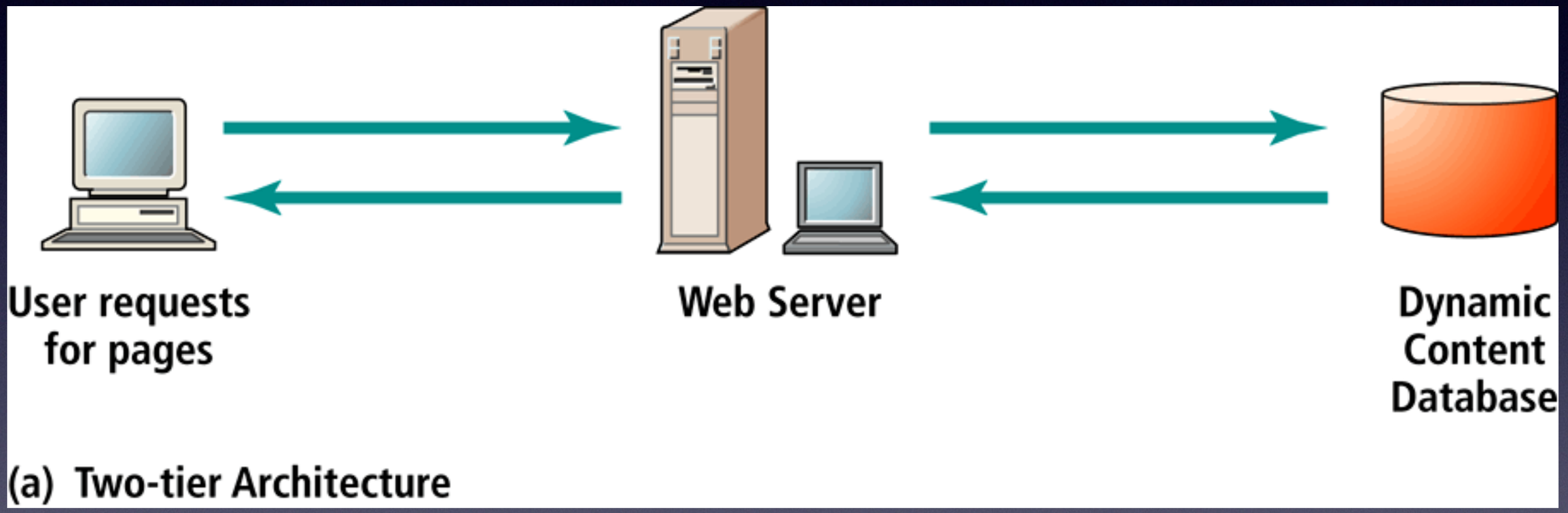
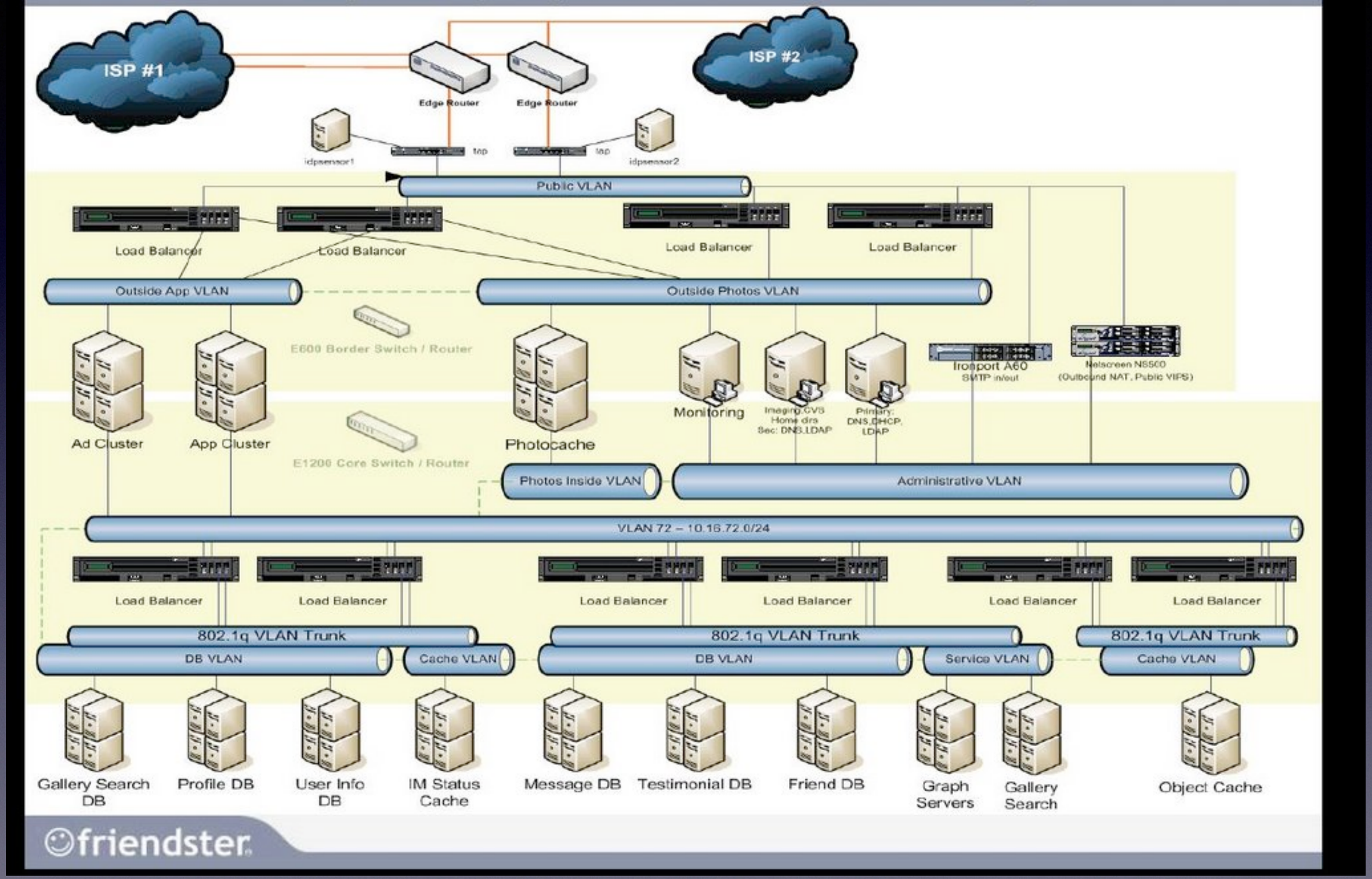
Das echte Leben ....

Typische Anwendung ...



[Graphic by S. Misra, CSU]

Traditional 3-tier architecture with a twist: Load Balanced mySQL (no persistent connections)



[http://mysqldba.blogspot.com/2006/11/unorthodox-approach-to-database-design.html]

ACID - gut und schön, aber ....

We‘re not in Kansas anymore

Zahlen

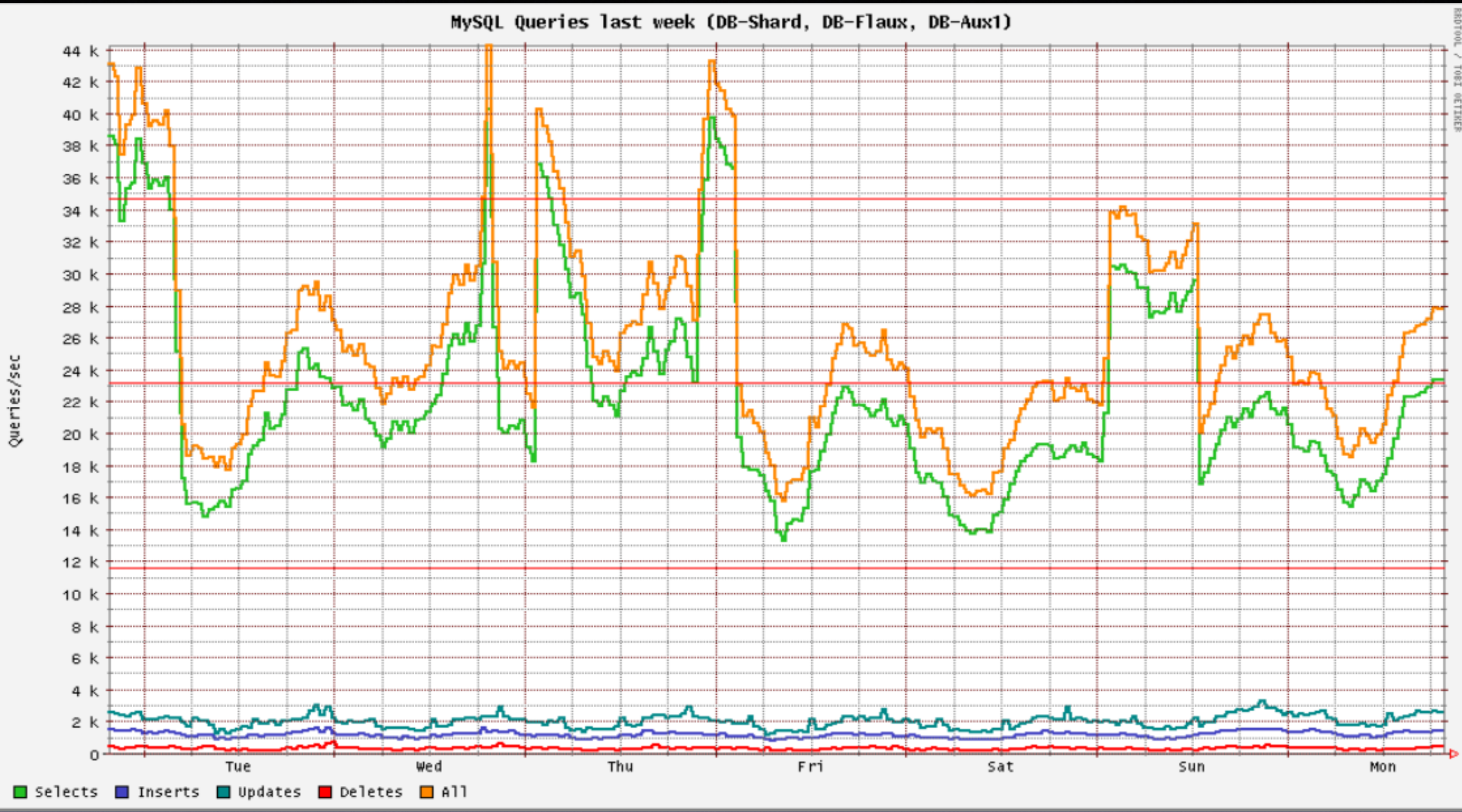
- Amazon
  - ca 150 Millionen Prime-Kunden weltweit (2020) 2,8 Millionen Bestellungen pro Tag in D, ca 2 Millionen aktive Verkäufer,
    100-150 service calls pro Seite
- Instagram
  - 80 Millionen Millionen Bilder / Tag
- ebay (Stand 2018)
- > 500 PB, ca. 200 Millionen aktive Nutzer, 300 Milliarden Suchen pro Tag
  - Spiegel (Januar 2020) ca. 20 Millionen Visits pro Tag

google

- richtig fette Datenmengen (Exabytes)
- keine Datenbanken (mehr möglich)
- „Sorting 1 PB with MapReduce. PB is not peanut-butter-and-jelly misspelled. It's 1 petabyte or 1000 terabytes or 1,000,000 gigabytes. *It took six hours and two minutes to sort 1PB (10 trillion 100-byte records) on 4,000 computers* and the results were replicated thrice on 48,000 disks.“
- „100k MapReduce jobs are executed each day; more than 20 petabytes of data are processed per day; more than 10k MapReduce programs have been implemented; machines are dual processor with gigabit ethernet and 4-8 GB of memory.“
- Details: http://research.google.com/archive/mapreduce.html

[http://highscalability.com/google-architecture]

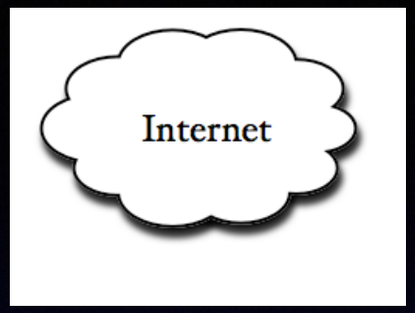
flickr Stand 2007

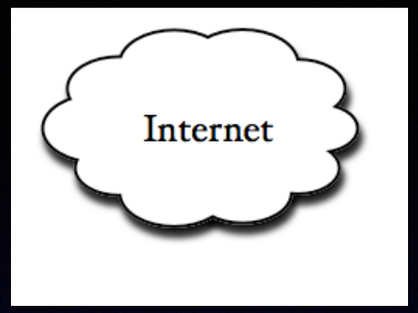


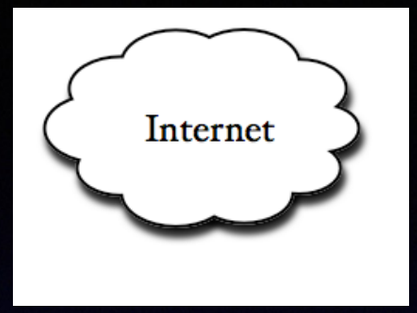
Skalierung

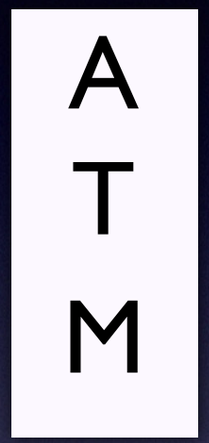
- vertikal (scaling up)
  - Hardware aufrüsten (mehr CPUs, Speicher, Platten usw.)
  - zunächst einfach ...
- horizontal (scaling out)
  - Verteilung auf mehrere Rechner
  - schwierig

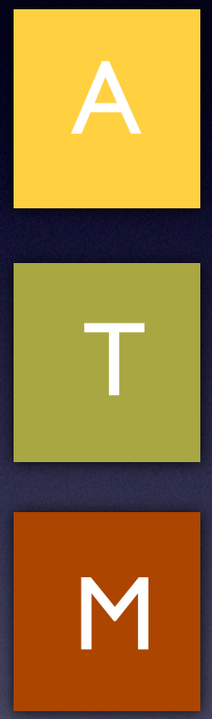
der einfache Fall





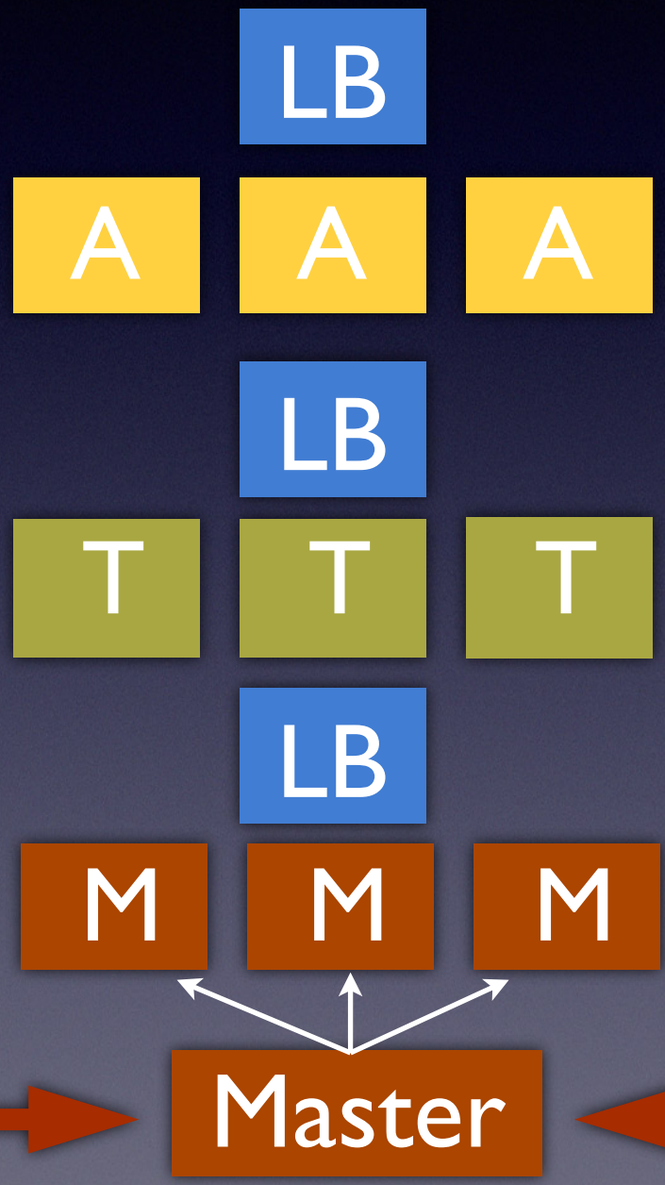






A = apache
T = tomcat
M = mysql

LB = Load Balancer

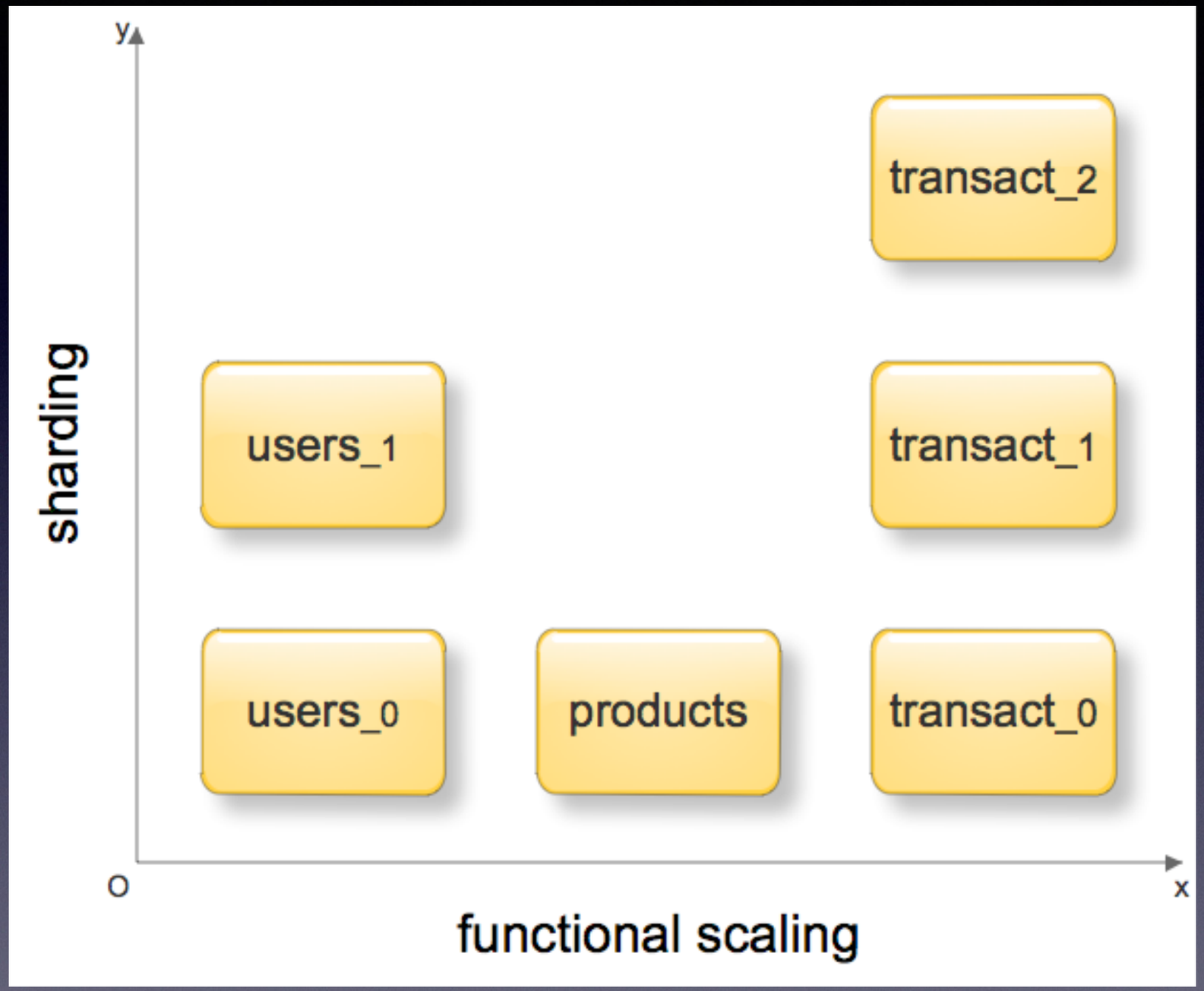


Master-Slave-Replikation

Replikation

- Replikation: Kopien der Daten auf mehreren Rechnern
  - Master-Slave: Schreiben nur auf einem Knoten
  - Multiple-Master: Schreiben auf mehreren (allen) Knoten
- „Multi-master replication is powerful, and it is complicated to create and monitor the replication environment. Because of this complexity, multi-master replication requires some additional planning“ (Oracle)
- synchron (verteilte Transaktionen, 2 phase commit)

Partitionierung

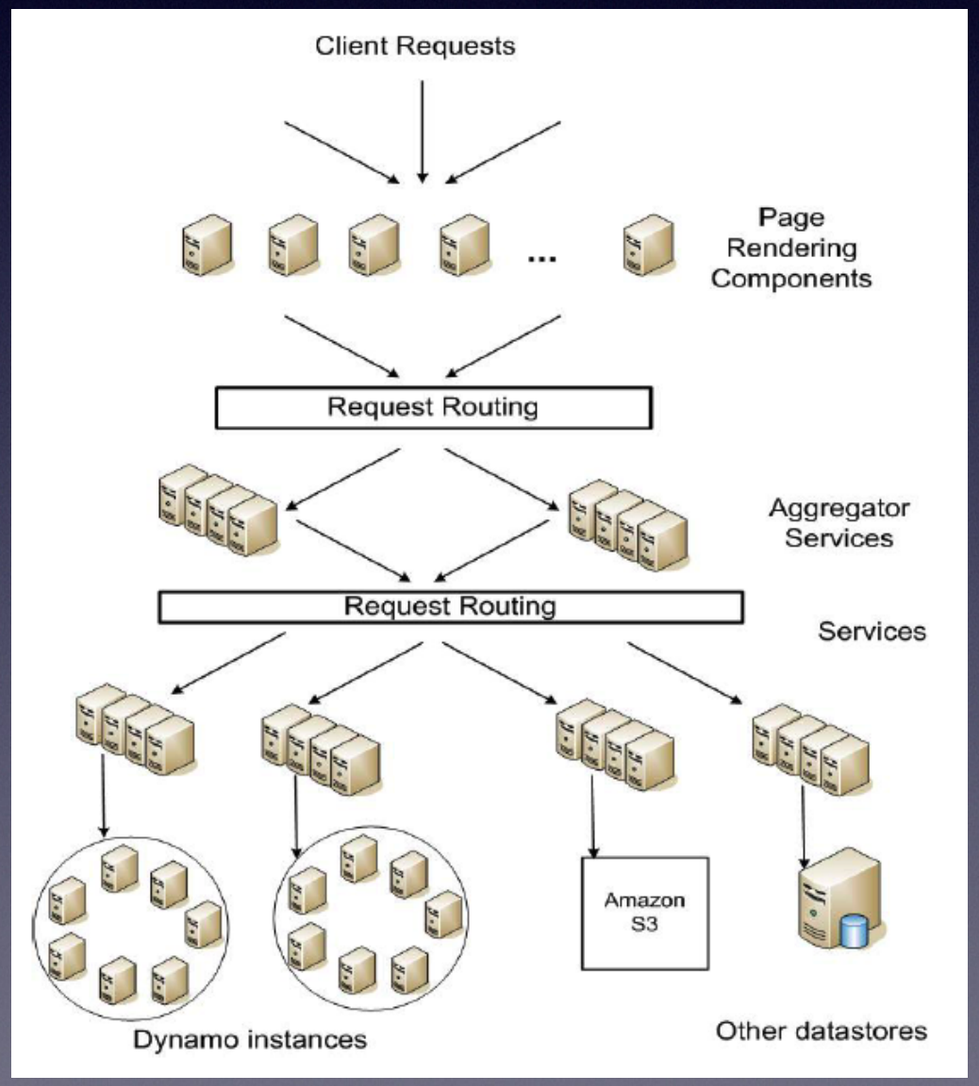


functional partitioning

- Daten zu verschiedenen Funktionalitäten in verschiedenen Datenbanken
- erhöht Skalierbarkeit, vermindert Kopplung (siehe auch Web-Services)
- Integrität ?
  - referential Integrity --> Kopplung
    --> keine Skalierbarkeit mehr :-(
  - Integrität muß in Applikation realisiert werden !
  - Problematisch !

Amazon: Dynamo

Hochgradig redundant, Replikation, extreme Verfügbarkeit, key-value-store



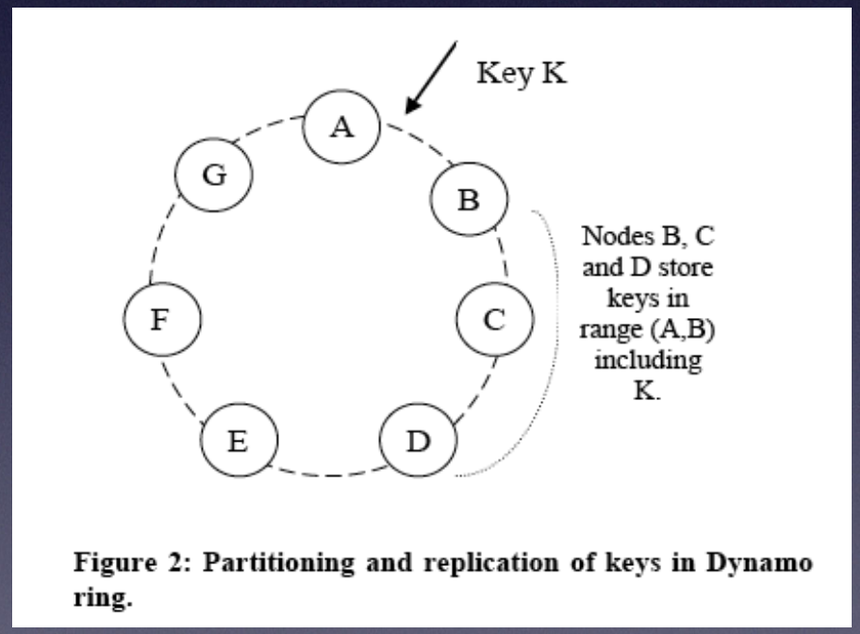


Figure 2: Partitioning and replication of keys in Dynamo ring.

http://s3.amazonaws.com/AllThingsDistributed/sosp/amazon-dynamo-sosp2007.pdf

Ergebnis:

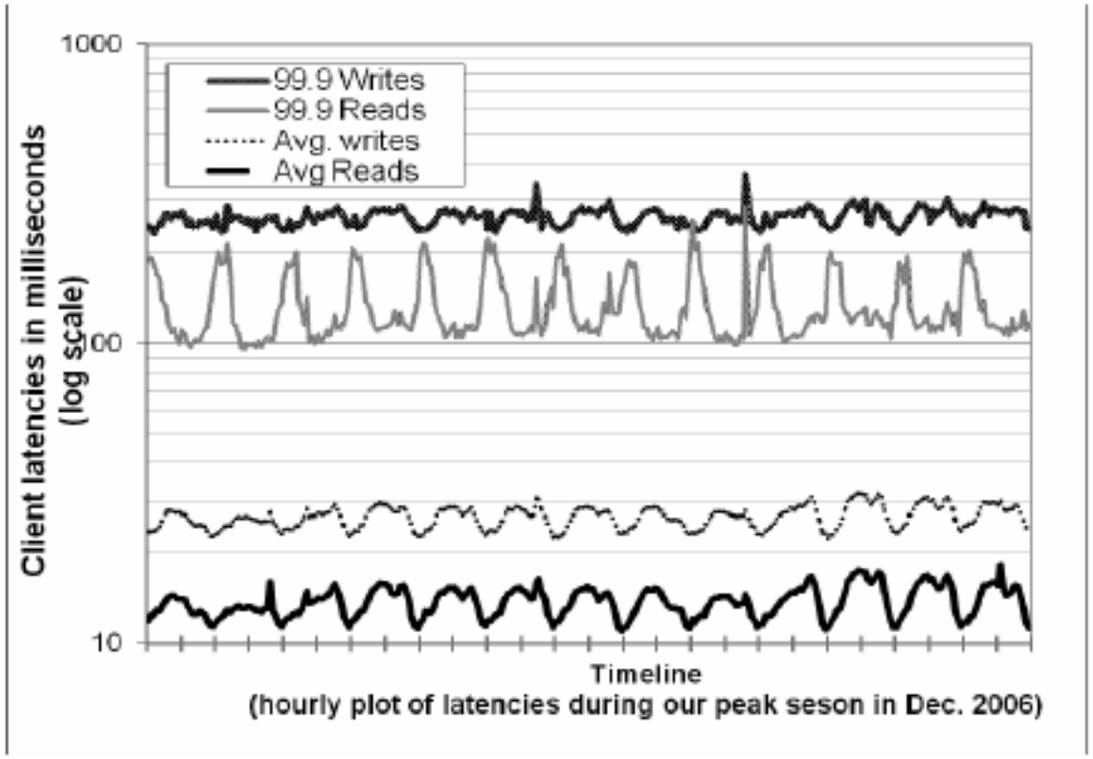


Figure 4: Average and 99.9 percentiles of latencies for read and write requests during our peak request season of December 2006. The intervals between consecutive ticks in the x-axis correspond to 12 hours. Latencies follow a diurnal pattern similar to the request rate and 99.9 percentile latencies are an order of magnitude higher than averages

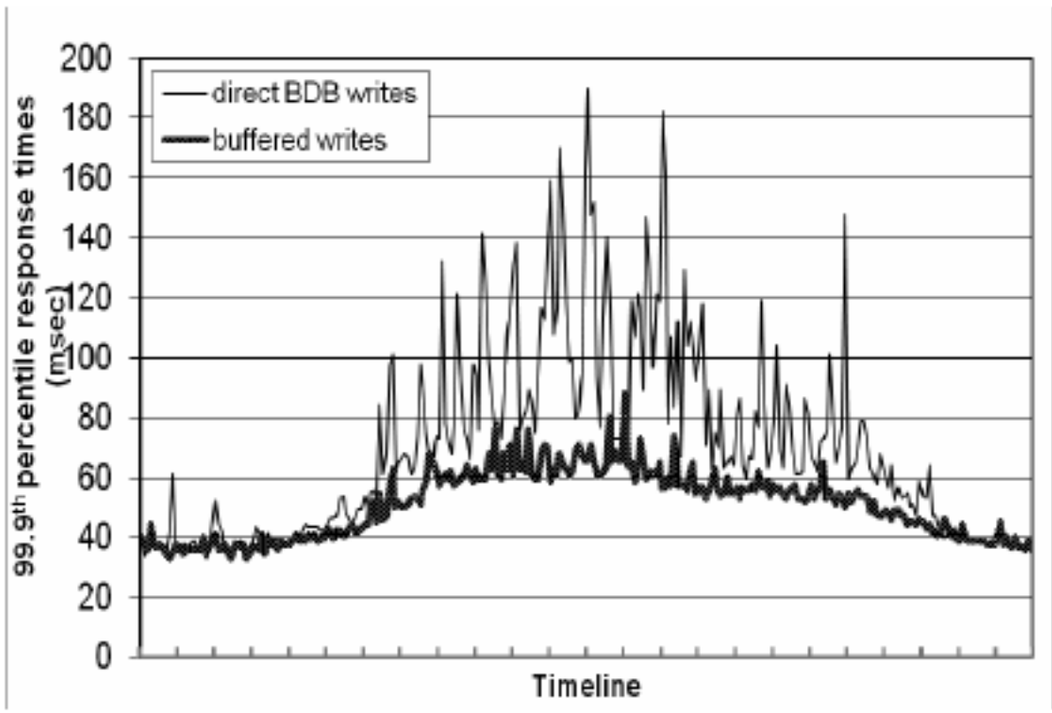


Figure 5: Comparison of performance of 99.9th percentile latencies for buffered vs. non-buffered writes over a period of 24 hours. The intervals between consecutive ticks in the x-axis correspond to one hour.

Zusammenfassung

- Replikation
  - synchron: Integrität, Konsistenz gut, Verfügbarkeit problematisch
  - asynchron: Verfügbarkeit gut, Konsistenz problematisch
- Partitionierung
  - Verfügbarkeit, Performance gut, Integrität problematisch
  - Spezialfall sharding

Nochmal dynamo

- „always writeable“
  - write request werden geschrieben, auch wenn nicht alle Replikas verfügbar sind
- Konfliktauflösung beim Lesen, Versionierung

Versionen	Anteil
1	99,94 %
2	0,00057 %
3	0,00047 %
4	0,00009 %

Anzahl d. Versionen bei shopping cart Funktion, 24 Stunden Mittel