

# PACKET NETWORK QoS

13 NOVEMBER 2024



**AT HANURANTO,**

Senior Lecturer FTE

Director of Endowment Fund Management – Telkom University

Telkom  
University,  
your beloved  
campus right  
now!

TELKOM UNIVERSITY TELAH MERAIH REKOGNISI:



per



per September 2023



Tahun 2023



per Juli 2024



per Oktober 2024



Telkom University meraih  
**peringkat 2 Indonesia's SDGs  
Action Awards 2024**

**2024**

**85.875** Alumni

**34.078** Mahasiswa kampus utama

**1.433** Mahasiswa kampus Jakarta

**3.346** Mahasiswa kampus Surabaya

**5.951** Mahasiswa kampus Purwokerto

**44.808** Total Mahasiswa TUNC

+



Telkom University  
Landmark Tower

Telkom University National Campus  
Road to **World Class University**



Telkom University Kampus Jakarta



Telkom University Kampus Surabaya



Telkom University Kampus Purwokerto

# EMPLOYABILITY

## Why Graduate Employability Excellence is *important & urgent*?

- Honesty
- Discipline
- Easy to get along
- Creative thinking
- Critical thinking
- Hard work
- Leadership
- Ability to sell idea
- Loving for his career
- Competitive personality
- Persistence
- Being well organized
- Etc.

1 Amanah Utama Orang Tua, Masyarakat & Bangsa

2 Bagian dari University Business Process Chain  
Calon Mahasiswa – Mahasiswa - Alumni

3 Alumni Tetap Relevan di tengah industri yang disruptive

4 Alumni mempunyai Transisi Karier yang lebih mudah

5 Indikator Kinerja Utama (IKU) PT no. 1  
Lulusan Mendapatkan Pekerjaan yang Layak

6 Memperkuat Industrial Relationship

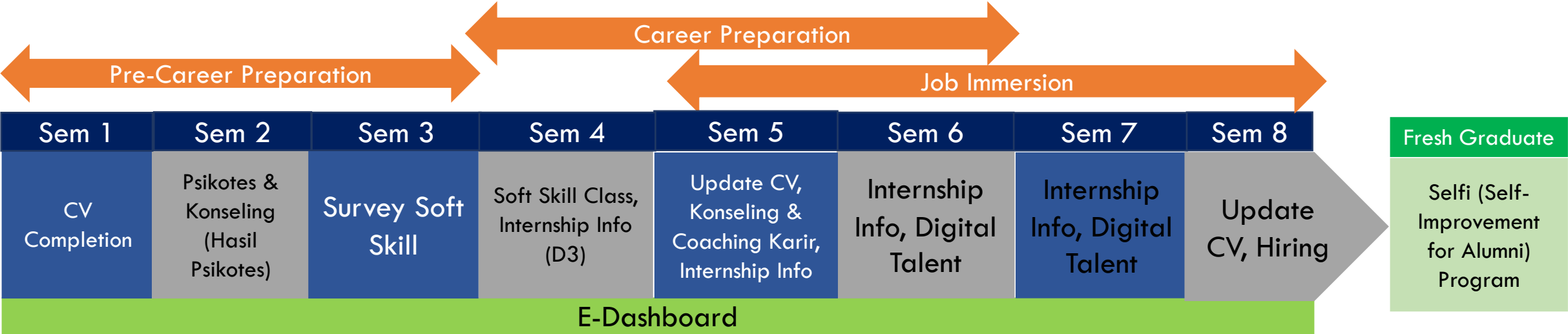
7 Amal Jariyah Civitas Telkom University

Part of  
Telkom University  
Sustainability  
Growth



# STUDENT EMPLOYABILITY JOURNEY

Rangkaian kegiatan persiapan & pengembangan karir bagi mahasiswa Telkom University yang meliputi **pra-career preparation, career preparation, & job immersion** yang dilaksanakan secara terencana & terintegrasi.

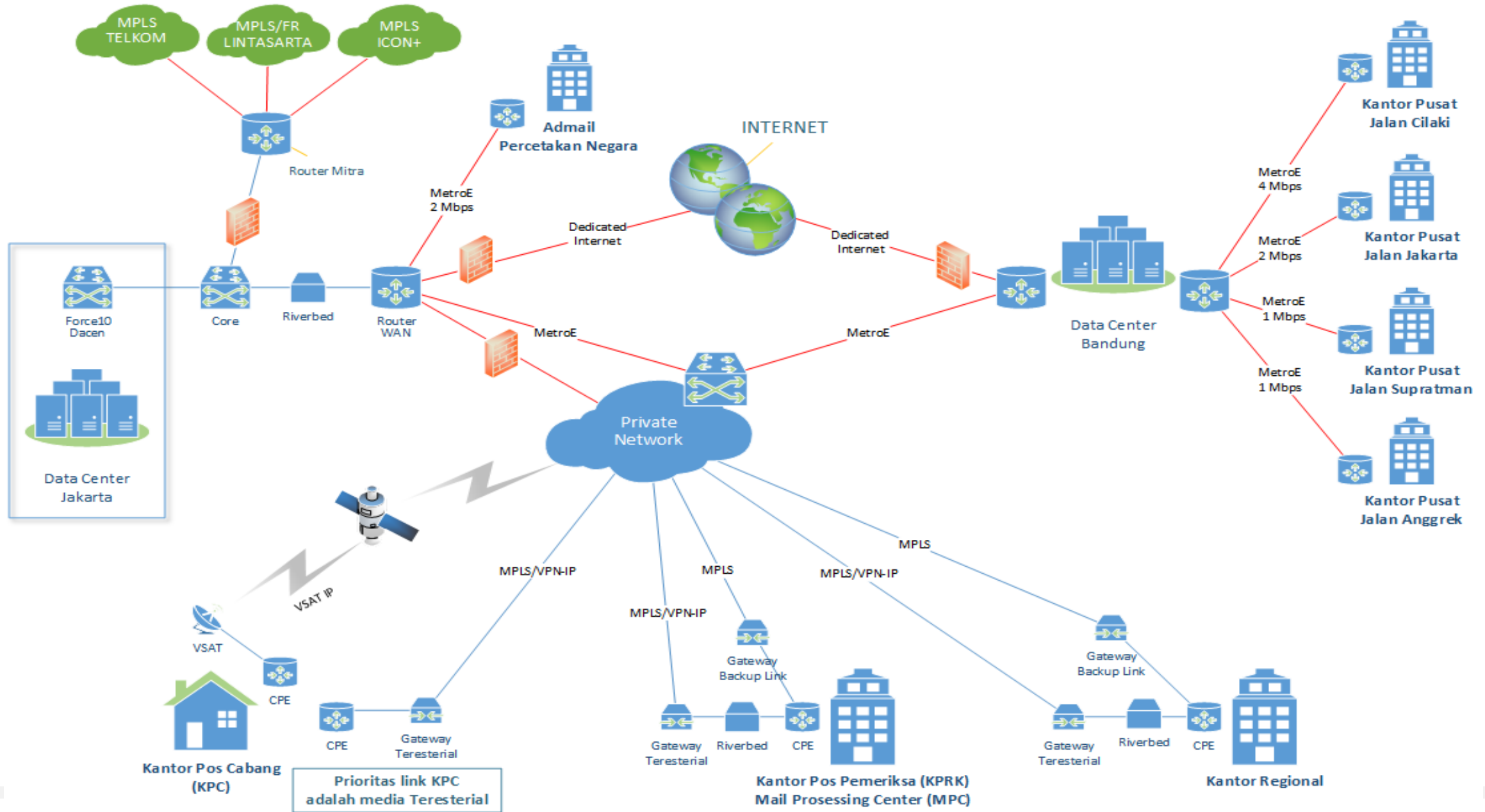


## Manfaat STAR:

- Membuat kegiatan persiapan dan pengembangan karir mahasiswa Universitas Telkom menjadi lebih terencana dan terukur.
- Mengetahui tingkat *employability* mahasiswa Universitas Telkom sejak dini.
- Membentuk *talent pool* mahasiswa Universitas Telkom.







There are **2 disadvantages** of circuit switched

- Communication networks are found to be idle frequently
- Circuit switched network provides constant data rate even in idle condition

## Packet Switched

The packet is the format of data sent in the packet switched network



- Header & Trailer contain information of control/signaling
- Each packet passes through some nodes (forwarding/ routing)
- In the node, each packet is treated with these actions (Receive, Store and Forward Networks)
- There is no pre-emption / interruption in the transmission process
- There is no certain capacity allocated in the transmission process

Therefore, the packet switch network approach provides some advanced solutions

- Higher network efficiency
- Providing several data rate services
- While experiencing higher traffic, the data will not be blocked / dropped → buffering mechanism
- It is equipped with data priority

# PACKET SWITCH QoS

QoS (Quality of Service) dalam jaringan paket adalah teknologi yang mengelola lalu lintas data untuk mengurangi packet loss, latency, dan jitter. QoS bertujuan untuk memastikan kinerja aplikasi penting dengan kapasitas jaringan terbatas.

QoS dapat mengelola parameter-parameter berikut: Bandwidth, Delay, Packet loss, Jitter, Congestion,

Beberapa aspek terkait layanan jaringan yang sering dipertimbangkan untuk mengukur kualitas layanan secara kuantitatif adalah:

Packet loss, Bit rate, Throughput, Transmission delay, Availability, Jitter, Congestion.

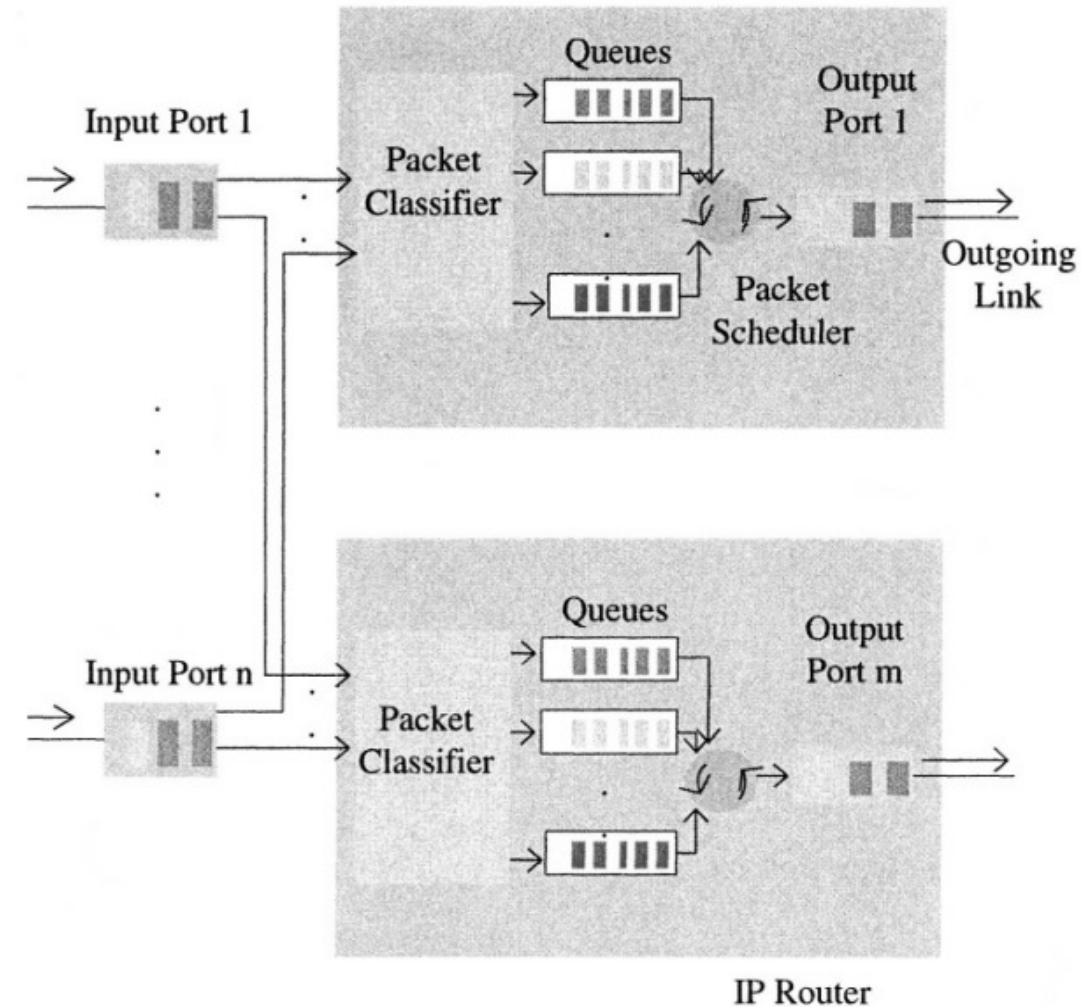
# PACKET SWITCH QoS IMPROVEMENTs

1. **TRAFFIC PRIORITY:** Menggunakan protokol seperti Multi-Protocol Label Switching (MPLS) yang menentukan jalur prioritas untuk jenis trafik tertentu, seperti VoIP atau video conferencing, agar mendapat prioritas lebih tinggi daripada trafik data biasa.
2. **BANDWIDTH MANAGEMENT:** Mengimplementasikan Traffic Shaping dan Resource Reservation (Integrated Service) Policing untuk mengontrol kecepatan data yang masuk dan keluar dari jaringan, memastikan bahwa aplikasi penting mendapat bandwidth yang cukup.
3. **PENGUNAAN CoS dan DiffServ:** CoS (Class of Service) digunakan pada jaringan LAN untuk membedakan jenis lalu lintas dan memberi prioritas, sedangkan DiffServ (Differentiated Services) digunakan pada jaringan WAN untuk mengklasifikasikan paket dan menetapkan prioritas berdasarkan DSCP (Differentiated Services Code Point).
4. **QUEUEING / SCHEDULING MANAGEMENT :** Menggunakan algoritma antrian seperti Weighted Fair Queuing (WFQ) atau Class-Based Weighted Fair Queuing (CBWFQ) untuk memastikan bahwa trafik dengan sensitif waktu dijalankan terlebih dahulu.
5. **MENGURANGI LATENCY & JITTER:** Memasang WAN Optimizers yang dapat mengurangi latency dengan teknik seperti deduplikasi data dan mengoptimalkan protokol, serta mengimplementasikan jitter buffer pada sistem VoIP untuk menyamakan interval kedatangan paket.



# SCHEDULING CONCEPTS

- Packet scheduling merupakan proses “pengaturan jadwal” paket-paket dalam antrian dengan cara yang membagi secara adil dan optimal bandwidth yang tersedia pada port keluar di antara kelas-kelas berbeda dari aliran lalu lintas masuk yang diarahkan ke port keluar tersebut.
- Packet scheduling merupakan inti dari mekanisme Quality of Service (QoS) dan memberikan perbedaan teknologi di antara produk-produk dari produsen yang berbeda.
- Gambar di samping menunjukkan contoh packet scheduling: packet scheduling diterapkan secara individual pada setiap port keluar pada router








# SCHEDULING CONCEPTS

1. Objective: maximum resource utilization

2. Benefit v.s. Cost ???

3. Scheduling Criteria:

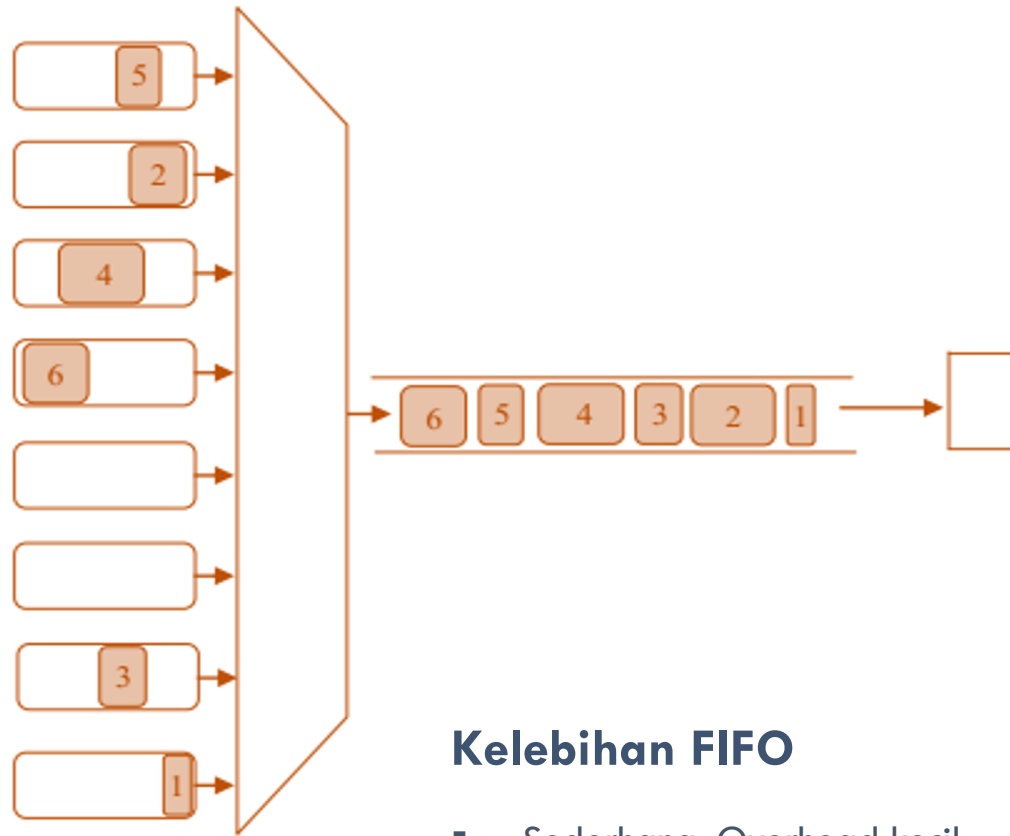
- Resources Utilization 
- Throughput 
- Turnaround Time 
  - Waiting Time 
  - Response Time 

4. Scheduling Algorithm

- First-Come First-Serve
- Shortest-Job First
- Round-Robin
- Priority
- Multilevel Queue

## FIFO (First In First Out)

- FIFO = FCFS (First Come First Serve)
- metode penjadwalan dimana semua paket diperlakukan dengan menempatkannya pada sebuah antrian lalu dilayani sesuai dengan urutan yang sama ketika paket-paket tersebut memasuki antrian
- paket yang datang terlebih dahulu, akan dilayani terlebih dahulu pula
- Jika antrian sudah penuh, paket yang datang selanjutnya ke antrian akan di *drop*
- Kelemahan algoritma ini adalah *waiting time* rata-rata yang cukup lama

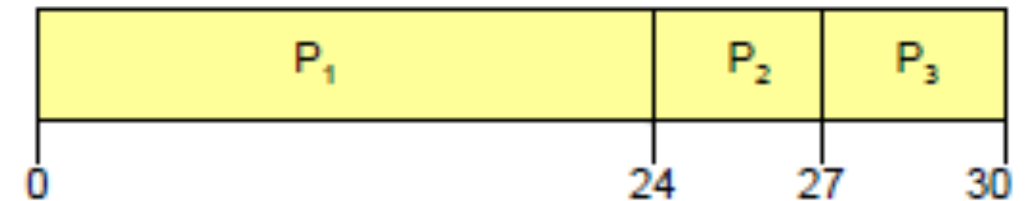


## Kelebihan FIFO

- Sederhana, Overhead kecil
- Dapat mencegah starvation, yaitu proses di mana terjadi kekurangan *resource* (biasanya terjadi karena *deadlock*) sehingga antrian tidak akan pernah mendapat *resource* yang dibutuhkan.

Process	Burst Time
$P_1$	24
$P_2$	3
$P_3$	3

Suppose that the processes arrive in the order:  $P_1, P_2, P_3$   
The Gantt Chart for the schedule is:



Waiting time for  $P_1 = 0$ ;  $P_2 = 24$ ;  $P_3 = 27$

Average waiting time:  $(0 + 24 + 27)/3 = 17$

## Kelemahan FIFO

- Proses yang pendek dapat dirugikan, bila urutan eksekusinya setelah proses yang panjang

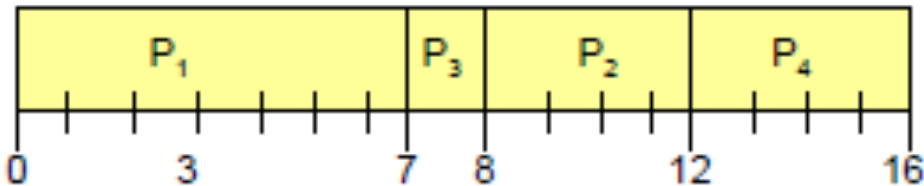
# *Shortest-Job First (SJF) Scheduling*

- Associate with each process the length of its next CPU burst. Use these lengths to schedule the process with the shortest time
- Two schemes:
  - nonpreemptive – once CPU given to the process it cannot be preempted until completes its CPU burst
  - preemptive – if a new process arrives with CPU burst length less than remaining time of current executing process, preempt. This scheme is known as the Shortest-Remaining-Time-First (SRTF)
- SJF is optimal – gives minimum average waiting time for a given set of processes

# Non-Preemptive SJF Scheduling

Process	Arrival Time	Burst Time
$P_1$	0.0	7
$P_2$	2.0	4
$P_3$	4.0	1
$P_4$	5.0	4

SJF (non-preemptive)

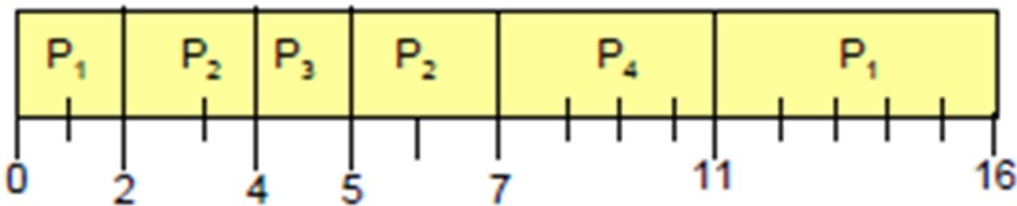


Average waiting time =  $(0 + 6 + 3 + 7)/4 = 4$

# Preemptive SJF (SRTF) Scheduling

Process	Arrival Time	Burst Time
$P_1$	0.0	7
$P_2$	2.0	4
$P_3$	4.0	1
$P_4$	5.0	4

SJF (preemptive)



Average waiting time =  $(9 + 1 + 0 + 2)/4 = 3$



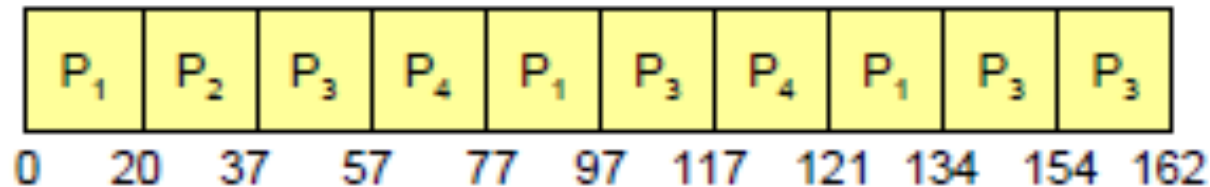
# Round-Robin Scheduling

- RR = Fair Time Sceduling
- Algoritma ini mengatur proses yang ada di antrian. Proses akan mendapat jatah waktu yg disebut *time quantum/time slice*.
- Umumnya time quantum antara 10-100ms
- Jika *time quantum*-nya habis atau proses sudah selesai maka penjadwalan dialokasikan ke proses berikutnya.
- Tak ada proses yang diprioritaskan
- Jika *time quantum* terlalu besar, sama saja dengan algoritma *first-come first-served*, jika terlalu kecil akan semakin banyak peralihan proses sehingga banyak waktu terbuang

# Round-Robin

<u>Process</u>	<u>Burst Time</u>
$P_1$	53
$P_2$	17
$P_3$	68
$P_4$	24

The Gantt chart is:



Typically, higher average turnaround than SJF, but better *response*

## Kelebihan RR

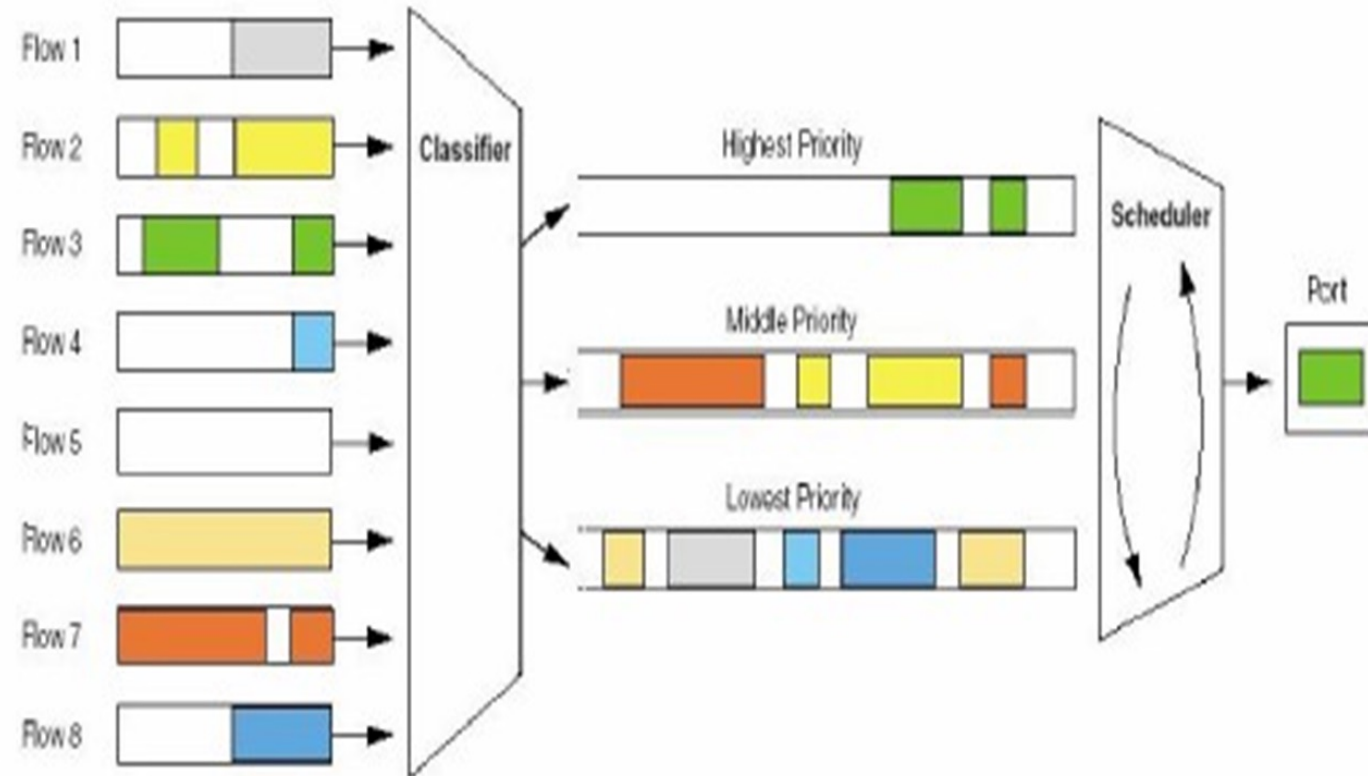
- Dapat menghindari ketidakadilan layanan terhadap proses kecil seperti yang telah terjadi pada FCFS
- Response time lebih cepat untuk proses yang berukuran kecil
- dapat mencegah *starvation*
- Overhead kecil, jika ukuran proses rata-rata lebih kecil dibandingkan quantum / slot

## Kelemahan RR

- Performa lebih buruk dibandingkan FCFS jika ukuran slot lebih besar daripada ukuran proses terbesar
- Dapat terjadi overhead berlebihan jika ukuran setiap slot terlalu kecil

# Priority Queueing (PQ)

- algoritma dasar yang didesain untuk mendukung penyediaan antrian dengan kelas layanan yang berbeda
- paket-paket yang datang akan dikelompokkan ke dalam kelas-kelas antrian tertentu sesuai dengan prioritas yang dimilikinya
- prioritas paling tinggi akan dilayani terlebih dahulu
- prioritas paling tinggi dimiliki oleh layanan yang bersifat *real-time* dan rentan *delay* seperti layanan suara atau *voice* → layanan video, → non *real-time* seperti layanan data
- paket dengan prioritas antrian sama akan dilayani berdasarkan algoritma FIFO

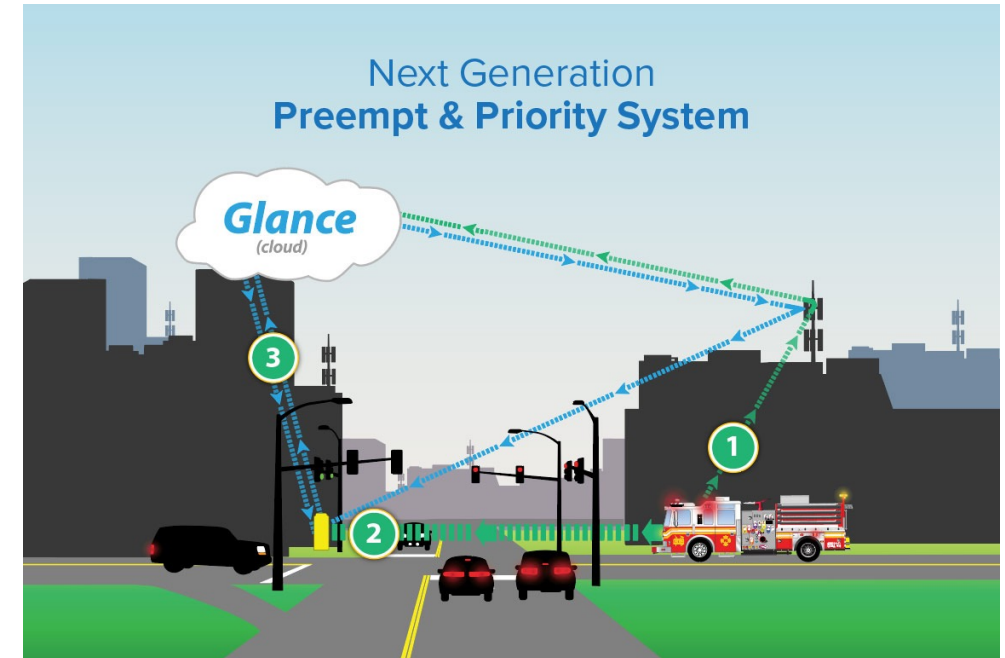


# Kelebihan PQ

- PQ memungkinkan klasifikasi paket dan trafik dengan banyak layanan yang berbeda

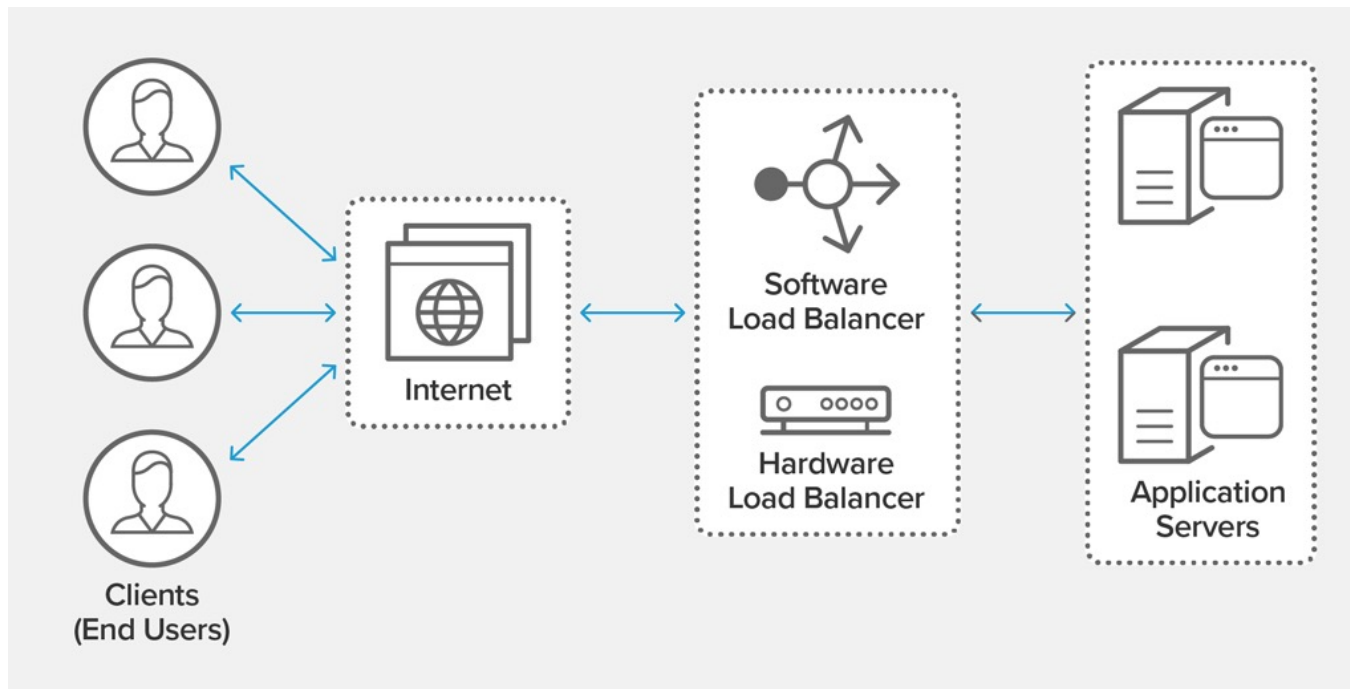
# Kelemahan PQ

- pada trafik dengan prioritas rendah akan mengalami delay yg signifikan ketika menunggu trafik prioritas tinggi dijalankan
- jika volume trafik prioritas tinggi meningkat, trafik prioritas rendah akan mengalami drop karena buffer berlebihan

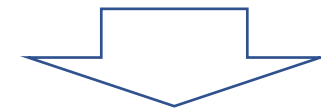


# LOAD BALANCING

- Load balancing adalah proses mendistribusikan beban kerja (workload) secara merata di antara beberapa sumber daya komputasi, seperti server, jaringan, atau sistem komputasi lainnya.
- Tujuannya adalah untuk mengoptimalkan kinerja sistem dengan memastikan bahwa setiap sumber daya digunakan secara efisien dan tidak terbebani secara berlebihan.



## SUMBER DAYA



- Pemerataan Beban Kerja
- Menjaga Ketersediaan yang tinggi
- Skalabilitas

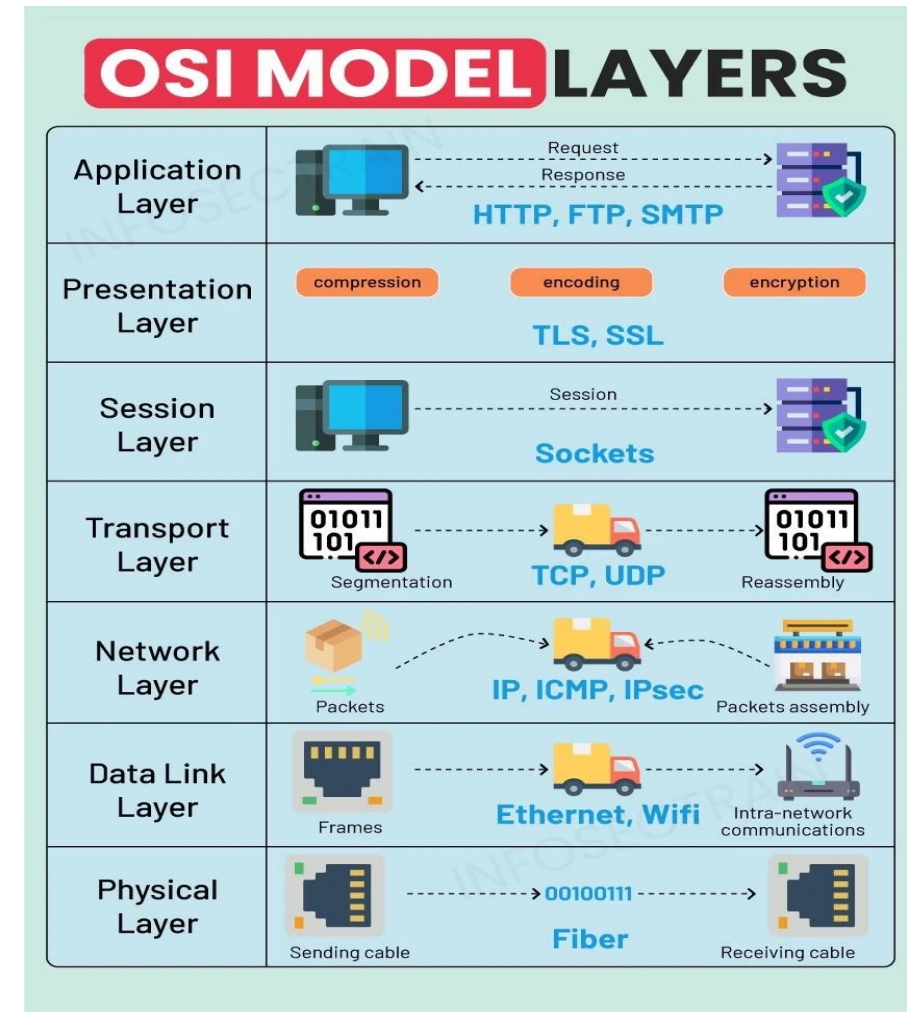


## Layer 4 Load Balancer:

- Beroperasi pada lapisan jaringan dan transport dalam model OSI.
- Menggunakan informasi dari protokol IP, TCP, FTP, UDP untuk mengatur lalu lintas.
- Tidak memperhatikan konten aplikasi atau data spesifik di dalam permintaan.
- Fokus pada faktor jaringan dan transport, seperti alamat IP dan port tujuan.

## Layer 7 Load Balancer:

- Beroperasi pada lapisan aplikasi tertinggi dalam model OSI.
- Menganalisis data dalam protokol lapisan aplikasi seperti HTTP.
- Mendistribusikan permintaan berdasarkan data aplikasi, seperti header HTTP, metode permintaan, dan isi konten.
- Mempertimbangkan data di tingkat aplikasi, seperti jenis konten dan perilaku aplikasi.

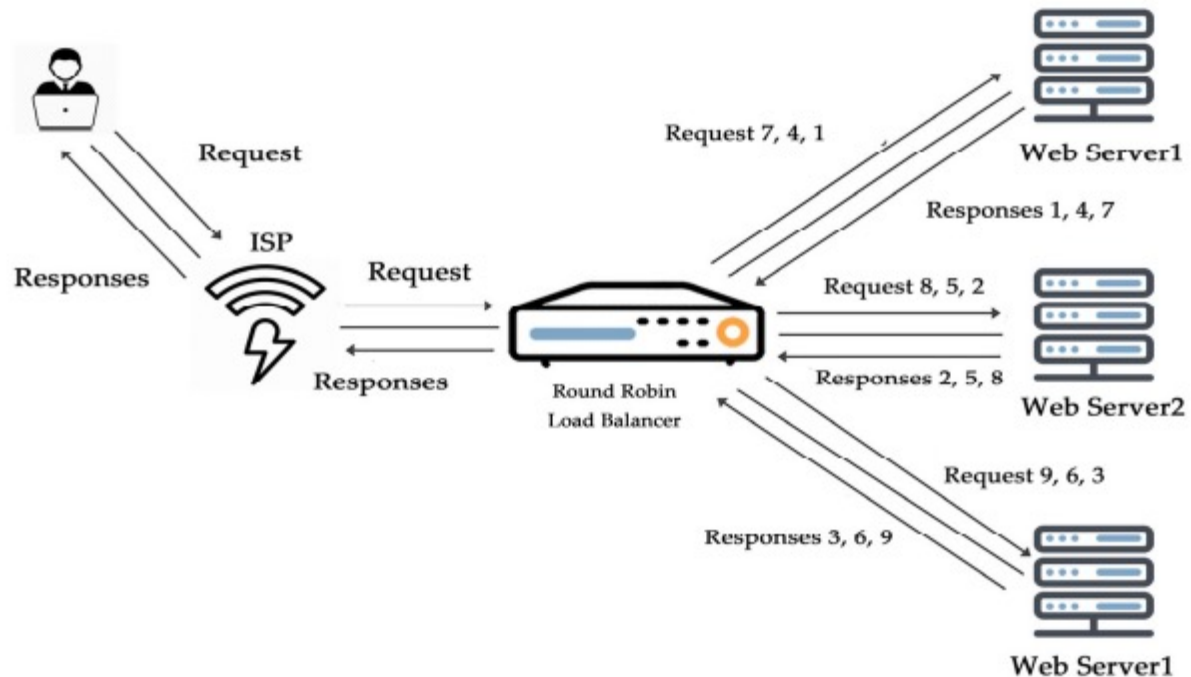


# ALGORITMA LOAD BALANCING

Image Source : Alankar B,  
Sharma G, Kaur H,  
Valverde R, Chang V.  
Experimental Setup for  
Investigating the Efficient  
Load Balancing  
Algorithms on Virtual  
Cloud. Sensors. 2020;  
20(24):7342

Algorithm	Request Serve Method	Nature	Advantage	Disadvantage
Round Robin	The request is given to the server in a rotated Sequential Manner	Dynamic	Easy to configure, deploy and widely used algorithm	Servers can overload and crash if they have different resources and processing capacities.
Least Connections	The request is given to the server with the lowest number of connections.	Dynamic	Avoids overloading a server by verifying the number of server connections	When calculating the number of existing connections, the server capacity cannot be considered.
Weighted Round Robin	Every server is used, in turn, by weight.	Static	Can send more requests to more capable and loaded servers	All the estimation requires implementing this algorithm, and this is a major drawback and also requires estimating IP networks with different packet sizes, which are difficult to do.
Source Hash	The source IP address shall be hashed and divided by the total number of servers operating to decide which server the request receives.	Static	Users connect to a still active session after disconnection and reconnection. It will increase performance.	Internet Service Provider (ISP) provides dynamic IP addresses, so it is difficult to maintain them.
Least Response Time	The request is given to the server with the lowest response time.	Dynamic	Consider both the server's capacity, response time and the number of current connections to avoid overload and crash.	Simple performing virtual machines are used, then the unequal route of traffic might be shown and this algorithm is not recommended for cookie-based session applications.
Least Bandwidth	Every server is used, in turn, by network bandwidth.	Static	Can send more requests to more capable and network bandwidth loaded servers.	It requires approximate network bandwidth, which is difficult to do in networks where the packet size of the data varying, and network bandwidth might be exhausted.

# Round Robin (RR) Algorithm

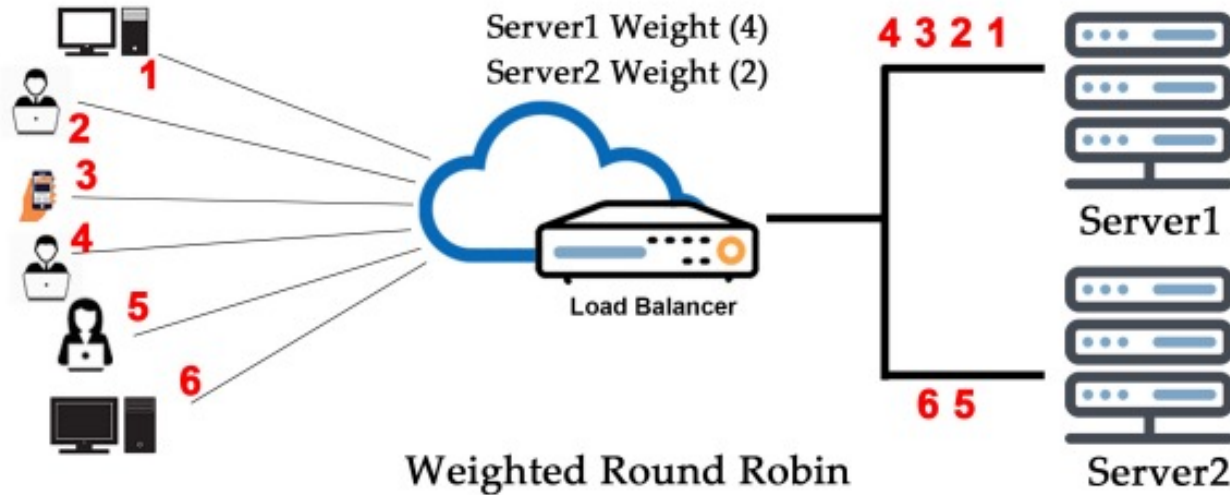


Algoritma round robin bergantung pada sistem rotasi atau pemilihan server secara berputar.

Setiap permintaan masuk secara bergantian diteruskan ke server yang tersedia secara berurutan.

Setelah melayani permintaan, server dipindahkan ke akhir antrian untuk memberi giliran pada server lain.

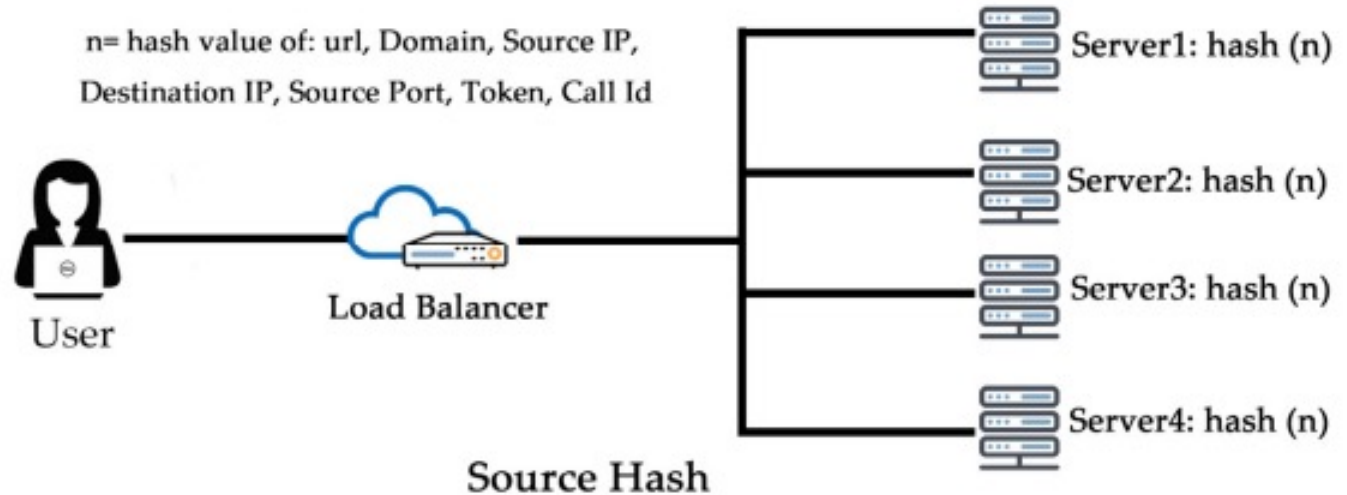
# Weighted Round Robin (WRR) Algorithm



- Algoritma Weighted Round Robin mirip dengan algoritma round robin, tetapi bersifat statis dan tidak mengubah bobot server secara dinamis.
- Jumlah server tidak terbatas dan sering direplikasi setelah peta penuh dipublikasikan ulang saat ada penambahan server.
- Algoritma ini menggunakan lebih sedikit CPU untuk berjalan.

# Source Hash Algorithm

Source hash adalah kunci (*unique identifier*) yang digunakan untuk mengalokasikan klien ke server tertentu. Source hash juga dapat digunakan sebagai metode penyeimbangan beban

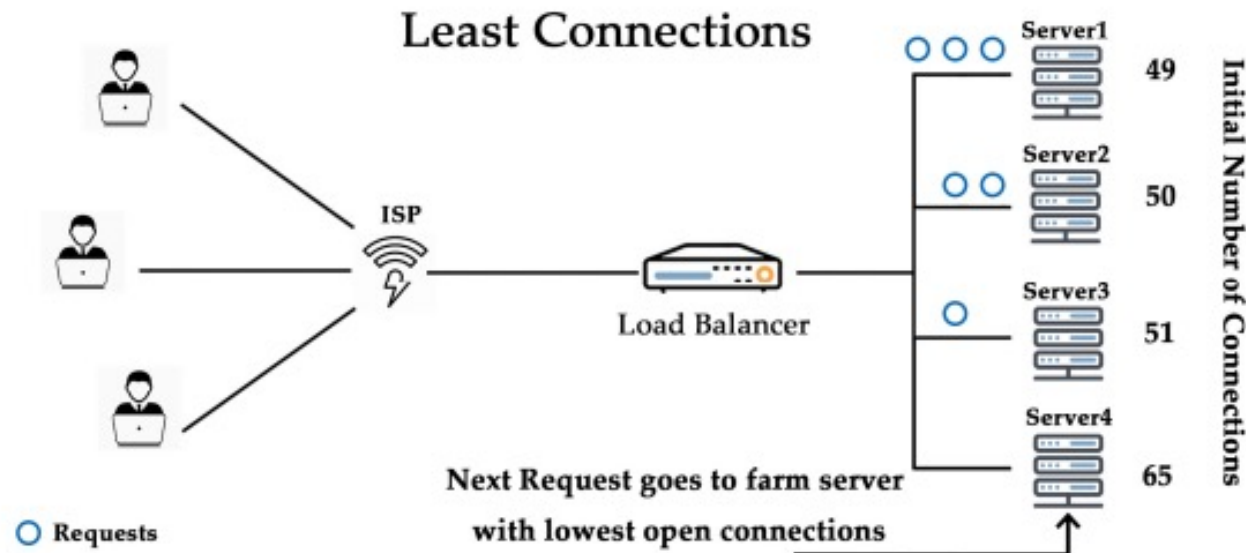


- Algoritma load balancing ini menggunakan alamat IP pengguna (sumber) untuk menentukan server yang akan menerima permintaannya.
- Tujuan dari algoritma ini adalah agar pengguna yang sama selalu diarahkan ke server yang sebelumnya telah melayani permintaannya.



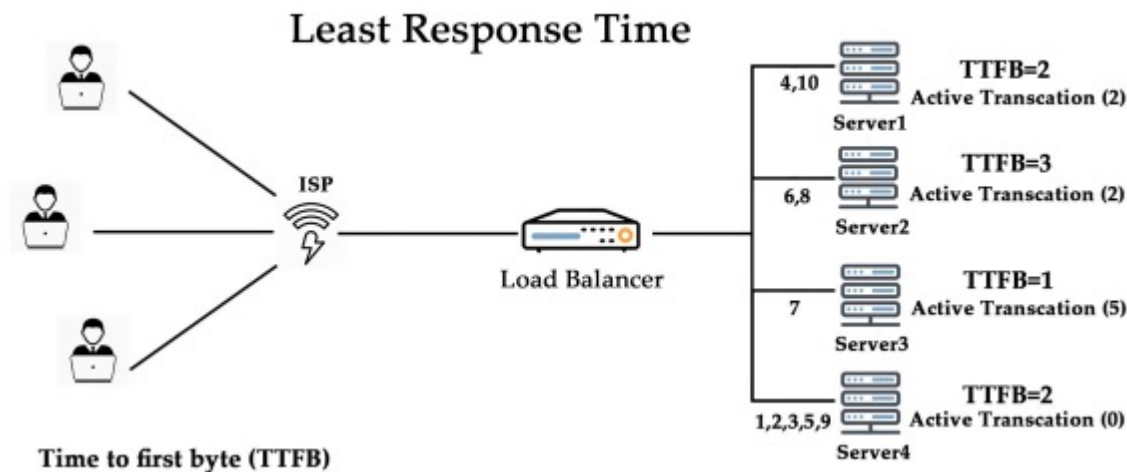
# Least Connection Algorithm

- Algoritma least connection mengarahkan lalu lintas ke server yang memiliki jumlah koneksi aktif terendah saat sedang menghadapi beban lalu lintas yang tinggi.
- Algoritma ini memastikan distribusi yang merata di antara semua server yang tersedia.
- Algoritma ini tidak direkomendasikan untuk aplikasi dengan sesi yang singkat seperti HTTP.



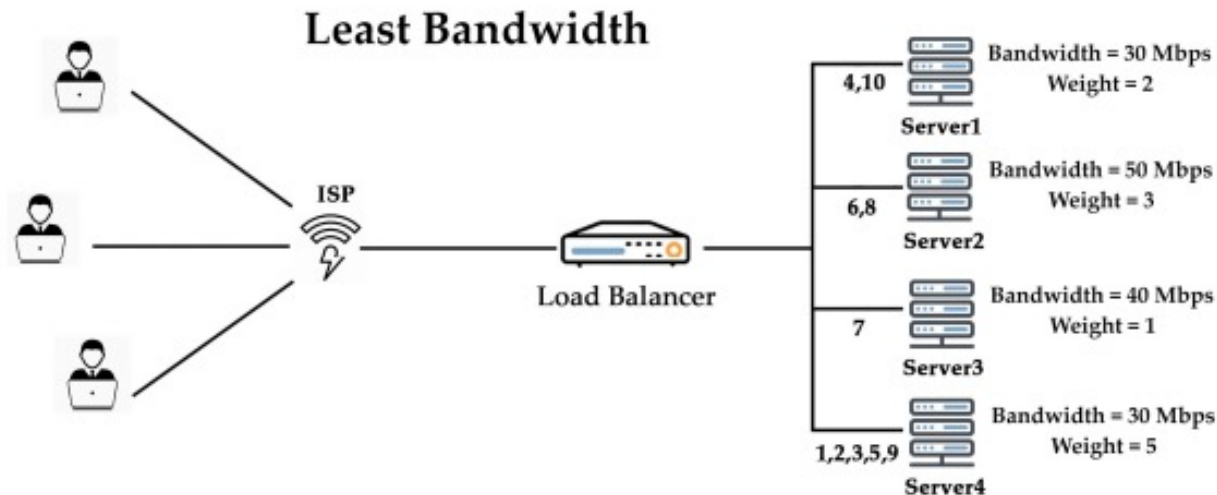
# Least Response Time Algorithm

- Algoritma least response time mengarahkan lalu lintas ke server dengan jumlah koneksi aktif terendah dan waktu respons yang paling rendah.
- Algoritma ini memadukan dua parameter, yaitu jumlah koneksi aktif terendah dan waktu respons terendah, per virtual machine (VM).
- Digunakan ketika mesin virtual dengan performa tinggi digunakan, karena jika mesin virtual dengan performa rendah digunakan, maka rute lalu lintas yang tidak merata mungkin terjadi.



# Least Bandwidth Algorithm

- Algoritma least bandwidth mengolah lalu lintas dalam megabit per detik (Mbps), mengirim permintaan pengguna ke server dengan Mbps pengunjung terkecil.
- Algoritma ini sangat berguna ketika semua mesin virtual dalam infrastruktur cloud memiliki bandwidth jaringan yang berbeda-beda.
- Algoritma ini membutuhkan perkiraan bandwidth jaringan yang hampir akurat, namun sulit dilakukan dalam jaringan di mana ukuran paket data bervariasi dan bandwidth jaringan dapat habis.



# ASSIGNMENT / CHALLENGE #5

1. Baca Referensi di bawah dengan Seksama dan Pilih SALAH SATU Metoda QOS Improvement (Scheduling or Load balancing)
2. Buat Analisis ADVANTAGES & DIS-ADVANTAGES untuk Salah satu Algorithm !

## **QOS IN PACKET NETWORKS**

by

Kun I. Park, Ph.D.

The MITRE Corporation USA

International Journal "Information Technologies & Knowledge" Volume 9, Number 4, 2015

## **METHODS AND ALGORITHMS OF LOAD BALANCING**

Igor Ivanisenko



**TERIMA KASIH**