THÈSE DE DOCTORAT

de l'UNIVERSITÉ DENIS DIDEROT – PARIS VII

Spécialité

Informatique Fondamentale

présentée par

STÉPHANE ERANIAN

Sujet de la thèse

Un service de synchronisation distribuée tolérant les pannes: implantation dans CHORUS

soutenue le 29 Septembre 1995 devant le jury composé de:

Paul Feautrier	Président
Guy Bernard	Rapporteurs
Michel Raynal	
François Armand	Examinateurs
Jean-Charles Fabre	
Hugues Fauconnier	
Jean-Marie Rifflet	Directeur de Recherche

Un service de synchronisation distribuée tolérant les pannes : implantation dans CHORUS

Stéphane ERANIAN

Résumé - L'évolution des environnements informatique et notamment le développement des réseaux ouvre de nouvelles possibilités au niveau de la construction des systèmes d'exploitations : les systèmes répartis. Les nouvelles architectures sont basées sur des micro noyaux fournissant des abstractions et services de base à partir desquels on construit des systèmes complets via un ensemble de serveurs, disséminés sur plusieurs sites, communicants par échanges de messages. Cette répartition des tâches impose une certaine coordination afin de maintenir une cohérence d'ensemble. Ce maintien repose sur l'utilisation de synchronisations. Notre environnement cible est constitué d'un ensemble de sites fonctionnant sous le système micro noyau réparti CHORUS.
Le but de cette thèse est de fournir un service de synchronisation distribuée générique basée sur la notion de jeton pour garantir la propriété d'exclusion mutuelle répartie.Deux qualités de service sont fournies. La première est basée sur un algorithme centralisé, la seconde utilise un algorithme distribué, dérivé de celui de Naimi et Trehel, basé sur une structure arborescente dynamique Nous présentons des résultats sur les performances obtenus à partir d'un prototype.
Dans un second temps, nous montrons comment ce service est rendu tolérant aux pannes franches de site sans perte de performances. Les services sont régénérés dynamiquement, les clients sont perdus et les sites peuvent être réintégrés après réparation. Nous n'utilisons aucune technique à base de points de contrôle ou transactions. Notre architecture repose sur un service générique externe, aM ou gestionnaire de pannes réparti. Il fournit des services de détection de pannes, de notifications sur pannes et d'aide au recouvrement. Le serveur de remplacement est choisi dynamiquement suivant un principe d'hospitalité. Le recouvrement exploite la nature répartie de l'environnement, il est basé sur la collecte de l'état auprès des clients survivants en vue de sa reconstruction dans le serveur. La tolérance est supportée dans les deux qualités de services.

Abstract - The evolution of computing environments and especially networkss opens new opportunities in the way of building operating system : distributed systems. New architectures are based on micro kernels which provide basic abstractions and services on which one can build a complete system made of a set of servers, running on different sites, which communicate through a message passing facility. In this context, the distribution of work requires a certain degree of coordination in order to maintain a global consistency. This, in turn, requires the use synchronization primitives. Our target environment is the CHORUS micro kernel system.
The purpose of this PhD thesis is to provide a distributed synchronization service based on the token notion to ensure the distributed mutual exclusion property. Two qualities of service are provided. The first is based on a simple centralized algorithm, th second uses a distributed algorithm, derived from the one by Naimi and Trehel, based on a dynamic tree structure. We present a set of performance results gathered from our prototype.
In the second part, we explain how this service can support fail-stop site failure without loss in performances. Service is regenerated dynamically on another site, clients are lost and site may be restarded after repair. We do not use any checkpoints nor transactions based techniques. Our architecture is based on an external service, the aM or failure manager. It provides services for failure detection, failure notification and recovery help. The backup server is selected dynamically using a hospitality principle. The recovery takes fully benefit of the distributed nature of the environment , it is based on a distributed state collect from the remaining clients. Fault tolerance is supported for both qualities of service.

Table des matières