

nodigt U graag uit op de openbare verdediging van het proefschrift van

## Jim Bauwens

ter behaling van de graad van Doctor in de wetenschappen

Titel van het proefschrift:

**FLEXIBLE CRDTS FOR A DEMANDING WORLD**  
An open and Systematic Implementation Approach  
for CRDT Development

### Curriculum vitae

Promotor:  
Prof. dr. Elisa Gonzalez Boix

De verdediging heeft plaats op  
Vrijdag 23 augustus 2024 om 17u in aula  
I.2.01

### Samenstelling van de jury

Prof. dr. Bas Ketsman (VUB, voorzitter)  
Prof. dr. Paul Van Eecke (VUB, secretaris)  
Prof. dr. Bruno da Silva Gomes (VUB)  
Prof. dr. Carlos Baquero (University of Porto,  
Portugal)  
Prof. dr. Annette Bieniusa (University of  
Kaiserslautern-Landau, Duitsland)

Jim Bauwens behaalde zijn Bachelor-diploma aan de Artesis Plantijn Hogeschool in Antwerpen in 2015, en zijn Master-diploma aan de Vrije Universiteit Brussel (VUB) in 2019. Daarna startte hij een doctoraat aan het Software Languages Lab (SOFT) ondersteund door een FWO-SB beurs.

Zijn onderzoek heeft zich voornamelijk gericht op software engineering technieken voor gerepliceerde datatypes om ontwikkelaars te helpen bij het implementeren van beschikbare systemen. Jim's werk resulteerde in drie publicaties in internationale peer-reviewed tijdschriften en conferenties, en zes bijdragen in internationale peer-reviewed workshops.

### Abstract van het doctoraatsonderzoek

De groei van apparaten verbonden met het internet, waaronder smartphones en IoT-systemen, heeft het landschap van gedistribueerde systemen veranderd. Om beschikbaarheid te garanderen, worden gegevens vaak gerepliceerd tussen servers of systemen aan de rand via een local-first benadering. Dit verbetert prestaties, maar brengt uitdagingen met zich mee omdat verschillende kopieën van gegevens up-to-date moeten worden gehouden. Conflict-free Replicated Data Types (CRDT's) kunnen worden gebruikt om de convergentie van gerepliceerde gegevens te garanderen. Replica's kunnen tijdelijk van elkaar verschillen, maar zodra updates stoppen zullen ze uiteindelijk convergeren.

We hebben tijdens ons onderzoek vier belangrijke uitdagingen geïdentificeerd met betrekking tot het gebruik van CRDTs. Ten eerste is geheugenbeheer gecompliceerd door de accumulatie van meta-data die causaliteit tussen operaties vastlegt. Ten tweede is er beperkte ondersteuning voor dynamische netwerken waar peers kunnen toetreden en vertrekken. Ten derde kan de afhankelijkheid op Reliable Causal Broadcasting (RCB) in veel CRDTs tot onnodige vertragingen leiden. Operaties worden gebufferd tot dat alle causale afhankelijkheden zijn toegepast, zelfs als die afhankelijkheden niet logisch gerelateerd zijn. Ten vierde is de bestaande ondersteuning voor het combineren van CRDTs inflexibel, waardoor ad-hoc implementaties nodig zijn.

Dit proefschrift onderzoekt oplossingen voor deze problemen met behulp van Flec — een framework die we hebben ontwikkeld voor het bouwen van eventueel consistente applicaties — als laboratorium voor het experimenteren met CRDT-ontwerpen en technieken. Flec biedt een modulaire aanpak voor het ontwikkelen van CRDT's door middel van een open implementatie.

Om de onbegrensde accumulatie van meta-data aan te pakken, introduceren we eerst een techniek die communicatiebevestigingen gebruikt om snel causale stabiliteit te bepalen, waardoor meta-data vroegtijdig verwijderd kan worden. Vervolgens wordt deze techniek aangepast om dynamische netwerken te ondersteunen en laten we zien hoe nieuwe peers asynchroon een correcte staat kunnen verkrijgen. Ten tweede stellen we een uitbreiding voor op pure operation-based CRDT's die hun reactievermogen verbetert. Door toe te staan dat pending operaties die zijn opgeslagen in de RCB buffer gedeeltelijk kunnen worden toegepast voordat alle causale afhankelijkheden zijn gearriveerd, wordt de reactiviteit van replica's verbeterd. Tenslotte stellen we geneste, pure operation-based CRDT's voor, een nieuw framework voor het bouwen van geneste, gerepliceerde structuren. Met onze aanpak kunnen CRDT's compositie ondersteunen zonder semantische veranderingen of structurele beperkingen.

We valideren onze bijdragen door een reeks gestructureerde RDT's te implementeren binnen Flec. Elk hoofdstuk bevat een validatiesectie die onze voorgestelde technieken onderbouwt en aantoont dat onze bijdragen het gebruik van CRDTs vereenvoudigt.