

Cloud Lab 5

Traffic Shifting:

Here I am running this command to route all traffic to v1

```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl apply -f samples/bookinfo/networking/virtual-service-all-v1.yaml
```

```
virtualservice.networking.istio.io/productpage created
virtualservice.networking.istio.io/reviews created
virtualservice.networking.istio.io/ratings created
virtualservice.networking.istio.io/details created
```

I then got it back up and running locally on my browser

BookInfo Sample [Sign In](#)

The Comedy of Errors

Wikipedia Summary: The Comedy of Errors is one of **William Shakespeare's** early plays. It is his shortest and one of his most farcical comedies, with a major part of the humour coming from slapstick and mistaken identity, in addition to puns and word play.

[Learn more about Istio →](#)

Book Details

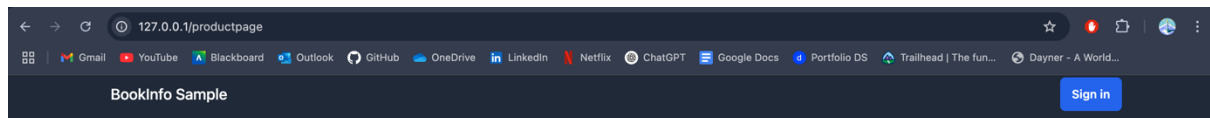
ISBN-10	Publisher	Pages	Type	Language
1234567890	PublisherA	200	paperback	English

Book Reviews

Then I ran this, now the stars will show up 50% of the time because they are split between v1 and v3

```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl apply -f samples/bookinfo/networking/virtual-service-reviews-50-v3.yaml
```

```
virtualservice.networking.istio.io/reviews configured
aaron@Aarons-MacBook-Air istio-1.25.0 %
```



The Comedy of Errors

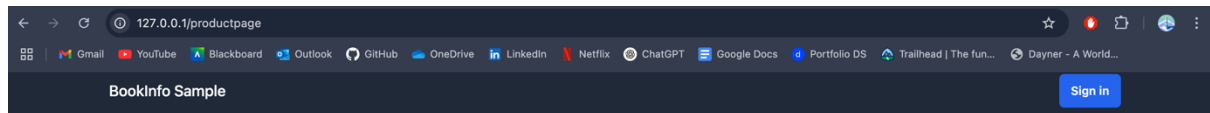
Wikipedia Summary: The Comedy of Errors is one of **William Shakespeare's** early plays. It is his shortest and one of his most farcical comedies, with a major part of the humour coming from slapstick and mistaken identity, in addition to puns and word play.

[Learn more about Istio](#) →

Book Details

ISBN-10	Publisher	Pages	Type	Language
1234567890	PublisherA	200	paperback	English

Book Reviews



The Comedy of Errors

Wikipedia Summary: The Comedy of Errors is one of **William Shakespeare's** early plays. It is his shortest and one of his most farcical comedies, with a major part of the humour coming from slapstick and mistaken identity, in addition to puns and word play.

[Learn more about Istio](#) →

Book Details

ISBN-10	Publisher	Pages	Type	Language
1234567890	PublisherA	200	paperback	English

Book Reviews

★★★★★

★★★★☆

Now I am routing 100% of the traffic to v3 so the stars will be there every time

```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl apply -f samples/bookinfo/networking/virtual-service-reviews-v3.yaml
```

```
virtualservice.networking.istio.io/reviews configured
```

```
aaron@Aarons-MacBook-Air istio-1.25.0 %
```

127.0.0.1/productpage

BookInfo Sample

Sign in

The Comedy of Errors

Wikipedia Summary: The Comedy of Errors is one of **William Shakespeare's** early plays. It is his shortest and one of his most farcical comedies, with a major part of the humour coming from slapstick and mistaken identity, in addition to puns and word play.

[Learn more about Istio →](#)

Book Details

ISBN-10	Publisher	Pages	Type	Language
1234567890	PublisherA	200	paperback	English

Book Reviews

★★★★★

★★★★☆

Request Routing

Here I am routing everything to v1

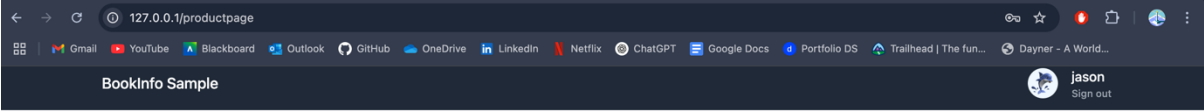
```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl apply -f samples/bookinfo/networking/virtual-service-all-v1.yaml

virtualservice.networking.istio.io/productpage unchanged
virtualservice.networking.istio.io/reviews configured
virtualservice.networking.istio.io/ratings unchanged
virtualservice.networking.istio.io/details unchanged
```

Here I am enabling user based routing as can be seen this is for users with the name jason

```
destination":{"host":"reviews","subset":"v2"}}}],{"route":[{"destination":{"host":
:"reviews","subset":"v1"}}]]]]}
creationTimestamp: "2025-03-26T11:48:56Z"
generation: 9
name: reviews
namespace: default
resourceVersion: "297926"
uid: 7b1d7533-d238-4356-9b1b-cfb64ac08af9
spec:
  hosts:
  - reviews
  http:
  - match:
    - headers:
      end-user:
        exact: jason
    route:
    - destination:
      host: reviews
      subset: v2
    - route:
      - destination:
        host: reviews
        subset: v1
```

As can be seen in the screenshots below when I am logged in as a user that isn't jason it doesn't show the stars but when I am jason they appear



The Comedy of Errors

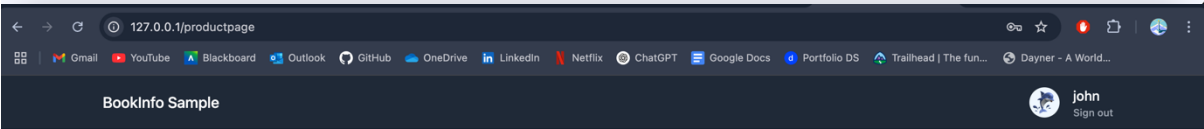
Wikipedia Summary: The Comedy of Errors is one of **William Shakespeare's** early plays. It is his shortest and one of his most farcical comedies, with a major part of the humour coming from slapstick and mistaken identity, in addition to puns and word play.

[Learn more about Istio](#) →

Book Details

ISBN-10	Publisher	Pages	Type	Language
1234567890	PublisherA	200	paperback	English

Book Reviews



The Comedy of Errors

Wikipedia Summary: The Comedy of Errors is one of **William Shakespeare's** early plays. It is his shortest and one of his most farcical comedies, with a major part of the humour coming from slapstick and mistaken identity, in addition to puns and word play.

[Learn more about Istio](#) →

Book Details

ISBN-10	Publisher	Pages	Type	Language
1234567890	PublisherA	200	paperback	English

Book Reviews

Fault Injection

Here I am creating a fault injection rule to delay traffic for users called jason

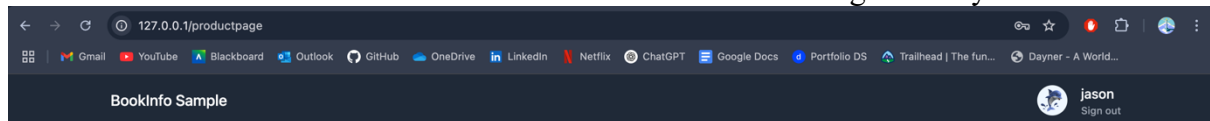
```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl apply -f samples/bookinfo/networking/virtual-service-ratings-test-delay.yaml

virtualservice.networking.istio.io/ratings configured
aaron@Aarons-MacBook-Air istio-1.25.0 %
```

We can see its working here with the fixed delay spec:

```
hosts:
- ratings
http:
- fault:
    delay:
      fixedDelay: 7s
      percentage:
        value: 100
```

This took about 7 seconds to load which indicates that it's working correctly.



The Comedy of Errors

Wikipedia Summary: The Comedy of Errors is one of **William Shakespeare's** early plays. It is his shortest and one of his most farcical comedies, with a major part of the humour coming from slapstick and mistaken identity, in addition to puns and word play.

[Learn more about Istio →](#)

Book Details

ISBN-10	Publisher	Pages	Type	Language
1234567890	PublisherA	200	paperback	English

Error fetching product reviews

Sorry, product reviews are currently unavailable for this book.

Then I am going to use the following command to create a fault injection rule to send a HTTP abort for users called Jason:

```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl apply -f samples/bookinfo/networking/virtual-service-ratings-test-abort.yaml

virtualservice.networking.istio.io/ratings configured
aaron@Aarons-MacBook-Air istio-1.25.0 %
```

As can be seen here this gives the user Jason a 500 fault:

```
spec:
  hosts:
    - ratings
  http:
    - fault:
        abort:
          httpStatus: 500
          percentage:
            value: 100
```

Now when logged in as Jason you can't see the reviews but when I log in as any other user they are visible

The screenshot shows a web browser at 127.0.0.1/productpage. The page title is "BookInfo Sample" and the user is logged in as "tom". The main content is for "The Comedy of Errors". Below the book details, there is a "Book Reviews" section. It displays two reviews: "An extremely entertaining play by Shakespeare. The slapstick humour is refreshing!" by Reviewer1 and "Absolutely fun and entertaining. The play lacks thematic depth when compared to other plays by Shakespeare." by Reviewer2. Both reviews are attributed to "Reviews served by: reviews-v1-848b87496f-6rsj6".

The screenshot shows the same web browser at 127.0.0.1/productpage, but the user is now logged in as "jason". The page content is identical to the previous screenshot, but the "Book Reviews" section now displays "Ratings service is currently unavailable" in red text for both review slots.

Circuit Breaking

First of all I am creating a destination rule to apply circuit breaking settings when calling the httpbin service

```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl apply -f - <<EOF
apiVersion: networking.istio.io/v1
kind: DestinationRule
metadata:
  name: httpbin
spec:
  host: httpbin
  trafficPolicy:
    connectionPool:
      tcp:
        maxConnections: 1
      http:
        http1MaxPendingRequests: 1
        maxRequestsPerConnection: 1
    outlierDetection:
      consecutive5xxErrors: 1
      interval: 1s
      baseEjectionTime: 3m
      maxEjectionPercent: 100
EOF
destinationrule.networking.istio.io/httpbin created
aaron@Aarons-MacBook-Air istio-1.25.0 %
```

Now I am going to inject the client with the istio sidecar proxy

```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl apply -f samples/httpbin/sample-
client/fortio-deploy.yaml

service/fortio created
deployment.apps/fortio-deploy created
```


Now I'm logging into client pod and using the fortio tool to use httpbin

```
aaron@Aarons-MacBook-Air istio-1.25.0 % export FORTIO_POD=$(kubectl get pods -l app=fortio -o 'jsonpath={.items[0].metadata.name}')
kubectl exec "$FORTIO_POD" -c fortio -- /usr/bin/fortio curl -quiet http://httpbin:8000/get

HTTP/1.1 200 OK
access-control-allow-credentials: true
access-control-allow-origin: *
content-type: application/json; charset=utf-8
date: Wed, 26 Mar 2025 14:16:18 GMT
content-length: 621
x-envoy-upstream-service-time: 17
server: envoy

{
  "args": {},
  "headers": {
    "Host": [
      "httpbin:8000"
    ],
    "User-Agent": [
      "Fortio.org/fortio-1.66.5"
    ],
    "X-Envoy-Attempt-Count": [
      "1"
    ],
    "X-Forwarded-Client-Cert": [
      "By=spiffe://cluster.local/ns/default/sa/httpbin;Hash=b6b3f57c73b08c5d09f8444ea984bb8c7f4f582829b4dd678354591a9d84d8d5;Subject=\"\";URI=spiffe://cluster.local/ns/default/sa/default"
    ],
    "X-Forwarded-Proto": [
      "http"
    ],
    "X-Request-Id": [
      "bleicd5-919b-94eb-920f-1bc2118a8299"
    ]
  },
  "method": "GET",
  "origin": "127.0.0.6:49153",
  "url": "http://httpbin:8000/get"
}
aaron@Aarons-MacBook-Air istio-1.25.0 %
```

Now I'm going to call the service with 2 concurrent connections and send 20 requests

```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl exec "$FORTIO_POD" -c fortio -- /usr/bin/fortio load -c 2 -qps 0 -n 20 -loglevel Warning http://httpbin:8000/get

{"ts":1742998685.317694,"level":"info","r":1,"file":"logger.go","line":290,"msg":"Log level is now 3 Warning (was 2 Info)"}
Fortio 1.66.5 running at 0 queries per second, 8->8 procs, for 20 calls: http://httpbin:8000/get
Starting at max qps with 2 thread(s) [gomax 8] for exactly 20 calls (10 per thread + 0)
{"ts":1742998685.335738,"level":"warn","r":24,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998685.337209,"level":"warn","r":23,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
{"ts":1742998685.338926,"level":"warn","r":23,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
{"ts":1742998685.341891,"level":"warn","r":24,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998685.342809,"level":"warn","r":23,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
{"ts":1742998685.343908,"level":"warn","r":24,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998685.344415,"level":"warn","r":24,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998685.346212,"level":"warn","r":23,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
{"ts":1742998685.347313,"level":"warn","r":24,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
Ended after 23.485ms : 20 calls. qps=851.61
Aggregated Function Time : count 20 avg 0.0021445499 +/- 0.002245 min 0.000154375 max 0.009770667 sum 0.042890998
# range, mid point, percentile, count
>= 0.000154375 <= 0.001 , 0.000577188 , 35.00, 7
> 0.001 <= 0.002 , 0.0015 , 65.00, 6
> 0.002 <= 0.003 , 0.0025 , 80.00, 3
> 0.003 <= 0.004 , 0.0035 , 90.00, 2
> 0.005 <= 0.006 , 0.0055 , 95.00, 1
> 0.009 <= 0.00977067 , 0.00938533 , 100.00, 1
# target 50% 0.0015
# target 75% 0.00266667
# target 90% 0.004
# target 99% 0.00961653
# target 99.9% 0.00975525
Error cases : count 9 avg 0.00089863422 +/- 0.000945 min 0.000154375 max 0.002908875 sum 0.008087708
# range, mid point, percentile, count
>= 0.000154375 <= 0.001 , 0.000577188 , 66.67, 6
> 0.001 <= 0.002 , 0.0015 , 88.89, 2
> 0.002 <= 0.00290888 , 0.00245444 , 100.00, 1
# target 50% 0.000746313
# target 75% 0.001375
# target 90% 0.00209089
# target 99% 0.00282708
# target 99.9% 0.0029007
# Socket and IP used for each connection:
[0] 5 socket used, resolved to 10.108.18.5:8000, connection timing : count 5 avg 0.000578433 +/- 0.000642 min 2.825e-05 max 0.001544833 sum 0.002892165
[1] 5 socket used, resolved to 10.108.18.5:8000, connection timing : count 5 avg 5.54666e-05 +/- 2.905e-05 min 3.3459e-05 max 0.00011275 sum 0.000277333
Connection time (s) : count 10 avg 0.0003169498 +/- 0.0005243 min 2.825e-05 max 0.001544833 sum 0.003169498
Sockets used: 10 (for perfect keepalive, would be 2)
Uniform: false, Jitter: false, Catchup allowed: true
IP addresses distribution:
```

Now I'm bringing the connections up to 3

```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl exec "$FORTIO_POD" -c fortio -- /usr/bin/fortio load -c 3 -aps 0 -n 30 -loglevel Warning http://httpbin:8000/get

{"ts":1742998743.727517,"level":"info","r":1,"file":"logger.go","line":290,"msg":"Log level is now 3 Warning (was 2 Info)"}
Fortio 1.66.5 running at 0 queries per second, 8->8 procs, for 30 calls: http://httpbin:8000/get
Starting at max aps with 3 thread(s) [gomax 8] for exactly 30 calls (10 per thread + 0)
{"ts":1742998743.737246,"level":"warn","r":13,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
{"ts":1742998743.739243,"level":"warn","r":15,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":2,"run":0}
{"ts":1742998743.739249,"level":"warn","r":13,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
{"ts":1742998743.740191,"level":"warn","r":15,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":2,"run":0}
{"ts":1742998743.740519,"level":"warn","r":14,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998743.741105,"level":"warn","r":14,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998743.741177,"level":"warn","r":15,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":2,"run":0}
{"ts":1742998743.742310,"level":"warn","r":14,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998743.743080,"level":"warn","r":14,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998743.743419,"level":"warn","r":13,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
{"ts":1742998743.743916,"level":"warn","r":13,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
{"ts":1742998743.744594,"level":"warn","r":15,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":2,"run":0}
{"ts":1742998743.744668,"level":"warn","r":13,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
{"ts":1742998743.746171,"level":"warn","r":15,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":2,"run":0}
{"ts":1742998743.747011,"level":"warn","r":15,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":2,"run":0}
{"ts":1742998743.747161,"level":"warn","r":14,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998743.747821,"level":"warn","r":15,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":2,"run":0}
{"ts":1742998743.749139,"level":"warn","r":15,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":2,"run":0}
{"ts":1742998743.750208,"level":"warn","r":14,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":1,"run":0}
{"ts":1742998743.751415,"level":"warn","r":13,"file":"http_client.go","line":1151,"msg":"Non ok http code","code":503,"status":"HTTP/1.1 503","thread":0,"run":0}
Ended after 16.972875ms : 30 calls. aps=1767.5
Aggregated Function Time : count 30 avg 0.0014980889 +/- 0.0011194 min 0.000202167 max 0.003989751 sum 0.044942667
# range, mid point, percentile, count
>= 0.000202167 <= 0.001 , 0.000601084 , 53.33, 16
> 0.001 <= 0.002 , 0.0015 , 63.33, 3
> 0.002 <= 0.003 , 0.0025 , 83.33, 6
> 0.003 <= 0.00398975 , 0.00349488 , 100.00, 5
# target 50% 0.000946811
# target 75% 0.00258333
# target 90% 0.0033959
# target 99% 0.00393037
# target 99.9% 0.00398381
Error cases : count 20 avg 0.0007462729 +/- 0.000495 min 0.000202167 max 0.002072208 sum 0.014925458
# range, mid point, percentile, count
>= 0.000202167 <= 0.001 , 0.000601084 , 80.00, 16
> 0.001 <= 0.002 , 0.0015 , 95.00, 3
> 0.002 <= 0.00207221 , 0.0020361 , 100.00, 1
# target 50% 0.000680867
# target 75% 0.000946811
# target 90% 0.00166667
```

Here is some more info from the istio proxy stats:

```
aaron@Aarons-MacBook-Air istio-1.25.0 % kubectl exec "$FORTIO_POD" -c istio-proxy -- pilot-agent request GET stats | grep httpbin | grep pending

cluster.outbound|8000||httpbin.default.svc.cluster.local; circuit_breakers.default.remaining_pending: 1
cluster.outbound|8000||httpbin.default.svc.cluster.local; circuit_breakers.default.rq_pending_open: 0
cluster.outbound|8000||httpbin.default.svc.cluster.local; circuit_breakers.high.rq_pending_open: 0
cluster.outbound|8000||httpbin.default.svc.cluster.local; upstream_rq_pending_active: 0
cluster.outbound|8000||httpbin.default.svc.cluster.local; upstream_rq_pending_failure_eject: 0
cluster.outbound|8000||httpbin.default.svc.cluster.local; upstream_rq_pending_overflow: 29
cluster.outbound|8000||httpbin.default.svc.cluster.local; upstream_rq_pending_total: 22
aaron@Aarons-MacBook-Air istio-1.25.0 %
```