

LINUXDAY 2020
LinuxDay Online

HackLab2020

Linuxshell Italia Presenta

Raccogliere 4Mln di dati utente online con un budget modesto, in sicurezza e affidabilità.

Documento rilasciato con licenza copy zero X v2.3

Linuxshell è stata la prima associazione a Roma, in ordine temporale, ad occuparsi di software libero e open source. Dal 1995 i professionisti di Linuxshell propongono formazione e soluzioni informatiche basate sui sistemi Gnu/Linux.

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



Linuxshell Italia, da oltre venti anni, si occupa di promuovere alfabetizzazione informatica, e informatica basata su software open e free, collabora con enti e istituzioni, università e centri di ricerca e tiene occasionalmente seminari gratuiti, workshop e corsi di aggiornamento per informatici oltre a erogare le Certificazioni LSI e LCP.

Linuxshell Servizi Informatici in oltre venti anni di attività ha preso parte:

Al Tavolo dell'Altra Economia per il progetto Città dell'Altra Economia del Comune di Roma.

Alla consulenza per la realizzazione dell'Incubatore Open Source del comune di Roma.

Alla formazione della Scuola di Giornalismo "Scrivi di diritto" della Fondazione Internazionale

Lelio Basso, la prima ad avere nei suoi programmi formazione basata su software libero.

Al progetto Respet del Comune di Roma.

A varie iniziative didattiche e formative per i giovani come quelle svolte per l'ITC Lombardo

Radice, L'Ist Lombardi o il Centro Nuove Tecnologie delle Biblioteche di Roma.

All'attività di sviluppo della comunità internazionale per il software libero e la open source initiative.

Collabora e accetta Tirocini con gli studenti dell'Università Roma Tre

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3





**INGEGNERIA
INFORMATICA**

RomaTre Ingegneria Informatica

Linuxshell App Center Istruzione

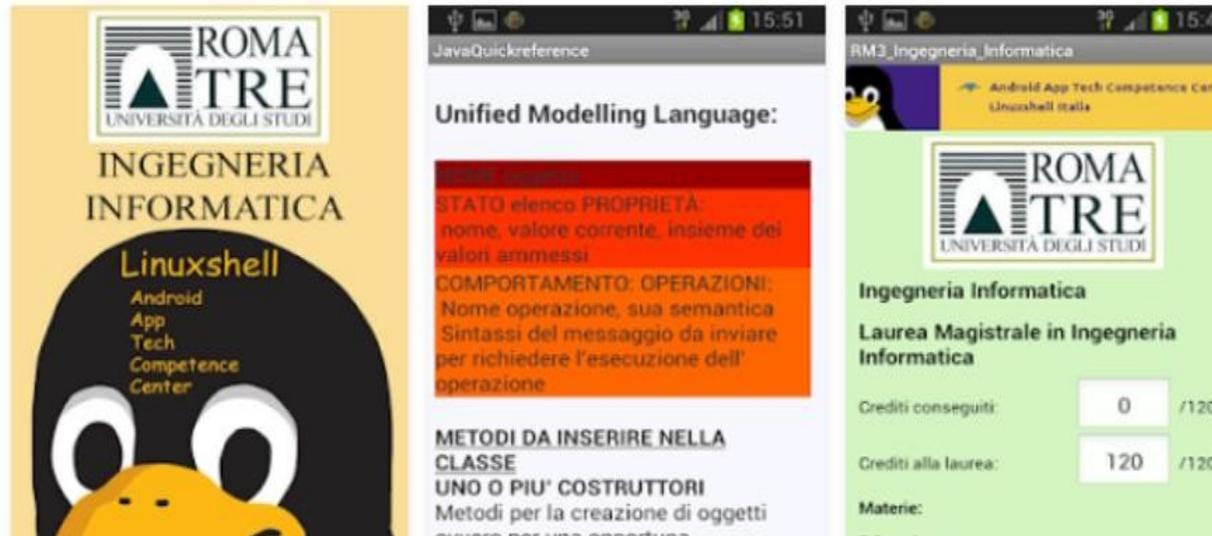
★★★★★ 19

PEGI 3

Contiene annunci

L'app è compatibile con tutti i tuoi dispositivi.

Installata



Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



INFO@LINUXSHELL.IT

<http://www.linuxshell.it/>

<http://www.certificazionilinux.com/>

<https://it-it.facebook.com/linuxshellitalia/>

info@linuxshell.it

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



Come raccogliere ingenti quantità di dati online in maniera veloce, sicura ed originale

Questo studio prende spunto dalle recenti esigenze di raccolta online di dati utente da parte della pubblica amministrazione e di enti privati

Abbiamo cercato un sistema veloce alternativo e sicuro come valida alternativa ai fallimenti che si sono svolti online durante l'emergenza covid-19

Il nostro obiettivo era quello di dimostrare che si possono raccogliere milioni di dati online con un budget modesto mantenendo standard di sicurezza ed affidabilità elevati ed usando software open-source

Gli step che andiamo a dimostrare sono i seguenti :

- Organizzare un sistema che simulasse la ricezione massiccia di centinaia di migliaia di connessioni simultanee senza bloccarsi

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



- Organizzare e salvare i dati
- Proteggere i dati

1) La parte Sistemistica

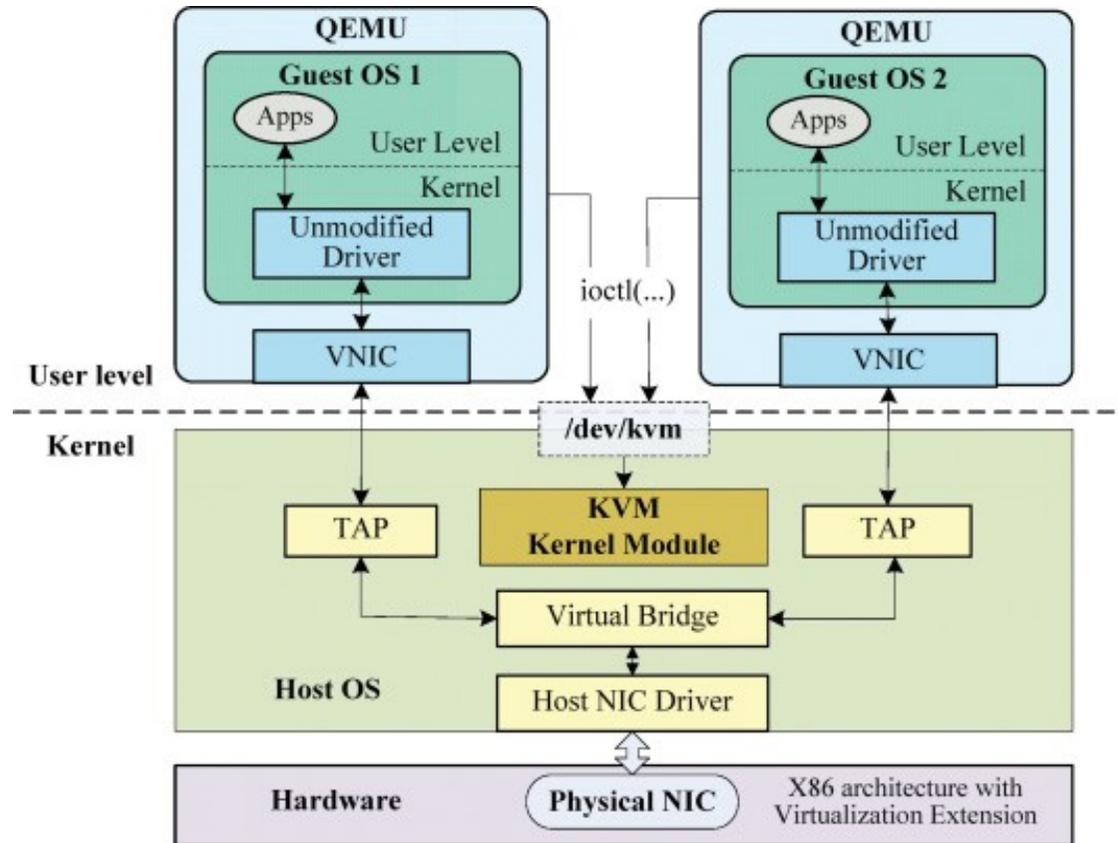
Sono stati configurati i seguenti sistemi:

- VM Ubuntu 20.04 Server con 4 Vcpu, 4 GB di ram e 40 GB di disco
- Virtualizzatore KVM su Server Debian 11 (Testing)
- Server web Apache
- Base di dati Mysql

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



- Python + Html parte sviluppo web
- sistema di controllo NET-DATA



Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3

	 Linuxshell Italia	Formazione	Linux System
		WebDesign	Administrators
		Consulenza	Sviluppo App
		Reti e Sicurezza	Android

La parte sistemistica usa macchine virtuali software open source e simula sia i client che i server.

Il punto cruciale della simulazione è il superamento del flood che si crea quando al server arrivano centinaia di migliaia di richieste generando il dos sia sul server web che sulla base dati.

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



La proof of concept riesce a confermare che saltando il passaggio della connessione al db e interrogando il server web per scrivere dati sul disco attraverso python e apache si riesce egregiamente attraverso i meccanismi di sistema build in nel sistema operativo Gnu/Linux a gestire una mole impressionante di connessioni e dati senza problemi, scrivendo i dati nella base sql in differita. Dando così modo anche di controllo sui dati e validazione ed eseguendo procedure di disaster recovery e sicurezza avanzate. Nella nostra simulazione non abbiano preso nemmeno in considerazione il balancing delle connessioni in cluster di server e tutto viene eseguito su server singolo.

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



2) Lo sviluppo software

parte in linguaggio bash scripting e parte in python

```
script 1 random_words.sh
```

iniziamo a creare la simulazione dei dati da inserire nella base crea 4 files contenenti le informazioni sulle persone (nome, cognome, matricola, iban), in input prende il numero di persone che si vogliono creare

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



script 2 invio.sh

serve per fare un invio massivo di entry dai files creati con random_words.sh la parte principale è il curl verso il webserver in ascolto.

Script inserisci_db3.sh

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



**inserisce nel db i dati ricevuti tramite web server, è
inserito nel crontab in modo da essere scadenzato ogni
tot minuti**

Script Python

form CGI.py

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



Questo script in funzione sul server acquisisce i dati dal form html di inserimento e scrive poi il tutto in un file.

HTML

pagina3.html

Front-end usato per l'inserimento dati, si può visualizzare con un normale browser

3) base dati

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



mysql contenente 1 schema ed 1 tabella – tab_anagrafica e 4 campi



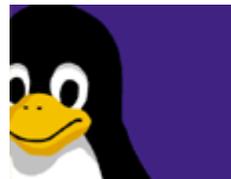
Struttura della tabella



Vista relazioni

#	Nome	Tipo	Codifica caratteri	Attributi	Null	Predefinito	Commenti	Extra	Azione
<input type="checkbox"/>	1	nome	varchar(26)	utf8mb4_0900_ai_ci		No	Nessuno		Modifica Elimina Più
<input type="checkbox"/>	2	cognome	varchar(26)	utf8mb4_0900_ai_ci		No	Nessuno		Modifica Elimina Più
<input type="checkbox"/>	3	cod_fisc	varchar(16)	utf8mb4_0900_ai_ci		No	Nessuno		Modifica Elimina Più
<input type="checkbox"/>	4	iban	int			No	Nessuno		Modifica Elimina Più
<input type="checkbox"/>	5	matricola	int			No	Nessuno		Modifica Elimina Più

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



Linuxshell
Italia

Formazione
WebDesign
Consulenza
Reti e Sicurezza

Linux System
Administrators
Sviluppo App
Android

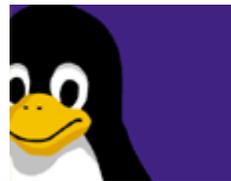
Il Database viene popolato secondo le informazioni scritte nel file di acquisizione dati: eccone un esempio

andrea	prova1	asd	1	12
dams	utmosts	WARTKM52P09L765T	83634	13
berthing	cool	JTXEFE02S94U382P	70467	14
surplice	deodorant	FDOHJX97B17P118Y	37534	15
spiky	saucers	ILZJBK17G16Y676N	46597	16
distentions	catalyzing	ERFMMN27B24W232C	28082	17
davenport	Oxfords	ALJHIO72D92F336T	60237	18

**il campo
matricola è un
numero
progressivo
che riceve la
entry nel**

momento in cui viene scritta.

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



Linuxshell
Italia

Formazione
WebDesign
Consulenza
Reti e Sicurezza

Linux System
Administrators
Sviluppo App
Android

le query che vengono invece fatte sul db sono:

```
INSERT INTO tab_anagrafica (matricola, nome, cognome, iban,  
cod_fisc) VALUES ('$numero_elem', '${ar[0]}', '${ar[1]}', '${ar[3]}', '${  
ar[2]}');
```

inserisci una entry nel DB

```
numero_elem=`mysql -u root -p'andrea' -D db_test -B -N -e "select  
count(*) from tab_anagrafica"
```

calcola il numero di entry nel DB

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



4) Considerazioni sulla sicurezza

Nel sito è installato un certificato self signed di tipo openssl in questo modo ogni transazione viene criptata

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



5) Disaster Recovery

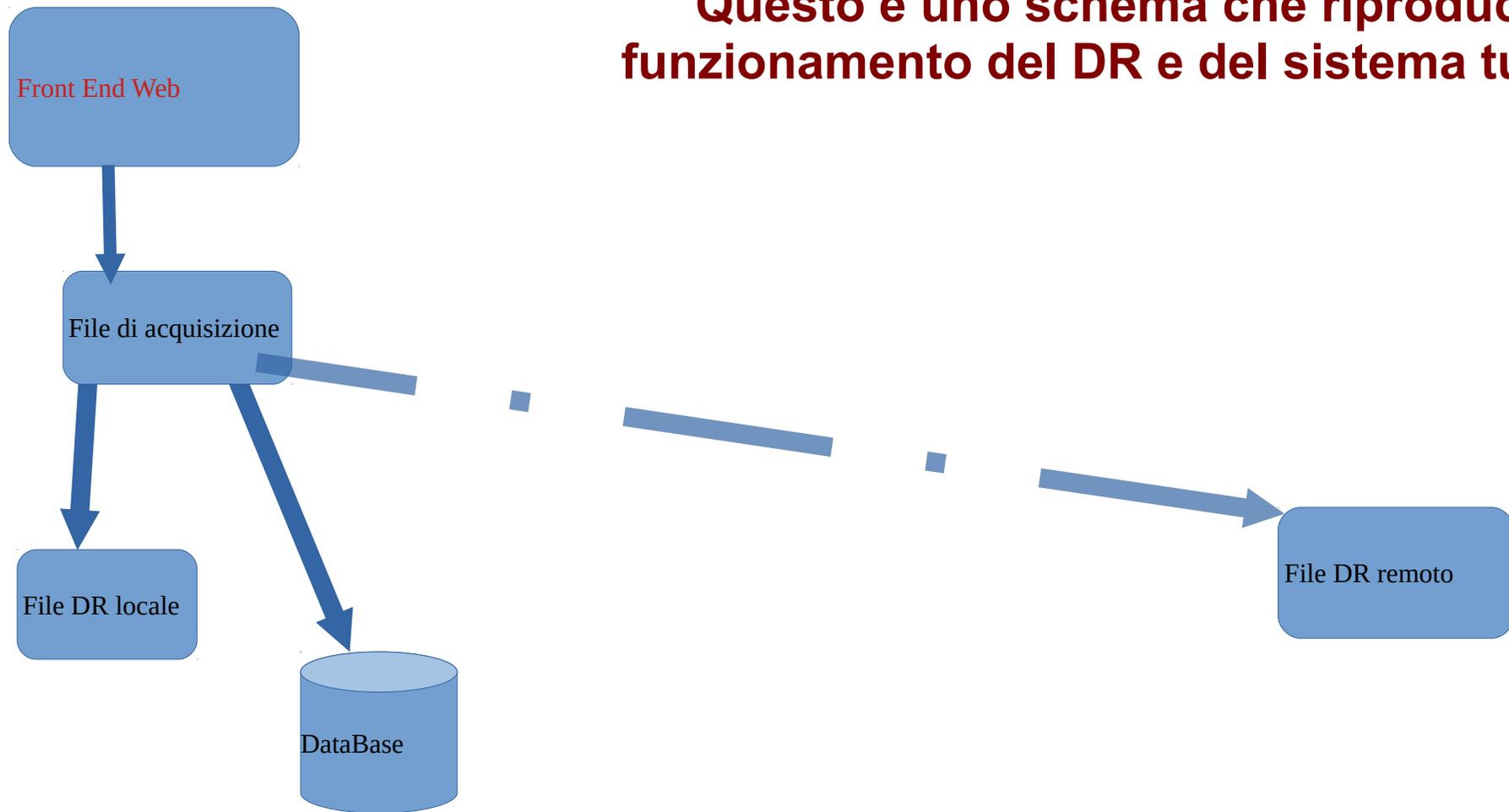
Una parte importante è sicuramente quella del disaster recovery (DR), i dati acquisiti tramite la piattaforma vengono salvati anche in dei files, prima di essere scritti con una query di insert nel database.

Questi files sono catalogati per giorno e scritti sulla macchina primaria e su una macchina esterna per avere la doppia copia (locale-remota) come da documentazione standard DR. La proprietà del DR permette qualora il DB non possa partire di avere comunque l'acquisizione dei dati da parte del sistema e una volta partito il DB di fare una insert partendo da questi files.

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



Questo è uno schema che riproduce il funzionamento del DR e del sistema tutto

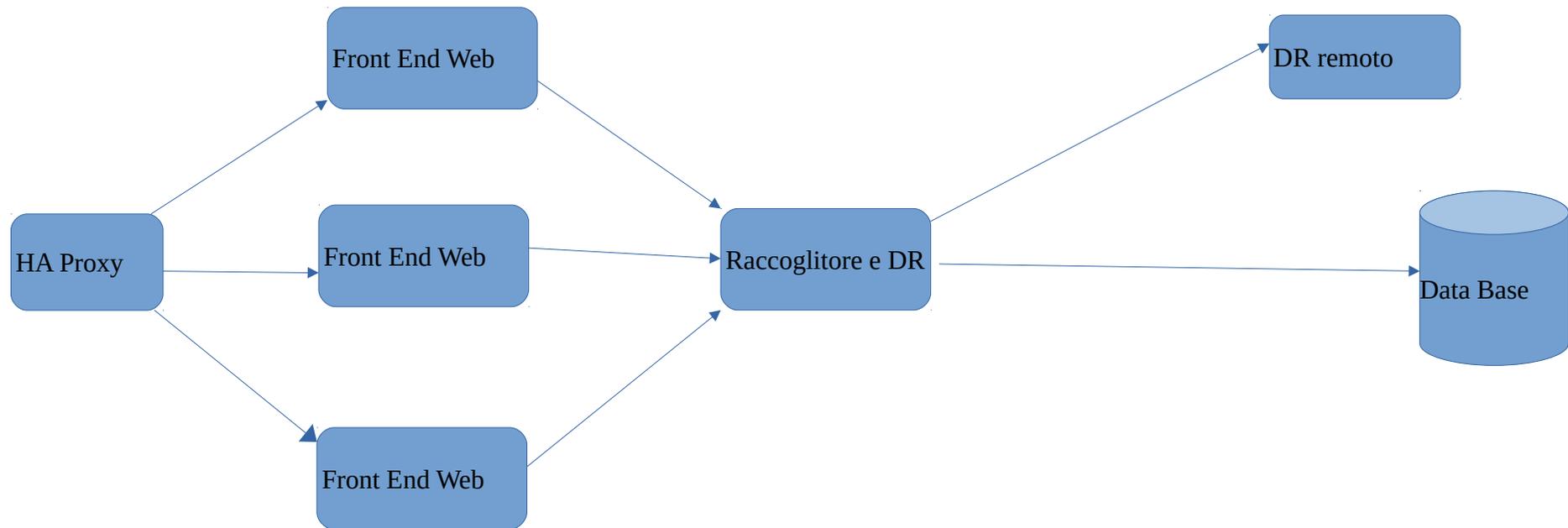


Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



6) Clustering

Questo sistema si può estendere per creare un sistema clusterizzato, ovvero formato da più istanze delle varie parti che lo compongono



Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



In questo caso è stata clusterizzata la parte di raccolta dati da parte della parte web. All'inizio viene inserito un sistema che smista le chiamate verso i web servers, in questo caso è stato usato HA PROXY ma anche Apache2 e Nginx hanno la modalità ad alta affidabilità, poi il sistema raccogliitore costruisce i files che verranno poi scritti tramite l'insert nel DB. Il sistema raccogliitore fa anche da primo sistema di DR e naturalmente c'è un DR remoto

I test sono stati verificati e la matematica necessaria è stata applicata per il calcolo dell'obiettivo dichiarato l'esito dei test è stato positivo ed è stato dimostrata la proof of concepts.

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



PER INFO
INFO@LINUXSHELL.IT

www.linuxshell.it
www.certificazionilinux.com

GRAZIE DI AVERCI SEGUITO
BUON LINUXDAY 2020

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



LICENZA COPYZERO X v.2.3

Il/I sottoscritto/i **Marco Pantò**, in qualità di autore/i dell'Opera rappresentata nel documento informatico corrispondente al **digest SHA-1 digest SHA-1**, dichiara/dichiarano di concedere in licenza (non esclusiva, non revocabile, a titolo gratuito, a tempo indeterminato) tale Opera, nei termini e alle condizioni di seguito specificati.

1. Definizioni

a. Comunicazione: l'impiego di uno dei mezzi di diffusione a distanza, quali il telegrafo, il telefono, la radiodiffusione, la televisione, la rete telematica ed altri mezzi analoghi; la comunicazione al pubblico via satellite e la ritrasmissione via cavo; la messa a disposizione del pubblico dell'Opera in maniera che ciascuno possa avervi accesso dal luogo e nel momento scelti individualmente.

b. Documento informatico: la rappresentazione informatica (file) dell'Opera.

c. Diritti connessi: i diritti degli artisti interpreti ed esecutori, dei produttori fonografici, degli organismi di radiodiffusione; altri diritti indicati nel Titolo II della legge 22 aprile 1941, n. 633.

d. Distribuzione: la messa in commercio o in circolazione, o comunque a disposizione del pubblico, con qualsiasi mezzo ed a qualsiasi titolo, dell'Opera o degli esemplari di essa.

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



e. Elaborazione: la modificazione di carattere creativo dell'Opera.

f. Licenziante/i: la/le persona/e che concede/concedono in licenza l'Opera.

g. Licenziatario: la persona che acquista diritti nei termini e alle condizioni della presente licenza.

h. Misure tecnologiche di protezione: tutte le tecnologie, i dispositivi o i componenti che, nel normale corso del loro funzionamento, sono destinati a impedire o limitare l'esercizio di diritti di utilizzazione economica e/o l'esercizio di diritti connessi.

i. Noleggio: la cessione in uso dell'Opera o di copie o di supporti dell'Opera, fatta per un periodo limitato di tempo ed ai fini del conseguimento di un beneficio economico o commerciale diretto o indiretto.

j. Opera: l'Opera o le Opere dell'ingegno (di cui al primo comma dell'art. 1, ai numeri 1, 2, 3, 4, 5, 6, 7 e 10 dell'art. 2, all'art. 3 e all'art. 4 della legge 22 aprile 1941, n. 633) rappresentata/e nel documento informatico corrispondente al digest SHA-1 indicato nella presente licenza.

k. Opera derivata: una trascrizione, una traduzione o un'elaborazione dell'Opera.

l. Prestito: la cessione in uso dell'Opera o di copie o di supporti dell'Opera, fatta per un periodo di tempo limitato ed a fini diversi da quelli relativi al noleggio.

m. Riproduzione: la moltiplicazione in copie (diretta o indiretta, temporanea o permanente, parziale o totale) dell'Opera, in qualunque modo o forma, come la copiatura a mano, la stampa, la litografia, l'incisione, la fotografia, la fonografia, la cinematografia, la duplicazione informatica ed ogni altro procedimento di riproduzione.

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



n. Traduzione: la traduzione dell'Opera in altra lingua o dialetto.

o. Trascrizione: l'uso dei mezzi atti a trasformare l'Opera orale in Opera scritta o riprodotta.

2. Diritti di utilizzazione economica

Il/I licenziante/i autorizza/autorizzano il licenziatario ad esercitare il/i diritto/i selezionato/i con una "X":

X diritto di riprodurre l'Opera;

X diritto di distribuire l'Opera;

X diritto di pubblicare l'Opera;

X diritto di pubblicare l'Opera in raccolta;

X diritto di comunicare l'Opera al pubblico;

- **diritto di eseguire, rappresentare o recitare l'Opera in pubblico;**

- **diritto di dare in prestito l'Opera;**

- **diritto di noleggiare l'Opera;**

- **diritto di utilizzare l'Opera a scopo di lucro diretto;**

- **diritto di utilizzare l'Opera a scopo di lucro indiretto (fine promozionale);**

- **diritto di trascrivere l'Opera;**

X diritto di tradurre l'Opera;

- **diritto di elaborare l'Opera;**

- **diritto di modificare l'Opera.**

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



Per esercitare il/i diritto/i non selezionato/i e i diritti non presenti nell'elenco sopraindicato occorre l'autorizzazione del/dei licenziante/i.

3. Diritti connessi e diritto al compenso

Il/i licenziante/i dichiara/dichiarano di essere il/i titolare/i dei diritti connessi (relativi al documento informatico corrispondente al digest SHA-1 indicato nella presente licenza) ed autorizza/autorizzano il licenziatario ad esercitarli.

Fatta eccezione per i compensi irrinunciabili di cui alla legge 22 aprile 1941, n. 633, il/i licenziante/i rinuncia/rinunciano al diritto esclusivo di riscuotere i compensi a lui/loro attribuiti dalla medesima legge.

4. Opere derivate

Se a sinistra compare una "X", l'Opera derivata, se destinata alla pubblica utilizzazione, dovrà essere concessa in licenza negli stessi termini e alle stesse condizioni della presente licenza.

Qualora l'autore dell'Opera derivata sia iscritto ad un organismo di intermediazione, l'Opera derivata non potrà essere concessa in licenza e potrà essere oggetto di solo utilizzo privato.

Il titolo dell'Opera derivata e il/i nome/i del/dei suo/suoi autore/i devono precedere il titolo dell'Opera originaria e il/i nome/i del/dei suo/suoi autore/i, secondo lo schema riportato nella "Nota di Copyright" della presente licenza.

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



5. Misure tecnologiche di protezione

X Se a sinistra compare una "X", il licenziatario non può apporre sull'Opera misure tecnologiche di protezione.

6. Organismi di intermediazione

Il/i licenziante/i dichiara/dichiarano di non essere iscritto/i ad organismi di intermediazione nazionali (quali SIAE, NUOVOIMAIE, SCF) o esteri.

7. Risoluzione della licenza

La presente licenza si intenderà risolta e i diritti con essa concessi si estingueranno automaticamente in caso di inadempimento della presente licenza (in particolare, delle disposizioni di cui ai paragrafi 2, 4 e 5) da parte del licenziatario.

8. Limitazione di responsabilità

Il/i licenziante/i non presta/prestano alcuna garanzia con riguardo all'Opera e si esime/esimono da ogni responsabilità che possa derivare dall'uso dell'Opera o dalla presente licenza.

La presente clausola non ha lo scopo di escludere o di limitare la responsabilità del/dei licenziante/i in caso di dolo o colpa grave e nei casi in cui tale

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



esclusione o limitazione non è consentita in forza di inderogabile disposizione di legge.

9. Legge applicabile

La presente licenza è soggetta alle leggi della Repubblica Italiana.

10. Foro competente

Per ogni disputa relativa alla presente licenza, sarà competente il Foro italiano, specializzato in materia di proprietà intellettuale, più vicino al domicilio del licenziante unico o primo firmatario.

Nota di Copyright

Titolo Opera: **Raccogliere 4Mln di dati utente online con un budget modesto, in sicurezza e affidabilità.**

Anno di creazione: **2020**

Nome/i Autore/i: **Marco Pantò**

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3



Titolo Opera originaria: **Raccogliere 4Mln di dati utente online con un budget modesto, in sicurezza e affidabilità.**

Anno di creazione: **2020**

Nome/i Autore/i: **Marco Pantò**

Data

L'Opera è stata rilasciata con la presente licenza in data **20/10/2020**

Linuxshell Italia – Documento rilasciato con licenza copy zero X v2.3

