

Content-based Publish/Subscribe with P4

Dissertation

zur

Erlangung des akademischen Grades

Doktor-Ingenieur (Dr.-Ing.)

der Fakultät für Informatik und Elektrotechnik

der Universität Rostock

vorgelegt von

M.Sc. Christian Wernecke, geb. am 31. Juli 1982 in Berlin

Rostock, 05. Januar 2025

Abstract

Publish/subscribe is a flexible communication pattern for loosely coupled distributed applications. In the content-based variant, each published notification is matched against active subscriptions to determine a set of interested subscribers to which the notification is to be delivered. However, since the set of receivers can be any combination of subscribers, distributing notifications from publishers to subscribers with matching subscriptions is complex. Traditional approaches rely on *broker-based networks* to deliver each message to a unique set of recipients. However, brokers make forwarding decisions at the application-layer and typically perform application-layer multicast by sending multiple unicast packets, which increases network load and delays delivery.

With the advent of *Software-Defined Networking (SDN)*, new distribution strategies are possible that combine the efficiency of *network-assisted multicast* with the flexibility of *application-layer multicast*. In this thesis, we explore how SDN and the P4 programming language can be used to implement fast and flexible notification distribution. We present fundamental strategies that embed the entire delivery tree of a notification directly into the packet header, which conserves bandwidth and reduces latency compared to traditional methods.

Furthermore, we present advanced P4-based forwarding schemes that combine static routing rules in network switches with dynamically encoded additional forwarding information in the packet header. This hybrid approach allows precise control over the trade-off between compact headers and flexibility in notification distribution. Evaluations in an emulated data center network demonstrate how these strategies can significantly reduce header size and network traffic.

To further improve delivery efficiency, we introduce innovative distribution strategies that use *virtual trees*. These virtual trees map frequently and simultaneously used paths in the switch infrastructure as partial delivery trees. A specially designed encoding algorithm dynamically constructs notification-specific delivery trees by merging multiple virtual trees and encoding additional routing information in the notification header to extend, connect, or prune virtual trees. Evaluation results confirm that the integration of combined virtual trees with routing information in the header significantly reduces header size and network traffic.

Finally, this work introduces installation strategies that dynamically deploy virtual trees in the switch infrastructure based on various levels of application-dependent knowledge and runtime statistics. These strategies exploit topological information, publisher-subscriber relationships, and notification frequencies to construct efficient distribution trees that require minimal additional routing information in notification headers. Evaluation results highlight the efficiency of these strategies under different churn rate scenarios in real-world network topologies.

By leveraging SDN and P4, this work enables optimized, scalable, and high-performance notification distribution in content-based publish/subscribe systems, significantly reducing delivery latency and bandwidth consumption.

Zusammenfassung

Publish/Subscribe ist ein flexibles Kommunikationsmuster für lose gekoppelte verteilte Anwendungen. In der inhaltsbasierten Variante wird jede veröffentlichte Benachrichtigung mit aktiven Abonnements abgeglichen, um die interessierten Teilnehmer zu ermitteln, an die die Benachrichtigung zugestellt werden soll. Da die Empfänger jedoch jede beliebige Kombination von Abonnenten sein können, ist die Verteilung von Benachrichtigungen von Publishern an Subscriber mit passenden Abonnements komplex. Traditionelle Ansätze basieren auf brokerbasierten Netzwerken, die für jede Nachricht eine individuelle Empfängergruppe adressieren. Broker treffen Weiterleitungsentscheidungen auf Anwendungsebene und führen typischerweise anwendungsseitiges Multicast durch, indem sie mehrere Unicast-Pakete senden. Dies erhöht die Netzwerklast und verzögert die Zustellung.

Mit dem Aufkommen von *Software-Defined Networking (SDN)* eröffnen sich neue Möglichkeiten für Verteilungsstrategien, die die Effizienz von netzwerkbasiertem Multicast mit der Flexibilität des anwendungsseitigen Multicasts kombinieren. In dieser Arbeit untersuchen wir, wie SDN und die *Programmiersprache P4* eingesetzt werden können, um eine schnelle und flexible Benachrichtigungsverteilung zu implementieren. Wir präsentieren grundlegende Strategien, die den gesamten Verteilungsbaum einer Benachrichtigung direkt in den Paketheader einbetten, was im Vergleich zu herkömmlichen Methoden die Bandbreite schont und die Latenz reduziert.

Darüber hinaus entwickeln wir in dieser Arbeit fortschrittliche P4-basierte Weiterleitungsschemata, die statische Routing-Regeln in Netzwerk-Switches mit dynamisch kodierten zusätzlichen Weiterleitungsinformationen im Paket-Header kombinieren. Dieser hybride Ansatz erlaubt eine präzise Kontrolle des Kompromisses zwischen kompakten Headern und Flexibilität bei der Benachrichtigungsverteilung. Anhand von Evaluationen in einem emulierten Rechenzentrumsnetzwerk wird gezeigt, wie diese Strategien die Headergröße und den Netzwerkverkehr signifikant reduzieren können.

Um die Effizienz der Verteilung weiter zu steigern, führen wir innovative Verteilungsschemata ein, die *virtuelle Bäume* nutzen. Diese virtuellen Bäume bilden häufig und gleichzeitig genutzte Pfade in der Switch-Infrastruktur als partielle

Verteilungsbäume ab. Ein speziell entwickelter Kodieralgorithmus konstruiert dynamisch benachrichtigungsspezifische Zustellungsbäume, indem er mehrere virtuelle Bäume zusammenführt und zusätzliche Routing-Informationen im Benachrichtigungsheader kodiert, um virtuelle Bäume zu erweitern, zu verbinden oder zu beschneiden. Die Ergebnisse der Evaluierung bestätigen, dass durch die Integration von virtuellen Bäume mit Routing-Informationen im Header die Headergröße und der Netzwerkverkehr deutlich reduziert werden können.

Abschließend untersucht diese Arbeit Installationsstrategien, die virtuelle Bäume in der Switch-Infrastruktur dynamisch bereitstellen, basierend auf verschiedenen Niveaus von anwendungsabhängigem Wissen und Laufzeitstatistiken. Diese Strategien nutzen topologische Informationen, spezifische Beziehungen zwischen Publishern und Subscribern sowie Benachrichtigungsfrequenzen, um effiziente Zustellungsbäume zu konstruieren, die nur minimale zusätzliche Routing-Informationen in den Benachrichtigungs-Headern benötigen. Die Evaluationsergebnisse unterstreichen die Effizienz dieser Strategien bei sich ändernden Abonnements in realen Netzwerk-Topologien.

Durch den Einsatz von SDN und P4 ermöglicht diese Arbeit eine optimierte, skalierbare und leistungsfähige Benachrichtigungsverteilung in inhaltsbasierten Publish/Subscribe-Systemen, die die Zustelllatenz und Bandbreitennutzung deutlich reduziert.