

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

XML PARSING ΜΕ ΧΡΗΣΗ WEB WORKERS

Του φοιτητή
Μυλώση Δ. Κωνσταντίνου

Επιβλέπων καθηγητής
Κασκάλης Θεόδωρος

Εισαγωγή



Εντοπισμός προβλήματος

- ✿ Αύξηση απαιτήσεων μεταδιδόμενης πληροφορίας
- ✿ Εκθετική αύξηση όγκου πληροφορίας
- ✿ Ανάγκη ανάπτυξης νέων αλγορίθμων επεξεργασίας
- ✿ Ανάγκη για επεξεργασία δεδομένων σε πραγματικό χρόνο
- ✿ Ανάγκη για περισσότερο εύρος χρησιμοποιούμενου δικτύου
- ✿ Υπερφόρτωση διακομιστών
- ✿ Μη αξιοποίηση υπολογιστών χρηστών

Σημαντικότητα της μελέτης

- ❁ Ελάφρυνση διακομιστών
- ❁ Περισσότερη αξιοποίηση υπολογιστών χρηστών
- ❁ Βελτίωση εμπειρίας του χρήστη
- ❁ Εκμετάλλευση πολυνηματικής επεξεργασίας σε περιβάλλον περιηγητών
- ❁ Εξεύρεση στρατηγικών επεξεργασίας για βελτιστοποίηση

Πρότυπα ανταλλαγής δεδομένων



CSV



JSON



YAML



XML

```
<people>
  <person>
    <name>Kostas</name >
    <surname>Milo</surname>
    <age>26</age>
  </person>
</people>
```

Περιγητής - DOM



Κορμός διαδικτυακών εφαρμογών



Δενδροειδής δομή όλων των αντικειμένων μιας ιστοσελίδας

Parsing



Στάδια



Προεπεξεργασία, επικύρωση, συγχώνευση



Σειριακή, παράλληλη και υβριδική



Στατικό και δυναμικό Partitioning

Έρευνες



Skeleton



Accelerator



State Machine

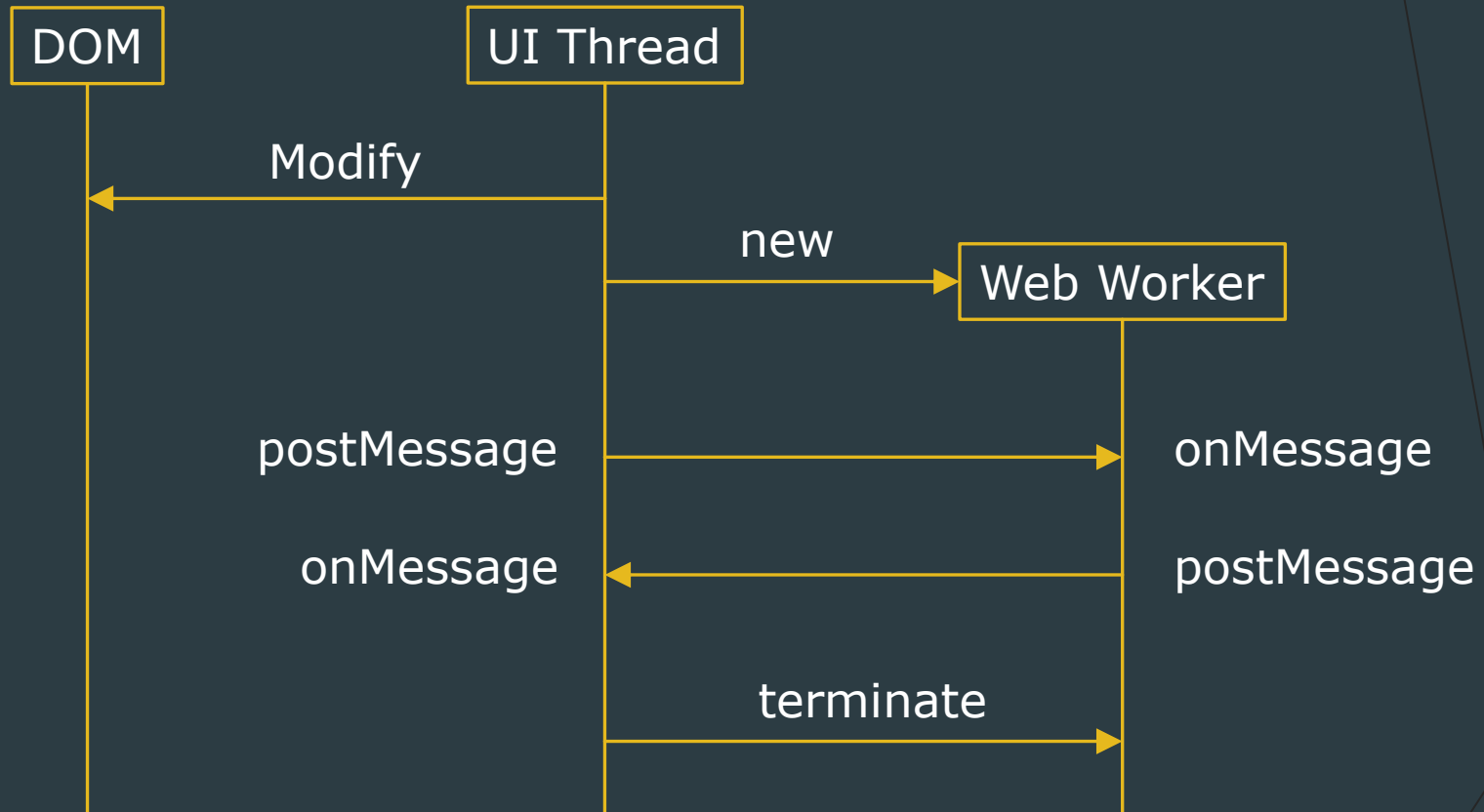


Work Stealing



Field-Programmable Gate Array

Web worker



Μεθοδολογία



Προεργασία



Ανάπτυξη προγράμματος



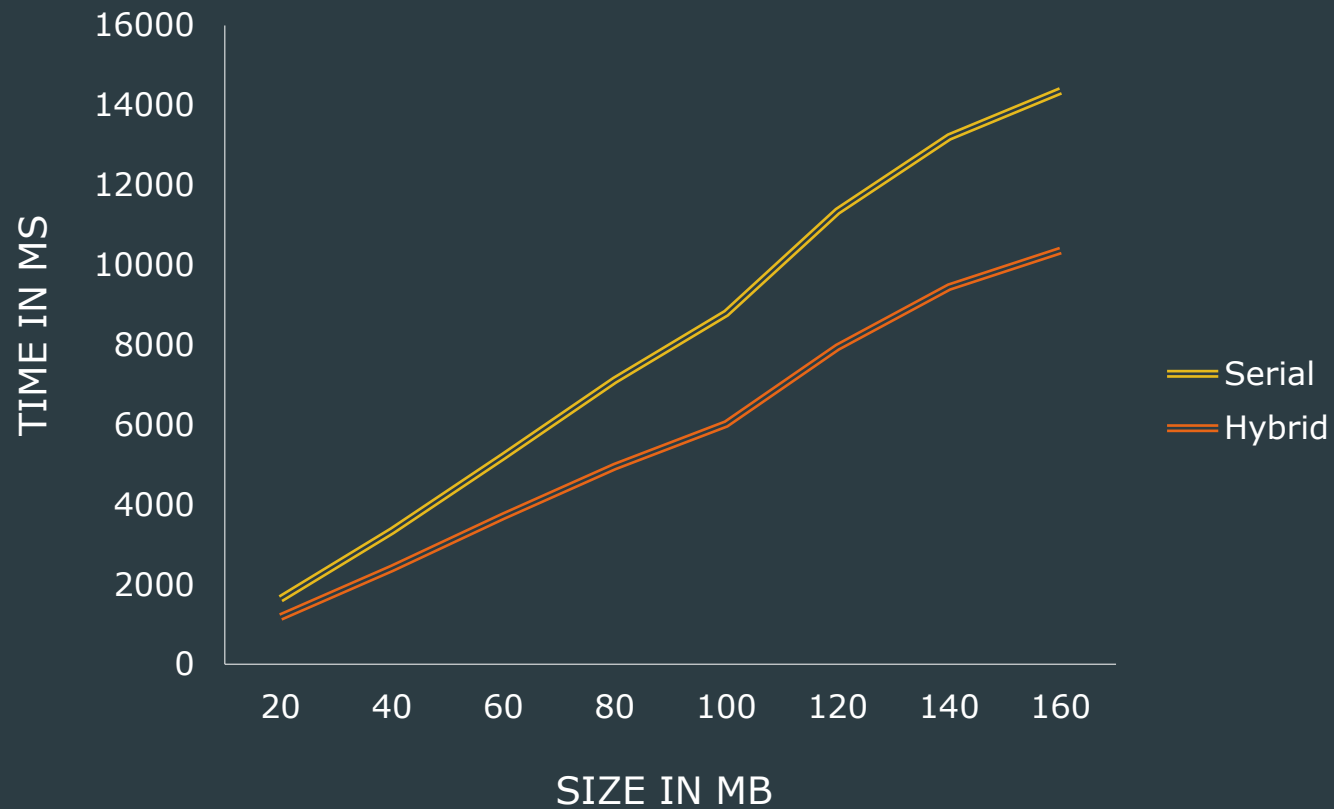
Εξοπλισμός

Κώδικας

```
var ww = new Worker('parseChunk.js', { type: "module" });
ww.postMessage({ dd: { cnt: (content, [content]), lvl0Rsize: 'lvl',
| | | | |      widx: widx++, tag: tag, startsArr: startsArr[i] } }));
ww.onmessage = function (e) {
|   processMessage(e, startsArr.length);
}
}
```

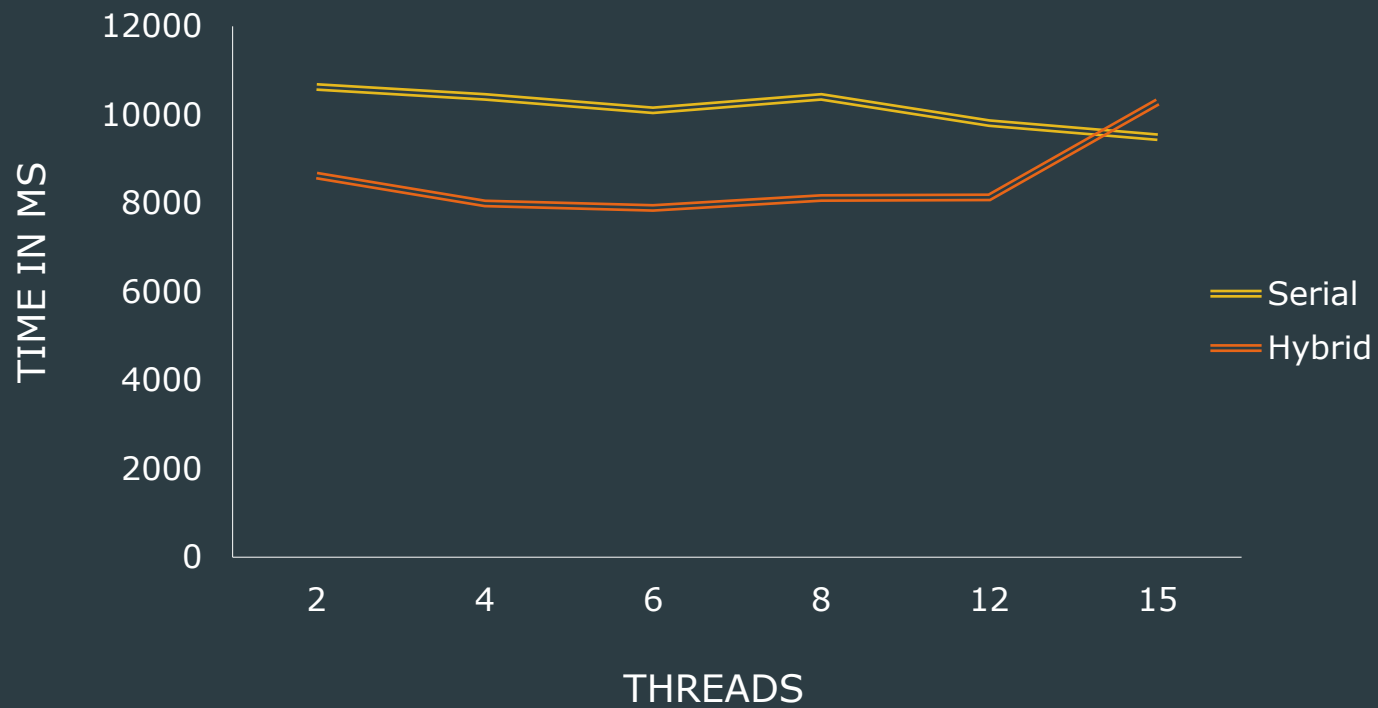
Αποτελέσματα

8 THREADS



Αποτελέσματα

120 MB



Συμπεράσματα

- ⚙ Παράγοντες
- ⚙ Μεγάλος όγκος πληροφορίας - μεγάλο όφελος
- ⚙ Εμφανές από τα πρώτα δύο νήματα
- ⚙ Μέχρι περίπου 80% των νημάτων, σταθερό κέρδος
- ⚙ Όριο στη χρήση web workers
- ⚙ Ταχύτερη επεξεργασία, βελτίωση εμπειρίας χρήστη

Όρια και περιορισμοί



Περιορισμοί συστήματος



Προσομοίωση περιβάλλον χρήστη και διακομιστή



Χρήση κατασκευασμένων αρχείων



Δυσκολία δοκιμής σε κατανεμημένο σύστημα



Λίγες μελέτες για Web Workers

Μελλοντικές προτάσεις

- ⚙ Εξολοκλήρου παραλληλισμός
- ⚙ Παράλληλη επεξεργασία σε πραγματικό χρόνο
- ⚙ Παράλληλη επεξεργασία με κατανεμημένο τρόπο
- ⚙ Αξιοποίηση σύγχρονων καρτών γραφικών (web GPU)

Σας ευχαριστώ