

Your Own Distributed System using Consensus Protocols

Bellagio Casinó roulette multiplayer

Stefano Agostini
agostinistefano1991@gmail.com

Salomé Paolo
paolosalome@gmail.com

Valenti Alessandro
alessandro.valenti1991@gmail.com

ABSTRACT

This paper provides a sample of a \LaTeX document which conforms, somewhat loosely, to the formatting guidelines for ACM SIG Proceedings. It is an *alternate* style which produces a *tighter-looking* paper and was designed in response to concerns expressed, by authors, over page-budgets. It complements the document *Author's (Alternate) Guide to Preparing ACM SIG Proceedings Using \LaTeX 2_ε and Bib \TeX* . This source file has been written with the intention of being compiled under \LaTeX 2_ε and Bib \TeX .

The developers have tried to include every imaginable sort of “bells and whistles”, such as a subtitle, footnotes on title, subtitle and authors, as well as in the text, and every optional component (e.g. Acknowledgments, Additional Authors, Appendices), not to mention examples of equations, theorems, tables and figures.

To make best use of this sample document, run it through \LaTeX and Bib \TeX , and compare this source code with the printed output produced by the dvi file. A compiled PDF version is available on the web page to help you with the ‘look and feel’.

1. INTRODUZIONE

Le tecnologie Cloud, esplose in questi ultimi anni, hanno permesso di trasferire il peso della computazione e dell’approvvigionamento delle risorse informatiche verso internet. Perciò la rete come strumento di sviluppo oggi offre scenari ampi nella direzione della distribuzione delle risorse e della computazione: la distribuzione, come nel caso che poniamo in esame, deve rappresentare un’avanguardia per lo sviluppo di applicazioni con caratteristiche di scalabilità e di alta disponibilità per l’utente finale. Gli sviluppatori utilizzando i servizi cloud possono gestire in maniera semplificata l’immagazzinamento di informazioni e l’elaborazione di molti dati, concentrandosi principalmente sullo sviluppo della loro applicazione piuttosto che nell’utilizzo di un sistema ortodosso e dispendioso come avveniva nel passato. Aziende come *Spotify*, *Netflix* ed altre sfruttano il servizio

cloud per rendere fruibile le loro applicazioni agli utenti di tutto il mondo, ottenendo molteplici consensi. Molte persone all’interno del mondo non sanno che queste applicazioni sono così efficienti proprio perché sfruttano una tecnologia cloud, che per quanto possa agevolare gli utenti e gli sviluppatori, presenta sempre qualche difficoltà da gestire:

1. garantire l’accesso concorrente ad una vasta molteplicità di utenti.
2. la possibilità di connettersi da ogni angolo del mondo, senza ledere la qualità del servizio.
3. la sincronizzazione degli eventi associati ad ogni applicativo.
4. la reperibilità dei dati.

1.1 Obiettivi

L’obiettivo preposto consiste nell’elaborare un gioco multiplayer che supporta molteplici entità autonome che si contendono risorse condivise, l’aggiornamento in tempo reale di una qualche forma di stato condiviso, essere distribuito su molteplici nodi (eventualmente distribuiti geograficamente) e supportare la scalabilità. E’ proprio con l’obiettivo di rispettare tali punti cardine che abbiamo proceduto nello sviluppo della nostra applicazione distribuita sulla piattaforma Cloud: abbiamo operato nell’intenzione di venire incontro alle esigenze degli utenti, sempre più abituati ad avere interazioni costanti e ripetute con le applicazioni di uso comune. Il *Bellagio Casinó* è l’applicazione (accessibile mediante interfaccia Web) da noi proposta che consente agli utenti di giocare ad una versione semplificata di una roulette francese condividendo un unico tavolo di gioco. Tale applicazione permette di effettuare particolari puntate in base alla propria disponibilità di credito.

2. ARCHITETTURA

In questa sezione verrà illustrata a livello logico la nostra architettura descrivendo i servizi *Amazon* utilizzati e i moduli sviluppati. I componenti architetturali sono:

- un Cluster composto da istanze *EC2*, situato nella regione Francoforte.
- due Cluster composti da istanze *EC2*, situate nella regione Eire.
- Client composto da una applicazione Web.

i servizi *AWS* sfruttati dai componenti sono:

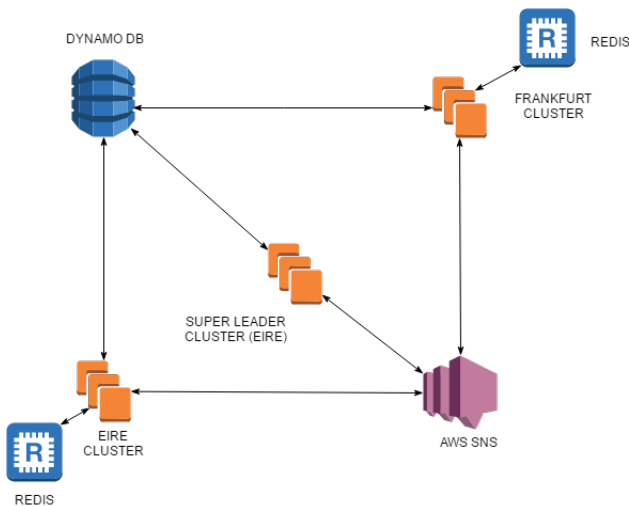


Figure 1: Architettura

- *SNS* per la comunicazione inter-cluster.
- *DynamoDB* per la persistenza dei dati situato nella regione Eire.
- *Redis* per supportare la logica del gioco situato sia nella regione Eire che Francoforte.
- *LoadBalancer* per distribuire il carico.

La motivazione che ci ha spinto nell'adottare più cluster su più aree geografiche europee è stata la possibilità di realizzare una scalabilità di tipo geografico. Ciò non toglie che tale realizzazione si possa estendere anche all'interno di altre regioni geografiche messe a disposizione da AWS. La motivazione che ci ha spinto a questa scelta è legata a fattori di latenza ma anche economici, facilitando lo sviluppo dell'applicazione e la fase di testing.

2.1 Cluster EC2

All'interno di questa sezione andiamo a trattare nel dettaglio quali sono le funzionalità e le componenti dei Cluster EC2.

I Cluster istanziati sono soggetti ad una gerarchia dettata dalla realizzazione di uno stato condivisibile, che corrisponde alle fasi di gioco dell'applicazione. Pertanto possiamo suddividere tali cluster in base alla loro mansione :

- *SuperLeader* Cluster: identifica il Cluster per la gestione dello stato di gioco, e per le azioni di registrazione e di accesso sito in Irlanda.
- *Frankfurt/Eire* Cluster: identifica uno dei Cluster di gioco sito sia in Irlanda che a Francoforte.

All'interno di ogni Cluster per mantenere lo stato corrente del gioco, abbiamo adottato *Raft* come protocollo del consenso. Abbiamo utilizzato 3 istanze EC2 per Cluster, in quanto è il numero minimo di nodi necessario per raggiungere il consenso mediante tale protocollo.

3. IMPLEMENTAZIONE

3.1 Metodologie di sviluppo

L'intero progetto è stato realizzato seguendo la metodologia di sviluppo *Scrum*, la quale prevede di dividere il progetto in blocchi rapidi di lavoro (Sprint), alla fine di ciascuno dei quali si deve creare un incremento del software. L'intera stesura del progetto, dalla progettazione all'implementazione, è stata quindi suddivisa in sprint con scadenza molto stretta assegnati ai singoli componenti del team. Per la suddivisione dei task sono stati effettuati incontri giornalieri, mentre per la sincronizzazione e per il controllo di versione è stato utilizzato il servizio di *Google Drive*. Il periodo di sviluppo, è stato caratterizzato da incontri giornalieri (Daily Scrum) incentrati sulla tematica del giorno, con la presenza del team al completo. A supporto dello sviluppo è stato previsto, fin dal primo istante, un ulteriore incontro prima della partenza di ogni nuovo sprint. La metodologia utilizzata si è dimostrata efficiente poiché ha permesso di progettare e realizzare un'architettura ben congegnata in un arco temporale ridotto.

3.2 Tecnologie utilizzate

Il progetto è stato interamente realizzato utilizzando il framework *Node.js*, che consente di sfruttare il linguaggio *JavaScript*, tipicamente utilizzato nella client-side, anche per la scrittura di applicazioni server-side. La caratteristica principale di *Node.js* risiede nella possibilità di accedere alle risorse del sistema operativo in modalità event-driven evitando il modello basato su processi o thread concorrenti utilizzato dai classici web server. Per facilitare l'uso di *Node.js* è stato utilizzato il framework *Express.js*, per la realizzazione semplificata del middleware http. L'applicazione client è costituita da una singola pagina *HTML* con contenuti *Javascript* ed elaborata attraverso l'uso della libreria *JQuery* rendendo l'interfaccia dinamica e consistente e di facile utilizzo.

3.3 Raft

Raft è un protocollo del consenso in alternativa a *Paxos* che offre la possibilità di condividere uno stato comune tra i vari nodi del Cluster. Questi si suddividono in leader, candidate e follower. Viene raggiunto il consenso attraverso l'elezione di un leader, la cui mansione principale è quella di propagare il log dei cambiamenti di stato ai propri follower. Inoltre per ribadire la propria leadership deve informare periodicamente i propri follower attraverso dei brevi messaggi. Ogni follower deve ricevere questi messaggi all'interno di un arco temporale fissato. Se ciò non dovesse avvenire verrà richiesta una nuova elezione di leader; colui che indice la nuova elezione cambia il proprio ruolo in candidate. Per far sì che *Raft* tolleri K guasti occorrono almeno $2K + 1$ nodi all'interno del Cluster.

Per l'implementazione di tale protocollo è stata adottata la libreria *Skiff*. Implementa la persistenza dello stato attraverso un database *in-memory* di tipo *MemDown* il quale utilizza un'interfaccia *LevelUp*. Per quanto concerne la comunicazione del log e dell'heartbeat utilizza server *TCP*, in ascolto su ogni nodo, attraverso chiamate di tipo *RPC*.

Le motivazioni che ci hanno condotto a questa scelta sono da addurre alla semplicità di esecuzione e di condivisione del log tra i vari nodi del Cluster, la cui propagazione avviene in tempi irrisori. L'utilizzo di *Skiff* con l'architettura dei Cluster da noi ideata consente di raggiungere una certa tolleranza a guasti di vario genere sulle varie istanze EC2. In caso di un guasto ad una macchina, se questa risulta essere un leader quello che avviene è una nuova elezione, con conseguente

riallineamento dei nodi del Cluster allo stato corrente del gioco. Nel caso di un guasto nei confronti di un follower bisogna monitorare il numero di nodi presenti nel Cluster, come precedentemente illustrato.

3.4 Comunicazione

Per quanto riguarda la comunicazione, vogliamo porre l'attenzione verso due tipologie presenti all'interno del nostro elaborato:

- Comunicazione *intra-Cluster*: sfruttiamo le chiamate *RPC*, come descritto nel paragrafo precedente.
- Comunicazione *inter-Cluster*: Sfruttiamo il servizio *AWS* di *SNS*.

Amazon SNS é un servizio di notifiche push rapido, flessibile e completamente gestito che ti consente di inviare messaggi individuali o collettivi a un numero elevato di destinatari. Grazie a questa soluzione, inviare notifiche push a utenti di dispositivi mobili, destinatari di posta o addirittura altri servizi distribuiti é semplice e conveniente. Per comunicare lo stato del gioco a tutti nodi dei vari Cluster sono stati creati 5 *Topic* per permettere lo scambio di informazioni e di cambiamento di stato, ciò verrà chiarificato successivamente.

La motivazione per cui é stato scelto questo servizio sono le seguenti:

- *SNS* permette la realizzazione di un servizio di tipo *Publish/Subscribe*.
- Sfrutta endpoint di tipo *http* per la ricezione di notifiche.
- Permette la comunicazione tra piú aree geografiche.
- Garantisce un certo grado di riservatezza.

Generalmente *SNS* viene utilizzato in abbinamento con *SQS*, però nel nostro caso non avevamo una necessità di permanenza e di ordinamento dei messaggi.

3.5 Persistenza

Poiché la nostra applicazione é un gioco multiplayer e dovendo gestire la registrazione e l'accesso di piú giocatori necessitiamo di un servizio di persistenza dati. In particolare la persistenza é stata realizzata sfruttando due tipi di servizio:

- *Redis*: é un servizio di cache distribuito di *Amazon AWS*.
- *DynamoDB*: é un database *NoSQL* distribuito di *Amazon AWS*.

Nel dettaglio *Redis* viene sfruttato dai Cluster di gioco per poter immagazzinare le ultime informazioni relative alle puntate degli utenti, nell'ultima mano di gioco, rendendo tali dati accessibili per l'elaborazione delle vincite da parte del sistema. Alla fine di questa operazione il sistema prevederà all'eliminazione di tali dati per permettere l'aggiornamento dei dati coerentemente con la mano del gioco.

Per quanto concerne *DynamoDB* sono state create due tabelle

- una per conservare i dati relativi agli utenti (credenziali di accesso e credito).

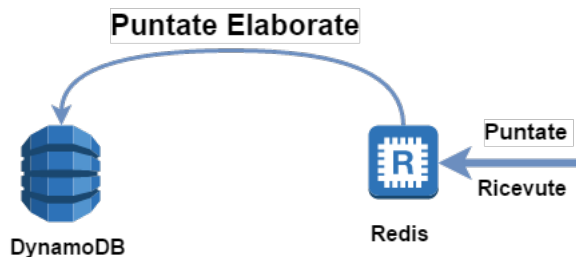


Figure 2: Processo scrittura delle puntate

- una per conservare le puntate elaborate dal sistema, precedentemente contrassegnate come vincenti o perdenti.

Quando un'applicazione effettua delle scritture su una tabella in *DynamoDB* i dati impiegano del tempo per propagarsi in ogni locazione di memoria della regione *AWS* corrente. I dati saranno consistenti in tutte le locazioni di memoria in un periodo lungo ~ 1 sec o anche meno. Dato che la nostra applicazione non effettua chiamate al Database frequentemente, la scelta dell'*eventual consistency* ci é sembrata la piú opportuna anche in termini di costi.

3.6 Type Changes and Special Characters

We have already seen several typeface changes in this sample. You can indicate italicized words or phrases in your text with the command `\textit`; boldening with the command `\textbf` and typewriter-style (for instance, for computer code) with `\texttt`. But remember, you do not have to indicate typestyle changes when such changes are part of the *structural* elements of your article; for instance, the heading of this subsection will be in a sans serif¹ typeface, but that is handled by the document class file. Take care with the use of² the curly braces in typeface changes; they mark the beginning and end of the text that is to be in the different typeface.

You can use whatever symbols, accented characters, or non-English characters you need anywhere in your document; you can find a complete list of what is available in the *L^AT_EX User's Guide*[?].

3.7 Math Equations

You may want to display math equations in three distinct styles: inline, numbered or non-numbered display. Each of the three are discussed in the next sections.

3.7.1 Inline (In-text) Equations

A formula that appears in the running text is called an inline or in-text formula. It is produced by the `\math` environment, which can be invoked with the usual `\begin. . .\end` construction or with the short form `\$. . .\$`. You can use any of the symbols and structures, from α to ω , available in L^AT_EX[?]; this section will simply show a few examples of in-text equations in context. Notice how this equation: $\lim_{n \rightarrow \infty} x = 0$, set here in in-line math style, looks slightly different when set in display style. (See next section).

¹A third footnote, here. Let's make this a rather short one to see how it looks.

²A fourth, and last, footnote.

3.7.2 Display Equations

A numbered display equation – one set off by vertical space from the text and centered horizontally – is produced by the **equation** environment. An unnumbered display equation is produced by the **displaymath** environment.

Again, in either environment, you can use any of the symbols and structures available in L^AT_EX; this section will just give a couple of examples of display equations in context. First, consider the equation, shown as an inline equation above:

$$\lim_{n \rightarrow \infty} x = 0 \quad (1)$$

Notice how it is formatted somewhat differently in the **displaymath** environment. Now, we'll enter an unnumbered equation:

$$\sum_{i=0}^{\infty} x + 1$$

and follow it with another numbered equation:

$$\sum_{i=0}^{\infty} x_i = \int_0^{\pi+2} f \quad (2)$$

just to demonstrate L^AT_EX's able handling of numbering.

3.8 Citations

Citations to articles [?, ?, ?, ?], conference proceedings [?] or books [?, ?] listed in the Bibliography section of your article will occur throughout the text of your article. You should use BibTeX to automatically produce this bibliography; you simply need to insert one of several citation commands with a key of the item cited in the proper location in the .tex file [?]. The key is a short reference you invent to uniquely identify each work; in this sample document, the key is the first author's surname and a word from the title. This identifying key is included with each item in the .bib file for your article.

The details of the construction of the .bib file are beyond the scope of this sample document, but more information can be found in the *Author's Guide*, and exhaustive details in the *L^AT_EX User's Guide*[?].

This article shows only the plainest form of the citation command, using \cite. This is what is stipulated in the SIGS style specifications. No other citation format is endorsed or supported.

3.9 Tables

Because tables cannot be split across pages, the best placement for them is typically the top of the page nearest their initial cite. To ensure this proper “floating” placement of tables, use the environment **table** to enclose the table's contents and the table caption. The contents of the table itself must go in the **tabular** environment, to be aligned properly in rows and columns, with the desired horizontal and vertical rules. Again, detailed instructions on **tabular** material is found in the *L^AT_EX User's Guide*.

Immediately following this sentence is the point at which Table 1 is included in the input file; compare the placement of the table here with the table in the printed dvi output of this document.

To set a wider table, which takes up the whole width of the page's live area, use the environment **table*** to enclose the table's contents and the table caption. As with a

Table 1: Frequency of Special Characters

Non-English or Math	Frequency	Comments
Ø	1 in 1,000	For Swedish names
π	1 in 5	Common in math
\$	4 in 5	Used in business
Ψ ₁ ²	1 in 40,000	Unexplained usage

Figure 3: A sample black and white graphic.

single-column table, this wide table will “float” to a location deemed more desirable. Immediately following this sentence is the point at which Table 2 is included in the input file; again, it is instructive to compare the placement of the table here with the table in the printed dvi output of this document.

3.10 Figures

Like tables, figures cannot be split across pages; the best placement for them is typically the top or the bottom of the page nearest their initial cite. To ensure this proper “floating” placement of figures, use the environment **figure** to enclose the figure and its caption.

This sample document contains examples of .eps files to be displayable with L^AT_EX. If you work with pdfL^AT_EX, use files in the .pdf format. Note that most modern T_EX system will convert .eps to .pdf for you on the fly. More details on each of these is found in the *Author's Guide*.

As was the case with tables, you may want a figure that spans two columns. To do this, and still to ensure proper “floating” placement of tables, use the environment **figure*** to enclose the figure and its caption. and don't forget to end the environment with figure*, not figure!

3.11 Theorem-like Constructs

Other common constructs that may occur in your article are the forms for logical constructs like theorems, axioms, corollaries and proofs. There are two forms, one produced by the command \newtheorem and the other by the command \newdef; perhaps the clearest and easiest way to distinguish them is to compare the two in the output of this sample document:

This uses the **theorem** environment, created by the \newtheorem command:

THEOREM 1. *Let f be continuous on $[a, b]$. If G is an antiderivative for f on $[a, b]$, then*

$$\int_a^b f(t)dt = G(b) - G(a).$$

The other uses the **definition** environment, created by the \newdef command:

Definition 1. *If z is irrational, then by e^z we mean the unique number which has logarithm z :*

$$\log e^z = z$$

Figure 4: A sample black and white graphic that has been resized with the includegraphics command.

Table 2: Some Typical Commands

Command	A Number	Comments
<code>\alignauthor</code>	100	Author alignment
<code>\numberofauthors</code>	200	Author enumeration
<code>\table</code>	300	For tables
<code>\table*</code>	400	For wider tables

Figure 5: A sample black and white graphic that needs to span two columns of text.

Figure 6: A sample black and white graphic that has been resized with the `includegraphics` command.

Two lists of constructs that use one of these forms is given in the *Author's Guidelines*.

There is one other similar construct environment, which is already set up for you; i.e. you must *not* use a `\newdef` command to create it: the **proof** environment. Here is an example of its use:

PROOF. Suppose on the contrary there exists a real number L such that

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = L.$$

Then

$$l = \lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} \left[g(x) \cdot \frac{f(x)}{g(x)} \right] = \lim_{x \rightarrow c} g(x) \cdot \lim_{x \rightarrow c} \frac{f(x)}{g(x)} = 0 \cdot L = 0,$$

which contradicts our assumption that $l \neq 0$. \square

Complete rules about using these environments and using the two different creation commands are in the *Author's Guide*; please consult it for more detailed instructions. If you need to use another construct, not listed therein, which you want to have the same formatting as the Theorem or the Definition[?] shown above, use the `\newtheorem` or the `\newdef` command, respectively, to create it.

A Caveat for the T_EX Expert

Because you have just been given permission to use the `\newdef` command to create a new form, you might think you can use T_EX's `\def` to create a new command: *Please refrain from doing this!* Remember that your L^AT_EX source code is primarily intended to create camera-ready copy, but may be converted to other forms – e.g. HTML. If you inadvertently omit some or all of the `\defs` recompilation will be, to say the least, problematic.

4. CONCLUSIONS

This paragraph will end the body of this sample document. Remember that you might still have Acknowledgments or Appendices; brief samples of these follow. There is still the Bibliography to deal with; and we will make a disclaimer about that here: with the exception of the reference to the L^AT_EX book, the citations in this paper are to articles which have nothing to do with the present subject and are used as examples only.

5. ACKNOWLEDGMENTS

This section is optional; it is a location for you to acknowledge grants, funding, editing assistance and what have you.

In the present case, for example, the authors would like to thank Gerald Murray of ACM for his help in codifying this *Author's Guide* and the `.cls` and `.tex` files that it describes.

APPENDIX

A. HEADINGS IN APPENDICES

The rules about hierarchical headings discussed above for the body of the article are different in the appendices. In the **appendix** environment, the command **section** is used to indicate the start of each Appendix, with alphabetic order designation (i.e. the first is A, the second B, etc.) and a title (if you include one). So, if you need hierarchical structure *within* an Appendix, start with **subsection** as the highest level. Here is an outline of the body of this document in Appendix-appropriate form:

A.1 Introduction

A.2 The Body of the Paper

A.2.1 Type Changes and Special Characters

A.2.2 Math Equations

Inline (In-text) Equations.

Display Equations.

A.2.3 Citations

A.2.4 Tables

A.2.5 Figures

A.2.6 Theorem-like Constructs

A Caveat for the T_EX Expert

A.3 Conclusions

A.4 Acknowledgments

A.5 Additional Authors

This section is inserted by L^AT_EX; you do not insert it. You just add the names and information in the `\additionalauthors` command at the start of the document.

A.6 References

Generated by bibtex from your `.bib` file. Run latex, then bibtex, then latex twice (to resolve references) to create the `.bbl` file. Insert that `.bbl` file into the `.tex` source file and comment out the command `\thebibliography`.

B. MORE HELP FOR THE HARDY

The sig-alternate.cls file itself is chock-full of succinct and helpful comments. If you consider yourself a moderately experienced to expert user of L^AT_EX, you may find reading it useful but please remember not to change it.