Data Replication and Placement

Jean-Marie, Baert, Boudet, Roche

Variations on the data replication and placement problem

Alain Jean-Marie[1,2] Anne-Elisabeth Baert[2]
Vincent Boudet[2] Xavier Roche[2]

[1]INRIA
LIRMM CNRS/Univ. Montpellier 2

[2]LIRMM CNRS/Univ. Montpellier 2

Seminaire TREC
Paris, 3 April 2012

Outline

Progress

Motivation: Data placement and replication

Origin of the problem: the VoDDnet (a.k.a. “Peerates”) company.

Distributed Storage and Download System:

- many storage sites, each with limited capacity
- potentially unavailable
- existence of a permanent backup data storage

Questions

Among the design issues for such a system

- how much replication of the data?
- where to place it?
- what download strategy?

Answers in this talk

We fix a particular download strategy. We consider the download time as the metric of interest.

1. Documents with different popularity:
   how much replication of the data?
   $\rightarrow$ combinatorial optimization problem
2. Documents with identical popularity:
   where to place the replications?
   - theoretical properties
   - experimental investigation

Progress

The Document Model

Document model:

- K distinct documents
- p_k, popularity (probability) of document k
- T_k, number of unit-size blocks of data in document k
- r_k, replication factor for document k (same for every block)
- bound N on the replication number

The Download Model

Choice of a two-phase download algorithm

1. choose a download site for each block
2. try to download in parallel from available sites
3. download the rest from the backup storage

Network Model

Network/Storage Assumptions:

- sites available with uniform probability δ, independently
- download from distributed sites ("clients") at data rate θ_c (blocks/s)
- download from central server at data rate θ_s (blocks/s)
- global storage capacity S

Availability

One block is available if at least one of its replications is online:

$$P(\text{block available}) \;=\; 1 \;-\; (1-\delta)^{r_k} \;=\; 1 \;-\; \gamma^{r_k}\;.$$

Let Λ_k be the number of available blocks in the document:

Document Availability

Λ_k is a Binomial Random Variable with parameters T_k and $1-\gamma^{r_k}$. In particular

$$\begin{aligned} \mathbb{E}\Lambda_k &= T_k\,(1-\gamma^{r_k}) \\ \mathbb{E}(T_k-\Lambda_k) &= T_k\,\gamma^{r_k}\;. \end{aligned}$$

Response time

Individual response time

Response time for document k

$$R_k = \frac{\Lambda_k}{\theta_c} + \frac{T_k - \Lambda_k}{\theta_s}$$

Average response time

Taking into account popularity

$$\begin{aligned} \mathbb{E}R &= \sum_{k=1}^{K} p_k \mathbb{E}R_k \\ &= \frac{1}{\theta_c} \sum_k p_k T_k + \left(\frac{1}{\theta_s} - \frac{1}{\theta_c}\right) \sum_k p_k T_k \gamma^{r_k} . \end{aligned}$$

Progress

Optimization problem

Minimizing the expected download time:

Minimum download time

$$\min_{r_k} \sum_k p_k T_k \gamma^{r_k},$$

$$\text{subject to } \sum_{k=1}^{K} T_k r_k \leq S$$

$$\text{and } 0 \leq r_k \leq N \text{ for all } k\,,$$

where $\gamma \in [0,1]$, $0 \leq p_k \leq 1$, $T_k \geq 0$, $\forall k$.

In the relaxation of the problem, $r_k \in [0, N]$. In the real problem, $r_k \in [0..N]$.

Algorithmic Complexity

Limiting cases are difficult problems:

$\gamma \sim 0$ In that case, $r_k \in \{0, 1\}$ and $\gamma^{r_k} \sim 1 - r_k$.

Knapsack problem

$$\min_{r_k} \sum_k p_k T_k (1 - r_k) \qquad \text{s.c.} \sum_{k=1}^{K} T_k r_k \leq S$$
$$\text{and } r_k \in \{0, 1\} \text{ for all } k.$$

$\gamma \sim 1$ In that case, $\gamma^{r_k} \sim 1 - \delta r_k$.

Bounded knapsack problem

$$\min_{r_k} \sum_k p_k T_k (1 - \gamma r_k) \qquad \text{s.c.} \sum_{k=1}^{K} T_k r_k \leq S$$
$$\text{and } r_k \in \{0, \ldots, N\} \text{ for all } k.$$

Lagrangian formulation

Lagrangian:

$$
\begin{aligned}
L(r_k, \lambda_k^-, \lambda_k^+, \lambda) &= \sum_{k=1}^{N} p_k T_k \gamma^{r_k} - \sum_{k=1}^{N} \lambda_k^- r_k \\
&\quad - \sum_{k=1}^{N} \lambda_k^+ (N - r_k) - \lambda \left(S - \sum_{k=1}^{N} T_k r_k \right) .
\end{aligned}
$$

First-order conditions:

$$
\left\{
\begin{array}{l}
\dfrac{\partial L}{\partial r_k} = p_k T_k \log(\gamma) \gamma^{r_k} - \ \lambda_k^- + \lambda_k^+ + \lambda T_k = 0, \ \forall k \ , \\
\lambda_k^- r_k = \lambda_k^+ (N - r_k) = 0, \ \forall k, \\
\lambda \left(S - \sum_{k=1}^{N} T_k r_k \right) = 0, \\
\lambda_k^- \geq 0, \quad \lambda_k^+ \geq 0, \ \forall k \ , \\
\lambda \geq 0 \ .
\end{array}
\right.
$$

Solution of the relaxed problem

Assuming $\lambda > 0$ is known:

$$r_k \;=\; f_k(\lambda) \;:=\; \max\left(\min(\frac{\log(-\lambda/p_k/\log(\gamma))}{\log(\gamma)}, N), 0\right).$$

Plugging this expression in the constraint

$$S = \sum_{k=1}^{N} T_k r_k$$

we get:

$$S = \sum_{k=1}^{N} T_k \max\left(\min(\frac{\log(-\lambda/p_k/\log(\gamma))}{\log(\gamma)}, N), 0\right),$$

to be solved for λ. The solution exists if

$$S \leq N \sum_{k=1}^{N} T_k \; .$$

Practical solution

Define the sets:

$$\begin{aligned}\mathcal{A}(\lambda) &= \{k \text{ such that } \lambda > -p_k \log(\gamma)\}\\ \mathcal{B}(\lambda) &= \{k \text{ such that } -p_k \log(\gamma) \geq \lambda > -p_k {\gamma}^N \log(\gamma)\}\\ \mathcal{C}(\lambda) &= \{k \text{ such that } -p_k {\gamma}^N \log(\gamma) \geq \lambda \geq 0\}\ .\end{aligned}$$

With these notations, we have

$$f_k(\lambda) = \begin{cases} 0 & \text{if } k \in \mathcal{A}(\lambda)\,, \\ \dfrac{1}{\log(\gamma)}\left(\log(\lambda) - \log(-p_k \log(\gamma))\right) & \text{if } k \in \mathcal{B}(\lambda)\,, \\ N & \text{if } k \in \mathcal{C}(\lambda)\,. \end{cases}$$

and

$$f(\lambda) = N \sum_{k \in \mathcal{C}(\lambda)} T_k + \sum_{k \in \mathcal{B}(\lambda)} \frac{T_k}{\log(\gamma)}(\log(\lambda) - \log(-p_k log(\gamma)))$$

$\rightarrow$ solution of $f(\lambda) = S$ by dichotomy.

Approximations for the discrete problem

The simple "floor" algorithm:

- calculate the solution of the relaxed problem $\{r_1^*, \cdots, r_N^*\}$
- for each k, $r_k \leftarrow \lfloor r_k^* \rfloor$

A greedy improvement

- Set $\overrightarrow{r^e} = \{r_1, \cdots, r_N\}$ given by the "floor" algorithm
- $\mathcal{F} = \{\text{relevant documents}\}$: can increase replication without violating constraints
- while $\mathcal{F} \neq \emptyset$:
  - find $j = \arg\max_j (p_k T_k \gamma^{r_k})$
  - $r_j = r_j + 1$
  - Update $\mathcal{F}$

Progress

Minimizing Variance

Documents have now identical popularity $\rightarrow$ same replication factor k. The question is: how to place them on v hosts? Consider just one document. Such a placement is a block design $\mathcal{B}$ of b "blocks", each being a k-subset of $[1..v]$.
Le Λ, $\overline{\Lambda}$ the number of (un)available blocks.

Theorem

$$\mathbb{E}(z^{\overline{\Lambda}}) = \sum_{\mathcal{S}\subset\mathcal{B}} (z-1)^{|\mathcal{S}|} \; \bar{\delta}^{|\cup_{B\in\mathcal{B}} B|} .$$

The first moments of Λ and $\overline{\Lambda}$ are given by:

$$\begin{aligned} \mathbb{E}(\overline{\Lambda}) &= b\,\bar{\delta}^k \\ \mathbb{E}(\Lambda) &= b\,(1-\bar{\delta}^k) \\ \mathbb{V}(\overline{\Lambda}) = \mathbb{V}(\Lambda) &= \sum_{B,B'} \left(\bar{\delta}^{|B\cup B'|} - \bar{\delta}^{2k} \right) . \end{aligned}$$

The MinVar Problem

$$
\begin{aligned}
\mathbb{V}(\Lambda) &= \sum_{B,B'} \left(\bar{\delta}^{|B \cup B'|} - \bar{\delta}^{2k} \right) \\
&= \sum_{B,B'} \bar{\delta}^{|B|+|B'|-|B\cap B'|} - b(b-1)\bar{\delta}^{2k} \\
&= \sum_{B,B'} \bar{\delta}^{2k} \bar{\delta}^{-|B\cap B'|} - b(b-1)\bar{\delta}^{2k}
\end{aligned}
$$

MinVar Problem

Let γ be a real number, $\gamma \geq 1$, and b, v, k be integers. The MinVar(γ, b, v, k) problem: find one design $\mathcal{B}$ with $|\mathcal{B}| = b$, $|\mathcal{V}| = v$ and $|B| = k$ for all $B \in \mathcal{B}$, which minimizes the function:

$$
J(\mathcal{B}, \gamma) := \sum_{B \neq B' \in \mathcal{B}} \gamma^{|B \cap B'|} .
$$

A Greedy Algorithm

This algorithm begins with an empty graph and, at each iteration, finds the edge which minimizes $J(\mathcal{B},\gamma) + J(\mathcal{B}^c,\gamma)$, where $\mathcal{B}^c$ is the complementary design of $\mathcal{B}$.

- $\mathcal{B} = \emptyset$
- for $\ell \in \{1 \cdots k\}$
  - for $i \in \{1 \cdots b\}$
    - $j = arg \min_j \{J(\mathcal{B} \cup (i,j),\gamma) + J((\mathcal{B} \cup (i,j))^c),\gamma)\}$
    - $\mathcal{B} = \mathcal{B} \cup (i,j)$

Random Designs

Random Algorithm

Choose each block uniformly at random over $\binom{[1..v]}{k}$.

Define the function

$$\pi(\gamma) = \binom{v}{k}^{-1} \sum_{j=0}^{k} \binom{k}{j} \binom{v-k}{k-j} \gamma^j .$$

It is the generating function of $X_{BB'} = |B \cap B'|$, where B and B' are two uniformly chosen random blocks.

Theorem

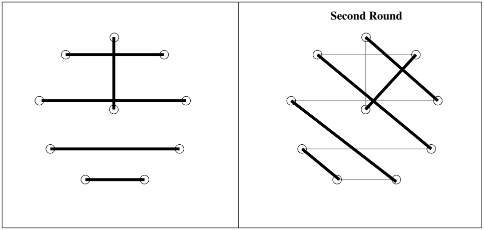
If $\mathcal{B}$ is a design generated by the Random algorithm, then

$$\begin{aligned} \mathbb{E}(J(\mathcal{B},\gamma)) &= b(b-1)\,\pi(\gamma) \\ \mathbb{V}(J(\mathcal{B},\gamma)) &= 2b(b-1)\left(\pi(\gamma^2) - \pi^2(\gamma)\right) . \end{aligned}$$

Solutions

Solutions can be constructed:

- systematically for $k = 2$:



- for Steiner systems
- for many values of b when $k = 3$ (work in progress)