



TEL AVIV UNIVERSITY

מערכות הפעלה

מרצה: ערן טרומר
סמסטר א' תשע"ב

הרצאה 10

רשתות תקשורת:
IP (המשך) ו-TCP

פרוטוקול ARP

- ❖ פרוטוקול לתרגום כתובות IP לכתובות פיזיות של מחשבים במקטע אתרנט
- ❖ ממומש כחלק משכבת המיקשר, לתרגום כתובות רשת לכתובות מיקשר; פרוטוקולים דומים דרושים למימושי מיקשר אחרים
- ❖ מחשב שצריך לבצע תרגום כזה שולח הודעת שאילתה לכל המחשבים במקטע (הודעת broadcast)
- ❖ מחשב שמזהה את כתובת ה-IP שלו בשאילתה כזו עונה בהודעה עם כתובת האתרנט שלו (MAC address)
- ❖ שכבת המיקשר מטמינה את התרגומים הללו בטבלה לזמן מה על מנת לחסוך בהודעות ARP (שמעכבות משלוח מנות)

דוגמה למטמון ARP

```
$ /sbin/arp -n
```

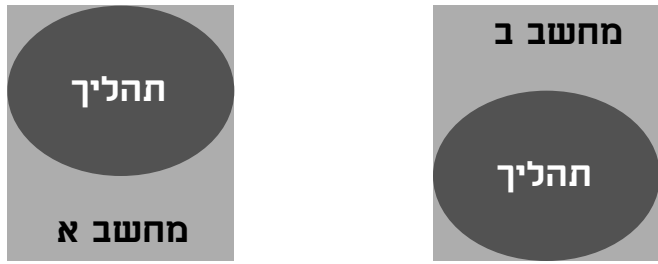
Address	HWtype	HWaddress	Flags	Mask	Iface
192.168.0.1	ether	30:46:9a:96:10:b5	C		wlan0
192.168.0.5	ether	00:2d:0a:53:4d:a4	C		wlan0

תרגום כתובות IP (NAT)

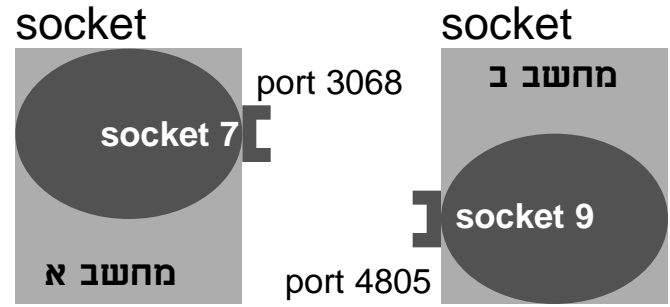
- ❖ להרבה רשתות אין מספיק כתובות IP ציבוריות (כלומר שאפשר לנתב אליהן מנה מכל מחשב ברשת)
- ❖ למשל לבתים פרטיים מוקצה כתובת אחת לכל היותר
- ❖ הרשת משתמשת בטווח של כתובות שמותר להשתמש בהן שימוש חוזר בתוך ארגונים אבל אסור להן להופיע באינטרנט הציבורי (private addresses)
- ❖ הנתב שמחבר את הרשת הפרטית לאינטרנט מתרגם את הכתובות הפרטיות בכל מנה יוצאת לכתובת שלו וממציא מספר שער
- ❖ הוא זוכר את המיפוי ומתרגם בחזרה מנות נכנסות
- ❖ בעייתי ודורש קינפוג כאשר המנה הראשונה בקשר מגיעה מבחוץ
- ❖ טווח כתובות שמורים נפוצים: 192.168.0.0/16 ו- 10.0.0.0/8
- ❖ קונפיגורציה אוטומטית משתמשת בכתובות 169.254.0.0/16

מבוא לשכבת ההעברה: שקעים

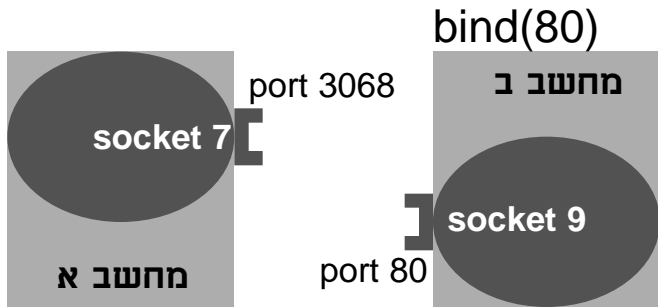
- ❖ כתובת IP מציינת מחשב; על מנת ליצור קשר בין תהליכים צריך כתובת לתהליך
- ❖ תהליך יכול ליצור שקע, עצם שניתן לשלוח ולקבל דרכו מידע
- ❖ שקע חדש מקבל כתובת שרירותית, אבל ניתן לקשור אותו לכתובת מסוימת; הכתובת נקראת **שער**
- ❖ בפרוטוקול UDP ניתן לשלוח מנדעים (מעין מברקים) לשער מסוים בכתובת IP מסויימת
- ❖ בפרוטוקול TCP ניתן לחבר שני שקעים בקשר קבוע ואמין



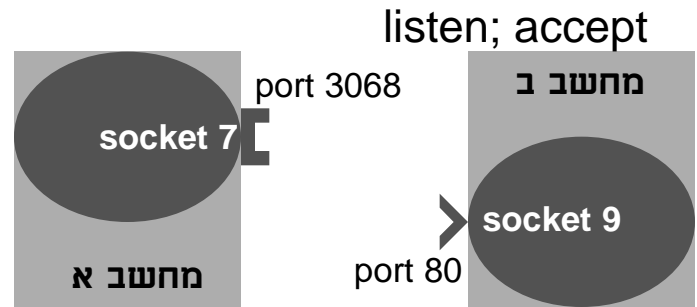
1



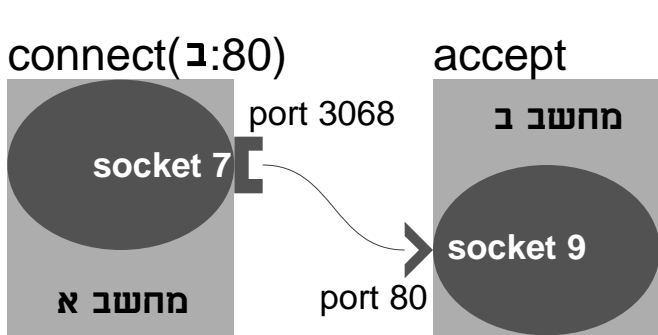
2



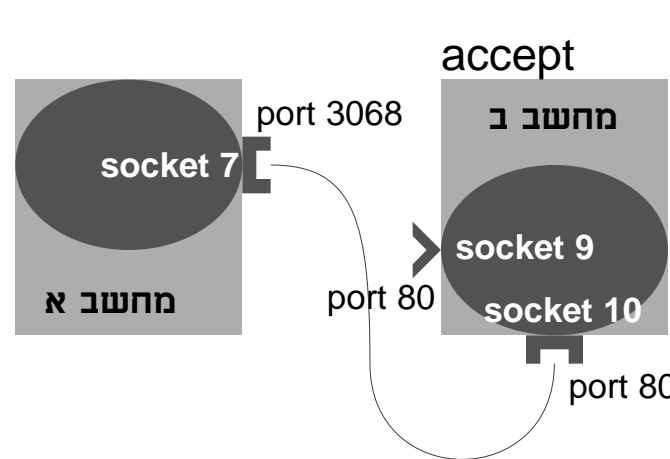
3



4



5



6

זרימת מידע במחסנית הפרוטוקולים

- ❖ קריאת מערכת מעבירה מידע מתהליך (שכבת היישום) לשכבת ההעברה דרך שקע
- ❖ קריאת המערכת מזהה את פרוטוקול ההעברה על פי סוג השקע וקוראת לשגרה מתאימה של אותו פרוטוקול
- ❖ השגרה הזו מפרקת את המידע למנות או מצרפת אותו למידע קודם או מתייחסת אליו כמנה, מוסיפה למנות תחילית עם שדות של הפרוטוקול, ומעבירה אותן לשגרה של שכבת הרשת
- ❖ השגרה של שכבת הרשת מחליטה איך לנתב את המנה, מוסיפה לה תחילית IP, ומעבירה אותה לשגרה של פרוטוקול המיקשר המתאים
- ❖ השגרה של שכבת המיקשר מתרגמת את כתובת ה-IP לכתובת פיזית של מחשב שמחובר לקשת היוצאת, מוסיפה תחילית, ומעבירה לבקר לשליחה

פרק 8

פרוטוקול TCP

תכונות הפרוטוקול

- ❖ קשר דו כיווני אמין וסדור בין תהליכים במחשבים שונים
- ❖ העברה מהירה ככל האפשר של מידע
- ❖ מניעת עומס מיותר על הרשת, עומס שעלול לנבוע ממשלוח מנות שיאבדו וייווצר צורך לשלוח שוב
- ❖ יצירת ופירוק קשר באופן מפורש ומוסכם על ידי שני הצדדים
- ❖ תכונות מתקדמות שלא נדון בהן, כגון הודעות דחופות

בזכות התכונות הללו, פרוטוקול ההעברה הנפוץ ביותר

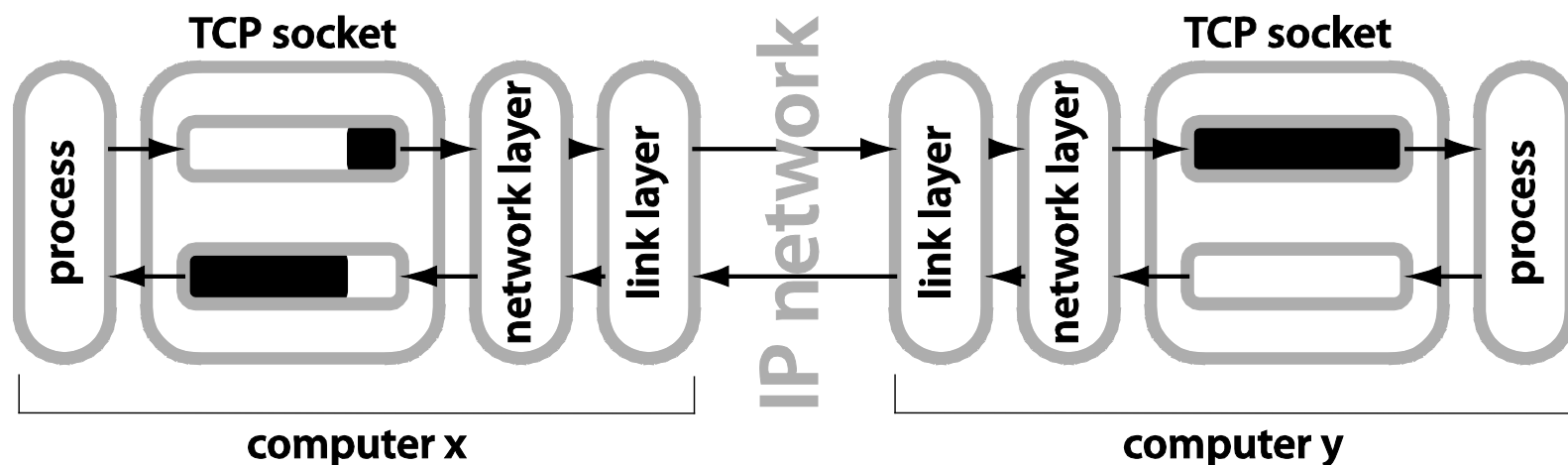
המימוש מורכב, מכיון ש...

- ❖ פרוטוקול הרשת IP אינו אמין ואינו סדור, גם אם פרוטוקולי המיקשר אמינים וסדורים
- ❖ נתב שמקבל מנות בקצב גבוהה מזה שהוא יכול לטפל בהן (למשל בניתוב מערוץ מהיר לאיטי) פשוט מוחק אותן
- ❖ IP שולח מנות בדידות וכל מנה מנותבת בנפרד; כאשר טבלאות הניתוב משתנות, מסלולים משתנים, ומנה מאוחרת ברצף המידע עשוייה להקדים מנה מוקדמת

נגזרות מהדרישות

- ❖ מספור סודר על מנות על מנת שניתן יהיה לשחזר בצד המקבל את הרצף לפי סדר
- ❖ משלוח חוזר של מנות שאבדו
- ❖ אישורי קבלה על מנת שניתן יהיה לזהות מנות שלא הגיעו אחרי פרק זמן סביר ולשלחן שוב
 - מכיון שאי אפשר לדעת האם מנה אבדה או רק מתעכבת הרבה, מנות עלולות להגיע יותר מפעם אחת!
- ❖ חוצצים למשלוח מנות באופן יעיל

חוצים ב-TCP

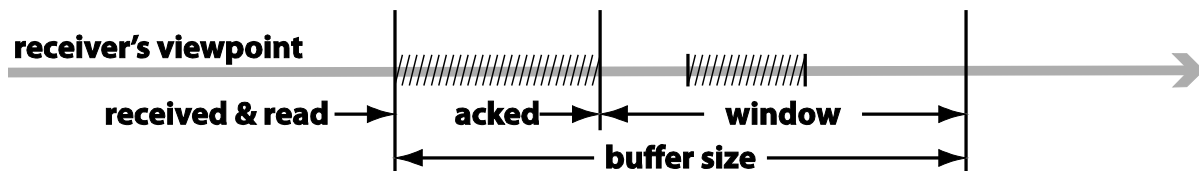
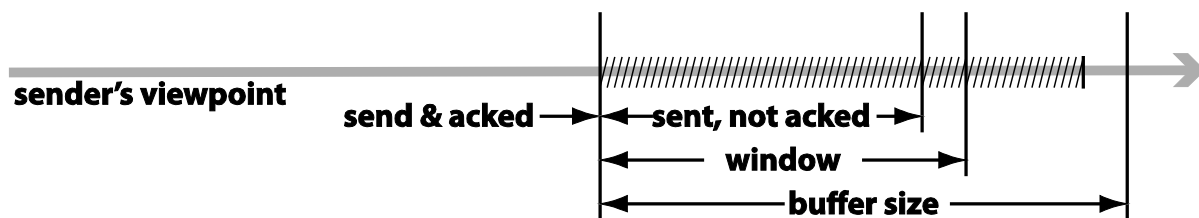


- ❖ החוצץ בצד המקבל מאפשר לשלוח לשלוח כמות גדולה של מידע לפני שהוא מקבל אישור למנה כלשהי; בהעדר חוצץ כזה
 - השולח יכול היה לשלוח מידע שאולי אין היכן לאחסן: עומס ללא צורך
 - לשלוח מנה רק כשהקודמת מאושרת: לא מנצל את רוחב הפס של הרשת
- ❖ החוצץ הזה מאפשר גם לקבל הודעות שהתהליך המקבל עסוק
- ❖ החוצץ בצד השולח מאפשר לתהליך לרוץ בשטף ולנצל את הרשת

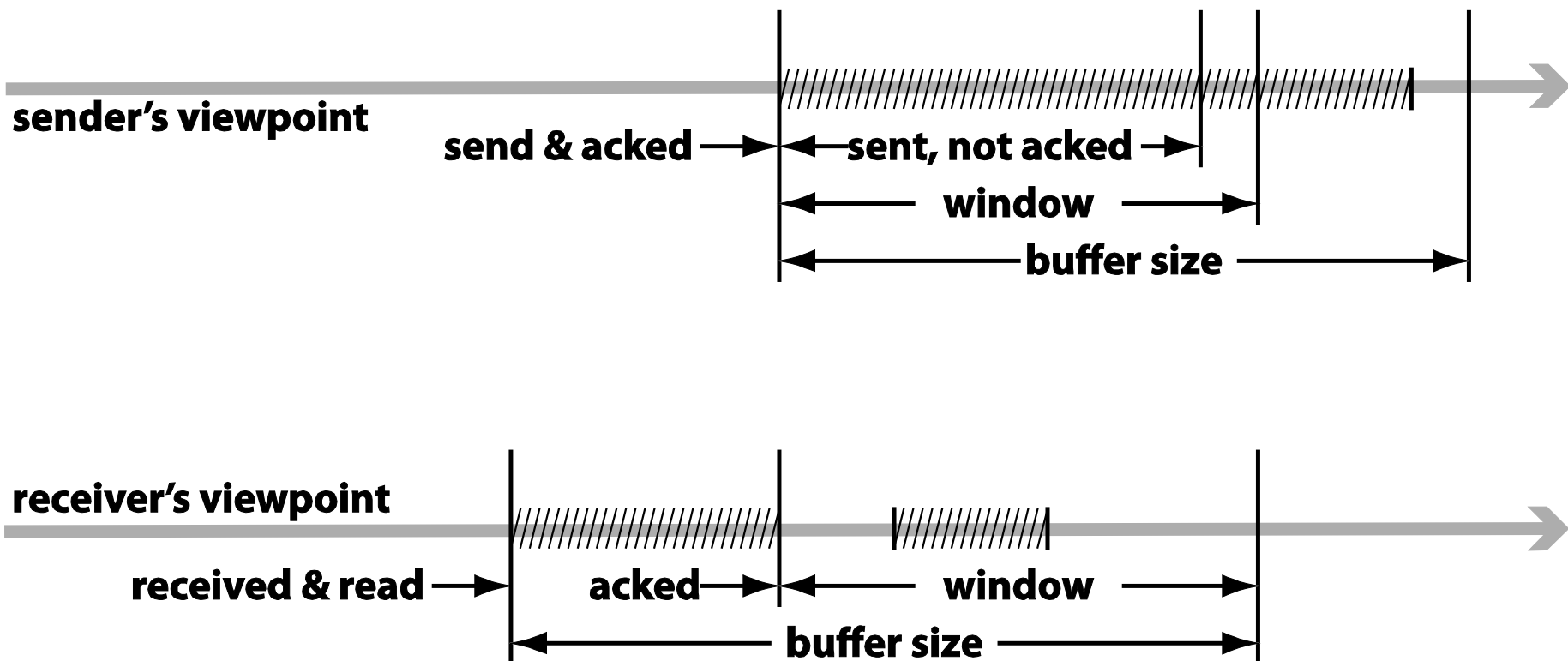
תחילת TCP

❖ לכל מנת מידע מצורפת תחילית עם השדות הבאים

- כמות המידע שהמנה נושאת, בבתים, 0 הוא ערך מותר
- מיקום המידע ברצף מודולו 2^{32} ; השדה הזה מהווה את השדה הסודר
- מיקום הבית האחרון שנקלט ללא חורים על ידי הצד ששולח את המנה (אין דרך לאשר קבלת מנה שנושאת נתונים שנתונים קודמים להם חסרים)
- גודל חלון מודיע לצד השני כמה בתים שלא אושרו מהמשך הרצף חוצץ הקבלה מסוגל להכיל; שדה של 16 סיביות, בבתים או יחידות גדולות יותר
- סכום בדיקה (CRC checksum) לגילוי שגיאות
- שדות שמציינים מצבים מיוחדים

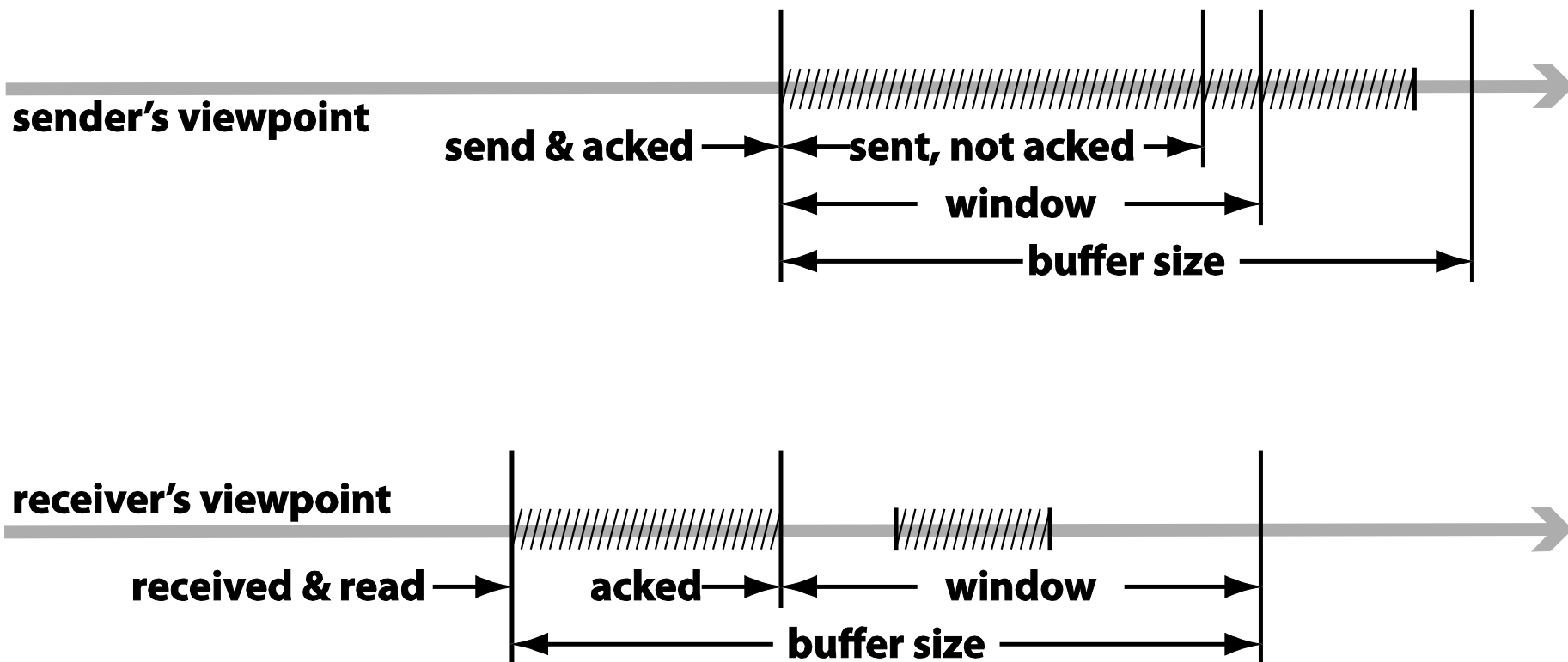


חלונות מחליקים



מצב פשוט: מצב הקשר נראה ל שני הצדדים (לא טיפוסים)

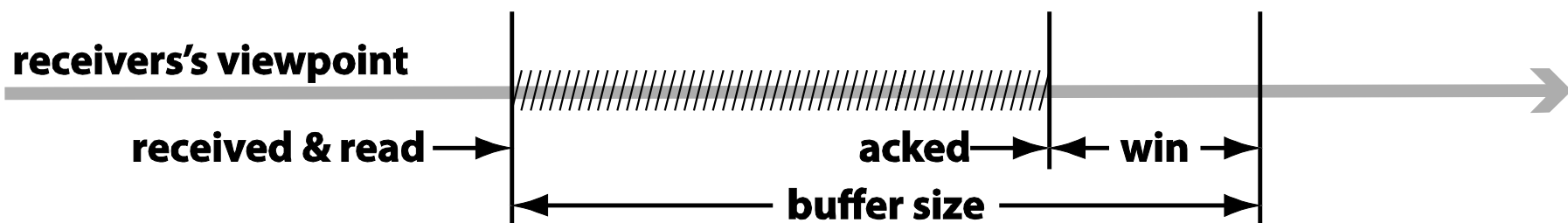
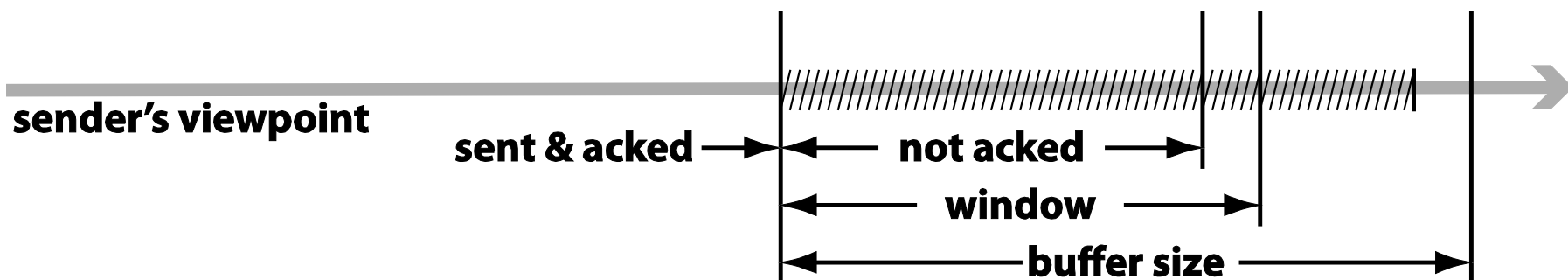
חלונות מחליקים



מצב פשוט: מצב הקשר נראה זהה לשני הצדדים (לא טיפוסי)

מה קורה כאשר מתקבלת מנה שממלאה את החור?

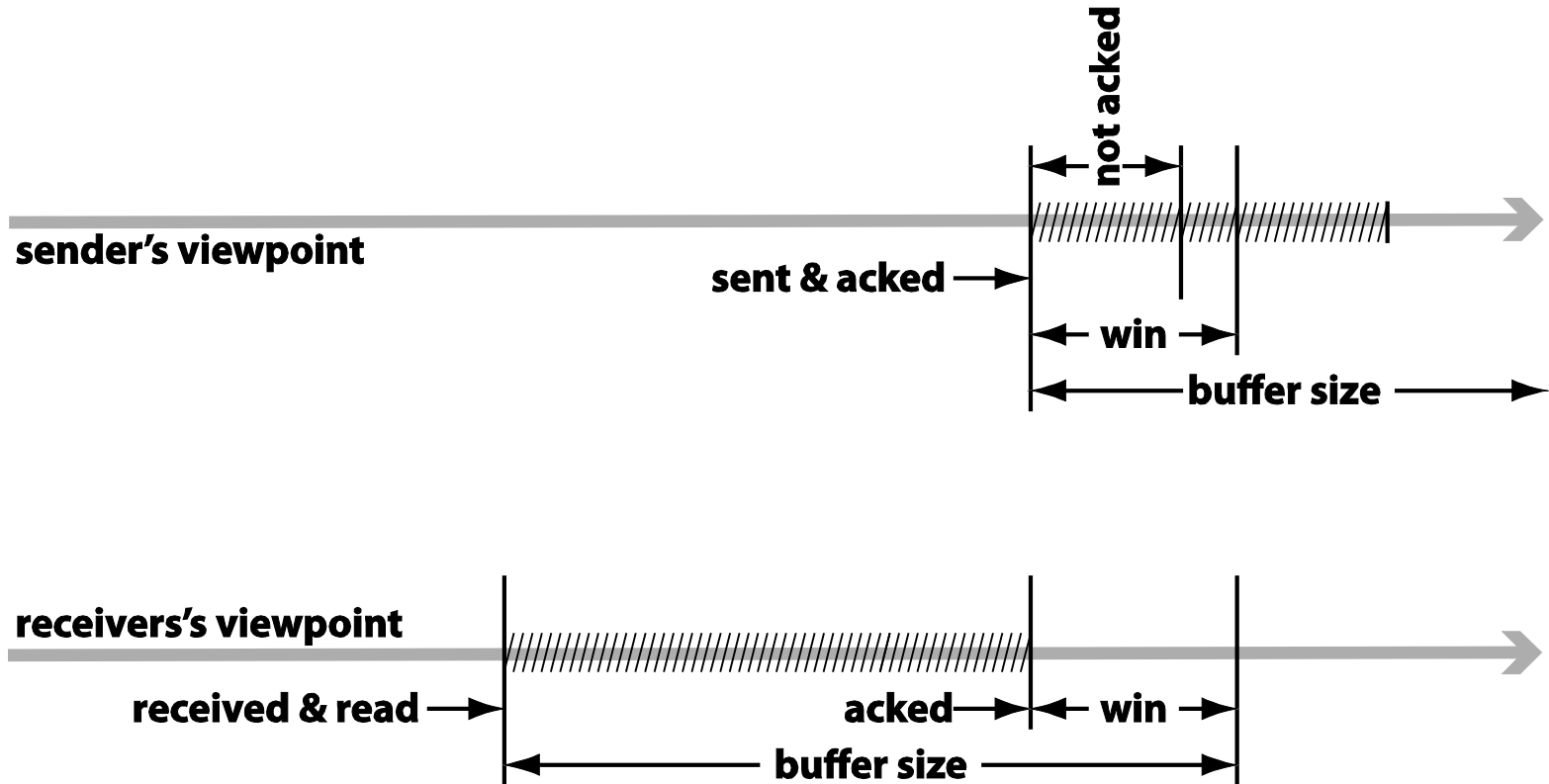
החור התמלא



המקבל שולח אישור (עדיין לא התקבל)

מה יקרה כאשר השולח יקבל את האישור?

האישור התקבל

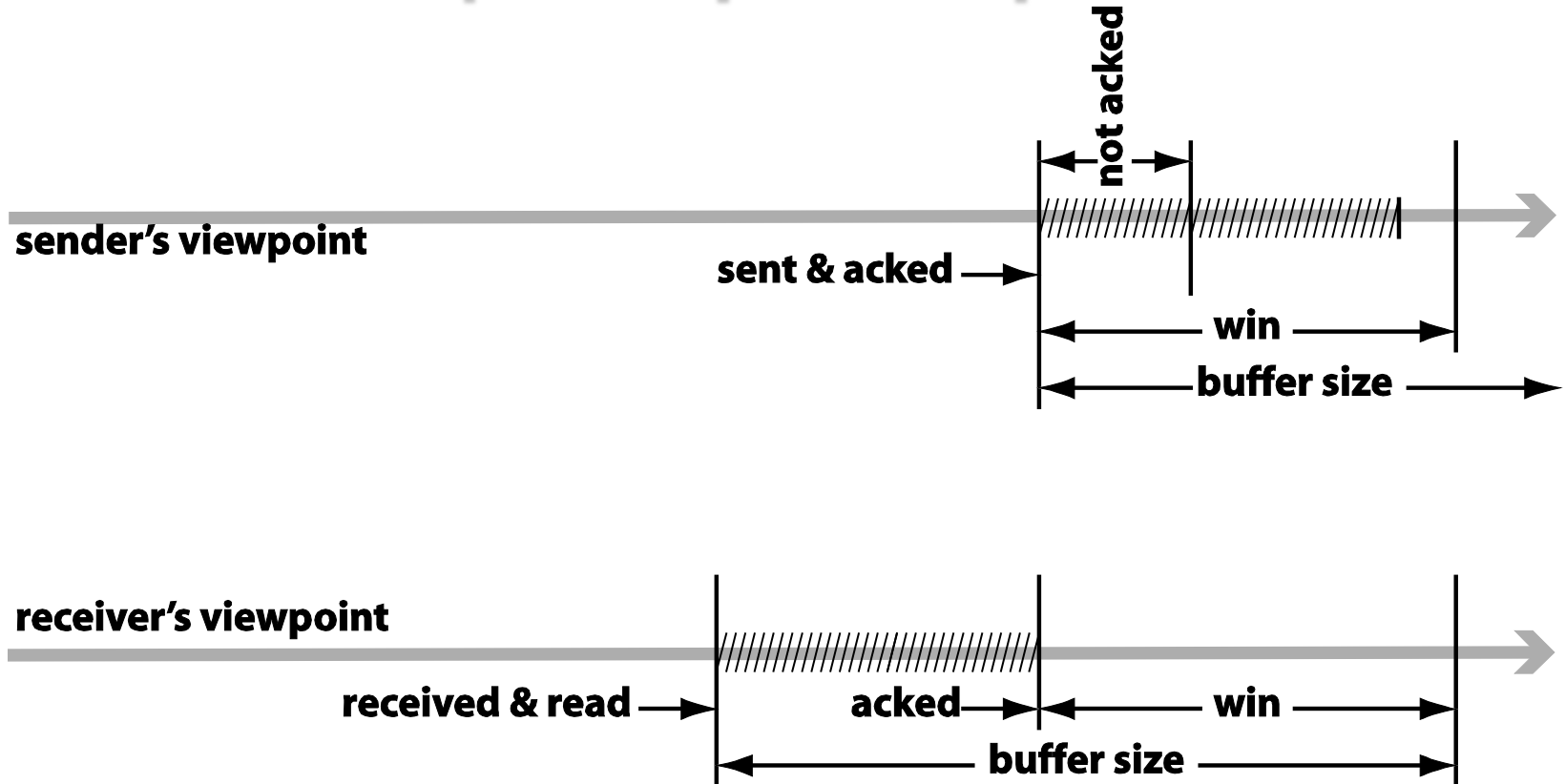


כאשר המקבל יקרא מהחוצץ, הוא ישלח עדכון חלון

מה קורה אם המקבל קורא מעט מדי או יותר מדי נתונים?

מה יקרה כאשר השולח יקבל את עדכון החלון?

עדכון החלון התקבל



מה יקרה כאשר השולח יעביר נתונים נוספים לשקע?

אובדן מנות

- ❖ שירות יקיצה (timer) מזכיר לשולח לשלוח שוב מנות שלא אושרו (שירות שידור חוזר, retransmit)
- שליחת מנה כאשר השירות כבוי מפעילה אותו לנקודת זמן בעתיד
- כאשר הוא פוקע שולחים שוב את המנה הישנה ביותר שלא אושרה ומתזמנים יקיצה לזמן רחוק יותר
- פרק הזמן ליקיצה גדל פי קבוע בכל פעם, עם גג
- הגדלת פרק הזמן מיועדת למנוע שידור חוזר תכופ מדי בערוץ איטי
- כאשר מנה מאושרת, מתזמנים שוב יקיצה עבור המנה הישנה ביותר שעדיין לא אושרה

גם אישורים הולכים לאיבוד

- ❖ התוצאה של אישור אבוד הוא חלון סגור
- ❖ שירות יקיצה בשם persist מזכיר לשולח לברר האם החלון נפתח
- ❖ השירות מופעל כאשר יש מידע לשלוח והחלון סגור (או קטן מדי)
- ❖ בזמן יקיצה, השולח שולח בית בודד אם החלון סגור וכמה שאפשר אם הוא פתוח, ומתזמן יקיצה נוספת
- ❖ כאשר המנה הזו מגיעה, המקבל מאשר אותה עם גודל חלון עדכני
- ❖ שירות היקיצה מכובה כל אימת שמגיע עדכון חלון שמאפשר לשלוח מנה שלמה
- ❖ המקבל מאשר גם מידע שכבר נתקבל ואושר, מחשש שהאישור המקורי אבד

אופטימיזציות 1:

סינדרום החלון האילי

- ❖ חלון כמעט סגור מלמד בדרך כלל שהחוצץ בצד המקבל מלא במידע שהתהליך המקבל לא קרא עדיין
- ❖ במצב כזה אין טעם לשלוח מידע נוסף: זה לא יועיל למקבל ומנות קטנות מעמיסות על המחשבים והרשת
- ❖ על מנת למנוע משלוח מנות קטנות:
 - המקבל אינו מכריז על חלון קטן (פחות ממנה סטנדרטית או פחות מרבע גודל החוצץ שלו) אלא על 0
 - השולח משהה משלוח מנות קטנות עד ליקיצת `persist`
- ❖ התנהגות השולח מעכבת תקשורת ללא צורך כאשר למקבל חוצץ קטן מאוד: כדי למנוע זאת, שולחים מייד אם גודל החלון לפחות חצי מהחלון הגדול ביותר שהוכרז עד כה (שהוא לכל היותר גודל החוצץ בצד המקבל)

אופטימיזציות 2: האלגוריתם של Nagle

- ❖ השולח משהה משלוח מידע חדש שאינו ממלא מנה שלמה כאשר לא כל המנות שנשלחו אושרו
- ❖ כאשר נשלחות כמויות מידע גדולות, העיכוב במשלוח קטן ומונע עומס
- ❖ כאשר ערוץ התקשורת מהיר (רשת מקומית) ואישורים מתקבלים כמעט מייד, אין בדרך כלל עיכוב כלל
- ❖ האלגוריתם הזה פוגע בביצועים של יישומים אינטראקטיביים מסויימים וניתן לכבות אותו על ידי קריאה ל-`setsockopt`

אופטימיזציות 3:

השהיית אישורים ועדכוני חלון

❖ מנה שאינה נושאת נתונים אלא רק אישור ועדכון חלון מעוכבת
עד ש:

- נוצר צורך לשלוח מנה של נתונים, או
- חלון הקבלה גדל למנה שלמה, או
- חלפה חמישית שניה

❖ האופטימיזציה הזו מונעת משלוח מספר עצום של עדכוני חלון
כאשר התהליך הקורא מבקש נתונים מועטים בכל קריאת
מערכת

❖ ההשהיה מוגבלת על מנת למנוע מנות כתוצאה מיקיצות
persist-ו retransmit בצד השני

אופטימיזציות 4:

מניעת עומס

- ❖ עומס על נתבים וקווי תקשורת גורם לאובדן מנות
- ❖ יקיצות persist ו-restrasmit מעידות בדרך כלל על עומס (ורק לעיתים רחוקות על השחתת מידע בתווך)
- ❖ השולח מקטין את קצב שליחת החבילות כאשר יקיצות כאלה גורמות לו לחשוב שהרשת עמוסה
- ❖ המנגנון מבוסס על התנהגות כאילו החלון קטן ממה שהוא באמת, אך לא נתאר אותו באופן מפורט

אופטימיזציות 5:

שליחה חוזרת מהירה

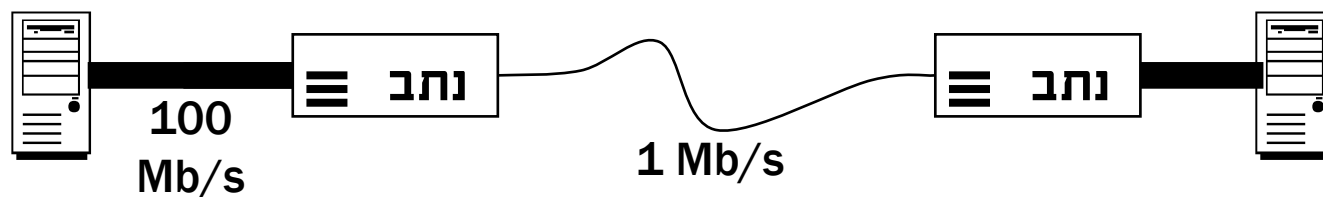
- ❖ השולח מסיק שמנה בודדת אבדה כאשר הוא מקבל שלושה אישורים זהים ברצף
- ❖ ההנחה כאן היא שמנה אבדה ומנות עוקבות התקבלו וגרמו למשלוח אישורים
- ❖ במקרה כזה השולח משגר שוב את המנה שכנראה אבדה ואינו מקטין את קצב השליחה (מהאופטימיזציה הקודמת)

אופטימיזציות 6:

קביעת גודל חוצים

- ❖ בהנחה שהתהליך הקורא לא מעכב מידע בחוצץ, תפקיד החוצץ הוא לאפשר לשולח לשלוח מידע לפני קבלת אישורים
- ❖ מה צריך להיות גודל החוצץ? קצב הערוץ כפול פעמיים ההשהיה
- ❖ למשל,
 - ערוץ לווני שמנה חוצה ברבע שניה ברוחב פס של 100 Mb/s
 - אישור למנה מגיע חצי שניה אחרי שליחתה
 - השולח צריך לשלוח 50 Mb שהם 6 MB לפני קבלת אישור
 - מאותו רגע הוא מקבל אישורים בקצב השליחה
 - החוצץ צריך להיות בגודל 6 MB על מנת לנצל את רוחב הפס של הערוץ
- ❖ חוצץ של 64 KB מאפשר לשלוח רק במאית מרוחב הפס

קביעת גודל חוצים – חוצץ גדול מדי (bufferbloat)



- ❖ חוצץ קבלה גדול מדי מאפשר לשולח להציף נתבים במנות שהנתב אינו יכול להעביר
- ❖ מנגנון מניעת העומס לא ידע על צוואר הבקבוק עד שהחוצץ מלא
- ❖ משהחוצץ בנתב התמלא, הוא יישאר בתפוסה מלאה
- ❖ תוצאה: עיקוב (latency) ארוך מאד בנתב
 - פוגע ביישומים אינטראקטיביים
 - פוגע במנגנון מניעת העומס, שכן זמן התגובה לעומס עולה
- ❖ החוצץ צריך להתאים בגודלו לרוחב הפס האפקטיבי של הערוץ מקצה לקצה ולא לרוחב הפס של מקטעים בדרך

אופטימיזציות 7:

התחלה איטית

- ❖ במצב יציב של הקשר מנה נשלחת כל אימת שמגיע אישור שפותח את החלון למנה נוספת
- ❖ בתחילת הקשר החלון פתוח לרווחה והשולח עלול לשלוח חלון מלא בבת אחת לנתב קרוב
- ❖ אם הנתב לא יכול לשלוח את המנות במהירות, חלקן יאבד
- ❖ על מנת למנוע את זה, השולח משתמש בחלון עומס שמגביל את כמות המידע שנשלח ולא אושר
- ❖ חלון העומס מתחיל במנה אחת וגדל באחת כל אימת שמתקבל אישור
- ❖ קצב שליחת המנות גדל כך אקספוננציאלית לקצב של הערוץ