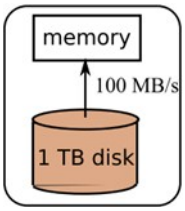
Family name: …… Given name: ……

Family name: …… Given name: ……

Family name: …… Given name: ……



Type	Latency	Bandwidth
Disk	$\approx 5 \times 10^{-3}$s (5 millisec.);	At best 100 MB/s
LAN	$\approx 1 - 2 \times 10^{-3}$s (1-2 millisec.);	$\approx$ 1GB/s (single rack); $\approx$ 100MB/s (switched);
Internet	Highly variable. Typ. 10-100 ms.;	Highly variable. Typ. a few MB/s.;

Bottom line (1): it is approx. one order of magnitude faster to exchange main memory data between 2 machines in a data center, that to read on the disk.

Bottom line (2): exchanging through the Internet is slow and unreliable with respect to LANs.

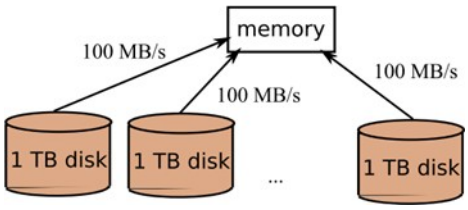
a) *How long would it take to read 1TB with sequential access? (in secs)*

b) How long would a single random access (i.e., reading one tuple, of few bytes, through an index) take? (in secs)

Family name: …… Given name: ……

Family name: …… Given name: ……

Family name: …… Given name: ……



Type	Latency	Bandwidth
Disk	$\approx 5 \times 10^{-3}$s (5 millisec.);	At best 100 MB/s
LAN	$\approx 1 - 2 \times 10^{-3}$s (1-2 millisec.);	$\approx$ 1GB/s (single rack); $\approx$ 100MB/s (switched);
Internet	Highly variable. Typ. 10-100 ms.;	Highly variable. Typ. a few MB/s.;

Bottom line (1): it is approx. one order of magnitude faster to exchange main memory data between 2 machines in a data center, that to read on the disk.

Bottom line (2): exchanging through the Internet is slow and unreliable with respect to LANs.

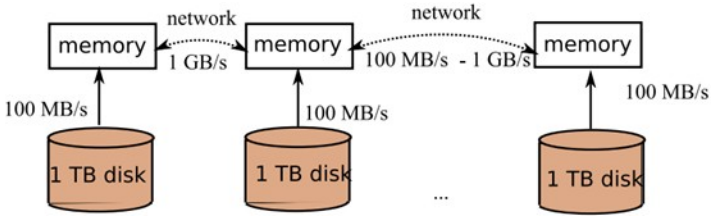
c) *How long would it take to read 1TB with parallel access (fig. b)? Assume 100 disks (i.e., 100 replicas of the whole data) on the same machine with shared-memory and infinite CPU capacity.*

d) How long would a single random access (i.e., reading one tuple, of few bytes, through an index) take? (in secs)

Family name: …… Given name: ……

Family name: …… Given name: ……

Family name: …… Given name: ……



Type	Latency	Bandwidth
Disk	$\approx 5 \times 10^{-3}$s (5 millisec.);	At best 100 MB/s
LAN	$\approx 1 - 2 \times 10^{-3}$s (1-2 millisec.);	$\approx$ $1\mathrm{GB/s}$ (single rack); $\approx 100\mathrm{MB/s}$ (switched);
Internet	Highly variable. Typ. 10-100 ms.;	Highly variable. Typ. a few MB/s.;

Bottom line (1): it is approx. one order of magnitude faster to exchange main memory data between 2 machines in a data center, that to read on the disk.

Bottom line (2): exchanging through the Internet is slow and unreliable with respect to LANs.

e) *How long would it take to read 1TB with distributed access (fig. c)? Assume 100 shared-nothing machines (with all data replicated in each of them) in a star-shape LAN in a single rack where all data is sent to the center.*

f) How long would a single random access (i.e., reading one tuple, of few bytes, through an index) take? (in secs)