

Final Report

Arianna Soriani
arianna.soriani@studio.unibo.it

November 28, 2023

Il progetto nasce allo scopo di realizzare il materiale didattico da presentare in un'ipotetica lezione universitaria. Il topic di discussione è la "Replication".

Contents

1	Obiettivi del progetto	2
1.1	Argomenti trattati	2
2	Realizzazione	3
2.1	Introduzione	3
2.1.1	The 5Ws	3
2.1.2	Failures	3
2.1.3	Limit and costs	4
2.2	Cosa riguarda	4
2.2.1	Replication of Data	5
2.2.2	Replication of Processes	5
2.2.3	Replication of Objects	5
2.2.4	Replication of Messages	5
2.3	Heterogeneous, Mobile, and Large-Scale Systems	6
3	Conclusion	6

1 Obiettivi del progetto

L'obiettivo principale del progetto vede la realizzazione di un insieme di slide, atte a sostenere una lezione didattica della durata di 4/6 ore circa. Il tema di discussione riguarda la Replicazione in Ambienti Distribuiti.

1.1 Argomenti trattati

Le slide tratteranno diversi argomenti inerenti al mondo della replicazione, allo scopo di descriverne dettagliatamente diversi aspetti. In particolare, la presentazione tratterà i seguenti macro-argomenti:

- Perché nasce la replicazione
 - Replication for failure management
 - Replication for Availability
 - Reliability x Availability = Dependability
 - Replication for Performance
 - Benign Fault vs Malicious Fault
- Che cos'è la replicazione
- Costi e limiti
- Replication of Data
 - Consistency Models: Model of Distributed Database System
 - Replication Techniques: ROWA & QC
- Replication of Processes
- Replication of Objects
- Replication of Messages
- The Future of Replication

2 Realizzazione

La presentazione è strutturata nei seguenti punti:

- The 5Ws
- Failures
- Limits and Costs
- Replication of Data
- Replication of Processes
- Replication of Objects
- Replication of Messages
- Replication in Heterogeneous, Mobile, and Large-Scale Systems
- The future of Replication

2.1 Introduzione

I primi tre macro-argomenti introducono il tema della replicazione, con definizioni ed esempi.

2.1.1 The 5Ws

Attraverso la regola delle 5W viene descritta nel complesso la replicazione, rispondendo alle domande:

- Why
- Where
- What
- Who
- When

2.1.2 Failures

Il tema degli "Errors and Failures" è uno degli argomenti cardini della replicazione, che ne accompagna l'intera evoluzione.

Per questo motivo, in questa prima fase introduttiva, ne vengono descritti i principali aspetti. Questa sezione tratta i seguenti punti:

- Descrizione di che cosa si intende per "*failures*" e la distinzione dagli *errors*

- Da dove derivano i failures e come poterli gestire
- Viene affrontato quindi il tema del fault management: *fault-isolation e fault-tolerance*
- Vengono distinti i fault in: *Benign fault e Malicious fault*

Successivamente, dopo questa breve fase introduttiva, vengono definite le metriche che un sistema distribuito deve ottemperare:

- **Reliability**
- **Availabilty**
- **Dependability**

Per ciascuna di esse, viene data una definizione, spiegato il legame che le unisce e definito il motivo per cui esse sono considerate tanto importanti quanto ostacolanti nella replicazione dei sistemi distribuiti.

Infine, vengono introdotte alcune delle strategie più rilevanti per la replicazione, distinguendole in base allo scopo per cui viene introdotto questo meccanismo nel sistema distribuito, e quindi:

- Replication for failure management
- Replication for performance

2.1.3 Limit and costs

Dopo aver introdotto la replicazione descrivendone i principali benefici, obiettivi e i motivi che ne inducono l'utilizzo nei sistemi distribuiti, vengono evidenziati gli svantaggi e i problemi che comporta questa strategia, definendone i limiti e i costi.

2.2 Cosa riguarda

Nella seconda parte, vengono analizzate nel dettaglio le strategie di replicazione relative agli elementi più importanti e caratterizzanti dei sistemi distribuiti, andando perciò a trattare diversi aspetti delle seguenti aree:

- **Replication of Data**
- **Replication of Processes**
- **Replication of Objects**
- **Replication of Messages**

2.2.1 Replication of Data

I dati sono l'area per la quale vengono riportate nella presentazione strategie e protocolli con maggior livello di dettaglio, poichè risulta da sempre la più studiata e documentata. I protocolli più importanti che vengono analizzati sono:

- **ROWA e ROWA-A**
- **QC**
- **Hybrid ROWA/QC**

2.2.2 Replication of Processes

Vengono evidenziati diversi approcci, a seconda che la replicazione venga realizzata allo scopo di fault-tolerance o di performance.

Gli approcci riportati si distinguono in due categorie:

- **Modular Redundancy**
- **Primary/standby schemes**

Al primo, appartengono due modelli molto importanti e che vengono riportati nel dettaglio, ossia:

- **Circus approach**
- **Cloud Approach**

2.2.3 Replication of Objects

In questa sezione, vengono trattate le strategie e gli schemi di replicazione, tra i quali si possono distinguere:

- **Class-specific**
- **Instance-specific**
- **Instance-set-specific**

2.2.4 Replication of Messages

Per ultimo, vengono descritte le caratteristiche dei protocolli di replicazione dei messaggi. In particolare dei:

- **Reliable Broadcast Protocol**
- **Quorum Multicast Protocol**

2.3 Heterogeneous, Mobile, and Large-Scale Systems

In conclusione, vengono riportate le difficoltà che si riscontrano nell'implementazione di sistemi di replicazione negli attuali ambienti distribuiti, in particolare:

- **Eterogeneità**
- **Mobilità**
- **Larga-scala**

3 Conclusion

Nella parte conclusiva si dà volutamente un incentivo allo studio di questa area molto importante dei Sistemi Distribuiti.

Ad oggi, sono ancora molti gli aspetti che devono essere studiati e documentati, allo scopo di risolvere e realizzare sistemi sempre più prestanti e adatti allo specifico sistema distribuito.