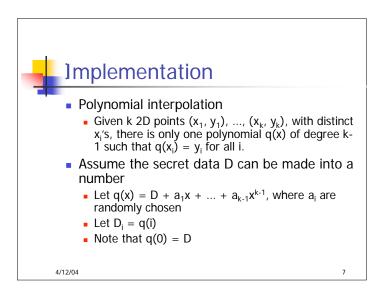
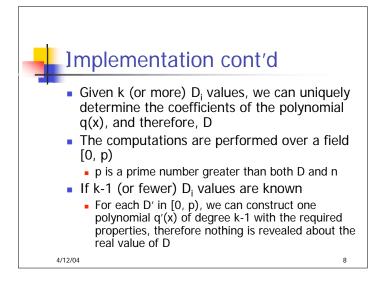


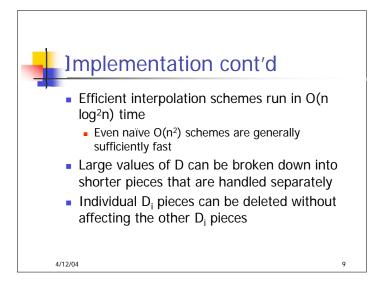
Applications cont'd

4/12/04

- Threshold schemes are ideal when mutually suspicious individuals, with conflicting interests, must cooperate
- A sufficiently large majority can take action
- A sufficiently large minority can veto









- All the D_i pieces can be changed at once without affecting the original D
- The D_i pieces can be shared differently based on their importance
 - The CEO gets 3 pieces
 - The VPs get 2 pieces

4/12/04

• The middle managers get 1 piece

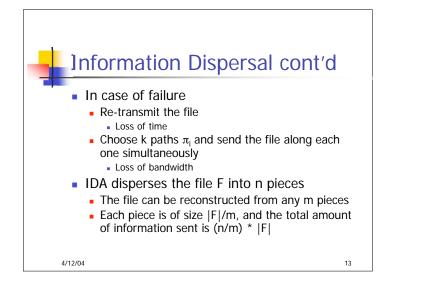
Efficient Dispersal of Information for Security, Load Balancing, and Fault Tolerance

> Michael O. Rabin, Harvard Journal of the ACM April 1989, Vol. 36, Nr. 2

> > 11

4/12/04

<text><list-item><list-item><list-item><list-item></table-row>



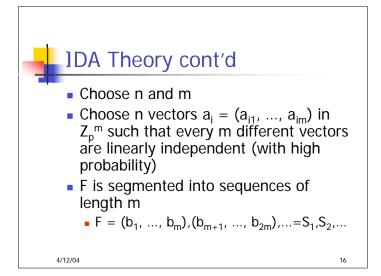
Information Dispersal cont'd

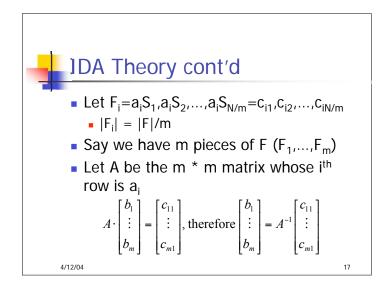
- Space efficiency
 - We can choose n and m such that (n/m)~1, therefore the overhead is low
- Time efficiency

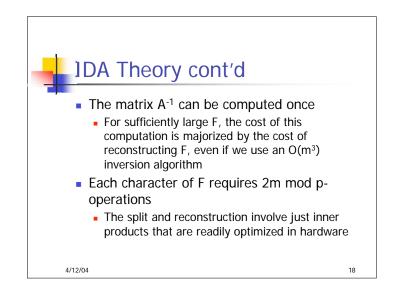
4/12/04

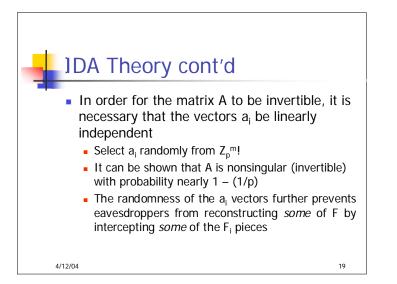
- The splitting and reconstruction algorithms are efficient (more later)
- File pieces can be transmitted in parallel, which better utilizes network resources

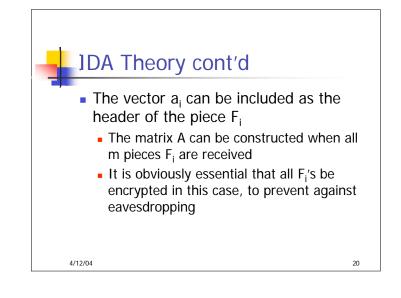
DA Theory
Let F=b₁b₂...b_N be a file, where b_i are in the range [0, B]
We want to disperse pieces of F with the assumption that no more than k pieces will be lost in transmission
Choose p such that p > B
If the file consists of bytes, p = 257
All the following computations are in Z_p

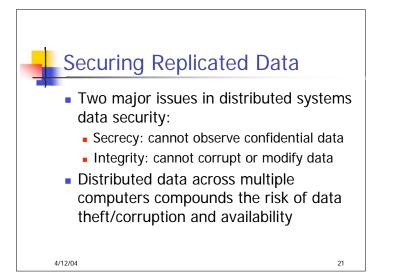








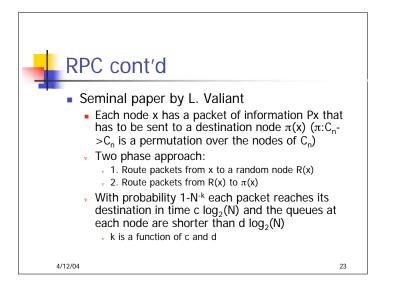


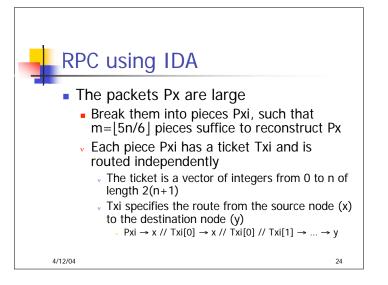


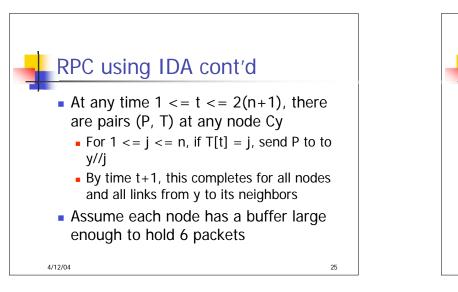


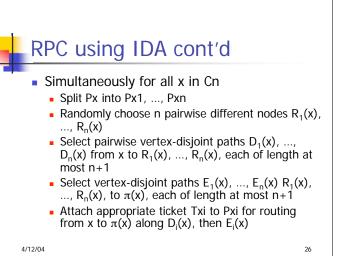
- PC_n = parallel computer with N=2ⁿ nodes
 - Each node x contains a processor Cx and memory Mx
 - A node is of the form {0, 1}ⁿ
 E.g. {0, 1, 0, 1}
 - The notation x//i means that we flip bit i
 {0, 1, 0, 1}//2 = {0, 0, 0, 1}
 - Each node x is connected by two-way links to each of the nodes x//i, where 1<=i<=n</p>

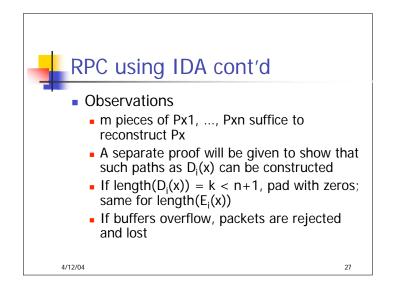


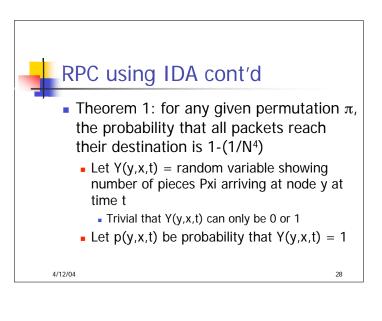


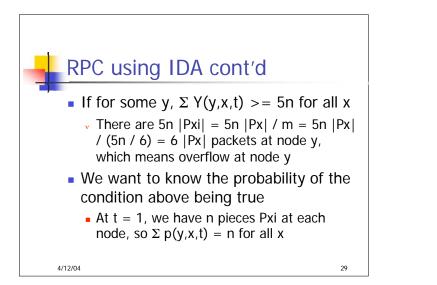




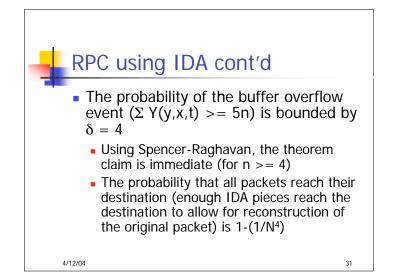


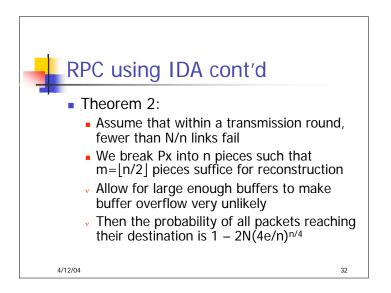


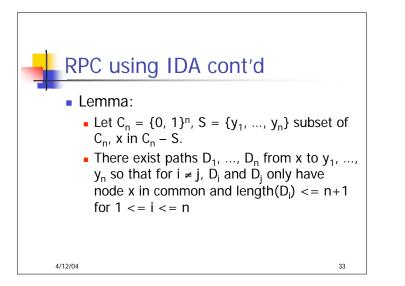


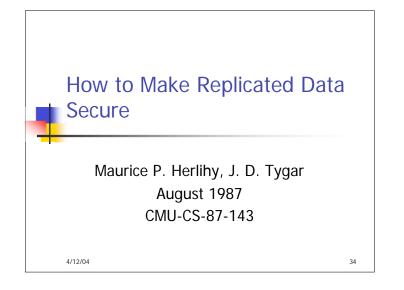


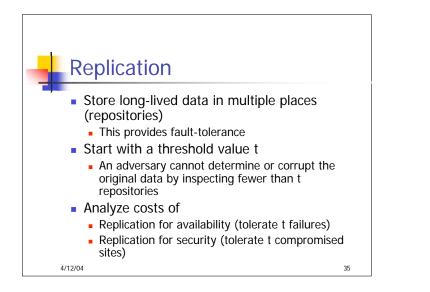
• The random variables Y(y,x,t) are pairwise independent • Use Raghavan-Spencer theorem • If Y1, ..., Yn are independent Bernoulli trials with expected sum n, then for $\delta > 0$: $Pr\left(\sum_{i} Y_{i} \ge (1+\delta)n\right) \le \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{n}$

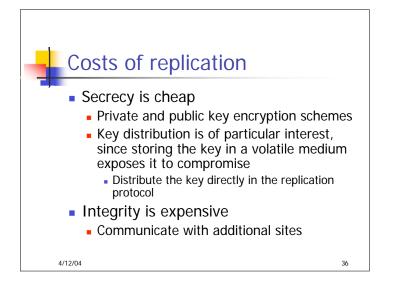


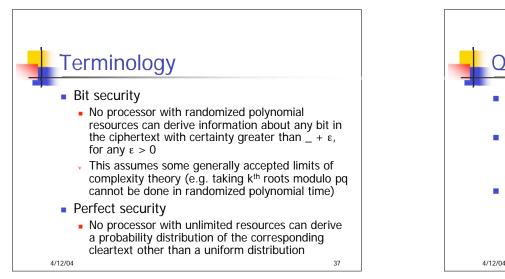














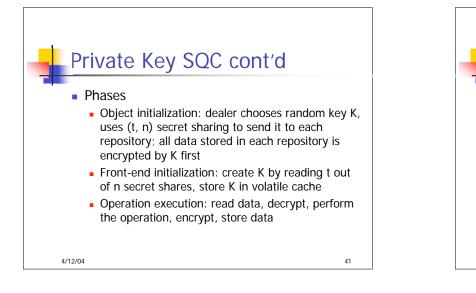
- Repository
 - Long term storage for object state
- Quorum
 - A set of repositories whose cooperation suffices for an operation
- Assignment
 - Associate an operation with a set of quorums

Ouorum Consensus Repl.
 Replicated file

 A collection of timestamped versions
 To read: take latest version from read quorum
 To write: generate new time-stamp, record new version at write quorum

 An assignment is correct iff each read quorum has a non-empty intersection with each write quorum

 Private Key Secure Quorum Consensus (SQC)
 Protect secrecy against an adversary who can observe < t repositories
 Depends on a bit-secure, probabilistic private key encryption scheme
 Implementation
 Front-ends: clients, volatile store
 Repositories: connected, long-term store
 Dealer: communicate with repositories, has a source of random bits, volatile store

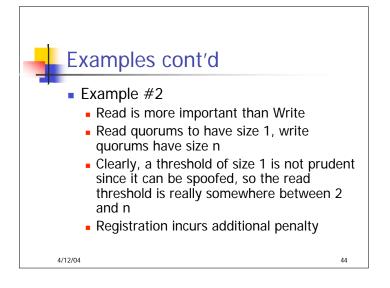


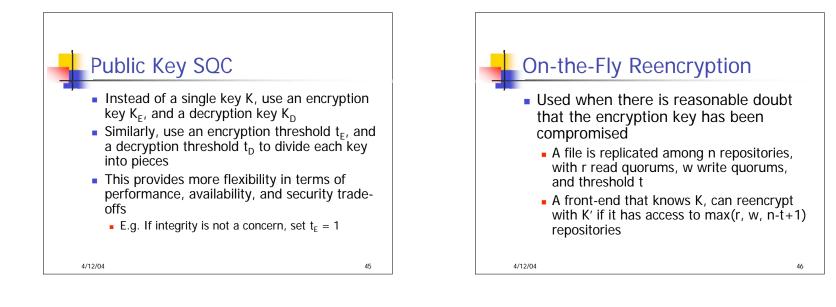
Private Key SQC cont'd

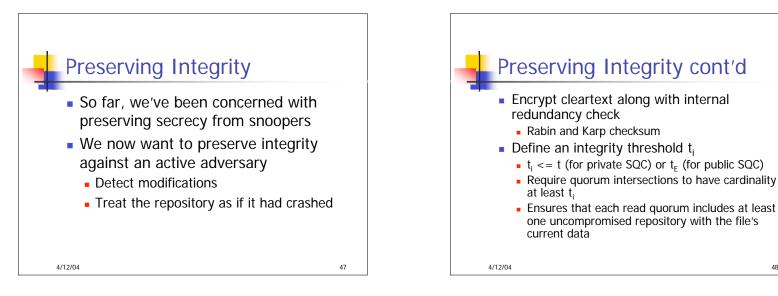
4/12/04

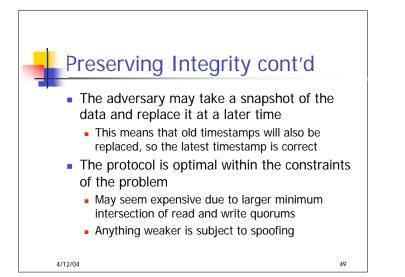
- An adversary can still glean data
 - For example, if the log timestamp entries are not encrypted, they can provide hints
 - The frequency of read/write operations can also provide hints
- If the threshold is set to the smallest quorum, there is no availability penalty

```
    Examples
    Example #I
    Read and write operations are equally important
    Read and write quorums, as well as the share threshold require a majority of [(n+1)/2] repositories
    Registration does not incur an additional penalty, since it can be done at the first quorum
    Up to [(n+1)/2] may fail or be compromised by an adversary
```











File replicated at n repositories

4/12/04

- Read quorums of size r, write quorums of size w, read intersect write at x repositories: r+w-x=n
- R, W, X are disjoint sets of repositories of sizes r-x, w-x, and x
 - The repositories in X are controlled by an adversary

