

# Load Balancing

- *Load Balancing* merupakan teknik mendistribusikan beban secara merata diantara banyaknya sumber daya komputasi seperti processor, memori, dsb.
- *Load Balancing* sangat penting untuk aplikasi yang sering berurusan dengan jumlah permintaan layanan yang besar dan tidak terprediksi.
- **Tujuan** : mengoptimalkan pemanfaatan sumber daya, memaksimalkan *throughput*, meminimalkan *respons time* dan menghindari kelebihan beban pada beberapa sumber daya.

# Load Balancing...

- Ketika satu komponen gagal karena suatu alasan, *Load Balancing* dapat mengalihkan beban ke komponen operasi lainnya
- Mekanisme *Load Balancing* yang efisien akan meningkatkan pemanfaatan sumber daya → meningkatkan kinerja seluruh sistem

# Pentingnya Load Balancing dalam Cloud Computing

- *Load Balancing* merupakan salah satu masalah utama Cloud Computing → kelebihan pada sumber daya apapun dapat menyebabkan kinerja sistem yang buruk
- Untuk mendapatkan kinerja optimal dari suatu sistem, beban kerja harus didistribusikan dengan benar pada sumber daya yang tersedia.

# Pentingnya Load Balancing dalam Cloud Computing...

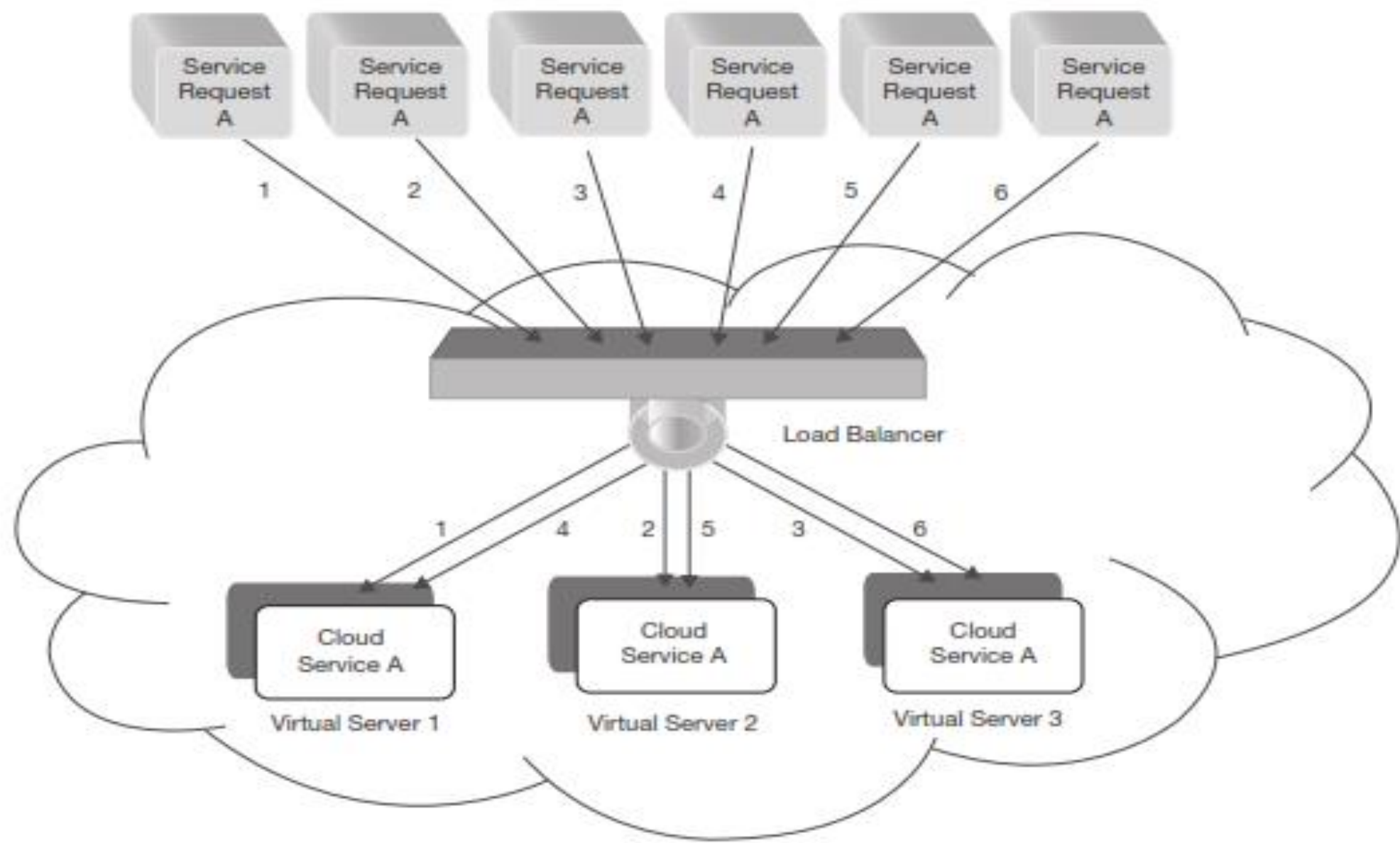
1. Memberikan fleksibilitas arsitektur dan bantuan yang penting dalam membuat arsitektur komputasi yang *scalable*
2. Memastikan pemanfaatan sekumpulan sumber daya yang sejenis secara efisien
3. Pemanfaatan sumber daya yang efisien secara otomatis dapat meningkatkan performa sistem secara keseluruhan
4. Teknik memisahkan aplikasi dari implementasi fisik selama eksekusi. Hal ini menciptakan lapisan abstraksi yang meningkatkan keamanan aplikasi dan sistem.
5. Pemisahan dari sumber daya fisik dari akses langsung aplikasi juga membuat Cloud Computing lebih toleran jika terjadi kegagalan komponen

# Mekanisme Load Balancing dalam Cloud Computing

- Mekanisme *Load Balancing* mendistribusikan permintaan layanan di seluruh aplikasi Cloud yang digunakan di *data center* yang tersebar di seluruh dunia.
- Setiap *data center* Cloud harus memiliki *Load Balancer* sendiri untuk menjadwalkan permintaan layanan masuk ke sumber daya yang tepat.
- *Service Load Balancing* mendistribusikan permintaan layanan aplikasi di sumber daya sangat penting untuk keberhasilan Cloud Computing.

# Mekanisme Load Balancing dalam Cloud Computing...

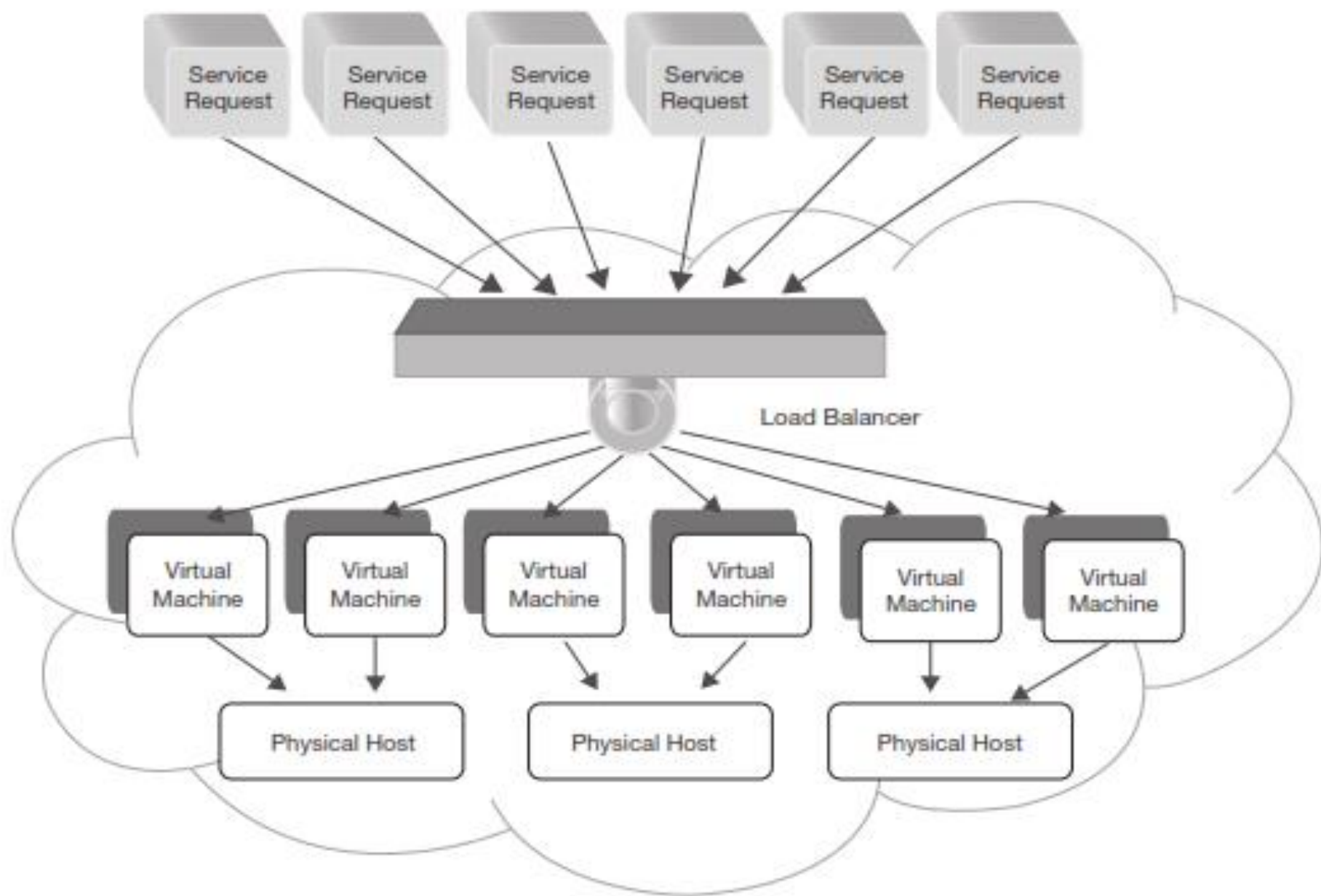
1. Permintaan layanan tertentu dari *client* yang masuk akan menuju *Load Balancer*
2. *Load Balancer* di *front-end node* menerima beberapa permintaan layanan
3. Permintaan didistribusikan menuju *virtual server (back-end node)* yang tersedia menggunakan algoritma penjadwalan.



# Mekanisme Load Balancing dalam Cloud Computing...

- Cloud *provisioning* merupakan teknik mengalokasikan sumber daya penyedia Cloud kepada *client*
- Ketika *front-end node* meneruskan permintaan layanan, semua aktivitas terjadi di *back-end node*.
- Penugasan dan eksekusi permintaan ini terjadi dalam dua fase :
  - VM Provisioning
  - Resource Provisioning





# Mekanisme Load Balancing dalam Cloud Computing...

Sumber daya dan layanan yang bertanggung jawab dalam *Load Balancing* :

- *Application server instances* (termasuk Processor dan Memori)
- *Storage resources*
- *Network switches*
- *Services* seperti DNS, TCP, HTTP, dan HTTPS