

Project: Validating a TCP Packet 16

בפרויקט זה תכתוב קוד שמאמת חבילה של TCP, תוך ווידוא שהיא לא נפגמה במהלך ההעברה.

קלט: סדרת זוגות של קבצים:

- אחד מכיל את כתובת ה-IP של המקור והיעד בפורמט של נקודות ומספרים.
- השני מכיל את חבילת ה-TCP הגולמית, כולל כותרת ה-TCP והנתונים שלה.

תוכל להוריד את קבצי הקלט מתוך תיקיית התרגילים.

פלט: לכל זוג קבצים, הדפס "PASS" אם סכום הבדיקה של ה-TCP נכון. אחרת, הדפס "FAIL". יש הרבה חלקים בפרויקט זה, ולכן מומלץ לכתוב ולבדוק תוך כדי עבודה.

עליך להבין את המפרט הזה ב-100% לפני שתתחיל לתכנן את הגישה שלך! קבל הבהרה לפני המעבר לשלב התכנון!

החלק הקשה ביותר בפרויקט הזה הוא להבין אותו! הקוד שלך לא יעבוד לפני שתבין אותו!

הפתרון המודלי הוא 37 שורות קוד! (לא כולל רווחים והערות). זה לא מספר שצריך לשבור, אלא אינדיקציה לכמות המאמץ שתצטרך להשקיע בהבנת המפרט הזה לעומת הקלדת קוד!

16.1 פונקציות אסורות

אסור לך להשתמש בשום דבר בספריית הסוקט.

16.2 איך לכתוב את הקוד הזה

אתה יכול לעשות זאת איך שאתה רוצה, אבל אני ממליץ על הסדר הבא. פרטים לכל אחד מהשלבים האלה כלולים בחלקים הבאים.

1. קרא את קובץ `tcp_addrs_0.txt`.
2. חילק את השורה לשניים: כתובת המקור וכתובת היעד.
3. כתוב פונקציה שממירה כתובת IP בפורמט של נקודות ומספרים למיתרים של בתים.
4. קרא את קובץ `tcp_data_0.dat`.
5. כתוב פונקציה שמייצרת את המיתרים של כותרת ה-IP הפיקטיבית מהכתובת של ה-IP מקובץ `tcp_addrs_0.txt` ואורך ה-TCP מקובץ `tcp_data_0.dat`.
6. בנה גרסה חדשה של נתוני ה-TCP שבה סכום הבדיקה מוגדר לאפס.
7. חבר את כותרת ה-IP הפיקטיבית עם נתוני ה-TCP שיש בהם סכום בדיקה לאפס.
8. חשב את סכום הבדיקה של החיבור הזה.
9. חילץ את סכום הבדיקה מהנתונים המקוריים בקובץ `tcp_data_0.dat`.
10. השווה את שני סכומי הבדיקה. אם הם תואמים, זה עובד!
11. שנה את הקוד שלך כדי להריץ אותו על כל 10 קבצי הנתונים. חמשת הקבצים הראשונים צריכים להכיל סכומי בדיקה תואמים! חמשת הקבצים הבאים לא! כלומר, חמשת הקבצים האחרונים מדמים שהנתונים נפגמו במהלך ההעברה.

16.3 סכום הבדיקה באופן כללי

סכום הבדיקה ב-TCP הוא ערך בן 16 ביט שמייצג "סכום כולל" של כל הבתים בחבילה. (ותוספות נוספות. וזה לא רק סכום הבתים – אל תסכם אותם סתם ככה!!! פרטים למטה.)

כותרת ה-TCP עצמה מכילה את סכום הבדיקה שנשלח מה-host השולח.

ה-host המקבל (שאתה מתיימר להיות כאן) מחשב את סכום הבדיקה שלו מהנתונים שמגיעים, ומוודא שהוא תואם לערך שבחבילה.

אם הוא תואם, החבילה טובה. אם לא, זה אומר שמשונו נפגם. או שכתובת ה-IP של המקור או היעד שגויה, או שכותרת ה-TCP נפגמה, או שהנתונים לא נכונים. או שסכום הבדיקה עצמו נפגם.

במקרה כזה, תוכנת ה-TCP במערכת ההפעלה תבקש לשלוח מחדש את הנתונים אם סכומי הבדיקה לא תואמים.

המשימה שלך בפרויקט הזה היא לחשב מחדש את סכום הבדיקה של הנתונים הנתונים, ולוודא שהוא תואם (או לא) לסכום הבדיקה שכבר קיים בנתונים הנתונים.

16.4 פרטי קבצי קלט

הורד את קובץ ה-ZIP7 עם קבצי הקלט.

ישנם 10 סטים של קבצים. חמשת הקבצים הראשונים מכילים סכומי בדיקה נכונים. חמשת הקבצים האחרונים נפגמו.

אם לא שמת לב, השורה הקודמת מסבירה לך מה עליך לעשות כדי להשיג 100% בפרויקט הזה!

שמות הקבצים הם כך:

tcp_addrs_0.txt tcp_addrs_0.dat

tcp_addrs_1.txt tcp_addrs_1.dat וכן הלאה, עד מספר אינדקס 9. כל זוג קבצים קשור זה לזה.

16.4.1 קובץ ה-txt

תוכל לצפות בקבצים tcp_addrs_n.txt בעורך טקסט ותראה שהם מכילים זוג כתובות IP אקראיות, דומה לדוגמה הבאה:

Copy code
192.0.2.244 192.0.2.207

אלו הן כתובות ה-IP של המקור וכתובת ה-IP של היעד עבור חבילת ה-TCP הזאת.

למה אנחנו צריכים לדעת מידע על ה-IP אם מדובר רק בסכום בדיקה של TCP? הישארו מעודכנים!

16.4.2 קובץ ה-dat

זהו קובץ בינארי המכיל את כותרת ה-TCP הגולמית ואחריה נתוני ה-payload. זה ייראה כמו ג'יבריש בעורך טקסט. אם יש לך תוכנה להדפסה ב-hex, תוכל לצפות בבתים דרך התוכנה הזאת. לדוגמה, פלט מהפקודה hexdump:

```
Copy code
hexdump -C tcp_data_0.dat
3f d7 c9 c5 ed d8 23 52 6a 15 32 96 50 d9 78 d8 00000000
|. |?.....#Rj.2.P.x
be ba aa 2a 63 25 2d 7c 4f 2a 39 52 69 4b 75 67 00000010
| |g...*c%-|0*9RiKu
|B9S| 53 39 42 00000020
00000023
```

אבל עבור פרויקט זה, הדברים היחידים שחשוב לך בקובץ הזה הם:

- האורך (בבתים) של הנתונים.
- סכום הבדיקה (16 ביט) שנמצא במיקומים 16-17 בקובץ (ב-endian גדול).

הערה: קבצים אלו מכילים "כותרות TCP חצי נכונות". כל החלקים שם, אך הערכים השונים (במיוחד בשדות הדגלים והאפשרויות) עשויים לא להיות הגיוניים.

16.5 איך לעזאזל מחשבים סכום בדיקה ב-TCP?

זה לא קל.

סכום הבדיקה ב-TCP נועד לאמת את שלמותם של מספר דברים:

- כותרת ה-TCP
- נתוני ה-TCP
- כתובת ה-IP של המקור והיעד (כדי להגן מפני נתונים מנותבים שגויים).

החלק האחרון הוא מעניין, כי כתובת ה-IP לא נמצאת בכלל בכותרת ה-TCP או בנתונים, אז איך אנחנו מכניסים אותם לסכום הבדיקה?

סכום הבדיקה של TCP הוא מספר של שני בתים שמחושב כך, בהינתן נתוני כותרת TCP, נתוני payload וכתובת ה-IP של המקור והיעד:

- בנה רצף של בתים המייצגים את כותרת ה-IP הפיקטיבית (ראה למטה).
- הגדר את סכום הבדיקה הקיים של כותרת ה-TCP לאפס.
- חבר את כותרת ה-IP הפיקטיבית עם כותרת ה-TCP ונתוני ה-payload.
- חשב את סכום הבדיקה של החיבור הזה.

ככה מחשבים סכום בדיקה ב-TCP.

אבל יש הרבה פרטים.

16.6 כותרת ה-IP הפיקטיבית

מכיוון שאנחנו רוצים לוודא שכתובת ה-IP נכונה עבור נתונים אלו, אנחנו צריכים לכלול את כותרת ה-IP בסכום הבדיקה.

רק שאנחנו לא כוללים את כותרת ה-IP האמיתית. אנחנו בונים אחת פיקטיבית. היא נראית כך (גנוב ישירות מה- RFC של TCP):

```
+-----+-----+-----+-----+
| כתובת ה-IP של המקור |
+-----+-----+-----+-----+
| כתובת ה-IP של היעד |
+-----+-----+-----+-----+
| אפס | PTCL | אורך ה-TCP |
+-----+-----+-----+-----+
```

הערה: כאן ה-PTCL הוא 6 (ל-TCP).

4o mini

○

ChatGPT can make mistake