

Interactive / complex / 1

IC 1

IC 2

IC 3

IC 4

IC 5

IC 6

IC 7

IC 8

IC 9

IC 10

IC 11

IC 12

IC 13

IC 14v1

IC 14v2

query	Interactive / complex / 1																																																																					
title	Transitive friends with a certain name																																																																					
pattern	<pre>graph LR person[person: Person id = \$personId] -- knows*1..3 --> otherPerson[otherPerson: Person firstName = \$firstName id lastName birthday creationDate gender browserUsed locationIP email speaks] otherPerson -- isLocatedIn --> locationCity[locationCity: City name] otherPerson -- «opt» workAt --> company[company: Company name] otherPerson -- «opt» studyAt --> university[university: University name] locationCity -- isLocatedIn --> companyCountry[companyCountry: Country name] university -- isLocatedIn --> universityCity[universityCity: City name]</pre>																																																																					
description	Given a start Person with ID \$personId, find Persons with a given first name (\$firstName) that the start Person is connected to (excluding start Person) by at most 3 steps via the knows relationships. Return Persons, including the distance (1..3), summaries of the Persons workplaces and places of study.																																																																					
params	<table><tr><td>1</td><td>\$personId</td><td>ID</td><td></td></tr><tr><td>2</td><td>\$firstName</td><td>String</td><td></td></tr></table>					1	\$personId	ID		2	\$firstName	String																																																										
1	\$personId	ID																																																																				
2	\$firstName	String																																																																				
result	<table><tr><td>1</td><td>otherPerson.id</td><td>ID</td><td>R</td><td></td></tr><tr><td>2</td><td>otherPerson.lastName</td><td>String</td><td>R</td><td></td></tr><tr><td>3</td><td>distanceFromPerson</td><td>32-bit Integer</td><td>C</td><td></td></tr><tr><td>4</td><td>otherPerson.birthday</td><td>Date</td><td>R</td><td></td></tr><tr><td>5</td><td>otherPerson.creationDate</td><td>DateTime</td><td>R</td><td></td></tr><tr><td>6</td><td>otherPerson.gender</td><td>String</td><td>R</td><td></td></tr><tr><td>7</td><td>otherPerson.browserUsed</td><td>String</td><td>R</td><td></td></tr><tr><td>8</td><td>otherPerson.locationIP</td><td>String</td><td>R</td><td></td></tr><tr><td>9</td><td>otherPerson.email</td><td>{Long String}</td><td>R</td><td></td></tr><tr><td>10</td><td>otherPerson.speaks</td><td>{String}</td><td>R</td><td></td></tr><tr><td>11</td><td>locationCity.name</td><td>String</td><td>R</td><td></td></tr><tr><td>12</td><td>universities</td><td>{<String, 32-bit Integer, String>}</td><td>A</td><td>{<university.name, studyAt.classYear, universityCity.name>}</td></tr><tr><td>13</td><td>companies</td><td>{<String, 32-bit Integer, String>}</td><td>A</td><td>{<company.name, workAt.workFrom, companyCountry.name>}</td></tr></table>					1	otherPerson.id	ID	R		2	otherPerson.lastName	String	R		3	distanceFromPerson	32-bit Integer	C		4	otherPerson.birthday	Date	R		5	otherPerson.creationDate	DateTime	R		6	otherPerson.gender	String	R		7	otherPerson.browserUsed	String	R		8	otherPerson.locationIP	String	R		9	otherPerson.email	{Long String}	R		10	otherPerson.speaks	{String}	R		11	locationCity.name	String	R		12	universities	{<String, 32-bit Integer, String>}	A	{<university.name, studyAt.classYear, universityCity.name>}	13	companies	{<String, 32-bit Integer, String>}	A	{<company.name, workAt.workFrom, companyCountry.name>}
1	otherPerson.id	ID	R																																																																			
2	otherPerson.lastName	String	R																																																																			
3	distanceFromPerson	32-bit Integer	C																																																																			
4	otherPerson.birthday	Date	R																																																																			
5	otherPerson.creationDate	DateTime	R																																																																			
6	otherPerson.gender	String	R																																																																			
7	otherPerson.browserUsed	String	R																																																																			
8	otherPerson.locationIP	String	R																																																																			
9	otherPerson.email	{Long String}	R																																																																			
10	otherPerson.speaks	{String}	R																																																																			
11	locationCity.name	String	R																																																																			
12	universities	{<String, 32-bit Integer, String>}	A	{<university.name, studyAt.classYear, universityCity.name>}																																																																		
13	companies	{<String, 32-bit Integer, String>}	A	{<company.name, workAt.workFrom, companyCountry.name>}																																																																		
sort	<table><tr><td>1</td><td>distanceFromPerson</td><td>↑</td><td></td></tr><tr><td>2</td><td>otherPerson.lastName</td><td>↑</td><td></td></tr><tr><td>3</td><td>otherPerson.id</td><td>↑</td><td></td></tr></table>					1	distanceFromPerson	↑		2	otherPerson.lastName	↑		3	otherPerson.id	↑																																																						
1	distanceFromPerson	↑																																																																				
2	otherPerson.lastName	↑																																																																				
3	otherPerson.id	↑																																																																				
limit	20																																																																					
CPs	2.1, 5.3, 8.2																																																																					
relevance	This query is a representative of a simple navigational query. It is interesting for several aspects. (1) It requires for a complex aggregation for returning the concatenation of universities, companies, languages and email information of the Person. (2) It tests the ability of the optimizer to move the evaluation of sub-queries functionally dependant on the Person, after the evaluation of the top-k. (3) Its performance is highly sensitive to properly estimating the cardinalities in each transitive path, and paying attention not to explore already visited Persons.																																																																					

Interactive / complex / 2

IC 1	query	Interactive / complex / 2				
IC 2	title	Recent messages by your friends				
IC 3	pattern					
IC 4						
IC 5						
IC 6						
IC 7						
IC 8						
IC 9						
IC 10	description	Given a start Person with ID \$personId, find the most recent Messages from all of that Person’s friends (friend nodes). Only consider Messages created before the given \$maxDate (excluding that day).				
IC 11	params	1	\$personId	ID		
IC 12		2	\$maxDate	Date		
IC 13	result	1	friend.id	ID	R	
IC 14v1		2	friend.firstName	String	R	
IC 14v2		3	friend.lastName	String	R	
		4	message.id	ID	R	
		5	message.content or message.imageFile (for photos)	Text	R	
		6	message.creationDate	DateTime	R	
	sort	1	message.creationDate	↓		
		2	message.id	↑		
	limit	20				
	CPs	1.1, 2.2, 2.3, 3.2, 8.5				
	relevance	This is a navigational query looking for paths of length two, starting from a given Person, going to their friends and from them, moving to their published Posts and Comments. This query exercises both the optimizer and how data is stored. It tests the ability to create execution plans taking advantage of the orderings induced by some operators to avoid performing expensive sorts. This query requires selecting Posts and Comments based on their creation date, which might be correlated with their identifier and therefore, having intermediate results with interesting orders. Also, messages could be stored in an order correlated with their creation date to improve data access locality. Finally, as many of the attributes required in the projection are not needed for the execution of the query, it is expected that the query optimizer will move the projection to the end.				

Interactive / complex / 3

IC 1
IC 2
IC 3
IC 4
IC 5
IC 6
IC 7
IC 8
IC 9
IC 10
IC 11
IC 12
IC 13
IC 14v1
IC 14v2

query	Interactive / complex / 3				
title	Friends and friends of friends that have been to given countries				
pattern	<p>The diagram illustrates the query pattern. It starts with a person: Person entity (orange box) with attribute <code>id = \$personId</code>. This person knows*1..2 otherPerson: Person entities (orange box). The otherPerson entity has attributes <code>id</code>, <code>firstName</code>, and <code>lastName</code>. The otherPerson entity hasCreator Message entities (purple box). There are two Message entities shown, each with a filter: <code>\$startDate ≤ creationDate < \$startDate + \$durationDays</code>. The Message entities isLocatedIn Country entities (yellow box). The Country entities are named <code>countryX: Country</code> and <code>countryY: Country</code>, with attributes <code>name = \$countryXName</code> and <code>name = \$countryYName</code> respectively. The Country entities isPartOf City entities (green box). The City entities have a filter: <code>«neg» isPartOf</code>.</p>				
description	Given a start Person with ID <code>\$personId</code> , find Persons that are their friends and friends of friends (excluding the start Person) that have made Posts / Comments in both of the given Countries (named <code>\$countryXName</code> and <code>\$countryYName</code>), within <code>[\$startDate, \$startDate + \$durationDays)</code> (closed-open interval). Only Persons that are foreign to these Countries are considered, that is Persons whose location Country is neither named <code>\$countryXName</code> nor <code>\$countryYName</code> .				
params	1	<code>\$personId</code>	ID		
	2	<code>\$countryXName</code>	String	In SNB Interactive v2, this query has two variants: (a) Correlated Countries (b) Anti-correlated Countries	
	3	<code>\$countryYName</code>	String		
	4	<code>\$startDate</code>	Date	Beginning of requested period	
	5	<code>\$durationDays</code>	32-bit Integer	Duration of requested period, in days. The interval <code>[\$startDate, \$startDate + \$durationDays)</code> is closed-open	
result	1	<code>otherPerson.id</code>	ID	R	
	2	<code>otherPerson.firstName</code>	String	R	
	3	<code>otherPerson.lastName</code>	String	R	
	4	<code>xCount</code>	32-bit Integer	A	Number of Messages from Country named <code>\$countryXName</code> created by the Person within the given time
	5	<code>yCount</code>	32-bit Integer	A	Number of Messages from Country named <code>\$countryYName</code> created by the Person within the given time
	6	<code>count</code>	32-bit Integer	A	<code>count = xCount + yCount</code>
sort	1	<code>count</code>	↓		
	2	<code>otherPerson.id</code>	↑		
limit	20				
CPs	2.1, 3.1, 5.1, 8.2, 8.5				
relevance	This query looks for paths of length two and three, starting from a Person, going to friends or friends of friends, and then moving to Messages. This query tests the ability of the query optimizer to select the most efficient join ordering, which will depend on the cardinalities of the intermediate results. Many friends of friends can be duplicate, then it is expected to eliminate duplicates and those people prior to access the Post and Comments, as well as eliminate those friends from Countries named <code>\$countryXName</code> and <code>\$countryYName</code> , as the size of the intermediate results can be severely affected. A possible structural optimization could be to materialize the number of Posts and Comments created by a Person, and progressively filter those people that could not even fall in the top 20 even having all their posts in the Countries named <code>\$countryXName</code> and <code>\$countryYName</code> .				

Interactive / complex / 4

IC 1

IC 2

IC 3

IC 4

IC 5

IC 6

IC 7

IC 8

IC 9

IC 10

IC 11

IC 12

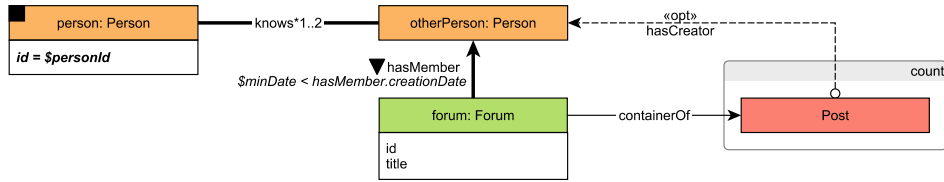
IC 13

IC 14v1

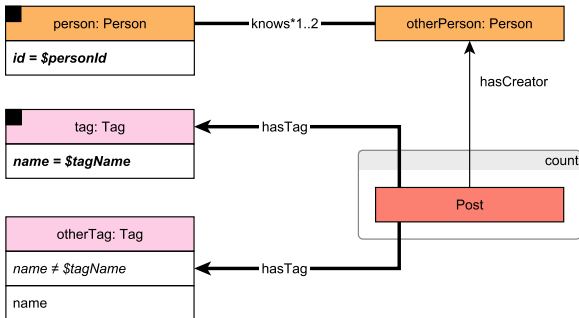
IC 14v2

query	Interactive / complex / 4			
title	New topics			
pattern	<pre> graph LR Person1[Person] -- knows --> person[person: Person id = \$personId] person -- knows --> friend[friend: Person] friend -.-> «opt» hasCreator Post2[Post \$startDate ≤ creationDate < \$startDate + \$durationDays] Post2 -- postCount = count --> Post1[Post creationDate < \$startDate] Post1 -- «neg» hasTag --> Tag[tag: Tag] friend -- hasTag --> Tag </pre>			
description	<p>Given a start Person with ID \$personId, find Tags that are attached to Posts that were created by that Person's friends. Only include Tags that were attached to friends' Posts created within a given time interval [\$startDate, \$startDate + \$durationDays) (closed-open) and that were never attached to friends' Posts created before this interval.</p>			
params	1	\$personId	ID	
	2	\$startDate	Date	
	3	\$durationDays	32-bit Integer	Duration of requested period, in days. The interval [\$startDate, \$startDate + \$durationDays) is closed-open
result	1	tag.name	Long String	R
	2	postCount	32-bit Integer	A
	Number of Posts made within the given time interval that have tag			
sort	1	postCount	↓	
	2	tag.name	↑	
limit	10			
CPs	2.3, 8.2, 8.5			
relevance	<p>This query looks for paths of length two, starting from a given Person, moving to Posts and then to Tags. It tests the ability of the query optimizer to properly select the usage of hash joins or index based joins, depending on the cardinality of the intermediate results. These cardinalities are clearly affected by the input Person, the number of friends, the variety of Tags, the time interval and the number of Posts.</p>			

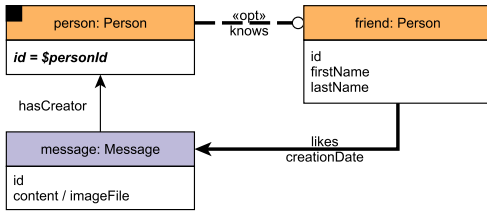
Interactive / complex / 5

IC 1	query	Interactive / complex / 5			
IC 2	title	New groups			
IC 3	pattern	 <pre> graph LR P1[person: Person id = \$personId] -- "knows*1..2" --> P2[otherPerson: Person] P2 -.-> «opt» hasCreator Post[Post] P2 -- "hasMember \$minDate < hasMember.creationDate" --> F[forum: Forum id title] F -- "containerOf" --> Post Post -- "count" --> C[count] </pre>			
IC 4	description	<p>Given a start Person with ID \$personId, denote their friends and friends of friends (excluding the start Person) as otherPerson.</p> <p>Find Forums that any Person otherPerson became a member of after a given date (\$minDate). For each of those Forums, count the number of Posts that were created by the Person otherPerson.</p>			
IC 5	params	1	\$personId	ID	
IC 6		2	\$minDate	Date	
IC 7	result	1	forum.title	Long String	R
IC 8		2	postCount	32-bit Integer	A
IC 9		Number of Posts made in forum that were created by the Person otherPerson			
IC 10	sort	1	postCount	↓	
IC 11		2	forum.id	↑	
IC 12	limit	20			
IC 13	CPs	2.3, 3.3, 8.2, 8.5			
IC 14v1	relevance	<p>This query looks for paths of length two and three, starting from a given Person, moving to friends and friends of friends, and then getting the Forums they are members of. Besides testing the ability of the query optimizer to select the proper join operator, it rewards the usage of indices, but their accesses will be presumably scattered due to the two/three-hop search space of the query, leading to unpredictable and scattered index accesses. Having efficient implementations of such indices will be highly beneficial.</p>			
IC 14v2					

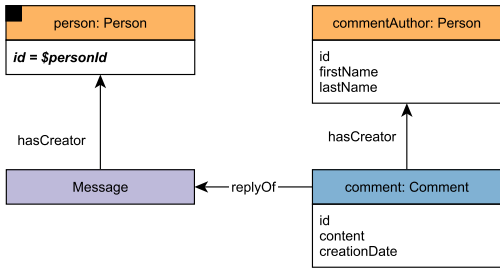
Interactive / complex / 6

query	Interactive / complex / 6													
title	Tag co-occurrence													
pattern														
description	Given a start Person with ID <code>\$personId</code> and a Tag with name <code>\$tagName</code> , find the other Tags that occur together with this Tag on Posts that were created by start Person’s friends and friends of friends (excluding start Person). Return top 10 Tags, and the count of Posts that were created by these Persons, which contain both this Tag and the given Tag.													
params	<table><tr><td>1</td><td><code>\$personId</code></td><td>ID</td><td></td></tr><tr><td>2</td><td><code>\$tagName</code></td><td>Long String</td><td></td></tr></table>	1	<code>\$personId</code>	ID		2	<code>\$tagName</code>	Long String						
1	<code>\$personId</code>	ID												
2	<code>\$tagName</code>	Long String												
result	<table><tr><td>1</td><td><code>otherTag.name</code></td><td>Long String</td><td>R</td><td></td></tr><tr><td>2</td><td><code>postCount</code></td><td>32-bit Integer</td><td>A</td><td>Number of Posts that were created by friends and friends of friends, which have the Tag <code>otherTag</code></td></tr></table>	1	<code>otherTag.name</code>	Long String	R		2	<code>postCount</code>	32-bit Integer	A	Number of Posts that were created by friends and friends of friends, which have the Tag <code>otherTag</code>			
1	<code>otherTag.name</code>	Long String	R											
2	<code>postCount</code>	32-bit Integer	A	Number of Posts that were created by friends and friends of friends, which have the Tag <code>otherTag</code>										
sort	<table><tr><td>1</td><td><code>postCount</code></td><td>↓</td><td></td></tr><tr><td>2</td><td><code>otherTag.name</code></td><td>↑</td><td></td></tr></table>	1	<code>postCount</code>	↓		2	<code>otherTag.name</code>	↑						
1	<code>postCount</code>	↓												
2	<code>otherTag.name</code>	↑												
limit	10													
CPs	5.1, 8.2													
relevance	This query looks for paths of lengths three or four, starting from a given Person, moving to friends or friends of friends, then to Posts and finally ending at a given Tag.													

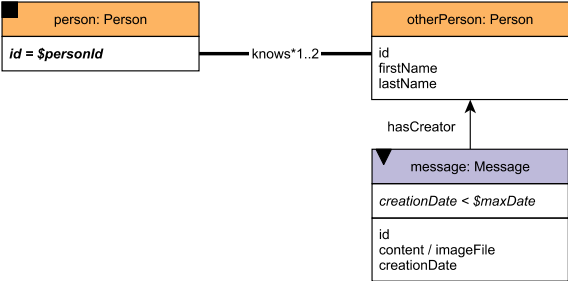
Interactive / complex / 7

IC 1	query	Interactive / complex / 7																																											
IC 2	title	Recent likers																																											
IC 3	pattern																																												
IC 4																																													
IC 5																																													
IC 6																																													
IC 7																																													
IC 8	description	<p>Given a start Person with ID <code>\$personId</code>, find the most recent <code>likes</code> on any of start Person’s Mes-sages. Find Persons that liked (<code>likes</code> edge) any of start Person’s Messages, the Messages they liked most recently, the creation date of that like, and the latency in minutes (<code>minutesLatency</code>) between creation of Messages and like. Additionally, for each Person found return a flag indicating (<code>isNew</code>) whether the liker is a friend of start Person. In case that a Person liked multiple Messages at the same time, return the Message with lowest identifier.</p> <p><i>Validation rule:</i> Depending on whether the system-under-test supports leap seconds or uses UTC-SLS (UTC with Smoothed Leap Seconds), a difference of 1 minute can occur between the <code>minutesLatency</code> results of two correct implementations when the time interval includes June 30, 2012, when there was a leap second. Therefore, the <code>minutesLatency</code> value is validated using a tolerance of 1 minute.</p>																																											
IC 10																																													
IC 11																																													
IC 12																																													
IC 13																																													
IC 14v1	params	<table><tr><td>1</td><td><code>\$personId</code></td><td>ID</td><td></td></tr></table>				1	<code>\$personId</code>	ID																																					
1		<code>\$personId</code>	ID																																										
IC 14v2																																													
result	<table><tr><td>1</td><td><code>friend.id</code></td><td>ID</td><td>R</td><td><code>friend.id = personId</code> is allowed</td></tr><tr><td>2</td><td><code>friend.firstName</code></td><td>String</td><td>R</td><td></td></tr><tr><td>3</td><td><code>friend.lastName</code></td><td>String</td><td>R</td><td></td></tr><tr><td>4</td><td><code>likes.creationDate</code></td><td>DateTime</td><td>R</td><td></td></tr><tr><td>5</td><td><code>message.id</code></td><td>ID</td><td>R</td><td></td></tr><tr><td>6</td><td><code>message.content</code> or <code>message.imageFile</code> (for photos)</td><td>Text</td><td>R</td><td></td></tr><tr><td>7</td><td><code>minutesLatency</code></td><td>32-bit Integer</td><td>C</td><td>Duration between the creation of the Message and the creation of the like, in minutes.</td></tr><tr><td>8</td><td><code>isNew</code></td><td>Boolean</td><td>C</td><td>False if person and friend know each other, True otherwise</td></tr></table>					1	<code>friend.id</code>	ID	R	<code>friend.id = personId</code> is allowed	2	<code>friend.firstName</code>	String	R		3	<code>friend.lastName</code>	String	R		4	<code>likes.creationDate</code>	DateTime	R		5	<code>message.id</code>	ID	R		6	<code>message.content</code> or <code>message.imageFile</code> (for photos)	Text	R		7	<code>minutesLatency</code>	32-bit Integer	C	Duration between the creation of the Message and the creation of the like, in minutes.	8	<code>isNew</code>	Boolean	C	False if person and friend know each other, True otherwise
	1	<code>friend.id</code>	ID	R	<code>friend.id = personId</code> is allowed																																								
	2	<code>friend.firstName</code>	String	R																																									
	3	<code>friend.lastName</code>	String	R																																									
	4	<code>likes.creationDate</code>	DateTime	R																																									
	5	<code>message.id</code>	ID	R																																									
	6	<code>message.content</code> or <code>message.imageFile</code> (for photos)	Text	R																																									
	7	<code>minutesLatency</code>	32-bit Integer	C	Duration between the creation of the Message and the creation of the like, in minutes.																																								
8	<code>isNew</code>	Boolean	C	False if person and friend know each other, True otherwise																																									
sort	<table><tr><td>1</td><td><code>likes.creationDate</code></td><td>↓</td><td></td></tr><tr><td>2</td><td><code>friend.id</code></td><td>↑</td><td></td></tr></table>					1	<code>likes.creationDate</code>	↓		2	<code>friend.id</code>	↑																																	
	1	<code>likes.creationDate</code>	↓																																										
2	<code>friend.id</code>	↑																																											
limit	20																																												
CPs	2.2, 2.3, 3.3, 5.1, 8.1, 8.3																																												
relevance	<p>This query looks for paths of length two, starting from a given Person, moving to its published messages and then to Persons who liked them. It tests several aspects related to join optimization, both at query optimization plan level and execution engine level. On the one hand, many of the columns needed for the projection are only needed in the last stages of the query, so the optimizer is expected to delay the projection until the end. This query implies accessing two-hop data, and as a consequence, index accesses are expected to be scattered. We expect to observe variate cardinalities, depending on the characteristics of the input parameter, so properly selecting the join operators will be crucial. This query has a lot of correlated sub-queries, so it is testing the ability to flatten the query execution plans.</p>																																												

Interactive / complex / 8

IC 1	query	Interactive / complex / 8			
IC 2	title	Recent replies			
IC 3	pattern	 <pre> graph TD P1[person: Person id = \$personId] M[Message] C[comment: Comment id content creationDate] PA[commentAuthor: Person id firstName lastName] M -- hasCreator --> P1 C -- hasCreator --> PA M -- replyOf --> C </pre>			
IC 4	description	Given a start Person with ID \$personId, find the most recent Comments that are replies to Messages of the start Person. Only consider direct (single-hop) replies, not the transitive (multi-hop) ones. Return the reply Comments, and the Person that created each reply Comment.			
IC 5	params	1	\$personId	ID	
IC 6	result	1	commentAuthor.id	ID	R
IC 7		2	commentAuthor.firstName	String	R
IC 8		3	commentAuthor.lastName	String	R
IC 9		4	comment.creationDate	DateTime	R
IC 10		5	comment.id	ID	R
IC 11		6	comment.content	Text	R
IC 12	sort	1	comment.creationDate	↓	
IC 13		2	comment.id	↑	
IC 14v1	limit	20			
IC 14v2	CPs	2.4, 3.3, 5.3			
	relevance	This query looks for paths of length two, starting from a given Person, going through its created Messages and finishing at their replies. In this query there is temporal locality between the replies being accessed. Thus the top-k order by this can interact with the selection, i.e. do not consider older Posts than the 20th oldest seen so far.			

Interactive / complex / 9

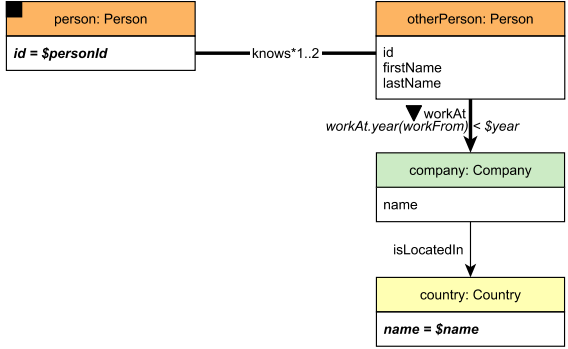
IC 1	query	Interactive / complex / 9			
IC 2	title	Recent messages by friends or friends of friends			
IC 3	pattern	 <pre> graph LR person[person: Person] -- "knows*1..2" --> otherPerson[otherPerson: Person] otherPerson -- "hasCreator" --> message[message: Message] person -- "id = \$personId" --> person message -- "creationDate < \$maxDate" --> message message -- "id" --> message message -- "content / imageFile" --> message message -- "creationDate" --> message </pre>			
IC 4					
IC 5					
IC 6					
IC 7					
IC 8					
IC 9					
IC 10					
IC 11					
IC 12	description	Given a start Person with ID \$personId, find the most recent Messages created by that Person's friends or friends of friends (excluding the start Person). Only consider Messages created before the given \$maxDate (excluding that day).			
IC 13	params	1	\$personId	ID	
IC 14v1		2	\$maxDate	Date	
IC 14v2	result	1	otherPerson.id	ID	R
		2	otherPerson.firstName	String	R
		3	otherPerson.lastName	String	R
		4	message.id	ID	R
		5	message.content or message.imageFile (for photos)	Text	R
		6	message.creationDate	DateTime	R
	sort	1	message.creationDate	↓	
		2	message.id	↑	
	limit	20			
	CPs	1.1, 1.2, 2.2, 2.3, 3.2, 3.3, 8.5			
	relevance	This query looks for paths of length two or three, starting from a given Person, moving to its friends and friends of friends, and ending at their created Messages. This is one of the most complex queries, as the list of choke points indicates. This query is expected to touch variable amounts of data with entities of different characteristics, and therefore, properly estimating cardinalities and selecting the proper operators will be crucial.			

Interactive / complex / 10

IC 1
IC 2
IC 3
IC 4
IC 5
IC 6
IC 7
IC 8
IC 9
IC 10
IC 11
IC 12
IC 13
IC 14v1
IC 14v2

query	Interactive / complex / 10			
title	Friend recommendation			
pattern	<p>The diagram illustrates the query pattern. It starts with a <code>person: Person</code> box with a variable <code>id = \$personId</code>. This box is connected via a <code>knows*2..2</code> relationship to a <code>foaf: Person</code> box. The <code>foaf: Person</code> box has a filter: <code>(month(birthday) = \$month and day(birthday) ≥ 21) or (month(birthday) = \$month+1 and day(birthday) < 22)</code>. Below this box are variables <code>id</code>, <code>firstName</code>, <code>lastName</code>, and <code>gender</code>. The <code>foaf: Person</code> box is connected via an <code>isLocatedIn</code> relationship to a <code>city: City</code> box, which has a variable <code>name</code>. Below the main pattern are two sub-patterns: common: Shows <code>person: Person</code> connected to <code>Tag</code> via <code>hasInterest</code>, and <code>foaf: Person</code> connected to <code>Post</code> via <code>hasCreator</code>. <code>Post</code> is connected to <code>Tag</code> via <code>hasTag</code>. A <code>count</code> box is associated with <code>Post</code>. uncommon: Shows <code>person: Person</code> connected to <code>Tag</code> via a red arrow labeled <code>«neg» hasInterest</code>, and <code>foaf: Person</code> connected to <code>Post</code> via <code>hasCreator</code>. <code>Post</code> is connected to <code>Tag</code> via <code>hasTag</code>. A <code>count</code> box is associated with <code>Post</code>.</p>			
description	<p>Given a start Person with ID <code>\$personId</code>, find that Person’s friends of friends (<code>foaf</code>) – excluding the start Person and his/her immediate friends –, who were born on or after the 21st of a given <code>\$month</code> (in any year) and before the 22nd of the following month. Calculate the similarity between each friend and the start person, where <code>commonInterestScore</code> is defined as follows:</p> <ul style="list-style-type: none"> <code>common</code> = number of Posts created by friend, such that the Post has a Tag that the start person is interested in <code>uncommon</code> = number of Posts created by friend, such that the Post has no Tag that the start person is interested in <code>commonInterestScore</code> = <code>common</code> - <code>uncommon</code> 			
params	1	<code>\$personId</code>	ID	
	2	<code>\$month</code>	32-bit Integer	Between 1 and 12. Implementations may also pass the next month as an additional <code>\$nextMonth</code> parameter
result	1	<code>foaf.id</code>	ID	R
	2	<code>foaf.firstName</code>	String	R
	3	<code>foaf.lastName</code>	String	R
	4	<code>commonInterestScore</code>	32-bit Integer	A
	5	<code>foaf.gender</code>	String	R
	6	<code>city.name</code>	String	R
sort	1	<code>commonInterestScore</code>	↓	
	2	<code>foaf.id</code>	↑	
limit	10			
CPs	2.3, 3.3, 4.1, 4.2, 5.1, 5.2, 6.1, 7.1, 8.6			
relevance	<p>This query looks for paths of length two, starting from a Person and ending at the friends of their friends. It does widely scattered graph traversal, and one expects no locality of in friends of friends, as these have been acquired over a long time and have widely scattered identifiers. The join order is simple but one must see that the anti-join for “not in my friends” is better with hash. Also the last pattern in the scalar sub-queries joining or anti-joining the Tags of the candidate’s Posts to interests of self should be by hash.</p>			

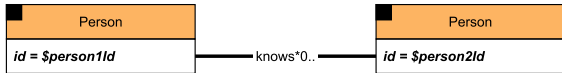
Interactive / complex / 11

IC 1	query	Interactive / complex / 11			
IC 2	title	Job referral			
IC 3	pattern	 <pre> graph TD P1[person: Person id = \$personId] -- knows*1..2 --> P2[otherPerson: Person id firstName lastName] P2 -- "workAt workAt.year(workFrom) < \$year" --> C[company: Company name] C -- isLocatedIn --> CO[country: Country name = \$name] </pre>			
IC 4					
IC 5					
IC 6					
IC 7					
IC 8					
IC 9					
IC 10					
IC 11					
IC 12					
IC 13					
IC 14v1	description	Given a start Person with ID \$personId, find that Person's friends and friends of friends (excluding start Person) who started working in some Company in a given Country with name \$countryName, before a given date (\$workFromYear).			
IC 14v2	params	1	\$personId	ID	
		2	\$countryName	String	
		3	\$workFromYear	32-bit Integer	
	result	1	otherPerson.id	ID	R
		2	otherPerson.firstName	String	R
		3	otherPerson.lastName	String	R
		4	company.name	String	R
		5	workAt.workFrom	32-bit Integer	R
	sort	1	workAt.workFrom	↑	
		2	otherPerson.id	↑	
		3	company.name	↓	
	limit	10			
	CPs	1.3, 2.3, 2.4, 3.3, 4.2			
	relevance	This query looks for paths of length two or three, starting from a Person, moving to friends or friends of friends, and ending at a Company. In this query, there are selective joins and a top-k order by that can be exploited for optimizations.			

Interactive / complex / 12

IC 1	query	Interactive / complex / 12			
IC 2	title	Expert search			
IC 3	pattern				
IC 4	description	<p>Given a start Person with ID <code>\$personId</code>, find the Comments that this Person's friends made in reply to Posts, considering only those Comments that are direct (single-hop) replies to Posts, not the transitive (multi-hop) ones. Only consider Posts with a Tag in a given TagClass with name <code>\$tagClassName</code> or in a descendant of that TagClass. Count the number of these reply Comments, and collect the Tags that were attached to the Posts they replied to, but only collect Tags with the given TagClass or with a descendant of that TagClass. Return Persons with at least one reply, the reply count, and the collection of Tags.</p>			
IC 5	params	1	<code>\$personId</code>	ID	
IC 6		2	<code>\$tagClassName</code>	Long String	
IC 7	result	1	<code>friend.id</code>	ID	R
IC 8		2	<code>friend.firstName</code>	String	R
IC 9		3	<code>friend.lastName</code>	String	R
IC 10		4	<code>tagNames</code>	{Long String}	A
IC 11		5	<code>replyCount</code>	32-bit Integer	A
IC 12	sort	1	<code>replyCount</code>	↓	
IC 13		2	<code>friend.id</code>	↑	
IC 14v1	limit	20			
IC 14v2	CPs	3.3, 7.2, 7.3, 8.2			
	relevance	<p>This query starts at a Person, moves to its friends, and then to their Comments and their root Posts. Then, it gets the Tag of each Post and checks whether it (directly or transitively) belongs to the specified TagClass. This can be thought of as a bidirectional search between the Person and the TagClass. The difficulty of this query is determining the optimal direction of this traversal.</p>			

Interactive / complex / 13

query	Interactive / complex / 13				
title	Single shortest path				
pattern					
description	<p>Given two Persons with IDs \$person1Id and \$person2Id, find the shortest path between these two Persons in the subgraph induced by the knows edges. Return the length of this path:</p> <ul style="list-style-type: none">• -1: no path found• 0: start person = end person• > 0: path found (start person ≠ end person)				
params	1	\$person1Id	ID	In SNB Interactive v2, this query has two variants: (b) Guaranteed that there is no path between the two Persons (b) Guaranteed that there is a 4-hop path between the two Persons	
	2	\$person2Id	ID		
result	1	shortestPathLength	32-bit Integer	C	
CPs	3.3, 7.2, 7.3, 7.5, 7.8, 8.1, 8.6				
relevance	This query looks for a variable length path, starting at a given Person and finishing at an another given Person. Proper cardinality estimation and search space pruning, will be crucial. This query also allows for possible parallel implementations.				

Interactive / complex / 14v2

IC 1
IC 2
IC 3
IC 4
IC 5
IC 6
IC 7
IC 8
IC 9
IC 10
IC 11
IC 12
IC 13
IC 14v1
IC 14v2

query	Interactive / complex / 14v2			
title	Trusted connection paths (v2)			
pattern	<div> <div> <p>Find a cheapest path on edges where $\text{numInteractions} \geq 1$, using edge weight = $\max(\text{round}(40 - \sqrt{\text{numInteractions}}), 1)$</p> </div> <div> <p>$\text{numInteractions} = \text{count}(c)$</p> </div> <div> <p>Example for finding a path between person1 and person2</p> </div> </div>			
description	<p><i>This query is used in SNB Interactive v2.</i></p> <p>Find a cheapest path between two given Persons with IDs $\\$person1Id$ and $\\$person2Id$ in the interaction subgraph. If there are multiple cheapest paths, any of them can be returned. Do not return any rows if there is no path between the Persons. The interaction subgraph is based on a projection of the Person-knows-Person graph. In this projection, only those knows edges are kept whose endpoint Persons have at least one interaction between them. An interaction is defined as a direct reply Comment (by one of the Persons) to a Message (by the other Person). The weights are defined as: $\max(\text{round}(40 - \sqrt{\text{numInteractions}}), 1)$</p> <p><i>Note:</i> Interactions are counted both ways, e.g. if Alice knows Bob, Alice writes 2 reply Comments to Bob's Messages and Bob writes 3 reply Comments to Alice's Messages, their total number of interactions is 5 and the weight of the knows edge is 38.</p> <p><i>Remark:</i> Determinism is ensured by using square root followed by rounding. For all integers between 1 and 100 000, the square root's fractional part is more than 10e-5 from 0.5, where the rounding could be non-deterministic based on floating point inaccuracies. As 10e-5 is significantly larger than the machine epsilon of IEEE 754 floats (both 32- and 64-bit), the floating point inaccuracies have no chance to affect the derived integer edge weights.</p>			
params	1	$\$person1Id$	ID	(b) There are no paths between the two Persons (b) There is a 4-hop path between the two Persons
	2	$\$person2Id$	ID	
result	1	personIdsInPath	[ID]	C Identifiers representing an ordered sequence of the Persons in the path
	2	pathWeight	64-bit Integer	C
CPs	3.3, 5.3, 7.6, 7.7, 7.8, 8.1, 8.2, 8.3, 8.6			
relevance	This query tests the performance of cheapest path (weighted shortest path) computation.			