Youssef waleed
2205065

# *NASA-HTTP*

**Description:**
These two traces contain two month's worth of all HTTP requests to the NASA Kennedy Space Center WWW server in Florida.

## *Analysis insights*

### 1. Request Counts

| | July | August |
|---|---|---|
| Total Requests | 1891714 | 1569898 |
| Get Requests | 1887646 | 1565812 |
| Post Requests | 111 | 111 |

**Observation:**
The vast majority of the traffic consisted of GET requests (~99.74%), withPOST requests being
negligible in volume

---

### 2. Unique IP Addresses

Total unique ips: **Jul:81983, Aug: 75060**

**Observation:**
A wide distribution of IPs hence it's a globally accessed system, exposed to the public

---

### 3. Failed Requests

| | July | August |
|---|---|---|
| Failed Requests | 10893 | 10251 |
| Failure Rate | 0.58% | 0.65% |

**Observation:** Failure rate in August is Slightly worse but still pretty minimal  which is a positive indicator of system stability and reliability.

---

## 4. Top User (Most Active IP)

**Most Active IP**: (July) piweba3y.prodigy.com (17572 requests)
(Aug) edams.ksc.nasa.gov (6530 requests)

**Observation**:
This IP  especially in **July** represents a high request volume, which may indicate internal automated access or scraping behavior and can be a security concern to validate it

---

## 5. Daily Request Averages

**Average Requests per Day:**(July) 65231
(Aug) 52329

**Observation:**
Traffic is consistently high. System performance under sustained load appears to be adequate given the low failure rate since only 0.61% of total requests failed

---

## 6. Top 10 Days with the Most Failures

| July | August |
|---|---|
| 19/Jul (670 Failure) | 30/Aug (601 Failures) |
| 06/Jul (656 Failure) | 31/Aug (541 Failures) |
| 07/Jul (583 Failure) | 07/Aug (538 Failures) |
| 13/Jul (543 Failure) | 29/Aug (453 Failures) |
| 03/Jul (543 Failure) | 25/Aug (444 Failures) |
| 05/Jul (520 Failure) | 24/Aug (441 Failures) |
| 25/Jul (510 Failure) | 27/Aug (423 Failures) |

| | |
|---|---|
| 18/Jul (509 Failure) | 28/Aug (417 Failures) |
| 12/Jul (488 Failure) | 08/Aug (409 Failures) |
| 11/Jul (479 Failure) | 06/Aug (378 Failures) |

**Observation:** Late August had a noticeable cluster of failure spikes. These could be linked to system updates, content changes, or malicious activity so it needs to be checked if there was an new update published on these spikes days like (19/jul and 30/Aug) also on the 7th of each month there was a noticeable spike in the number of failures as well, so these failures needs to be checked carefully incase if it was an attempt to attack the server.

---

## 7. Request Distribution by Hour (24 Hour Format)

| July | August |
|---|---|
| 12:00 122085 Requests | 12:00 105143 Requests |
| 14:00 122479 Requests | 13:00 104536 Requests |
| 13:00 120814 Requests | 15:00 109465 Requests |
| 15:00 121200 Requests | 14:00 101394 Requests |
| 11:00 115720 Requests | 16:00 99527 Requests |
| 16:00 118037 Requests | 11:00 95344 Requests |
| 10:00 105507 Requests | 10:00 88309 Requests |
| 09:00 99969 Requests | 09:00 78695 Requests |
| 08:00 83750 Requests | 08:00 65443 Requests |
| 07:00 54017 Requests | 07:00 47386 Requests |
| 00:00 62450 Requests | 00:00 47862 Requests |
| 01:00 53066 Requests | 01:00 38531 Requests |
| 02:00 45297 Requests | 02:00 32508 Requests |
| 03:00 37398 Requests | 03:00 29995 Requests |
| 06:00 35253 Requests | 06:00 31287 Requests |
| 04:00 32234 Requests | 04:00 26756 Requests |
| 05:00 31919 Requests | 05:00 27587 Requests |
| 17:00 97609 Requests | 17:00 80834 Requests |
| 18:00 79282 Requests | 18:00 66809 Requests |
| 21:00 71922 Requests | 19:00 59315 Requests |
| 19:00 71776 Requests | 20:00 59944 Requests |
| 22:00 70759 Requests | 22:00 60673 Requests |
| 20:00 69809 Requests | 21:00 57985 Requests |
| 23:00 69362 Requests | 23:00 54570 Requests |

**Observation**: even distribution with most requests are at peak hours like 12:00 and a noticeable drop in the numbers at time like 23:00, and one of the **suggestions** if the company is using cloud service is to set up automatic scaler to scale up in the peak hours and then scale down when the server is idle to save computing power and also to make sure servers are meeting all requests and prevent any kind of Attack like DDos that could happen at peak hours.

### 8. Status Code Breakdown

| July | August |
|---|---|
| 200: 1697914 | 200: 1,396,473 |
| 304:132626 | 304: 134138 |
| 404: 10,714 | 302: 26422 |
| 302: 46549 | 404: 9978 |
| 786: 244 | 403: 171 |
| 500->599: 1,253 | 500->599:99 |

**Observation:**

Successful responses (200) dominate, but notable 404 codes indicate broken or restricted links. The presence of 500 errors hints at some server-side issues.

So we need to investigate recurring 404 errors, consider creating redirects or updating broken urls and regarding 500 errors we need to review internal logs to determine their root causes like if it's a database failure or app crashes and work on them to get it fixed.

---

### 9. Most Active User by Method

Most GETs: (**July)** piweba3y.prodigy.com (17572), (**Aug**) edams.ksc.nasa.gov (6528)

Most POSTs: (**July)** 163.205.1.45 (21), (**Aug**) seabrk.mindspring.com (8)

**Observation:**

GET usage Dominates POST activity, consistent with the request count data.But we need to monitor these users to make sure no malicious behaviour is happening especially if most of their request are failing or getting blocked due to restricted access

### 10. Failure Patterns by Hour

| July | August |
|---|---|
| 15:00 (841 failures) | 12:00 (670 failures) |
| 14:00 (753 failures) | 02:00 (618 failures) |

| | |
|---|---|
| 11:00 (732 failures) | 13:00 (616 failures) |
| 12:00 (649 failures) | 17:00 (590 failures) |
| 10:00 (643 failures) | 16:00 (583 failures) |
| 16:00 (643 failures) | 15:00 (549 failures) |
| 17:00 (616 failures) | 14:00 (525 failures) |
| 13:00 (531 failures) | 10:00 (488 failures) |
| 18:00 (494 failures) | 23:00 (487 failures) |
| 22:00 (487 failures) | 22:00 (459 failures) |
| 09:00 (481 failures) | 20:00 (457 failures) |
| 23:00 (467 failures) | 19:00 (444 failures) |
| 21:00 (446 failures) | 21:00 (434 failures) |
| 00:00 (429 failures) | 11:00 (429 failures) |
| 19:00 (411 failures) | 18:00 (429 failures) |
| 20:00 (380 failures) | 00:00 (371 failures) |
| 08:00 (358 failures) | 03:00 (365 failures) |
| 01:00 (319 failures) | 09:00 (352 failures) |
| 02:00 (276 failures) | 08:00 (340 failures) |
| 03:00 (242 failures) | 01:00 (332 failures) |
| 07:00 (240 failures) | 07:00 (225 failures) |
| 04:00 (167 failures) | 04:00 (182 failures) |
| 05:00 (155 failures) | 05:00 (171 failures) |
| 06:00 (133 failures) | 06:00 (135 failures) |

Most failures are at peak time which is a normal thing given the count of request but some of the failures like at 02:00 are a bit high considering that it's not one of the top hours for requests so these failures needs to be investigated and check if it's was due an internal error or malicious activity was happening.

—----------

# Conclusion

The system handled over **1.5 million** requests with excellent stability and only **0.65% failure rate**. Usage patterns are consistent and global. There are some failure spikes and a few heavily active users worth monitoring, but overall, the log data indicates healthy traffic and operational performance.

**Acknowledgements**

The logs were collected by Jim Dumoulin of the Kennedy Space Center, and contributed by Martin Arlitt (*mfa126@cs.usask.ca*) and Carey Williamson (*carey@cs.usask.ca*) of the University of Saskatchewan.