The MIND Lab

The Maryland Information and Network Dynamics Lab

System Performance Using Operational Analysis

Ashok Agrawala

View of a system

- A collection of resources
- A processing request uses these resources in arbitrary order
- Operational view is based on measured quantities

Operational Quantities

- Measured during a finite observation period T
- System has number of device.
- For device i
  - Number of Arrivals (requests) Ai
  - Number of Completions (responses) Ci
  - Busy time Bi

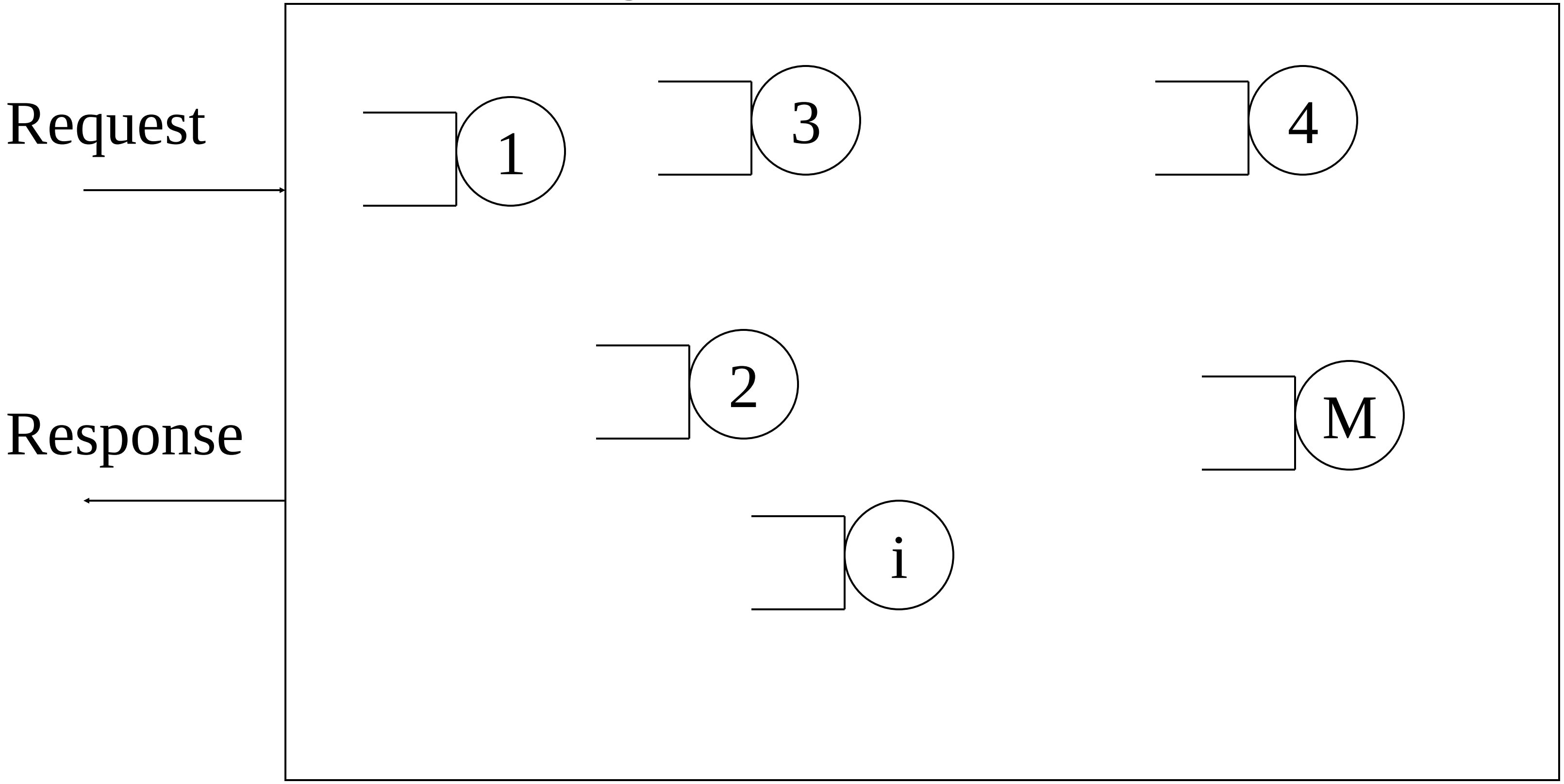
Derived Quantities

- Arrival Rate l i = Ai/T
- Throughput Xi = Ci/T
- Utilization Ui = Bi/T
- Mean service time Si = Bi/Ci

Utilization Law

- $U_i = B_i/T = (C_i/T)(B_i/C_i) = X_i S_i$
- Example – At a gateway packets arrive at a rate of 125 pps and the gateway takes 2 ms to process them.
  - $X_i = 125$ pps, $S_i = 0.002$ sec
  - $U_i = 125 \times 0.002 = 0.25 = 25\%$

System View



Assumption

- Job Flow Balance
- $A_i = C_i$
- Number of jobs at the device same at the beginning and the end of the observation period

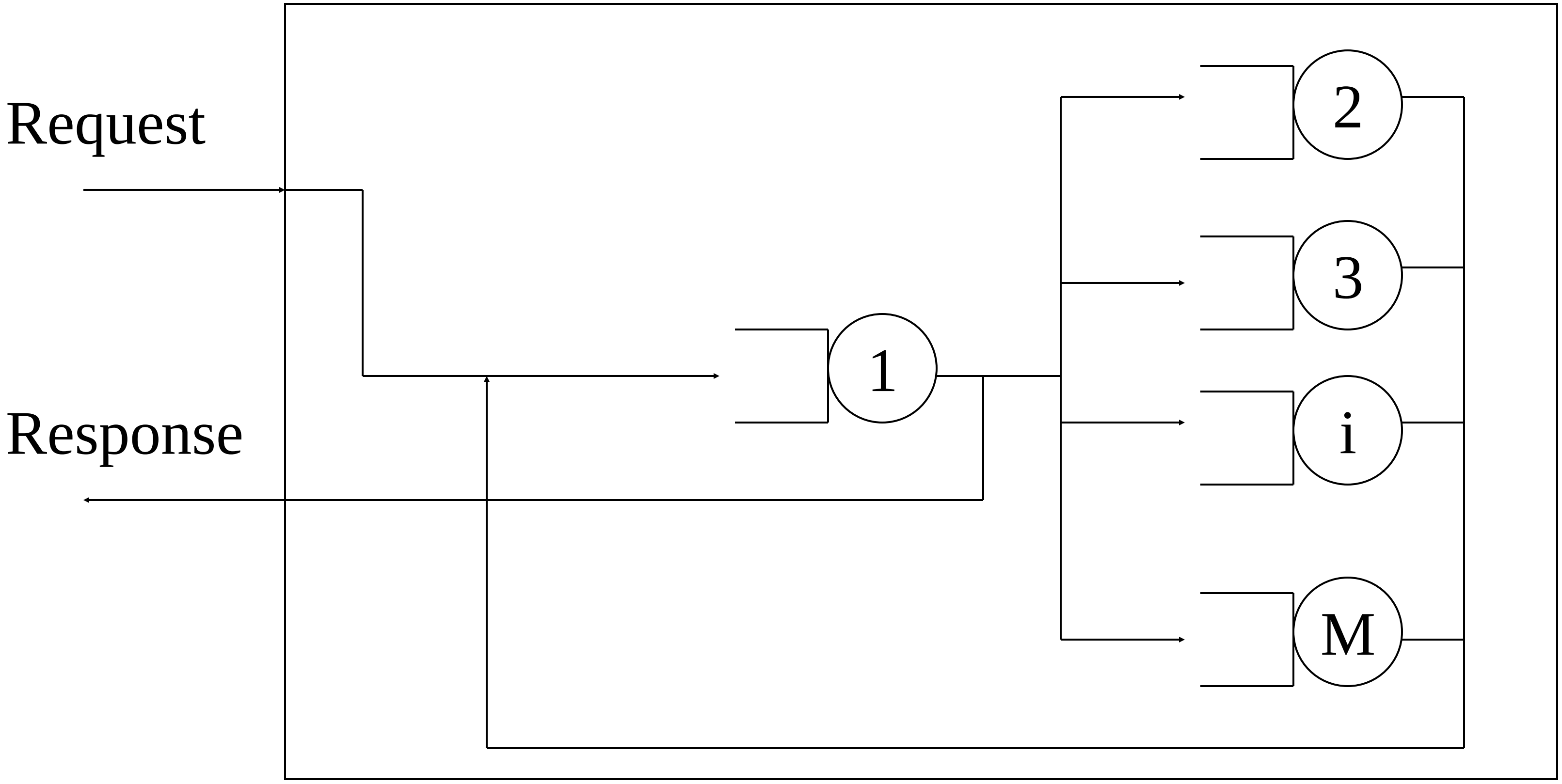
Visit Counts

- Each request makes V_i visits to device i
- C_0 = Number of outside requests
- Then $C_i = C_0 V_i$
- Or $V_i = C_i / C_0$
- System Throughput $X = C_0 / T$
- Device throughput $X_i = C_i / T = C_i / C_0 \times C_0 / T$
- Or $X_i = X V_i$
- Forced Flow Law !!

Device Utilization

- $U_i = X_i S_i = X V_i S_i = X D_i$
  - Where $D_i = V_i S_i$
  - The total service demand for the device i by a request
- Utilization of devices is proportional to D_i
- Device with the highest D_i is the Bottleneck device – It saturates first.

Central Server Model



Little's Law

- Queue length at device i = Qi
- Response Time at device i = Ri

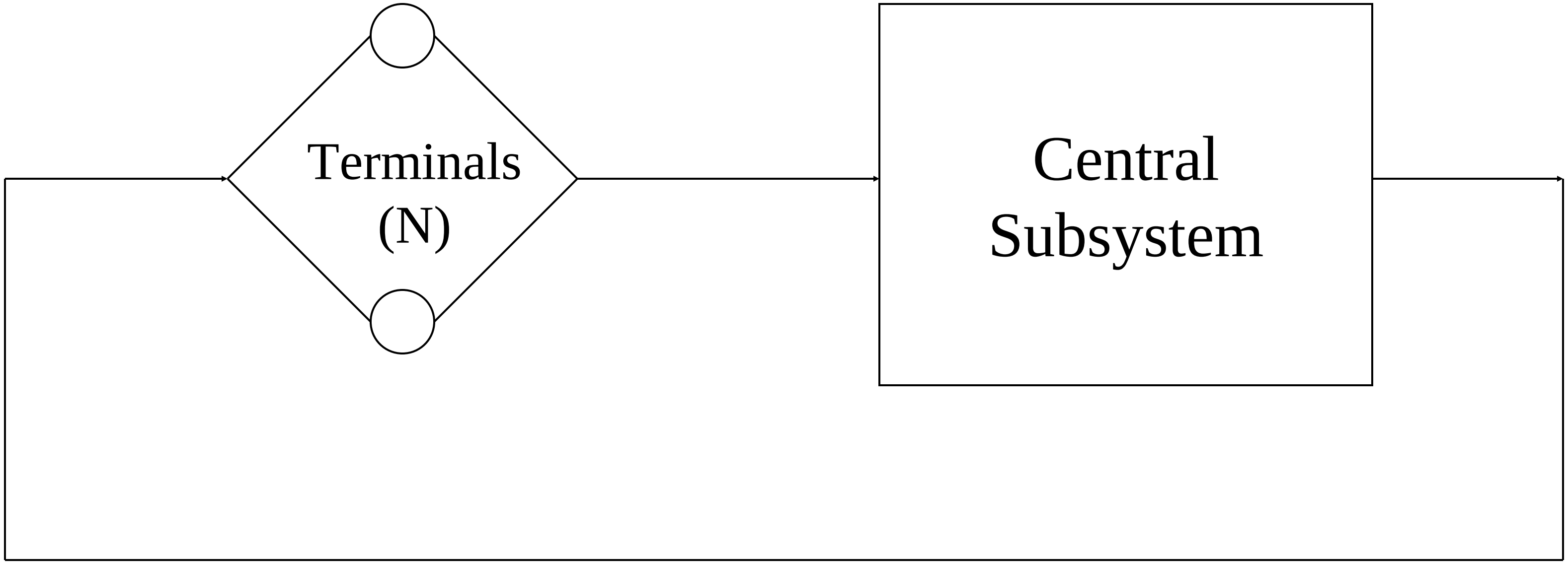
- Qi = XiRi
- Can be applied to any part of the system as long as the flow balance holds for that part

General Response Time Law

- $Q = XR$
- $Q = Q_1 + Q_2 + \dots + Q_M$
- $XR = X_1R_1 + X_2R_2 + \dots + X_MR_M$
- $R = V_1R_1 + V_2R_2 + \dots V_MR_M$
- $= D_1 + D_2 + \dots + D_M$

The MIND Lab
The Maryland Information and Network Dynamics Lab

Interactive System



Interactive Response Time Law

- Let Z be the think time
- Cycle time = R+Z
- X = (Total Number of Requests)/(Total Time)
- = N[T/(R+Z)]/T
- = N/(R+Z)
- R = (N/X) -Z

Bottleneck Analysis

- Bottleneck Device – Device that saturates first – Utilization goes to 1.
- $U_i = X D_i$
- Device with highest D_i is the Bottleneck device

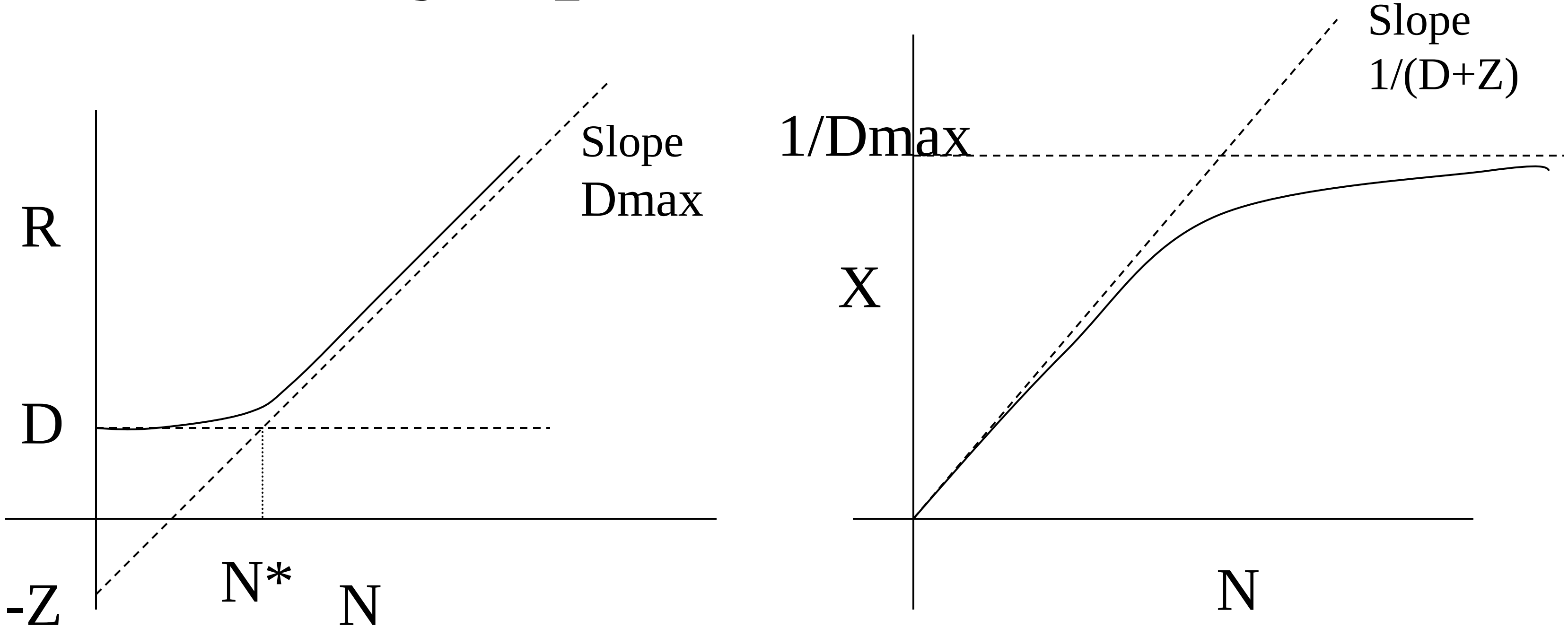
Asymptotic Bounds

- $X(N) \leq \min \{ 1/Dmax , N/(D+Z) \}$

- $R(N) \geq \max \{ D, (NDmax – Z)\}$
  - Where $D = D1+D2+ \dots +DM$

The Maryland Information and Network Dynamics Lab

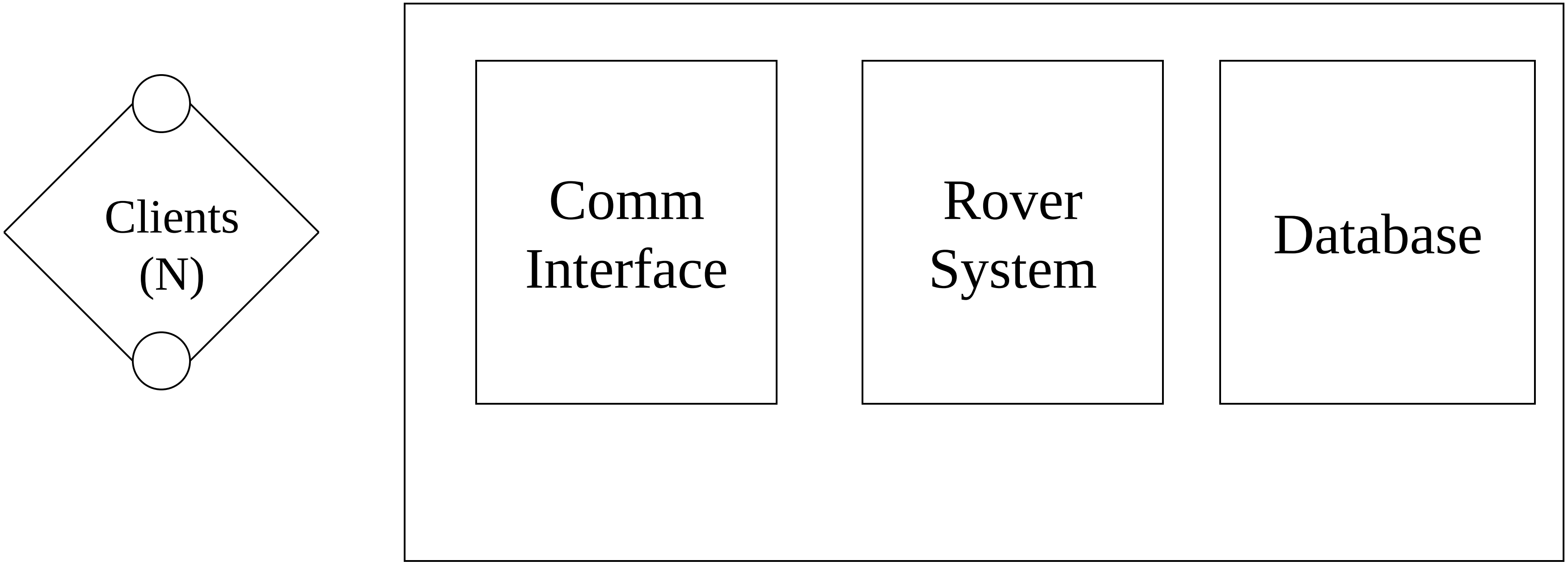
Asymptotic Bounds



N* = (D+Z)/Dmax

The Maryland Information and Network Dynamics Lab

Rover Model



Rover Performance

- Script
  - Operations with defined Think Time
- Three Scripts
  - Vector Map
  - Get All Login (GAL)

Client Script

- Send OP Request
- Wait for Response
- Think
- Loop

Server Actions for Vector Map

- Simple Pass Through – to DB
- Simple I/O Response Handler – From DB to Op State
- Simple Reply - send info to Comm Manager
- Simple Cleanup

Vector Map Script

- Single Client Times – Z=2
  - Database 531 ms
  - Server 621 ms
  - Client 669 ms

The Maryland Information and Network Dynamics Lab

Vector Map

DB

Z	1	2	3
1	535	531	537
2	847	532	534
3	997	685	535
4	1277	907	801
5	1807	1049	847
6	2292	1494	1027
7	2943	1991	1207
8	3448	2566	1700
9	4100	3156	2195
10	4681	3672	2804

Server

z	1	2	3
1	714	621	616
2	1161	618	623
3	1240	775	621
4	1404	998	890
5	1983	1156	939
6	2496	1635	1127
7	3139	2175	1327
8	3655	2747	1858
9	4290	3345	2380
10	4862	3891	3000

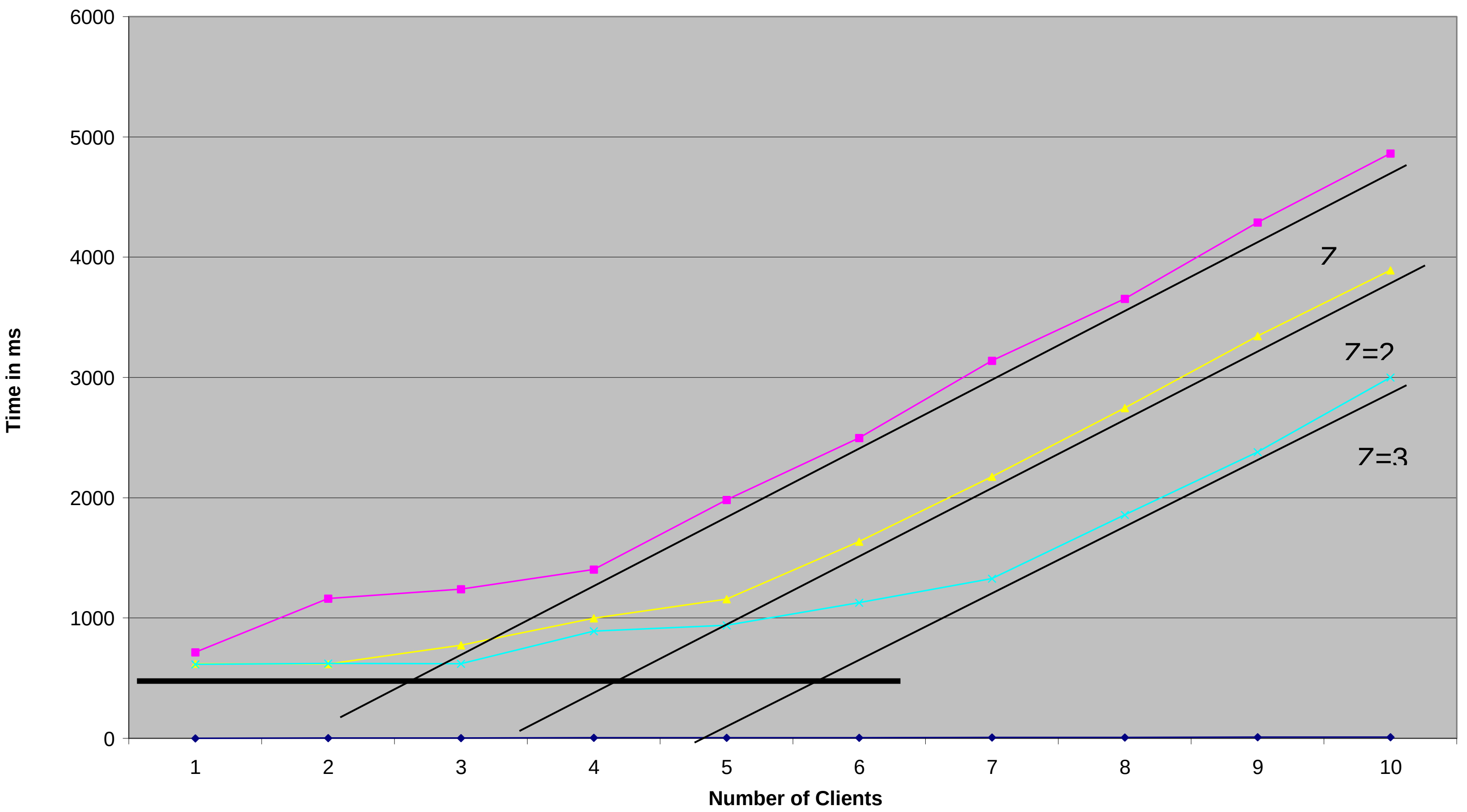
Client

z=	1	2	3
1	818	669	685
2	1298	679	688
3	1393	839	682
4	1458	1067	964
5	2035	1220	1020
6	2550	1696	1195
7	3186	2224	1390
8	3704	2797	1907
9	4359	3393	2431
10	4920	3933	3046

MIND
The MIND Lab
The Maryland Information and Network Dynamics Lab

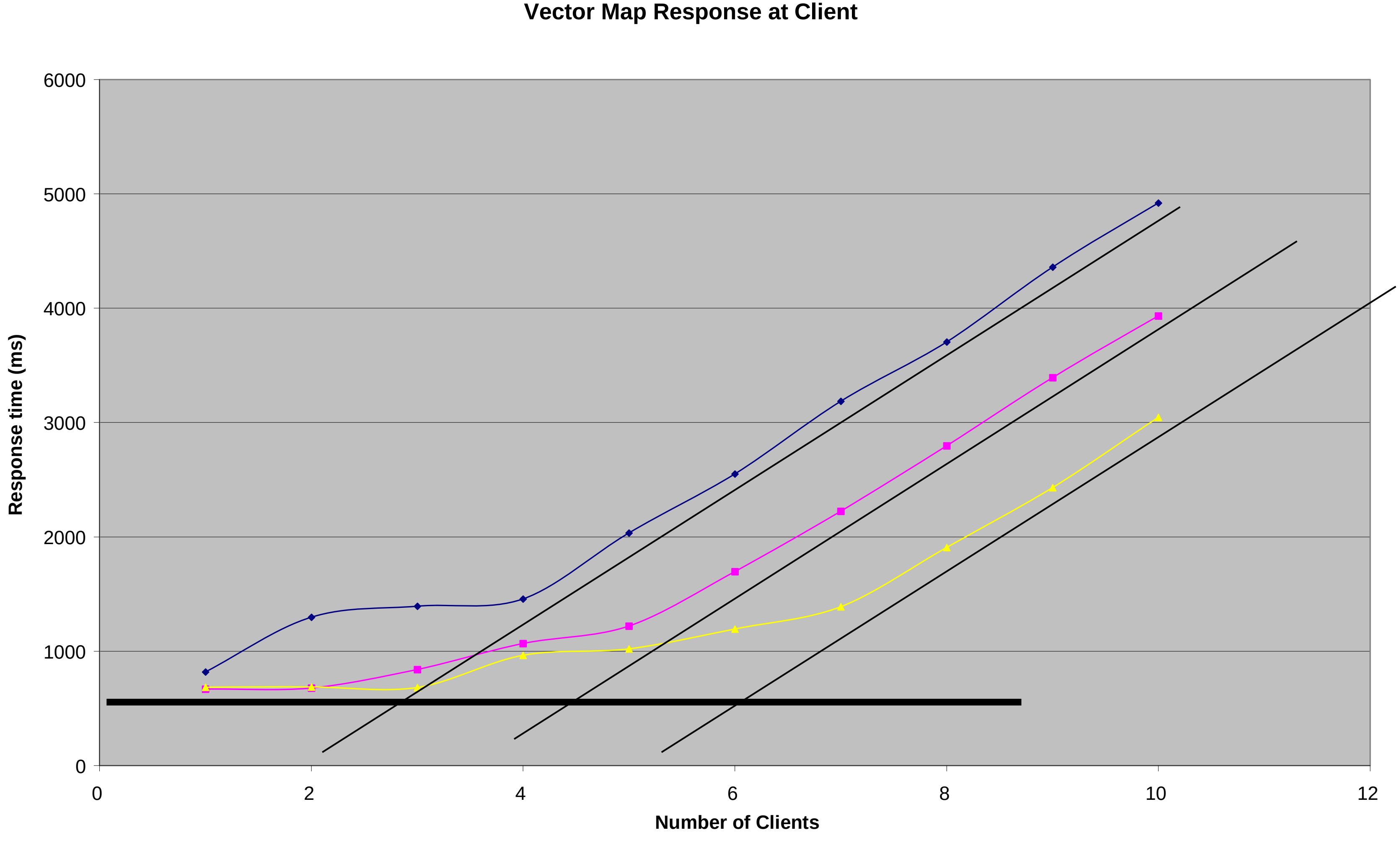
Vector Map Response Time at Server

Response time at Server for Vector Map

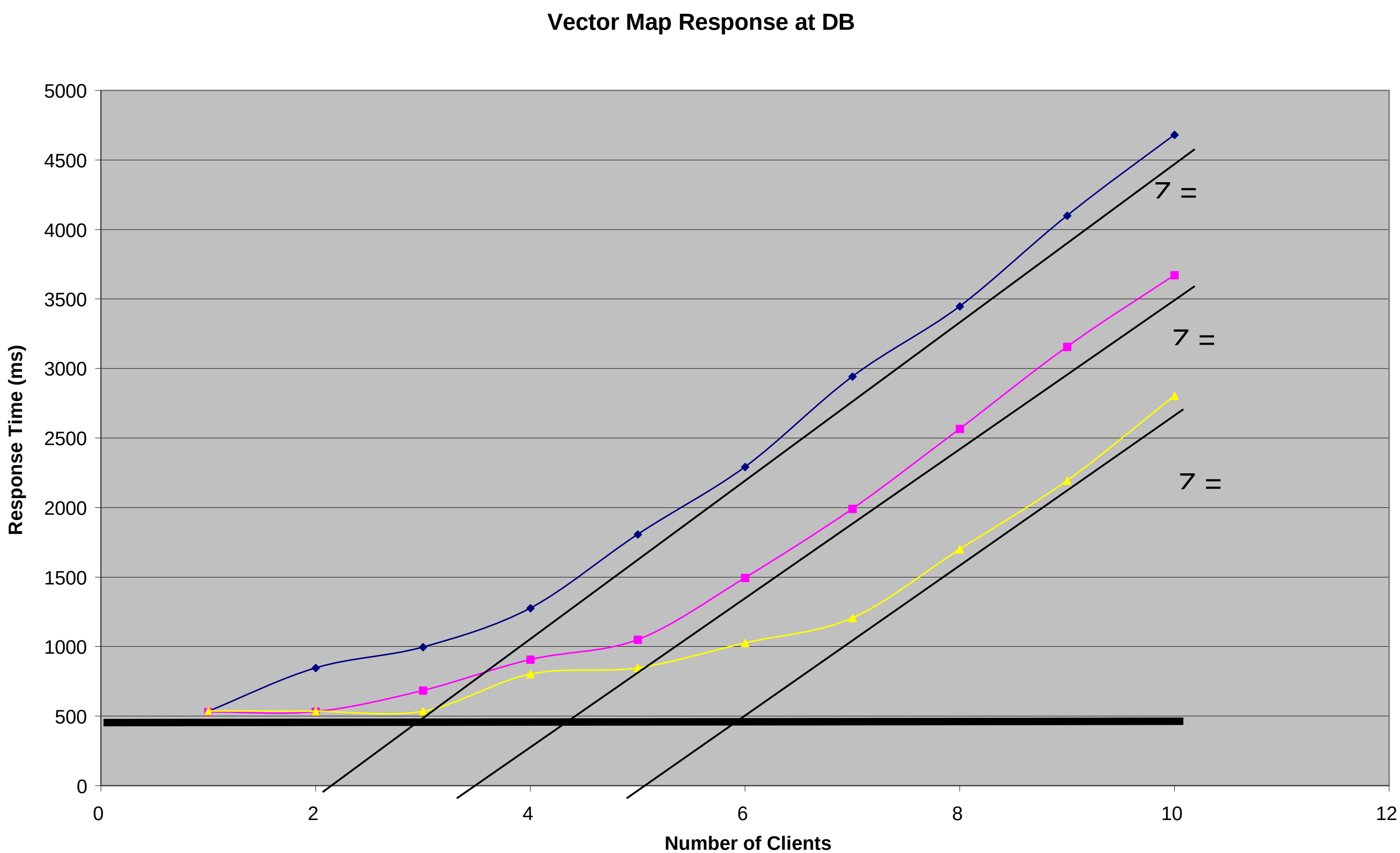


The Maryland Information and Network Dynamics Lab

Vector Map Response Time at Client



Vector Map Response at DB



GAL Script

- Single Client
  - Server 296 ms
  - Client 2,912 ms
- Communications Time of 2 ms
  - Using 802.11 connection

The Maryland Information and Network Dynamics Lab

GAL Response Time

	Cl 100	Cl 200	Sr 100	Sr 200
1	2912	4755	296	325
3	4265	3836	276	276
5	3647	4295	299	314
10	4513	6261	417	358
15	17346	9948	415	409
20	23863	14245	447	491
30	28566	23783	741	809
50	136606	50443	1644	1211
100	785828		2120	

GAL Response Time