analitiche
- tecniche simulative: eseguire esperimenti usando l'implementazione del modello
- misure: nel processo di monitoraggio di un sistema si collezionano una quantità di misure che sono molto importanti per capire il comportamento del sistema.

Tutto ciò che verrà derivato dalle due tecniche (simulative e analitiche) andrà verificato.