



## Διάλεξη 6 Εργαλεία Ωφελιμότητας Unix: Sed

Δημήτρης Ζεϊναλιπούρ



# Επεξεργασία Ρεύματος

- **sed** (**S**tream **E**Ditor)
- **awk** (Alfred **A**ho, Peter **W**einberger, and Brian **K**ernighan) ➔ Επόμενη Διάλεξη

Μελετήστε επίσης τα εγχειρίδια που υπάρχουν στην ιστοσελίδα και το *man* στο UNIX.

# Επεξεργαστής Ρεύματος: sed



- **Επεξεργαστής ρεύματος για φιλτράρισμα και μετασχηματισμό του κειμένου που διαβάζεται από το προκαθορισμένο ρεύμα εισόδου ή αρχείο ή από διοχέτευση.**
  - Γράφει τη **μεταλλαγμένη μορφή** του κειμένου στο **προκαθορισμένο ρεύμα εξόδου** ή σε **αρχείο**
  - Αποδοτική τροποποίηση (φιλτράρισμα) εφόσον η είσοδος σαρώνεται ΜΟΝΟ μια φορά.

# Ψευδοκώδικας Λειτουργίας



```
while (read line){
```

1) **sed διαβάσει μια γραμμή εισόδου** από το προκαθορισμένο ρεύμα εισόδου ή από αρχείο, μια-μια κάθε φορά, στο **χώρο προτύπων (*pattern space*)**

*pattern space* = a data buffer = το υφιστάμενο κείμενο όπως επεξεργάζεται

2) Η **sed εκτελεί μια σειρά από εντολές επεξεργασίας** (που έχουν προσδιοριστεί από το χρήστη) **επάνω στο *pattern space***

3) Το ***pattern space*** γράφεται στο **ρεύμα εξόδου**

```
}
```

# Αντικατάσταση (Substitution) με την Sed



- Η πιο ουσιαστική εντολή της sed: **s** (*substitution* - αντικατάσταση)
  - Σύνταξη: ***s/pattern/replacement/<flags>***
    - ***pattern***: πρότυπο ανεύρεσης
    - ***replacement***: συμβολοσειρά αντικατάστασης του προτύπου
    - ***flags***: προαιρετικά
      - ***g*** : καθολική (global): αντικατάσταση όλων των εμφανίσεων του προτύπου στη γραμμή που διαβάζεται (εναλλακτικά αλλάζεται μόνο η πρώτη εμφάνιση)
      - ***p*** : εκτύπωση του περιεχομένου του *pattern space*
      - ***n*** : ένας αριθμός από 1-512 που προσδιορίζει ποια εμφάνιση του προτύπου θα αντικατασταθεί

# Αντικατάσταση (Substitution)

## Παραδείγματα



- Παραδείγματα:

```
bash-3.1$ cat -n example.sed
```

```
1 This is the first line of an example text.  
2 It is a text with erors.  
3 Lots of erors.  
4 So much erors, all these erors are making me sick.  
5 This is a line not containing any errors.  
6 This is the last line.
```

# Αντικατάσταση (Substitution)

## Παραδείγματα



- Παραδείγματα:

```
bash-3.1$ sed 's/erors/errors/' example.sed
```

This is the first line of an example text.

It is a text with **errors**.

Lots of **errors**.

So much **errors**, all these erors are making me sick.

This is a line not containing any errors.

This is the last line.

*Η δεύτερη εμφάνιση δεν μεταβάλλεται*

# Αντικατάσταση (Substitution) Παραδείγματα



- Παραδείγματα:

αντικατάσταση όλων των  
εμφάνισεων του προτύπου στη  
γραμμή που διαβάζεται

```
bash-3.1$ sed 's/erors/errors/g' example.sed
```

```
This is the first line of an example text.
```

```
It is a text with errors.
```

```
Lots of errors.
```

```
So much errors, all these errors are making me sick.
```

```
This is a line not containing any errors.
```

```
This is the last line.
```

```
bash-3.1$
```



# Αντικατάσταση (Substitution) Παραδείγματα



- Παραδείγματα:

2<sup>η</sup> εμφάνιση του προτύπου στη  
γραμμή κειμένου που διαβάζεται

```
bash-3.1$ sed 's/erors/errors/2' example.sed
```

```
This is the first line of an example text.
```

```
It is a text with erors.
```

```
Lots of erors.
```

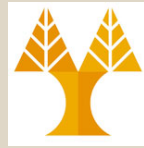
```
So much erors, all these errors are making me sick.
```

```
This is a line not containing any errors.
```

```
This is the last line.
```

# Αντικατάσταση (Substitution)

## Παραδείγματα



### Αναδρομικότητα της Sed;

- Η εντολή sed ΔΕΝ είναι αναδρομική.
- Επομένως, η ακόλουθη εντολή ΔΕΝ θα καταλήξει σε ένα ατέρμον βρόχο επανάληψης
- `sed 's/loop/loop the loop/g' <old >new`
- Το πιο πάνω θα αντικαταστήσει απλά την κάθε εμφάνιση της λέξης “loop” με “loop the loop” από αριστερά δεξιά

# SED (BRE vs ERE)



- BRE: `.*[]^$ \ + \| \?` BRE Char: `\ . \| * \| ^ \| $ \| \| \| \ + \| \| \?`
- ERE: `.*[]^$ \ + \| \?` ERE Char: `\ . \| * \| ^ \| $ \| \| \| \| \ + \| \| \| \?`

- BRE Syntax: **sed**

- `$ echo "aab." | sed -n '/a\ + b\ . /p'`  
aab
  - `$ echo "aab+" | sed -n '/a\ + b\ + /p'`  
aab+

Δεν εκτυπώνει τις γραμμές που επεξεργάζεται η sed (απλά εκτυπώνει αυτά που αναγνωρίζονται από το pattern).

- ERE Syntax: **sed -E**

- `echo "aab." | sed -E -n '/a + b\ . /p'`  
aab.
  - `echo "aab+" | sed -E -n '/a + b\ + /p'`  
aab+

# Αντικατάσταση (Substitution)

## Παραδείγματα



- Παραδείγματα:

Αντικατάσταση αρχής γραμμής κειμένου με τους χαρακτήρες "> "

bash-3.1\$ **sed 's/^/> /' example.sed**

> This is the first line of an example text.  
> It is a text with errors.  
> Lots of errors.  
> So much errors, all these errors are making me sick.  
> This is a line not containing any errors.  
> This is the last line.

# Αντικατάσταση (Substitution)

## Παραδείγματα



- Παραδείγματα:

τέλος γραμμής κειμένου που  
διαβάζεται

```
bash-3.1$ sed 's/$/ ENDOFLINE/' example.sed
```

```
This is the first line of an example text. ENDOFLINE
```

```
It is a text with errors. ENDOFLINE
```

```
Lots of errors. ENDOFLINE
```

```
So much errors, all these errors are making me sick. ENDOFLINE
```

```
This is a line not containing any errors. ENDOFLINE
```

```
This is the last line. ENDOFLINE
```

# Αντικατάσταση (Substitution)

## Παραδείγματα



- Παραδείγματα (Πολλαπλές εντολές) – διαφορετικό από το `sed -E` :

```
bash-3.1$ sed -e 's/erors/errors/g' -e 's/last/final/g'  
example.sed
```

ή

```
sed 's/erors/errors/g; s/last/final/g' example.sed
```

This is the first line of an example text.

It is a text with **errors**.

Lots of **errors**.

So much **errors**, all these **errors** are making me sick.

This is a line not containing any errors.

This is the **final** line.

# Αντικατάσταση (Substitution)

## Εναλλακτικοί Οριοθέτες



- Εάν χρειάζεται να γίνει χρήση του **slash /** (το οποίο **χρησιμοποιείται** ως οριοθέτης της sed) τότε υπάρχουν οι ακόλουθες **εναλλακτικές επιλογές**:

- Χρήση του **escaped-slash** (δηλ., **\**)

- Π.χ., Εάν πρέπει να αλλάξει το /usr/local/bin σε common/bin:

```
sed 's/\usr/local/bin/common/bin/'
```

- Χρήση άλλων οριοθετών: underscore, κόμμα, το | ή το #

```
sed 's#/usr/local/bin#/common/bin#'
```

# Αντικατάσταση (Substitution): Απομνημόνευση Pattern με &



- **Ειδικός χαρακτήρας &**

- Απομνημονεύει στο πρότυπο που έχει ταιριάξει-βρεθεί για να επανα-χρησιμοποιηθεί.

*Επανάληψη χαρακτήρων a-z, A-Z μια ή περισσότερες φορές*

```
bash-3.1$ sed 's/[a-zA-Z]\+\. /(&)/g' example.sed
```

```
This is the first line of an example (text.)
```

```
It is a text with (erors.)
```

```
Lots of (erors.)
```

```
So much erors, all these erors are making me (sick.)
```

```
This is a line not containing any (errors.)
```

```
This is the last (line.)
```

→ χαρακτήρας



# Κανονικές Εκφράσεις (Regular Expressions)



## POSIX Class Definitions

Value	Meaning
<code>[ :digit: ]</code>	Only the digits 0 to 9
<code>[ :alnum: ]</code>	Any alphanumeric character 0 to 9 OR A to Z or a to z.
<code>[ :alpha: ]</code>	Any alpha character A to Z or a to z.
<code>[ :blank: ]</code>	Space and TAB characters only.
<code>[ :xdigit: ]</code>	Hexadecimal notation 0-9, A-F, a-f.
<code>[ :punct: ]</code>	Punctuation symbols . , " ' ? ! ; : # \$ % & ( ) * + - / < > = @ [ ] \ ^ _ { }   ~
<code>[ :print: ]</code>	Any printable character.
<code>[ :space: ]</code>	Any whitespace characters (space, tab, NL, FF, VT, CR). Many system abbreviate as \s.
<code>[ :graph: ]</code>	Exclude whitespace (SPACE, TAB). Many system abbreviate as \W.
<code>[ :upper: ]</code>	Any alpha character A to Z.
<code>[ :lower: ]</code>	Any alpha character a to z.
<code>[ :cntrl: ]</code>	Control Characters NL CR LF TAB VT FF NUL SOH STX EXT EOT ENQ ACK SO SI DLE DC1 DC2 DC3 DC4 NAK SYN ETB CAN EM SUB ESC IS1 IS2 IS3 IS4 DEL.

# Αντικατάσταση (Substitution): Απομνημόνευση Pattern με &



- Ειδικός χαρακτήρας & *Any character NOT whitespace (e.g., alphanumeric or symbols)*

```
$ sed 's/[^\ ]\+/(&)/' example.sed
```

```
(This) is the first line of an example text.
```

```
(It) is a text with errors.
```

```
(Lots) of errors.
```

```
...
```

```
$ sed 's/[^\ ]\+/(&)/g' example.sed
```

```
(This) (is) (the) (first) (line) (of) (an) (example) (text.)
```

```
(It) (is) (a) (text) (with) (errors.)
```

```
(Lots) (of) (errors.)
```

```
...
```

```
$ sed 's/[^\ ]/(&)/' example.sed
```

```
(T)his is the first line of an example text.
```

```
(I)t is a text with errors.
```

```
(L)ots of errors.
```

```
...
```

```
bash-3.1$ echo "123 abc" | sed 's/[0-9]*/& &/'
```

```
123 123 abc
```

**ΙΔΙΟ με /[^\ ]\+/(&)/ ή [^[:space:]]\+**



# Αντικατάσταση (Substitution):

## Απομνημόνευση Πολλαπλών Patterns με /1, /2, ..

- Στα πλαίσια κανονικών εκφράσεων συχνά επιθυμούμαι να απομνημονεύουμε πολλαπλά patterns.
- Π.χ., Κάθε εμφάνιση της λέξης «**is**» να ανταλλάξει θέση με την λέξη «**erors**» (εφόσον εμφανίζονται και οι δυο λέξεις)

- `sed 's/\(is\) \(.*\) \(erors\) /\3\2\1/g' a.txt`  
Copy to Var1      Copy to Var2      Copy to Var3      Display Sequence



This is the first line of an example text.

It **erors** a text with **is**.

Lots of errors.

So much errors, all these errors are making me sick.

This is a line not containing any errors.

This is the last line.

# Εκτύπωση (Print) με /p



- Εκτύπωση γραμμών που περιέχουν το πρότυπο
- Σημειώστε ότι εξ' ορισμού εκτυπώνονται όλες οι γραμμές
  - έκτος και εάν η sed πάρει το όρισμα **-n** (silent mode, οπότε εκτυπώνονται μόνο οι γραμμές που ικανοποιούν το πρότυπο)

```
bash-3.1$ sed '/errors/p' example.sed
```

```
This is the first line of an example text.
```

```
It is a text with errors.
```

```
It is a text with errors.
```

```
Lots of errors.
```

```
Lots of errors.
```

```
So much errors, all these errors are making me sick.
```

```
So much errors, all these errors are making me sick.
```

```
This is a line not containing any errors.
```

```
This is the last line.
```

# Εκτύπωση (Print) με /p



Δεν εκτυπώνει τις γραμμές που επεξεργάζεται η sed.

Εκτύπωση των γραμμών που βρέθηκαν

```
bash-3.1$ sed -n '/errors/p' example.sed
```

```
It is a text with errors.
```

```
Lots of errors.
```

```
So much errors, all these errors are making me sick.
```

Προκαλεί την εκτύπωση των γραμμών που **δεν** έχουν βρεθεί

```
bash-3.1$ sed -n '/errors/!p' example.sed
```

```
This is the first line of an example text.
```

```
This is a line not containing any errors.
```

```
This is the last line.
```

Χωρίς p δεν θα εμφάνιζε τίποτα λόγω του -n (silent mode)

# Εκτύπωση γραμμών που μεταβλήθηκαν.

```
bash-3.1$ sed -n 's/errors/errors/gp' example.sed
```

```
It is a text with errors.
```

```
Lots of errors.
```

```
So much errors, all these errors are making me sick.
```

# Εκτύπωση (Print) με /p



- Περιορισμός σε μια γραμμή κειμένου

Περιορισμός  
εκτέλεσης της εντολής  
μόνο στη γραμμή 4

2<sup>η</sup> εμφάνιση του προτύπου στη  
γραμμή κειμένου που  
διαβάζεται

```
# Διόρθωση της δεύτερης εμφάνισης errors στη γραμμή 4.
bash-3.1$ sed '4 s/erors/errors/g2' example.sed
This is the first line of an example text.
It is a text with erors.
Lots of erors.
So much erors, all these errors are making me sick.
This is a line not containing any errors.
This is the last line.
```

# Εκτύπωση (Print) με /p



- Περιορισμός σε πεδίο γραμμών κειμένου
  - το πεδίο γραμμών ορίζεται τοποθετώντας ένα κόμμα (,) μεταξύ των αριθμών.

Περιορισμός  
εκτέλεσης της εντολής  
στις γραμμές 1 - 3

```
bash-3.1$ sed '1,3 s/erors/errors/g' example.sed
```

This is the first line of an example text.

It is a text with **errors.**

Lots of **errors.**

So much **erors**, all these **erors** are making me sick.

This is a line not containing any errors.

This is the last line.

# Περιορισμός Εκτέλεσης Pattern



- Περιορισμός εκτέλεσης του pattern σε γραμμές που περιέχουν συγκεκριμένο πρότυπο

```
bash-3.1$ cat example2.sed
```

```
#This is the first line of an example text.
```

```
#It is a text with errors.
```

```
#Lots of errors.
```

```
So much errors, all these errors are making me sick.
```

```
This is a line not containing any errors.
```

```
This is the last line.
```

└─ Περιορίζει την αντικατάσταση στις γραμμές που ξεκινούν από #

```
bash-3.1$ sed '/^#/ s/errors/errors/g' example2.sed
```

```
#This is the first line of an example text.
```

```
#It is a text with errors.
```

```
#Lots of errors.
```

```
So much errors, all these errors are making me sick.
```

```
This is a line not containing any errors.
```

```
This is the last line.
```



# Περιορισμός Εκτέλεσης Pattern



- Περιορισμός εκτέλεσης pattern μεταξύ START και STOP pattern ***sed '/START/,/STOP/ command'***

```
bash-3.1$ cat example3.sed
```

```
/* This is a comment
```

```
*
```

```
*
```

```
*/
```

```
This is the first line of an example text.
```

```
This is the last line.
```

*Substitute one or more  
characters with nothing  
Or use delete: /\.+/d (see  
next)*

```
bash-3.1$ sed '/\/*/,/\*\/ s/.\+/' example3.sed
```

Πρότυπο που δείχνει την αρχή-πρώτη γραμμή για εκτέλεση της sed εντολής

Πρότυπο που δείχνει το τέλος-τελευταία γραμμή εκτέλεσης της sed εντολής

```
This is the first line of an example text.
```

```
This is the last line.
```

Θα δούμε σε λίγο πώς να  
αφαιρέσουμε τις κενές γραμμές

# Διαγραφή (Delete) με /d



- εντολή: **d** (*delete* - διαγραφή)

## Παραδείγματα:

- **d** διαγράφει **ΟΛΕΣ** τις γραμμές
- **6d** διαγράφει τη **γραμμή 6**
- **/^\$/d** διαγράφει όλες τις **κενές γραμμές**
- **/^\./d** διαγράφει **όλες τις γραμμές** που **ξεκινούν με . (lite)**
- **1,10d** διαγράφει τις **γραμμές 1 -10**
- **1,/^\$/d** διαγράφει από τη 1<sup>η</sup> γραμμή μέχρι και την **πρώτη κενή γραμμή**
- **/^\$/, \$d** διαγράφει από την **πρώτη κενή γραμμή** μέχρι και τη **τελευταία γραμμή**
- **/^\$/, 10d** διαγράφει από την **πρώτη κενή γραμμή** μέχρι και την 10<sup>η</sup> γραμμή
- **/^ya\*y/,/[0-9]\$/d** διαγράφει από την πρώτη γραμμή που ξεκινά με το πρότυπο **ya\*y** μέχρι και την πρώτη γραμμή που τελειώνει με το πρότυπο **[0-9]**
  - **ab\*c** matches "ac", "abc", "abbc", "abbbc", and so on.
  - **a.\*c** matches "ac", "a<ANY>c", "a<ANY><ANY>c", "a<ANY><ANY><ANY>c", and so on.

# Διαγραφή (Delete) με /d



- Διαγραφή – παράδειγμα

```
bash-3.1$ cat example3.sed
```

```
/* This is a comment
```

```
*
```

```
*
```

```
*/
```

```
This is the first line of an example text.
```

```
This is the last line.
```

*Del pattern within  
START TO START..TO  
Del empty  
lines*

```
bash-3.1$ sed -e '/\//\*/,/\/\*\// s/.\+//' -e '/^$/d'  
example3.sed
```

```
This is the first line of an example text.
```

```
This is the last line.
```

# Διαγραφή (Delete) με /d



- Διαγραφή των **C comments** και των **tabs** που εμφανίζονται στην αρχή κάθε γραμμής.

```
bash-3.1$ cat example3spaces.sed
```

```
/* This is a comment
 *
 *
 */
```

```
This is the first line of an example text.
```

```
\t
```

```
This is the last line.
```

*Del words within /\* \*/*

*Del one or more  
space|tab*

*Del empty lines*

```
$ sed -e '/\/*/,/\*\/ s/.\|+//' -e 's/^[ \t]\+//' -e '/^$/d'
example3spaces.sed
```

```
This is the first line of an example text.
```

```
This is the last line.
```

Not same with '/^[  
\t]\+/' which

deletes the whole line

# Δομή Εντολής sed



- **sed <options>** # επιλογές εκτέλεσης
  - '[FROM,[TO]]** # περιορισμός εύρους εκτέλεσης  
(π.χ., εκτέλεση σε γραμμές 1,3)
  - [command]/** # εντολή εκτέλεσης π.χ., s
  - pattern** # πρότυπο αναζήτησης  
(απλή ή επεκταμένη κανονική έκφραση)
  - [/replacement]** # αντικατάσταση προτύπου  
(προαιρετικό)
  - [/option2]** # επιπλέον επιλογές  
(προαιρετικό)
  - /**
  - infilename** # όνομα αρχείου επεξεργασίας  
(εναλλακτικά από STDIN)

# Δομή Εντολής sed



- **<options>: Επιλογές Εκτέλεσης:**

Επιλογή	Επίδραση
<b>-e</b> <i>command</i>	<b>Πολλαπλές Επιλογές</b> Π.χ., -e 'cmd1' -e 'cmd2' ... -e 'cmdN'
<b>-f</b> SCRIPT-FILE	<b>Ανάγνωση Επιλογών από Αρχείο.</b> Π.χ., <pre># This file is named "sample.sed" # comments appear at the beginning s/color/colour/g s/flavor/flavour/g s/theater/theatre/g # Remember to chmod u+x sample.sed</pre>
<b>-n</b>	<b>Silent mode</b> (only print lines specified with the print command). By default each processed line is displayed on STDOUT.