2 case studies
models with
# atypical behavior

09/11/08

case study
1 class of users
influence of
think time on performance

objectives

- analyze the influence of think time Z on the system performance (response time R, throughput X, utilization U, power X/R, number of users in think state and in execution)

- the workload, number of users of the system, is fixed (N=300), the service requests and the service times are kept constant

- think time Z changes from 1 to 1000 sec
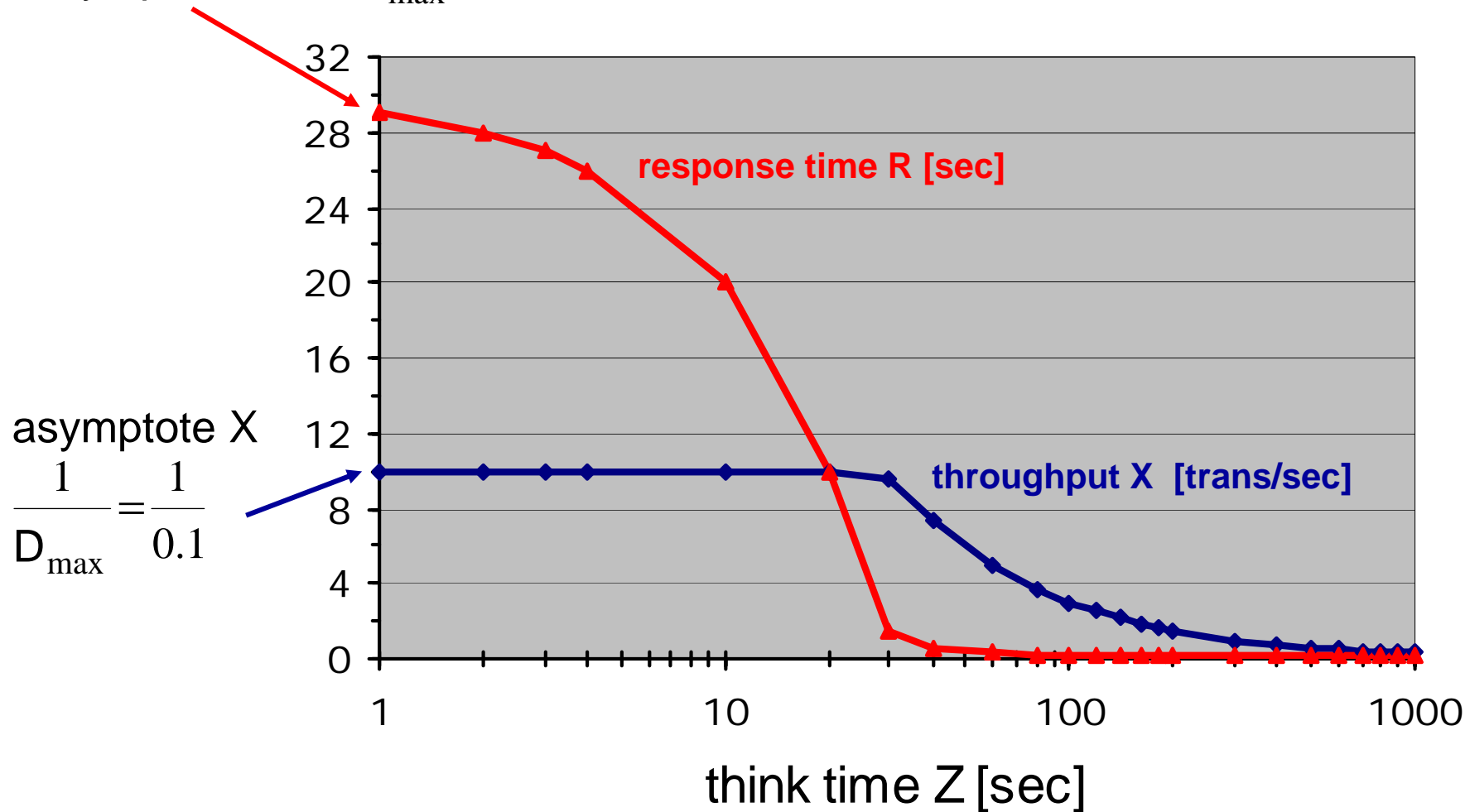
system topology
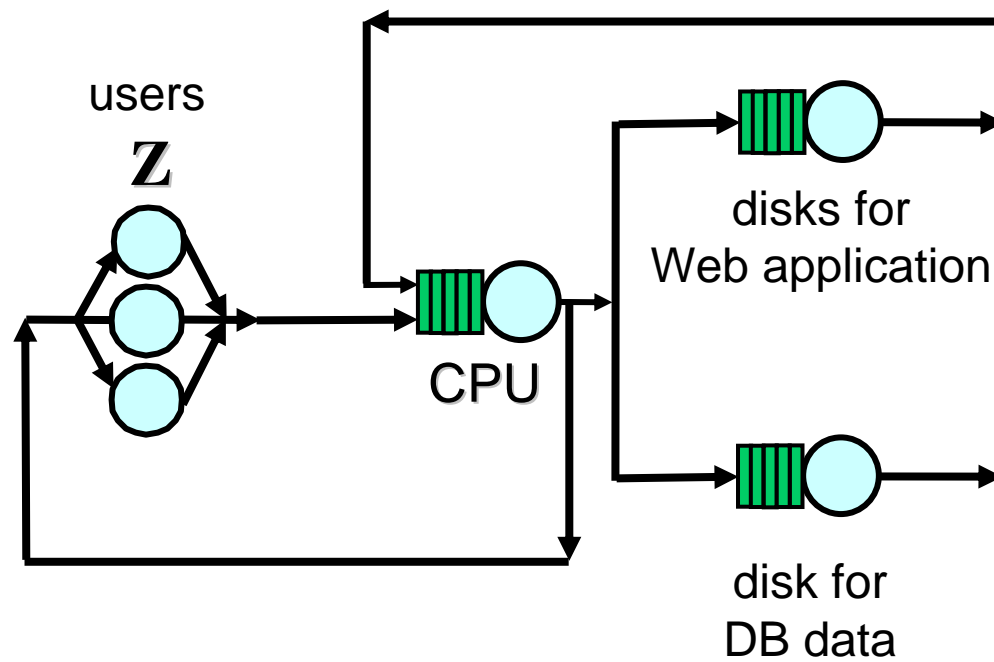
workload (D vector)

workload vector
$D_i = V_i * S_i$ [sec]

resources	D [sec]
think time (Z)	1-1000
CPU	0.100
Web applic.	0.060
DB data	0.020

bottleneck
D_{max}

throughput X and response time R vs Z

asymptote $R = ND_{max} - Z = 300*0.1 - 1 = 29\,\text{sec}$

asymptote X
$$\frac{1}{D_{max}} = \frac{1}{0.1}$$

response time R [sec]

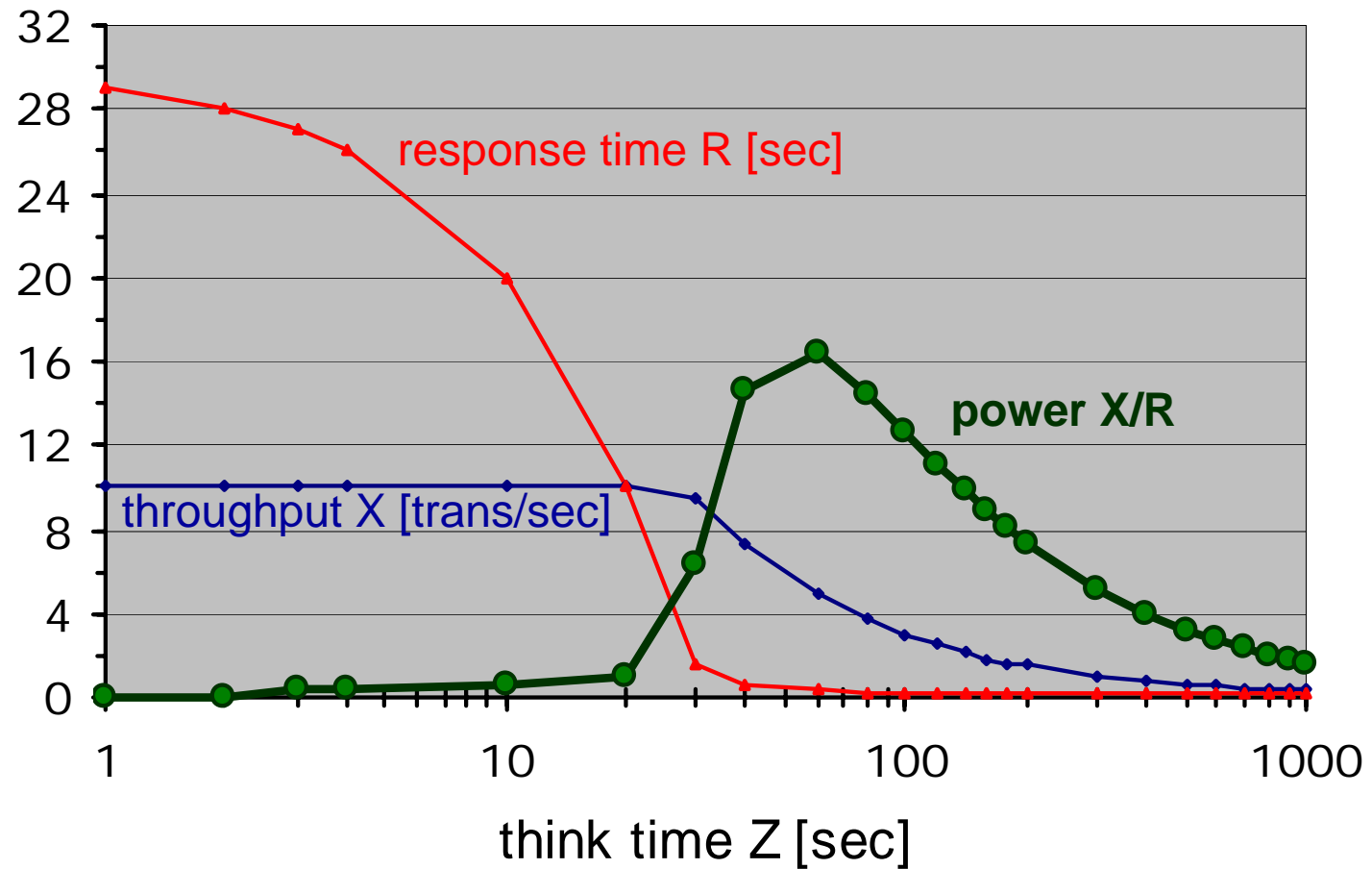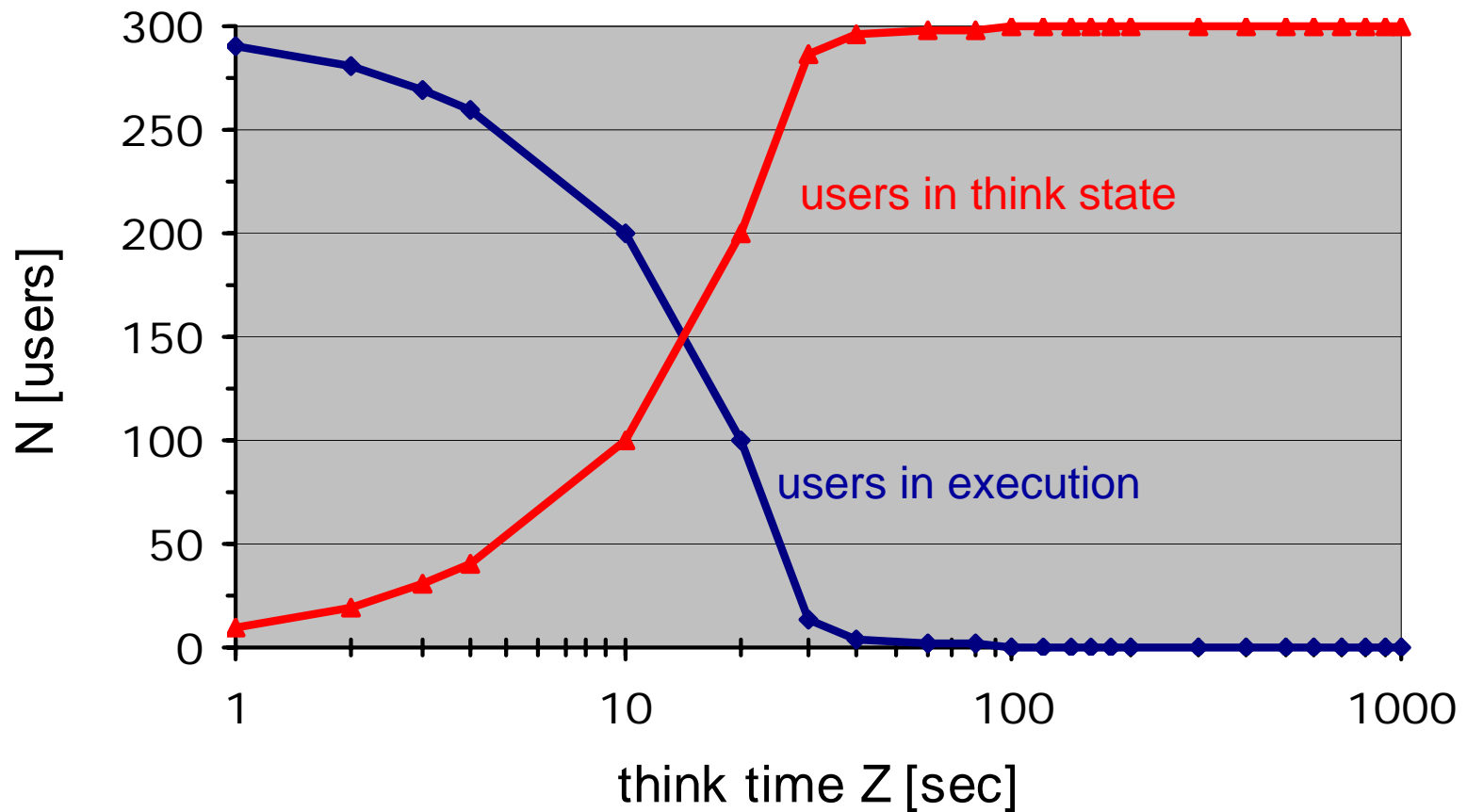throughput X [trans/sec]

think time Z [sec]

utilizations
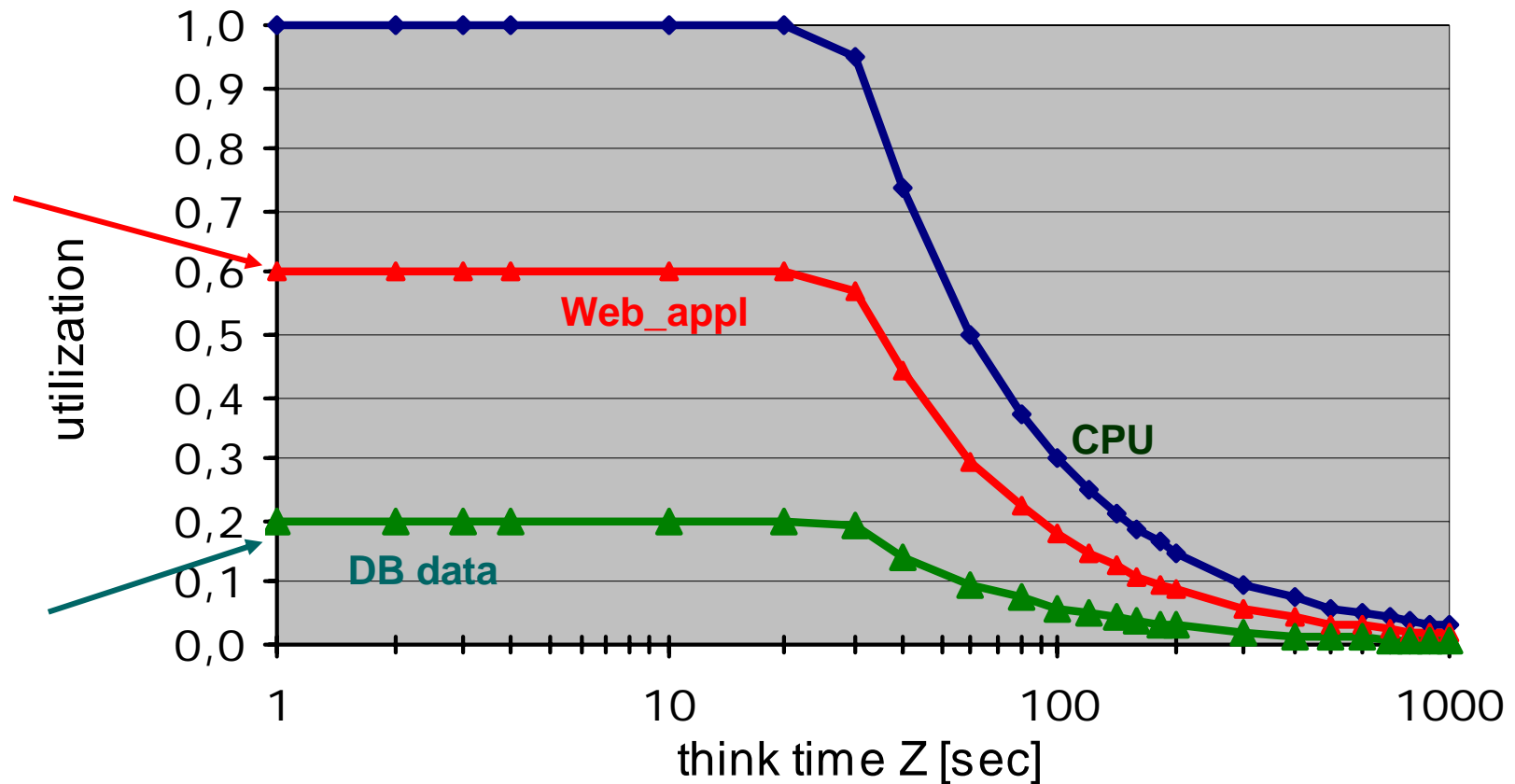
$$U_{CPU} = X D_{CPU} \qquad U_{wappl} = X D_{wappl} \qquad U_{db} = X D_{db}$$

$$\frac{U_{wappl}}{U_{CPU}} = \frac{D_{wappl}}{D_{CPU}} = \frac{0.06}{0.10} = 0.6 \qquad \frac{U_{DB}}{U_{CPU}} = \frac{D_{DB}}{D_{CPU}} = \frac{0.02}{0.10} = 0.2$$

system power: X/R vs Z

response time R [sec]

power X/R

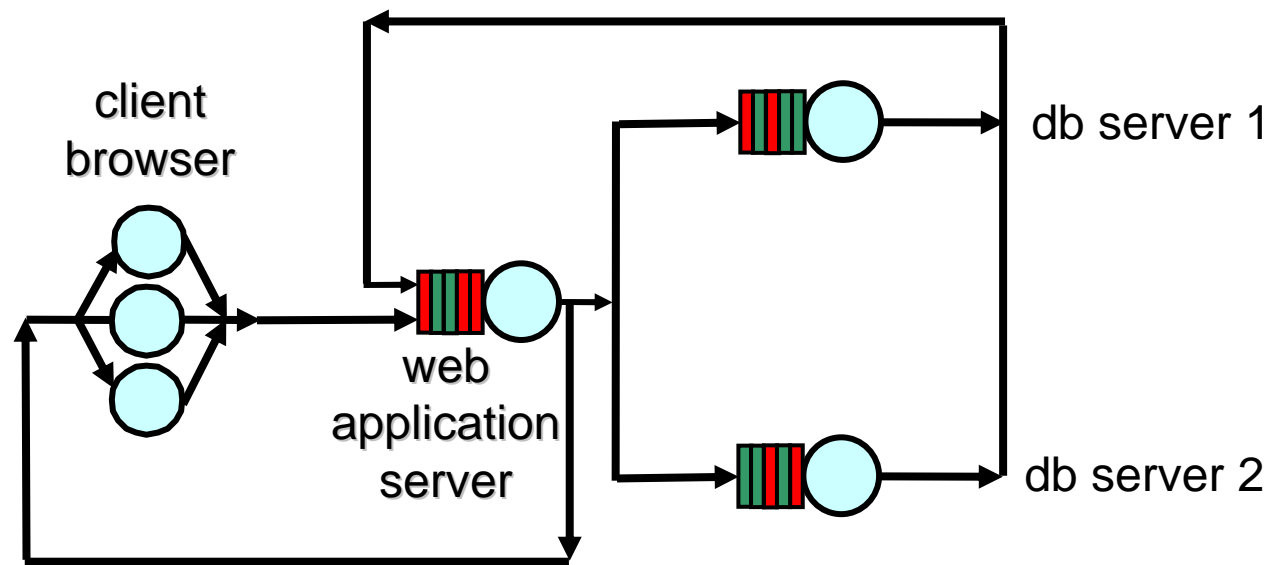throughput X [trans/sec]

think time Z [sec]

users in think state and in execution vs Z

case study
2 classes of requests
unbalanced workload increase

system topology

client
browser

web
application
server

db server 1

db server 2

workload behavior

- two classes of web requests: light e heavy

- balanced growth, up to N=200 requests (50% cl.1, 50% cl.2)

- unbalanced workload growth after 200 requests

  1. only light class users increase

  2. only heavy class users increase

- evaluate the performance indexes behavior (utilization, throughput, response time)

service demands matrix D (2 classes)

resources	class 1 light users	class 2 heavy users
browser think time Z	5	60
web application server	0.100	0.030
DB server 1	0.060	0.040
DB server 2	0.020	0.200

natural bottleneck
class 1
$D_{CPU}=D_{max}=0.1$

natural bottleneck
class 2
$D_{DB}=D_{max}=0.2$

unbalanced growth: utilization asymptotes

only light users increase

$$U_{appl.server} = XD_{appl.server} = 1 \qquad U_{DB1} = XD_{DB1} \qquad U_{DB2} = XD_{DB2}$$

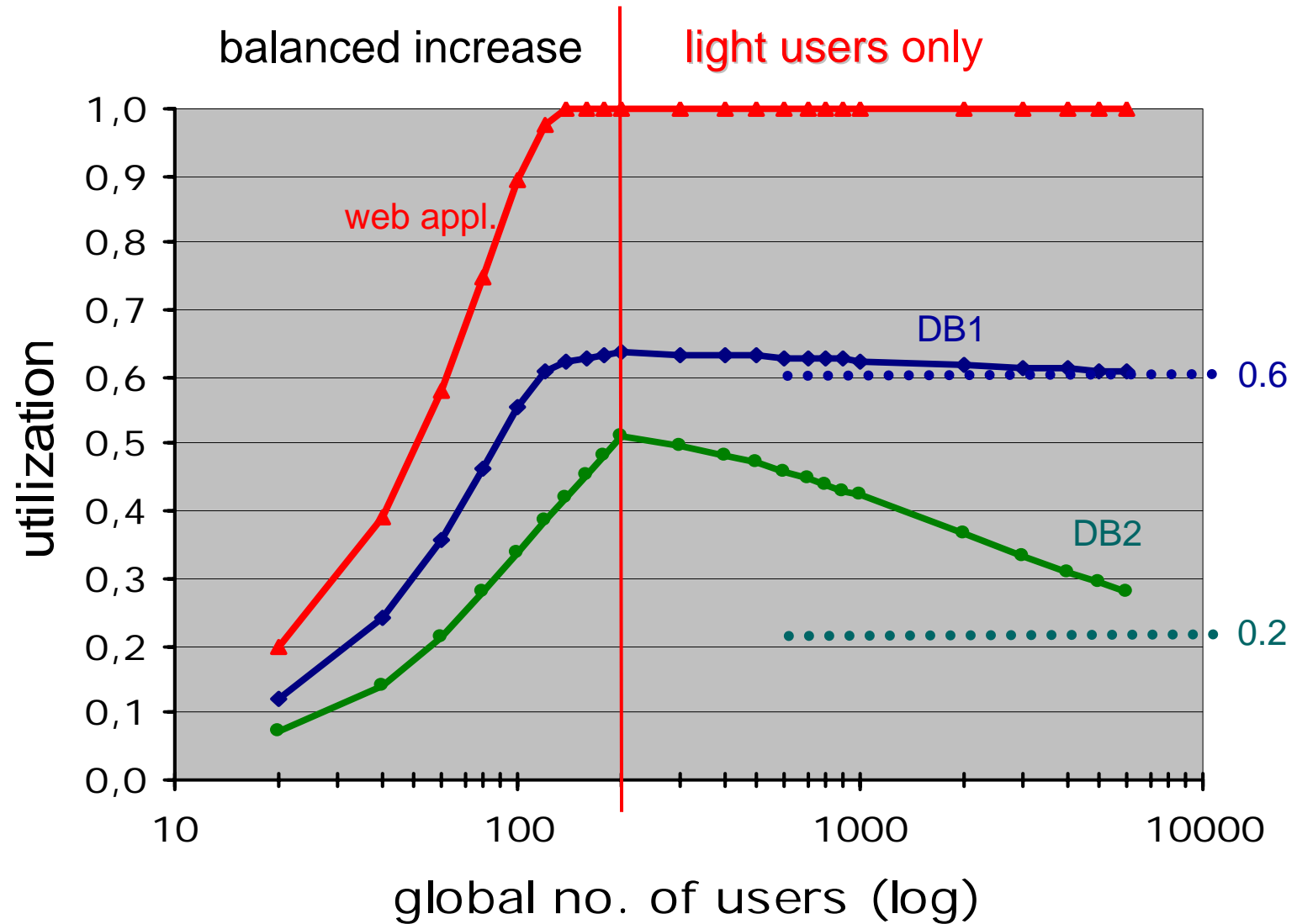$$U_{DB1} = U_{appl.server} \frac{D_{DB1}}{D_{appl.server}} = \frac{0.060}{0.100} = 0.6$$

$$U_{DB2} = U_{appl.server} \frac{D_{DB2}}{D_{appl.server}} = \frac{0.020}{0.100} = 0.2$$
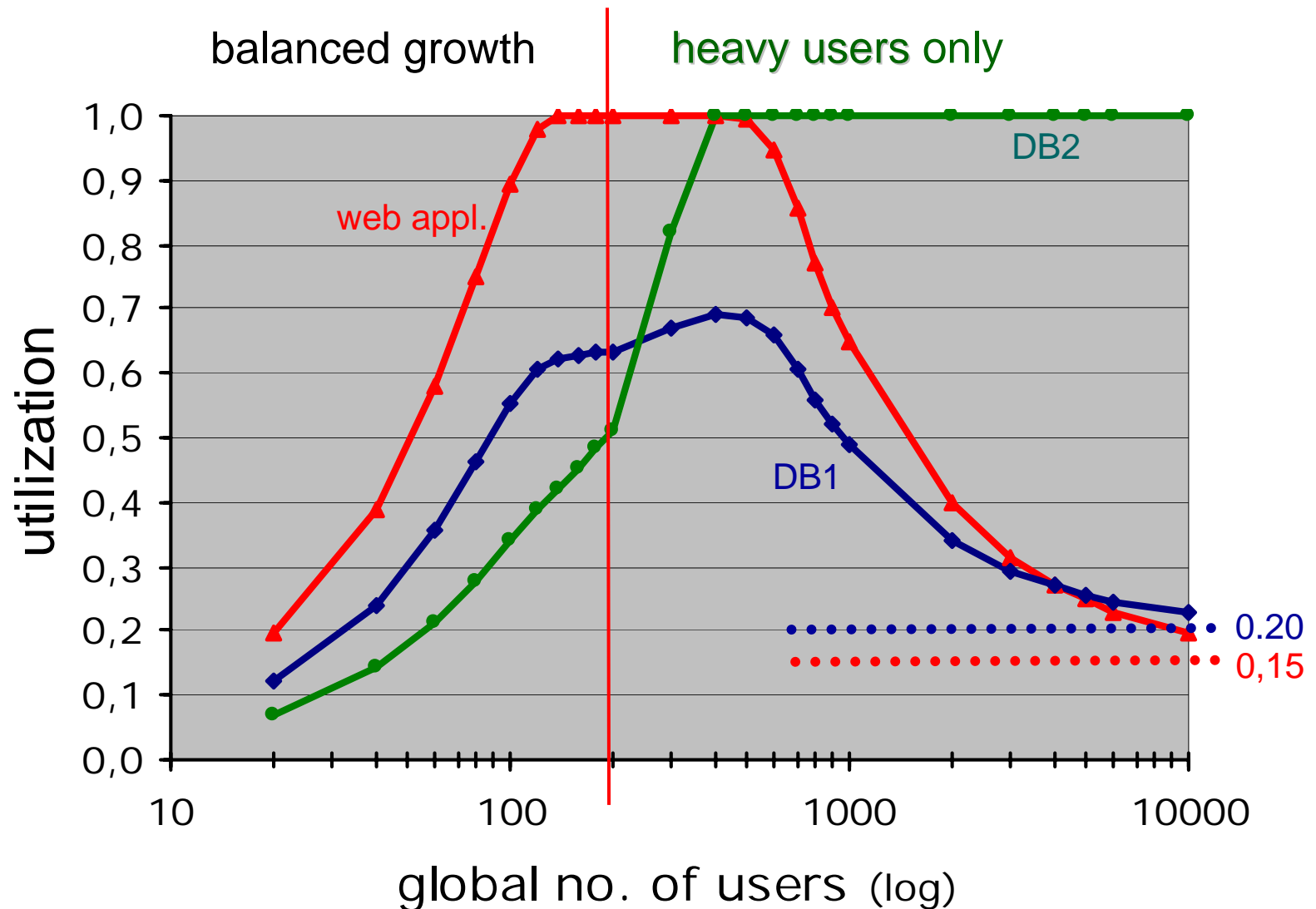
only heavy users increase

$$U_{DB2} = 1 \qquad \frac{U_{DB1}}{U_{DB2}} = \frac{D_{DB1}}{D_{DB2}} \qquad U_{DB1} = U_{DB2} \frac{D_{DB1}}{D_{DB2}} = \frac{0.040}{0.200} = 0.2$$

$$\frac{U_{appl.server}}{U_{db}} = \frac{D_{appl.server}}{D_{db}} \qquad U_{appl.server} = U_{DB2} \frac{D_{appl.server}}{D_{DB2}} = \frac{0.030}{0.200} = 0.15$$
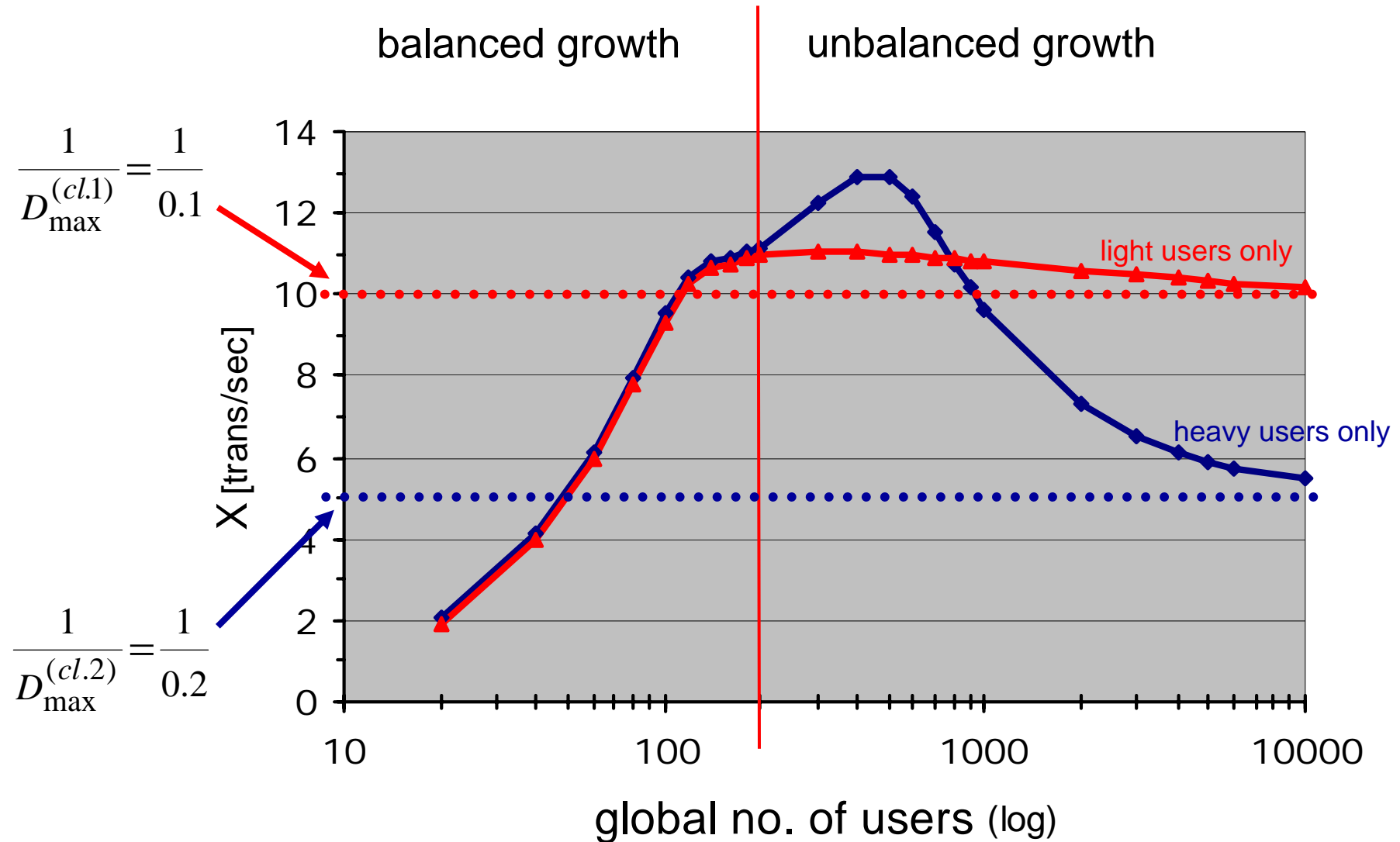
utilizations – only class 1 increases (light users)

utilizations – only class 2 increase (heavy users)

system throughput X [trans/sec]

balanced growth | unbalanced growth

$$\frac{1}{D_{\max}^{(cl.1)}} = \frac{1}{0.1}$$

$$\frac{1}{D_{\max}^{(cl.2)}} = \frac{1}{0.2}$$

light users only

heavy users only

X [trans/sec]

global no. of users (log)

response time R [sec]

balanced growth | unbalanced growth

light requests only

heavy requests only

R [sec]

R heavy

R light

R heavy

R light

global no. of users (log)

2 Atypical models

response time R [sec]

balanced growth | unbalanced growth

······· light requests only

——— heavy requests only