

Università degli studi di Modena e Reggio Emilia

Dipartimento di Scienze Fisiche, Informatiche e Matematiche

Corso di laurea in Informatica A.A. 2016-2017

Improving the BFQ I/O Scheduler

From porting to Android to analysis and boosting of performance

Candidato: Luca Miccio

Relatore: Dott. Paolo Valente

Correlatore: Dott. Mauro Andreolini



I/O Scheduler

- Sceglie l'ordine di servizio delle richieste di I/O
- Obiettivi
 - Alto I/O throughput, bassa latenza, alta reattività, fairness
- Algoritmi attuali in Linux
 - BFQ, Noop, CFQ, Deadline

Budget Fair Queueing

- Scheduler gerarchico
 - Scelta processo da servire
 - Scelta della quantità di servizio in base al budget
- Pro
 - Bassa latenza
 - Alto throughput

Problemi

● Funzionali

- BFQ non presente in Android

● Prestazionali

- Basso **throughput** su dispositivi flash storage
- Basso **latenza** non sempre rispettata
- Throughput max. sostenibile più basso degli altri algoritmi

Approccio risolutivo

- Integrazione BFQ in Android
- Analisi prestazioni e profiling
- Risoluzione problemi incontrati
 - Correzione meccanismo idling
 - Disattivazione write-back throttling
 - Overhead

A decorative background featuring a network diagram with nodes and connecting lines. Some nodes are highlighted with blue circles or dots. The diagram is composed of grey lines and dots, with blue accents. It is positioned in the top-left and bottom-right corners of the slide.

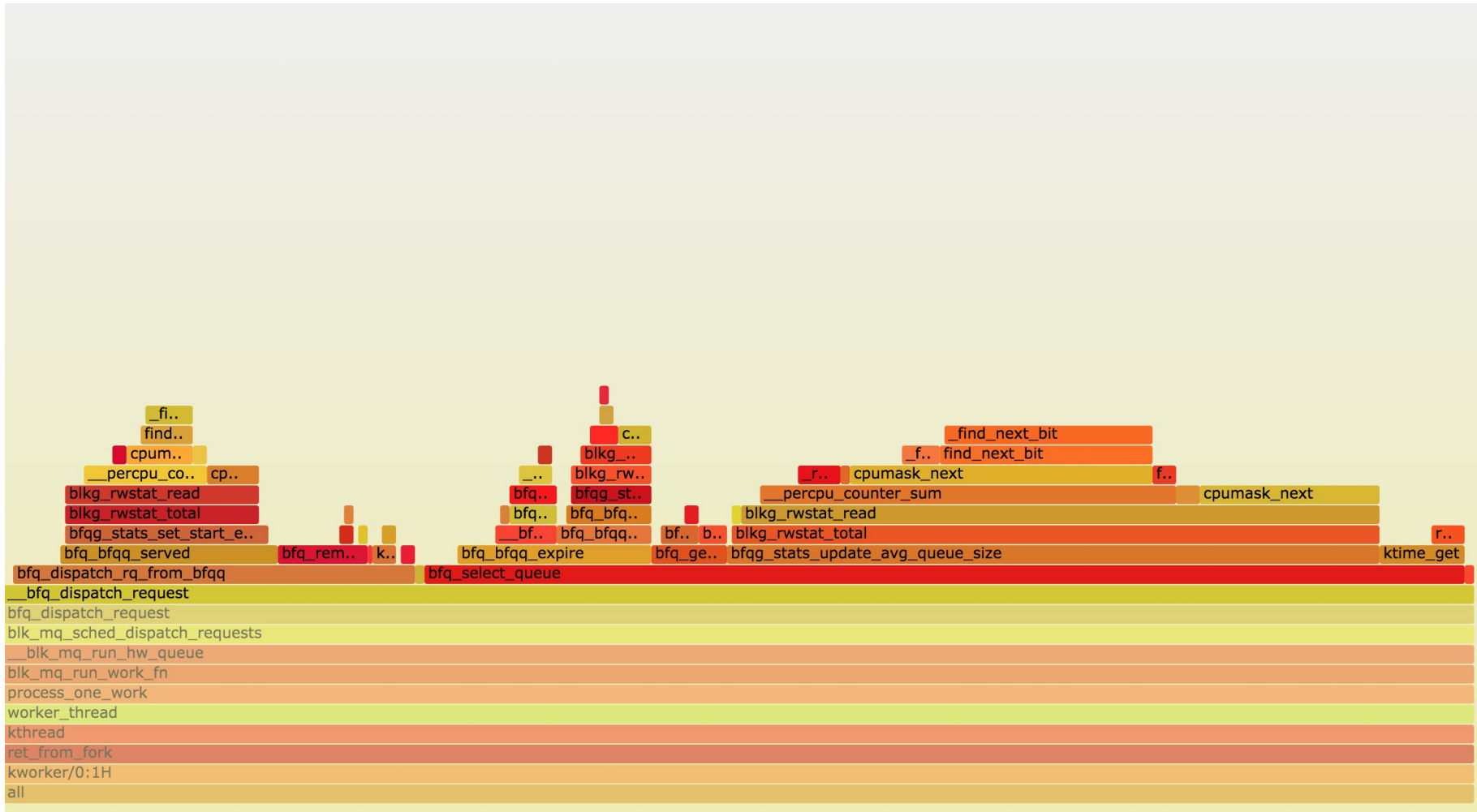
Demo

“Se un’immagine vale più di $e^{6,90778}$ parole,
un video è un vero e proprio tesoro”

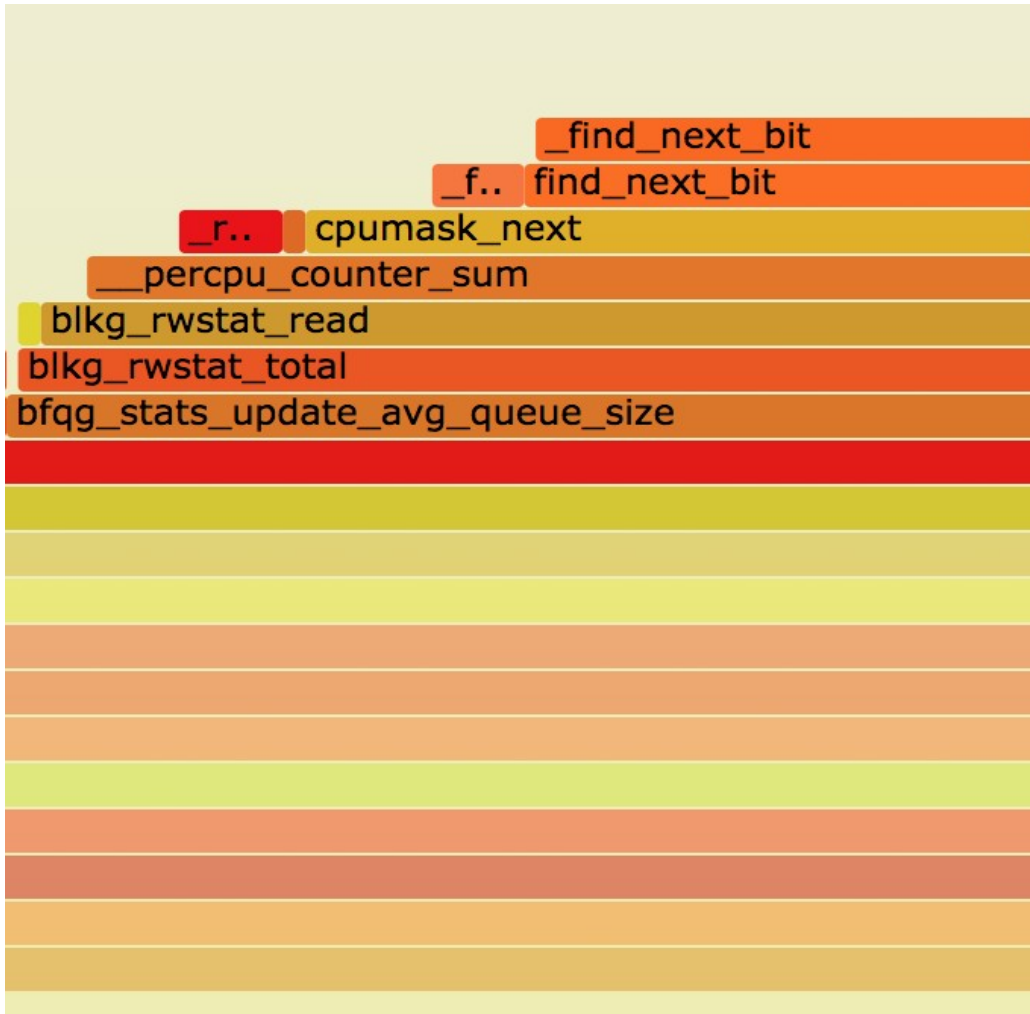
Aumento throughput

- Analisi overhead BFQ
 - Obiettivo: aumento throughput max.
 - Strumenti: Perf, eBPF, Flame Graph
- Benchmark: IOSpeed
 - Implementato da zero (C, BASH)
 - Generatore carico sequenziale e random
 - Distribuzione carico su N thread

Flame Graph kernel

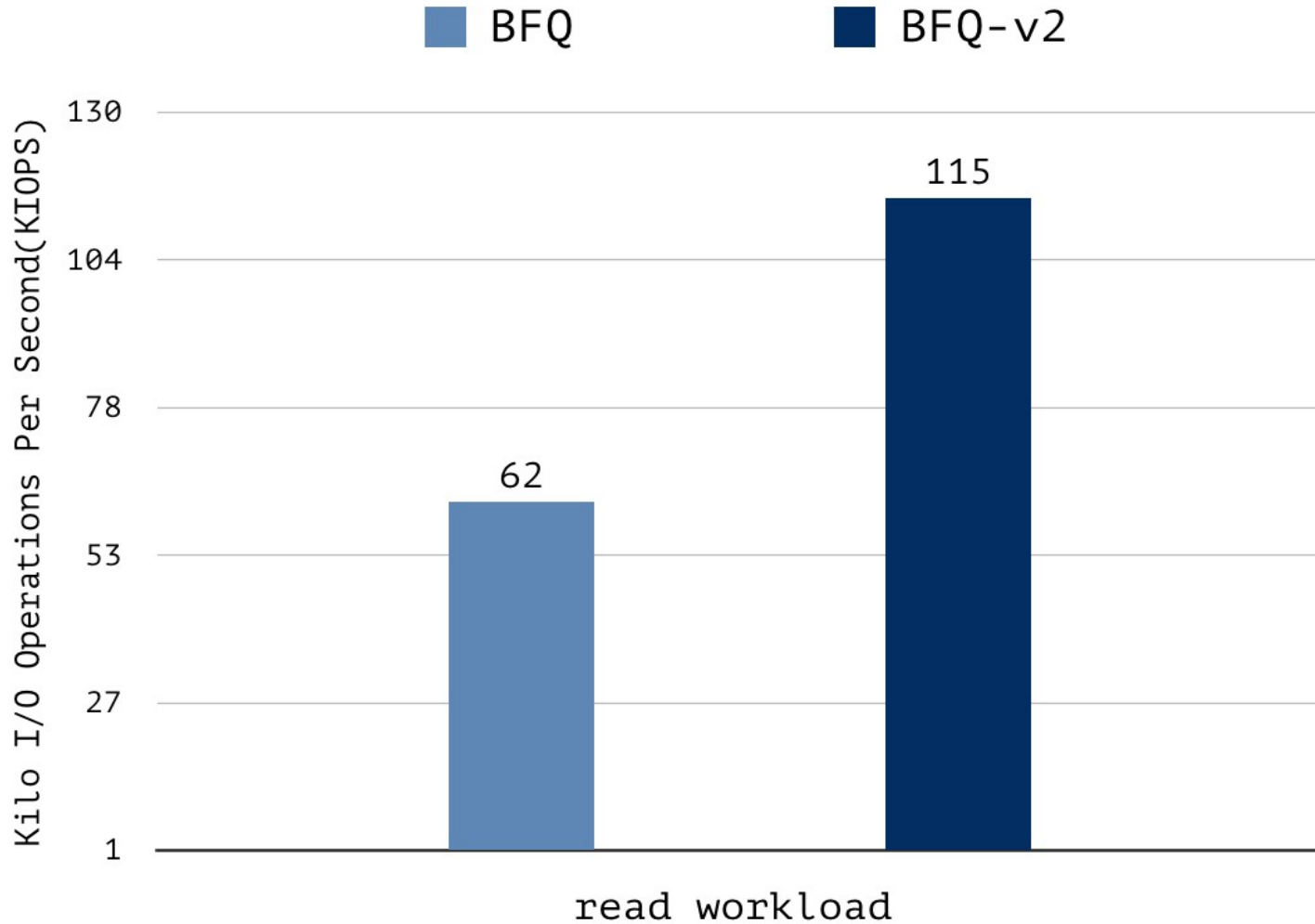


Overhead: *_find_next_bit*



- Costo: circa il 50% del tempo della funzione chiamante

Aumento throughput - ARM



Aumento throughput - Desktop



`"Honestly, this code is too ugly
to live..."`



“...but the rest of the pull looks ok, so I’ll take it anyway”



Conclusioni

- Integrazione con Android
- Aumento throughput fino al **125%** su flash storage
- Aumento throughput massimo sostenibile dallo scheduler fino al **90%**



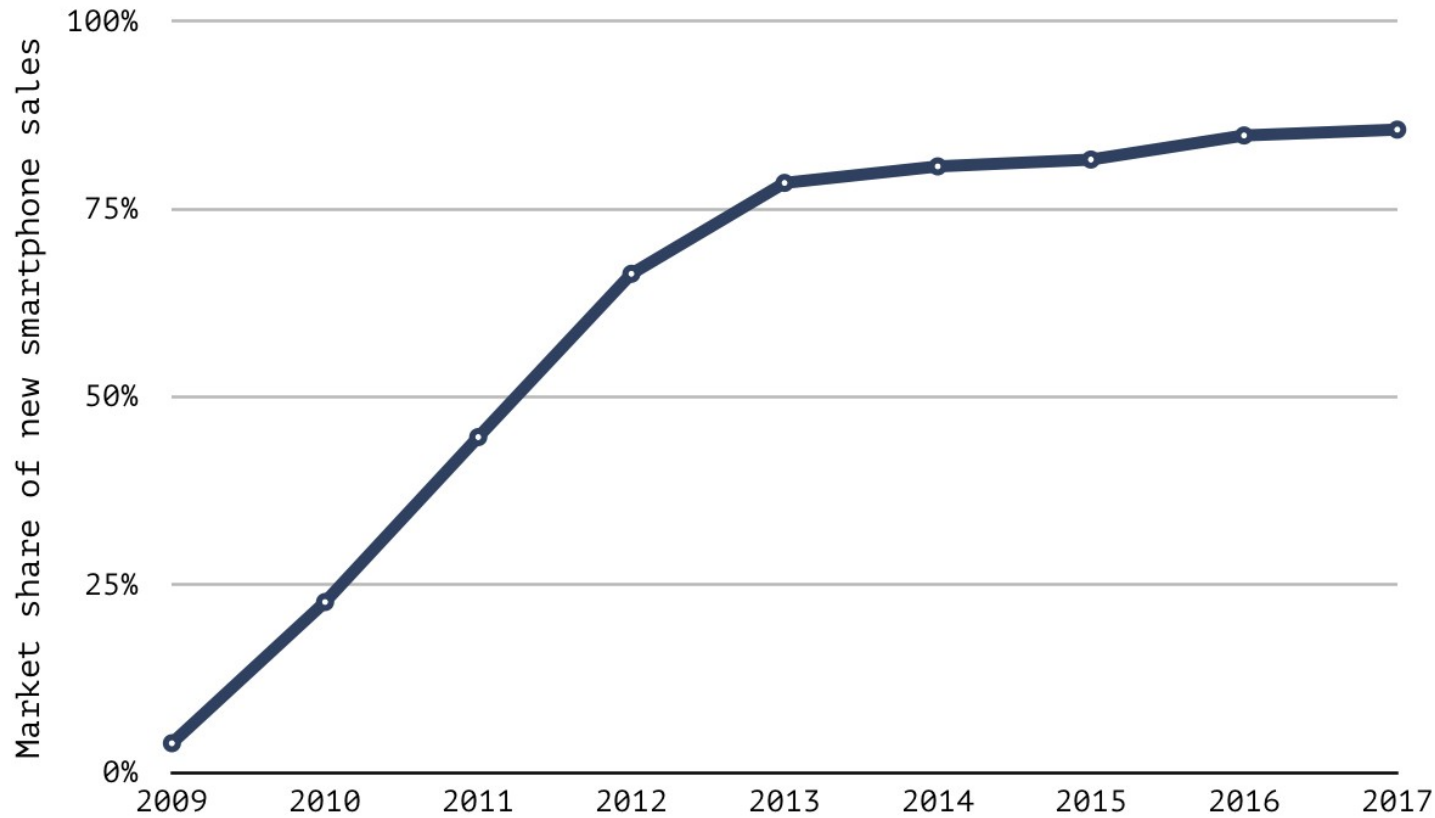
**Grazie
dell'attenzione**

A decorative background featuring a network diagram. It consists of numerous nodes, represented by circles of varying sizes and shades of gray, connected by thin, light gray lines. Some nodes are highlighted with a blue outline, and a few are solid blue. The network is more densely packed in the top-left and bottom-right corners, with the text 'Backup slide' centered in the middle.

Backup slide

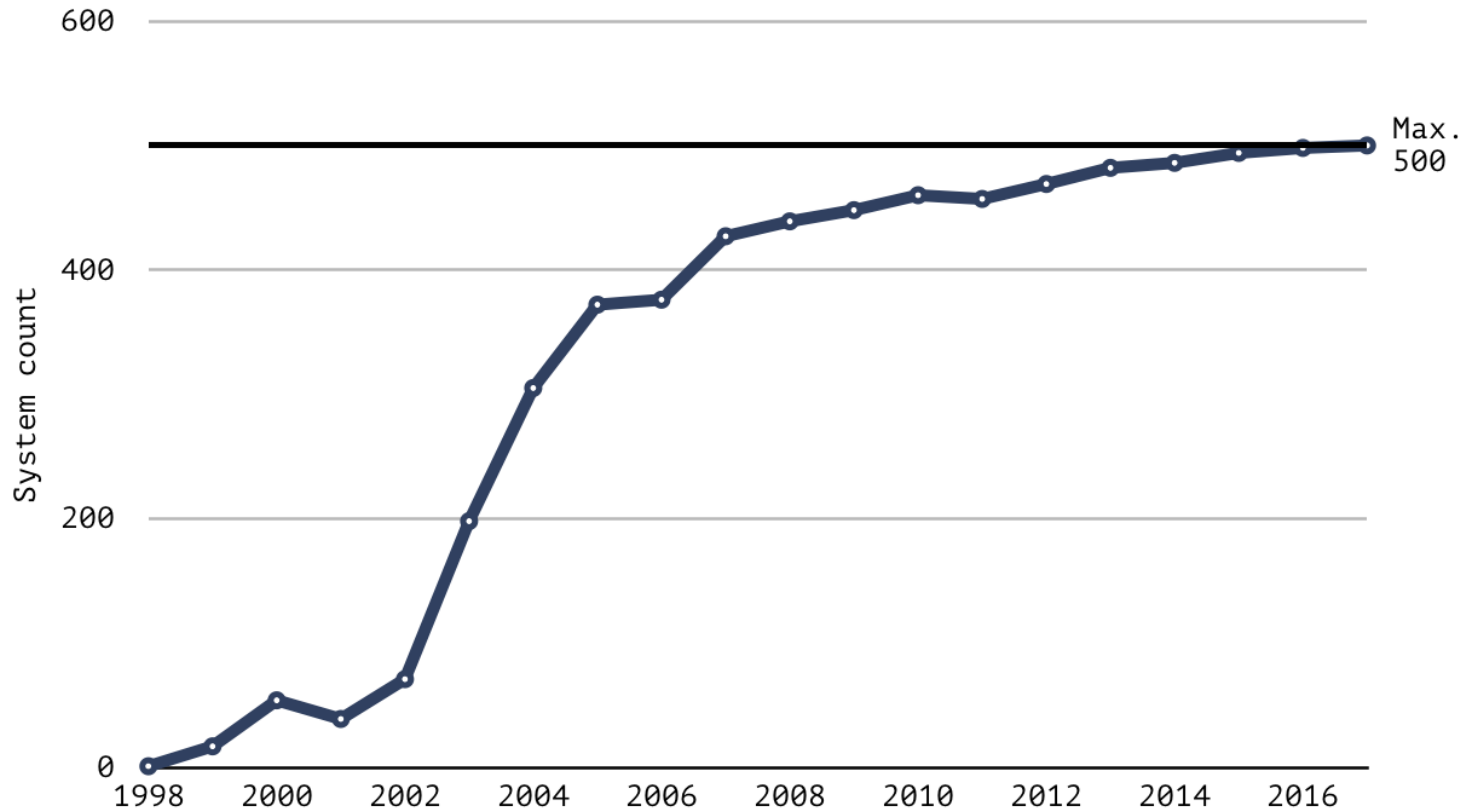
Linux in Android

Android OS market share



Linux in HPC

Linux systems in Top 500 HPC

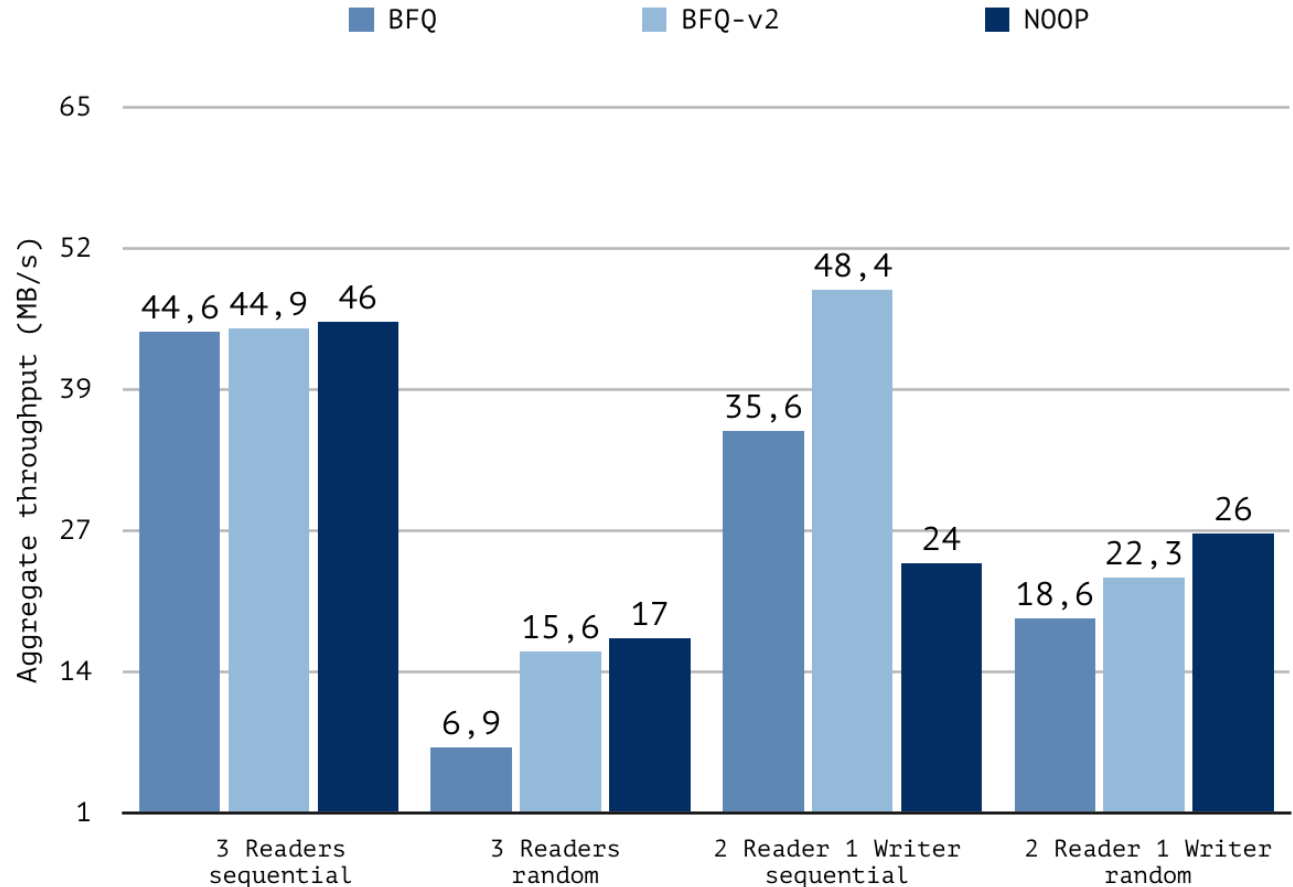


Write-back throttling

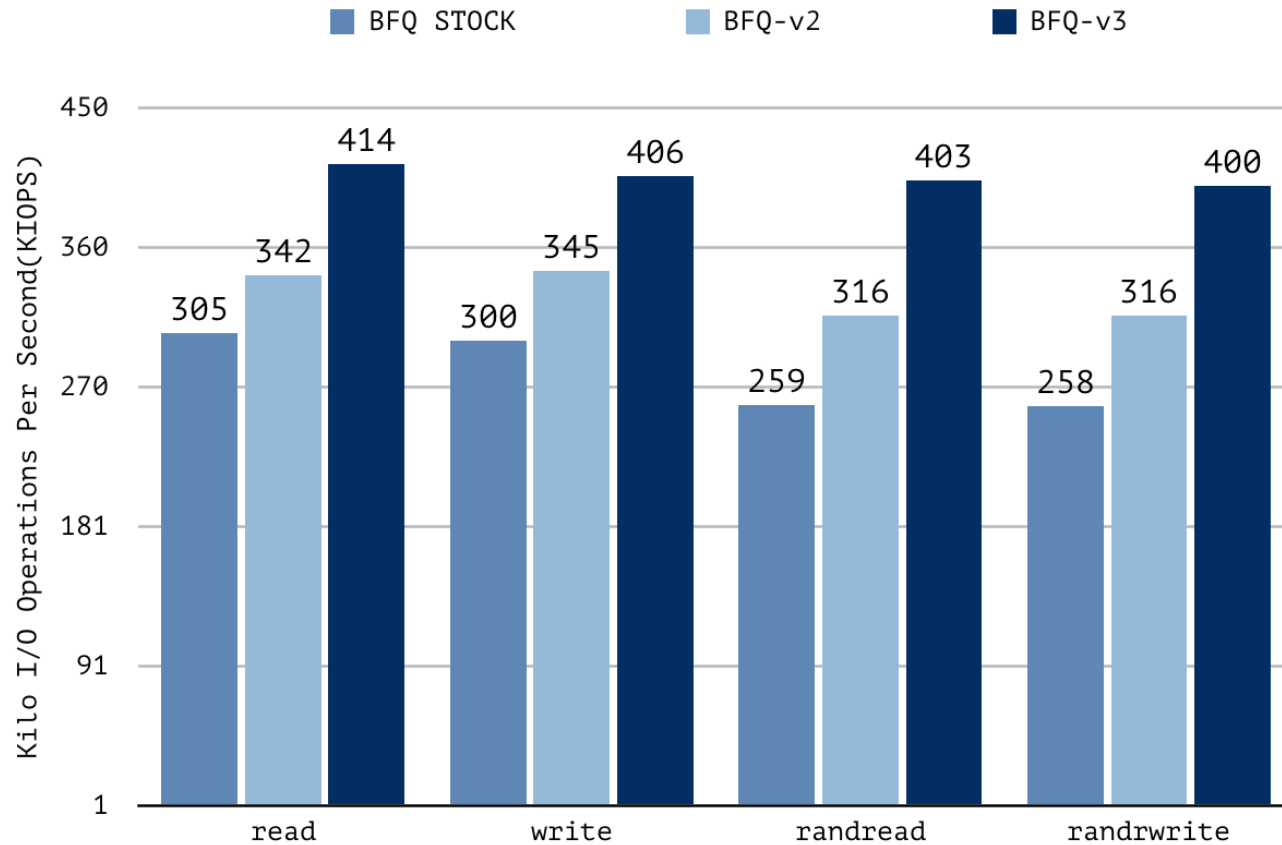
- Write-back throttling vs. euristiche BFQ
- Soluzione necessaria: disattivazione write-back throttling
- Nessuna perdita delle garanzie di BFQ
- 1 patch in mainline dalla versione 4.14 del kernel Linux

Boost throughput - risultati

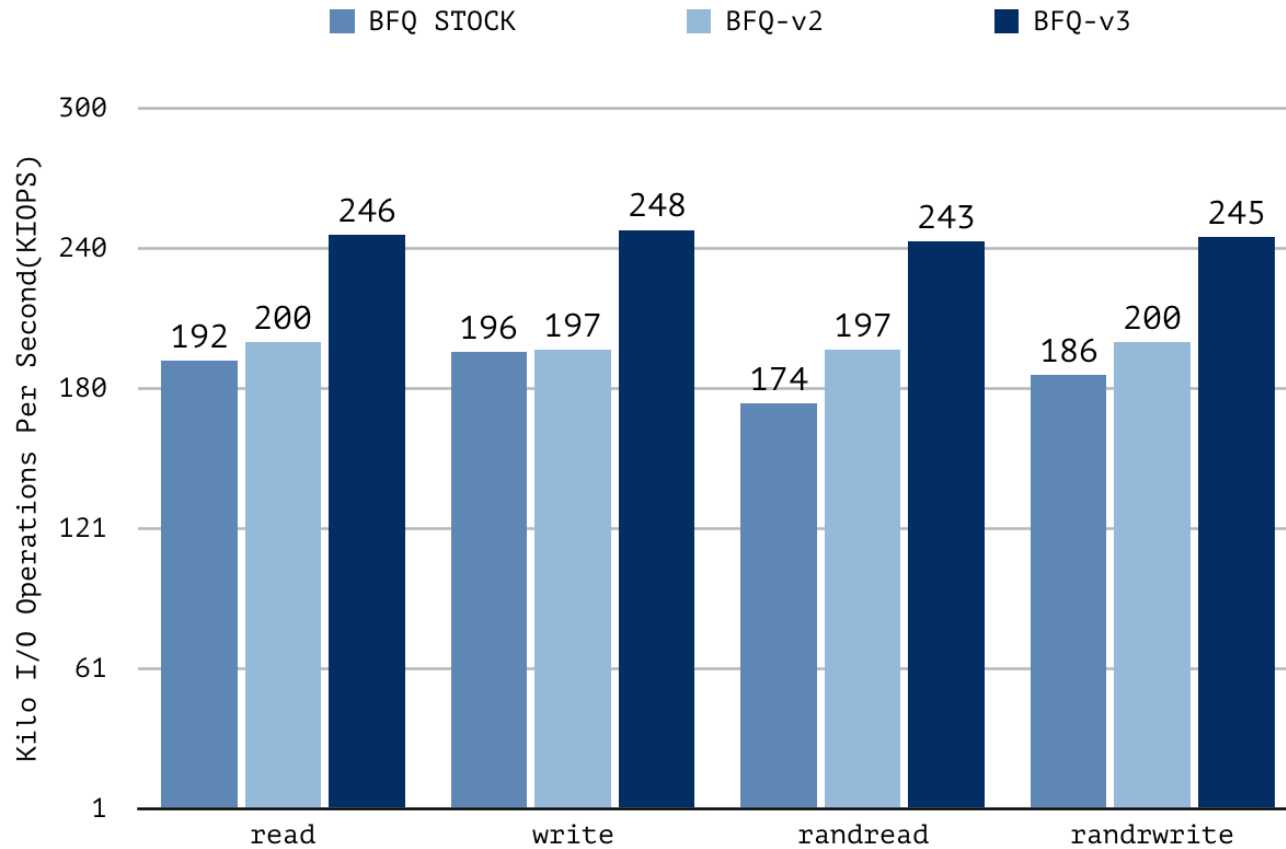
- 2 patch in mainline dalla versione 4.13 del kernel Linux



Risultati - INTEL



Risultati - AMD



Risultati - Hikey Board

