Skalierbare Performance-Simulation von fehlertoleranten verteilten Systemen

Christian Berger
Universität Passau
Passau, D
cb@sec.uni-passau.de

Sadok Ben Toumia
Universität Passau
Passau, D
bentou01@ads.uni-passau.de

Hans P. Reiser
Reykjavik University
Reykjavik, IS
hansr@ru.is

1 MOTIVATION UND ZIELE

Die Bewertung der Leistungsfähigkeit von Byzantinisch fehlertoleranten (BFT) Protokollen erfordert eine sorgfältige Evaluierung, was in einem größeren Maßstab eine signifikante Herausforderung darstellt. Experimente mit einer echten Instantiierung eines vollständigen Produktivsystems bieten in der Regel realistische Ergebnisse, sind sie aber kostenaufwändig und zeitintensiv und darüber hinaus nichtdeterministisch und dadurch schlecht reproduzierbar. Bei rein modellbasierten Simulationen dagegen führen Vereinfachungen, die im Modell gegenüber der Wirklichkeit gemacht werden, zu ungenauen oder gar realitätsfernen Ergebnissen.

In diesem Beitrag präsentieren wir unsere Entwicklung von *Delphi-BFT* [2][1], einem Werkzeug zur Bewertung der Performance von BFT-Protokollen für großskalierte Systeme auf der Basis der Simulationssoftware Phantom [3]. Mit diesem Werkzeug zielen wir auf die folgenden Fragestellungen:

- Wie gut lässt sich das reale Verhalten eines BFT-Systems durch geeignete Simulation präzise nachbilden?
- Welche Skalierbarkeit in Hinblick auf eine große Anzahl an teilnehmenden Replikaten kann dabei erzielt werden?

Darüber hinaus gehen wir auf offene Herausforderungen ein, zu denen wir aktuell neue Lösungsansätze entwickeln.

2 ARCHITEKTUR VON DELPHI-BFT

Delphi-BFT (s. Abb. 1) folgt einer modularen Architektur, bestehend aus einem Orchestrator, einem Umgebungsgenerator, Protokollkonnektoren, einem Ressourcenmonitor und einem Plotter. Der Orchestrator erwartet eine simple Spezifikation des Experiments als Input und verwaltet alle Werkzeuge: Beispielsweise bereitet er die Netzwerkumgebung vor, konfiguriert Laufzeit-Artefakte für ein BFT-Protokoll und initialisiert den Ressourcenmonitor. Der Umgebungsgenerator erstellt Netzwerktopologien, die realistische Szenarien für eine LAN- oder WAN-Umgebung widerspiegeln.

Protokollkonnektoren sind für die Erstellung von Protokollkonfigurationsdateien für jede BFT-Protokollimplementierung zuständig. Durch unsere modulare Architektur kann ein neues Protokoll zu unserem Werkzeug hinzugefügt werden, indem ein neuer Konnektor implementiert wird (nur ca. 100–200 Codezeilen). Der Ressourcenmonitor sammelt Informationen über Ressourcenverbrauch während der Simulationsläufe, einschließlich der zugewiesenen Speicher- und CPU-Zeit sowie die gesamte Simulationszeit.

Die Ergebnisse der Simulationen werden von Phantom gespeichert und können aggregiert und auf spezifische Diagramme für bestimmte Metriken wie Latenz oder Durchsatz abgebildet werden.

[1]Gefördert von der Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 446811880 (BFT2Chain)

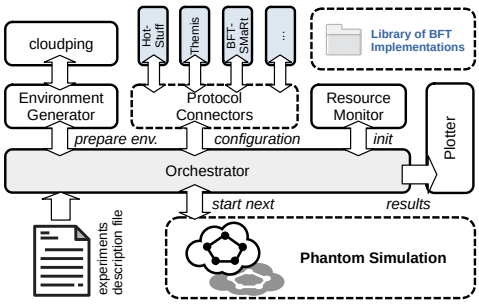

Abbildung 1: Delphi-BFT-Toolchain auf Basis von Phantom

3 ERGEBNISSE

Wir können mit Delphi-BFT die Leistung eines BFT-Protokolls genau vorhersagen, während wir die Umgebung (Anzahl Knoten, geografische Verteilung) experimentell skalieren. Unser Ansatz ist ressourcenschonend und bewahrt die Realitätsnähe der Anwendungen, da die Implementierung bestehender BFT-Frameworks als Linux-Prozesse in die Simulationsumgebung eingestöpselt wird, ohne erforderliche Codeänderungen oder Neuimplementierung.

Ein interessantes Experiment, das mit Delphi-BFT durchgeführt wurde, vergleicht die Ergebnisse von Simulationen von PBFT, HotStuff und BFT-SMaRt mit zunehmender Systemgröße. Die Ergebnisse zeigen, dass PBFT zunächst besser abschneidet als HotStuff, aber bei einer höheren Anzahl an Replikaten schlechter skaliert. Ein Vergleich unserer Simulationsergebnissen mit Messungen echter, cloud-basierter Protokollausführungen suggeriert eine höhere Genauigkeit mit wachsender Systemgröße, da die Leistung von BFT Protokollen dann zunehmend durch das Netzwerk bestimmt wird.

4 AUSBLICK

Aktuell arbeiten wir an der Einführung eines Angreifer-Modells, um Auswirkungen von Angriffen auf die Systemleistung zu untersuchen. Dafür orientieren wir uns an der „Twins“-Methodik, einem neuen Ansatz zur Validierung von BFT-Protokollen, bei dem Unit-Tests mittels Simulation byzantinischer Angriffe durch Duplizieren von kryptografischen Replikatidentitäten erzeugt werden [1].

LITERATUR

[1] Bano, S., Sonnino, A., Chursin, A., Perelman, D., Li, Z., Ching, A., and Malkhi, D. Twins: BFT Systems Made Robust. In *25th Int. Conf. on Principles of Distributed Systems* (Dagstuhl, Germany, 2022), vol. 217, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, pp. 7:1–7:29.

[2] Berger, C., Toumia, S. B., and Reiser, H. P. Does my BFT protocol implementation scale? In *Proc.of the 3rd Int. Workshop on Distributed Infrastructure for the Common Good* (New York, NY, USA, 2022), DICG '22, ACM, pp. 19–24.

[3] Jansen, R., Newsome, J., and Wails, R. Co-opting linux processes for high-performance network simulation. In *USENIX ATC* (2022), pp. 327–350.