UNIVERSITÀ DEGLI STUDI DI
CASSINO E DEL LAZIO MERIDIONALE

DIPARTIMENTO DI INGEGNERIA ELETTRICA
E DELL'INFORMAZIONE "MAURIZIO SCARANO"

CORSO DI LAUREA IN INGEGNERIA INFORMATICA
E DELLE TELECOMUNICAZIONI

Assessment delle performance di Elixir nell'ambito IOT.

Relatore:
Prof. Ciro D'Elia

Candidato:
Nico Fiorini

ANNO ACCADEMICO 2022/2023

Questa è una dedica

La perfezione non è il nostro obbiettivo ma la nostra tendenza
Omar Palermo

Abstract

L' industria del software si trova a fronteggiare la necessità di sviluppare software sempre più scalabili e performanti per fronteggiare l'aumento degli utenti e di servizi che ne fanno utilizzo. In questo contesto, Elixir, un linguaggio di programmazione funzionale e concorrente basato su Erlang, emerge come una scelta promettente per la costruzione di sistemi altamente affidabili e reattivi, semplificando di molto lo sviluppo di software concorrenziale.

Questo studio si propone di analizzare le caratteristiche di Elixir e le sue performance attraverso una serie di esperimenti empirici esplorando diversi aspetti delle performance mettendo in rilievo vantaggi e svantaggi nell'adottarlo.

I risultati di questa ricerca forniranno una comprensione approfondità delle capacità di Elixir in termini di prestazioni e affidabilità consentendo agli sviluppatori di fare una scelta pensata alle esigenze dei loro progetti.

Indice

Capitolo 1

Introduzione

Elixir è un linguaggio di programmazione dinamico e funzionale sviluppato nel 2012 da José Valim, con l'obbiettivo di favorire una maggiore scalabilità e produttività nella macchinia virtuale di Erlang, mantenendo al contempo la compatibilità con l'ecosistema di Erlang[1]. Elixir si è affermato come una promettente scelta nell'industria del software,specialmente in contesti dove è richiesta scalabilità, tolleranza agli errori e reattività grazie al suo approccio concorrenziale.

In particolare, Elixir può risultare vantaggioso nel campo dell'IoT per diversi motivi:

1. **Concorrenza**: Nell'ambito dell'IoT, la gestione simultanea di dispositivi è essenziale. Elixir, grazie alla sua capacità di gestire facilmente la concorrenza, consente il monitoraggio e il controllo efficiente di numerosi dispositivi contemporaneamente.

2. **Fault Tolerance**: Data la natura degli ambienti IoT, dove i dispositivi possono guastarsi improvvisamente, Elixir offre strumenti per la supervisione e la gestione degli errori, garantendo la continuità delle operazioni anche in caso di fallimenti.

3. **Sviluppo Rapido e Manutenzione**: Elixir è un linguaggio moderno che offre una sintassi efficiente e snella, oltre a strumenti di sviluppo come Mix per la gestione delle dipendenze e l'ambiente interattivo iex. La presenza di un package manager (Hex)[**Hex63:online**] e la possibilità di generare automaticamente la documentazione facilitano il processo di sviluppo e manutenzione del codice.

Il trattato esplora Elixir concentrandosi su due aspetti principali: la semplicità e le performance. Si analizzano i punti di forza di un linguaggio funzionale

e come questi sono sfruttati in Elixir, con un focus sulla concorrenza. Nella scelta di un linguaggio, la semplicità è fondamentale e deve essere accessibile a tutti i programmatori. Tuttavia, l'efficienza è altrettanto importante, quindi vengono condotti test empirici per valutare le performance di Elixir.

In particolare il lavoro effettuato è così ripartito:

- Nel capitolo 2 si discute del linguaggio funzionale, esaminando le astrazioni offerte da Elixir per lo svilluppo di codice affidabile, si tratta la concorrenza e come la Erlang VM si occupa della gestione dei processi.
- Nel capitolo 3 si spiega il lavoro sperimentale svolto e i risultati ottenuti (continuare)

Capitolo 2

Caratteristiche di Elixir

2.1 Introduzione

In questo capitolo, esamineremo le caratteristiche distintive di Elixir, un linguaggio di programmazione funzionale e concorrente che sfrutta appieno la potenza della piattaforma OTP (Open Telecom Platform), non si vuole coprire ogni dettaglio del linguaggio, ma mettere in evidenza le caratteristiche fondamentali per iniziare a capire come pensare il codice con questo linguaggio.

Elixir, scritto in Erlang ed eseguito sulla macchina virtuale Erlang (BEAM), eredita gli obiettivi di Erlang, ma apporta miglioramenti significativi per rendere il linguaggio più appetibile e moderno.

Erlang, nato nel 1986, è stato progettato per semplificare lo sviluppo di software concorrente e robusto. Elixir si basa su queste fondamenta solide, offrendo un'API più pulita e astrazioni avanzate che consentono ai programmatori di ragionare a un livello più elevato, facilitando la scrittura di codice concorrente in modo intuitivo.

Una delle massime principali di Erlang e, di conseguenza, di Elixir, è "Let it crash" (Lascia che si schianti), che riflette l'approccio alla gestione degli errori nei sistemi concorrenti, incoraggiando la gestione degli errori tramite il rilancio e la supervisione anziché il blocco del processo.

Per capire come lavorare con questo linguaggio, bisogna affrontare un po' di questioni e farsi un po' di domande. Bisogna capire come la macchina virtuale Beam affronta la concorrenza, Elixir in particolare è un linguaggio orientato alla concorrenza e le astrazioni che fornisce sono proprio per far sì che si programmi in modo concorrenziale portando ad avere un codice responsivo e gestendo bene i processi attraverso il meccanismo di Supervision, il software diventa anche robusto. Un altro punto da affrontare è l'immutabilità dei dati,

è un concetto chiave in Elixir ed Erlang, è proprio questa caratteristica che ci semplifica la programmazione concorrenziale.

2.2 Il paradigma funzionale

Come già accennato Elixir è un linguaggio funzionale, dove il concetto di funzione ricopre il ruolo di protagonista, i dati sono immutabili e il codice è dichiarativo.

Questo modo di vedere le cose deriva dal Lambda calcolo o λ-calcolo [5] un sistema formale definito da Alonzo Church nel 1936, sviluppato per definire formalmente le funzioni e il loro calcolo.

In un paradigma basato su stati come la programmazione ad oggetti spesso si hanno variabili condivise mutabili, ovvero, più parti del codice possono riferirsi alla stessa variabile, e questo complica la programmazione multithreading dovendosi preoccupare di meccanismi come il blocco sincronizzato o il locking per evitare le race condition tra più parti del codice, e non è immediato scrivere del codice concorrenziale sicuro e spesso si riscontrano comportamenti indeterminati. In un paradigma funzionale si predilogono le variabili immutabili che aggirano questo problema riducendo il rischio di scrivere codice concorrenziale non sicuro.

Cambiare paradigma non è immediato, un paradigma si può dire che definisce il modo di pensare al problema, nella programmazione ad oggetti per esempio si definiscono le cosidette classi, pensando al problema come oggetti che possono comportarsi in un determinato modo attraverso le funzioni definite su di esso. Perciò si pensa ad un oggetto che ha un comportamento e che cambia il suo stato nel tempo, un modo di sviluppare intuitivo ma non sempre ottimale per la risoluzione di problemi. Nella programmazione funzionale si cambia prospettiva, ovvero si ha un input, si passa l'input alla funzione e si ottiene la trasformazione dell'input ottenendo l'output.

In poche parole un linguaggio funzionale assume che scrivere un software complesso sia più facile nel momento in cui il codice ha queste proprietà:

- I dati sono immutabili
- Le funzioni sono pure, ovvero, il risultato di una funzione dipende soltanto dai suoi parametri in input.
- Le funzioni non generano effetti oltre il suo valore restituito.

Con queste proprietà si ha più controllo del flusso del programma, anche se non sempre possono essere soddisfatte.

2.2.1 Struttura di un progetto Elixir

Elixir è un linguaggio moderno, e come ogni linguaggio moderno che si rispetti fornisce un tool per la creazione e configurazione di progetti, questo tool si chiama **Mix**.

Il tool Mix

È possibile creare un progetto con il comando:

```
mix new <nome-progetto>
```

Verrà creata una struttura per il progetto come nell'esempio 2.1, il codice sarà organizzato nella cartella **lib**, viene creata una cartella **test**, e il file per la configurazione del progetto **mix.exs**. Da notare che anche la configurazione del progetto avviene tramite funzioni.

```
.
>build
>deps
  -->...
>lib
  -->example.ex
mix.exs
README.md
>test
  -->example.exs
```

Esempio 2.1: Struttura progetto

Come sappiamo Elixir fornisce anche un ambiente interattivo (**iex**) per testare il nostro codice, ed è consentito avviare questo ambiente nel dominio della nostra applicazione con il comando:

```
iex -S mix
```

Si può compilare il progetto con:

```
mix compile
```

Con Mix possiamo includere e scaricare facilmente anche librerie esterne attraverso il package manager.[3]

Moduli

Elixir organizza il codice in Moduli, permettendo di definire le funzioni dentro dei namespace, così da separare le responsabilità delle funzioni.

Ci sono varie cose che si possono definire dentro un modulo, si possono definire delle **struct** ma cosa più importante si possono definire i cosiddetti **Behaviour**, un modo per definire un interfaccia Api, Elixir fornisce delle

astrazioni proprio attraverso questi. Ciò che si vuole evidenziare ora è che il progetto è definito in moduli, il modo che Elixir fornisce per organizzare il codice.

2.2.2 Basi dichiarative

Come già accennato, Elixir adotta un approccio dichiarativo nella definizione delle funzioni. Questo si contrappone all'approccio imperativo, che si concentra su "come posso risolvere questo problema?", mentre quello dichiarativo si pone la domanda "come posso definire un problema?".

Nell'esempio 2.2 è presentato un approccio imperativo al problema "somma dei primi n elementi", mentre nell'esempio 2.3 è presentato l' approccio dichiarativo con Elixir.

```
int sum_first_n(n){
  int sum=0;
    for(int i=1;i++;i<=n){
      sum+=i;
    }
    return sum;
}
```

Esempio 2.2: Somma N elementi

```
defmodule Sum do
  def sum_recursive(0), do: 0
  def sum_recursive(n), do: n + sum_recursive(n - 1)
end
```

Esempio 2.3: Somma N elementi

In particolare questo approccio è rafforzato tramite il meccanismo del **Pattern Matching**, infatti in Elixir l'operatore =, non è un operatore di assegnazione, ma è comparabile all'equivalente algebrico. Quest'operatore ci permette di scrivere dell'equazioni che condizionano il flusso del codice. Questo è molto utile per l'approccio dichiarativo, infatti possiamo definire più corpi della stessa funzione ed Elixir capisce quale funzione invocare in base al valore dei suoi parametri attraverso questo meccanismo, è già stato utilizzato implicitamente nell'esempio 2.3, dove viene invocata la funzione "sum_recursive(0)" quando il valore dell'argomento vale 0, altrimenti chiamerà quello con l'argomento "n" assegnando il valore dell'argomento alla variabile n.

Elixir cerca di assegnare il valore a destra dell'equazione al valore di sinistra cercando di risolvere l'equazione tentando di fare un assegnazione dove possibile. Questo ci permette di usare quest'operatore per poter scomporre un dato, un esempio può essere usato usando l'ambiente interattivo *iex*

```
# Lists
iex> list = [1, 2, 3]
[1, 2, 3]
iex> [1, 2, 3] = list
[1, 2, 3]
iex> [] = list
** (MatchError) no match of right hand side value: [1, 2, 3]
iex> [1 | tail] = list
[1,2,3]
iex> tail
[2,3]
```

Esempio 2.4: Pattern Matching

2.3 Concorrenza

La concorrenza è un concetto fondamentale nell'ambito dello sviluppo software, si riferisce alla capacità di eseguire più attività contemporaneamente. Questo è particolarmente importante in applicazione che devono gestire molteplici operazioni in parallelo come ad esempio un server web, applicazioni di elaborazione dati.

La concorrenza consente a un programma di sfruttare appieno le risorse disponibili, aumentando l'efficienza e migliorando le prestazioni complessive. Inoltre, permette di creare sistemi più reattivi e scalabili, in grado di gestire un numero crescente di richieste senza compromettere le prestazioni.

In generale, la gestione della concorrenza può essere complessa e soggetta a errori o comportamenti indeterminati, infatti, più processi o thread possono interagire tra loro in modo imprevedibile, essendo così costretti a usare tecniche di locking in concomitanza di memoria condivisa trà più thread.

Nel contesto di Elixir, la concorrenza è un concetto centrale e viene gestita attraverso il modello di programmazione basato su processi leggeri, tutti isolati tra loro con il proprio stack ed il proprio heap.

2.3.1 La concorrenza in Beam

La concorrenza gioca un ruolo chiave per un software che vuole essere altamente responsivo. La concorrenza fa uso dei cosiddetti processi leggeri nella piattaforma Erlang, non sono processi del sistema operativo, ma processi della VM, chiamato processo e non thread in quanto non condividono memoria tra di loro e sono completamente isolati.

Un server tipico deve gestire migliaia di richieste, e gestirle concorrenzialmente è essenziale per non far rimanere in attesa il client. Quello che si vuole è gestirli parallelamente il più possibile sfruttando più risorse della

Cpu disponibile. Quello che fa la macchina virtuale per noi è permetterci la scalabilità,più richieste allora più risorse da allocare.

Inoltre, siccome il processo è isolato, un errore in una richiesta può essere localizzato senza avere impatto sul resto del sistema, così creando anche un sistema robusto agli errori.

Un processo appena creato occupa in memoria 326 words [2], quindi in una macchina a 64 bit occupa 2608 byte. Lo si può vedere in Elixir facilmente nell'esempio

```
1  defmodule Examples.Memory do
2    def mypid do
3      receive do
4        :stop -> :exit
5        _ -> mypid()
6      end
7    end
8
9    def benchmark do
10     pid = spawn(fn -> mypid() end)
11     {_,byte_used} = :erlang.process_info(pid,:memory)
12     IO.puts("La memoria usata dal processo e': #{byte_used} byte")
13     send pid, :stop
14   end
15 end
```

Esempio 2.5: Memoria in un processo

Infatti l'output ottenuto dall'esempio 2.5

```
iex> Examples.Memory.benchmark
La memoria usata dal processo e': 2640 byte
:stop
```

Si nota dall'output dell'esempio che la memoria utilizzata è leggermente superiore della memoria minima dichiarata, infatti il processo non fa nulla di particolare oltre ad aspettare il messaggio di ":stop". È da notare che è proprio questa leggerezza nei processi che permette al linguaggio di essere orientato alla concorrenza, e poter usare i processi con più leggerezza rispetto ad altri meccanismi di altri linguaggi che usano i thread, Un altro punto da notare è che si può migliorare significativamente la reattività del programma ma non l'efficienza totale del sistema, infatti non tutti i processi sono eseguiti in parallelo, quindi in una macchina con quattro processori non si possono eseguire più di 4 processi per volta ed in un normale software Elixir è normale avere migliaia di processi che lavorano.

Facendo un esempio con una Cpu dual-core, i processi vengono eseguiti concorrenzialmente e gestiti dagli Scheduler della VM, in figura 2.1 è mostrato come la VM gestisce i processi di default.

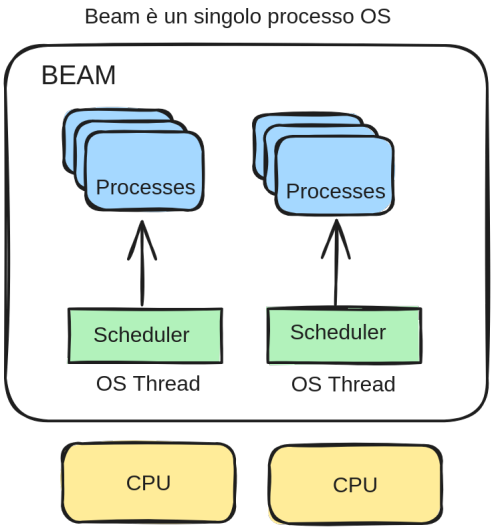

Figura 2.1: Concorrenza nella VM Beam [4]

2.3.2 Concorrenza basata su attori

Elixir usa un modello di concorrenza basato su attori, gli attori sono entità di elaborazione indipendenti che eseguono operazioni in modo asincrono, questi attori non sono altro che processi che vengono identificati attraverso un **PID** univoco. Come già detto sono isolati l'uno dall'altro e comunicano solo attraverso lo scambio di messaggi, questo scambio avviene attraverso dei canali di comunicazione detti **"mailbox"**. Ogni processo ha una propria mailbox dove avviene la ricezione del messaggio da parte di altri processi.

Conoscendo il PID di un processo può avvenire la comunicazione attraverso la primitiva fornita dal linguaggio **send/2** che permette di inviare un messaggio ad un processo come avviene nell'esempio 2.5, dove il processo principale crea un processo che rimane in ascolto tramite il blocco **receive/1**, fin quando non riceve il messaggio di **:stop** e il processo viente terminato.

Se non ci sono messaggi nella mailbox, il processo aspetta fino a quando non arriva un messaggio, in particolare, nell'esempio tutti i messaggi che non siano :stop, verranno ignorati continuando ad ascoltare altri messaggi.

Process Linking

I processi sono isolati, ma se creo un processo con spawn/1 spesso voglio sapere se va in errore e magari propagare l'errore, per questo in Elixir c'è il concetto di linking, ovvero si può legare un processo ad un altro, in modo

che se uno va in errore, i processi legati vengono notificati con un messaggio propagando l'errore ai processi legati.

Spesso i processi vengono legati ad un Supervisor, che lascerà andare in errore i processi figli, e si occuperà semplicemente di riavviarli seguendo una strategia scelta. Questa è la filosofia "Let it Crash" di Erlang, che si contrappone alla gestione delle eccezioni di altri linguaggi. Si veda l'esempio **??**

```
iex>spawn(fn -> raise "oops" end)
#PID<0.58.0>

[error] Process #PID<0.58.00> raised an exception
** (RuntimeError) oops
        (stdlib) erl_eval.erl:668: :erl_eval.do_apply/6
```

Esempio 2.6: Process linking

2.4 Supervision Tree

Capitolo 3

Performance - Test sperimentali

3.1 Introduzione

Capitolo 4

Conclusioni

Bibliografia

[1] *Elixir (programming language) - Wikipedia*. https://en.wikipedia.org/wiki/Elixir_(programming_language).

[2] *Erlang – Processes*. https://www.erlang.org/doc/efficiency_guide/processes.

[3] *HexDocs*. https://hexdocs.pm/.

[4] Saša Juric. «Elixir in Action». In: 2019. Cap. 5.

[5] *Lambda calcolo - Wikipedia*. https://it.wikipedia.org/wiki/Lambda_calcolo.