

Monitoring Analysis

omitted

2024-04-05

Total analyzed commits

```
nrow(monitoring_data)
```

```
## [1] 5899
```

Total time to execute all commits analysis

```
total_in_seconds <- sum(monitoring_data$duration_in_seconds)
total_days <- total_in_seconds / (60*60*24)
total_days
```

```
## [1] 59.87519
```

Execution duration of commits

```
monitoring_data %>%
  select(duration_formatted) %>%
  summary()
```

```
## duration_formatted
## Min. :2024-04-22 00:00:41.55
## 1st Qu.:2024-04-22 00:02:27.47
## Median :2024-04-22 00:06:32.00
## Mean :2024-04-22 00:06:18.07
## 3rd Qu.:2024-04-22 00:08:51.38
## Max. :2024-04-22 06:20:19.39
```

Execution duration of commits by repo

```

monitoring_data %>%
  group_by(repo) %>%
  pivot_wider(names_from = repo, values_from = c("duration_formatted")) %>%
  select("accumulo", "cayenne", "commons-collections", "commons-io", "commons-lang", "helix", "httpcomponents-client", "httpcomponents-struts-core", "httpcomponents-zookeeper", "maven-surefire", "opennlp", "wicket")
summary()

```

##	accumulo		cayenne
##	Min. :2024-04-22 00:05:30.20		Min. :2024-04-22 00:03:00.17
##	1st Qu.:2024-04-22 00:06:58.23		1st Qu.:2024-04-22 00:03:49.05
##	Median :2024-04-22 00:08:19.11		Median :2024-04-22 00:04:06.30
##	Mean :2024-04-22 00:08:47.95		Mean :2024-04-22 00:04:17.55
##	3rd Qu.:2024-04-22 00:10:09.04		3rd Qu.:2024-04-22 00:04:25.73
##	Max. :2024-04-22 00:20:27.66		Max. :2024-04-22 00:19:08.52
##	NA's :2953		NA's :5759
##	commons-collections		commons-io
##	Min. :2024-04-22 00:01:08.41		Min. :2024-04-22 00:00:41.55
##	1st Qu.:2024-04-22 00:01:15.11		1st Qu.:2024-04-22 00:01:03.00
##	Median :2024-04-22 00:01:18.18		Median :2024-04-22 00:01:10.96
##	Mean :2024-04-22 00:01:20.54		Mean :2024-04-22 00:01:10.50
##	3rd Qu.:2024-04-22 00:01:25.75		3rd Qu.:2024-04-22 00:01:14.70
##	Max. :2024-04-22 00:01:45.30		Max. :2024-04-22 00:01:51.75
##	NA's :5816		NA's :5664
##	commons-lang		helix
##	Min. :2024-04-22 00:01:11.76		Min. :2024-04-22 00:02:01.16
##	1st Qu.:2024-04-22 00:01:22.26		1st Qu.:2024-04-22 00:02:31.40
##	Median :2024-04-22 00:01:28.88		Median :2024-04-22 00:02:54.90
##	Mean :2024-04-22 00:01:35.70		Mean :2024-04-22 00:07:20.25
##	3rd Qu.:2024-04-22 00:01:39.55		3rd Qu.:2024-04-22 00:13:31.50
##	Max. :2024-04-22 00:02:53.32		Max. :2024-04-22 06:20:19.39
##	NA's :5571		NA's :4989
##	httpcomponents-client		maven-surefire
##	Min. :2024-04-22 00:01:02.61		Min. :2024-04-22 00:02:35.27
##	1st Qu.:2024-04-22 00:01:08.94		1st Qu.:2024-04-22 00:02:44.84
##	Median :2024-04-22 00:01:15.72		Median :2024-04-22 00:02:50.78
##	Mean :2024-04-22 00:01:22.78		Mean :2024-04-22 00:03:06.88
##	3rd Qu.:2024-04-22 00:01:30.86		3rd Qu.:2024-04-22 00:03:00.57
##	Max. :2024-04-22 00:12:42.53		Max. :2024-04-22 00:05:43.16
##	NA's :5594		NA's :5844
##	opennlp		struts
##	Min. :2024-04-22 00:01:05.98		Min. :2024-04-22 00:02:16.51
##	1st Qu.:2024-04-22 00:02:08.06		1st Qu.:2024-04-22 00:02:33.50
##	Median :2024-04-22 00:02:17.51		Median :2024-04-22 00:02:41.18
##	Mean :2024-04-22 00:02:20.69		Mean :2024-04-22 00:02:55.38
##	3rd Qu.:2024-04-22 00:02:26.25		3rd Qu.:2024-04-22 00:03:12.41
##	Max. :2024-04-22 00:04:54.98		Max. :2024-04-22 00:10:30.79
##	NA's :5689		NA's :5473
##	wicket		zookeeper
##	Min. :2024-04-22 00:03:06.64		Min. :2024-04-22 00:01:24.05
##	1st Qu.:2024-04-22 00:03:38.10		1st Qu.:2024-04-22 00:01:27.50
##	Median :2024-04-22 00:03:51.43		Median :2024-04-22 00:01:35.00
##	Mean :2024-04-22 00:03:52.19		Mean :2024-04-22 00:01:34.81
##	3rd Qu.:2024-04-22 00:04:02.02		3rd Qu.:2024-04-22 00:01:36.78
##	Max. :2024-04-22 00:06:08.21		Max. :2024-04-22 00:02:27.00

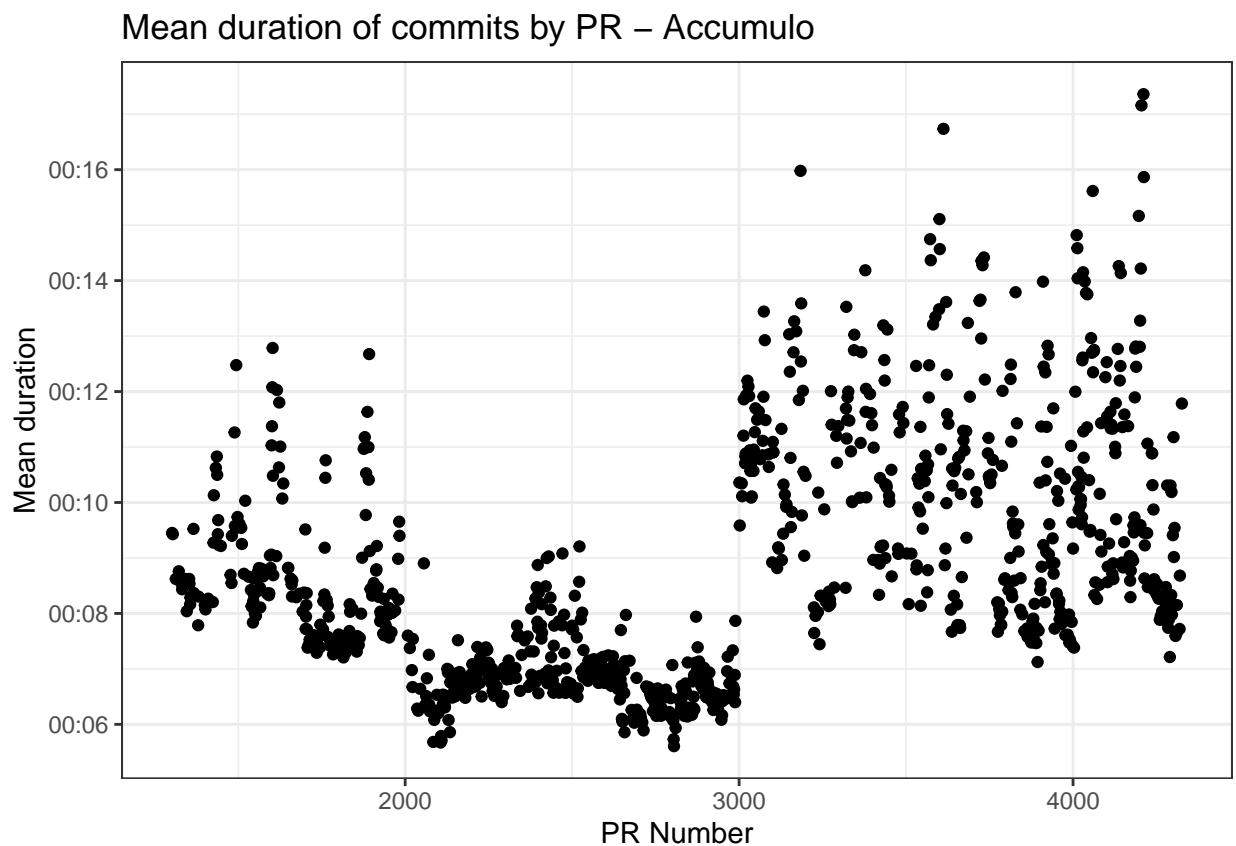
NA's :5760

NA's :5777

Mean duration of commits by PR - Accumulo

PRs with pr_number ≥ 2000 and pr_number < 3000 was executed on Intel Core i5-12500H cores, while the rest on Ryzen 5 5600G cores.

```
monitoring_data %>%
  filter(repo == "accumulo") %>%
  group_by(pr_number) %>%
  summarise(mean_duration = mean(duration_formatted)) %>%
  ggplot(aes(y=mean_duration, x=pr_number)) +
  geom_point() +
  labs(title = "Mean duration of commits by PR - Accumulo",
       x = "PR Number",
       y = "Mean duration")
```



Mean duration of commits by PR - Helix

PRs with pr_number ≤ 2549 was executed on Intel Core i5-12500H cores, while the rest on Ryzen 5 5600G cores.

```
monitoring_data %>%
  filter(repo == "helix") %>%
  group_by(pr_number) %>%
  summarise(mean_duration = mean(duration_formatted)) %>%
  ggplot(aes(y=mean_duration, x=pr_number)) +
  geom_point() +
  labs(title = "Mean duration of commits by PR - Helix",
       x = "PR Number",
       y = "Mean duration")
```

